ANALES DE ECONOMÍA APLICADA

ANAIS DE ECONOMIA APLICADA

2009

José Pires Manso
João Dionísio Monteiro
(Directores)

DELTA PUBLICACIONES

sepelt Internacional Economía Aplicada

ANALES DE ECONOMÍA APLICADA

ANAIS DE ECONOMIA APLICADA

2009

Número XXIII

José R. Pires Manso
João Dionísio Monteiro
(Directores)

DELTA PUBLICACIONES

sepelt Internacional Economía Aplicada

ASEPELT son las siglas de la Asociación Científica Internacional de Economía Aplicada. El objetivo de la Asociación, recogido en sus estatutos y que inspira su actividad, es organizar, promover y favorecer trabajos originales de carácter científico en el dominio de la Economía Aplicada.

Para ello, desde la Asociación se definen las siguientes líneas de actuación:

1. El intercambio de información, entre sus miembros, de sus trabajos de investigación.
2. La publicación de una Revista Científica: "Estudios de Economía Aplicada".
3. La publicación de selecciones de artículos o de obras colectivas.
4. La organización de seminarios, coloquios o congresos.

La Asociación se convierte, de esta manera, en un foro abierto al intercambio y debate de las distintas ideas y aportaciones científicas, que se desarrollan tanto en el ámbito universitario como en el empresarial, dentro del campo de la Economía Aplicada.

Las Reuniones Anuales de la Asociación, punto de encuentro de un elevado número de investigadores involucrados en estas cuestiones, proporcionan un medio excepcional para el mejor conocimiento mutuo, embrión de futuras colaboraciones.

Esta publicación ANALES DE ECONOMÍA APLICADA es la expresión del camino que están tomando las nuevas iniciativas en el ámbito de la investigación y de la innovación en Economía Aplicada.

ASEPELT são as siglas da Associação Científica Internacional de Economia Aplicada. O objectivo da Associação, transcrito nos seus estatutos, e que inspira a sua actividade, é organizar, promover e favorecer trabalhos originais de carácter científico no domínio da Economia Aplicada.

Para isso, a Associação define as seguintes linhas de actuação:

1. O intercâmbio de informação, entre seus membros, de seus trabalhos de investigação.
2. A publicação de uma Revista Científica: “Estudios de Economía Aplicada”.
3. A publicação de selecções de artigos ou de obras colectivas.
4. A organização de seminários, colóquios ou congressos.

A Associação converte-se, desta forma, num fórum aberto ao intercâmbio e debate de distintas ideias e contributos científicos, que se desenvolvem tanto no âmbito universitário como no empresarial, dentro do campo da Economia Aplicada.

As Reuniões Anuais da Associação, ponto de encontro de um elevado número de investigadores envolvidos na discussão destas questões, proporcionam um meio excepcional para o melhor conhecimento mútuo, embrião de futuras colaborações.

Esta publicação ANAIS DE ECONOMIA APLICADA é a expressão do caminho que estão a tomar as novas iniciativas no âmbito da investigação e da inovação na Economia Aplicada.

ANALES DE ECONOMÍA APLICADA
ANAIS DE ECONOMIA APLICADA
Año/Ano 2009 - Número XXIII

© 2009 ASEPELT
© 2009 Delta Publicaciones Universitarias

Reservados todos los derechos. El contenido de esta publicación,
tanto de la obra escrita como electrónica, puede ser utilizado,
de común acuerdo con ASEPELT y DELTA PUBLICACIONES,
para usos exclusivamente particulares y/o profesionales
y, en ningún caso, comerciales.

Reservados todos os direitos. O conteúdo desta publicação
tanto da obra escrita como eletrônica, pode ser utilizado,
de comum acordo com ASEPELT e DELTA PUBLICACIONES,
para usos exclusivamente particulares e/ou profissionais
e, em nenhum caso, comerciais.

ISBN: 978-84-92453-69-6
Depósito Legal:

Preimpresión/Preimpressão: Delta Publicaciones
Impresión/Impressão: FER Impresores

Comité Académico de ASEPELT

Presidente JESÚS BERNARDO PENA TRAPERO
Universidad de Alcalá

Vocal: **ANTONIO GARCÍA LIZANA**
Universidad de Málaga

Vocal: **ANTONIO PULIDO SANROMÁN**
Universidad Autónoma de Madrid

Vocal: **ESTELA BEE DAGUM**
Universidad de Bolonia

Vocal: **JOAQUÍN ARANDA GALLEGO**
Universidad de Murcia

Vocal: **JOSÉ GARCÍA PÉREZ**
Universidad de Almería

Vocal: **JOSEFA E. FERNÁNDEZ ARUFE**
Universidad de Valladolid

Vocal: **LAWRENCE R. KLEIN**
Pennsylvania University

Vocal: **RAFAEL HERRERÍAS PLEGUEZUELO**
Universidad de Granada

Comité Científico

Adolfo Rodero Franganillo
Universidad ETEA Córdoba-ESPAÑA
Agustín Hernández Bastida
Universidad de Granada-ESPAÑA
Alcino F.P. Couto
FCSH/UBI-PORTUGAL
Alfredo Iglesias Suárez
Universidad de Castilla–La Mancha-ESPAÑA
Álvaro Martins Monteiro
ISEG-UTL-PORTUGAL
Amelia Pérez Zabaleta
UNED-ESPAÑA
Ana Paula Matias Gama
FCSH/UBI-PORTUGAL
Anabela Almeida
FSCG/UBI-PORTUGAL
Anabela Dinis
FSCG/UBI-PORTUGAL
António F. de Matos
FSCG/UBI-PORTUGAL
Antonio García Lizana
Universidad de Málaga-ESPAÑA
António L. Silvestre
ISEG/UTL-PORTUGAL
Antonio Macías Hernández
Universidad de La Laguna-ESPAÑA
António Marques
FSCG/UBI-PORTUGAL
Antonio Pulido San Román
Universidad Autónoma de Madrid-MADRID
Bala Veeramacheneni
State University of New York-USA
Beatriz González López-Valcárcel
Universidad de Las Palmas de Gran Canaria-ESPAÑA
Carles Murillo Fort
Universidad Pompeu Fabra-ESPAÑA
Carlos Barros
ISEG/UTL-PORTUGAL
Carlos Braumann
U. Evora-PORTUGAL
Carlos Legna Verna
Universidad de La Laguna-ESPAÑA
Carlos Moslares García
IQS. Universitat Ramon Llull-ESPAÑA
Carmen Delia Dávila Quintana
Universidad de Las Palmas de Gran Canaria-ESPAÑA
Christopher Gerry
UTAD-PORTUGAL
Diego Martínez Perán
Intereconomía-ESPAÑA
Estela Bee Dagum
Universidad de Bolonia-ITALIA
Eugeni Aguiló Pérez
Universidad de las Islas Baleares-ESPAÑA
Fco. Javier Callealta Barroso
Universidad de Alcalá de Henares-ESPAÑA
Fernanda Nogueira
UTAD-PORTUGAL
Fernando Cortina García
Instituto Nacional de Estadística-ESPAÑA
Flor María Guerrero Casas
Universidad Pablo Olavide-ESPAÑA
Francisco Diniz
UTAD-PORTUGAL
Francisco José Vázquez Polo
Universidad de Las Palmas de Gran Canaria-ESPAÑA
Ginés Guirao Pérez
Universidad de La Laguna-ESPAÑA
Helena Corrales Herrero
Universidad de Valladolid-ESPAÑA
Isabel Soares
FE/UP-PORTUGAL
Jacques Soichot
UB-Dijón
Jesús Bernardo Pena Trapero
Universidad de Alcalá de Henares-ESPAÑA
Jesús Cavero Álvarez
Universidad de Valladolid
Jesús Esteban García
Universidad de Valencia-ESPAÑA
João Albino M. Silva
FE/Ualg-PORTUGAL
João C. C. Leitão
IST/UTL-PORTUGAL
João Carlos Matias
FE/UBI-Portugal
João D. Monteiro
FCSH/UBI-PORTUGAL
João José M Ferreira
FCSH/UBI-Portugal
João Rebelo
UTAD-PORTUGAL
João Sousa Andrade
FE/UC-PORTUGAL
Joaquim Ramos Silva
ISEG/UTL-Portugal
Joaquín Aranda Gallego
Universidad de Murcia-ESPAÑA
Joaquín Turmo Garuz
Universidad de Barcelona-ESPAÑA
Jorge Miguel Silva
FCE/UBI-PORTUGAL
José A. Cadima Ribeiro
EEG/UM-PORTUGAL
José A. S. Rodrigues Fuinhas
FCSH/UBI-PORTUGAL
José Alberto Soares da Fonseca
FE/UC-Portugal
José Boza Chirino
Universidad de Las Palmas-ESPAÑA
José Callejon Cespedes
Universidad de Granada-España
José García Pérez
Universidad de Almería-ESPAÑA
José L Miralles Quirós
Universidad de Extremadura-ESPAÑA
José Luis Rojo García
Universidad de Valladolid-ESPAÑA
José Manuel Díez Modino
Universidad de León-ESPAÑA
José María Gil Roig
Universidad de Politécnica de Cataluña-ESPAÑA
José María Moreno Jiménez
Universidad de Zaragoza-ESPAÑA
José María Sarabia
Universidad de Cantabria-ESPAÑA
Jose Miguel Casas Sánchez
Universidad de Alcalá de Henares-ESPAÑA
José R. Pires Manso
FCSH/UBI-PORTUGAL
Josefa E. Fernández Arufe
Universidad de Valladolid-ESPAÑA
Josefina Martínez Barbeito
Universidad de La Coruña-ESPAÑA
Josep García Blandón
Universidad Ramon Llull-ESPAÑA

Josep Lladós Masllorenç
UOC-ESPAÑA
Josep Maria Argilés Bosch
Universidad de Barcelona-ESPAÑA
Juan José Cáceres Hernández
Universidad de La Laguna-ESPAÑA
Juan Manuel Rey Julia
Universidad de Cádiz-ESPAÑA
Julián Ramajo Hernandez
Universidad de Extremadura-ESPAÑA
Justo Sotelo Navalpotro
Universidad de San Pablo CEU-ESPAÑA
Lawrence R. Klein
Pennsylvania University-USA
Lluís Jovell I Turró
Universidad Autónoma de Barcelona-ESPAÑA
Lucía Navarro Gómez
Universidad de Málaga-ESPAÑA
Lucinio González Sabaté
Universidad de Ramon Llull-ESPAÑA
Luís Lourenço
FCSH/UBI-PORTUGAL
Luis Mendes
FCSH/UBI-Portugal
Mª Isabel Sánchez y Sánchez Amaya
Universidad del País Vasco-ESPAÑA
Manuel A. Ferreira
ISCTE-PORTUGAL
Manuel Antonio Muñiz Pérez
Universidad de Oviedo-ESPAÑA
Manuel Coelho
ISEG/UTL-Portugal
Manuel Díez Modino
Universidad de León-ESPAÑA
Manuel Navarro Ibáñez
Universidad de La Laguna-ESPAÑA
Manuela Hill
ISCTE-PORTUGAL
Mar Guitert Catasús
Universidad de Ramon Llull-ESPAÑA
Margarida Proença
EEG/UM-PORTUGAL
Margarida Vaz
FCSH/UBI-PORTUGAL
María del Carmen Guisán Seijas
Universidad de Santiago de Compostela-ESPAÑA
Maria do Céu Alves
FCSH/UBI-Portugal
Maria Fátima David
ESTG/IPG-Portugal
Maria José Silva
FCSH/UBI-PORTUGAL
Maria Teresa Noronha
FE/Ualg-PORTUGAL
Mariana Bosch Casabó
Universidad de Ramon Llull-ESPAÑA
Mário B. Raposo
FCSH/UBI-PORTUGAL
Mário J. B. Franco
FCSH/UBI-Portugal
Miguel Angel Fajardo Caldera
Universidad de Extremadura-ESPAÑA
Milagros Dones Tacero
Universidad de Autónoma de Madrid-ESPAÑA
Mònica Martínez Blasco
Universidad de Ramon Llull-ESPAÑA
Nuno Valério
ISEG/UTL-PORTUGAL
Paulo Duarte
FCSH/UBI-PORTUGAL
Paulo G. Pinheiro
FCSH/UBI-PORTUGAL
Paulo M. Nunes
FCSH/UBI-PORTUGAL
Pedro Guedes de Carvalho
FCSH/UBI-Portugal
Pedro P. Barros
FE/UNL-PORTUGAL
Rafael Amiel Sánchez
Global Insight-ESPAÑA
Rafael Caballero
Universidad de Málaga-ESPAÑA
Rafael Herrerías Pleguezuelo
Universidad de Granada-ESPAÑA
Ricardo Alves
FCSH/UBI-PORTUGAL
Ricardo G. Rodrigues
FCSH/UBI-Portugal
Ricardo Úbeda Sales
Universidad de Ramón Llull-ESPAÑA
Rigoberto Pérez Suárez
Universidad de Oviedo-ESPAÑA
Rui Junqueira Lopes
UE-PORTUGAL
Santiago Rodríguez Feijoo
Universidad de Las Palmas de Gran Canaria-ESPAÑA
Susana C. Gabriel
FCSH/UBI-PORTUGAL
Susana P. Garrido
FCSH/UBI-Portugal
Tiago N. Sequeira
FCSH/UBI-PORTUGAL
Victor Cano Fernández
Universidad de La Laguna-ESPAÑA
Vitor Cavaleiro
FE/UBI-Portugal
Xavier Tomàs Morer
Universidad de Ramon Llull-ESPAÑA
Zélia Serrasqueiro
FCSH/UBI-PORTUGAL

Comité de Honor/Comité de Honra

D. **Victor Constâncio**
Governador do Banco de Portugal
D. **Castro Guerra**
Secretário de Estado da Indústria e Energía
D. **Emílio Rui Vilar**
Presidente da Fundacão C. Gulbenkian
D. **João Senteiro**
Presidente da FCT
D. **Manuel J. Santos Silva**
Reitor da Universidade da Beira Interior
D. **Luís Lourenço**
Presidente da Faculdade das Ciências Sociais e Humanas
D. **Ginés Guirao Pérez**
Presidente de ASEPELT
D. **Murteira Nabo**
Bastonário da Ordem dos Economistas
D. **Victor Santos**
ISEG/UTL e Presidente da ERSE
D. **Simões Lopes**
Ex-bastonário da Ordem dos Economistas e ex-Reitor da Universidade Técnica de Lisboa
D. **Carlos Pinto**
Presidente da Câmara da Covilhã
Dª. **Alzira Serrasqueiro**
Governadora Civil do Distrito de Castelo Branco
D. **Artur Santos Silva**
Presidente do Conselho de Administracão do Banco Português de Investimento (BPI)
D. **Júlio Fermoso**
Presidente do Conselho de Administracão da Caja Duero

Comité de Organizador

José R. Pires Manso (Chairman)
Alcino P. Couto
António F. de Matos
António Marques
João D. Monteiro
José A. S. Rodrigues Fuinhas
Margarida M. Vaz
Paulo M. Nunes
Susana C. Gabriel
Tiago N. Sequeira

Comité Ejecutivo/Comité Executivo de ASEPELT

Presidente
Ginés Guirao Pérez

Vicepresidente
Adolfo Rodero Franganillo

Secretario general
Carlos Moslares García

Vocales
Justo Sotelo Navalpotro
José Manuel Diez Modino
José María Moreno Jiménez
Flor Mª Guerrero Casas
Miguel Ángel Fajardo Caldera
José Boza Chirino
Fernando Cortina García
José María Sarabia Alegría

Reuniones Anuales celebradas por la Asociación ASEPELT-España

Reunión I	1987	Barcelona
Reunión II	1988	Valladolid
Reunión III	1989	Sevilla
Reunión IV	1990	Murcia
Reunión V	1991	Las Palmas de Gran Canaria
Reunión VI	1992	Granada
Reunión VII	1993	Cádiz
Reunión VIII	1994	Mallorca
Reunión IX	1995	Santiago de Compostela
Reunión X	1996	Albacete
Reunión XI	1997	Bilbao
Reunión XII	1998	Córdoba
Reunión XIII	1999	Burgos
Reunión XIV	2000	Oviedo
Reunión XV	2001	La Coruña
Reunión XVI	2002	Madrid
Reunión XVII	2003	Almería
Reunión XVIII	2004	León
Reunión XIX	2005	Badajoz
Reunión XX	2006	Tenerife
Reunión XXI	2007	Valladolid
Reunión XXII	2008	Barcelona
Reunión XXIII	**2009**	**Covilhã (Portugal)**

APRESENTACÃO

A ciência económica evolui muito rapidamente e cada vez mais para áreas interdisciplinares; e essa evolução tem tido reconhecimento público como é o caso da recente atribuição do prémio Nobel da Economia a um investigador da área da psicologia económica; mas outras áreas há que se têm vindo a afirmar progressivamente como é o caso da economia da energia e do ambiente, da economia da saúde, da economia experimental, da economia da estratégia, ou da economia da inovação, entre outras.

Neste XXIII Congresso Internacional de Economia Aplicada, o XXIII Encontro Anual e primeiro da ASEPELT Internacional, dada a sua recente mudança de nome, a par das tradicionais áreas da economia, abriu-se o leque a algumas destas novas áreas científicas e a verdade é que os investigadores corresponderam totalmente à nossa expectativa apresentando trabalho de indubitável interesse e em quantidades dignas de registo, como o prova o resultado da dupla revisão a que os mesmos artigos foram sujeitos e que aprovou a apresentação de um razoável número desses artigos. Esperemos que seja também essa a opinião dos investigadores que se dignarem assistir ao XXIII congresso na Covilhã.

Ao todo as áreas científicas a concurso foram dezasseis, das mais tradicionais a outras mais recentes como as já referidas áreas da economia da energia e do ambiente, da estratégia económica, da economia da inovação, da economia do conhecimento. O número de artigos recebidos foi de várias centenas, dos quais se seleccionaram, após a referida dupla revisão por avaliadores anónimos e independentes, cerca de duas centenas que estão englobados no CD dos Anais de Economia Aplicada que a Delta Publicaciones acaba de editar. Alguns desses artigos, distinguidos pelos próprios avaliadores, merecem um destaque especial e por isso se publicam como 'ponencias' na versão em livro dos mesmos Anais, a par dos resumos de todas as comunicações seleccionadas para apresentação no congresso; quer um quer o outro desses documentos apresenta ISBN.

No âmbito desta nota introdutória cabe um agradecimento especial ao imenso trabalho desenvolvido pelos inúmeros revisores anónimos e independentes, espanhóis e portugueses, de todos os artigos recebidos, alguns deles em áreas bem densas e complexas.

Também queremos destacar o trabalho do Comité Organizador constituído por colegas da Universidade da Beira Interior, realçando apenas o nome do colega João Dionísio Monteiro pela sua total disponibilidade e empenho desde o início deste processo, e o trabalho do secretariado com destaque para a técnica superior Elisabete

Ramos, e ainda a Delta Publicaciones que pacientemente foi aguardando o envio dos nossos trabalhos, apesar do extremamente curto espaço de tempo de que dispunha.

Uma palavra breve para agradecer ao Presidente da ASEPELT Internacional, Prof. Ginés Guirao, pelo apoio sempre disponibilizado e por ter confiado nas nossas mãos a organização deste congresso, o congresso da internacionalização da ASEPELT.

Como é evidente o maior realce vai para as centenas de autores dos artigos seleccionados; são eles os verdadeiros responsáveis por esta obra e pelo seu mérito científico.

Um obrigado a todos pela colaboração dispensada.

Covilhã, Junho de 2009

José Ramos Pires Manso
João Dionísio Monteiro
(directores)

PRESENTACIÓN

La ciencia económica evoluciona muy rápidamente y cada vez más para áreas interdisciplinares; y esa evolución ha tenido reconocimiento público como es el caso de la reciente atribución del Premio Nobel de Economía a un investigador del área de la Psicología Económica; pero hay otras áreas que se están afirmándose progresivamente como es el caso de la Economía de la energía y del medioambiente, Economía de la Salud, Economía experimental, Economía de la estrategia o la Economía de la innovación, entre otras.

En este XXIII Congreso Internacional de Economía Aplicada, o XXIII Encuentro Anual y primero de ASEPELT Internacional, dado su reciente cambio de nombre, junto a las áreas tradicionales de la Economía, se abre el abanico a algunas de estas nuevas áreas científicas y la verdad es que los investigadores correspondieron totalmente a nuestra expectativa, presentando trabajos de indudable interés y en cantidades dignas de mención, como prueba el resultado de la doble revisión a la que los mismos artículos fueron sujetos y que aprobó un razonable número de ellos. Esperemos también que esta sea también la opinión de los investigadores que se dignaron a asistir al XXIII Congreso en Covilhã.

Finalmente las áreas científicas a concurso fueron dieciséis, desde las más tradicionales a otras más recientes como las ya referidas áreas de Economía de la energía y del medioambiente, Economía de la estrategia económica, Economía de la innovación y Economía del conocimiento. Fueron cientos los artículos recibidos, de los cuales se seleccionaron tras la referida revisión doble por evaluadores anónimos e independientes, cerca de dos centenares que están recogidos en el CD de los Anales de Economía Aplicada que Delta Publicaciones acaba de editar. Algunos de estos artículos, distinguidos por los propios evaluadores, merecen una relevancia especial y por ello se publican como "ponencias" en la versión en libro de los mismos Anales, así como los resúmenes de todas las comunicaciones seleccionadas para su presentación en el Congreso; en todo caso, cualesquiera de esos documentos presenta ISBN.

En el ámbito de esta nota introductoria, cabe un agradecimiento especial al inmenso trabajo desarrollado por los innumerables revisores anónimos e independientes, españoles y portugueses, de todos los artículos recibidos, algunos de ellos en áreas muy densas y complejas.

También queremos agradecer el trabajo del Comité Organizador constituido por colegas de la Universidade da Beira Interior, destacando apenas el nombre del colega João Dionísio Monteiro por su total disponibilidad y empeño desde el inicio de este proceso, y el trabajo de secretariado, con especial mención a la labor técnica superior

de Elisabete Ramos, e incluso a Delta Publicaciones que pacientemente aguardó el envío de nuestros trabajos, a pesar del extremadamente corto espacio de tiempo del que disponía.

Unas breves palabras también para agradecer al Presidente de ASEPELT Internacional, Prof. Ginés Guirao, por el apoyo siempre dispuesto y por haber confiado a nuestras manos la organización de este congreso, el congreso de la internacionalización de ASEPELT.

Como es evidente, el mayor realce va para los centenares de autores de los artículos seleccionados; son ellos los verdaderos responsables de esta obra y por su mérito científico.

Agradecemos a todos la colaboración dispensada

Covilhã, Junho de 2009

José Ramos Pires Manso
João Dionísio Monteiro
(directores)

ÁREA I
ECONOMÍA INDUSTRIAL
ECONOMIA INDUSTRIAL

EFICIÊNCIA E CONCENTRAÇÃO NA INDÚSTRIA BANCÁRIA

FRANCISCO VITORINO DA SILVA MARTINS
E-mail: diamar@netcabo.pt ou vmartins@fep.up.pt
Universidade do Porto, Faculdade de Economia
MARIA CLARA DIAS PINTO RIBEIRO
E-mail: diamar.clara@netcabo.pt ou mclara@iscap.ipp.pt
Instituto Politécnico do Porto, ISCAP

Resumo

O objectivo deste trabalho é a análise da eficiência produtiva e dos efeitos da concentração sobre os custos bancários, tendo por base a indústria bancária portuguesa. O carácter multiproduto da empresa bancária sugere a necessidade de se adoptar formas multiproduto da função custo (tipo Fourier). Introduzimos variáveis de homogeneidade e de estrutura que permitem o recurso a formas funcionais uniproduto (Cobb-Douglas) à banca. A amostra corresponde a 22 bancos que operavam em Portugal entre 1995-2001, base não consolidada e dados em painel. Para o estudo da ineficiência recorreu-se ao modelo estocástico da curva fronteira (SFA), para as duas especificações. Na análise da concentração, introduziram-se variáveis binárias que pretendem captar os efeitos durante quatro anos após a concentração. Tanto no caso da SFA como no da concentração, os resultados encontrados são sensíveis à especificação funcional adoptada. Concluindo, o processo de concentração bancário parece justificar-se pela possibilidade da diminuição da ineficiência-X.

Palavras-chave: Indústria bancária; X-ineficiência; fusões e aquisições; dados em painel
Área temática: Economia Industrial

Abstract

This study addresses the productive efficiency and the effects of concentration over the banking costs, stressing its focus on the Portuguese banking market. The multiproduct character of the banking firm suggests the use of functional forms as Fourier. The introduction of variables of structure and of homogeneity allows the association of the banking activity (multiproduct) with a single product function (Cobb-Douglas type). The sample covers 22 banks which operated in Portugal from 1995-2001, non consolidated base with a panel data structure. The study about inefficiency is elaborated through the stochastic frontier model (SFA), for the two specifications selected. As a methodology to analyze the concentration, we introduced binary variables, which intend to catch the effects through four years after the concentration process. The results obtained, through SFA and concentration approach, are influenced by the kind of specifications selected. Summing up, the concentration process of the Banking Industry sounds to be justified by the possibility of the X-inefficiency.

Key words: Banking Industry; X-Inefficiency; Mergers and Acquisitions; Panel Data Analysis.
Thematic Area: Industrial Economics.

INTRODUÇÃO

O objectivo deste trabalho é o estudo da eficiência produtiva e dos efeitos da concentração sobre os custos bancários, tendo como base a indústria bancária portuguesa.

O conceito de ineficiência-X ou ineficiência custo, da área da Economia Industrial, introduzido por H. Leibenstein (1966), foi adoptado pela investigação em economia bancária. Para o cômputo da ineficiência custo recorre-se a diferentes métodos, paramétricos e não paramétricos, que pretendem medir a distância entre a fronteira de produção eficiente e a situação particular da empresa bancária. Para o estudo da eficiência, recorre-se ao modelo de fronteira estocástica (SFA), com a análise comparativa das especificações Cobb-Douglas e Fourier da função custo. Os efeitos sobre a eficiência decorrentes do processo de concentração têm sido objecto de estudo em diversos trabalhos de economia bancária. Os estudos pioneiros de economia bancária decorreram, precisamente, do acelerado processo de fusões e aquisições[1] ocorrido, a partir de meados dos anos cinquenta, nos Estados Unidos. Expressões como concentração, fusões e aquisições (F&A), "takeover", participação ou agrupamentos contratuais não são coincidentes e deverão ser diferenciadas (Hirshleifer (1993)). A diferença entre fusão e aquisição caracteriza-se, sinteticamente, na permanência de todos os accionistas (no caso de uma fusão) ou do abandono dos accionistas da empresa alvo (no caso de uma aquisição). Quando se fala de concentração ─ fundamentalmente, de fusões ou de aquisições ─ o número de instituições (ou grupos) operacionais é reduzido, havendo lugar a processos de reestruturação significativos. Para o cômputo dos efeitos da concentração sobre os custos bancários foram consideradas quatro variáveis *dummy* (que pretendem captar possíveis efeitos sobre os custos no próprio ano, um ano, dois anos e três anos após a concentração), associados aos modelos Cobb-Douglas e Fourier. Os modelos propostos incluem quatro inputs para melhor interpretar os recursos produtivos que mais contribuem para a eficiência produtiva e para detectar eventuais falhas na gestão ineficiente desses recursos.

A amostra corresponde a vinte e dois bancos operando em Portugal em 31 de Dezembro, dados não consolidados, no período de 1995 a 2001. O processo de estimação releva a estrutura dos dados em painel. Esta análise surge como uma das mais inovadoras e promissoras linhas de desenvolvimento econométrico, tanto ao nível das técnicas de estimação como dos resultados teóricos. O recurso a dados em painel permite considerar a existência de efeitos não observáveis, específicos a cada um dos bancos, conseguindo ultrapassar, ainda que parcialmente, o enviesamento no processo de estimação decorrente da existência de variáveis omitidas.

Este trabalho é dividido em três pontos. O primeiro discute os diferentes conceitos de eficiência e a sua mensuração, por via de modelos estocásticos. O segundo ponto analisa as tendências e consequências da concentração sobre os custos bancários, bem como a abordagem econométrica para a mensuração dos efeitos dessa concentração sobre os custos bancários. No terceiro ponto estima-se a eficiência produtiva e os efeitos da concentração sobre os custos bancários, com base numa amostra do sector bancário português de dados em painel, para duas formas funcionais da função custo bancária. Finalmente, são apresentadas as principais conclusões que podem ser emanadas sobre a banca nacional com base nas aplicações empíricas.

1. A ESTIMAÇÃO DA EFICIÊNCIA NO SECTOR BANCÁRIO

O conceito de eficiência-X foi introduzido por Leibenstein (1966) com o objectivo de analisar a eficácia do funcionamento das empresas. Contrariando a teoria neo-clássica da empresa Leinbenstein afirma que as empresas que operam segundo a minimização dos custos (sobre a curva fronteira) são excepções.

Diferenças de comportamento de gestão no controlo dos custos (ou de maximização dos rendimentos), imperfeições do mercado ou imposições regulamentares parecem, no seu conjunto, ser relativamente mais importantes para o estudo da eficiência do que a escolha da

[1] Ao longo do texto sinonimizamos as expressões "fusões e aquisições" e "concentração".

escala ou da gama. Berger, Hunter e Timme(1993)[2], na sua revisão dos estudos sobre eficiência, afirmam que a ineficiência-X é responsável por mais de 20% da totalidade dos custos bancários (enquanto que as ineficiências ligadas à escala e à gama não ultrapassavam os 5% dos custos).[3]

O processo de estimação baseado nas formas funcionais custo não permite uma separação directa entre a ineficiência decorrente da escala e a ineficiência-X. Por outro lado, o maior problema para o cômputo da ineficiência é conseguir separar a efectiva ineficiência de outros factores aleatórios que afectem o comportamento dos custos.[4] Para tentar resolver esta dificuldade no cômputo da eficiência-X, foram utilizados na literatura bancária diferentes métodos.[5]

Neste trabalho recorremos ao método estocástico que é usado, mais recorrentemente, nos estudos sobre a empresa bancária para a estimação das funções custo fronteira - SFA (stochastic frontier approach). O método SFA (stochastic frontier approach) especifica que os custos observados num banco podem desviar-se da curva fronteira eficiente devido quer à ineficiência quer a flutuações aleatórias; para separar estas duas parcelas assume-se que o termo da ineficiência segue, habitualmente, uma distribuição assimétrica semi-normal e os erros aleatórios seguem uma distribuição simétrica normal (Aigner, Lovell e Smith (1977)). Pode ser aplicada a dados cross section ou a dados em painel.

Nesta aproximação, a função custo tem três parcelas: a primeira correspondente aos custos dos produtores mais eficientes, a segunda derivada da ineficiência-X (montante dos custos associados ao desperdício de recursos) e a terceira decorrente dos efeitos aleatórios não controlados pela empresa.

(1) $CT = CT(Q_i, W_i, B_i) + u_i + v_i$

A $\mathrm{CT}(\mathrm{Q_i}, \mathrm{W_i}, \mathrm{B_i})$ designa-se por curva custo fronteira ou fronteira estocástica – ou seja, são estimados os custos necessários para atingir um determinado nível de produção sem desperdícios de recursos. A curva fronteira representa a relação entre custos, níveis de produções e preços dos factores produtivos dos bancos relativamente mais eficientes, ou seja, os de "melhor prática". Se $v_i = 0$, o modelo da fronteira estocástica (SFA) coincide com o modelo da fronteira determinística.

A u_i impõe-se que seja estritamente positivo, uma vez que a ineficiência aumenta os custos; habitualmente assume-se que u_i segue uma distribuição semi-normal. A v_i associa-se uma distribuição normal de média zero e com desvio-padrão de σ_v. Assume-se, ainda, que o erro da função custo, ε, é dado por $\varepsilon = u + v$ e u e v são variáveis aleatórias independentes.

A sua função densidade de probabilidade será dada por

$$(2)\ g(\varepsilon) = \frac{2}{\sigma} f\left(\frac{\varepsilon}{\sigma}\right)\left[1 - F\left(\varepsilon\lambda / \sigma\right)\right]$$

sendo $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$, $\lambda = \frac{\sigma_u}{\sigma_v}$, f(.) é a função densidade da normal reduzida e F(.) é a função distribuição da normal reduzida.

Jondrow, Lovell, Materov e Smidt (1982) mostraram que o rácio da variação das variáveis ν e υ, seja λ, pode ser tomada como uma medida da ineficiência relativa de um banco:

$$(3)\ \lambda = \frac{\sigma_u}{\sigma_v}$$

[2]Berger, Hunter e Timme (1993), p. 222

[3] Contrariando as conclusões de Berger, Hunter e Timme(1993), Altunbas, Liu, Molyneux e Seth (2000), para o caso específico do Japão, obtiveram como resultado que as ineficiências ligadas à escala superavam a ineficiência - X.

[4] No nosso trabalho vamos trabalhar exclusivamente com funções custo e não com funções lucro. Apesar das vantagens que o recurso a funções lucro pode trazer, em particular para a análise das fusões e aquisições (Berger (1995)), a especificação da função lucro exige que se conheçam os preços dos inputs e dos outputs.

[5] Para mais desenvolvimentos ver Resti (2000).

A estimação da ineficiência produtiva de um banco, dada pelo parâmetro λ, pode ser feita de uma forma directa, através da maximização da função de log-verosimilhança ($\ln\varphi$), decorrente de $g(\varepsilon)$ - estimação pelo método da máxima verosimilhança:

$$(4)\ \ln\varphi = \frac{N}{2}\ln\frac{2}{\pi} - N\ln\sigma + \sum_{i=1}^{N}\ln\left[F\left(\varepsilon_i\lambda/\sigma\right)\right] - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\varepsilon_i^2$$

em que N corresponde ao número de observações, $\lambda = \sigma_u/\sigma_v$, $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$, $\varepsilon_i = u_i + v_i$ e F(.) é a função distribuição da normal reduzida.

Os resíduos estimados ε_i podem ser decompostos nas suas parcelas, através do método apresentado por Jondrow, Lovell, Materov e Schmidt (1982), segundo o qual a esperança matemática da distribuição condicionada para o modelo semi-normal é dada por

$$(5)\ E(u_i\left|\varepsilon_i\right.) = \frac{\sigma_u\sigma_v}{\sigma}\left[\frac{f(\varepsilon_i\lambda/\sigma)}{F\left(\varepsilon_i\lambda/\sigma\right)} + \frac{\varepsilon_i\lambda}{\sigma}\right]$$

$E(u_i\left|\varepsilon_i\right.)$ é um estimador cêntrico, mas não consistente do parâmetro u_i, porque, independentemente de N, a variância do estimador é positiva[6].

A eficiência produtiva global (eficiência-X), definida como o rácio entre o mínimo e o custo actual, é dada por[7]:

$$(6)\ EP-X = \frac{CT(\ln q_i, \ln w_i)}{CT_i} = e^{-u_i - v_i}$$

onde q_i e w_i são vectores incluindo os logaritmos dos outputs e os preços dos inputs do i-ésimo produtor. Ignorando o termo de perturbação estocástico v_i, o indicador da eficiência produtiva global obter-se-á por:

$(7)\text{EP-X} = e^{-u}$

A sensibilidade dos resultados estimados para a ineficiência aos diferentes modelos é comummente aceite[8]. No entanto, a ordenação dos produtores pela sua ineficiência individual (dos produtores ou por classes) não parece ser sensivelmente afectada pelo modelo adoptado. De facto, parece haver uma correlação muito elevada entre os resultados obtidos, assumindo diferentes distribuições para u[9], o que permite, nestas circunstâncias, optar por uma distribuição simples para u (como a semi-normal) em vez de distribuições mais complexas. De realçar que o recurso a dados em painel permite a estimação das componentes cross-section e temporais de ε (Schmidt e Sickles (1984)[10]).

2. A ESTIMAÇÃO DA CONCENTRAÇÃO NO SECTOR BANCÁRIO

O movimento de concentração bancária parece ter tendência crescente. A desregulamentação e o progresso tecnológico são factores determinantes para a diminuição das barreiras à entrada e tornam o mercado bancário mais contestável.

A diminuição do número de bancos pode ser explicada, em parte, pelo processo de concentração ocorrido, nestas duas décadas, no conjunto dos países. A tendência para o aumento do processo de fusões e aquisições está patente no número de fusões e aquisições registado, nos últimos anos da década de noventa, comparativamente a anos anteriores. No entanto, este movimento de concentração, no espaço europeu, apresenta um menor dinamismo do que o processo nos Estados Unidos da América, quer em termos de número de fusões e aquisições,

[6] Greene (1993), p.80-82
[7] Resti(2000), p. 562 e Resti(1997), p.235
[8] Kumbhakar e Lovell (2003), p.90
[9] Kumbhakar e Lovell (2003), a propósito da estimação de uma curva fronteira custo proposta por Greene(1990) recorrendo a diferentes distribuições para u , chegaram a coeficientes de correlação dos resultados estimados de 0,7467 (entre a exponencial e a gama) a 0,9803 (entre a normal truncada e a semi-normal).
[10] Citados por Mendes e Rebelo (1999), p.515

quer em valor (Berger, Demsetz e Strahan(1999),p.142); o que parece sugerir um excesso de capacidade, no espaço europeu, e a continuação da dinâmica de fusões e aquisições na Europa (Molyneux (2001)). Uma especificidade da concentração europeia é a da nítida predominância da concentração a nível nacional, havendo uma fraca apetência pelas fusões e aquisições regionais ou transnacionais, ao contrário do que tem ocorrido nos Estados Unidos da América. Com efeito, o aumento das fusões e aquisições europeias (com uma desaceleração do processo de concentração a partir de 2001), continua a situar-se a nível nacional e, sobretudo, no seio dos pequenos países (Banco Central Europeu (2000)). A seriação das consequências do processo de concentração está longe de constituir uma tarefa pacífica, tendo em conta os estudos com conclusões contraditórias. A alteração do poder de mercado por via da alteração dos preços parece ser verdadeira apenas ao nível local, havendo tendência para a uniformização das taxas nos mercados regionais e nacionais. A melhor ou pior performance resultante da concentração aparece como contraditória, ao longo dos diferentes estudos. Também as conclusões acerca de possíveis implicações da concentração sobre a eficiência são díspares: desprezibilidade do efeito das concentrações na eficiência, quando a amostra respeita aos anos oitenta nos Estados Unidos e resultados diferentes entre si, quando a amostra respeita aos anos noventa; aumento da eficiência nos mercados europeus (embora as variáveis "envolventes", caracterizadoras da especificidade de cada país, interfiram com os resultados). A melhoria dos serviços e produtos bancários disponibilizados por via da concentração a pequenos clientes parece depender do tipo de concentração: se os bancos envolvidos são pequenos bancos, essa melhoria tende a ocorrer, mas no caso da concentração envolver grandes bancos a pequena clientela deixa de ser uma opção para a empresa fusionada. A concentração parece promover a maior segurança financeira das instituições, pela dupla via da diversificação e da maior atenção que a entidade reguladora de cada Estado terá sobre bancos de maior dimensão.

A mensuração empírica das consequências da concentração bancária tem como suporte a comparação dos dados contabilísticos ou, alternativamente, do valor de mercado dos bancos, antes da fusão e após a fusão. Simultaneamente, o estudo empírico da concentração pode ser feito *ex-ante*, ou seja, estudando as consequências de uma hipotética fusão ou aquisição, ou *ex-post*, analisando as implicações de uma fusão ou aquisição que ocorreu. Uma outra segmentação no estudo empírico da concentração bancária é o método e a forma funcional adoptadas (SFA, DFA, TFA, DEA, ou outro, por um lado; especificação Cobb-Douglas, Translog, Fourier ou outra, por outro).

Os estudos sobre fusões e aquisições podem dividir-se em dois grandes grupos: aqueles que comparam a situação, antes e após a concentração, baseando-se em dados contabilísticos (nomeadamente, dados do Balanço e da Demonstração de Resultados) e outros que baseiam a sua análise nas variações das cotações, no mercado de capitais, antes do anúncio da fusão ou aquisição e após o anúncio da sua concretização.

A primeira abordagem tem a vantagem de se basear em dados estabilizados contabilisticamente, não assentando em expectativas de performance. Por outro lado, apresenta a desvantagem de haver um "lag" temporal, entre a ocorrência da concentração e a disponibilização pública das contas das instituições de crédito, o que pode explicar, em certa medida, as discrepâncias entre o entendimento dos dirigentes bancários e os resultados empíricos destes estudos. Exemplos deste tipo de abordagem foram desenvolvidos por vários autores, havendo um predomínio dos estudos de F&A, no espaço norte-americano: Srinivasan e Wall (1992), Berger e Humphrey (1992) adquirentes, Berger, Demsetz e Strahan.(1999) ─ para o mercado norte-americano; Focarelli, Panetta e Salleo(1998), Resti (1998), Molyneux, Altunbas e Gardener(1996), Haynes e Thompson (1999), Cuesta e Orea (2002), Mendes e Rebelo (2001), Barros (1996) ─ para o mercado europeu. Genericamente, o conjunto dos estudos, para o mercado bancário europeu, vai no sentido da existência de aumento da eficiência decorrente da concentração, enquanto que os estudos sobre o mercado norte-americano ou não apresentam ganhos com a concentração (anos oitenta) ou apresentam resultados inconclusivos (dados dos anos noventa).

O estudo da eficiência por via da abordagem estocástica é aplicável à análise das consequências das fusões, sendo a forma eleita no presente estudo, no seguimento de Humphrey e Vale (2003). Estes autores recorreram a formas funcionais alternativas para a função custo da

banca ─ Translog, Fourier e Spline ─ para estudar quer a existência de economias à escala quer as alterações decorrentes de fusões (dados reportados ao sector bancário norueguês nos anos 1987-1998). A mensuração dos efeitos das fusões é feita pelo recurso a variáveis binárias para separar a situação sem fusão (ou antes da fusão) da de com fusão (ou após fusão). Assim, os autores recorreram a uma variável dummy M, que assume o valor de zero, antes da fusão e o valor de um, após a fusão. Se uma segunda fusão ocorreu no intervalo da amostra para um determinado banco, então *M* passa a assumir, nesse ano e nos seguintes, o valor de dois (e assim sucessivamente). Humphrey e Vale adicionaram aos três modelos de custo (funções custo logaritmizadas Translog, Fourier e Spline) os termos em $\mathrm{M} - \theta M$ e $\sum \theta_i M \ln Q_i$. Desta aproximação resultou uma redução dos custos por via das fusões.

A extracção das consequências sobre a eficiência das fusões pelo método dos rácios e pelos métodos econométricos parecem ser fortemente correlacionadas, permitindo uma análise complementar ou, mesmo, alternativa (Peristiani (1997)).

3. APLICAÇÕES EMPÍRICAS: O CASO DA BANCA PORTUGUESA

Vão ser analisadas a eficiência, recorrendo a modelos de fronteira estocástica, e a concentração da banca portuguesa, com base em dados não consolidados, de 1995 a 2001, para um conjunto de vinte e dois bancos, tendo por base dados em painel. Na análise da ineficiência-X agregada são comparados os resultados obtidos pelas duas especificações avançadas, Cobb-Douglas e Fourier. Os efeitos da concentração são estimados por via da introdução de variáveis binárias que tentarão captar os efeitos sobre os custos no próprio ano da fusão e ao longo dos três anos subsequentes, para as duas formas funcionais citadas.

3.1. Os Modelos

Na sua formulação generalizada a função Cobb-Douglas, para um processo produtivo incluindo n factores produtivos (Chajai (1986)), obtendo-se uma expressão para a função custo logaritmizada

$$(8)\ \ln CT = \ln A + \frac{1}{\mu}\ln Q + \sum_{i=1}^{n}\frac{\alpha_i}{\mu}\ln w_i$$

onde $\mathrm{w_i}$ representa os preços do input i, α_i a elasticidade da produção em relação ao input i, e as outras variáveis com os significados habituais. Ou seja, a função custo logaritmizada é o desenvolvimento da série de ordem um, na vizinhança de zero, em relação às variáveis $\ln \mathrm{Q}$ e $\ln \mathrm{w_i}$; nesta forma funcional, não são considerados os termos da função de produção de grau igual ou superior a dois.

A função Fourier é uma forma funcional mista ou semi-não paramétrica, construída a partir de uma Translog completa e incluindo os termos trigonométricos de primeira, segunda e terceira ordem. No nosso caso, consideramos apenas os termos trigonométricos de primeira ordem.

$$(9)\quad \ln CT = \alpha_0 + \sum_{i=1}^{m}\alpha_i \ln Q_i + \sum_{j=1}^{n}\beta_j \ln W_j + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\delta_{ij}\ln Q_i \ln Q_j +$$

$$+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma_{ij}\ln W_i \ln W_j + \sum_{i=1}^{m}\sum_{j=1}^{n}\rho_{ij}\ln Q_i \ln W_j + \sum_{i=1}^{m}(\sigma_i \cos z_i + \theta_i \sin z_i)$$

em que CT representa o custo total, $\mathrm{Q_i}$ o nível do output i, $\mathrm{W_i}$ o preço do input i e z_i os valores "ajustados" de $\ln \mathrm{Q_i}$ de forma a variarem no intervalo $[0;2\pi]$[11]. O cálculo de z_i é dado

[11] Berger, Leusner e Mingo.(1997) fazem variar os z_i apenas no intervalo $[0{,}1.2\pi;\ 0{,}9.2\pi]$ para evitar problemas de aproximação junto dos extremos ─ vamos seguir esta sugestão na formulação do cálculo de z_i. De igual forma os autores somam uma unidade aos outputs para evitar problemas no cálculo dos logaritmos.

por $z_i = 0{,}2\pi - \mu \times a + \mu \times \ln Q_i$, onde $[a;b]$ é o intervalo de variação de $\ln Q_i$ e $\mu = \dfrac{0{,}9 \times 2\pi - 0{,}1 \times 2\pi}{b-a}$.

Gallant (1981, 1982) demonstrou que se consegue obter uma boa aproximação da verdadeira função custo, recorrendo a um número limitado de termos trigonométricos[12]. De facto, foi mostrado por Tolstov (1962) que uma combinação linear das funções seno e coseno (séries Fourier) pode ajustar exactamente o comportamento de uma função multivariável (uma vez que as funções seno e coseno são mutuamente ortogonais no intervalo $[0; 2\pi]$).

No caso da forma funcional Fourier, as equações dos "shares" podem ser explicitadas, atendendo ao Lema de Shephard [13]:

$$(10)\; S_i = \beta_i + \sum_{j=1}^{n} \gamma_{ij} \ln W_j + \sum_{i=1}^{m} \rho_{ij} \ln Qi$$

De igual forma, sendo n o número de factores de produção e dado que a soma dos "shares" tem de ser um, apenas n-1 equações podem ser estatisticamente independentes.

Uma vez que o teorema da dualidade exige que a função custo seja linearmente homogénea em relação aos preços dos inputs, os parâmetros têm de obedecer às seguintes restrições:

$$(11)\begin{cases} \sum_{j=1}^{n} \beta_j = 1 \\ \sum_{j=1}^{n} \gamma_{ij} = 0, \quad \forall\, i \\ \sum_{i=1}^{n} \rho_{ij} = 0, \quad \forall\, j \end{cases}$$

As condições de simetria dos parâmetros têm de estar asseguradas na Fourier ($\delta_{ij} = \delta_{ji}$ para i>1, m>j e $\gamma_{ij} = \gamma_{ji}$ para i>1, n>j).

Altunbas, Liu, Molyneux e Seth (2000) sustentam que a função Fourier tem propriedades matemáticas e estatísticas desejáveis, uma vez que é capaz de representar qualquer função com exactidão (mesmo truncada, a forma Fourier pode representar razoavelmente qualquer função). Ao contrário da forma funcional Translog, em que se assume que a função custo da indústria bancária tem, obrigatoriamente, a forma em U (do que, a não ser verdade, decorrerão erros na especificação), a formulação Fourier permite que os dados determinem a forma da função custo.

Um problema que se pode colocar com a aplicação da forma funcional tipo Fourier é a dificuldade (previsivelmente acrescida em relação aos resultados obtidos através da formulação tipo Translog) na interpretação dos resultados, uma vez que a existência de um número tão elevado de variáveis implica um elevado grau de multicolinearidade.

3.2. A definição do Output e dos Inputs

No âmbito deste trabalho adoptou-se a abordagem pela intermediação para a definição de empresa bancária, considerando, assim, a totalidade dos custos como variável explicada pelo modelo. Ou seja, além dos custos operacionais são incluídos os custos financeiros.

É comummente aceite o carácter multiproduto da empresa bancária. Grande parte dos estudos[14] assume como output bancário a totalidade das aplicações creditícias, das aplicações noutras instituições de crédito e das aplicações em títulos.

Mas, independentemente do output escolhido (crédito, total do activo, número de contas), dois problemas decorrem da especificação Cobb-Douglas.

[12] Altunbas, Gardener, Molyneux e Moore (2001), p.1938.
[13] Alguns estudos estimam a função Fourier não considerando as equações de "shares"(cfr. Altunbas, Liu, Molyneux e Seth(2000)) apenas por uma questão de simplicidade, não havendo nenhum suporte teórico para este procedimento.
[14] Mendes (1994), Kolari e Zardhoohi (1987)

Um primeiro refere-se ao facto desta especificação forçar um cariz homogéneo para o produto bancário: uma conta muito movimentada acarreta maiores custos que uma conta mais estável. Daí se ter incorporado na função custo, no seguimento de Sassenou (1992), além das variáveis tradicionais (produto, custos unitários dos factores de produção), uma variável de homogeneidade (quota de depósitos à ordem na totalidade do activo), com vista a diferenciar as várias instituições.

Uma segunda limitação da especificação Cobb-Douglas é a de não diferenciar os bancos quanto aos aspectos organizacionais. De facto, uma instituição de crédito pode aumentar a sua produção, através do crescimento do número dos seus balcões ou mantendo-o. A introdução de variáveis de estrutura como o número de balcões permitirá minimizar este problema.

A definição dos inputs apresenta algumas dificuldades.

Quanto ao factor de produção trabalho, é relativamente pacífico que o cômputo do custo unitário é feito através da relação entre a totalidade das despesas com o pessoal e o número de efectivos (dada a inexistência de informação quanto ao número de horas de trabalho).

No referente ao factor capital financeiro, vamos defini-lo como o conjunto dos saldos das contas de depósitos, seguindo trabalhos anteriores[15]. A totalidade dos depósitos pode ser segmentada em depósitos à ordem, a prazo e de poupança e acrescentada de outros recursos equiparados (débitos representados por títulos). O preço unitário do capital financeiro será resultante do quociente entre juros e custos equiparados e os depósitos totais e débitos representados por títulos.

Quanto ao capital físico, a sua definição é objecto de algumas leituras não coincidentes. No referente à banca, o produto bancário[16] é utilizado para remunerar o conjunto dos factores de produção: o capital físico, o capital financeiro e a mão-de-obra. Daí que alguns autores[17] proponham como preço unitário do capital físico a relação entre o custo do capital físico (amortizações adicionadas dos encargos e alugueres decorrentes de operações de leasing e similares, dada a inexistência de informação estatística acerca do preço por metro quadrado de cada estabelecimento) e o stock do factor (somatório das imobilizações líquidas, dos stocks de capital amortizado e das operações de leasing e similares). É esta a solução que se propõe no presente trabalho.

Em síntese, na especificação Cobb-Douglas, em relação ao modelo base, vão ser relevadas as seguintes variáveis:

CT ≡ Custos totais, englobando: os custos operacionais — CO (Custos com Pessoal e Outros Custos); os custos financeiros — CF (Juros e Custos Equiparados, Comissões e Prejuízos de Operações Financeiras); e as Amortizações do Exercício;

QCOBB ≡ Output, medida da produção bancária, composto pelo conjunto de Créditos sobre Instituições de Crédito, Créditos sobre Clientes e Aplicações em Títulos;

W1 ≡ preço unitário do factor trabalho, definido como a relação entre Custos com Pessoal e o Número de Trabalhadores;

W2 ≡ preço unitário do capital financeiro, definido como a relação entre Juros e Custos Equiparados e a soma de Débitos para com Instituições de Crédito, Débitos para com Clientes e Débitos Representados por Títulos;

W3 ≡ preço unitário do capital físico, definido como a relação entre Amortizações do Exercício e Imobilizações Brutas;

W4 ≡ preço unitário de outros inputs, definido como a soma de Prejuízos de Operações Financeiras, Comissões e Outros Custos de Exploração ponderada pelo Activo Líquido;

BALC ≡ Número de balcões (total doméstico) — Variável de estrutura ("a control variable to capture the influence on cost changes in the structure of the banking markets"[18]);

COTADV ≡ cota dos Depósitos à Vista no Activo Total.

Na especificação Fourier, e decorrente do que foi exposto a propósito da especificação Cobb-Douglas, vamos considerar as seguintes variáveis:

[15] Mendes e Rebelo (1997).

[16] Produto bancário = (juros e proveitos equiparados - juros e custos equiparados) + (rendimento de títulos+comissões+lucros em operações financeiras+outros proveitos de exploração) -(comissões+prejuízos em operações financeiras+impostos+outros custos de exploração)

[17] Sassenou (1992), pág.281

[18] Molyneux, Altunbas e Gardener (1996), p.206

$CT \equiv$ custo total, incluindo custos operacionais, custos financeiros e as amortizações do exercício.

Produtos:

CC ≡ Crédito sobre Clientes;

CIC ≡ Créditos sobre Instituições de Crédito;

APT ≡ Aplicações em Títulos.

Factores produtivos e respectivos preços unitários

1. Trabalho

$X_1 \equiv$ Número de trabalhadores;

$W_1 \equiv$ preço de X_1, definido como a relação entre despesas com pessoal e o número de trabalhadores.

2. Capital financeiro

$X_2 \equiv$ Saldo das contas de depósitos e débitos representados por títulos;

$W_2 \equiv$ Preço de X_2, dado pelo quociente entre juros e custos equiparados e os depósitos totais e débitos representados por títulos.

3. Capital físico

$X_3 \equiv$ Saldo das imobilizações líquidas e dos stocks de capital amortizado;

$W_3 \equiv$ Preço de X_3, dado pelo quociente entre as amortizações do exercício do factor e as imobilizações brutas.

4. Outros inputs

$X_4 \equiv$ Activo Líquido;

W4 ≡ preço unitário de outros inputs, definido como a soma de Prejuízos de Operações Financeiras, Comissões e Outros Custos de Exploração ponderada pelo Activo Líquido;

BALC ≡ Número de balcões (total doméstico).

Além destas variáveis vão ser considerados os "shares" para cada um dos inputs.

Na especificação Fourier acrescentaram-se as variáveis seno (siz) e coseno (csz) de primeira ordem, de forma a conseguir um melhor ajustamento do modelo aos dados.

Além destes modelos base, acrescentamos outras variáveis que permitem o estudo da concentração — variáveis dummy .

3.3. A Amostra

As estimações vão ter por base uma amostra de 22 bancos a operar em Portugal, entre 31 de Dezembro de 1995 e 31 de Dezembro de 2001, correspondendo a um total de 127 observações de registos anuais, após eliminação das observações para as quais não havia dados disponíveis (bancos sujeitos a fusão ou aquisição), tendo o número de observações inicial de 154 observações anuais (22 bancos durante 7 anos) sido reduzido. Os dados referem-se à actividade não consolidada dos bancos.

Os bancos incluídos na amostra são: Banco BPI (BPI), Banco Borges & Irmão (BBI), Banco Fonsecas & Burnay (BFB) e Banco de Fomento e Exterior (BFE) — Grupo BPI; Banco Comercial dos Açores (BCA) e BANIF — Grupo BANIF; Banco Santander (Portugal) (BSP), Banco Totta e Açores (BTA) e Crédito Predial Português (CPP) — Grupo Totta/Santander; Banco Internacional de Crédito (BIC) e Banco Espírito Santo (BES) — Grupo BES; Banco Comercial Português (BCP), Banco Pinto & Sotto Mayor (BPSM), Banco Português do Atlântico (BPA), Banco Mello (BMello) e Expresso Atlântico (BEXA) — Grupo Millennium BCP; Caixa Geral de Depósitos (CGD) e Banco Nacional Ultramarino (BNU) — Grupo CGD; Barclays Bank (BARC), Banque National de Paris (BNP), Deutsche Bank (DB) e Montepio Geral (MG) — outros. As instituições incluídas na amostra representavam cerca de 88% da

totalidade do Activo Líquido e de 81% da totalidade das Aplicações Creditícias das instituições de crédito que operavam em Portugal, em 31 de Dezembro de 2001[19].

3.4. Estimação da fronteira estocástica com dados em Painel

O processo de estimação vai basear-se, em grande parte, em modelos econométricos que interpretam a estrutura de dados em painel da amostra identificando o efeito individual (banco).

Nestes modelos, consideram-se os dados com dupla atribuição (dois índices). Os dados são relativos a cada uma de várias unidades seccionais (*cross-section*) observadas, durante vários períodos de tempo. No estudo, as diferentes unidades seccionais são os bancos (no total de 22 bancos) e os períodos de tempo são os anos (7 anos, de 1995 a 2001). Isto significa que para as variáveis observáveis presentes nos modelos se obtêm dados que são os valores numéricos que cada um dos 22 bancos regista, ao longo de 7 anos.

No estudo desta estrutura de dados, vai dar-se ênfase às unidades seccionais, no sentido de que se admitem comportamentos diferentes para as diferentes unidades seccionais, mas cujo comportamento, todavia, não variará com o tempo — é, neste aspecto, constante. Repare-se que o número de períodos é relativamente curto em relação ao número de bancos (que não é, contudo, um número elevado para efeitos estatísticos).

Consideremos a equação em que a variável dependente(Y) — no estudo trata-se do logaritmo (neperiano) do custo — é determinada por três grandes factores: o conjunto de variáveis explicativas (assumidas como determinísticas) agrupadas no vector X, que tem associado o vector de parâmetro β' (e que varia consoante as especificações são Cobb-Douglas, Translog ou Fourier); o efeito específico por cada unidade seccional (bancos) não observável (também designado por variável latente), α ; e o termo de perturbação estocástico relativo aos erros não observáveis, ε .

Assim o modelo econométrico é o seguinte:

$$(12)\, Y_{it} = \beta' X_{it} + \alpha_i + \varepsilon_{it}$$

Sendo, ainda :

$Y_{it} = Ln(CT_{it})$ ≡ logaritmo neperiano do custo total do banco i no ano t

X_{it} ≡ vector de variáveis explicativas, relativas ao banco i no ano t (tipicamente calculadas em logaritmos neperianos)

i= 1,2,….,n_b; o índice i representa cada uma das unidades seccionais, ou seja cada um dos bancos, que são no total 22 bancos ($n_b = 22$);

t= 1, 2, …7; o índice t representa o tempo, neste caso o ano a que diz respeito a informação; no estudo consideram-se 7 anos, de 1995 a 2001.

Nesta formulação está considerado o efeito específico não observável (ou latente) α_i que desempenha uma função importante na lógica do modelo e que tem consequência sobre os procedimentos de estimação. Note-se, pois, que este termo pode ser visto como substituindo os factores individuais característicos de cada banco que não estão disponíveis (como variáveis explicativas) ou não são directamente observáveis. Quer dizer que, ao ter este efeito específico em consideração e sob o ponto de vista econométrico, se evita assim um erro de omissão de variáveis explicativas relevantes, o que significa que a estimação convencional pelo método dos mínimos quadrados ordinário (MQO) não seria adequada, atendendo ao enviesamento, ineficiência e inconsistência dos estimadores de MQO, ao não considerar o efeito não observável.

Coloca-se, contudo, a questão de saber como incorporar este efeito específico latente, não observável, no modelo. Assumir-se-ão dois modelos: o modelo de efeitos fixos e o modelo de

[19] Cálculo com base nos dados consolidados do Boletim Informativo da Associação Portuguesa de Bancos, de Junho de 2002

efeitos aleatórios, como é típico na literatura especializada — veja-se, por exemplo, Greene (2003), Arellano (2003) e Johnston e DiNardo(1997).

Modelo de efeitos fixos (*Fixed Effects Model*)
Considere-se o seguinte modelo:

$$(13)\, Y_{it} = \alpha_i + \beta' X_{it} + \varepsilon_{it}$$

Como já foi dito, o modelo considera para cada unidade seccional (cada banco) um termo constante específico, α_i, que não varia com o tempo. Assume-se que o factor específico seccional α_i é correlacionado com as variáveis explicativas observáveis presentes na parte $\beta' X_{it}$.

Para a estimação do modelo, usam-se estimadores (presentes no software LIMDEP, versão 7 e versão 8) que são consistentes, conforme Greene (1995), Johnston (1997), Arellano (2003) e Greene (2003). Estes estimadores usam a técnica das variáveis binárias (dummy), para dar conta dos efeitos individuais das unidades seccionais. Note-se que a convencional utilização de estimadores de MQO não seria adequada, uma vez que os estimadores não são consistentes (nem cêntricos, nem eficientes), pois não têm em atenção os factores não observáveis, omitindo-os.

Repare-se, no entanto, que os estimadores para os efeitos individuais não podem ser consistentes, pois esta propriedade exige um aumento das observações, o que por definição significaria também aumentar, na mesma medida, o número de parâmetros (os efeitos individuais seccionais).

Modelo de efeitos aleatórios (*Random Effects Model*)
Considere-se a seguinte especificação:

$$(14)\, Y_{it} = \alpha + \beta' X_{it} + (u_i + \varepsilon_{it})$$

Sendo $\alpha_i = \alpha + u_i$, considera-se que o factor específico é a soma de uma constante representativa do valor médio referente às unidades seccionais e de um desvio específico por cada unidade seccional que não é correlacionado com as variáveis explicativas. Neste modelo, a especificidade das unidades seccionais (relembre-se que estas especificidades são não observáveis) é incluída no termo de perturbação – e não no termo independente como no modelo de efeitos fixos. O (novo) termo de perturbação passa a ser $v_{it} = u_i + \varepsilon_{it}$ assume-se que não é correlacionado com as variáveis explicativas presentes na parte $\beta' X_{it}$.

No processo de estimação, usam-se estimadores consistentes (presente no software LIMDEP – versões 7 e 8) conforme Greene (1995), Johnston (1997), Arellano (2003) e Greene (2003). A orientação, neste caso, é dar conta da heterogeneidade individual através do comportamento do termo de perturbação, ou seja, é usada a lógica presente nos estimadores de mínimos quadrados generalizados.

Modelo fronteira de custo estocástica (*Stochastic frontier model)*
Considere-se a seguinte especificação:

$$(15)\quad \begin{aligned} & Y_{it} = \alpha + \beta' X_{it} + (v_{it} + u_i) \\ & u_i = |U_i|; \quad U_i \approx N(0, \sigma_u^2); \quad v_{it} \approx N(0, \sigma_v^2); \quad \left[Var(u_i) = (1 - \frac{2}{\pi})\sigma_u^2 \right] \end{aligned}$$

O termo de perturbação é dado por $\varepsilon_{it} = v_{it} + u_i$, representado o termo de perturbação u_i o efeito específico da ineficiência de cada banco. A fronteira estocástica é dada por

$\alpha + \beta' X_{it} + v_{it}$ (o qualificativo estocástico é justificado pelo termo aleatório v_{it} na fronteira de eficiência de custo). Este modelo terá em consideração ainda os efeitos fixos ou aleatórios, como apresentado anteriormente, para dar conta da estrutura de dados em painel. Note-se que sendo u_i uma variável aleatória semi-normal positiva, obtida a partir da normal U_i, com média nula e variância σ_u^2, a sua variância é dada por $Var(u) = (1 - \frac{2}{\pi})\sigma_u^2$

Neste modelo, são relevantes a ineficiência para cada banco conforme definido por Jondrow, $E(u_i \mid \varepsilon_{it})$, e a proporção da variância devido à ineficiência

$$(16)\ \frac{Var(u)}{Var(\varepsilon)} = \frac{Var(u)}{Var(u) + Var(v)}.$$

O método de estimação utilizado é o da máxima verosimilhança (presente no software LIMDEP — versão 7 e versão 8) que sendo um método não linear permite ganhos de eficiência (Greene (2003)).

3.5. Resultados de Estimação

3.5.1. Análise da eficiência

3.5.1.1. A especificação Cobb-Douglas

Recorrendo ao conceito de fronteira estocástica com a especificação custo Cobb-Douglas, da estimação com dados em painel e efeitos aleatórios[20], pode concluir-se que existe ineficiência de cerca de 12%, que corresponde à capacidade de diminuição dos custos totais, para a média dos bancos da amostra, entre 1995 e 2001.

Quadro 1 — Resultados da estimação da fronteira estocástica

Variável	Coeficiente	Desvio-Padrão	Coeficiente/Desvio-Padrão	P($\|Z\| > z$)
Constante	1,040	0,157	6,622	0,000
Ln QCOBB	0,996	0,016	61,513	0,000
LnW1	-0,022	0,074	-0,292	0,770
LnW2	0,614	0,038	16,195	0,000
LnW3	0,092	0,033	2,806	0,005
LnW4	0,276	0,017	16,652	0,000
LnBALC	-0,009	0,016	-0,564	0,573
LnCOTADV	0,107	0,035	3,089	0,002

σ_u^2 =0,288 σ_v^2 =0,002	$Var(u)$ =0,105 $\frac{Var(u)}{\sigma_v^2}$ =49,837	$\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$ =0,539 $\lambda = \frac{\sigma_u}{\sigma_v}$ =11,711

Note-se que, a parte da variância devida à ineficiência, é muito superior (49,837) à variância da outra componente do erro total.

Quadro 2 — Medidas estatísticas descritivas da ineficiência

	Média	Máximo	Mínimo	Desvio-Padrão
Taxa de Ineficiência	12,2	115,3	1,0	15,8

[20] Os modelos de efeitos fixos não são convergentes

Se se analisar a eficiência por banco, obtemos os seguintes resultados:

Quadro 3 — Ineficiência por banco (média do período)

Banco	Taxa de Ineficiência (média)[21]	Banco	Taxa de Ineficiência (média)
CGD	42,34[22]	CPP	10,89
BNU	12,89	BSP	5,22
BCP	10,63	BANIF	14,06
BPA	4,19	BCA	8,34
BEXA	11,74	BFE	7,27
BPSM	8,02	BPI	29,81
BMELLO	21,07	BBI	12,41
BES	17,73	DB	13,71
BIC	5,30	BFB	6,41
BTA	3,88	MG	16,04
BARC	5,13	BNP	6,89

Se se testar a hipótese nula de as taxas médias de ineficiência serem iguais entre os bancos (contra a hipótese alternativa de elas serem diferentes) rejeita-se H_0 a 5% (estatística F de 2,593). Assim, diferentes bancos são, como esperado, diferentemente eficientes. O que confirma, agora nos modelos de ineficiência, a importância da modelização sugerida, em que se deu relevância aos efeitos individuais.

3.5.1.2. A Especificação Fourier

Os resultados da estimação para a especificação Fourier encontram-se no quadro seguinte.

Quadro 4 — Resultados de estimação da fronteira estocástica[23]

Variáveis	Coeficientes	Desvio-padrão	b/Desvio-padrão	|P[|T|>t]
Constante	31,117	15,130	2,057	0,040
LN2APT	0,030	0,029	1,003	0,316
LN2CC	0,468	0,172	2,711	0,007
LN2CIC	0,092	0,093	0,988	0,323
LNAPT	0,486	0,293	1,660	0,097
LNCC	-3,143	2,045	-1,537	0,124
LNCCAPT	-0,079	0,028	-2,827	0,005
LNCCCIC	-0,073	0,024	-3,080	0,002
LNCIC	-0,972	1,232	-0,789	0,430
LNCICAPT	0,032	0,024	1,362	0,173
LN2W1	-0,161	0,335	-0,482	0,630
LN2W2	0,276	0,156	1,769	0,077
LN2W3	-0,220	0,157	-1,400	0,162
LN2W4	0,210	0,013	15,701	0,000
LNW1	0,153	1,023	0,150	0,881
LNW1W2	0,057	0,198	0,289	0,773
LNW1W3	0,008	0,210	0,038	0,970
LNW1W4	0,009	0,069	0,128	0,899
LNW2	0,812	0,915	0,888	0,375

[21] Médias calculadas com base no quadro 35.

[22] A taxa de ineficiência média da CGD está muito influenciada por dois valores (dos anos de 1996 e 1997); se se recalcular a taxa de ineficiência média sem estes dois valores extremos obtém-se 15,51.

[23] Neste modelo não foram consideradas as variáveis associadas a BALC uma vez que não foi possível obter resultados com o modelo alargado (não há convergência no processo de estimação.

Quadro 4 (cont.)

Variáveis	Coeficientes	Desvio-padrão	b/Desvio-padrão	|P[|T|>t]
LNW2W3	0,095	0,093	1,022	0,307
LNW2W4	-0,151	0,062	-2,453	0,014
LNW3	-0,777	0,736	-1,056	0,291
LNW3W4	0,005	0,067	0,073	0,942
LNW4	0,840	0,275	3,057	0,002
LNAPTW1	-0,010	0,058	-0,179	0,858
LNAPTW2	0,331	0,047	0,709	0,478
LNAPTW3	-0,013	0,042	-0,300	0,764
LNAPTW4	-0,003	0,015	-0,175	0,861
LNCCW1	-0,122	0,123	-0,990	0,322
LNCCW2	0,056	0,055	1,015	0,310
LNCCW3	0,070	0,091	0,771	0,441
LNCCW4	-0,005	0,021	-0,234	0,815
LNCICW1	0,161	0,093	1,734	0,083
LNCICW2	-0,082	0,047	-1,736	0,083
LNCICW3	-0,026	0,064	-0,396	0,692
LNCICW4	-0,006	0,022	-0,253	0,800

Variáveis	Coeficientes	Desvio-padrão	b/Desvio-padrão	|P[|T|>t]
CSZAPT	0,083	0,345	0,240	0,811
CSZCC	-0,699	0,368	-1,898	0,058
CSZCIC	-0,010	0,162	-0,062	0,951
SIZAPT	0,025	0,254	0,100	0,921
SIZCC	-0,766	0,402	-1,905	0,057
SIZCIC	-0,092	0,086	-1,062	0,288

$\sigma_u^2 = 0{,}00282$ $\sigma_v^2 = 0{,}00166$	$Var(u) = 0{,}001025$ $\frac{Var(u)}{\sigma_v^2} = 0{,}617$	$\sigma = \sqrt{\sigma_u^2 + \sigma_v^2} = 0{,}0669$ $\lambda = \frac{\sigma_u}{\sigma_v} = 1{,}303$

Das duas partes que contribuem para a variância total a relativa à ineficiência é, apenas, 0,617 da outra fonte.

A estimação da fronteira estocástica com a especificação Fourier sugere a possibilidade de poupança de custos totais, a não haver ineficiência, de 4,2%, em termos médios — ineficiência estimada claramente inferior à que se obteve com a especificação Cobb-Douglas.

Quadro 5 — Medidas estatísticas descritivas da ineficiência

	Média	Máximo	Mínimo	Desvio-Padrão
Taxa de Ineficiência	4,2	13,4	1,0	2,1

Se se analisar a eficiência por banco, obtemos os seguintes resultados:

Quadro 6 — Ineficiência por banco (média do período)

Banco	Taxa de Ineficiência (média)	Banco	Taxa de Ineficiência (média)
CGD	4,66	**CPP**	3,46
BNU	3,34	**BSP**	4,31
BCP	7,61	**BANIF**	5,11
BPA	3,47	**BCA**	4,57
BEXA	3,62	**BFE**	3,55
BPSM	5,85	**BPI**	3,72
BMELLO	5,38	**BBI**	3,40
BES	3,72	**DB**	3,07
BIC	2,88	**BFB**	4,10
BTA	3,57	**MG**	3,64
BARC	4,39	**BNP**	3,95

Se se testar a hipótese nula de as taxas médias de ineficiência serem iguais entre os bancos (contra a hipótese alternativa de elas serem diferentes) rejeita-se H_0 a 5% (estatística F de 1,792).

Como se pode ver, a estimação das taxas de ineficiência é fortemente influenciada pela especificação escolhida. As 127 observações encontram-se seriadas para as duas especificações, no quadro seguinte, notando-se um grande afastamento dos valores estimados para a especificação Cobb-Douglas e para a especificação Fourier.

Quadro 7 — Taxas de ineficiência por função e por observação

Obs	TICB	TIF	Obs	TICB	TIF
1	4.31	4.43	66	6.86	4.73
2	103.47	5.29	67	5.30	4.51
3	115.34	5.12	68	5.65	7.92
4	4.38	3.24	69	4.94	2.26
5	26.79	5.77	70	5.75	2.23
6	19.85	4.81	71	6.67	4.14
7	22.22	3.93	72	3.48	2.90
8	14.11	5.25	73	4.75	6.18
9	9.64	3.29	74	12.21	6.40
10	7.75	2.53	75	18.68	4.81
11	8.14	2.13	76	16.27	4.16
12	14.18	3.11	77	13.74	3.48
13	23.50	3.75	78	18.72	5.09
14	19.37	12.0	79	12.37	5.25
15	13.46	13.4	80	6.40	6.55
16	18.19	13.2	81	6.75	2.19
17	6.19	5.50	82	12.97	5.62
18	4.05	3.53	83	10.13	3.69
19	7.65	2.96	84	7.98	5.89
20	5.48	2.53	85	7.96	6.32
21	4.66	4.07	86	4.25	3.68
22	4.90	5.75	87	12.47	4.14
23	4.99	3.48	88	4.45	3.28
24	2.82	2.10	89	4.88	3.22
25	3.57	1.95	90	83.76	5.89
26	3.60	3.72	91	16.50	2.55
27	2.53	3.31	92	5.57	3.07
28	29.43	3.77	93	13.39	3.38
29	9.97	4.13	94	9.35	2.44
30	13.16	3.19	95	4.84	1.86
31	3.96	6.55	96	23.03	5.90
32	10.47	7.61	97	16.80	5.12

Quadro 7 (cont.)

Obs	TICB	TIF	Obs	TICB	TIF
33	8.81	5.11	98	4.90	1.94
34	9.02	3.47	99	19.42	2.14
35	7.82	6.53	100	14.33	3.84
36	14.59	3.28	101	9.44	4.89
37	19.26	5.50	102	6.85	3.96
38	27.05	8.61	103	4.35	4.48
39	23.39	4.12	104	2.81	2.60
40	8.51	5.25	105	3.48	4.77
41	4.66	4.53	106	3.64	4.19
42	6.68	3.79	107	8.94	8.83
43	27.42	2.45	108	8.17	2.88
44	27.12	3.94	109	35.92	2.54
45	20.17	2.63	110	12.05	3.50
46	29.56	3.44	111	8.81	2.18
47	4.34	1.52	112	3.64	3.75
48	15.56	4.29	113	34.77	1.77
49	7.12	3.25	114	2.71	6.57
50	2.92	2.33	115	7.07	4.56
51	2.21	2.31	116	13.01	3.47
52	1.58	2.48	117	5.86	4.56
53	3.34	3.97	118	3.65	2.37
54	3.88	3.39	119	2.06	3.02
55	3.68	2.40	120	1.53	6.17
56	0.98	1.01	121	10.49	2.21
57	5.75	2.36	122	4.70	2.36
58	4.45	2.70	123	5.40	4.79
Obs	TICB	TIF	Obs	TICB	TIF
59	4.32	5.71	124	9.58	4.36
60	4.08	7.44	125	2.98	1.75
61	13.56	3.38	126	4.00	2.76
62	16.33	3.72	127	11.06	9.43
63	12.27	4.05			
64	7.64	2.76			
65	8.68	2.12			

TICB ≡ Taxa de Ineficiência calculada para a função Cobb-Douglas
TIF ≡ Taxa de Ineficiência calculada para a função Fourier
Observações: 1-7 (CGD); 8-13 (BNU); 14-20 (BCP); 21-25 (BPA); 26-30 (BEXA); 31-35 (BPSM); 36-39 (BMELLO); 40-46 (BES); 47-53 (BIC); 54-60 (BTA); 61-66 (CPP); 67-73 (BSP); 74-80 (BANIF); 81-86 (BCA); 87-89 (BFE); 90-93 (BPI); 94-96 (BBI); 97-99 (BFB); 100-106 (MG); 107-113 (DB); 114-120 (BARC); 121-127 (BNP).

3.5.2. Análise da Concentração

3.5.2.1. A especificação Cobb-Douglas

Considerando a função custo Cobb-Douglas e a estimação com dados em painel e efeitos fixos, a variável que capta os efeitos da aquisição ou fusão sobre os custos após um ano — variável binária D1DM — revela-se estatisticamente significativa, o mesmo acontecendo com os efeitos, dois e três anos após a concentração (variáveis binárias D2DM e D3DM).

Quadro 8 — Resultados da estimação da concentração (dados em painel, efeitos fixos)

Variável	Coeficiente	Desvio-Padrão	t-Ratio	P($\|T\| > t$)
Ln QCOBB	0,907	0,063	14,318	0,000
LnW1	0,026	0,052	0,508	0,613
LnW2	0,547	0,049	11,060	0,000
LnW3	0,073	0,058	1,251	0,214
LnW4	0,354	0,022	15,859	0,000
LnBALC	0,037	0,081	0,460	0,647
LnCOTADV	0,040	0,057	0,703	0,483
DZM	-0,003	0,117	-0,023	0,982
D1DM	-0,314	0,148	-2,115	0,037
D2DM	-0,603	0,213	-2,824	0,006
D3DM	-0,498	0,213	-2,340	0,021

F(31,95) =419,28 (*p-value*=0,000)

Efeitos fixos estimados

Grupo	Coeficiente	Desvio-padrão	t-ratio
1	2,308	0,814	2,834
2	2,033	0,735	2,766
3	2,139	0,779	2,746
4	1,947	0,785	2,479
5	1,676	0,541	3,099
6	2,003	0,737	2,717
7	2,060	0,710	2,903
8	2,075	0,790	2,628
9	1,30	0,702	2,747
10	1,908	0,768	2,484
11	2,034	0,718	2,833
12	1,872	0,670	2,793
13	2,062	0,685	3,008
14	1,909	0,611	3,123
Grupo	**Coeficiente**	**Desvio-padrão**	**t-ratio**
15	2,020	0,728	2,775
16	2,576	0,731	3,524
17	2,085	0,722	2,889
18	2,096	0,710	2,954
19	1,973	0,709	2,784
20	2,113	0,747	2,828
21	1,799	0,660	2,726
22	2,012	0,625	3,219

As reduções de custos estimadas são de 31%, 60% e 50%, um, dois e três anos, respectivamente, após a fusão e em comparação com os bancos que não fizeram a fusão. Em relação à redução de custos operada, no próprio ano da fusão, o coeficiente estimado é reduzido e negativo (-0,3%) e não é estatisticamente significativo.

De notar que estas estimativas devem ser avaliadas com precaução, pois são calculadas com um número de casos (observações) muito reduzido, sobretudo 2 e 3 anos após a fusão, e sofrem ainda da interferência individual dos bancos ao serem considerados os efeitos fixos (que captam a especificidade de cada um dos 22 bancos). Estes efeitos fixos são, agora, positivos (ao contrário de resultados anteriores), a que não será estranha a presença de variáveis binárias relativas à concentração; mas mantêm-se estatisticamente significativos quer ao nível individual quer ao nível global (neste aspecto, resultado idêntico ao da análise efectuada no ponto relativo às economias de escala e de gama, com a função custo Cobb-Douglas, mas sem a consideração

das variáveis *dummy* relativas às fusões – DZM, D1DM, D2DM e D3DM - como variáveis explicativas dos custos.

A utilização de um modelo idêntico, mas em que se admitem efeitos aleatórios, permite obter resultados semelhantes aos, agora, indicados.

Quadro 9 — Resultados da estimação da concentração (dados em painel, efeitos aleatórios)

Variável	Coeficiente	Desvio-Padrão	t-Ratio	P($\|T\| > t$)
Ln QCOBB	0,969	0,026	36,587	0,000
LnW1	0,111	0,072	1,545	0,122
LnW2	0,602	0,044	13,688	0,000
LnW3	0,126	0,054	2,340	0,019
LnW4	0,348	0,018	19,528	0,000
LnBALC	0,029	0,026	1,094	0,274
LnCOTADV	-0,001	0,044	-0,012	0,990
DZM	0,059	0,090	0,656	0,512
D1DM	-0,167	0,104	-1,615	0,106
D2DM	-0,345	0,149	-2,318	0,021
D3DM	-0,251	0,149	-1,677	0,094
Constante	1,271	0,287	4,422	0,000

Procedeu-se, ainda, à análise da influência de grupo sobre os custos (usando variáveis *dummy*, com o prefixo DG) para os 6 principais grupos bancários portugueses – CGD, Millennium BCP (BCP), BES, Santander/Totta (BS), BPI e BANIF — recorrendo à especificação Cobb-Douglas e a um modelo de efeitos aleatórios (a influência de grupo sobre os custos já foi contemplada nos efeitos fixos). Resulta um modelo estimado em que os coeficientes estatisticamente significativos (a 10%) são os dos grupos CGD, BPI e BANIF, com custos estimados, superiores aos dos restantes bancos, de respectivamente 17,9%, 11% e 12%.

Quadro 10 — Resultados da estimação da concentração, por grupos (dados em painel, efeitos aleatórios)

Variável	Coeficiente	Desvio-Padrão	b/desvio-padrão	P($\|T\| > t$)
Ln QCOBB	0,967	0,023	41,141	0,000
LnW1	0,096	0,077	1,251	0,211
LnW2	0,579	0,046	12,451	0,000
LnW3	0,140	0,055	2,569	0,010
LnW4	0,344	0,019	18,311	0,000
LnBALC	0,021	0,022	0,654	0,340
LnCOTADV	-0,025	0,042	-0,596	0,551
DGCGD	0,179	0,068	2,628	0,009
DGBPI	0,110	0,058	1,907	0,057
DGBANIF	0,120	0,068	1,757	0,079
Constante	1,226	0,242	5,057	0,000

3.5.2.2. A especificação Fourier

Considerando a função Fourier anteriormente apresentada, mas com dados não escalonados, e tendo-se acrescentado, em modelos distintos, uma variável *dummy* — para medir o efeito das fusões e aquisições – ou quatro outras variáveis *dummy* que consideram os efeitos das fusões e aquisições respectivamente no próprio ano em que se realizaram, 1 ano, 2 anos e 3 anos após — conclui-se, em ambos os casos, que as variáveis adicionadas não têm significância estatística, pelo que não se detectam efeitos significativos das fusões e aquisições sobre os custos, quando se admite como especificação a função Fourier (Ribeiro, 2006). A especificação desta função (com um número importante de variáveis explicativas adicionais) e a consideração dos efeitos individuais por grupo bancário são uma possível explicação para esta diferença em relação ao modelo Cobb-Douglas.

CONCLUSÃO

Neste estudo procedeu-se ao cômputo da ineficiência produtiva e à análise dos efeitos da concentração sobre os custos bancários portugueses.

Foram adoptadas duas formas funcionais de funções custo bancária — a especificação Cobb-Douglas e a especificação Fourier — tendo por base dados em painel, para os anos de 1995 a 2001, comportando 22 instituições bancárias. Dado o carácter multiproduto da empresa bancária introduziram-se variáveis de ajustamento (variáveis de estrutura e de homogeneidade) na especificação Cobb-Douglas.

A estimação da ineficiência, recorrendo à especificação Cobb-Douglas (dados em painel, efeitos aleatórios), gera um valor elevado de quase 12%, em média, para o período de 1995 a 2001. Este valor é claramente superior ao que se estimou por via da especificação Fourier (dados em painel, efeitos aleatórios) — taxas de ineficiência média, para o período amostral, de cerca de 4%. De toda a forma, conclui-se ser possível que as instituições portuguesas reduzam os seus custos totais, diminuindo a ineficiência-X.

A análise dos efeitos da concentração ensaiada neste estudo, recorrendo a variáveis binárias para captar os efeitos no próprio ano, um ano após, dois anos e três anos após a fusão (uma vez que a literatura é unânime na afirmação de que os efeitos da concentração, a haver, se diluem após três anos) sugerem efeitos na redução de custos muito elevados na especificação Cobb-Douglas (dados em painel, efeitos fixos e efeitos aleatórios), em particular nos grupos CGD, BPI e BANIF. Já recorrendo à especificação Fourier as estimativas não sugerem quaisquer efeitos significativos do processo de fusões e aquisições sobre os custos (quer globalmente, quer repartindo os possíveis efeitos entre o próprio ano da concentração e os três anos subsequentes). Esta contradição dos resultados, consoante a especificação adoptada, pode estar relacionada com a própria amostra: o número muito reduzido de observações após concentração (em particular no período de dois e de três anos após a concentração).

BIBLIOGRAFIA

AIGNER, DENNIS; LOVELL, KNOX C.A. E SCHMIDT, PETER (1977), "Formulation and estimation of stochastic frontier production function models", Journal of Econometrics, 6, pp. 21-37

ALTUNBAS, Y., GARDENER, E.P.M., MOLYNEUX, P. E MOORE, B. (2001), "Efficiency in European banking", European Economic Review, 45, pp. 1931-1955

ALTUNBAS, YENER, LIU, MING-HAU, MOLYNEUX, PHILIP E SETH, RAMA (2000), "Efficiency and risk in Japonese Banking", Journal of Banking & Finance, 24, 1605-1628

ARELLANO, MANUEL(2003), Panel Data Econometrics, Oxford University Press

BARROS, PEDRO PITA (1996), "Approval rules for sequential horizontal mergers", Working Paper nº 287, Faculdade de Economia da Universidade Nova de Lisboa

BERGER, A.N. (1995), "The profit-structure relationship in banking — tests of market-power and efficient-structure hypothesis", Journal of Money, Credit and Banking, 27(2), pp. 404-431

BERGER, A., W.C. HUNTER E S.G. TIMME (1993), "The efficiency of financial institutions: a review and preview of research past, present and future", Journal of Banking and Finance, 17 (2-3), pp. 221-50

BERGER, A.N., HUMPHREY, D.B.(1992), "Megamergers in banking and the use of cost efficiency as an antitrust defence", Antitrust Bulletins, 37, pp. 541-600

BERGER, ALLEN N., DEMSETZ, REBECA E STRAHAN, PHILIP E. (1999), "The consolidation of financial services industry: Causes, consequences, and implications for the future", Journal of Banking & Finance, vol. 23, pp. 135-194

BERGER, ALLEN N., LEUSNER, JOHN H. E MINGO, JOHN J. (1997), "The efficiency of bank branches", Journal of Monetary Economics, 40, pp. 141-162

CUESTA, RAFAEL A. E OREA, LUÍS (2002), "Mergers and technical efficiency in Spanish saving banks: A stochastic distance function approach", Journal of Banking & Finance, 26, pp. 2231-2247

FOCARELLI, D., PANETTA, F., SALLEO, C. (1998), "Why do banks merge: Some empirical evidence from Italy", Journal of Money, Credit, and Banking, In Press

GALLANT, A.R.(1981), "On the bias in flexible functional forms and essentially unbiased form: The Fourier flexible form", Journal of Econometrics, 15, pp. 211-245

GALLANT, A.R.(1982), "Unbiased determination of production technologies", Journal of Econometrics, 20, pp. 285-324

GREENE, W.M. (1990),"A gamma-distributed stochastic frontier model", Journal of Econometrics, 46, pp. 141-163
GREENE, W. H. (1993), Econometric, Second Edition, Analysis, MacMillan, New York, (cap.16)
GREENE, W.M. (1993), "The econometric approach to efficiency analysis", in H.O.Fried, C.A.Lovell e P. Schmidt (eds.), The Measurement of Productive Efficiency: Techniques and Applications, Oxford University Press
GREENE, W.M. (2003), Econometric Analysis, Prentice Hall, 5th edit.
GREENE, W.M. (1995), LIMDEP — version 7 — User's Manual , Econometric Software, Inc.
HAYNES, MICHELLE E THOMPSON, STEVE (1999), "The productivity effects of bank mergers: Evidence from the UK building societies", Journal of Banking & Finance, 23, pp. 825-846
HIRSHLEIFER, DAVID (1993), Takeovers, in P. Newman, M. Milgate e J. Eatwell (eds), The New Palgrave Dictionary of Money and Finance, New York, Stockton Press
HUMPHREY, DAVID B. E VALE, BENT (2003), "Scale economies, bank mergers, and electronic payments: A spline function approach", Journal of Banking & Finance (article in press)
JOHNSTON, J., DINARDO, J. (1997), Econometric Methods, McGraw Hill, 4td Edition
JONDROW,J., C.A. LOVELL, I.S. MATEROV, SCHMIDT (1982), "On estimation of thecnical inefficiency in the stochastic frontier production model", Journal of Econometrics, 19, pp.233-38
KOLARY, J., ZARDKOOHI, A. (1987), Bank Cost, Struture and Performance, Lexington Books
KUMBHAKAR, SUBAL C., LOVELL, C. A. KNOX (2003), Stochastic Frontier Analysis, Cambridge University Press
LEIBENSTEIN, H. (1966), "Allocative efficiency vs X-efficiency", American Economic Review, 56, pp. 392-15
MENDES, VICTOR (1994), "Eficiência produtiva no sector bancário: uma aplicação do método DEA aos anos 1990-92", Trabalhos em Curso, 42, Faculdade de Economia da Universidade do Porto
MENDES, VICTOR E REBELO, JOÃO (2001), "The effect of bank M&AS on efficiency: the Portuguese experience", in What Financial System for the Year 2000?, Margarida Abreu e Victor Mendes (org.), Ed. BVLP — Soc. Gestora de Mercados Regulamentados, S.A., pp.277-294
MENDES, VICTOR E REBELO, JOÃO (1997), "Relações custo-produção e eficiência produtiva no sistema integrado de crédito agrícola mútuo no período 1990-1995", Estudos de Economia, Vol. XVI-XVII, nº2, pp.117-35
MENDES, VICTOR E REBELO, JOÃO (1999), "Productive efficiency, technological change and productivity in Portuguese banking", Applied Financial Economics, 9, pp. 513-521
MOLYNEUX, P., ALTUNBAS, Y. E GARDENER, E. (1996), Efficiency in European banking, Ed. John Willey & Sons, Lda, England
MOLYNEUX, PHILIP (2001), "Does Size matter? Financial Restructuring under EMU", in What Financial System for the Year 2000?, Margarida Abreu e Victor Mendes (org.), Ed. BVLP — Soc. Gestora de Mercados Regulamentados, S.A., pp.229-250
PERISTIANI, S. (1997), "Do mergers improve the X-efficiency and scale efficiency of U.S. banks? Evidence from the 1980s", Journal of Money, Credit and Banking, vol.29 (3), pp. 326-337
RESTI, ANDREA (1997), "Evaluating the cost-efficiency of the italian banking system: what can be learned from the joint application of parametric and non-parametric techniques", Journal of Banking & Finance, 21, pp. 221-50
RESTI, ANDREA (1998), "Regulation can foster mergers, can mergers foster efficiency? The Italian case", Journal of Economics and Business, 50, pp. 157-169
RESTI, ANDREA (2000), "Efficiency measurement for multi-product industries: A comparison of classic and recent techniques based on simulated data", European Journal of Operational Research, 121, pp. 559-578
RIBEIRO, M.C.P. (2006), "Economias de Escala e de Gama e os efeitos da concentração na eficiência bancária", Tese doutoramento, Universidade do Minho
SASSENOU, MOHAMED (1992), "Economies des coûts dans les banques et les caisses d'épargne, impact de la taille et de la variété de produits", Revue Économique, vol.43, nº2, pp. 277-50
SRINIVASAN, A., Wall, L.D. (1992), "Cost saving associated with bank mergers", Working Paper, Federal Reserve Bank of Atlanta, Atlanta
TOLSTOV, G.P. (1962), Fourier Séries, Prentice-Hall, London

Outras Publicações:

ASSOCIAÇÃO PORTUGUESA DE BANCOS, Boletim Informativo (vários números)
BANCO CENTRAL EUROPEU (2000), "Fusões e Aquisições envolvendo o sector bancário da União Europeia — Factos e implicações", Frankfurt, 20 de Dezembro
RELATÓRIOS E CONTAS DE INSTITUIÇÕES DE CRÉDITO (1995-2001)

DETERMINANTS OF ENTRY AND EXIT IN PORTUGUESE REGIONS

CLÁUDIA REIS DUARTE
E-mail: claudia_duart3@portugalmail.pt
ANTÓNIO CARDOSO MARQUES
E-mail: amarques@ubi.pt
Universidade da Beira Interior

Abstract

In the literature regarding the entry and exit of firms, there is growing emphasis on the interrelationship between these two phenomena. This article studies the determinants of the entry and exit of firms by region in Portugal. We contribute the following points to the literature: the control for new factors in entry/exit regional dynamics; the use of both static and dynamic panel models; the study of a country with no administrative decentralisation. Our results show that each region has its own specificities and entry and exit are both persistent. The results support the multiplier effect for entry and the competition effect for exit. We conclude that both the minimum wage and the unemployment rate discourage entry, while population density does not influence a firm's decision to leave the market. The agglomeration factor is relevant in explaining exit and contributes negatively towards this decision occurring.

Keywords: Entry, Exit, Panel data, Portuguese regions.
Thematic Area: Industrial Economics.

Resumo

Neste trabalho são estudados os determinantes da entrada e saída de empresas por região em Portugal. Contribuímos para a literatura existente com: o controlo para novos factores na dinâmica regional de entrada e saída; com o uso de estimadores de painel estáticos e dinâmicos; com o estudo de um país sem descentralização administrativa. Os resultados mostram que existem especificidades por região, bem como mostram que a entrada e a saída são fenómenos persistentes. Concluímos que o estudo da entrada/saída sem considerar como variável explicativa a saída/entrada pode criar estimativas inconsistentes. Os resultados sugerem que o salário mínimo e a taxa de desemprego desincentivam a entrada ao mesmo tempo que a densidade populacional não influencia a decisão das empresas de sair do mercado. O factor de aglomeração é relevante e contribui negativamente para a taxa de saída.

Palavras-chave: entrada e saída de empresas, dados em painel, regiões de Portugal.
Áreas temáticas: Economia Industrial.

1. INTRODUCTION

This paper is part of the literature on the determinants of entry and exit in a regional context. The interest in analysing regional aspects in the setting up of companies, as well as in their exit, stems from the fact shown in the literature that economic growth is stimulated by entrepreneurship and by the setting up of companies, which are fundamental factors for creating the wealth and well-being of economic agents.

Recently, attention has been paid to the spatial dimension in the entry and exit dynamic of companies in various countries and regions, such as in Western Germany (Brixy and Grotz, 2007), Sweden (Nyström, 2007c) and Italy (Carree *et al*, 2008). In Portugal, authors that have studied the entry and exit dynamic of companies have been essentially José Mata, Pedro Portugal, Paulo Guimarães and Octávio Figueiredo. The first two have done so from the point of

view of how many companies go into and out of business (e.g. Mata and Portugal, 2004) while the others have focused on the decision for company location in terms of industrial district (e.g. Guimarães *et al*., 2004). To the best of our knowledge, no study has yet been made that gives a full picture of the patterns of entry and exit per NUTS III region in Portugal.

The purpose of this paper is to provide empirical evidence on the relationship between regional entry and exit rates in a spatial context of a country with no administrative and political decentralisation, as is the case of Portugal. Since the distribution of companies per region in this country is known to be heterogeneous, it is important to understand how regions and their decision-makers can intervene by promoting and enhancing the attractiveness of each geographical area. Our aim is then to understand the determinants that explain a company's decision to enter or exit. Methodologically, this paper contributes to the empirical literature by applying both static and dynamic panel data methods.

In section 2 of this paper, a short review of the relevant literature is given. Section 3 presents the method and the variables used in the empirical analysis. The results of both the static and dynamic estimators are presented in the following section. The paper concludes with section 5, in which the results are discussed and the main conclusions drawn.

2. MOTIVATION

The literature and empirical work are paying increasing attention to the interrelationship between entry and exit. Johnson and Parker (1996) relate entries and exits in the light of the multiplier effect and the competition effect. The former, the multiplier effect, means that entries cause the entry of other companies, resulting from two effects: the demonstration effect and the income effect. The demonstration effect is related to the theories of clusters and occurs because the entry of a new company encourages others to do the same. In this way, companies can benefit from human capital by sharing the market of qualified workers, through access to local production inputs and through the effects of knowledge spillovers. The income effect presents a more direct incentive mechanism since the entry of companies can create additional income for entrepreneurs and their employees. The competition effect suggests that companies on the market cause exits in the future. This effect may occur because entry causes more competition, excess capacity on the market or because the entering companies are more efficient or produce better quality products.

In addition to the interrelation between entry and exit, regional determinants have always merited attention in the literature. Authors such as Johnson and Parker (1996), Hicks and Sutaria (2004), Maria *et al.* (2005) or Nyström (2007b) have worked on the three main determinants that influence the different spatial dimensions: local demand, entrepreneurship and the agglomeration of existing companies. These three categories of factors serve to reflect different perspectives in the process of setting up new companies or in their winding-up. They prove to be very pertinent to an entrepreneur's decision about whether to locate his/her company in a certain region. In analysing the conditions of local demand, the literature fundamentally considers these variables: population and population density (e.g. Nyström, 2007b); GDP and GDP *per capita* (e.g. Maria *et al.* (2005), Arauzo *et al* (2007) and Nyström (2007b); salaries (e.g. Reynolds *et al*., 1994 and Lay, 2003) and productivity (e.g. Dunne *et al*, 2005). Changes in population imply alterations in the determinants in the choice of a certain region (Nyström, 2007b), while Gross Domestic Product (GDP) and GDP *percapita* measure economic activity and may be a factor in attracting new investors. Nyström (2007b) suggests that high levels of demand and growing revenue make it easier for companies to survive, so the exit rate will fall.

Similarly to Dunne *et al.* (2005), we use productivity as a way of explaining the exit ratio. High levels of income increase demand while one source of this income (the salary) constitutes company costs. These costs could be synonymous with exit for a company with intensive labour. This variable may also be considered in the factor that follows – entrepreneurship –

since this influences the decision of an individual to create his/her own work post or to occupy an existing one, because the individual compares the salary from being an employee with the expected payoff from opening his/her own business.

In entrepreneurship, the literature has fundamentally used these variables: employment (e.g. Audrestch and Fritisch, 1999 and Carree and Dejardin, 2007); schooling (e.g. Evans and Leighton, 1990 and Delmar *et al*, 2003) and interest rate (e.g. Dunne, *et al.*, 2005). With regard to employment, there is no agreement in the literature on the nature of its influence. On the one hand, Storey (1991), Audretsch and Fritsch (1999) and Carree and Dejardin (2007) sustain that high rates of unemployment may cause a greater exit ratio since the unemployed tend to create their own employment. On the other hand, there may be a negative relationship between high rates of unemployment and the entry of companies onto the market, based on the fact that employment is an indicator of the region's economic situation. In this way, if the economic situation deteriorates, there may be less demand for employment, since workers will try to find better opportunities outside the region.

With regard to education, there is a long list of authors who have analysed its impact on the levels of entrepreneurship. The theory suggests that people with a higher level of education are better prepared to make the most of new opportunities (business advantages). Delmar *et al* (2003) concluded that in Sweden, most people with higher levels of academic qualifications would rather be employees than create their own business. The justification lies in the fact that higher income is generated as an employee than as an entrepreneur. They also noticed that the number of companies set up by this group of society was no higher than that of any other group. It may also be related to the fact that one region has a more qualified labour force than another. This suggests that there is a positive relationship between the level of schooling in a certain region and the entry of companies onto the market.

As for the income earned as an employee being higher than that of being one's own boss, it is expected that if there is an opportunity to find a job with someone else, the entrepreneur will decide to close the firm. With regard to the population's level of education, people with higher qualifications are expected to have greater capacity to assess the future and recognise a more viable idea when they are faced with a situation of the "death" of a firm. In Portugal, however, the educational level of entrepreneurs and the population in general is relatively low, which may affect the relationship between education and entrepreneurship.

It is common for capital to be seen both as a barrier to entry and as a cause of exit. According to Mata (1991), this is due to imperfections on the capital market. The author explains that new entrants may well have great difficulty in obtaining funding for their projects. This means that there are many possible entrepreneurs that never become entrepreneurs because they lack the financial capability. It therefore makes sense to assess how the reference interest rate of loans can influence the entry and exit of companies in the region. Storey (1982), for example, was quick to mention this evidence, showing that future entrepreneurs often mortgage their own properties at banks so as to obtain the loan to set up their business. The reason for exit of entrepreneurs is often the consequence of decisions taken by the creditors themselves.

Seeing as the purpose of this work is to analyse the density and diversity of a region instead of a particular industry, we will only consider the variable urbanisation economies as an agglomeration factor. We aim to assess the existence of positive externalities of the location such as lower transport costs and greater proximity of suppliers and customers, which reduce costs and improve the quality of the services provided. Depending on the region's population density, negative externalities will also occur, such as traffic congestion, an increase in rent price and salary increases. For companies that are already established, the positive effects of urbanisation can increase their chances of survival, while the negative effects may lead to their exiting the market. According to Krugman (1998), there may be a cause-effect whenever new companies are set up in or move to a certain region, because conditions of greater demand for goods and services, better accesses and more human capital will be created.

3. METHOD AND VARIABLES

The database used was constructed by gathering information from several sources, namely the National Institute of Statistics Directory, publications from the Ministry of Employment and Social Security and Bank of Portugal statistics. The data refer to the period from 1999 to 2005 and are aggregated through the NUTS III system. This system consists of 30 units, 28 in continental Portugal and 2 in the Autonomous Regions of the Azores and Madeira. Because of the geographical characteristics of the islands, we will only consider 28 units of observation in continental Portugal. Although some literature focuses on each sector individually (e.g. Brixy and Grotz, 2007), in this paper we chose to aggregate all the sectors. This is because the only way we could have a long enough period of time to undertake dynamic panel data analysis was by considering the aggregate information. Moreover, information is not always available for each separate NUTS III, and the aim of this work is based on the analysis of the capacity to attract companies from each of the 28 regions in continental Portugal.

The method of measuring the number of entries and/or exits is not consistent throughout the literature. For Geroski (1991), entries and exits can be measured through net or gross entries. Sustenance is in the volume of sales or in the number of companies. Other forms of measurement are also traditionally used. Authors such as Baldwin and Gorecki (1987), Günalp and Cilasun (2006) used the number of entrants to measure entry. Storey and Jones (1987) and Acs and Audretsch (1989) used market share, while Dunne, *et al.* (2005), Berglund and Brannas (2001) used a logarithmic transformation of entry. Like us, Love (1996), Acs and Storey (2004) and Nyström (2007b) used rate of entry, or entry ratio, which is the number of entrants divided by the initial number of established companies, in line with the ecological approach.

Our aim is to understand the factors that influence the entry and exit of companies per region in Portugal, using the ratios of entry and exit as variables. As explanatory variables, we controlled for various factors, aggregated in the main determinants summarised in table 1.

TABLE 1 – Definition of variables and expected signs

Variable	Definition	Expected sign
$Entry_{i,t}$ $E_{i,t}$	Entry rate: number of firms entering the region in relation to total number of firms in the region in year t	
$Exit_{i,t}$ $S_{i,t}$	Exit rate: number of firms exiting the region in relation to total number of firms in the region in year t	
Local Demand		
$POP_{i,t}$ $P_{i,t}$	Population: population of region i in year t	Entry: + Exit: -
$POPDEN_{i,t}$ $PD_{i,t}$	Population Density: number of inhabitants per km² in year t	Entry: + Exit: -
$GDP_{i,t}$	Gross Domestic Product: Gross Domestic Product in region i in year t	Entry: + Exit: -
$GDPpercap_{i,t}$ $GDPpc_{i,t}$	Gross Domestic Product *per capita*: Gross Domestic Product in region i in year t divided by the population	Entry: + Exit: -
$PRODUC_{i,t}$ $Pd_{i,t}$	Productivity: Gross added value of region i at the beginning of year t divided by the number of employees of region i in year t.	Entry: + Exit: -
$WAGE_t$ W_t	National minimum wage in year t	Entry: +/- Exit: +/-
Entrepreneurship		
$UNEMPR_{i,t}$ $UR_{i,t}$	Unemployment rate: number of unemployed in region i in year t divided by the active population of region i in year t.	Entry: +/- Exit: +/-
$EDUELE_{i,t}$ $EE_{i,t}$	Proxy to basic level education: number of students at basic level in region i divided by the active population in region in year t.	Entry: - Exit: +
$EDUSECD_{i,t}$ $ES_{i,t}$	Proxy to secondary level education: number of students at secondary level in region i divided by the active population in the region in the year t.	Entry: +/- Exit: +/-
$EDUUNI_{i,t}$ $EU_{i,t}$	Proxy to higher level education: number of students at higher level in region i divided by the active population in the region in the year t.	Entry: +/- Exit: +/-
$INTRATE_t$ IR_t	Interest rate: weighted average rate relative to new loan operations denominated in euros, granted by banks to residents in the euro zone	Entry: - Exit: -
Agglomeration		
$URB_{i,t}$ $U_{i,t}$	Urbanisation economies: number of companies from region i divided by the number of inhabitants of region *i*	Entry: + Exit: -

In addition to indicating the variables used, this table also shows how they were obtained and the expected signs. Table A.1, presented in Appendix A, shows the descriptive statistics of the information used. As expected there is a positive correlation between the rate of entry and the rate of exit (0.7442). We use the STATA data analysis and statistical software package.

3.1 Analysis

The data collected is panel data, combining cross-section and time-series. In this way, suitable econometric techniques were used to deal with this type of data. With few exceptions, such as Nyström (2007a), panel data has been used very little by the literature. Unlike the author that uses only the Arellano and Bond (1991) estimator, in this paper we use the Blundell and Bond (1998) estimator, which extends the GMM (1991) estimator by proposing a dynamic estimator. This simultaneously uses the variables in the equation to estimate in levels and first differences, and the same occurs with the instruments used.

The panel data suggests that there are different characteristics between the individual units (regions), which may or may not be unchangeable over time, in such a way that time or sectional studies that do not take into account such heterogeneity will almost always produce highly biased results. On the other hand, panel data predicts more information, greater variability of data, less collinearity between the variables, a higher number of degrees of freedom and more efficiency in the estimate[1] (Greene, 2003). The use of both static and dynamic estimators allows the specific, individual and unobservable effects of regions to be eliminated as well as controlling the possible endogeneity between explicative variables, since their delayed values are used as instruments.

We test and compare the results from both static and dynamic estimators. Under the static estimator framework, these region-specific effects may be assumed to be either random or fixed. On the one hand, if the individual specific effects are assumed as random, then they are not correlated with the other explanatory variables. On the other hand, in the fixed effects model, the region-specific effects are assumed to be correlated with the explanatory variables.

The model for entry, without exit as an explanatory variable is:

$$\begin{aligned} E_{i,t} &= \alpha + \beta_1 P_{i,t} + \beta_2 PD_{i,t} + \beta_3 GDP_{i,t} + \beta_4 GDPpc_{i,t} + \beta_5 W_t + \beta_6 UR_{i,t} + \\ &+ \beta_7 EE_{i,t} + \beta_8 ES_{i,t} + \beta_9 EU_{i,t} + \beta_{10} IR_t + \beta_{11} U_t + d_t + v_{i,t} + e_{i,t} \end{aligned} , \qquad (1)$$

where *i* represents each of the regions, *t* represents the year, d_t is a time variable that measures the possible impact of time changes on the entry/exit by region, $v_{i,t}$ are regions' non-observable effects and $e_{i,t}$ is the error, which is assumed to have normal distribution. Model (2) includes the exit rate $S_{i,t}$ as an explanatory variable in model of entry (1). The same procedure was done for the exit rate. In model (3):

$$\begin{aligned} S_{i,t} &= \alpha + \beta_1 P_{i,t} + \beta_2 PD_{i,t} + \beta_3 GDP_{i,t} + \beta_4 GDPpc_{i,t} + \beta_5 Pd_{i,t} + \beta_6 W_t + \\ &+ \beta_7 UR_{i,t} + \beta_8 EE_{i,t} + \beta_9 ES_{i,t} + \beta_{10} EU_{i,t} + \beta_{11} IR_t + \beta_{12} U_t + d_t + v_{i,t} + e_{i,t} \end{aligned} . \qquad (3)$$

Finally, model (4) explains the exit rate, including entry rate $E_{i,t}$ as an explanatory variable.

The use of dynamic estimators allows us to see whether the entry and exit by region phenomenon is related to what is verified in the previous season. We see that the autocorrelation of order 1 in the dependent variables ENTRY and EXIT is high (0.5 and 0.57 respectively), and the number of time-series observations is small (7). Blundell and Bond (1998) show that when the dependent variable is persistent and when the number of periods considered is not very high, then the Arellano and Bond dynamic difference GMM estimator (1991) is inefficient. The instruments used are generally weak, since this estimator uses several instruments in relation to

[1] Despite the correlation between the variables GDP, GDP*percapita*, POP and POPDEN, we chose to include them simultaneously in the models, as did Gambardella *et al*. (2008). Taking into consideration the characteristics of the panel data predictions with and without the variables simultaneously, the estimates are not substantially affected.

the size, thereby increasing the bias. There is a *trade-off* between efficiency and bias, or rather, increasing the number of instruments increases efficiency but can be problematic and cause bias. Blundell and Bond (1998) extend the GMM (1991) estimator proposing a dynamic estimator, which simultaneously uses the variables in the equation to estimate in levels and first differences, and the same occurs with the instruments used. For the variables in levels in the equation to be estimated, the instruments come in first differences. For the variables in first differences in the equation to be estimated, the respective instruments come in levels. In this way, through the systems of equations in levels and first differences, the problem of persistence of the dependent variable is attenuated, making the instruments used more robust. It is important to note, however, that these estimators can only be considered valid after two conditions have been validated. The first is to check whether the restrictions created as a consequence of the use of instruments are valid and the second is to check whether there is no second-order autocorrelation.

In this way, considering that the autocorrelation of order 1 in the dependent variables ENTRY and EXIT is high and the number of time-series observations is small, the most appropriate estimator is that of the Blundell and Bond (1998) system, in detriment to the Arellano and Bond (1991) difference estimator. To test the validity of the restrictions, we used the Hansen test, with the null hypothesis that the restrictions imposed by the use of the instruments are valid. The non-existence of autocorrelation of the second order was tested by the m2 test, with the null hypothesis of non-second-order autocorrelation. If H_0 is rejected, it is concluded that the estimators are not robust.

Since the database is small, the dynamic estimator GMM System (1998) can, however, generate bias of the estimated parameters. Bruno (2005) therefore proposed an estimator to resolve this problem, an estimator designed by LSDVC (Least Square Dummy Variable Corrected), fixed effect convergence regression. The LSDVC estimator serves to analyse the robustness of the results obtained from the application of the dynamic estimators GMM Difference and GMM System. This estimator, proposed by Bruno (2005), enables possible biased estimates obtained from smaller samples to be reduced and the solidity of the results obtained from other estimators to be checked.

The dynamic model for entry, without considering exit as an explanatory variable is:

$$E_{i,t} = \alpha + \delta E_{i,t-1} + \beta_1 P_{i,t} + \beta_2 PD_{i,t} + \beta_3 GDP_{i,t} + +\beta_4 GDPpc_{i,t} + \beta_5 W_t + \\ + \beta_6 UR_{i,t} + \beta_7 EE_{i,t} + +\beta_8 ES_{i,t} + \beta_9 EU_{i,t} + \beta_{10} IR_t + \beta_{11} U_{i,t} + d_t + v_{i,t} + e_{i,t} \quad (5)$$

Model (6) includes the exit rate of the previous period $S_{i,t-1}$ as an explanatory variable in entry model (5). The same procedure is performed for the exit rate. Model (7) gives:

$$S_{i,t} = \alpha + \delta S_{i,t-1} + \beta_1 P_{i,t} + \beta_2 PD_{i,t} + \beta_3 GDP_{i,t} + +\beta_4 GDPpc_{i,t} + \beta_5 W_t + \\ + \beta_6 Pd_{i,t} + \beta_7 UR_{i,t} + +\beta_8 EE_{i,t} + \beta_9 ES_{i,t} + \beta_{10} EU_{i,t} + \beta_{11} IR_t + \beta_{12} U_{i,t} + \\ + d_t + v_{i,t} + e_{i,t} \quad (7)$$

Finally, model (8) explains the exit rate, including Entry rate $E_{i,t-1}$ of the previous period as an explanatory variable.

4. RESULTS

4.1 Static models

Tables B.1 and B.2 (see appendix B) show the results of the estimation of the static panel models[2]. We can see that the OLS estimation is preferable to the random effects model. From

[2] We estimate the models excluding exit (in the model of entry) and entry (in the model of exit) as explanatory variables, i.e. models (1) and (3). The results are available from the authors upon request.

the LM test, we do not reject the null hypothesis of no relevance of individual effects, thus we reject the random effects estimator as a good estimator. Through the Wald test, we reject the null hypothesis of non-significance as a whole of the parameters of the independent variables. In order to check the relevance of the non-observed individual effects, the Hausman test was performed. The null hypothesis is rejected by the test, so the panel model of fixed effects is the most appropriate model for this estimation. The results therefore suggest that there is empirical evidence of particularities in regions with regard to their influence on the entry and exit rates. Using the results of the application of static estimators to the entry and exit of companies, we can therefore summarise that:

■ entry and exit rates are statistically related to each other. This confirms what is well documented in the literature that markets with more entries also have more exits.

■ there is a positive and statistically significant relationship between population density and entry;

■ there is a negative and statistically significant relationship between the region's GDP and the exit rate;

■ there is no statistical evidence of a relationship between minimum wage and entry rate. However, there is a negative relationship between WAGE and exit rate in model (4);

■ the parameter that measures the effect of unemployment on entry is negative and statistically significant, suggesting that higher unemployment rates may discourage entrepreneurship. As for the exit of companies, model (4) shows that there is no statistical evidence on the relationship between unemployment and exit. This suggests two effects: greater unemployment reduces the climate of trust in business initiatives; the entrepreneur is not resourceful, or rather, he is not enterprising because he is unemployed. Nevertheless, this requires confirmation;

■ a negative and statistically significant relationship can be seen between the proxy for lower level of education EDUELE and the entry rate. This suggests that the low level of schooling does not encourage entrepreneurship;

■ there is a negative and statistically significant relationship between the urbanisation variable and the dependent exit variable.

4.2 Dynamic models

As mentioned above, the dependent variables entry and exit are persistent, a fact that is also confirmed by the results of the estimation models with fixed effects with autocorrelation of the first order. For all these reasons, we must estimate using dynamic panel estimators. In estimating the dynamic models, the GMM system (1998) and the LSDVC (2005) were used. As mentioned, taking into consideration the characteristics of the available panel, the method proposed by Bruno (2005) is the best estimation method for analysing this dynamic panel problem. This estimator can thus produce more robust results[3]. Tables 2 and 3 show the results of the dynamic estimators for entry and exit, respectively.

[3] Like Nyström (2005) or Cilasun and Burak (2006) we did not include time *dummies* since they are not significant.

Table 2 - Dynamic estimators – Entry (1999 – 2005)

Dependent variable – entry rate				
Independent variables	Entry(5) GMMsys	Entry(5) LSDVC	Entry(6) GMMsys	Entry(6) LSDVC
$ENTRY_{i,t-1}$	-0.1948549	0.2189168***	0.2144385	0.46766***
	(0.1400819)	(0.0358348)	(0.1452551)	(0.0510886)
$EXIT_{i,t-1}$			-0.51038**	-0.50344***
			(0.2191854)	(0.1793534)
Local demand				
$POP_{i,t}$	5.52e-08	2.01e-08	1.30e-07	1.18e-08
	(1.17e-07)	(8.65e-08)	(9.52e-08)	(1.04e-07)
$POPDEN_{i,t}$	0.0000418	0.0000463	-0.0000351	0.0000409*
	(0.0000473)	(0.0000595)	(0.0000433)	(0.0000569)
$GDP_{i,t}$	-1.94e-12	8.21e-13	-4.09e-12	-3.13e-12
	(4.93e-12)	(3.62e-12)	(2.97e-12)	(4.58e-12)
$GDPpercap_{i,t}$	-4.01e-06	-1.66e-06***	-1.66e-06	-1.46e-06***
	(2.48e-06)	(5.50e-09)	(3.33e-06)	(6.37e-08)
$WAGE_t$	-0.0006552	-0.000382***	0.000239*	-0.00064**
	(0.0006496)	(0.0001561)	(0.0001329)	(0.0002927)
Entrepreneurship				
$UNEMPR_{i,t}$	-0.801697**	-0.48594***	-1.2798***	-0.3761***
	(0.346326)	(0.0340708)	(0.2549549)	(0.0680606)
$EDUELE_{i,t}$	-0.939785*	-0.122771***	-0.1913192	-0.1424***
	(0.5503911)	(0.0135575)	(0.2581263)	(0.0093299)
$EDUSECD_{i,t}$	-1.009903	-0.4499965	1.10771**	-0.4908557
	(1.483156)	(0.4042661)	(0.5376609)	(0.4736671)
$EDUUNI_{i,t}$	-0.0486067	-0.0227674	0.1192687*	-0.0048357
	(0.1013788)	(0.0407967)	(0.0703688)	(0.0420735)
$INTRATE_t$	-0.0040077	-0.0041554	-0.0053203	-0.008890*
	(0.0042517)	(0.0027366)	(0.0025035)	(0.0002927)
Agglomeration				
$URB_{i,t}$	-0.0772911	-0.020927***	-0.0583214	-0.0242***
	(0.0681871)	(0.0018159)	(0.0533029)	(0.0032954)
Instruments	21		21	
Observations	168		168	
CrossSection	28		28	
Wald (χ^2)	226.43***		187.68	
Hansen N(0,1)	9.93		8.54	
m1 (0,1)	-2.08**		-	
m2 (0,1)	-1.28		-0.77	

Notes: The Hansen test of over-identifying restrictions is distributed as $N(0,1)$ under the H_0 of instrument validity, used in the one-step Blundell and Bond GMM system (1998); the m1 test is a test for first-order autocorrelation of residuals and is distributed as $N(0,1)$, testing H_0 of no first-order autocorrelation; the m2 test is a test for second-order autocorrelation of residuals and is distributed as $N(0,1)$, under null hypothesis of no second-order autocorrelation. ***, **, *, denote significance at 1, 5 and 10% significance level, respectively. Heteroskedasticity is consistent and asymptotic robust standard deviations are reported in brackets.

Table 3 - Dynamic estimators – Exit (1999 – 2005)

Dependent variable – exit rate				
Independent variables	Exit(7) GMMsys	Exit(7) LSDVC	Exit(8) GMMsys	Exit(8) LSDVC
$EXIT_{i,t-1}$	-0.0036327	0.362576***	-0.76773***	-0.21427***
	(0.2707994)	(0.179216)	(0.2759814)	(0.0007881)
$ENTRY_{i,t-1}$	-	-	0.568622***	0.500911***
			(0.146571)	(0.0870415)
Local demand				
$POP_{i,t}$	-5.80e-08	-6.44e-09	-7.43e-08	-4.75e-09
	(1.32e-07)	(9.84e-08)	(5.00e-08)	(8.23e-08)
$POPDEN_{i,t}$	5.53e-06	-0.0000118	7.26e-06	-4.75e-06
	(0.0000499)	(0.0000592)	(0.0000273)	(0.0000594)
$GDP_{i,t}$	5.33e-12	-2.36e-12	4.37e-12**	8.21e-13
	(4.64e-12)	(3.43e-12)	(2.08e-12)	(2.26e-12)
$GDPpercap_{i,t}$	-0.00001***	-4.17e-06***	-7.24e-06**	2.92e-06***
	(3.46e-06)	(7.49e-07)	(3.62e-06)	(4.67e-07)
$WAGE_t$	-0.001607**	-0.000762***	-0.00128***	4.40e-07*
	(0.00073)	(0.000207)	(0.0003727)	(0.0002164)
$PRODUC_{i,t}$	1.83e-06*	8.68e-07***	5.04e-07	-0.000647**
	(9.70e-07)	(3.28e-07)	(7.42e-07)	(2.30e-07)
Entrepreneurship				
$UNEMPR_{i,t}$	-0.2769302	-0.219694***	0.047566	-0.14878***
	(0.3807169)	(0.0331194)	(0.1952522)	(0.0169978)
$EDUELE_{i,t}$	-0.4921235	0.055777***	0.0816352	0.0234748**
	(0.352032)	(0.0105215)	(0.3885114)	(0.0114214)
$EDUSECD_{i,t}$	-2.618412	-0.7574111**	-1.510876	-0.6820273*
	(1.834017)	(0.3761837)	(0.9255693)	(0.3782719)
$EDUUNI_{i,t}$	0.0021536	-0.0100203	-0.0185046	-0.0127782
	(0.111061)	(0.0363292)	(0.0578093)	(0.0357266)
$INTRATE_t$	-0.0129***	-0.008971***	-.016764***	-0.01015***
	(0.0053781)	(0.0032868)	(0.0039988)	(0.0036908)
Agglomeration				
$URB_{l,t}$	-0.0367623	-0.200328***	-0.0166693	-0.02160***
	(0.0504569)	(0.0019183)	(0.0203975)	(0.001596)
Instruments	22		22	
Observations	168		168	
CrossSection	28		28	
Wald (χ^2)	201.3***		728.03***	
HansenN(0,1)	12.64		8.22	
m1 (0,1)	-1.48		-1.65*	
m2 (0,1)	0.39		0.5	

Notes: The Hansen test of over-identifying restrictions is distributed as $N(0,1)$ under the H_0 of instrument validity, used in the one-step Blundell and Bond GMM system (1998); the m1 test is a test for first-order autocorrelation of residuals and is distributed as $N(0,1)$, testing H_0 of no first-order autocorrelation; the m2 test is a test for second-order autocorrelation of residuals and is distributed as $N(0,1)$, under null hypothesis of no second-order autocorrelation. ***, **, *, denote significance at 1, 5 and 10% significance level, respectively. Heteroskedasticity is consistent and asymptotic robust standard deviations are reported in brackets.

In both cases, the Hansen test allows us to conclude that we cannot reject the null hypothesis of the validity of the instruments used. Thus, we conclude that the use of the instruments and the consequent restrictions is not inappropriate. The m2 test shows us that we cannot reject the null hypothesis of the absence of autocorrelation of the second order, thus showing the validity conditions of the estimator. Since these conditions are verified, we conclude that the results of the GMM system dynamic estimator are robust and are more likely to produce consistent and efficient estimates than OLS and fixed-effect estimation.

As stated above, when we consider the number of available observations and the high number of instruments generated by the GMM system estimator (though fewer than the number of cross sections), which could contribute to the bias of the estimated parameters, we present the Bruno (2005) dynamic estimator in order to test the robustness of the results obtained from the GMM system (1998) dynamic estimator. Although many results are similar, some bias can be seen of the parameters estimated by the application of the GMM system (1998). Considering the results of the LSDVC Bruno estimator (2005), it is therefore important to summarise that:

■ there is a positive and statistically significant relationship between rate of entry in the previous period and that of the current period. The same relationship was found by various authors, such as Hicks and Sutaria (2004), justifying it by the multiplier effect. More entries encourage new entries because the market seems more attractive;

■ there is a negative and statistically significant relationship between rate of exit in the previous period and the rate of entry in the current period;

■ there is a negative and statistically significant relationship between rate of exit in the previous period and that of the current period. This relationship is justified by the effect of adjustment to the new market conditions. The more exits there are in the previous period, the fewer exits there will be in the present period given that the adjustment had already been partially made;

■ there is a positive and statistically significant relationship between rate of entry in the previous period and rate of exit in the current period;

■ there is a negative and statistically significant relationship between minimum wage and rate of entry, while there is a positive and statistically significant relationship between minimum wage and rate of exit only in model (8). This is consistent with the theory. The increase in minimum salary makes the labour market firmer at the same time as guaranteeing great income, thus discouraging enterprising activities;

■ the parameter that measures the effect of productivity on the exit rate is negative and statistically significant;

■ there is no statistically significant relationship between either population or population density and exit rate, confirming the results of Love (1996). The same conclusion is obtained in the static estimators;

■ a negative and statistically significant relationship can be seen between GDP *per capita* and entry rate. The inverse relationship is observed in the exit rate, in model (8);

■ the parameter that measures the effect of productivity on the exit rate is negative and statistically significant. This is the expected effect and not what is seen in model (3), where it is not considered as an explanatory variable for entry;

■ there is a negative and statistically significant relationship between the rate of unemployment and the dependent variables. As presented by Nyström (2007b), this relationship is negative, in as much as if the level of unemployment is high, entrepreneurs decide not to close their firm because there are few employment alternatives available and the price of labour will not increase;

■ the parameter that measures the effect of the lower level of education (EDUELE) in the entry rate is negative and statistically significant and it is positive and statistically significant in the exit rate;

■ a negative and statistically significant relationship can be seen between interest rate and exit rate. The same relationship is observed in the entry rate at the 10% of significance in model (6);

■ there is a negative and statistically significant relationship between the effect of agglomeration (URB) and the dependent variables - a relationship already shown by the static estimators.

5. DISCUSSION AND CONCLUSIONS

The study of the determinants of entry and exit of companies per region has been the focus of a lot of literature, since entrepreneurship encourages economic growth and the well-being of a community. The literature and empirical work are now paying more attention to the interrelationship between entry and exit. In this article, we have contributed to the methodology by applying static and dynamic panel models and comparing them to the dynamic of entry and exit of companies in the regions in Portugal.

The results suggest that in understanding exit/entry, entry/exit must be taken as an explanatory factor since not including it may imply inconsistent estimators. The estimated parameters of the variables are generally significant; they show the expected sign and are in accordance with economic theory, which could suggest that the methodology we use is relevant. We use static estimators in order to understand whether there is statistical relevance of non-observable individual effects. A fixed effects panel was used, which showed that the inclusion of specific effects of non-observable regions explains some difference between the exit and entry rates of the regions. We therefore conclude that there really are specific characteristics for each region and this naturally has policy implications, so much so that policymakers must take into consideration that a single policy of entrepreneurship for all regions may have different consequences in each of them.

We saw that the dependent variables entry and exit are persistent in the regions in Portugal. This means that the number of entries/exits may influence the entries/exits in the present period, which leads us to use dynamic estimators. In entry, the empirical results support the multiplier effect in that entry stimulates more entry in the future. In addition to this, this effect is also verified because more exits lead to more entries in the future, which may result either from the demonstration effect or from the income effect. However, we do not confirm the competition effect since past exit does not induce additional entry. In exit, we confirm the competition effect since entries cause more exits in the future. In the same way, we see that more exits in the past cause fewer exits in the future, suggesting that by decreasing competition or excess capacity, companies stay on the market longer. With the exception of the competition effect on entry, our results are essentially similar to those of Nyström (2007a) in the Swedish Manufacturing Industries.

From our models we conclude that the rate of unemployment discourages both the entry and exit decision. We do not obtain the positive effect observed by Audretsch and Fritsch (1999), but our results do concur with those of Reynolds *et al*. (1994), Davidson et al (1994), Hicks and Sutaria (2004) and more recently Nyström (2007b). The same conclusion was obtained by Carree *et al*. (2008) in 103 Italian provinces, suggesting that there is a lack of dynamism in the regional labour markets in Italy. It should also be pointed out that Fritsch and Muller (2003), who use the full information maximum likelihood method, did not find any relationship between the rate of unemployment and creation of new companies.

The results show that the increase in minimum salary does not encourage entrepreneurship, suggesting a substitution effect between creating one's own job and being an employee. The more attractive the latter becomes, the less appealing it is to be enterprising. We can see that if

entrepreneurship were essentially technology-based, this effect would be expected to be irrelevant. In the Portuguese situation, this relationship between wages and entry can also be explained by the fact that incentives to entrepreneurs are subordinate to the creation of work posts, thus causing two effects. On the one hand, employment increases in the business initiatives already established. On the other, this type of intervention does not encourage the setting up of companies.

We do not confirm the income measured by the GDP *per capita* as an attraction factor for new investors. In a first approach, and taking into consideration the literature, this would not be an expected result. However, upon analysing the nature of the Portuguese production structure, where labour intensive industries still predominate, we find a possible explanation for this fact. In these industries, companies seek to make the most of low cost labour, and the positive relationship between income and exit rate is therefore understandable. Once this competitive advantage has been exploited and when income rises, companies abandon the market. This intuition is confirmed both by static estimators and in the dynamics that show a positive relationship between population density and firms' entry.

Both estimators show a negative relationship between the agglomeration effect and the dependent variables. This result is expected for exit but not for entry. It may be due to the fact that many Portuguese regions have a small number of firms, and particularly related and support industries. We also confirm the results obtained by Dunne *et al.* (2005) for productivity, both in static and dynamic estimators.

The evidence of the schooling effect is consistent with the tradition of the studies on this subject. Both in static and dynamic estimators, we confirm the expected result that a low level of schooling does not promote entrepreneurship. Nor does interest rate promote entry, as expected, while an increase in interest rate also leads to a decrease in exit rate. The same conclusion was reached by Dunne *et al.* (2005). In future research, it will be important to ascertain the reasons for this effect. Possible causes could be the fact that periods with higher interest rates are contemporaneous with higher market prices. The authority of the Euro zone monetary policy drives the definition of guiding rates precisely in accordance with "inflation targeting". Another possible explanation could also include entry costs, or rather, the costs of initial investment. If the reference rates for funding increase, it is more likely that firms will stay on the market to generate resources to pay off this debt.

An interesting extension of this work would be to attempt a direct assessment of the interrelationship between exit and entry at an even more geographically specific level - at the level of local councils. In Portugal, public policies have encouraged the cluster of renewable energies. The conditions of territoriality, the physical surroundings and the edaphic and climatic characteristics of each region may be crucial in the entry and exit dynamic in this cluster. This will also be a subject for future research.

REFERENCES

ACS, Z. and AUDRETSCH D. 1989 Small firm entry in U.S. manufacturing, *Economica*, 222(56): 255-265.

ACS, Z. and STOREY D. 2004 Introduction: entrepreneurship and economic development, *Regional Studies*, 38: 871–877.

AUDRETSCH, D. B and FRITSCH, M. 1999 The industry component of regional new firm formation processes, *Review Industrial Organization*, 15: 239–252.

ARAUZO J. M., MANJÓN M., MARTÍN M. and SEGARRA A. 2007 Regional and sector-specific determinants of industry dynamics and the displacement-replacement effects, *Empirica* 34 (2): 89-115.

ARELLANO, M., and BOND, S. 1991 Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies*, 58: 277-297.

ARMINGTON C. and ACS Z. 2002 The determinants of regional variation in new firm formation, *Regional Studies*, 36: 33–45.

BALDWIN, J. and GORECKI P. 1987 Plant creation versus plant acquisition: the entry process in Canadian manufacturing*, International Journal of Industrial Organization,* 5: 27-41.

BERGLUND, E. and BRANNAS, K. 2001 Plant's entry and exit in Swedish municipalities, *Journal of Regional Science*, 35: 431-448.

BLUNDELL, M. and BOND, S. 1998 Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, *Journal of Econometrics*, 87: 115–143.

BRIXY, U. and GROTZ, R. 2007 Regional patterns and determinants of birth and survival of new firms in Western Germany, *Entrepreneurship & Regional Developments*, 19: 293-312.

BRUNO, G. 2005 Approximating the Bias of LSDV Estimation the Bias of LSDV Estimator for Dynamic Unbalanced Panel Data Models, *Economic Letters*, 87: 361-366.

CARREE M and DEJARDIN, M. 2007 Entry Thresholds and Actual Entry and Exit in Local Markets, *Small Business Economics*, 29(1): 203-212,

CARREE, M., SANTARELLI, E. and VERHEUL, I. 2008 Firm entry and exit in Italian provinces and the relationship with unemployment, *International Entrepreneurship Management Journal*, 4: 171-186.

DELMAR, F., SJÖBERG, K. and WIKLUND, J. 2003 The Involvement in Self-Employment among the Swedish Science and Technology Labour-Force between 1990 and 2000, ITPS, Stockholm.

DUNNE T., KLIMEK S. D. and ROBERTS M. J. 2005 Exit from regional manufacturing markets: The role of entrant experience, *International Journal of Industrial Organization*, 23: 399–421.

EVANS, D. and LEIGHTON, L. 1990 Small business formation by unemployed and employed workers, *Small Business Economics*, 2: 319–330.

GAMBARDELLA, A.; MARIANI, M. and TORRISI, S. 2008 How 'Provincial' is your Region? Openness and Regional Performance in Europe, *Regional Studies*, forthcoming.

GEROSKI, P. 1991 *Market Dynamics and Entry,* Blackwell, Massachusetts Gilbert, R. (1989). *"Mobility barriers and the value of incumbency",* in Schmalensee, R and Willing, R., Handbook of Industrial Organization, 1: 537-96 (Amsterdam: North-Holland Elsevier Science Publishers).

GREENE, W. 2003 *Econometric Analysis,* (Prentice Hall, Englewood Cliffs 5th ed.).

GUIMARÃES, P. FIGUEIREDO, O. and WOODWARD, D. 2004 Industrial Location Modelling: Extending the Random Utility Framework, *Journal of Regional Science*, 44(1): 1-20.

GÜNALP, B. and CILASUN, S. 2006 Determinants of Entry in Turkish Manufacturing Industries, *Small Business Economics*, 27: 275–287.

HICKS, A. and SUTARIA, V. 2004 New firm formation: Dynamics and determinants, *Journal of Regional Science*, 38: 241 – 262.

JOHNSON, P. and PARKER, S. 1996 Spatial variations in the determinants and effects of firm births and deaths*, Regional Studies*, 30: 679–688.

KRUGMAN, P. 1998 Space: The final frontier, *Journal of Economics Perspectives*, 12: 161–174.

LAY, T-J. 2003 The determinants of on Interaction between Entry and Exit in Taiwan's Manufacturing, *Small Business Economics*, 20:319-334.

LOVE, J. 1996 Entry and exit: a county level analysis, *Applied Economics*, 28: 441-451.

MARIA J., CAROD, A. and CARRIZOSA M. 2005 An Urban Approach to Firm Entry: The Effect of Urban Size, *Growth and Change*, 4(36): 508–528.

MATA, J. 1991 *Entrada e Mobilidade – Um estudo da Indústria Transformadora Portuguesa,* Universidade do Minho, Braga.

MATA, J. and PORTUGAL, P. 2004 Patterns of Entry, Post-Entry Growth and Survival, *Small Business Economics*, 22: 283–298.

NYSTRÖM, K. 2007a Interdependencies in the Dynamics of Firm Entry and Exit, *Journal of Industry, Competition and Trade*, 7(2): 113-130.

NYSTRÖM, K. 2007b An Industry Disaggregated Analysis of the Determinants of Regional Entry and Exit*, Annals of Regional Science*, 41: 877–896.

NYSTRÖM, K 2007c Patterns and determinants of entry and exit in industrial sectors in Sweden, *Journal of International Entrepreneurship*, 5: 85-110.

REYNOLDS P., STOREY, D. and WESTHEAD, P. 1994 Cross-national comparisons of the variation in new firm formation rates, *Regional Studies*, 28: 343 – 346.

STOREY, D. 1982 *Entrepreneurship and the New Firm,* Croom Helm, London.

STOREY, D. and JONES, A. 1987 New firm formation – a labour market approach to industrial entry*, Scottish Journal of Economic Policy*, 1(34): 37 – 51.

STOREY, D. 1991 The birth of new firms – Does unemployment matter? A review of the evidence*, Small Business Economics,* 3: 167-178.

Appendix A - Variables

Table A.1 – Descriptive statistics

Variable	*Obs.*	*Mean*	*SD*	*Min.*	*Max.*
$E_{i,t}$	196	0.022515	0.009513	0.001614	0.07109
$S_{i,t}$	196	0.021114	0.009691	0.00113	0.057717
Local Demand					
$POP_{i,t}$	196	351222.4	391485.3	42125	2012925
$POPDEN_{i,t}$	196	226.0814	389.6164	15.1	1739.4
$GDP_{i,t}$	196	4.06e+09	7.15e+09	2.95e+08	4.07e+10
$GDPpercap_{i,t}$	196	9458.816	2737.521	5252	20205
$PRODUC_{i,t}$	196	19556.63	5702.955	11000	50500
$WAGE_t$	196	354.7471	21.73936	31800023	385.9
Entrepreneurship					
$UNEMPR_{i,t}$	196	0.055311	0.016939	0.021516	0.110736
$EDUELE_{i,t}$	196	0.168613	0.022818	0.098252	0.278147
$EDUSECD_{i,t}$	196	0.057674	0.010262	0.035832	0.105748
$EDUUNI_{i,t}$	196	0.048625	0.044908	0	0.253783
$INTRATE_t$	196	3.017	0.880755	2.124	4.855
Agglomeration					
$URB_{i,t}$	196	0.113886	0.064792	0.080581	0.982186

Sources: National Institute of Statistics, Portugal, Directories from 1999 to 2005; Bank of Portugal, Statistical report; Ministry of Employment and Social Security, Department of Studies, Statistics and Planning, Demography of companies and establishments in Portugal; Ministry of Education, Information and Assessment Office of the Education System, Students registered according to level of education/training, per NUTS III: 1999 – 2005.

Appendix B – Static estimator results

Table B.1 – Static estimators – Entry with exit (1999 – 2005)

Dependent variable – entry rate				
Independent variables	*Entry (2) OLS*	*Entry (2) RE*	*Entry (2) FE*	*Entry (2) FE – AR(1)*
$EXIT_{i,,t}$	0.6545911	0.06529***	0.593328***	0.500582***
	(0.0439)	(0.044071)	(0.0544475)	(0.0883199)
Local Demand				
$POP_{i,t}$	1.39e-08	1.43e-08**	5.69e-09	5.16e-09
	(6.20e-09)	(6.55e-09)	(3.45e-08)	(5.12e-08)
$POPDEN_{i,t}$	3.72e-06	3.83e-06	0.00005***	0.0000495**
	(3.35 e-06)	(3.56e-06)	(0.000019)	(0.0000226)
$GDP_{i,t}$	-7.15e-13	-7.42e-13***	2.21e-12	2.80e-12
	(2.76e-13)	(2.92e-13)	(2.15e-12)	(2.85e-12)
$GDPpercap_{i,t}$	2.60e-07	6.03e-07**	9.39e-08	-2.08e-07
	(2.29e-06)	(2.76e-07)	(8.94e-07)	(1.07e-06)
$WAGE_t$	-0.1242072	0.000013	-0.0000242	-0.0000766
	(0.0256563)	(0.0000295)	(0.0000479)	(0.0001208)
Entrepreneurship				
$UNEMPR_{i,t}$	-0.0047509	-0.12928***	-0.27614***	-0.42050***
	(0.0242143)	(0.0270533)	(0.0608592)	(0.0836524)
$EDUELE_{i,t}$	0.0537441	-0.0083248	-0.123228**	-0.151645**
	(0.0603414)	(0.0252558)	(0.0558249)	(0.0662894)
$EDUSECD_{i,t}$	-0.0106596	0.0580607	-0.0749087	-0.030253
	(0.0111397)	(0.0628705)	(0.1376777)	(0.2114385)
$EDUUNI_{i,t}$	0.0000131	-0.0119168	-0.0226431	-0.0173319
	(0.0000293)	(0.0114327)	(0.0179374)	(0.0222198)
$INTRATE_t$	0.0011139	0.0010927*	-0.0004212	-0.000871
	(0.000629)	(0.0006239)	(0.0006317)	(0.0011373)
Agglomeration				
$URB_{i,t}$	-0.0078047	-0.0075053	-0.0060908	-0.0097349
	(0.006223)	(0.0062293)	(0.0065537)	(0.007196)
CONS	-0.0017609	-0.0011173	0.0316572	0.0556248
	(0.0130419)	(0.013243)	(0.0252348)	(0.049477)
Observations	196	196	196	168
R^2	0.7020		0.6834	0.6616
F-test (N(0,1))	35.93***		1.60**	1.44*
LM (χ^2)		0.11		
Hausman (χ^2)			24.04***	
Wald (χ^2)		418.77***		
DW				2.1523971

Notes: The Wald test has χ^2 distribution and tests the null hypothesis of non-significance of all coefficients of explanatory variables; the *F*-test is normally distributed *N*(0,1) tests the null hypothesis of non-significance as a whole of the estimated parameters; the Hausman-test has χ^2 distribution and tests the null of non-correlation of individual effects with other explanatory variables; the LM test has χ^2 distribution and tests the H_0 of non-relevance of individual effects; the Durbin-Watson test has N(0,1) distribution and tests the existence of first order correlation, having as limits d_L=1,579 and d_U=1,82; ***, **, *, denote significance at 1, 5 and 10% significance level, respectively; robust standard deviations are reported in brackets; OLS – Ordinary Least Squares; AE – Random Effects; FE – Fixed effects.

Table B.2 - Static estimators– Exit with entry (1999 – 2005)

Dependent variable – exit rate				
Independent variables	*Exit (4) OLS*	*Exit (4) RE*	*Exit (4) FE*	*Exit (4) FE – AR(1)*
$ENTRY_{i,t}$	0.821696***	0.81754***	0.7147263***	0.3770111***
	0.0566298	(0.0568656)	(0.0672609)	(0.0687949)
Local demand				
$POP_{i,t}$	-1.12e-10	-1.24e-09	-1.03e-07***	-1.24e-08
	7.78e-09	(8.55e-09)	(3.78e-08)	(4.53e-08)
$POPDEN_{i,t}$	-3.76e-06	-4.28e-06	-0.0000364*	-0.0000311
	4.03e-06	(4.52e-06)	(0.0000214)	(0.0000199)
$GDP_{i,t}$	5.18e-14	2.57e-14	-5.69e-12**	-3.60e-12
	3.52e-13	(3.88e-13)	(2.35e-12)	(2.46e-12)
$GDPpercap_{i,t}$	3.70e-11	-3.63e-08	-1.15e-06	-3.99e-06***
	5.22e-07	(5.65e-07)	(1.19e-06)	(1.10e-06)
$PRODUC_{i,t}$	-3.36e-07*	-3.55e-07*	-4.25e-07	-9.25e-07***
	1.91e-07	(2.00e-07)	(2.90e-07)	(2.94e-07)
$WAGE_t$	-0.0000101	-4.97e-06	-0.000103**	-0.000707***
	0.0000349	(0.0000355)	(0.0000529)	(0.0001023)
Entrepreneurship				
$UNEMPR_{i,t}$	0.0800846**	0.072946**	-0.1049107	0.0034936
	0.0327426	(0.035985)	(0.0709864)	(0.0788838)
$EDUELE_{i,t}$	0.020569	0.0186496	-0.0497916	-0.0891789
	0.0272133	(0.0296881)	(0.0628697)	(0.0581503)
$EDUSECD_{i,t}$	-0.0556646	-0.0464661	-0.0715318	-0.618227
	0.0679246	(0.0741228)	(0.138576)	(0.1749688)
$EDUUNI_{i,t}$	0.0010658	0.0021322	-0.0020065	-0.0013603
	0.0125956	(0.0132452)	(0.0199625)	(0.0192689)
$INTRATE_t$	-0.0012757*	-0.001273*	-0.0011563*	-0.007245***
	0.0007114	(0.0006995)	(0.0006949)	(0.0009279)
Agglomeration				
$URB_{i,t}$	-0.0076388	-0.0087912	-0.014488**	-0.0102375
	0.0070157	(0.0070096)	(0.0071665)	(0.0062116)
CONS	0.0134395	0.0124291	-0.018579	0.358054***
	0.0154545	(0.015974)	0.0283913	(0.0419355)
Observations	196	196	196	168
R^2	0.6386		0.6612	0.7620
F-test (N(0,1))	24.74		1.88***	1.98***
LM (χ^2)		0.10		
Hausman (χ^2)			19.22**	
Wald (χ^2)		312.71***		
DW				1.8905909

Notes: The Wald test has χ^2 distribution and tests the null hypothesis of non-significance of all coefficients of explanatory variables; the *F*-test is normally distributed *N*(0,1) and tests the null hypothesis of non-significance as a whole of the estimated parameters; the Hausman-test has χ^2 distribution and tests the null of non-correlation of individual effects with other explanatory variables; the LM test has χ^2 distribution and tests the H_0 of non-relevance of individual effects; the Durbin-Watson test has N(0,1) distribution and tests the existence of first-order correlation, having as limits d_L=1,579 and d_U=1,82; ***, **, *, denote significance at 1, 5 and 10% significance level, respectively; robust standard deviations are reported in brackets; OLS – Ordinary Least Squares; AE – Random Effects; FE – Fixed effects.

EL PARTNERSHIP IRLANDÉS EN UN CONTEXTO DE CRISIS

JUAN MIGUEL BÁEZ MELIÁN
E-mail: jmbaez@unizar.es
Departamento de Economía y Dirección de Empresas
Universidad de Zaragoza

Resumen

La República de Irlanda ha sido uno de los países con mayor crecimiento económico durante la década de los noventa. Algunos autores defienden que ello es debido fundamentalmente a la política de partnership o de pacto social. Dicha política ha consistido en una serie de acuerdos a nivel nacional, entre el gobierno, los sindicatos y la patronal, implementados a partir de 1987. Los acuerdos tienen como pieza clave la tasa de crecimiento salarial, pero también recogen otros aspectos relevantes de política de rentas, así como de política fiscal. Este trabajo hace un análisis de dicho fenómeno, haciendo hincapié en sus consecuencias para el mundo empresarial, especialmente en lo que respecta a la extensión de dichos acuerdos al interior de las organizaciones empresariales.

Área Temática: Economía Industrial.

1. INTRODUCCIÓN

Desde 1987 se viene produciendo en Irlanda una serie de pactos sociales a tres bandas: el gobierno, la patronal y los sindicatos. La pieza clave de estos acuerdos es la tasa de crecimiento salarial, pero también incluyen aspectos relevantes de política fiscal y social.

Por otra parte, en los años noventa y en la primera mitad de la década posterior, la economía irlandesa ha crecido a un ritmo sin precedentes, bastante superior a la media de los países de la OCDE. Podríamos suponer que una de las principales causas de tan fuerte crecimiento ha sido precisamente la política de pacto social llevado a cabo. Pero también podemos situar la dirección de causalidad en sentido inverso.

En cualquier caso, más que la relación entre ambos fenómenos (fuerte crecimiento económico-política de pacto social), el objetivo de este trabajo es analizar el segundo de ellos. Sin embargo, para entenderlo mejor haremos una breve descripción de los logros económicos de Irlanda durante los últimos veinte años, así como de los otros dos hechos que nos parecen más relevantes de la historia económica más reciente de Irlanda: la decisión de "abrir" la economía en los últimos años 50 y la fuerte crisis de los años 80 (apartado 2). Los siguientes epígrafes los dedicaremos a la política de Partnership, propiamente dicha, la interpretación de la misma, su implementación a nivel de empresa, el papel de los sindicatos en dicha política y concluiremos con algunos apuntes sobre el futuro del mismo.

2. EL TIGRE CÉLTICO

El periodo posterior a la independencia se puede caracterizar como una economía protegida centrada en la sustitución de importaciones. Al igual que algunas economías latinoamericanas de la época, esta estrategia fracasó. Efectivamente, la industria nacional mejoró, pero los problemas estructurales que afectaban a la economía irlandesa no se modificaron, por lo que el nivel de vida de la población seguía estancado y la tradicional emigración irlandesa se presentaba como la única salida para una buena parte de los jóvenes.

A finales de los 50 se decide la apertura de la economía con un doble objetivo: modernizar y reorientar la industria nacional, por un lado, y atraer inversión extranjera, por otro. La estrategia funcionó, aunque generó y profundizó otros problemas estructurales que aún hoy en día siguen

sin resolverse. En esencia, la política económica ha estado caracterizada desde entonces por fuertes incentivos fiscales y monetarios, tanto para la inversión foránea como para la interna.

Los resultados de esa política fueron (NESC, 1996; citado en O'Donnell; 1999) los siguientes: fuerte crecimiento económico y demográfico; un importante ajuste estructural, reduciéndose el empleo del sector primario desde un 37% del total en 1960 a un 14% en 1987, lo que se tradujo un en rápido crecimiento de la productividad; aumento considerable en los niveles de vida y, por último, la convergencia con el Reino Unido en los niveles de salarios e ingresos.

Sin embargo, la excesiva orientación hacia afuera conllevó el asentamiento de una economía dual. Por un lado, estaban las empresas propiedad de las multinacionales extranjeras, caracterizadas por una alta productividad y calidad de los puestos de trabajo y, por otro lado, estaban las empresas nacionales, con menor productividad y peor clima laboral. La conexión entre ambos sectores no siempre ha sido fácil desde entonces, lo que ha caracterizado las relaciones industriales irlandesas durante las últimas décadas.

Esta dualización generaba problemas y, a pesar de los logros económicos conseguidos, a principios de los años 70 el National Economic Social Council (NESC) seguía hablando de restricciones estructurales (NESC, 2008): disminución de la población, especialización en productos agrícolas, debilidad del sector industrial y la consiguiente restricción en la balanza de pagos y un débil sistema nacional de innovación. Las dificultades económicas de los años 70 se intentaron resolver mediante la inflación y el endeudamiento del sector público. Esta mala resolución de la crisis de los 70 condujo a una situación crítica durante los 80.

Quizás lo más relevante, desde el punto de vista de la política económica, durante la crisis de los 80, fue el abandono de la paridad monetaria con la libra esterlina. En 1979, la República de Irlanda pasa a formar parte del tipo de cambio europeo, lo que va a conllevar un cierto desbloqueo de la economía inglesa. Pero los problemas no se resolvieron. El endeudamiento público se intentó solventar mediante el incremento de la presión fiscal, lo que significó un obstáculo más para el despegue de la actividad productiva. Al final de este periodo, en 1987, el desempleo alcanzó el 16.6% del total de la población activa (OCDE, 1995) y el ratio deuda/PNB del sector público se acercó al 130% (O'Donnell, 1999).

La crisis de los años 80 puede ser interpretada como un fuerte proceso de convergencia con los países de la Comunidad Europea. La debilidad crónica de la economía irlandesa y la menor escala de una buena parte de su tejido empresarial fueron los ingredientes principales. En cualquier caso, hemos encontrado otras posibles explicaciones en la literatura. Una buena síntesis de ellas es la debida a Mjoset (1992, citado en O'Donnell, 1999), para quien dicha crisis es producto de un conjunto de factores políticos, sociales y económicos. Para este autor el fracaso económico de Irlanda descansa en un círculo vicioso basado en dos hechos que definen la economía de este país durante las últimas décadas: su débil sistema nacional de innovación y la continua huída de la mano de obra vía emigración.

El siguiente periodo es el de *partnership*. Queremos advertir en primer lugar, que encontramos exagerado la importancia que algunos autores dan al *partnership* como causa principal de dichos logros económicos. Creemos que no está demostrada una relación causa-efecto, aunque sí somos de la opinión de que la firma de los pactos sociales a partir de 1987 facilitó considerablemente el impresionante despegue económico de Irlanda durante las dos últimas décadas.

CUADRO 1. Algunos datos económicos de las dos últimas décadas

	1987	1992	1997	2002	2006
Ratio de crecimiento de la población	0,17	0,82	1,05	1,82	0,49
Ratio neto de emigración	-10,57	0,48	5,07	8,39	..
PIBpp	9577	14334	21795	33036	40716
Balanza comercial	2,36	6,05	14,36	35,97	32,48
Desempleo	16,6	15,4	9,9	4,5	4,4
Gasto público social	21,16	17,06	14,34	15,49	..
Comparación con EEUU (PIBpp)	48,24	57,13	71,40	89,75	92,76
Comparación con EEUU (Productividad)	65,59	76,58	90,53	101,46	102,39

Construcción propia a partir de datos de la OCDE

Con el objetivo de ilustrar dicho despegue hemos construido el cuadro 1, en el que resaltamos las variables macroeconómicas que nos parecen más relevantes. Veamos, por ejemplo, el crecimiento de la población. Como es bien sabido, Irlanda ha sido desde hace bastante tiempo un país exportador neto de población. La válvula de escape de la crisis endémica que ha caracterizado a este país ha sido tradicionalmente la búsqueda de puestos de trabajo en el exterior (uno de los pilares del círculo vicioso de Mjoset, 1992, ya comentado). Por lo tanto, el ratio de crecimiento poblacional es un buen reflejo del ciclo económico. Así, durante toda la década de los 50 y hasta 1961, la población disminuye, llegando a poco más de 2.800.000 habitantes. A partir de ahí comienza un ciclo de crecimiento (reflejando los efectos positivos de las medidas de apertura tomadas), que solo se interrumpe durante los últimos aÿÿs 80. Enÿÿl periodo que aquí nos interesa, a partir de 1987, el ratio es casi siempre creciente, disminuyendo bruscamente en el año 2006.

Pero quizás, un mejor reflejo del ciclo económico es el ratio neto de emigración (segunda fila del cuadro 1). Este ratio es negativo durante casi toda la segunda mitad del siglo XX, excepto en los años 1971-1978, 1991-1992 y a partir de 1995 hasta nuestros días, llegando en el 2005 al 15.87 por cada mil habitantes. Irlanda, junto con el Reino Unido y Suecia, es uno de los países que permite en 2004 el libre acceso de los emigrantes procedentes de los 10 nuevos países miembros (EU25), por lo que, dada la boyante situación económica de entonces, se produce una llegada de inmigrantes de tal magnitud que obliga al gobierno irlandés a rectificar, restringiendo la libre entrada de trabajadores procedentes de Rumania y Bulgaria. Esta última oleada de mano de obra extranjera se diferencia de las anteriores, en que ocupan puestos de trabajo en sectores con menor cualificación, tales como la agricultura, la construcción y el comercio mayorista y minorista (OCDE, 2008).

Las otras variables que hemos destacado, y en las que también se reflejan los sorprendentes avances de las últimas décadas, son los siguientes: el PIB per cápita, que presenta una tendencia reciente durante toda la serie, llegándose a multiplicar por 4 en el periodo 1987-2006; la balanza comercial, que mejora sustancialmente durante el periodo 1992-2004 (empeorando los dos años siguientes), reflejando la vocación exportadora de la economía irlandesa; la tasa de desempleo, que ha ido disminuyendo desde 1994, pasando del 16,6% en 1987 al 4.4% en 2006 (téngase en cuenta, sin embargo, que una buena parte de este desempleo es de larga duración, una de los principales críticas que se le hace al periodo de *partnership*); y las comparaciones con EE.UU., tanto en término de ingreso (PIBpp) como de productividad (PIB por hora trabajada), que también mejoran de forma importante, llegando incluso a superar a EE.UU. en el caso de la productividad.

No obstante, otro ataque que recibe el *partnership* es su falta de inclusión social (que se ha intentado paliar en los últimos acuerdos, como veremos en el siguiente apartado). Esto se refleja en una continua reducción en el gasto público social (sexta fila del cuadro 1). Dicho gasto fue del 21.26% del PIB en 1987, fue reduciéndose hasta llegar al 13.64% en el año 2000, recuperándose ligeramente a partir de entonces.

Vemos por tanto que, a pesar de los increíbles avances económicos (Irlanda es, sin duda, uno de los países del mundo con mayor desarrollo durante los últimos años), existen problemas sin resolver que se ven agravados dada la coyuntura de crisis global en la que vivimos. Cierta parte de la literatura afirma que el contexto globalizador actual constituye una amenaza para los sectores sociales más desfavorecidos. En el caso de Irlanda, la continua disminución de los gastos sociales, así como los problemas de falta de inclusión social de una capa de la población no despreciable, apoyan esta idea. Además, la actual crisis financiera genera nuevas incertidumbres, especialmente en lo que respecta a los mercados laboral e inmobiliario. Este último tiene mucho que ver con la fuerte expansión de los últimos años y su desplome creemos que todavía no ha tocado fondo.

En cualquier caso, pensamos que dos son los pilares fundamentales de estas importantes transformaciones: por un lado, el proceso de integración de Irlanda en la UE, que ha mejorado su competitividad internacional; y, por otro lado, el ambiente de cohesión social que ha generado los acuerdos firmados a partir de 1987, que ha permitido presentar una estrategia

común a todos los agentes sociales, desarrollando las fortalezas de la economía irlandesa y solventado sus debilidades.

Por otra parte, la nueva estrategia económica tiene una clara vocación exportadora, que la asemejan considerablemente a la adoptada por los pequeños países asiáticos, también de reciente y sorprendente nivel de desarrollo. Sin embargo, el NESC intenta evitar esta semejanza, negando la idea de que la prosperidad irlandesa provenga exclusivamente del sector exportador, en su mayoría propiedad extranjera (NESC, 2008). Este organismo también hace otras tres recomendaciones que nos parecen importantes: desarrollar el sector servicios, centrándose en el comercio internacional, aunque no exclusivamente; hacer hincapié en el conocimiento y la formación y en el papel que puede jugar el sector público en la innovación; y tener en cuenta las implicaciones que para el bienestar social tiene la llegada de emigrantes, tanto en lo que respecta al retorno de emigrantes irlandeses, como a la nueva mano de obra procedente del exterior.

Asimismo, dicha institución deja vislumbrar cuáles son los principales objetivos de la política económica para los próximos años: mantener baja la inflación combinada con un crecimiento estable del output, política de ingresos que mejore la competitividad y dirija los conflictos distributivos, y una política de cambio estructural que también fomente la competitividad y elimine las barreras a la participación y cohesión social.

La historia económica de Irlanda de las últimas dos décadas resulta interesante al menos por cuatro razones (O'Donnell, 1999): es un interesante caso de estabilización y ajuste macroeconómico en una pequeña y extremadamente abierta economía, constituye una sorprendente transformación desde una economía periférica hacia un centro productivo de alta tecnología, forma parte del proceso de integración europea y, por último, es produce en un ambiente de concertación social y de innovación institucional. Estos cuatro aspectos van obviamente entrelazados y a nosotros nos interesa especialmente el último, pero es precisamente su vinculación con los otros tres lo que nos ha llevado a introducir este esbozo de la coyuntura económica irlandesa.

3. EL PARTNERSHIP

No cabe duda que durante la década de los 90 el partnership ha experimentado una considerable reactivación en Europa. De los quince estados miembros de la UE por aquel entonces, nueve habían visto el surgimiento o el resurgimiento de iniciativas de pacto social: Alemania, Bélgica, Finlandia, Grecia, España, Holanda, Irlanda, Italia y Portugal; mientras que Austria venía aplicándolas desde hacía varias décadas (O'Donnell, 2001). En ese mismo trabajo se citan otros donde se intenta encontrar las causas de este fenómeno: la reacción frente a los efectos nocivos de la liberalización y la globalización (Traxler; 1997), el mercado europeo y el proceso de integración monetaria (Rhodes, 1997) y el papel que puede jugar la concertación social en las reformas de sistemas de bienestar (Rhodes, 1997).

Con el objetivo de abordar el análisis del partnership europeo, O'Donnell parte de cuatro premisas que nos parece interesante recordar: no es muy productivo la creación de generalizaciones simplistas, dada la variedad de formas, métodos y temas implicados en los diferentes pactos sociales llevados a cabo; hay algunas características comunes en los partnerships de la UE: el criterio de Maastricht, el desempleo y el envejecimiento de la población, aunque este autor está más interesado en identificar factores propios del partnership en Irlanda; el contexto de mala performance macroeconómica en el que resurgió el partnership en muchos países europeos tiene poco que ver con la situación en Irlanda; y, por último, el caso irlandés puede verse reflejado en otros países, dada la relevancia internacional que ha ido adquiriendo su partnership. Las conclusiones de O'Donnell son las siguientes:

- La experiencia de los países europeos demuestra que el conocimiento compartido entre los diferentes agentes sociales juega un papel clave en la creación y sostenimiento del partnership. Donde éste existe, dicho conocimiento compartido abarca principalmente dos esferas de la economía: la integración monetaria europea y la interacción entre empleo, salarios, sistema impositivo y protección social.

- Hay una clara tendencia hacia la descentralización en los pactos sociales de los años 90. En la mayoría de los casos se ha tratado de procesos controlados de descentralización.
- El contenido social del partnership actual es diferente al de los años 60 ó 70: fortalecimiento del estado del bienestar con incremento del gasto y empleo público. Durante los 90, la moderación salarial formaba parte de una política de oferta, cuyos objetivos eran el empleo, la competitividad y el crecimiento, con una política macroeconómica basada en la moneda única y el saneamiento de las finanzas públicas.
- En algunos países, el partnership ha sido visto como la única vía de cambio, sobretodo en lo que se refiere al bienestar, dado el equilibrio de fuerzas políticas (Holanda, Suecia, Portugal y España), o la debilidad del sistema político (Italia).
- El partnership ha sido más predominante y exitoso en los pequeños países. Italia, y en menor medida España, ha sido el único país grande que ha concluido un pacto social y ha alcanzado una significativa mejora económica basada en el partnership.
- La performance económica ha sido mejor en aquellos países donde el partnership ha cambiado sustancialmente o se ha reinventado (Holanda, Irlanda, Suecia, Portugal, Finlandia, Italia y Dinamarca), que en aquellos donde ha tenido un carácter más continuista (Austria y Alemania). Esto sugiere que un aspecto clave debe ser encontrar un equilibrio entre las ventajas de un partnership estable, en términos de conocimiento compartido y confianza, y las desventajas de un modelo rígido, en el que ciertos intereses comienzan a atrincherarse, convirtiéndose en un importante obstáculo para el cambio.
- El gobierno es un elemento clave del partnership, incluso en aquellos países con acuerdos bilaterales.
- De todos los países de Europa, Irlanda ha sido el que ha ido más lejos en ampliar la participación de los grupos sociales marginales o menos favorecidos.

Por otra parte, podemos describir las relaciones laborales irlandesas de la segunda mitad de los años 80 a través de los siguientes rasgos (Teague, 2002). En primer lugar, tanto los sindicatos como las organizaciones empresariales se encontraban muy fragmentados. Pero las segundas carecían de la autoridad central que ejercía el Irish Congress of Trade Unions (ICTU) sobre el conjunto de los sindicatos. Aunque esto no impedía que los enlaces sindicales disfrutaran de la suficiente autonomía para marcar el ritmo y los contenidos de las negociaciones a nivel de empresa. En segundo lugar, el sistema de negociación colectiva carecía de orden. En el sector público, la negociación era centralizada; en algunas esferas de la economía, especialmente en el sector de la construcción, prevalecía la negociación a nivel sectorial; y, en general, la negociación centralizada era la norma en el sector privado. En tercer lugar, y a diferencia de la tónica general en Europa, Irlanda carecía de normativa legal que asegurara la negociación colectiva para la mayoría de los trabajadores. El salario mínimo no fue introducido hasta el año 2000. Además, desde los primeros años 80 se produce una constante disminución de la afiliación sindical, ocasionada probablemente por la llegada de nuevas multinacionales, especialmente de Estados Unidos.

En este contexto de relaciones laborales y con la profunda crisis económica de los años 80 de fondo, ya comentada en el apartado 2, comienza el periodo de partnership irlandés. En 1987 se firma el primero de los acuerdos, el *Programme for National Recovery (PNR)*, entre el gobierno, los sindicatos, la patronal y organizaciones agrícolas. Este pacto tenía como elemento clave, al igual que los siguientes, un límite máximo para el crecimiento de los salarios. En este caso, el 2.5% para el trienio 1988-1990 (véase el cuadro 2). La moderación salarial era vista como el factor clave para ganar competitividad internacional y lograr el saneamiento de las finanzas públicas. Pero este acuerdo y los siguientes también han incluido una serie de temas de política social y económica: reforma fiscal, sistema de pensiones, gasto sanitario, ajustes estructurales y la adherencia de Irlanda a la banda estrecha del Sistema Monetario Europeo y los criterios de Maastricht (O'Donnell, 2002). En el cuarto de los acuerdos, el *Partnership 2000*, la perspectiva de pacto social se amplió, añadiéndose a las negociaciones una serie de

organizaciones en representación de los grupos sociales menos favorecidos: el *National Women's Council of Ireland*, el *Irish National Organisation for the Unemployed (INOU)* y el *Council of Religions in Ireland (CORI).*

CUADRO 2. Principales acuerdos durante el periodo de Partnership (1987-2002)

Programa	Años	Acuerdos sobre salarios	Otros acuerdos relevantes
Programme for National Recovery (PNR)	1987-1990	2.5% annual	Reducción de la carga impositiva y mejoras de política social
Programme for Economic and Social Progress (PESP)	1991-1993	1991: 4% 1992: 3% 1993: 3.75%	Muy similares al anterior acuerdo
Programme for Competitiveness and Work (PCW)	1994-1996	1994: 2% 1995: 2.5% 1996: 2.5% para los 6 primeros meses y 1% para los 6 restantes	Similares a los anteriores
Partnership 2000 for Inclusion, Employment and Competitiveness	1997-1999	1997: 2.5% 1998: 2.25% 1999: 1.5% para los 9 primeros meses y 1% en los últimos 6 meses	Extensión de los acuerdos a sectores sociales menos favorecidos y al nivel de empresa
Programme for Prosperity and Fairness (PPF)	2000-2002	2000: 5.5% 2001: 5.5% 2002: 4% para los primeros 9 meses	Compromiso para mejorar los niveles de vida (especialmente para la gente de bajos ingresos) y promover el cambio organizacional y el partnership a nivel de empresa
Sustaining Progress	2003-2005	9 primeros meses: 3% 6 meses siguientes: 2% 3 meses siguientes: 2%	Similares a los anteriores
Towards 2016	2006-2015	6 primeros meses: 3% 9 meses siguientes: 2% (excepciones) 6 meses siguientes: 2.5% 6 meses siguientes: 2.5%	Similares a los anteriores

Construcción propia a partir de Gunnigle y Brosnan, Teague (2002) y los documentos de los seis acuerdos

Esta inclusión de los sectores marginados trataba de acallar las críticas sobre el reparto de las ganancias del Partnership entre los diferentes agentes sociales. Estas críticas tienen dos partes (Teague, 2002): por un lado, se dice que mientras la participación de los beneficios en el conjunto del PIB ha aumentado considerablemente, la participación de los salarios ha disminuido; por otro lado, también se afirma que la desigualdad de salarios ha aumentado con el partnership. En cuanto a la primera de estas críticas, Teague aporta un gráfico en donde demuestra que, efectivamente, la participación de los salarios en el total del output creado en Irlanda ha disminuido notablemente desde mediados de los años 80. En aquellos tiempos se situaba por encima del 70%, mientras que el en año 2000 bajó por debajo del 55%. Si las comparamos con la media de la UE las cifras son aún más escandalosas, ya que esta se sitúa algo por debajo del 70% durante los últimos años de la década de los 90. La segunda de las críticas es un problema serio para Teague, que puede poner en peligro la propia supervivencia del enfoque del Partnership en Irlanda. Sin embargo, los datos de la OCDE sobre la distribución de la renta disponible en Irlanda, no avalan este pesimismo: 33.1 de índice de Gini en 1985, 32.4 en 1995 y 30.4 en el año 2000. Por otra parte, los costes laborales por unidad de output producida ha disminuido en Irlanda durante el periodo 1988-1997, mientras que en el mismo periodo aumentaron en un 2.8% en la UE (Gunnigle y Brosna).

Estos dos autores resaltan otros logros económicos de este periodo: el ratio salario/hora aumentó un 44.3% y los ingresos semanales también lo hicieron en un 42.7% entre 1987 y junio de 1997, aunque este aumento fue superior en el sector público (un 55.8% entre marzo de 1988

y junio de 1997); la carga impositiva sobre el trabajador disminuyó desde el 35.4% de su renta bruta en 1987 al 26.2% una década más tarde; la inflación acumulada durante la década fue del 28.4%; los días perdidos debido a la conflictividad laboral durante 1988-1997 fueron menos de un tercio a los perdidos durante 1980-1987; y el crecimiento de la productividad fue del 8% anual durante el periodo 1991-1996, mientras que la media de la UE fue del 2%, aunque, recuerdan Gunnigle y Brosnan, la mayor parte del crecimiento de la productividad en manufacturas puede ser atribuida a la parte del sector de propiedad extranjera. Para estos autores, el impresionante desarrollo económico de Irlanda durante las dos últimas décadas no puede atribuirse exclusivamente al enfoque de Partnership predominante, aunque hay claras evidencias de que ha contribuido considerablemente.

En cualquier caso, y para terminar con este apartado, recordamos las razones que da O'Donnell (1999) para el éxito del Partnership irlandés:

- El consenso alcanzado en Irlanda, tanto entre agentes sociales como entre partidos políticos, sobre los requisitos para una participación exitosa en la economía europea y sobre el punto de vista de que no había forma de escapar de estos requisitos.
- El alto nivel de coordinación alcanzado sobre la evolución de los salarios.
- El consenso alcanzado sobre finanzas públicas. Esto era necesario no solo por los criterios de Maastricht, sino también por la estrecha relación existente entre dichas finanzas y la negociación colectiva y el tipo de cambio.
- Irlanda llevó a cabo una serie de políticas de oferta que aseguró la competitividad internacional y fortaleció el rápido crecimiento económico.

4. LA INTERPRETACIÓN DEL PARTNERSHIP IRLANDÉS

Gunnigle y Brosnan definen el Partnership como una actividad relacional, basada en el reconocimiento de intereses comunes para asegurar la competitividad, la viabilidad y la prosperidad de la empresa. Esto implica el compromiso por parte de los empleados para mejorar en calidad y eficiencia; y la aceptación por parte de los empresarios de que los empleados tienen derechos e intereses en la empresa, que deben ser considerados en lo que respecta a las principales decisiones que afecten a su empleo.

Gunnigle (1999) tiene una visión más macro: en términos generales, los acuerdos firmados desde 1987 tienen las principales características de los acuerdos corporativos. Están firmados por el gobierno, los sindicatos y la patronal y van más allá de las relaciones puramente industriales, incorporando amplios temas de carácter social y económico; representan un intercambio negociado entre las partes sociales por los que los sindicatos se comprometen a que sus afiliados aceptarán los acuerdos; y, por último y como parte de ese intercambio, tanto los sindicatos como las organizaciones empresariales tienen oportunidad de influir directamente en decisiones de política económica en áreas de gran importancia para sus afiliados: el empleo, la seguridad social y el sistema impositivo.

Sin embargo, O'Donnell (1999, 2002) no está de acuerdo con la asimilación de la experiencia irlandesa con el modelo clásico de corporativismo, desarrollado en Europa en los años posteriores a la Segunda Guerra Mundial. Para él, algunas circunstancia que se dieron en los países de corte neocorporativista (como el predomino de la socialdemocracia, la cohesión de las organizaciones empresariales y la existencia de un movimiento sindical con un alto nivel de autoridad centralizada), no se han dado en Irlanda. Por otra parte, O'Donnel también alega una característica del modelo irlandés que lo hace idiosincrásico: se refiere al enfoque de conocimiento compartido por los agentes sociales sobre los principales problemas sociales y económicos a los que ha de hacer frente el Partnership.

Pero ese conocimiento compartido va evolucionando con la experiencia que las partes van adquiriendo en el proceso de negociación. Es decir, los *partners* sociales no tienen sus preferencias definidas de antemano, sino que es durante el proceso de deliberación cuando tienen la posibilidad de diseñar y rediseñar sus conocimientos, identidades y preferencias. Las

partes se reúnen para resolver problemas comunes y cada una de ellas es consciente de que necesita a las otras para ello. Este enfoque de resolución de problemas resulta un aspecto esencial del Partnership irlandés para O'Donnel. Para él, las viejas formas de negociación (con predominio del enfrentamiento entre los diferentes puntos de vista) deben ser sustituidas por el análisis, el diálogo y el conocimiento compartido. El consenso y el conocimiento compartido, más que las premisas para el Partnership, son su resultado. El primero es necesario no sólo entre partners sociales, sino también entre partidos políticos, sobre política macroeconómica, lo que sugiere la idea de que, una vez logrado el consenso, el único margen que les queda al gobierno son políticas de oferta que permitan el cambio estructural y el desarrollo económico.

Por tanto, su visión del Partnership implica dos concepciones o dimensiones del mismo: por un lado, está la interdependencia funcional, expresada en la necesidad de negociar; y, por otro, está la visión de solidaridad, de inclusión social y participación en las decisiones tomadas colectivamente. Ambas son esenciales. Si sólo se tiene en cuenta la primera, el proceso sólo reflejaría el poder negociador que cada una de las partes tiene, y si sólo se considera la segunda se corre el riesgo de reducir el proceso a una sucesión de simples consultas. Dicho en otras palabras, cada una de ellas necesita de la otra. La tercera dimensión esencial del Partnership es su enfoque resolutivo de problemas, ya comentado.

Esta visión global y consensuada del Partnership implica tres requisitos de política económica y tres elementos constitutivos del modelo de concertación social irlandés (O'Donnel, 2002). En cuanto a los primeros, y dado el reducido mercado interno, el elevado grado de apertura y el proceso de integración europea de la economía irlandesa, la política macroeconómica debe:

- Garantizar una inflación baja y un crecimiento estable de la demanda agregada.
- Establecer una evolución de los ingresos que asegure la mejora en competitividad y que dirija los conflictos distributivos de forma que no interrumpan el funcionamiento regular de la economía.
- Establecer unas políticas complementarias que faciliten y promuevan el cambio estructural, de cara a mantener y mejorar la competitividad en un entorno internacional cambiante.

Los segundos son:

- Una orientación global, que parta de la creencia que la más alta participación en la actividad económica y social, por un lado, y la dirección política, por otro, son inseparables y requisitos esenciales para el bienestar de la sociedad irlandesa.
- Reconocimiento y aceptación sin ambigüedades de la participación de Irlanda en la economía internacional y en el proceso de integración europea. Esto implica que la competitividad de la economía irlandesa es prerrequisito indispensable para el logro de los otros objetivos sociales y económicos.
- La implantación de los tres requisitos de política económica vistos anteriormente ha estado fuertemente asociada a los programas negociados. Estos obligan las partes sociales ser consecuentes con ellos y permite al gobierno adoptar una perspectiva estratégica sobre los principales temas políticos

Por otra parte, O'Donnell (2001) resalta tres canales a través de los que el enfoque de Partenership ha tenido una influencia significativa sobre la economía irlandesa:

- La negociación colectiva: ya que el ambiente resultante de moderación salarial y de alta rentabilidad ha sido clave para la creación de empleo, la atracción de inversión extranjera y el éxito comercial de las compañías irlandesas.
- La consecuencia y coherencia de la política económica: que ha permitido la transición desde una economía de alta inflación, inestable y conflictiva a otra de baja inflación y estable. Esto se ha conseguido preservando un alto nivel de solidaridad social, que ha resultado ser un prerrequisito esencial para sostener las políticas redistributivas y dirigir las reformas y el cambio estructural necesario de forma no conflictiva.

- Los cambios productivos: que han partido de la idea de que hay una conexión cercana entre escenario macroeconómico y cuestiones distributivas, por un lado, y compromiso constructivo y problemas de oferta, por otro.

Pero el enfoque de Partnership también ha sido objeto de críticas. Las recibidas desde una perspectiva liberal, acusando a la negociación colectiva del mantenimiento e incremento del desempleo, O'Donnell no les da excesivo crédito, aludiendo que las mismas no están teniendo debidamente en cuenta las imperfecciones propias del mercado de trabajo. Desde el otro lado ideológico se afirma que el Partnership no ha resuelto, e incluso ha aumentado en algunos aspectos, la desigualdad existente. Esta crítica sí tiene un sustancial apoyo y una buena parte de los autores que hemos consultado reconoce que la falta de inclusión social es uno de los grandes problemas pendientes en Irlanda.

Otra crítica es el predominio de las empresas multinacionales sobre las irlandesas. O'Donnell contesta diciendo que es cierto que un porcentaje importante del ouput, el empleo y las exportaciones irlandesas proceden de las multinacionales, pero ello no es óbice para admitir que las mismas tienen un efecto positivo sobre el resto de la economía, lo que ha influido en una mejora significativa en la performance de las empresas locales. También se alega que los empleos generados en este periodo son de baja calidad. La respuesta es que, si bien se admite el incremento en las diferencias salariales, ello no es suficiente para asegurar que la expansión del empleo vivida en Irlanda durante las últimas décadas haya estado dominada por trabajo de mala calidad.

La última crítica que queremos comentar es sobre el alcance y la profundidad del cambio en las organizaciones irlandesas. O'Donnell califica de impresionante la proporción de organizaciones en las que ha habido algún de tipo de cambio, pero otros autores no son tan optimistas. Gunnigle y Brosnan , por ejemplo, afirman que una buena forma de comprobar si el enfoque de Partnership está llegando realmente a los lugares de trabajo es comprobar la extensión de los esquemas de participación financiera (participación en los beneficios y en la propiedad de la empresa), ya que ello nos dice en qué medida las ganancias originadas por las mejoras de la performance está siendo compartida entre empleadores y empleados. En el cuadro 3 tenemos algunos datos aportados por ellos.

CUADRO 3. Porcentajes de participación financiera en las organizaciones irlandesas

	Directivos	Técnicos/Profesionales	Oficina	Manuales
Participación en la propiedad	23	13.8	11.5	9.6
Participación en los beneficios	19.2	13.4	12.6	10

Fuente: Gunnigle et al (1997), citado en Gunnigle y Brosnan

Vemos que los porcentajes de participación financiera no son relevantes en Irlanda. Sólo en el caso de los cargos directivos este tipo de esquemas tiene una presencia algo significativa: alrededor del 20% de las organizaciones. Estos porcentajes descienden considerablemente en las otras categorías profesionales, llegando al 10% en el caso de los trabajadores manuales. Por otra parte, Gunnigle y Brosnan consideran el compromiso de seguridad de los puestos de trabajo. Éste, en ocasiones, es utilizado como moneda de cambio para asegurar el apoyo de los empleados a las iniciativas de cambio. Para ellos, el cumplimiento de este tipo de compromisos es altamente cuestionable y citan algunos ejemplos, como los casos de *Aughinish Alumina* y *Telecom Eireann*. En cualquier caso, su conclusión es que es demasiado pronto para valorar si estos acuerdos son un éxito o un fracaso del Partnership. En línea con otros autores, su posición es que lo único que podemos afirmar es que las mejoras en competitividad y empleo han tenido lugar en el contexto de la negociación colectiva, pero no podemos asegurar nada sobre algún tipo de relación causal, aunque existe alguna evidencia no conclusiva de un efecto beneficioso del ambiente de Partnership.

La valoración de McCarthy y Palcic (2007) tampoco es concluyente. En su trabajo sobre la *Eircom* ESOP comprobaron la evolución del número de empleados y de la productividad del

trabajo a partir de 1989 (cuadro 4). La plantilla disminuyó tanto antes como después de implantada la ESOP (*employee stock ownership plan*), en 1998, pero durante los nueve años previos lo hizo en un 15.23%, mientras que en los nueve años posteriores disminuyó en un 36.95% (última fila del cuadro 4). La productividad aumentó en ambos periodos, pero antes lo hizo en un 102.1% y después en un 132.1%. Parece claro que la ESOP agudizó ambas tendencias, pero ¿por qué ese mayor incremento en la productividad? Según el estudio de McCarthy y Palcic, el 60% de los trabajadores sienten que la ESOP no ha reducido su disposición a ir más allá de sus obligaciones laborales para ayudar a sus compañeros. Además, el 43% de ellos ponen más esfuerzo en su trabajo y son más productivos desde la implantación de la ESOP. Sin embargo, los autores no se atreven a discernir si la mejora en la productividad se debe a un cambio de comportamiento en los empleados o a la reducción de la plantilla, simplemente afirman que es posible que sea lo segundo.

CUADRO 4. Número de empleados y productividad en *Eircom*

Año	Nº empleados	Productividad	Año	Nº empleados	Productividad
1989	14.434	1.000	1998	11.497	1.000
1990	14.243	1.149	1999	11.323	1.148
1991	14.044	1.283	2000	11.689	1.177
1992	13.676	1.365	2001	11.593	1.340
1993	13.313	1.420	2002	10.338	1.505
1994	13.069	1.464	2003	9.129	1.742
1995	12.662	1.605	2004	8.306	1.836
1996	12.372	1.856	2005	7.595	2.098
1997	12.236	2.021	2006	7.249	2.321
Total acumulado	**-15.23%**	**+102.1%**		**-36.95%**	**+132.1%**

Fuente: McCarthy y Palcic (2007)

5. EL PARTNERSHIP A NIVEL DE EMPRESA

Hemos visto en el apartado anterior que una de las críticas que ha recibido el enfoque de Partnership es el de haberse atrincherado a nivel macro, es decir, a nivel de la negociación colectiva y la política macroeconómica. En otras palabras, que la perspectiva de Partnership no ha llegado a la empresa. Ya hemos considerado algunas posiciones y cifras al respecto que no son excesivamente optimistas. Veamos cuál es la situación en Irlanda con algo más de profundidad.

El énfasis de un ambiente de colaboración en el seno de la empresa está en buscar soluciones de ganancias mutuas, tanto para los empresarios como para los trabajadores. Este punto de vista se apoya, en parte, en las deficiencias percibidas por el enfoque de relaciones industriales de adversarios, especialmente en la aparente dominancia de la negociación colectiva en temas de corto plazo, con la perspectiva de reparto de recursos limitados. Este enfoque desarrolla posiciones de adversarios, basadas en la creencia de que sólo puede haber ganancias a costa de las pérdidas de la otra (Fischer y Ury, 1986; citado en Gunnigle, 1999). Según Gunnigle, este punto de vista se basa en que los trabajadores se comprometen a mejorar la productividad y, a cambio, los empleadores reconocen y facilitan la implicación de los trabajadores y sindicatos en la toma de decisiones estratégicas de la empresa y, además, garantizan que las ganancias obtenidas como consecuencia de la mayor productividad se repartan entre trabajadores y empresarios.

Pero, ¿este punto de vista es real en la práctica de las empresas irlandesas? Gunnigle se basa en los datos de la *Partnership Based Industrial Realtions Arrangements* (PBIRA), encuesta llevada a cabo por el *Irish Management Institute* (IMI) en 1997 (Hannigan, 1997), para afirmar que el enfoque de Partnership no está bien desarrollado en las empresas irlandesas. En concreto, ninguno de los indicadores directos de Partnership (es decir, aquellos que reflejan la existencia

del pacto comentado anteriormente), no obtienen el valor medio. Por ejemplo, en un rango de 1 a 5, la existencia de acuerdo formal de Partnership sólo obtiene 2.05; la implicación del empleado en la toma de decisiones estratégicas, 2.10; y el reparto de las ganancias del éxito, 2.50.

Pero Gunnigle cita otras fuentes de datos. La encuesta de la *University College Dublin* (UCD) (1998) aporta cifras sobre el cambio en los lugares de trabajo, tanto en las empresas con presencia sindical como en las que no. Para las primeras los datos indican que el Partnership es más la excepción que la norma y en aquellos lugares donde es implantado es más en relación a temas operativos que a estratégicos. Por el contrario, en las segundas existe una mayor implicación del empleado, tanto en decisiones operativas como estratégicas.

De estas y otras fuentes, Gunnigle concluye que la implicación directa del empleado en la empresa es limitada. En aquellos lugares donde existe, parece que la tónica general es facilitar la implicación de empleados individuales y de pequeños grupos sobre temas relacionados con su trabajo inmediato, dirigidos más a darles "voz" que a incrementar su influencia sobre temas estratégicos de la empresa.

Gunnigle y Brosnan citan a D'Art y Tunner (1999) para matizar la persistencia del "ellos y nosotros" en la relaciones industriales irlandesas. Estos autores llevaron a cabo una encuesta en nueve compañías con presencia sindical en el sector de manufacturas. Los resultados indican que no existen evidencias de una reducción de las actitudes del tipo "ellos y nosotros". Una mayoría de los empleados encuestados (56%) no estaba de acuerdo con la declaración de que los trabajadores y los managers están en el mismo lado, mientras que un 30% sí estaba de acuerdo. Por otra parte, Gunnigle y Brosnan opinan que las mejoras en competitividad se pueden lograr tanto a través de una negociación basada en un enfoque de Partnership, como a través de otra de tipo distributiva o de adversarios. Además, el acuerdo a nivel nacional *Partnership 2000* ha generado una serie de acuerdos a nivel de empresa, pero su número es todavía relativamente pequeño. En cualquier caso, los acuerdos colectivos implicando un juego de suma positiva, es decir, contribuyendo tanto a la creación de empleo como a la competitividad, son raramente encontrados en la práctica.

Según Teague (2004), en el sector privado, existen dos formas para medir la extensión o alcance del Partnership a nivel de empresa. En primer lugar, mediante el número de acuerdos establecidos a nivel de empresa en los que se contemplen algunos de los temas relacionados con este enfoque. Por ejemplo, en el cuadro 5 tenemos los acuerdos llevados a cabo por el *Services, Industrial, Professional and Technical Union* (SIPTU) bajo el auspicio de los acuerdos *Partnership 2000* (P2000) y *Programme for Prosperity and Fairness* (PPF), que ya hemos comentado.

CUADRO 5. Acuerdos alcanzados por el SIPTU dentro de los programas P200 y PPF

	P2000	**PPF**
Implicación del empleado	5	1
Formación y ayudas	32	2
Igualdad de oportunidades	39	4
Acuerdos representativos	39	3
Implicación financiera	83	15
Composición de la fuerza de trabajo	8	3
Nuevas formas de organización del trabajo	58	7
Sanidad y seguridad	14	3
Resolución de conflictos	7	3
Adaptabilidad, flexibilidad, etc…	25	1
Impuestos	2	1

Fuente: SIPTU, citado en Teague (2004)

A primera vista, las cifras parecen impresionantes. Pero Teague hace la siguiente advertencia: no todos los acuerdos indicados en la tabla han sido concluidos realmente. Muchos de los acuerdos denominados de Partnership pueden ser simplemente extensiones de acuerdos ya existentes, que fueron negociados desde la perspectiva tradicional de adversarios. Sin embargo, Teague también dice que esas cifras no son un inventario exhaustivo de la actividad empresarial de Partnership en Irlanda. En primer lugar, porque solamente incluye a los acuerdos firmados por el SIPTU en beneficio de sus miembros y, en segundo lugar, porque solamente incluye los acuerdos formalizados, por lo que no se está teniendo en cuenta mucha actividad de Partnership informal que existe a nivel de empresa.

Otra forma que Teague defiende para medir la extensión del Partnership es mediante la investigación del uso de algunas prácticas modernas de trabajo asociadas a unas relaciones más participativas por parte de los empleados. Veamos las cifras que aporta Teague basándose en cinco trabajos anteriores (cuadro 6). Dos cuestiones destaca Teague de esta tabla. Por un lado, que parece ser que hay una amplia experimentación con nuevas formas de prácticas de trabajo que implican algún tipo de Partnership. Por otro lado, que la mayoría de las prácticas son utilizadas en un mínimo de 35% de los establecimientos, porcentajes que aumentan considerablemente cuando se trata de prácticas relativas a la calidad, como al TQM y el control de calidad on-line.

CUADRO 6. Incidencia de las prácticas participativas en Irlanda (porcentajes)

	Rotación de trabajo	Trabajo en equipo	Prácticas consultivas	TQM	Just-in-time	Mantenimiento on-line	Calidad On-line
European Foundation (1997)	-	42.0	36.0	-	-	-	-
IBEC (1999)	-	63.0	37.0	37.0	33.0	-	54.0
Doherty (1998)	-	27.0	70.0	-	-	-	-
McCartney y Teague (2004)	56.9	27.5	44.1	68.7	48.3	39.2	57.8
Roche y Geary (2000)	-	59.0	45.5	71.1	-	-	-

Fuente: Teague (2004)

En lo que respecta a la profundidad de los cambios, Teague aporta otros dos cuadros, de los que concluye, por un lado, que la mayoría de los establecimientos que están experimentando con nuevas formas de organización del trabajo, implican a más de la mitad de sus empleados y, por otro lado, que los empleadores tienden más a delegar autoridad en asuntos operativos. Ellos están menos dispuestos a ceder el control en temas de asistencia, disciplina o presupuesto.

En cuanto al sector público, Teague opina que, en general, los acuerdos de Partnership se han quedado por debajo de las expectativas iniciales. Las opiniones difieren sobre las causas de esta diferencia entre expectativas y experiencia real. Una explicación plausible es que el desarrollo del Partnership estaba ligado a programas de modernización del sector público, sin estar suficientemente conectados a la prevalencia de la negociación colectiva. La gran lesión, para Teague, de la experiencia del Partnership en el sector público irlandés, es que los principios del Partnership han de ser aplicados clara y consistentemente a todos los niveles de la organización.

En otro trabajo, este mismo autor (Teague, 2002) hace la siguiente valoración sobre el Partnership realmente existente a nivel de empresa:

- Un problema importante son las reticencias dentro de las organizaciones: los managers porque piensan que puede minar su autoridad y los sindicatos porque puede debilitar la negociación colectiva.
- El Partnership es generalmente encontrado en aquellas organizaciones con una historia de relaciones de empleo difíciles. Esto significa que las prácticas de empleo asociadas con los acuerdos de Partnership tienden a no ser excesivamente innovadoras.

- La mayoría de los acuerdos se centran en asuntos operativos y, solo ocasionalmente, en asuntos estratégicos.
- Algunas veces, los acuerdos de Partnership con éxito no se han conectado apropiadamente con sistemas de implicación de los empleados.

La opinión de Roche y Geary (2000) es que el nivel de cambio en los lugares de trabajo y el rango de temas afectados parecen muy significativos, aunque es difícil estimar la profundidad de dichos cambios. En los lugares de trabajo con presencia sindical, los datos demuestran claramente que los empleadores en general han continuado regulando los mismos mediante la negociación colectiva y un control unilateral por parte de los managers. El Partnership en los lugares de trabajo sólo ha introducido cambios en un 20% de los casos. En los lugares sin presencia sindical, la incidencia de la implicación directa de los empleados para dirigir el cambio es realmente significativa, aunque en ocho de doce posibles áreas de cambio, y en todas las áreas de cambio estratégico, la mayoría de los establecimientos confiaban en una dirección exclusivamente de los directivos.

6. EL PAPEL DE LOS SINDICATOS

Como parece obvio, el papel que en general debe jugar los sindicatos en el enfoque de Partnership resulta crucial. Sin embargo, su postura inicial se puede calificar como de apatía. Tal apatía tiene su origen en las numerosas dudas sobre las implicaciones de las prácticas participativas en el papel de los sindicatos respecto a la negociación colectiva (Gunnigle, 1999). Los resultados preliminares de este autor dan cierto aval a estas reservas sindicales. Pero a mediados de los años noventa los sindicatos cambiaron su postura (ICTU, 1993, 1995). Estos documentos se centran especialmente en la participación de tareas a nivel de empresa.

Gunnigle cita el trabajo de D'Art y Turner (1999), consistente en una encuesta a 2000 afiliados de un gran sindicato en Irlanda. Dichos afiliados trabajaban en cuarenta y tres organizaciones y el ratio de respuesta de la encuesta fue del 29%. El 39% de las respuestas indicaban que en su puesto de trabajo no se había dado ninguna de las iniciativas investigadas en el estudio: reuniones sindicato-directiva, trabajo en equipo, participación en los beneficios, participación en la sociedad, mayor control a los empleados y mayor información financiera. Pero lo que nos interesa aquí es ver cuál es la percepción que tienen los encuestados sobre la influencia de las nuevas iniciativas sobre su trabajo (dos primeras filas del cuadro 7), sobre las relaciones institucionales dentro de la empresa (filas 3 y 4) y sobre sus ganancias (fila 5).

CUADRO 7. Influencia de las nuevas iniciativas (porcentajes)

	Fuerte disminución	Disminución	Sin cambios	Aumento	Fuerte aumento
Influencia en decidir cómo hacer tu trabajo	10	14	60	15	1
Cantidad de trabajo que tienes que hacer	1	2	32	42	23
Sentimiento de que el "ellos y nosotros" divide	5	12	53	18	12
Influencia del sindicato en tu lugar de trabajo	12	26	47	13	2
Un reparto más justo de los beneficios	12	10	63	15	0

Fuente: D'Art y Turner (1999), citado en Gunnigle (1999)

En lo que respecta a la influencia sobre su trabajo, estos resultados no son muy esperanzadores: por un lado, el 86% de los encuestados piensan que no ha cambiado o ha aumentado y, por otro lado, el 97% piensa que su carga de trabajo ha aumentado o permanecido

igual. Además, una de las premisas del Partnership, la reducción de la visión del "ellos y nosotros" tampoco se cumple, ya que el 83% cree que ha aumentado o permanecido constante. Tampoco se cumple la premisa de la decisión de participar en el enfoque de Partnership por parte de los sindicatos, el aumento de su influencia, ya que los resultados indican que el 85% opina que ha disminuido o permanecido igual. Esto da cierto apoyo a las reticencias iniciales de los sindicatos.

Según Gunnigle, hay amplia evidencia de iniciativas por parte de los empleadores para un mayor nivel de implicación directa de los empleados en la toma de decisiones operativas. Tales iniciativas tienen como objetivo fundamental la mayor implicación del trabajador individual y no necesariamente incluye a los sindicatos. En general, el *Irish Congress of Trade Unions* (ICTU) y el *Irish Business and Employers' Confederation* (IBEC) tienen una idea diferente sobre el Partnership. Mientras que el primero piensa en conseguir un mayor poder en la toma de decisiones estratégicas y el reparto de beneficios de la compañía, por parte de los trabajadores y sindicatos, el IBEC piensa más en un mayor poder en la toma de decisiones operativas, así en programas de implicación directa de los trabajadores, tales como el trabajo en equipo y las iniciativas de calidad. Los miembros del ICTU sienten que su sindicato ha incrementado su influencia a nivel nacional, en temas como el desempleo y los impuestos, sin embargo también sienten que la capacidad de los sindicatos para ganar reconocimiento por parte de los empleadores ha disminuido. En general, piensan que la influencia y la eficiencia del sindicato en la empresa se han reducido. Finalmente, Gunnigle cree que desde los años 80 ha habido un deterioro en la calidad de vida en el trabajo. Como muchos trabajadores temían, la carga y el ritmo del trabajo han aumentado, sin ninguna compensación en términos de una mayor influencia en la actividad diaria.

La paradójica evolución de los sindicatos es también comentada por Teague (2002). Los activistas sindicales son conscientes de que los diferentes gobiernos bajo el enfoque del Partnership, han permitido a los sindicatos acceder a la toma de decisiones relevantes sobre política social y económica. Sin embargo, los mismos gobiernos también han adoptado determinadas políticas que han dificultado el reclutamiento a nivel de empresa por parte de los sindicatos.

CUADRO 8. Tendencia en la densidad de los sindicatos

	Nº de miembros	Densidad (%)
1975	449.520	60.02
1980	527.960	61.93
1985	485.050	61.31
1990	474.590	57.10
1995	504.450	53.07
1999	561.800	44.50

Fuente: Roche y Ashmore (2001) y Gunnigle et al (2002), citado en D'Art y Turner (2003b)

La tesis de la pérdida de influencia sindical es corroborada por los datos del cuadro 11, en donde se demuestra la tendencia decreciente de la densidad de los sindicatos desde mediados de los años 80. Para D'Art y Turner (2003) esta pérdida de influencia no se ha detenido como consecuencia de la Industrial Relations (Amendment) Act 2001, ya que esta nueva legislación es poco probable que mejore el acceso de los sindicatos a los puestos de trabajo, especialmente donde los empleadores se oponen. De los 39 casos por ellos analizados, en sólo uno de ellos el empresario concedió el reconocimiento del sindicato. Otro defecto que ven ellos en esta enmienda es la ausencia de protección para los miembros de los sindicatos implicados en un

litigio de reconocimiento. Por consiguiente, para ellos algunas consecuencias no deseadas de la misma son la pérdida de determinación para el reconocimiento de los sindicatos, la exclusión de los sindicatos de los lugares de trabajo y la legitimización del estatus de empresa sin presencia sindical. La idea que puede estar implícita en esta enmienda es la pertinencia de la representación no sindical de los trabajadores. La posición de Gottlieb (1998) es que sólo la amenaza de la presencia sindical puede salvaguardar los intereses de los empleados. La representación no sindical puede funcionar, pero sólo en el contexto de una amenaza viable de presencia sindical.

7. EL FUTURO DEL PARTNERSHIP

El futuro del modelo irlandés de Partnership es incierto. O'Donnell (2002) destaca tres problemas que nosotros pensamos son de difícil solución en el contexto de crisis global actual: la debilidad de la empresa irlandesa frente a las multinacionales, el desempleo a largo plazo y el nivel de exclusión social que ello conlleva, y las dificultades para reflejar, en las relaciones industriales existentes en los lugares de trabajo, los acuerdos de Partnership alcanzados a nivel nacional.

Para este mismo autor, el futuro del Partnership tiene seis posibles opciones. La primera es el abandono del mismo, defendida por aquellos que piensan que distorsiona las fuerzas del libre mercado y/o dificulta el desarrollo de la democracia parlamentaria. Para O'Donnell la principal objeción que tiene esta alternativa es que no está teniendo en cuenta los problemas reales de coordinación existente en la economía. La experiencia ha demostrado que si cada individuo o grupo social se comporta como *free rider* algunos bienes públicos (incluidos la competitividad y el pleno empleo) pueden desaparecer.

La segunda opción es la vuelta al enfoque de Partnership original, del que nos hemos alejado a partir del PPF. Su principal limitación está en suponer que la única elección posible está entre un pacto social completamente centralizado o la salida del Partnership, e Irlanda puede perfectamente avanzar hacia un sistema de pactos parcialmente centralizado, lo que puede ser conveniente, tanto por razones macroeconómicas como para las propias empresas.

La tercera es volver a la negociación distributiva original, que incluye salarios, impuestos y seguridad social, ya que ella tendría un carácter más estratégico y sería más fácil de comunicar. Para O'Donnell una negociación centrada exclusivamente en esos aspectos no tiene por qué ser más estable. Dejar fuera el cambio organizacional probablemente no estabiliza el Partnership.

La cuarta es la opuesta a la tercera: un Partnership sin acuerdos sobre salarios. Pero, con o sin Partnership, Irlanda ha diseñado un apropiado sistema de negociación de salarios y la discusión sobre su inclusión o no en el Partnership parece estéril. Además, esta perspectiva ignora los beneficios que en materia de coordinación en determinadas áreas proporciona dicha negociación. Por otra parte, algunos agentes sociales estarían significativamente menos interesados en los acuerdos si estos no incluyen los pagos.

Las opciones quinta y sexta son las que O'Donnel considera más apropiadas: el usar el Partnership para alcanzar los cambios estructurales necesarios en la economía irlandesa y crear un nuevo modelo de gobernanza pública y de Partnership. Esta última implica un nuevo papel para los agentes sociales, en donde la información es el recurso clave y el análisis, el diálogo y el conocimiento compartido los procedimientos adecuados; y un nuevo papel para el gobierno central, consistente en una visión más emprendedora de la política y en facilitar la comunicación y la acción conjunta entre los diferentes intereses sociales, así como el apoyo de la formación de grupos de interés.

Turok (2001) también cree necesario una serie de reformas institucionales que mejoren el funcionamiento del Partnership, especialmente a nivel local:

- Un mayor apoyo a nivel nacional, así como la movilización de funcionarios y organizaciones a nivel local.

- Clarificar el modelo básico de gobernanza local. Especialmente en lo que se refiere a la ambigüedad actualmente existente sobre la importancia relativa de, y la relación entre, las autoridades locales y las organizaciones de los diferentes agentes sociales.
- Crear una red, lo más sencilla posible, de organizaciones locales que aumente la conciencia de las barreras a la inclusión, asegure un acceso igualitario a los servicios, asegure la calidad de los mismos y apoye el desarrollo de las organizaciones locales.
- Una mayor difusión de la experiencia adquirida en el Partnership, tanto a nivel local como nacional. Esto incluye tanto lo que ha funcionado, como lo que no; los proyectos que tengan espacio para un mayor desarrollo y aplicación; y el valor añadido que ha proporcionado el Partnership.
- Fortalecer la presencia de los sindicatos y empresarios en los Partnerships locales. Por otra parte, el modelo tripartito de Partnership local (sector público, comunidad y otros agentes sociales) podría ser revisado.

La visión de Gunnigle (1999) para el futuro tampoco es muy optimista. Los empleadores han sido tradicionalmente reacios a compartir decisiones de poder, especialmente las de carácter estratégico. Uno de las razones más relevantes de este miedo es la creencia de los empresarios de que reducirá la capacidad de la organización para la toma de decisiones importantes, por consiguiente su capacidad para hacer frente a los retos de una competitividad creciente. Por otra parte, están las mayores dificultades de implementar acuerdos de Partnership en las compañías de propiedad extranjera, de gran presencia en Irlanda, ya que sus principales decisiones estratégicas se toman fuera del país. El Partnership a nivel estratégico está más presente en las organizaciones irlandesas y, particularmente, en ciertas áreas del sector público. Aquí sí ha habido significativos avances, aunque hay que advertir que el Partnership se encuentra en estado embrionario, y sus perspectivas para el futuro dependen de cómo estas organizaciones se enfrente a los retos competitivos inminentes en sus mercados de productos respectivos.

REFERENCIAS BIBLIOGRÁFICAS

D'ART, D. and TURNER, T., 1999. *An Attitudinal Revolution in Irish Industrial Relations: The End of 'Them and Us'?* British Journal of Industrial Relations, Vol. 37, nº. 1, March, pp. 101-116.

D'ART, DARYL and THOMAS TURNER, 2003. Union Recognition in Ireland: One step forward or two steps back? Industrial Relations Journal, Vol. 34, Agosto, pp. 226-240.

GOTTLIEB TARAS, DAPHNE, 1998. *Nonunion Representation: Complement or Threat to Unions?* Industrial Relations Research Association Series, Reunión Anual, Vol. 1, Enero.

FISHER, R. & URY, W. 1986, Getting to Yes, London: Hutchinson.

GUNNIGLE, PADDY, 1999. *Involvement, participation and partnership: a review of the debate and some reflection on the Irish*. The 24th Countess Markeivicz Lecture, Noviembre

GUNNIGLE, PATRICK & KATHRYN BROSNAN. *Collective Bargaining, Employment Creation and Protection, and Competitiveness in the Republic of Ireland*. http://www.msu.edu/user/block/GunniglePaper.doc

GUNNIGLE, P., MORLEY, M., CLIFFORD, N. and TURNER, T., 1997. *Human Resource Management in Irish Organisations: Practice in Perspective*, Dublin: Oak Tree Press.

GUNNIGLE, P., O'SULLIVAN, M. and KINSELLA, M., 2002. *Organised Labour in the New Economy: Trade Unions and Public Policy in the Republic of Ireland*, en D'Art, D. and Turner, T. (eds.) *Irish Employment Relations in the New Economy*. Dublin: Blackhall Publishing.

HANNIGAN, K., 1997. *Partnership Survey*. Mimeo basado en la Irish Management Institute (IMI) survey on partnership. Dublin: Irish Management Institute.

IRISH CONGRESS OF TRADE UNIONS (ICTU), 1993. *New forms of work organisation: Options for unions.* Dublin: Irish Congress of Trade Unions.

IRISH CONGRESS OF TRADE UNIONS (ICTU), 1995. *Managing Change*, Dublin: Irish Congress of Trade Unions 1995.

MCCARTHY, DERMOT y DÓNAL PALCIC, 2007. *Employee Share Ownership Plans and Privatisation in Ireland.* http://www.irspm2007.org/Fullpaper/PT_PhD_McCarthy%20Palcic.pdf

MJOSET, L., 1992. *The Irish Economy in a Comparative Institutional Perspective*, Dublin: National Economic and Social Council, pp. 7-8.

NESC, 1996. *Strategy into the 21st Century*. Dublin, capítulo 1.

NESC, 2008. The Irish Economy in the Early 21st Century.

OCDE, 1995. *Profit Sharing in OECD Countries*, *OECD Employment Outlook*, Julio.

OCDE, 2008. International Migration Outlook.

O'DONNELL, RORY, 1999. *Social Partnership: Principles, Institutions and Interpretations,* en O'Connell, P.J., *Astonishing success: Economic Growth and the labour market in Ireland*, capítulo 4.

O'DONNELL, RORY, 2001. *The future of social partnership in Ireland.* Un trabajo preparado para el National Competitiveness Council, Mayo.

O'DONNELL, RORY, 2002. *Ireland's Economic Transformation. Industrial policy, European integration and social partnership*. University of Pittsburgh, Centre for West European Studies, Working Paper nº 2, Diciembre.

RHODES, R. A. W., 1997. *Understanding Governance: Policy Networks, Governance, Reflexivity and Accountability*, Open University.

ROCHE, W.K. y ASHMORE, J., 2001. Irish *Unions in the 1990s: Testing the Limits of Social Partnership*, en G. Griffin (ed.), *Changing Patterns of Trade Unionism*. Sydney: Mansell.

ROCHE, W. K. y J. F. GEARY, 2000. *"Collaborative Production" and the Irish Boom: Work Organisation, Partnership and Direct Involvement in Irish Workplaces.* The Economic Social Review, Vol. 31, nº 1, Enero, pp. 1-36.

TEAGUE, PAUL, 2002. *The Irish Experiment in Social Partnership*. Trabajo presentado al International Institute of Labour Studies, Génova.

TEAGUE, PAUL, 2004. *Social Partnership and the Enterprise Some Lessons From the Irish Experience*. European Political Economy Review, Vol. 2, nº1, pp. 6-35

TRAXLER, F, 1997. *The Logic of Social Pacts*, in Giusseppe Fajertag y Phillipe Pochet (eds.), *Social Pacts in Europe*, Brussels: European Trade Union Institute.

TUROK, IVAN, 2001. *Innovation in local Governance: The Irish partnership model*, en OCDE, *Local Partnerships for Better Governance*, París, pp. 135-173.

Buyer Power and Product Innovation with Partial Upstream Collusion*

Pedro Carreira†
Instituto Politécnico de Leiria
Universidade Nova de Lisboa

May 2009

Abstract

This paper aims to study the impact of retail concentration and increased buyer power on product variety and welfare. We provide a novel insight on this, by allowing for collusion between upstream manufacturer firms when choosing product location. We first show that, when retail stores conduct their sourcing strategy independently, manufacturers may have incentives to restrict product variety relatively to the social optimum, i. e., to produce a lower number of goods than what would be socially desirable. We next show that, if indeed manufacturers restrict product variety, some stores find it profitable to engage on mergers for buyer power. Finally, we show that buyer power is enhanced in such a way that manufacturers increase the number of goods produced. This theory may explain the recent "variety wars" (rival manufacturers expanding their product list) observed in food, beverage and some other markets, in many EU countries.

Keywords: Single Sourcing, Buyer Power, Upstream Collusion, Product Variety.

JEL Classification: L14

---

*This paper forms the first chapter of my Economics PhD dissertation at Universidade Nova de Lisboa (Faculdade de Economia). I gratefully acknowledge financial support by Fundação para a Ciência e a Tecnologia. I thank my PhD supervisor Steffen Hoernig for his guidance and helpful comments.

†Equivalent to assistant professor, Departamento de Gestão e Economia, Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Leiria. Adress: Campus 2, Morro do Lena - Alto do Vieiro, Apartado 4163, 2411-901 Leiria, Portugal (mailbox: D172). Phone Numbers: +351 244 820300 (ext: 3629), +351 93 4454540, +351 91 4258410. Fax : +351 244 820310. Email: pedro.carreira@estg.ipleiria.pt.

1 Introduction

It is a well-known fact that retail markets are becoming increasingly concentrated in many OECD countries and particularly in the European Union. As reported in Dobson (2002), the top ten retailers in the EU account now for more than 30% of all sales of food and dairy products. While retailers were traditionally seen as lacking in power on wholesale markets, this recent process of retail concentration has created market structures characterized by bilateral oligopolies, where each retailer accounts for a relatively large share of each supplier's sales. Furthermore, in many markets, retailers are able to achieve local differentiation and to enjoy significant market power at their outlets, facts that are mainly explained by the consumers' preferences for one-stop-shopping and by an increasing segmentation of retail formats.

A major welfare concern of increased buyer power is its impact on manufacturers' incentives to undertake welfare-enhancing activities, like product innovation or cost minimization. Even though Inderst and Wey (2003) show that incentives for upstream firms to reduce marginal costs may increase following a downstream merger, there is a generalized feeling that buyer power reduces such upstream incentives. Indeed, according to an often expressed view, buyer power, by reducing suppliers' profits, may be socially detrimental where it undermines the long-term viability of suppliers and their willingness to commit to new product and process investments. For example, Dobson et al. (1999) argue that retail concentration "can have an economic impact when [...] buyer power reduces prices for suppliers, and thus their income, making it difficult for them to finance required investments, which might then be postponed or even foregone completely". Hence, "buyer power can potentially undermine both upstream competition between suppliers and distort downstream competition when it serves to exaggerate differences in competitive positions", as reported in Dobson et al. (1998). Furthermore, using an analytical framework, Inderst and Shaffer (2007) show that, following a merger, a retailer may want to enhance its buyer power by committing to a single-sourcing purchasing strategy. In addition, they show that, anticipating further concentration in the retail industry, suppliers will strategically choose to produce less differentiated products, which further reduces product variety. If buyer-seller negotiations are efficient, the overall loss in product variety may reduce consumer surplus and total welfare, even though with linear tariffs there may be a countervailing effect as the more powerful retailer passes on lower prices to final consumers. Finally, Chen (2004) finds that increased countervailing/buyer power reduces the number of varieties produced by an upstream monopolist.

There are nevertheless many markets where the contrary can be observed. For example, in Portugal, after two decades of increasing retail concentration in food product markets, the two main beer brands Sagres and Super Bock have started a "variety war", with each of them adding many new beer varieties to its product list. Concretely, from only two beer products "Sagres" and "Sagres Preta" since 1940, Sagres recently added eight more beer varieties "Sagres Bohemia" (2005), "Sagres Zero" (2005), "Sagres Chopp" (2006), "Sagres Bohemia 1835" (2006, limited edition), "Sagres Limalight" (2007), "Sagres Zero Preta" (2007), "Sagres Bohemia D'Ouro" (2007) and "Sagres Zero Limalight" (2008). In the meantime, his rival Super Bock added to its product list "Super Bock Stout" (2003), "Super Bock Green" (2004), "Super Bock Twin" (2005), "Super Bock Abadia" (2006), "Super Bock Cool" (2006), "Super Bock Tango" (2006), "Super Bock Sem Álcool" (2007), "Super Bock Sem Álcool Preta" (2007) and "Super Bock Sem Álcool Pêssego" (2007).

Likewise, in the water market, five of the leading brands Luso, Vitalis, Vidago, Frize and Pedras have also greatly expanded their product list. Frize, a recent entrant (1994), rapidly added "Frize Limão" (2002), "Frize Tangerina" (2003), "Frize Figo" (2004), "Frize Ananás" (2004), "Frize Maracujá" (2005), "Frize Morango" (2006), "Frize Limão-Cola" (2006), "Frize Maçã Verde" (2007), "Frize Lima" (2007) and "Frize Groselha" (2007) to its product list. Luso created "Luso Fresh" (2005), "Luso Fresh Limão" (2005), "Luso Fresh Goiaba-Lima" (2005), "Luso Fresh Laranja/Framboesa" (2006), "Luso Fresh Pêssego/Ananás" (2006), "Formas Luso Limão" (2006), "Formas Luso Morango" (2006), "Formas Luso Manga" (2006), "Formas Luso Chá Verde e Pêra" (2007), "Ritmo Luso Limão" (2007), "Ritmo Luso Cereais" (2007), "Ritmo Luso Kiwi" (2007) and "Formas Luso Maçã Verde temperada com Framboesas" (2008, limited edition). In his hand, Vitalis added "Vitalis Sabores Limão e Magnésio" (2006), "Vitalis Sabores Maçã e Chá Branco" (2006) and "Vitalis Sabores Ananás e Fibras" (2006). Vidago added to its product list "Vidago Sabores Limão" (2003), "Vidago Sabores Maçã" (2004) and "Vidago Sabores Maracujá" (2005). Finally, Pedras created "Pedras Levíssima" (2005), "Pedras Sabores Limão & Chá Verde" (2006), "Pedras Sabores Pêssego & Chá Branco" (2006), "Pedras Sabores Framboesa & Ginseng" (2006), "Pedras Sabores Manga & Ginkgo Biloba" (2006), "Pedras Sabores Melão & Hortelã" (2006, limited edition) and "Pedras Sabores Lima & Menta" (2007, limited edition).

We argue that these "variety wars" may be explained by the retailers' increased buyer power resulting from the recent process of retail concentration and not only by market share or market development reasons. Our argument

is that, when manufacturers are able to choose the location of their products cooperatively, buyer power can erode the profitability potential from reducing product variety. Indeed, we first show that, when retail stores purchase as independent firms, manufacturers may have incentives to restrict product variety, i. e., to jointly agree on a reduction in the number of goods produced relatively to what would be socially desirable. We next show that, if indeed manufacturers restrict the number of goods produced, some retailers have incentives to merge in order to enhance their buyer power.[1] Finally, we show that the way retailers enhance their buyer power induces manufacturers to reposition their products by innovating and increasing the number of goods produced.

The possibility of upstream partial collusion on product location constitutes the critical assumption of our paper. We present two main reasons for such a consideration. First, "(...) collusion seems to be more prevalent in intermediate goods industries, perhaps because these industries are often more concentrated or sell goods which are more homogeneous", as reported in Nocke and White (2005). Indeed, many famous cases of collusion documented in the literature have involved intermediate goods industries. Second, as argued by Posada and Straume (2004), "(...) antitrust legislation may make price coordination infeasible, or at least difficult. It is reasonable to assume, though, that coordination of relocation decisions is much less likely to be prohibited by antitrust authorities". In fact, it may be difficult for authorities to detect and prove collusive behavior on product location. Furthermore, deviations from such collusive behavior usually involve large investments, which, in one hand, implies that they pay to a deviator firm only after a significant period of time and, in other hand, that they are easily detectable by the remaining collusion insiders. Hence, punishment may occur even before a deviator completely realizes its deviation profits.

The paper is organized as follows. In section 2, we present our bilateral oligopoly model, its main assumptions and our five-stage game. In section 3, we consider the case where each retail store conducts its sourcing strategy independently, i. e., where there is no retail chains. In section 4, we allow for retail mergers for buyer power and analyse its effects on product variety and welfare. In section 5, we analyse the robustness of the results, by discussing some of the main assumptions of the paper, and provide some concluding remarks.

[1] In truth, forming a buyer alliance would be sufficient. Indeed, in our paper, a buyer alliance is equivalent to a merger, given that we will assume separate retail markets. Throughout the paper, we will use the term merger though.

2 The Model

We consider a bilateral oligopoly model, constituted by four local monopolist retail stores $i = 1, 2, 3, 4$, each of which operates in local market i, and two upstream manufacturers $j = a, b$.

Manufacturers compete for supplying the retail stores but are able to coordinate their product location decisions. Each of them has limited capacity in the sense it can produce at most two varieties of a certain product. Formally, each supplier j may produce any subset X_j of varieties from the feasible set $X = \{x_1, x_2, x_3, x_4\}$, with $\#_{X_j} \leq 2$. Each feasible product variety x_t is assumed to maximize consumer valuation at local market i. Moreover, we assume that there are neither fixed nor variable costs of production, which is important for ensuring that any reduction on product variety undertaken by suppliers has purely strategic motives.

On the downstream side of the economy, in order to focus on variety rather than quantity, we assume that each local market is characterized by a representative consumer who buys at most one unit of only one variety of the product. We assume that the representative consumer of local market 1 prefers variety x_1 to variety x_2, variety x_2 to variety x_3, and variety x_3 to variety x_4. Likewise, the representative consumer of local market 4 prefers variety x_4 to variety x_3, variety x_3 to variety x_2, and variety x_2 to variety x_1. With respect to the representative consumer of local market 2, we assume that he prefers variety x_2 to varieties x_1 and x_3, which he values equally, and varieties x_1 and x_3 to variety x_4. Similarly, the representative consumer of local market 3 prefers variety x_3 to varieties x_2 and x_4, which he values equally, and varieties x_2 and x_4 to variety x_1.

Formally, let $v_i(x_t)$ be the valuation of the representative consumer of local market operated by store i when it sells variety x_t, $i, t = 1, 2, 3, 4$. We assume that

$$v_i(x_t) = v - \delta |t - i| \tag{1}$$

This pattern of consumer preferences is clearly valid for many markets, specially for food products. Indeed, it may be very easy to observe at an international level. As an example for Europe, german sausages are clearly valued the highest in Germany, followed by nearby countries. Likewise, one must expect portuguese sardines to be valued the highest in Portugal, followed by Spain and France. For more distant countries, they are likely to be perceived as of very low value.

By the same token, at a national level, one can think of traditional regional gastronomy. Regional bakery may be of high value only for local and nearby consumers. A different example may be the newspaper industry. A

local newspaper is probably of higher value than a national one. Furthermore, while a newspaper from a neighbour region may detain some significant value, a newspaper local to some distant region may be of very low value.

We now present our five-stage game.

In the first stage (merger stage), each retail store decides whether or not to make a merger proposal for *single sourcing* to at most one of the other stores. A merger occurs and a retail chain is formed whenever there are two matching merger proposals. For simplicity, we assume that, independently from the merger proposal it makes, a store believes that it is always matched and that the remaining two stores merge with a exogenous probability $\mu \in (0,1)$. Furthermore, if it indeed makes a merger proposal, it makes it to the store with which it achieves the highest increase in joint profit. Finally, by adopting *single sourcing*, we mean that a retail chain commits to stock both its stores with goods from a single supplier.[2] In section 3, where each outlet is assumed to conduct its purchasing strategy independently, this stage is obviously trivial. We thus consider it only at section 4.

In the second stage, manufacturers cooperatively choose how many and which varieties to produce. We consider the following collusion technology: manufacturers agree on the product location pattern that maximizes their joint profit if there is no symmetric pattern that gives a higher individual profit to at least one of them. If there is, they agree on the symmetric location pattern that induces the highest profit.[3,4] If indifferent between some location patterns, we assume that manufacturers prefer the pattern where they produce the lowest number of goods.

In the third stage, each manufacturer j competitively offers wholesale contracts (T_{ij}, x_{ij}) to each outlet i, where T_{ij} and x_{ij} represent the wholesale tariff and the product variety offered in the contract, respectively. When there is a pair of merged outlets adopting *single sourcing*, say outlets 1 and 2 without loss of generality, each manufacturer j offers a wholesale contract given by $\left(T_{(1+2)j}, x_{1j}, x_{2j}\right)$, where $T_{(1+2)j}$ stands for the wholesale tariff charged to the retail chain. Obviously, $x_{ij} \in X_j$.[5,6]

[2] This is precisely as in Inderst and Shaffer (2007).

[3] Symmetric location patterns are the ones where both manufacturers get the same individual profit.

[4] This is a collusion technology widely adopted in literature. It is specially reasonable if side payments are illegal and easily detectable by authorities.

[5] Note that, as we have considered separate local markets with unit demand, we have no need of specifying the wholesale tariff type. Indeed, wholesale tariffs such as T_{ij} and $T_{(1+2)j}$ may represent either linear or non-linear wholesale tariffs, including two-part tariffs.

[6] Note further that, by considering that manufacturers make take-it-or-leave-it offers

In the fourth stage, retailers choose one between the two wholesale contracts they are offered.[7] We assume that, in case of indifference, a retailer chooses the wholesale contract offer that induces a higher consumer valuation at its store. If still indifferent, we assume that it prefers to be supplied by manufacturer a, without loss of generality.

Finally, in the fifth stage, each retail store sets its retail price. All firms realize their profits.

3 Lack of Buyer Power Induces Minimal Product Diversity

In this section, we consider that retailers have no buyer power. There are no retail chains and each retail store conducts its sourcing strategy independently, i. e., manufacturers compete in separate auctions for supplying each of the stores. We show that, in this case, manufacturers restrict product variety, by producing a lower number of goods than in a competitive setting. We do it by solving our restricted game by backward induction.

3.1 Retail Prices

Clearly, given that it is a local monopolist, any retail store i sets

$$p_i = v_i(x_t) \tag{2}$$

where x_t is the variety contemplated in the wholesale contract it has accepted. Each store i thus sells one unit and has a profit given by

$$\pi_i = v_i(x_t) - T_i, \tag{3}$$

where T_i is the agreed wholesale tariff.

to supply the retail stores, we are assuming that they hold all bargaining power when negotiating wholesale contracts, which may be unrealistic given that retailers nowadays hold most of the bargaining power in vertical relations. Nevertheless, we allow for it in section 5 and show that the results of the paper remain unchanged.

[7] Vertical negotiations are thus equivalent to auctions. Indeed, by selling at most one unit, each retail store automatically commits to be supplied by a single manufacturer.

3.2 Contract Selection

Let store i receive wholesale contract offers (T_{ia}, x_{ia}) and (T_{ib}, x_{ib}) from manufacturers a and b, respectively.

It is optimal for store i to accept manufacturer a's contract if

$$v_i(x_{ia}) - T_{ia} > v_i(x_{ib}) - T_{ib}. \tag{4}$$

If $v_i(x_{ia}) - T_{ia} = v_i(x_{ib}) - T_{ib}$, store i prefers to be supplied by the manufacturer offering the variety that induces a higher v_i and, if still indifferent, it prefers to be stocked with a variety from manufacturer a, given our indifference assumptions.

3.3 Wholesale Contract Offers

At this stage, each manufacturer j set its wholesale contract offers (T_{ij}, x_{ij}) for each retail store i. Note however that X_a and X_b are still given here. Let the indice $-j$ represent the rival manufacturer.

Lemma 1 The following symmetric pair of strategies constitutes the unique Nash equilibrium for the wholesale contract offer game when retail store i conducts its sourcing strategy independently: manufacturer j sets

$$x_{ij}^* = \arg\max_{x \in X_j} v_i(x) \tag{5}$$

and

$$T_{ij}^* = \max\left\{v_i\left(x_{ij}^*\right) - v_i\left(x_{i(-j)}^*\right), 0\right\}. \tag{6}$$

Proof. See appendix A. ■

The equilibrium is intuitive. Given that manufacturers compete for supplying the retail stores, it is straightforward that each of them uses its "nearest" variety to a store when setting a wholesale contract offer to supply it. More formally, from X_j, manufacturer j offers to store i the good that maximizes local consumer valuation. Furthermore, the "nearest" manufacturer

to store i, i. e., the manufacturer j for which $v_i\left(x_{ij}^*\right)$ is higher, can set a strictly positive wholesale tariff and still supply the store, even when the other manufacturer offers a wholesale tariff of zero. How high can it be such that the "farthest" manufacturer does not have undercutting incentives?[8] By precisely the difference between the consumer valuations generated by the two varieties. By doing so, the "nearest" manufacturer assures that the "farthest" one can only supply store i if it offers a negative wholesale tariff, which is clearly not optimal since it results in losses.

More important, note that in equilibrium each store is stocked with the "nearest" good, from all the ones that are produced by the manufacturers, which is positive in terms of welfare.

3.4 Cooperative Choice of Product Location

At this stage, manufacturers cooperatively choose X_a and X_b. As a first step, we analyse the benchmark case of a competitive product location game.

Proposition 1 *In any Nash Equilibrium of the competitive product location game, manufacturers produce all four goods. Furthermore, any such product location pattern constitutes a NE. Welfare is maximized.*

Proof. See appendix B. ■

The first two results of the above proposition are easily explained. If one of the four varieties is not produced, there is at least one manufacturer which has incentives to produce it, either by replacing a former variety (product relocation) or by producing a new one (product innovation). Clearly, such deviating behaviour allows the deviator either to supply an extra store with a profit or to increase the profits it makes with the stores it already supplies. Moreover, when all four goods are produced, any possible individual deviation in location results either in a reduction on the number of stores supplied (it implies a profit loss with respect to these stores while not increasing the profits to be made with the stores it still supplies) or in lower profits made with the supplied stores.

Welfare is clearly maximized. Manufacturers set the maximum level of product variety, which assures that all stores sell the ideal product for its costumer.

[8] We call "farthest" manufacturer to outlet i to the manufacturer j for which $v_i\left(x_{ij}^*\right)$ is lower.

We now derive the coordinated choice of X_a and X_b and show that indeed manufacturers reduce product variety relatively to the competitive product location game.

Proposition 2 *If manufacturers coordinate their product location choice when no retail store is part of a retail chain and conducts its sourcing strategy independently, they agree on minimal product variety. More precisely, each of them produces a single good, ideally designed for the most distant local markets.*

Proof. See appendix C. ■

We thus observe that manufacturers reduce product variety if they are able to coordinate their product location choice, relatively to a competitive product location game. Intuitively, by not producing goods x_2 and x_3, manufacturers are able to enjoy from a higher differentiation. In particular, the consumer valuation induced by the variety offered by the farthest manufacturer to each of the stores 1 and 4 decreases. The outside option for each of these stores when negotiating with its nearest manufacturer thus looses some value. Hence, the nearest manufacturers are able to increase their wholesale tariffs and still supply the stores. More precisely, each of the nearest manufacturers to stores 1 and 4 is able to charge a wholesale tariff of 3δ.

Welfare decreases relatively to the social optimum by 2δ, since there are two local markets in which the stores are stocked with a good which their consumers value only at $v - \delta$ instead of v.

Propositions 1 and 2 together constitute the first main result of the paper. Upstream partial collusion on product location can harm welfare through a reduction in product variety relatively to a purely competitive setting. Another way to look at this is by arguing that manufacturers have little incentives to engage on product innovation or no incentives at all if able to partially collude on product location. Extended consumer preferences towards some new varieties of a certain product may thus not be attended by manufacturers, who may find it profitable to restrict to the two most differentiated varieties in order to soften their rivalry on the posterior wholesale contract offer game for supplying the retail stores.

4 Buyer Power Increases Product Diversity

Given the result in proposition 2, our questions for this section lie in whether retailers have incentives to merge for buyer power and in whether the way they do it induces manufacturers to increase product diversity, assuming that manufacturers are able to costlessly reposition their products or extend their list of varieties and that retailers anticipate their coordinated product relocation or innovation behaviour.

In order to answer such questions, we solve the unrestricted game.

4.1 Retail Prices and Contract Selection

Clearly, retail prices remain unchanged relatively to the previous section.

As for the contract selection stage, although the optimal strategy remains unchanged for a store that does not adopt *single sourcing*, i. e., which does not participate on a merger, there is a new equilibrium strategy for the ones that adopt it, i. e., for the ones participating on a merger.[9] Consider, without loss of generality, that stores 1 and 2 merge and adopt *single sourcing*. The merged firm accepts the wholesale contract offer of manufacturer a if

$$v_1(x_{1a}) + v_2(x_{2a}) - T_{(1+2)a} > v_1(x_{1b}) + v_2(x_{2b}) - T_{(1+2)b}. \tag{7}$$

Making again use of our indifference assumptions, if $v_1(x_{1a}) + v_2(x_{2a}) - T_{(1+2)a} = v_1(x_{1b}) + v_2(x_{2b}) - T_{(1+2)b}$, the merged retail firm accepts the wholesale contract offer inducing the highest $v_1(x_1) + v_2(x_2)$. If still indifferent, it prefers to be stocked with manufacturer a's good(s).

4.2 Wholesale Contract Offers

First of all, note here that we have a competitive wholesale contract offer game per retail firm and not per store.[10] Naturally, while we also have the same equilibrium for the stores not participating on a merger (described in lemma 1), there is a new equilibrium with respect to the merged stores. Consider, without loss of generality, a retail chain constituted by stores 1 and 2.

[9] We do not consider the possibility of mergers for *multiple sourcing*, since they would be perfectly equivalent to remain unmerged.

[10] In section 3, there was a perfect match between the number of stores and the number of retail firms.

Lemma 2 The following symmetric pair of strategies constitutes the unique Nash equilibrium for the wholesale contract offer game when stores $i = 1, 2$ are merged and adopt *single sourcing*: each manufacturer j sets

$$x_{ij}^{*} = \arg\max_{x \in X_j} v_i(x) \tag{7}$$

and

$$T_{(1+2)j}^{*} = \max\left\{v_1\left(x_{1j}^{*}\right) + v_2\left(x_{2j}^{*}\right) - v_1\left(x_{1(-j)}^{*}\right) - v_2\left(x_{2(-j)}^{*}\right), 0\right\}. \tag{8}$$

Proof. We omit it. It is very similar to the proof of lemma 1. ■

The intuition for this equilibrium is evident, as the one for lemma 1. Given the competition between manufacturers for supplying the stores, each of them offers the "nearest" good, from the ones it produces, to each store. Furthermore, the overall farthest manufacturer, i. e., the one for which $v_1\left(x_{1j}^{*}\right) + v_2\left(x_{2j}^{*}\right)$ is lower, undercuts until he arrives at $T_{(1+2)j} = 0$. Thus, the overall nearest manufacturer j may set a positive wholesale tariff and still supply the stores. The highest value for the wholesale tariff so as not to provide undercutting incentives to the farthest manufacturer $-j$, given the indifference assumptions, is $v_1\left(x_{1j}^{*}\right) + v_2\left(x_{2j}^{*}\right) - v_1\left(x_{1(-j)}^{*}\right) - v_2\left(x_{2(-j)}^{*}\right)$.

4.3 Cooperative Choice of Product Location

Proposition 3 *If retail stores* 1 *and* 4 *merge and adopt single sourcing, manufacturers produce all four goods, by setting* $X_j = \{x_1, x_4\}$ *and* $X_{-j} = \{x_2, x_3\}$*, and welfare is maximized. Otherwise, each of them produces a single good, ideally designed for the most distant local markets.*

Proof. We omit it here. It consists on multiple replications of the proof of proposition 2. It is available upon request to the author. ■

The proposition above states that manufacturers engage on product innovation and increase the number of goods produced relatively to the case

where each retail store conducts its sourcing strategy independently if and only if *single sourcing* is jointly adopted by stores 1 and 4. When stores 1 and 4 merge and adopt *single sourcing*, setting $X_j = \{x_1\}$ and $X_{-j} = \{x_4\}$ makes manufacturers not differentiated at all with respect to the retail chain, even though they are with respect to each of the stores separately. In this case, independently from the sourcing strategies adopted by the "central" stores, the best manufacturers can do is to differentiate the most with respect to the pairs of "margin" and "central" stores, i. e., to set $X_j = \{x_1, x_4\}$ and $X_{-j} = \{x_2, x_3\}$, which is clearly a NE for the competitive product location game. Hence, partial upstream collusion on product location looses its profitability whenever stores 1 and 4 merge.

4.4 Merger Stage

Using proposition 3, we now derive the equilibrium for the merger stage.

Proposition 4 *Two retail chains for single sourcing are formed. One constituted by stores 1 and 4 and other by stores 2 and 3.*

Proof. See appendix D. ■

The intuition of the result is as follows. If stores 1 and 4 do not merge together, manufacturers set $X = \{x_1, x_4\}$ and each of them gets a profit of $\pi_i = v - 3\delta$. In addition, if any of these stores makes a merger proposal to stores 2 or 3, manufacturers still set the same product location pattern, which implies that aggregate profits do not increase much, if they indeed do. However, if stores 1 and 4 merge together, manufacturers increase the number of goods produced by setting $X_j = \{x_1, x_4\}$ and $X_{-j} = \{x_2, x_3\}$, and they are able to increase their aggregate profits from $2v - 6\delta$ to $2v - 2\delta$. Clearly, making a merger proposal to store 1 is the equilibrium strategy for store 4, and vice-versa.

With respect to stores 2 and 3, they can never make a merger proposal that induces manufacturers to increase product variety. Thus, as we can see in appendix D, they benefit from the highest expected profit if leaving room for stores 1 and 4 to merge. Also, they can further increase their expected joint profit by making a merger proposal to each other, provided that they are not sure that stores 1 and 4 indeed merge.

We are now able to construct the following proposition.

Proposition 5 *When manufacturers are able to coordinate their product location decisions, buyer power is likely to increase product variety and welfare.*

Proposition 5 contains the second main result of the paper. Buyer power is likely to increase product innovation if manufacturers are able to coordinate their product location decisions.

If they anticipate manufacturers' collusive behaviour of increasing their differentiation with respect to their stores by not producing the two ideal goods for the "central" markets, retail stores 1 and 4 merge and adopt *single sourcing*. This greatly reduces the profitability potential for manufacturers of any product location pattern that features reduced product variety. More precisely, it is enough for inducing manufacturers to adopt a location pattern that maximizes product diversity and welfare, as it would arise in a competitive location game.

Enhanced buyer power through retail merges or buyer alliances may thus prevent the potential harmful effects of upstream collusion on product location and contribute to product innovation and, consequently, to product diversity and welfare.

5 Discussion and Concluding Remarks

In recent years, antitrust concerns on buyer power have significantly increased. The main reason is the often pronouced view (even though not properly formalized) that, by reducing ustream profits, increased buyer power may substantially erode upstream incentives to engage on welfare-enhancing activities, such as product innovation. Also, the Countervailing Power Hypothesis[11] has been questioned. Indeed, observed final prices and retailers' profits in many industries suggest that the potential benefits of increased buyer power are being mostly extracted by retailers themselves and are not being passed on to consumers at all.

Somewhat contrary to most of the literature, we have found that retail concentration may increase product variety and welfare, specially when manufacturers are able to coordinate their product location decisions.

We have first shown that when retail stores have no buyer power and conduct their purchasing strategy independently, manufacturers may have incentives to produce a lower number of goods than it would be socially desirable, which naturally implies welfare losses.

[11] First introduced by Galbraith (1952).

Next, we have shown that buyer power reduces those incentives, even though the way it is enhanced is through *single sourcing*, which consists on delisting some manufacturers. In concrete, we have found that *single sourcing* is indeed optimally adopted by some retail stores in such way that manufacturers relocate their products by engaging on product innovation and introducing some new goods, relatively to a competitive location game.

Our results depend however on some discussable assumptions. First, we have assumed that manufacturers can costlessly reposition their varieties or introduce new ones, which can be reasonable only for some products, and that retailers anticipate such behaviour, which may be reasonable only for large and powerful retailers. These two assumptions together allowed us to build the five-stage game as we did. If we had admitted a relocation or innovation cost, coordinated product relocation from $X_j \cup X_{-j} = \{x_1, x_4\}$ to $X_j = \{x_1, x_4\}$ and $X_{-j} = \{x_2, x_3\}$ after the merger for *single sourcing* between stores 1 and 4 would only occur for low enough repositioning and R&D costs. In other hand, if retailers were assumed not to foresee manufacturers' relocation or innovation behaviour, merger stage equilibria would not suffer relevant changes. Indeed, retail stores 1 and 4 would also end up merged.

Second, we have assumed separate local markets and no competition between retail stores. Allowing for it clearly originates substantial differences between linear and non-linear wholesale tariffs. With non-linear wholesale tariffs, pure strategy Nash equilibria for the wholesale contract offer game may not even exist. With linear wholesale tariffs, if the intensity of retail competition is strong enough, each manufacturer has incentives to decrease the wholesale prices charged to the stores it supplies. Therefore, when adopting location patterns that display reduced product variety, the stores for which manufacturers are less differentiated would get a higher relative demand, which clearly reduces their profitability when compared to the case of separate markets. Manufacturers would then probably adopt a welfare maximizing location pattern, even when all stores purchase independently. Nevertheless, if the intensity of retail competition is low enough, i. e., if each store has sufficient local market power, manufacturers may find it optimal to sustain the wholesale prices at a high level and sell low quantities to the stores they supply. In this case, the paper's results are likely to remain unchanged.

Third, by assuming that manufacturers make take-it-or-leave-it offers to retailers, we have given them all bargaining power in wholesale contract negotiations. However, as stated above, nowadays retailers detain most of

the bargaining power in vertical negotiations. Allowing retailers to hold some bargaining power as well would not change the paper's results though. In such more realistic setting, a nearest manufacturer has to share the "incremental surplus", i. e., the difference between the consumer valuation induced by its good and the farthest manufacturer's, with any store it supplies. For example, when there is not any retail chain and manufacturer a supplies store 1, the "incremental surplus" is $v_1\left(x_{1a}^*\right) - v_1\left(x_{1b}^*\right)$. If retail store 1 holds some bargaining power and is thus able to get a fraction $0 < \beta < 1$ of such surplus, manufacturer a extracts from its supplying relation with store 1 a profit of $(1-\beta)\left(v_1\left(x_{1a}^*\right) - v_1\left(x_{1b}^*\right)\right)$, which is lower than when it detains all bargaining power. Nevertheless, cooperative product location choice remains clearly unchanged at any stage of the paper. Moreover, as before, one can easily show that stores 1 and 4 merge.

Hence, all results of the paper are robust to variations on the distribution of bargaining power between manufacturers and retailers.

6 Appendix

6.1 Appendix A - Proof of Lemma 1

When competing for supplying retail store i, it is clear that each manufacturer contemplates in its wholesale contract offers, from the varieties it is producing, the one that induces the highest consumer valuation at the correspondent local market. Formally, when offering a wholesale contract to store i, each manufacturer j sets

$$x_{ij}^* = \arg\max_{x \in X_j} \; v_i\left(x\right).$$

Next, when competing for the wholesale tariff, the manufacturer j for which $v_i\left(x_{ij}^*\right)$ is lower goes until $T_{ij} = 0$. Thus, the manufacturer j for which $v_i\left(x_{ij}^*\right)$ is higher, in order to indeed supply the stores and maximize its profits, sets the highest wholesale tariff T_{ij} satisfying $T_{ij} \leq v_i\left(x_{ij}^*\right) - v_i\left(x_{i(-j)}^*\right)$,[12] which is $T_{ij}^* = v_i\left(x_{ij}^*\right) - v_i\left(x_{i(-j)}^*\right)$. If there is no manufacturer for which $v_i\left(x_{ij}^*\right)$ is higher (or lower), i. e., if both manufacturers induce the same $v_i\left(x_{ij}^*\right)$, then they both set $T_{ij}^* = 0$.

Hence, when offering a wholesale contract to store i, each manufacturer j sets the wholesale tariff according to

[12] We use here the assumption that, in case of indifference, a retailer chooses the wholesale contract inducing the highest consumer valuation.

$$T_{ij}^{*} = \max\left\{v_i\left(x_{ij}^{*}\right) - v_i\left(x_{i(-j)}^{*}\right), 0\right\}.$$

6.2 Appendix B - Proof of Proposition 1

It is straightforward to see that if one of the four goods is not produced, there is at least one manufacturer which has incentives to produce it, either by replacing a former good (product relocation) or by producing a new one (product innovation). Clearly, such deviating behaviour allows the deviator either to supply more stores or to increase the profits it makes with the stores it already supplies. Likewise, it is clear that when all four goods are produced, any possible individual deviation in location results either in a lower number of stores supplied (it implies a profit loss with respect to these stores while not increasing the profits to be made with the stores it still supplies) or in lower profits made with the supplied stores. Consequently, in any NE, all four goods are produced. In addition, any product location pattern featuring four varieties is a NE. We have thus three types of NE.

The first type of NE is characterized by $X_j = \{x_1, x_2\}$ and $X_{-j} = \{x_3, x_4\}$.

In such product location pattern, manufacturer j supplies stores 1 and 2 by setting wholesale contracts $(T_1, x_{1j}) = (2\delta, x_1)$ and $(T_{2j}, x_{2j}) = (\delta, x_2)$, and manufacturer $-j$ supplies stores 3 and 4 by setting wholesale contracts $\left(T_{3(-j)}, x_{3(-j)}\right) = (\delta, x_3)$ and $\left(T_{4(-j)}, x_{4(-j)}\right) = (2\delta, x_4)$, which implies $\pi_j = \pi_{-j} = 3\delta$.

The second type of NE is characterized by $X_j = \{x_1, x_3\}$ and $X_{-j} = \{x_2, x_4\}$.

In such product location pattern, manufacturer j supplies stores 1 and 3 by setting wholesale contracts $(T_{1j}, x_{1j}) = (\delta, x_1)$ and $(T_{3j}, x_{3j}) = (\delta, x_3)$, and manufacturer $-j$ supplies stores 2 and 4 by setting wholesale contracts $\left(T_{2(-j)}, x_{2(-j)}\right) = (\delta, x_2)$ and $\left(T_{4(-j)}, x_{4(-j)}\right) = (\delta, x_4)$, implying $\pi_j = \pi_{-j} = 2\delta$.

Finally, the third type of NE is characterized by $X_j = \{x_1, x_4\}$ and $X_{-j} = \{x_2, x_3\}$.

In this pattern, manufacturer j supplies stores 1 and 4 by setting wholesale contracts $(T_{1j}, x_{1j}) = (\delta, x_1)$ and $(T_{4j}, x_{4j}) = (\delta, x_4)$, and manufacturer $-j$ supplies stores 2 and 3 by setting wholesale contracts $\left(T_{2(-j)}, x_{2(-j)}\right) = (\delta, x_2)$ and $\left(T_{3(-j)}, x_{3(-j)}\right) = (\delta, x_3)$. Again, we have that $\pi_j = \pi_{-j} = 2\delta$.

In either case, total welfare is $4v$, provided that all product variety is maximized and all consumers are consuming their ideal good. Clearly, welfare is maximized.

6.3 Appendix C - Proof of Proposition 2

Without loss of generality, we identify the following thirteen candidate product location patterns, which we identify with the letters A to M.

In pattern A, $X_a = \{x_1\}$ and $X_b = \{x_4\}$. In pattern B, $X_a = \{x_1\}$ and $X_b = \{x_3\}$. In pattern C, $X_a = \{x_1\}$ and $X_b = \{x_2\}$. In pattern D, $X_a = \{x_2\}$ and $X_b = \{x_3\}$. These are the location patterns featuring minimum product variety, where each manufacturer produces only one good.

In pattern E, $X_a = \{x_1, x_2\}$ and $X_b = \{x_3\}$. In pattern F, $X_a = \{x_1, x_2\}$ and $X_b = \{x_4\}$. In pattern G, $X_a = \{x_1, x_3\}$ and $X_b = \{x_2\}$. In pattern H, $X_a = \{x_1, x_3\}$ and $X_b = \{x_4\}$. In pattern I, $X_a = \{x_1, x_4\}$ and $X_b = \{x_2\}$. In pattern J, $X_a = \{x_1, x_4\}$ and $X_b = \{x_3\}$. These are the asymmetric location patterns where there is one manufacturer producing two goods and another producing only one good.

In pattern K, $X_a = \{x_1, x_2\}$ and $X_b = \{x_3, x_4\}$. In pattern L, $X_a = \{x_1, x_3\}$ and $X_b = \{x_2, x_4\}$. Finally, in pattern M, $X_a = \{x_1, x_4\}$ and $X_b = \{x_2, x_3\}$. These are the symmetric location patterns featuring the competitive level of product variety.

We now use the single stage equilibria found for the third, fourth and fifth stages of the game in order to derive the optimal product location choice for manufacturers.

In location pattern A, a supplies stores 1 and 2 by $T_1 = 3\delta$ and $T_2 = \delta$ and b supplies stores 3 and 4 by $T_3 = \delta$ and $T_4 = 3\delta$. Each manufacturer thus gets a profit of $\pi_j = 4\delta$.

In location pattern B, a supplies stores 1 and 2 by $T_1 = 2\delta$ and $T_2 = 0$ and b supplies stores 3 and 4 by $T_3 = T_4 = 2\delta$. Manufacturer a gets a profit of $\pi_a = 2\delta$ and manufacturer b gets a profit of $\pi_b = 4\delta$.

In location pattern C, a supplies store 1 by $T_1 = \delta$ and b supplies stores 2, 3 and 4 by $T_2 = T_3 = T_4 = \delta$. Manufacturer a gets a profit of $\pi_a = \delta$ and manufacturer b gets a profit of $\pi_b = 3\delta$.

In location pattern D, a supplies stores 1 and 2 by $T_1 = T_2 = \delta$ and b supplies stores 3 and 4 by $T_3 = T_4 = \delta$. Each manufacturer j gets a profit of $\pi_j = 2\delta$.

In location pattern E, a supplies stores 1 and 2 by $T_1 = 2\delta$ and $T_2 = \delta$ and b supplies stores 3 and 4 by $T_3 = T_4 = \delta$. Manufacturer a gets a profit of $\pi_a = 3\delta$ and manufacturer b gets a profit of $\pi_b = 2\delta$.

In location pattern F, a supplies stores 1, 2 and 3 by $T_1 = 3\delta$, $T_2 = 2\delta$ and $T_3 = 0$ and b supplies store 4 by $T_4 = 2\delta$. Manufacturer a gets a profit of $\pi_a = 5\delta$ and manufacturer b gets a profit of $\pi_b = 2\delta$.

In location pattern G, a supplies stores 1, 3 and 4 by $T_1 = T_3 = T_4 = \delta$ and b supplies store 2 by $T_2 = \delta$. Manufacturer a gets a profit of $\pi_a = 3\delta$ and manufacturer b gets a profit of $\pi_b = \delta$.

In location pattern H, a supplies stores 1, 2 and 3 by $T_1 = 3\delta$, $T_2 = \delta$ and $T_3 = \delta$ and b supplies store 4 by $T_4 = \delta$. Manufacturer a gets a profit of $\pi_a = 5\delta$ and manufacturer b gets a profit of $\pi_b = \delta$.

In location pattern I, a supplies stores 1, 3 and 4 by $T_1 = \delta$, $T_4 = 2\delta$ and $T_3 = 0$ and b supplies store 2 by $T_2 = \delta$. Manufacturer a gets a profit of $\pi_a = 3\delta$ and manufacturer b gets a profit of $\pi_b = \delta$.

In location pattern J, a supplies stores 1, 2 and 4 by $T_1 = 2\delta$, $T_4 = \delta$ and $T_2 = 0$ and b supplies store 3 by $T_3 = \delta$. Manufacturer a gets a profit of $\pi_a = 3\delta$ and manufacturer b gets a profit of $\pi_b = \delta$.

In location pattern K, a supplies stores 1 and 2 by $T_1 = 2\delta$ and $T_2 = \delta$ and b supplies stores 3 and 4 by $T_3 = \delta$ and $T_4 = 2\delta$. Each manufacturer gets a profit of $\pi_j = 3\delta$.

In location pattern L, a supplies stores 1 and 3 by $T_1 = T_3 = \delta$ and b supplies stores 2 and 4 by $T_2 = T_4 = \delta$. Each manufacturer gets a profit of $\pi_j = 2\delta$.

In location pattern M, a supplies stores 1 and 4 by $T_1 = T_4 = \delta$ and b supplies stores 2 and 3 by $T_2 = T_3 = \delta$. Each manufacturer gets a profit of $\pi_j = 2\delta$.

It is straightforward to find that manufacturers agree on location pattern A. Indeed, it maximizes joint-profits and it is symmetric. For example, location patterns B to D would never be accepted by manufacturer a while location patterns E to J would never be accepted by manufacturer b.

6.4 Appendix D - Proof of Proposition 4

We define the product location patterns as in appendix C and make use of proposition 3.

If there is no pair of merged stores, manufacturers agree on location pattern A. Store 1 is supplied by manufacturer a by $T_1 = 3\delta$ and gets $\pi_1 = v - 3\delta$, store 4 is supplied by manufacturer b by $T_4 = 3\delta$ and gets $\pi_4 = v - 3\delta$, store 2 is supplied by manufacturer a by $T_2 = \delta$ and gets $\pi_2 = v - 2\delta$, and store 3 is supplied by manufacturer b by $T_3 = \delta$ and gets $\pi_3 = v - 2\delta$.

If stores 1 and 4 are merged and 2 and 3 don't, manufacturers agree on location pattern M. Retail chain $1 + 4$ is supplied by manufacturer a by $T_{1+4} = 2\delta$ and gets $\pi_{1+4} = 2v - 2\delta$ and stores 2 and 3 are supplied by manufacturer b by $T_2 = T_3 = \delta$ and get $\pi_2 = \pi_3 = v - \delta$.

If stores 2 and 3 are merged and 1 and 2 don't, manufacturers agree on location pattern A. Store 1 is supplied by manufacturer a by $T_1 = 3\delta$ and gets $\pi_1 = v - 3\delta$, store 4 is supplied by manufacturer b by $T_4 = 3\delta$ and gets $\pi_4 = v - 3\delta$ and retail chain $2 + 3$ is supplied by manufacturer a by $T_{2+3} = 0$ and gets $\pi_{2+3} = 2v - 3\delta$.

If stores 1 and 2 are merged and 3 and 4 don't, manufacturers agree on location pattern A. Store 3 is supplied by manufacturer b by $T_3 = \delta$ and gets $\pi_3 = v - 2\delta$, store 4 is supplied by manufacturer b by $T_4 = 3\delta$ and gets $\pi_4 = v - 3\delta$ and retail chain $1 + 2$ is supplied by manufacturer a by $T_{1+2} = 4\delta$ and gets $\pi_{1+2} = 2v - 5\delta$.

If stores 1 and 3 are merged and 2 and 4 don't, manufacturers agree on location pattern A. Store 4 is supplied by manufacturer b by $T_4 = 3\delta$ and gets $\pi_4 = v - 3\delta$, store 2 is supplied by manufacturer a by $T_2 = \delta$ and gets $\pi_2 = v - 2\delta$ and retail chain $1 + 3$ is supplied by manufacturer a by $T_{1+3} = 2\delta$ and gets $\pi_{1+3} = 2v - 4\delta$.

If both pairs of stores $1 + 4$ and $2 + 3$ are merged, manufacturers agree on location pattern M. Retail chain $1 + 4$ is supplied by manufacturer a by $T_{1+4} = 2\delta$ and gets $\pi_{1+4} = 2v - 2\delta$ and retail chain $2 + 3$ is supplied by manufacturer b by $T_{2+3} = 2\delta$ and gets $\pi_{2+3} = 2v - 2\delta$.

If both pairs of stores $1 + 2$ and $3 + 4$ are merged, manufacturers agree on location pattern A. Retail chain $3 + 4$ is supplied by manufacturer b by $T_{3+4} = 4\delta$ and gets $\pi_{3+4} = 2v - 5\delta$ and retail chain $1 + 2$ is supplied by manufacturer a by $T_{1+2} = 4\delta$ and gets $\pi_{1+2} = 2v - 5\delta$.

If both pairs of stores $1 + 3$ and $2 + 4$ are merged, manufacturers agree on location patterns A. If they set location pattern A, retail chain $2 + 4$ is supplied by manufacturer b by $T_{2+4} = 2\delta$ and gets $\pi_{2+4} = 2v - 4\delta$ and

retail chain $1+3$ is supplied by manufacturer a by $T_{1+3} = 2\delta$ and gets $\pi_{1+3} = 2v - 4\delta$.

Consider retail store 1. If it makes a merger proposal to store 2, it believes it achieves a joint profit after the merger of $2v - 5\delta$, which is not higher than the joint profit before the merger. If it makes a merger proposal to store 3, it expects a joint profit of $2v - 4\delta$, which is higher than the ex-ante joint profit of $2v - 5\delta$. If it makes a merger proposal to store 4, it expects a joint profit of $2v - 2\delta$, which is higher than the ex-ante joint profit of $2v - 6\delta$. Clearly, store 1 decides to make a merger proposal to store 4. By symmetry, store 4 makes a merger proposal to store 1. A retail chain constituted by stores 1 and 4 is thus formed.

Consider now store 2. If it makes a merger proposal to store 1, it believes it achieves a joint profit after the merger of $2v - 5\delta$, which is not higher than the joint profit before the merger. If it makes a merger proposal to store 4, it expects a joint profit of $2v - 4\delta$, which is higher than the ex-ante joint profit of $2v - 5\delta$. If it makes a merger proposal to store 3, it expects a joint profit of $2v - 3\delta + \mu\delta$, which is higher than the ex-ante joint profit of $2v - 4\delta$. Clearly, store 2 decides to make a merger proposal to store 3. By symmetry, store 3 makes a merger proposal to store 2. A retail chain constituted by stores 2 and 3 is thus formed.

References

[1] Adams, W. A. (1987), "Countervailing Power", in *The New Palgrave: A Dictionary of Economics*, edited by J. Eatwell, M. Milgate, and P. Newman, 4 volumes, Macmillan, London.

[2] Allain, M. - L. and Waelbroeck, P. (2006), "Music Variety and Retail Concentration", Working Paper, Telecom Paris.

[3] Berry, S. and Waldfogel, J. (2001), "Do Mergers Increase Product Variety? Evidence from Radio Broadcasting", *Quarterly Journal of Economics* 116, pp. 1009-25.

[4] Chae, S. and Heidhues, P. (2004), "Buyers' alliances for bargaining power", *Journal of Economics and Management Strategy*, Vol. 13, pp. 731-754.

[5] Chen, Z. (2004), "Monopoly and Product Diversity: The Role of Retailer Countervailing Power", Working Paper, Carleton University.

[6] Dobson, P. (2002), "Retailer buyer power in European markets: lessons from grocery supply", Loughborough University Business School, Working Paper No 2002-1.

[7] Dobson, P., Clarke, R., Davis, S., and Waterson, S. (1999), "Buyer Power and Its Impact on Competition in the Food Retail Distribution Sector of the European Union", Working Paper, Loughborough University.

[8] Dobson, P. and Waterson, M. (1999), "Retailer Power: Recent Developments and Policy Implications", *Economic Policy* 28, pp. 133-164.

[9] Dobson, P., Waterson, M. and Chu, A. (1998), "The Welfare Consequences of the Exercise of Buyer Power", *Office of Fair Trading*, Research Paper No 16, July.

[10] Dobson, P. W. and Waterson, M. (1997), "Countervailing Power and Consumer Prices", *Economic Journal*, Vol. 107, pp. 418-430.

[11] Galbraith, J. K. (1952), "American Capitalism: The Concept of Countervailing Power", Houghton Mifflin: New York.

[12] Götz, G. and Gugler, K. (2003), "Mergers and Product Variety under Spatial Competition: Evidence from Retail Gasoline", Manuscript, Department of Economics, University of Vienna.

[13] Horn, H. and Wolinsky, A. (1988), "Bilateral Monopolies and Incentives for Merger", *Rand Journal of Economics*, Vol. 19, pp. 408-419.

[14] Hunter, A. (1958), "Notes on Countervailing Power", *Economic Journal*, Vol. 68, pp. 89-103.

[15] Inderst, R. and Shaffer, G. (2007), "Retail Mergers, Buyer Power and Product Variety", *The Economic Journal* 117 (January), pp. 45-67.

[16] Inderst, R. and Wey, C. (2007), "Buyer power and supplier incentives", *European Economic Review*, Vol. 51, pp. 647-667.

[17] Inderst, R. and Wey, C. (2003), "Bargaining, mergers and technology choice in bilaterally oligopolistic industries", *RAND Journal of Economics*, Vol. 34, No 1, Spring, pp. 1-19.

[18] Inderst, R. and Wey, C. (2000), "Cross-Country Mergers and Countervailing Power", mimeo.

[19] Jensen et al. (2005), "Aggregation of retail stores", Physica A 351, pp. 551-570

[20] Lommerud, K. E. and Sørgard, L. (1997), "Merger and Product Range Rivalry", *International Journal of Industrial Organization* 16, pp. 21-42.

[21] Mazzarotto, N. (2003), "Retail mergers and buyer power", mimeo.

[22] Nocke, V. and White, L. (2005), "Do Vertical Mergers Facilitate Upstream Collusion?", PIER Working Paper.

[23] Posada, P. and Straume, O. (2004), "Merger, Partial Collusion and Relocation", *Journal of Economics*, Vol. 83, No 3, pp. 243-265.

[24] Rey, P. and Tirole, J. (2006), "A Primer on Foreclosure", *forthcoming in Handbook of Industrial Organization III.*

[25] Schwartz, M. (1999), "Buyer Power Concerns and the Aetna-Prudential Merger", U.S. Department of Justice, Washington D.C..

[26] Snyder, C. M. (1996), "A Dynamic Theory of Countervailing Power", *Rand Journal of Economics* 27, pp. 747-769.

[27] Snyder, C. M. (2005), "Countervailing Power", manuscript prepared for the New Palgrave Dictionary.

[28] Steptoe, M. L. (1993), "The Power-Buyer Defense in Merger Cases", *Antitrust Law Journal* 61, pp. 493-504.

[29] Vogel, L. (1998), "Competition Law and Buying Power: The Case for a New Approach in Europe", *European Competition Law Review* 1, pp. 4-11.

[30] von Ungern-Sternberg, T. (1996), "Countervailing Power Revisited", *International Journal of Industrial Organization* 14, pp. 507-520.

Private Labels, Retail Competition and Upstream Counterstrategies*

Pedro Carreira†
Universidade Nova de Lisboa
Instituto Politécnico de Leiria

March 2009

Abstract

This paper aims to study the impact of the development of private labels (PL) by some retailers on the selling strategies of national brand (NB) manufacturers. When NB manufacturers have alternative retailers for serving consumers, foreclosure may constitute an optimal counterstrategy. Moreover, by providing NB manufacturers a better control over retail competition, supplying the PL may be of strategic value. We indeed find that, after the development of a PL program by a retailer, a NB manufacturer forecloses its access to the NB when there is a high portion of non-loyal consumers and the NB is not competitive enough. Also, we show that the NB manufacturer may supply the PL even if it has a cost disadvantage relatively to a competitive industry producing the PL. Finally, we find that even though the adoption of such counterstrategies is in general socially harmful, the net welfare effect of a retailer's enhanced buyer power through introduction of a PL program is always positive, at least weakly.

Keywords: Private Labels, Retail Competition, Exclusive Dealing, Vertical Foreclosure.

Thematic Area: 1 (Industrial Economics)

*This paper forms the second chapter of my Economics PhD dissertation at Universidade Nova de Lisboa (Faculdade de Economia). I gratefully acknowledge financial support by Fundação para a Ciência e a Tecnologia. I thank my PhD supervisor Steffen Hoernig for his guidance and helpful comments.

†Equivalent to assistant professor, Departamento de Gestão e Economia, Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Leiria. Adress: Campus 2, Morro do Lena - Alto do Vieiro, Apartado 4163, 2411-901, Leiria, Portugal (mailbox: D172). Phone: +351 244 820300 (ext: 3629). Fax : +351 244 820310. Email: pedro.carreira@estg.ipleiria.pt.

Private Labels, Retail Competition and Upstream Counterstrategies

Pedro Manuel Rodrigues Carreira
pedro.carreira@estg.ipleiria.pt

Equiparado a Professor Assistente
Instituto Politécnico de Leiria

Aluno de Doutoramento em Economia
Universidade Nova de Lisboa

Resumo

Este artigo pretende estudar o impacto do desenvolvimento de marcas próprias por parte das empresas retalhistas nas estratégias de venda dos produtores das marcas nacionais. Quando os produtores dispõem de vários retalhistas para servir os mesmos consumidores, pode ser óptimo não vender a marca nacional a alguns deles. Para além disso, na medida em que pode proporcionar um melhor controlo sobre a concorrência no retalho, abastecer o programa de marca própria dos retalhistas pode ter valor estratégico. De facto, mostramos que um produtor monopolista de uma marca nacional deixa de vender a marca nacional a um retalhista que desenvolva um programa de marca própria se existe uma percentagem de consumidores desleais à marca nacional suficientemente grande e se a marca nacional não for suficientemente competitiva. Mostramos também que o produtor da marca nacional pode ser o fornecedor do programa de marca própria do retalhista mesmo que tenha uma desvantagem em custos em relação a outros possíveis produtores na produção da marca própria. Finalmente, mostramos que, embora estas estratégias sejam em geral nocivas para o bem-estar social, o efeito líquido no bem-estar resultante do aumento do poder de compra do retalhista após a introdução do programa de marca própria nunca é negativo.

1 Introduction

The process of retail concentration that has been occurring over the last decades in most industrialized countries has originated large retail chains, which benefit from extended buyer power in terms of profit appropriation in vertical structures. Indeed, contraposing to a past where upstream manufacturers and wholesalers detained most of the vertical bargaining power, the retail sector possesses now the major part of it. In addition, the constitution of large retail chains facilitates further expansion of retailers' bargaining power. A good example is the introduction and development of private label (PL) programs by most of the large food retailing companies around the highly developed countries.

The market share of PL products has been increasing substantially, both in terms of volume and value. In France for example, the PL market share in the grocery market increased from 18.7% in 1997 to 21.8% in 2000, as we can see in Bergès-Sennou et al. (2004). Likewise, "in the UK from 1975-1997 private label market share grew from 16.4% to 30%", as reported in Steiner (2004). Also, he states that "in Europe overall PL sales rose 28% from 1993-1998".

Naturally, economic literature on PLs has been growing too. The mainstream wave of literature deals with three main research topics, even though they are actually interrelated and jointly present in most papers. These are the reasons that lead retail firms to develop PL programs, its impact on vertical relations and why PL goods are being so successful in terms of market shares, specially for some markets. We present some of the most important papers for each of the research topics.

First, with respect to the factors that lead retail firms to develop PL programs, most of the authors refuted the assertion that the private label's primary attraction was the substantial price discount, relative to national brands, and directed their attention to bargaining power mutations and their impact on NB wholesale prices.

Lal (1990) shows that a strong private label is likely to increase retail pass-through of trade promotions as well as induce manufacturers to offer deeper trade discounts.

In one of the most important papers on private labels so far, Mills (1995) argues that the retailer engages on private label marketing so as to overcome the double-marginalization problem inherent in the distribution of well-known manufacturer brands.

Narasimhan and Wilcox (1998) show that "retailers introduce private labels (...) not only to gain profits directly from the private label but also to use as a strategic weapon to elicit concessions from the national brand manufacturers". Moreover, they find that the ability of the retailer to use the private label for that purpose is inversely related to the risks consumers associate with purchasing the PL.

In addition, Caprice (2000) shows that the retailer's bargaining power also increases even if the NB wholesaler can use two-part tariffs. More precisely, by selling a PL, the retailer improves the value of its outside option to the profit it makes if selling only the PL.

Second, about the impact of development of PL programs by retailers on vertical relations, Mills (1995) finds that: (1) if PL perceived quality relatively to the NB is too low, the retailer may not introduce the PL; (2) when the PL quality is above a certain threshold, the retailer might introduce the PL, but the NB manufacturer has an incentive to lower the NB's wholesale price so as to deter the retailer from selling the PL; (3) above a certain quality threshold, the NB manufacturer cannot deter PL's introduction and accommodates its entry on the market.

Bontems et al. (1999) allow the PL and the NB to differ in their marginal production costs, and allow these costs to increase with quality and find some different results. In short, when the PL has a cost-disadvantage relatively to the NB, if the PL quality is too low, the NB cannot deter PL's entry. For intermediate values of PL quality, the NB deters PL's entry and, above a certain quality threshold, the PL is not competitive. On the contrary, if the PL has a cost advantage, it is always introduced and the final NB price decreases as a function of PL quality.

In a different approach, Gabrielsen and Sorgard (2007) consider two groups of consumers, one constituted by consumers that are loyal to the NB and another constituted by non-loyal consumers who may buy the PL if priced lower than the NB. Conclusions are substantially different than the ones achieved by Mills (1995) and Bontems et al. (1999). When the number of loyal consumers is low, the PL is not introduced by the retailer given that the wholesaler sets a limit pricing strategy that deters PL's entry. When the number of loyal consumers increases, the limit pricing strategy becomes too costly and the wholesaler focus only on loyal consumers. Finally, when the number of loyal consumers is very high, introduction of the PL has no impact on the NB price.

This controversy shows that the structure of final demand is essential in predicting the impact of PL introduction.

Third, it should not be expected the PL market share to vary equally in all product categories.

Bergès-Sennou et al. (2004) report that "different factors influence the penetration of PLs. Some are related to supply (structure of demand, ease of entry, innovation policy, etc.) and others to the characteristics of demand. For example, goods that require a high level of consumer confidence (baby food, health and beauty products, etc.) generally exhibit low rates of PL penetration".

Furthermore, in a study over 34 categories of products in 106 locations in the U. S., Dhar and Hoch (1997) state that the main factors that favor a large market share for a PL are the high relative quality to the NB, the low variability of quality of PLs, the high product category sales, the high gross margins (in %), a small number of NB manufacturers operating in the product category and the low NB manufacturer advertising expenditures.

Finally, Raju et al. (1995) show that the introduction of PLs is more likely when the product market consists of a large number of national brands, while Morton and Zettelmeyer (2004) find that the introduction of PLs is more likely when the leading NB has a large market share.

There are nevertheless some less frequently adressed issues. Two of them are the questions of who produces the PL and how can the NB manufacturers react to PL development.

With respect to the first of these questions, we observe that sometimes it is a competitive manufacturer industry, constituted by medium and small firms, sometimes it is the own NB manufacturer and sometimes it is the retailer that integrates backwards and supplies its own PL. Quelch and Harding (1996) state that 50% of consumer package good producers also supply PLs. But when does a NB manufacturer have incentives to supply also a PL? Bell et al. (1997) point out that a NB manufacturer can supply the PL in order to use excess production capacity. Moreover, there are no implied marketing costs. In addition, Bergès-Sennou (2006) shows that, when the retailer's bargaining power is too low, he will prefer to entrust the production of his PL to an independent firm relatively to entrust it to the NB manufacturer. Also, he shows that the higher the consumer loyalty to the NB, the higher the probability for the retailer of entrusting the production of the PL to the NB manufacturer.

As for the second question, Mills (1999) suggests many manufacturer counterstrategies to fight the PLs, such as price promotions and increasing product differentiation, which may or not work effectively in order to offset the retailer bargaining power generated by the PL.

These papers work however with bilateral monopoly models, where a NB manufacturer obviously depends on the retail firm that develops a PL program in order to sell the NB product and is indeed very likely to react "friendly" towards such retail firm, by providing it some concessions, discounts or promotions. In a richer context of retail competition, where NB manufacturers may have several alternative distribution channels to serve each consumer, we can expect more "agressive" reactions from the NB manufacturer. We can even expect foreclosure to arise. Further, with respect to the question of who produces the PL, when there is retail competition, the NB manufacturer may have a strategic incentive to supply it, provided that it may constitute a way of regulating not only competition between the NB and the PL but also the competition between retail stores for consumers.

In our paper, we provide more complete answers to the above two questions by working on the richer context of retail competition, which has been generally neglected in the PL literature. We do it by studying optimal counterstrategy behaviour of a monopolist NB manufacturer towards a retailer that develops a PL program. In particular, we focus on the strategic factors that may induce the NB manufacturer to supply a PL and to engage in foreclosure.

We find that the NB manufacturer may adopt many different counterstrategies. If there is high consumer loyalty to the NB, foreclosure is not profitable. Nevertheless, the NB manufacturer preempts the entry of a competitive industry's PL if it can produce either the NB or a PL at a higher surplus, i. e., at a higher difference between consumer valuation and marginal cost. In this case, he supplies a PL if and only if it surplus enhancing relatively to the NB. If there is low consumer loyalty to the NB though, the NB manufacturer is likely to engage in foreclosure, specially if the PL is surplus enhancing. Further, we find that the NB manufacturer may supply a PL for motives other than cost saving.

The paper is organized as follows. In section 2, we present our analytical model and its main assumptions. In section 3, we analyse the benchmark case where no retailer develops a PL program. In section 4, we derive the optimal counterstrategy choice of the NB manufacturer. In section 5, we discuss the main assumptions of the paper and analyze the robustness of the results. In section 6, we provide some concluding remarks. The appendix constitutes section 7.

2 The Model

We consider two vertically related industries. The downstream sector is constituted by two retail stores $i = 1, 2$ (retailers). The stores possess some local market power but ultimately compete for the same consumers. Formally, we assume that the stores are located at the extreme points of a Hotelling linear city with total length of one. Without loss of generality, we assume that store 1 is the one located at the left end of the city and the one that develops a PL program. In concrete, it is able to obtain a PL product from a competitive supply industry at a wholesale marginal price of c.

Consumers are divided in two groups with respect to their loyalty to the NB. The first group, which we call "non-loyal", is constituted by the α consumers located the nearest to the left end of the city, i. e., the α ones whose preference for shopping at store 1 is stronger. We assume that these consumers attach to the PL of retailer 1 the same value that they do to the NB.[1] Clearly, if they are to make their shopping at store 1, they buy the PL if the PL is priced lower than the NB. If priced equally, we assume that they buy the PL. The second group, which we call "loyal", is constituted by the remaining $1 - \alpha$ consumers located the nearest to the right end of the city, i. e., the $1 - \alpha$ ones with a stronger preference towards store 2. We assume they attach no value to retailer 1's PL. All consumers attach a value of v to the NB, which we assume to be sufficiently high relatively to the Hotelling parameter t and marginal production costs.[2]

The upstream sector is constituted by a monopolist NB manufacturer who is able to produce also a PL. Its marginal costs for the NB and the PL are c_b and c_l, respectively. Alongside with the NB manufacturer, there is the competitive industry that may supply a PL to retailer 1.

We construct a two-stage game. In the first stage, wholesale contracts are agreed. We assume that the NB manufacturer is able to make publicly observed non-linear wholesale contract offers to the retail firms. This is specially reasonable for large and powerful monopolist NB manufacturers.[3] We

[1] Alternatively, we could obviously assume a lower valuation for the PL, i. e., some $_l$. However, we would loose simplicity while obtaining no new results.

[2] We assume a consumer located at incurs in a linear transport cost of if shopping at store 1 and of $(1 -$) if shopping at store 2. The assumption on a high enough is usual in related literature and it is made here for simplicity. More precisely, it assures that, when negotiating only with retailer 2, the NB manufacturer induces it to serve only the loyal group of consumers for any value of .

[3] Coca-cola may be a good example. In section 5 however, we discuss the case of a less powerful NB manufacturer which may not able to charge non-linear wholesale tariffs.

assume further that, when retailer 1 rejects its wholesale contract offer, the NB manufacturer is able to renegotiate the wholesale contract with retailer 2. This is meant to fix retailer 1's outside option. It is derived from the renegotiated wholesale contract between the NB manufacturer and retailer 2. Naturally, for any retailer, the NB manufacturer may also decide to offer no wholesale contract at all. Finally, we assume that, in case of indifference, retailers accept their offered wholesale contract.

In the second stage, given the wholesale contracts agreed in the first stage, retailers engage in price competition. All firms realize their profits.

3 Benchmark: Equilibrium Without Private Labels

Before entering the main section of the paper, consider first the benchmark case where no retailer develops a PL program.

Let p_1 and p_2 designate the NB retail prices charged by retailers 1 and 2 respectively, q_1 and q_2 the NB quantities sold by retailers 1 and 2 respectively, and π_1, π_2 and π_m the profits of retailer 1, retailer 2 and NB manufacturer, respectively. Finally, let (T_1, q_1) and (T_2, q_2) represent the non-linear wholesale contract offers made by the NB manufacturer to retailers 1 and 2, respectively, where T_i is the wholesale tariff charged to retailer i.

Given that the NB manufacturer is able to set non-linear wholesale contract offers and retailers have zero valued outside options, it is able to extract all industry surplus. Consequently, the NB manufacturer maximizes it, which is equivalent here to minimizing consumer transport costs and satisfying all demand, since v is high enough. Clearly, this is achieved if and only if each store serves half of the city. But then, if accepting their offered wholesale contracts, retailers set $p_i = v - \frac{t}{2}$, thus making a profit of $\pi_i = \frac{1}{2}\left(v - \frac{t}{2}\right) - T_i$.

Hence, the NB manufacturer sets

$$(T_1, q_1) = (T_2, q_2) = \left(\tfrac{1}{2}\left(v - \tfrac{t}{2}\right), \tfrac{1}{2}\right),$$

in order both retailers to indeed accept their wholesale contract offers, and it gets a profit of $\pi_m = v - \frac{t}{2} - c_b$. We construct the following proposition.

Proposition 1 *If there are no private labels, the NB manufacturer forecloses no retailer.*

The result in the proposition is obvious. If there are no private labels, the NB manufacturer is an unrestricted monopolist for both retail stores. Moreover, as each store detains some local market power, i. e., as some of the NB consumers achieve a higher surplus if shopping at retailer 1's while others achieve a higher surplus if shopping at retailer 2's, the NB manufacturer can charge higher global wholesale tariffs and make more profits when it supplies the NB product to both stores.

4 Private Labels and Counterstrategy Choice

Suppose now retailer 1 develops a PL program. In this section, we derive the optimal counterstrategy choice for the NB manufacturer, depending on consumer loyalty to the NB and on production costs.

We identify five equilibrium counterstrategies: *accommodating the PL from the competitive industry*, *tie-in*, *deter PL's entry*, *foreclosure* and *partial foreclosure*.

4.1 Accommodating Equilibria

In this subsection, we derive the conditions under which *accommodating the PL from the competitive industry* is the optimal counterstrategy for the NB manufacturer. We mean by this counterstrategy that the NB manufacturer keeps supplying the NB to retailer 1 while accommodating the PL from the competitive industry.

Proposition 2 *The NB manufacturer accommodates the PL from the competitive industry if the majority of consumers is loyal to the NB and* $c < \min\{c_b, c_l\}$.

Proof. See the Appendix. ■

The intuition is straightforward. When the majority of consumers is loyal to the NB ($\alpha < \frac{1}{2}$), some of the loyal consumers obtain a higher surplus (incur in a lower transport cost) if shopping at retailer 1's store. Therefore, the NB manufacturer can charge potentially higher wholesale tariffs for the NB product if he supplies it to both retail stores. In addition, if $c < \min\{c_b, c_l\}$, the PL from the competitive industry is more competitive than any of NB

manufacturers' products serving non-loyal consumers.[4] Thus, the NB manufacturer finds it unprofitable to deter the entry of the PL from the competitive industry and supplies the NB product only to be consumed by loyal consumers.

Retailer 1's enhanced buyer power is welfare improving, given that total production costs decrease. Moreover, welfare is maximized here, since both production and transport costs are minimized. There is a perfect match between the social and the NB manufacturer's private interests when the PL program of retailer 1 does not constitute a strong enough threat to NB market share, i. e., when the majority of consumers is loyal to the NB.

4.2 Tie-In Equilibria

Following last subsection, it is easy to figure out why the NB manufacturer may find it optimal to tie-in a PL to the NB when setting the wholesale contract offer to retailer 1.

Proposition 3 *The NB manufacturer ties-in the PL to the NB when setting retailer 1's wholesale contract offer if the majority of consumers is loyal to the NB and* $c_l < \min\{c, c_b\}$.

Proof. See the Appendix. ■

Similarly to above, when the majority of consumers is loyal, there are some loyal consumers who incur in a lower transport cost if shopping at retailer 1's store, which means that the NB manufacturer is able to charge higher wholesale tariffs if supplying the NB product to both stores. Furthermore, if $c_l < \min\{c, c_b\}$, it can "undercut" the wholesale price of the PL from the competitive industry. In other words, by supplying the PL the NB manufacturer increases industry surplus, which it can extract, through non-linear wholesale tariffs.

Welfare increases after the PL program introduction, given that there is no increase in transport costs and production costs are diminished. Furthermore, like above, it is clear to see that the NB manufacturer induces a welfare

[4] In a fixed demand model, a product is more competitive than other if it generates a higher surplus, i. e., a higher difference between consumer valuation and production cost. However, in our model, as we assume identical valuations for the NB and all PL, a product generates more surplus than other if and only if its production cost is lower.

maximizing outcome, provided that transport costs and production costs are minimized.

4.3 Deterrence Equilibria

The NB manufacturer may also find it optimal to *deter PL's entry*, i. e., to offer an NB exclusive dealing wholesale contract to retailer 1. It should be straightforward that this is optimal for the NB manufacturer only if the NB has a sufficiently low marginal production cost of.

Proposition 4 *The NB manufacturer deters PL's entry if* $c_b < \min\{c, c_l\}$ *or, when the majority of consumers is non-loyal and* $c < c_b < c_l$*, if the NB is competitive enough.*

Proof. See the Appendix. ■

The first part of the proposition is very intuitive. If the NB is the most competitive product, the NB manufacturer can profitably "undercut" the PL from the competitive industry using precisely the NB. The NB manufacturer then maximizes potential wholesale tariffs by maximizing consumer surplus (minimizing transport costs), which is achieved by inducing each retailer to serve half of the city. In this case, total welfare is again maximized, provided that both transport costs and production costs are minimized.

As for the second part of the proposition, note first that when $\alpha > \frac{1}{2}$, accommodating the PL from the competitive industry implies a reduction in potential wholesale tariffs. Indeed, given that the NB manufacturer sees it optimal to serve only loyal consumers, there would be consumers shopping at their least preferred store and consumer surplus would not be maximized. On the contrary, by deterring PL's entry, the NB manufacturer assures any consumer shops at its preferred store and it maximizes potential wholesale tariffs. This has however a cost for the NB manufacturer, given that it has to "undercut" a more efficient product with respect to serving non-loyal consumers. In the end, the NB manufacturer deters PL's entry if it is not a too costly counterstrategy, i. e., if the marginal cost difference between the NB and the PL from the competitive industry is sufficiently low. In detail, if $c_b - c < \frac{t}{2\alpha} - 2t(1-\alpha)$. With respect to welfare, looking at subsection 5 of the appendix, it should be clear that this counterstrategy is never welfare

maximizing in this case. More precisely, the socially optimal tradeoff between minimization of transport costs and of production costs would be achieved in this case when the indifferent consumer is located at $x = \frac{1}{4t}(2t + c_b - c)$ and not at $x = \frac{1}{2}$. Moreover, note that the NB manufacturer may end up deterring the entry of the PL from the competitive industry even when its entrance is socially desirable.

Yet, in either case, the enhanced buyer power of retailer 1 is clearly neither welfare improving nor welfare reducing here.

4.4 Partial Foreclosure Equilibria

By *partial foreclosure*, we mean the counterstrategy under which the NB manufacturer forecloses retailer 1's access to the NB but supplies it a PL under exclusive dealing. In this subsection, we derive the conditions under which such counterstrategy is optimal for the NB manufacturer.

Proposition 5 *The NB manufacturer partially forecloses retailer 1 if the majority of consumers is non-loyal and* $c_l < \min\{c, c_b\}$ *or, when* $c < c_l < c_b$*, if the NB is competitive enough.*

Proof. See the Appendix. ■

As we have already seen, when the majority of consumers is non-loyal and all loyal consumers achieve a higher surplus if shopping at retailer 2's store, the NB manufacturer is tempted to sell the NB only to retailer 2, specially if deterring the entry of the PL is too costly. Furthermore, it should be clear that the NB manufacturer can profitably supply the PL if it is the most competitive product. In this case, by looking at subsection 5 of the appendix, we are sure the NB manufacturer chooses the socially optimal solution, i. e., it sets the quantities that induce the indifferent consumer to be located at $x = \frac{1}{4t}(2t + c_b - c_l)$, which describes the best social tradeoff between the minimization of transport and production costs. Clearly, welfare increases after the PL program introduction by retailer 1.

More important, note that, when the majority of consumers is non-loyal, the NB manufacturer has an incentive to supply the PL even if it has a cost disadvantage in its production relatively to the competitive industry, if the NB is competitive enough. In detail, if $c_b < c_l + 2t(2\alpha - 1)$ and $c_b <$

$c+\frac{1}{8t\alpha}\left(2t+c_b-c_l\right)^2-2t\left(1-\alpha\right)$. This is because supplying the PL gives the NB manufacturer new options so as to balance more evenly retail competition and production costs. Clearly, if not supplying the PL, the NB manufacturer can only choose between two "extreme" options: it either maximizes potential wholesale tariffs (by deterring PL's entry) or it minimizes production costs (by serving only loyal consumers).

Now, the NB manufacturer indeed supplies the PL to retailer 1 under exclusive dealing if two conditions are verified. First, c_l cannot be much greater than c in order for such counterstrategy not to be too costly. Second, c_b cannot be much greater than c_l. If it is, the NB manufacturer induces equilibrium quantities in which the indifferent consumer is located at $x=\alpha$. But then supplying the PL would not increase consumer surplus and potential wholesale tariffs. It would only cause an increase in production costs. The NB manufacturer would be better off if serving only loyal consumers.

Welfare is clearly never maximized in this case. Indeed, production costs would only be minimized if the PL from the competitive industry enters. Yet, again, welfare increases after the introduction of a PL program by retailer 1. The saving effect on production costs is higher than the loss effect on transport costs. This may not be clear to the reader at first sight. However, note that the NB manufacturer is able to extract all industry surplus net from the retailers' outside options. Hence, if it *partially forecloses* rather than *deters PL's entry*, industry surplus is higher for sure.

4.5 Foreclosure Equilibria

We now derive the equilibria in which the NB manufacturer "totally" forecloses retailer 1, i. e., it does not sell it neither the NB nor the PL.

Proposition 6 *The NB manufacturer forecloses retailer 1's access to the NB if the majority of consumers is non-loyal and the NB is not sufficiently competitive.*

Proof. See the Appendix. ■

In detail, by looking at subsection 6 of the appendix, foreclosure equilibria arises if $\alpha>\frac{1}{2}$ and either $c<c_b<c_l$ and $c_b-c>\frac{t}{2\alpha}-2t\left(1-\alpha\right)$ or $c<c_l<c_b$ and $c_b>\min\left\{c_l+2t\left(2\alpha-1\right),c+\frac{1}{8t\alpha}\left(2t+c_b-c_l\right)^2-2t\left(1-\alpha\right)\right\}$. In short, these conditions require a high enough NB marginal production cost so that

the NB manufacturer does not *deters PL's entry* and a high enough PL marginal production cost so that *partial foreclosure* is too costly.

The first case is easier to explain. If the PL is a damaged good for the NB manufacturer, it never finds it optimal to supply it. Moreover, as we have seen above, when the majority of consumers is non-loyal, the NB manufacturer may have an incentive to deter the entry of the PL if the gain in potential wholesale tariffs is higher than the increased production costs. Hence, it forecloses retailer 1 if the marginal production cost disadvantage of the NB is sufficiently high, i. e., if deterring PL's entry is too costly. The NB manufacturer ends up serving only loyal consumers, through retailer 2's store, given that all loyal consumers achieve a higher surplus when shopping there.

Looking at subsections 5 and 6 of the appendix, we find that, in this case, foreclosure is welfare maximizing if and only if $c_b - c \geq 2t\,(2\alpha - 1)$. If $\frac{t}{2\alpha} - 2t\,(1-\alpha) < c_b - c < 2t\,(2\alpha - 1)$, the socially optimal indifferent consumer would be located at $x = \frac{1}{4t}\,(2t + c_b - c)$ and not at $x = \alpha$. More precisely, *foreclosure* may imply a too low consumed quantity of the NB, from a social point of view.

The NB manufacturer may also find it optimal to foreclose retailer 1 when $\alpha > \frac{1}{2}$ and it can produce a PL at a lower cost than the NB. As we have seen in the last subsection, if $\alpha > \frac{1}{2}$ and c_l is too great relatively to c, partial foreclosure is too costly, which clearly implies that deterring PL's entry is also a too costly counterstrategy. Moreover, provided that the majority of consumers is non-loyal, all consumers achieve a higher surplus if shopping at retailer 2's store. Hence, the NB manufacturer forecloses retailer 1 so that it can charge a higher wholesale tariff. Similarly, if $\alpha > \frac{1}{2}$ and c_b is too great relatively to c_l, the NB manufacturer induces a location for the indifferent consumer of $x = \alpha$ if supplying the PL under exclusive dealing. But, in this case, the unique effect of supplying the PL is the increase in production costs. Hence, the NB manufacturer totally *forecloses* retailer 1.

Welfare analysis is a bit more complex here. However, foreclosure may also not be welfare maximizing here. As before, it is welfare maximizing if and only if $c_b - c \geq 2t\,(2\alpha - 1)$.

Welfare is nevertheless increased after the PL program introduction, in both cases. The savings on production costs more than compensate the losses in transport costs. Again, just note that the NB manufacturer is able to extract all industry surplus net from the retailers' outside options and that, as a consequence, if it *forecloses* rather than *deters PL's entry*, industry surplus is higher.

5 Discussion

5.1 Differences With Linear Wholesale Tariffs

Consider at this section a less powerful NB manufacturer, which may not be able to set up non-linear wholesale contracts. Do the paper's main results remain valid in such setting?

The main difference is that *deterring PL's entry* looses some of its profitability, provided that it promotes homogeneity between retail stores. Indeed, if both stores sell the NB and only the NB, they compete directly for all consumers in the city. In order for retailer 1 to accept such exclusive dealing wholesale contract, the NB manufacturer may have to conceed it a high wholesale price discount, specially if there is a high portion of non-loyal consumers. Consequently, retailer 2 may no longer be competitive, unless it is conceeded a wholesale price discount too. Hence, the NB manufacturer is likely to supply the NB only for loyal consumers, in order to be able to charge high wholesale prices for it, and the PL is probably introduced, even when it is not surplus increasing.

The NB manufacturer is then tempted to foreclose retailer 1's access to the NB if there is a sufficiently large portion of non-loyal consumers. In such case, the consumers who achieve a higher surplus if shopping at retailer 1's store buy the PL, which implies that the NB manufacturer is probably able to charge higher wholesale prices and extract higher profits from the NB if supplying it only to retailer 2's store.

With respect to the question of who supplies the PL, it remains clear that the NB manufacturer supplies it if it can produce it at a lower cost. Moreover, whenever it is optimal not to foreclose retailer 1, the NB manufacturer has a strategic incentive to *tie-in* a PL to the NB when setting its wholesale contract offer. Indeed, the wholesale price of the PL may be used to extract retailer 1's profit from NB sales. In particular, the NB manufacturer may charge for the PL a higher wholesale price than c and still sell it. Naturally, the NB manufacturer *ties-in* the PL only if the maximum wholesale price it can charge in order to indeed sell α units of it is higher than c_l.

The main welfare change is that the PL may be introduced even though it is not surplus enhancing, which is socially harmful. Moreover, even though for different motives, we may still observe the NB manufacturer supplying the PL when $c_l > c$, which further reduces welfare.

5.2 Extensions

As for some possible extensions for the paper, note that we have considered a model with fixed total demand and high consumer valuations. This clearly limits our welfare analysis. It is reasonable to admit that the introduction of PL in a market raises welfare via consumption of low reservation value consumers who do not value the NB enough to pay its retail price. In a model where a PL can raise total equilibrium quantity, one should expect less often the NB manufacturer to *deter PL's entry*, specially with non-linear wholesale tariffs.

Also, we considered only differentiation between retail firms, not between products. It would be adequate to use a product differentiation model to treat the relation between the PL and the NB. Clearly, this would also reduce the profitability of *deterring PL's entry*, since total industry surplus would be higher when both products are sold.

Next, we have assumed that non-loyal consumers attach a value to the PL as high as to the NB. It is however likely that a PL product generates a lower consumer valuation than the NB. With non-linear wholesale tariffs, this clearly increases the incentives for the NB manufacturer to *deter PL's entry*, given that the PL is less likely to be surplus increasing.

Finally, we did not incorporate in our model an effect widely accepted in the literature on private labels, which is the reputation loss of the NB product if its manufacturer also supplies a PL. This clearly reduces NB manufacturer's incentives to supply the PL. If the reputation loss is sufficiently high, the NB manufacturer may decide not to supply the PL even if it is less costly to produce than the NB and it has a cost advantage on its production over the competitive industry.

6 Concluding Remarks

The purpose of our paper was to derive the optimal counterstrategy choice of a NB manufacturer to the development of PL programs by some retailers when there is retail competition, i. e., when the NB manufacturer has alternative retailers to serve each particular consumer. Most literature considers a one-to-one relationship between a NB manufacturer and a retailer. In these settings, it is straightforward that the NB manufacturer does not respond agressively to the development of a PL program, since it has no

valuable outside option. Usually, even if all supermarkets in a particular city detain some local market power, they ultimately compete for all consumers in the city. Foreclosure is thus likely to occur. Furthermore, when there is retail competition, supplying the PL may be of strategic value, since it may provide the NB manufacturer a better control over retail competition, which allows the NB manufacturer to charge higher wholesale prices in general.

The first main research question we answered was when should foreclosure be expected. We found that a NB manufacturer is likely to foreclosure a retailer with a PL program when the majority of consumers is non-loyal to the NB and the NB is not sufficiently competitive. Indeed, when the NB is not competitive enough, the NB manufacturer finds it too costly to deter PL's entry and supplies the NB to be consumed only by loyal consumers. But then, if the majority of consumers is non-loyal, the potential NB consumers are located nearest to the store without a PL program, which means that the NB manufacturer can charge higher wholesale prices for the NB if serving them only through that store.

Next, we answered the paper's second main research question of when should a NB manufacturer be expected to strategically supply the PL, i. e., to supply it even though having a cost disadvantage on its production relative to a competitive supply industry. We found that the NB manufacturer sees it profitable to strategically supply the PL when there is a majority of non-loyal consumers, its PL is surplus increasing relatively to the NB and the NB is competitive enough. Indeed, when $\alpha > \frac{1}{2}$ and $c < c_l < c_b$ but the costs differences are sufficiently low, balancing retail competition (minimizing consumer transport costs) gains importance for the NB manufacturer in order to raise its wholesale tariffs and profits. The NB manufacturer thus uses its PL to set an exclusive dealing wholesale contract contemplating a lower PL quantity than the one which would be bought by the retailer from the competitive industry. It is then able to supply a higher quantity of the NB product, which has a higher wholesale price.

Naturally, the NB manufacturer also finds it optimal to supply the PL if it is less costly to produce than the NB and the PL from the competitive industry, but by cost saving motives.

Finally, we have found that, independently of the NB manufacturer's optimal counterstrategy choice, the enhanced buyer power generated by the introduction of a PL program by a retailer improves welfare. Nevertheless, at the same time, we have find that *deterring PL's entry*, *foreclosure* and *partial foreclosure* may offset such welfare gains, at least partially. More

precisely, *deterring PL's entry* may be adopted when the PL's entry is socially desirable, *foreclosure* may be adopted when it would be preferable to have more consumers buying the NB, and *partial foreclosure* may be adopted when the competitive industry is more efficient than the NB manufacturer producing the PL.

7 Appendix

We calculate the NB manufacturer's profit for each of the counterstrategies in subsections 1-5 and derive the optimal choice in subsection 6.

7.1 Foreclosure

Let q_l and p_l represent the quantity sold and the price charged for the private label by retailer 1.

The NB manufacturer clearly sets $(T_2, q_2) = (p_2 q_2, q_2)$. It extracts all revenues made by retailer 2, provided that it has no valuable outside option. The NB manufacturer then maximizes

$$(p_2 - c_b)\, q_2.$$

When solving such problem, there are two options for the NB manufacturer. It may induce retailer 2 either to compete for the non-loyal consumers with retailer 1 or to behave as a monopolist with respect to the loyal consumers.

If the NB manufacturer induces retailer 2 to behave as a monopolist with respect to the loyal consumers, it solves

$$\begin{aligned} &\max_{p_2} (p_2 - c_b)\, \tfrac{v - p_2}{t}, \\ &\text{s. to } p_2 \geq v - t\,(1 - \alpha). \end{aligned}$$

The solution for the problem is $p_2 = \frac{1}{2}v + \frac{1}{2}c_b$.

However, if v is high enough (in particular, if $v > 2t\,(1-\alpha) + c_b$), the NB manufacturer prefers to induce

$$p_2 = v - t\,(1 - \alpha).$$

It then sets $q_2 = 1 - \alpha$ and $T_2 = (1 - \alpha)\,(v - t\,(1 - \alpha))$. Retailer 2 accepts the contract, sets indeed $p_2 = v - t\,(1 - \alpha)$ and gets zero profits. The NB manufacturer gets a profit of

$$\pi_w = (1-\alpha)(v - t(1-\alpha) - c_b) =$$
$$= v - t(1 - 2\alpha(1-\alpha)) - \alpha c - (1-\alpha)c_b - \alpha(v - c - t\alpha).$$

Retailer 1 then clearly orders from the competitive industry a quantity $q_l = \alpha$, sets $p_l = v - t\alpha$ and gets a profit of $\pi_1 = \alpha(v - c - t\alpha)$.

The NB manufacturer may however induce retailer 2 to serve some of the non-loyal consumers. In this case, the indifferent consumer's location is found by solving

$$v - p_l - tx = v - p_2 - t(1-x),$$

which has a solution given by

$$x = \tfrac{1}{2t}(t + p_2 - p_l).$$

Hence, retailer 1 solves

$$\max_{p_l}(p_l - c)x.$$

FOC is given by $t + c - 2p_l + p_2 = 0$.

Retailer 2 solves

$$\max_{p_2}(p_2 - w_2)(1-x).$$

FOC is given by $t + p_l - 2p_2 + w_2 = 0$.

Solving the system of two equations, one finds

$$p_l = t + \tfrac{2}{3}c + \tfrac{1}{3}w_2 \text{ and}$$
$$p_2 = t + \tfrac{1}{3}c + \tfrac{2}{3}w_2,$$

which implies $q_l = \frac{1}{6t}(3t + w_2 - c)$, $q_2 = \frac{1}{6t}(3t - w_2 + c)$,
$\pi_1 = \frac{1}{18t}(3t + w_2 - c)^2$ and $\pi_2 = \frac{1}{18t}(3t - w_2 + c)^2$.

The NB manufacturer then solves

$$\max_{w_2} w_2 q_2 + \pi_2.$$

The solution is given by $w_2 = \frac{3}{4}t + \frac{1}{4}c$, implying $\pi_w = \frac{1}{16t}(3t + c)^2$.
Hence, the NB manufacturer prefers retailer 2 to serve only loyal demand if

$$v > c_b + t(1-\alpha) + \tfrac{(3t+c)^2}{16t(1-\alpha)}.$$

Resuming, if $v > \max\left\{c_b + t(1-\alpha) + \tfrac{(3t+c)^2}{16t(1-\alpha)}, c_b + 2t(1-\alpha)\right\}$, the NB manufacturer sets $q_2 = 1-\alpha$ and $T_2 = (1-\alpha)(v - t(1-\alpha))$. Retailer 2 accepts the contract and sets $p_2 = v - t(1-\alpha)$. Retailer 1 orders a PL quantity $q_l = \alpha$ from the competitive industry, sets $p_l = v - t\alpha$ and gets a profit of $\pi_1 = \alpha(v - c - t\alpha)$ (this is the value of retailer 1's outside option after the development of the PL program).

The NB manufacturer gets a profit of

$$\pi_w^a = (1-\alpha)(v - t(1-\alpha) - c_b) =$$
$$= v - t(1 - 2\alpha(1-\alpha)) - \alpha c - (1-\alpha)c_b - \alpha(v - c - t\alpha).$$

7.2 Accommodating the PL from the Competitive Industry

When adopting this strategy, note first of all that the NB manufacturer sets $q_1 > 0$ if and only if $\alpha < \frac{1}{2}$. Indeed, if $\alpha \geq \frac{1}{2}$, any loyal consumer attaches a higher value to shopping at retailer 2's store. In such case, as the NB manufacturer can set non-linear wholesale contracts, it maximizes its profits when supplying the NB product only to retailer 2's store, i. e., when adopting "total" *foreclosure*.

Hence, we restrict to the case of $\alpha < \frac{1}{2}$.

Clearly, when accommodating the PL from the competitive industry and supplying only the NB, the NB manufacturer has no concern towards production costs. It thus worries only with minimization of consumer transport costs. Hence, it induces $p_l = p_1 = p_2 = v - \frac{t}{2}$, $q_1 = \frac{1}{2} - \alpha$ and $q_2 = \frac{1}{2}$. With respect to the wholesale tariffs, while it can extract all revenues from retailer 2, the NB manufacturer has to leave retailer 1 with a profit at least as high as the value of its outside option. More precisely, it sets

$$(T_1, q_1) = \left(\tfrac{1}{2}\left(v - \tfrac{t}{2}\right) - \alpha c - \alpha(v - c - t\alpha), \tfrac{1}{2} - \alpha\right) =$$
$$= \left(\tfrac{1}{2}\left(v - \tfrac{t}{2}\right) - \alpha(v - t\alpha), \tfrac{1}{2} - \alpha\right)$$
$$\text{and } (T_2, q_2) = \left(\tfrac{1}{2}\left(v - \tfrac{t}{2}\right), \tfrac{1}{2}\right).$$

Both retailers accept their wholesale contract offers and retail prices are indeed the target ones. The NB manufacturer gets a profit of

$$\pi_w^b = v - \tfrac{t}{2} - \alpha c - (1-\alpha)c_b - \alpha(v - c - t\alpha).$$

7.3 Deter PL Entry

If the NB manufacturer deters the entry of the PL from the competitive industry and supplies exclusively the NB to both retailers, its profits (and industry surplus) are again maximized when all demand is served and transport costs are minimized, i. e., when $p_l = p_1 = p_2 = v - \frac{t}{2}$, provided that v is high enough. The NB manufacturer thus sets

$$(T_1, q_1) = \left(\tfrac{1}{2}\left(v - \tfrac{t}{2}\right) - \alpha\left(v - c - t\alpha\right), \tfrac{1}{2}\right)$$
$$\text{and } (T_2, q_2) = \left(\tfrac{1}{2}\left(v - \tfrac{t}{2}\right), \tfrac{1}{2}\right).$$

All retailers accept their wholesale contract offers and set $p_l = p_1 = p_2 = v - \frac{t}{2}$. The NB manufacturer gets a profit of

$$\pi_w^c = v - \tfrac{t}{2} - c_b - \alpha\left(v - c - t\alpha\right).$$

7.4 Tie-In

Similarly to above, this counterstrategy makes sense if and only if $\alpha < \frac{1}{2}$. If $\alpha \geq \frac{1}{2}$, any loyal consumer attaches a higher value to shopping at retailer 2's store. In such case, as it can set non-linear wholesale contracts, the NB manufacturer would maximize its profits if supplying the NB product only to retailer 2's store, i. e., if adopting *partial foreclosure*.

We thus restrict to the case of $\alpha < \frac{1}{2}$.

Let $q_l^{'}$ be the PL quantity tied-in by the NB manufacturer to retailer 1's wholesale contract offer.

The NB manufacturer maximizes its profits when consumer transport costs are minimized. Thus, it sets

$$(T_1, q_1, q_l^{'}) = \left(\tfrac{1}{2}\left(v - \tfrac{t}{2}\right) - \alpha\left(v - c - t\alpha\right), \tfrac{1}{2} - \alpha, \alpha\right)$$
$$\text{and } (T_2, q_2) = \left(\tfrac{1}{2}\left(v - \tfrac{t}{2}\right), \tfrac{1}{2}\right).$$

All retailers accept their wholesale contract offers and set $p_l = p_1 = p_2 = v - \frac{t}{2}$. The NB manufacturer gets a profit of

$$\pi_w^d = v - \tfrac{t}{2} - \alpha c_l - (1 - \alpha)\, c_b - \alpha\left(v - c - t\alpha\right).$$

7.5 Partial Foreclosure

Contrary to above, this counterstrategy is reasonable if $c_l < c_b$ and $\alpha > \frac{1}{2}$. If $\alpha < \frac{1}{2}$, there are some loyal consumers who achieve a higher surplus if making their shopping at retailer 1's store. Hence, adopting *tie-in* would be strictly preferable. Furthermore, if $c_l > c_b$, the NB manufacturer could clearly do better by *deterring PL's entry.*

We thus restrict to such cases.

Here, the NB manufacturer faces a tradeoff. In one hand, by inducing a $q_l^i < \alpha$ and a $q_2 > 1 - \alpha$, its profit decreases via an increase in production costs. In other hand, by doing it, it increases consumer surplus and potential wholesale tariffs via a decrease in consumer transport costs. The NB manufacturer thus solves

$$\max_x x\left(v - tx\right) + \left(1 - x\right)\left(v - t\left(1 - x\right)\right) - xc_l - \left(1 - x\right)c_b,$$

which has a solution given by $x = \frac{1}{4t}\left(2t - c_l + c_b\right)$.
It is an interior solution ($x \leq \alpha$) if and only if $c_b - c_l \leq 2t\left(2\alpha - 1\right)$.

In this case, it sets

$$\left(T_1, q_1, q_l^i\right) = \left(x\left(v - tx\right) - \alpha\left(v - c - t\alpha\right), 0, x\right)$$

and

$$\left(T_2, q_2\right) = \left(\left(1 - x\right)\left(v - t\left(1 - x\right)\right), 1 - x\right).$$

Substituting, we obtain

$$\left(T_1, q_1, q_l^i\right) = \left(\begin{array}{c} \frac{1}{16t}\left(2t - c_l + c_b\right)\left(4v - 2t + c_l - c_b\right) \\ -\alpha\left(v - c - t\alpha\right), 0, \frac{1}{4t}\left(2t - c_l + c_b\right) \end{array} \right)$$

and

$$\left(T_2, q_2\right) = \left(\begin{array}{c} \frac{1}{16t}\left(2t - c_b + c_l\right)\left(4v - 2t + c_b - c_l\right), \\ \frac{1}{4t}\left(2t + c_l - c_b\right) \end{array} \right).$$

He gets a profit of $\pi_w^e = T_1 + T_2 - c_l q_l^i - c_b q_2 =$

$$\begin{array}{c} = \frac{1}{16t}\left(2t - c_l + c_b\right)\left(4v - 2t + c_l - c_b\right) - \alpha\left(v - c - t\alpha\right) \\ -\frac{1}{16t}\left(2t + c_l - c_b\right)\left(2t - 4v + c_l - c_b\right) \\ -\frac{c_l}{4t}\left(2t - c_l + c_b\right) - \frac{c_b}{4t}\left(2t + c_l - c_b\right). \end{array}$$

If however $c_b - c_l > 2t(2\alpha - 1)$, the NB manufacturer sets

$$(T_1, q_1, q_l^i) = (\alpha(v - t\alpha) - \alpha(v - c - t\alpha), 0, \alpha) = \\ = (\alpha c, 0, \alpha)$$

and

$$(T_2, q_2) = ((1-\alpha)(v - t(1-\alpha)), 1-\alpha)$$

and gets a profit of

$$\pi_w^e = v - t(1 - 2\alpha(1-\alpha)) - \alpha c_l - (1-\alpha)c_b - \alpha(v - c - t\alpha).$$

In either case, both retailers accept their wholesale contract offers, since they obtain a level of profits similar to the value of their respective outside options.

7.6 Counterstrategy Choice

Consider first that there is a majority of loyal consumers, i. e., that $\alpha < \frac{1}{2}$. In this case, foreclosure is never optimal provided that there are loyal consumers who achieve a higher surplus if shopping at retailer 1's store. Indeed, it is straightforward to show that $\pi_w^b > \pi_w^a$, if $\alpha < \frac{1}{2}$.

The NB manufacturer thus decides between *tie-in*, *deterring PL's entry* or *accommodating the PL from the competitive industry*. Clearly, the actual choice depends on costs and not on the portion of non-loyal consumers. The NB manufacturer adopts the counterstrategy that maximizes industry surplus, i. e., that minimizes production costs. Hence, if $c < \min\{c_b, c_l\}$, it *accommodates the entry of the PL from the competitive industry*. If $c_l < \min\{c, c_b\}$, it adopts *tie-in*. Finally, if $c_b < \min\{c, c_l\}$, it *deters PL's entry*.

Suppose now that $\alpha \geq \frac{1}{2}$. Here, only *tie-in* is automatically excluded as a potential equilibrium counterstrategy. Comparing profits, we easily derive the optimal counterstrategy choice for the NB manufacturer.

If $c_b < \min\{c, c_l\}$, the NB manufacturer *deters PL's entry*. Indeed, both transport and production costs are minimized, which means that the NB manufacturer achieves maximum wholesale tariffs at the expense of no cost at all.

If $c_l < \min\{c, c_b\}$, the NB manufacturer adopts *partial foreclosure*. To show that $\pi_w^e > \pi_w^c$, just notice that the NB manufacturer can do better than

π_w^c by maintaining quantities and transport costs minimized but by changing the product sold to retailer 1 from the NB to a PL, since production costs decrease. To show that $\pi_w^e > \pi_w^a$, just notice that the NB manufacturer can do better than π_w^a by maintaining quantities and supplying a PL to retailer 1 by $T_1 = \alpha (c - c_l)$.

If $c < c_b < c_l$, the NB manufacturer decides between "total" *foreclosure* and *deterring PL's entry*. This comes, for example, from the fact that $\pi_w^c > \pi_w^e$. Furthermore, we find that $\pi_w^a > \pi_w^c$ if and only if $c_b - c > \frac{t}{2\alpha} - 2t(1-\alpha)$.

If $c < c_l < c_b$, the NB manufacturer decides between *foreclosure* and *partial foreclosure*. To show that $\pi_w^e > \pi_w^c$, just notice that the NB manufacturer can do better than π_w^c by maintaining quantities and transport costs minimized but by changing the product sold to retailer 1 from the NB to a PL, since production costs decrease. Furthermore, we find that $\pi_w^e > \pi_w^a$ if and only if $c_b < c_l + 2t(2\alpha - 1)$ and $c_b < c + \frac{1}{8t\alpha}(2t + c_b - c_l)^2 - 2t(1-\alpha)$.

References

[1] Bell R., R. Davies and E. Howard (1997), *The changing structure of food retailing in Europe: the implications for strategy*, Long Range Planning, 30, pp 853-861.

[2] Bergès-Sennou F., P. Bontems P. and V. Réquillart (2004), *Economics of Private Labels: A Survey of Literature*, Journal of Agricultural & Food Industrial Organization, Vol. 2, Article 3.

[3] Bergès-Sennou F. (2006), *Store loyalty, bargaining power and the private label production issue*, European Review of Agricultural Economics, Vol. 33, N^o 3.

[4] Bontems P., S. Monier and V. Réquillart (1999), *Strategic effects of private labels*, European Review of Agricultural Economics, 26, pp 147-165.

[5] Caprice S. (2000), *Contributions à l'analyse de la puissance d'achat dans les relations verticales: interactions stratégiques et marques de distributeur*, PhD Thesis, University of Paris.

[6] Dhar S. K. and S. J. Hoch (1997), *Why Store Brand Penetration Varies by Retailer*, Marketing Science, 16, pp 208-227.

[7] Gabrielsen T. S. and L. Sorgard (2007), *Private Labels, price rivalry and public policy*, European Economic Review, Vol. 51, Issue 2, February, pp 403-424.

[8] Lal R. (1990), *Manufacturer trade deals and retail price promotions*, Journal of Marketing Research, 27 (November), pp 428-444.

[9] Mills D. E. (1995), *Why retailers sell Private Labels*, Journal of Economics and Management Strategy, Vol. 4, N^o 3, Fall, pp 509-528.

[10] Mills D. E. (1999), *Private labels and manufacturer counterstrategies*, European Review of Agricultural Economics, 26, pp 125-145.

[11] Morton F. S. and F. Zettelmeyer (2004), *The strategic positioning of store brands in retailer-manufacturer negotiations*, Review of Industrial Organization, Vol. 24, N^o 2, March, pp 161-194 (34).

[12] Narasimhan C. and R. T. Wilcox (1998), *Private Labels and the Channel Relationship: A Cross-Category Analysis*, Journal of Business, Vol. 71, N^o 4.

[13] Quelch J. A. and D. Harding (1996), *Brands Versus Private Labels: Fighting to Win*, Harvard Business Review, 74, pp 99-109.

[14] Raju J. S., R. Sethuraman and S. K. Dhar (1995), *The Introduction and Performance of Store Brands*, Management Science, 41, pp 957-978.

[15] Steiner R. L. (2004), *The Nature and Benefits of National Brand/Private Label Competition*, Review of Industrial Organization, 24, pp 105-127.

NO DEEP POCKETS: SOME STYLIZED RESULTS ON FIRMS' FINANCIAL CONSTRAINTS

FILIPE SILVA
E-mail: filipeourico@googlemail.com
Faculdade de Economia
Universidade de Coimbra

CARLOS CARREIRA
E-mail: ccarreir@fe.uc.pt
Faculdade de Economia
Universidade de Coimbra

Abstract
In this paper, we make a brief survey of recent empirical work on financing constraints. It is organized as a series of ten stylized results that reflect what is generally understood about the severity of financial constraints and effects that they have upon firms. The survey clearly shows that financial constraints is a widespread major concern for firms, hindering their ability to carry out their optimal investment and growth trajectories. The severity of such constraints depends on institutional and firm specific characteristics, as well as on the nature of investment projects.

Keywords: Survey; Financing constraints; Firm investment
Thematic Area: Industrial Economics

1. INTRODUCTION

What do we know about financing constraints? Theory is old and vast but only recently it did experience a widespread of empirical studies on financial constraints. Modigliani-Miller theorem tells us that external finance is a perfect substitute for internal finance, thus financial structure of firms and financial policy is irrelevant for its investment decisions. However, this will require the strong assumption of perfect capital markets which does not hold in reality – firms face difficulties when deciding to borrow (debt or equity). The existence of capital market imperfections due to information problems is well documented (see Stiglitz and Weiss, 1981 or Myers and Majluf, 1984). The rationale is based on the existence of credit risk and information asymmetries (moral hazard and adverse selection problems) that either set the price of money on above-optimal levels or rationalize the credit. This inefficiency will set a wedge between internal and external forms of firm finance – a financing hierarchy. As a result, firms will not be able to raise the necessary amounts to fulfil their investment and growth goals. The severity of constraints and their impact will depend on firms' and institutional characteristics as well as the type of investment.

Although the existing theoretical literature on the effect of financing constrains on growth is far extensive, empirical contributions seem to be rather scarce and recent. Nonetheless, the aim of this paper is to review the existing empirical literature on financial constraints, focusing on recent contributions. We shall make an incursion, though, on central theoretical results when needed.

The framework chosen to tackle this task has been to pinpoint and discuss stylized results that have gathered consensus among researchers. However, before we address these stylized results, a discussion must take place regarding theoretical aspects of financial constraints analysis. We shall provide the arguments, evidence and conclusions of both sides. The paper

will be organized as follows. Section II will prepare the ground for empirical evidence by providing a discussion over investment-cash flow sensitivities. Section III presents 10 stylized results, while Section IV raises some practical concerns that researchers in the field might find. Section V will pull the pieces together and concludes.

2. INVESTMENT-CASH FLOW SENSITIVITY: AN ONGOING DEBATE

Is Investment-Cash Flow Sensitivity (ICFS) higher for financial constrained or unconstrained firms? This is the same as asking whether or not is high ICFS a good measure of the presence of financing constraints. Fazzari, Hubbard and Petersen (1988), (hereafter FHP) investigate the impact of cash-flow sensitivities on investment by classifying firms according to their dividend policy.[1] The reason for this classification lies on the argument that "firms might pay low dividends [because] they require investment finance that exceeds their internal cash flow and retain all of the low-cost internal funds they can generate". Using a poled regression of investment/capital stock on the cash-flow/capital stock ratio, estimated Q (controls for investment opportunities)[2] and dummies for each firm and year, upon a sample consisting of 422 U.S. firms (1970-1984), they find that the coefficient of cash-flow for the low-dividend group is higher and statistically different than the coefficient for the high-dividend group. This suggests that low-dividend firms invest more of their extra cash-flow than high-dividend firms. Building on this study, a large body of literature investigates the presence and the impact of financial constraints through ICFS. Carpenter et al. (1998), who compare three measures of financing constraints (cash-flow sensitivity, cash stocks and coverage ratio), find evidence confirming cash-flow as preferred variable to test for the presence of financial constraints. Chapman et al. (1996) also find, for Australian firms, that investment is less sensitive to cash-flow when firms are financially unconstrained, while Bo et al. (2003), for a sample of Dutch listed firms, find that riskier firms face more severe financial constraints and argue that ICFS is a good proxy for financial constraints if "firms are classified by the degree of uncertainty they face and if the uncertainty originates from cost uncertainty".[3] In an interesting perspective, Almeida et al. (2004) propose the "propensity of firms to save cash out of cash-flows" (cash-flow sensitivity of cash) as a proxy for liquidity constraints, because only constrained firms will manage liquidity to maximise their value. They test if financially constrained firms exhibit high cash-flow sensitivities, while unconstrained firms exhibit no pattern. Results for 4/5 classification schemes for constrained/unconstrained firms confirm their hypothesis. Only for the classification based on Kaplan and Zingales (1997), (hereafter KZ) do the results differ.[4]

On the other hand, Kaplan and Zingales (1997) argue that cash-flow is not a good measure of the existence of financing constraints and FHP's a priori classification of firms is flawed. They instead classify firms according to information obtained from company annual reports and find evidence that constrained firms are the less sensitive to cash-flow.[5] This argument is also supported by Kadapakkam et al. (1998) and Cleary (1999). Almeida and Campello (2001) draw similar conclusions. Recently, Dasgupta and Sengupta (2007), for Japan, find that the response of investment to cash-flow shocks is non-monotonic, supporting Kaplan and Zingales (1998) and Cleary (1999). Notwithstanding, Allayannis and Mozumdar (2002) present an explanation

[1] Firms that pay low dividends are classified as financially constrained, while high-dividend firms are classified as unconstrained.

[2] Q measures the increase in the present value of a firm's profits resulting from a unit increase in the firm's capital stock and, as a result, Q is the market value of a unit of capital. A firm invests until the cost of acquiring capital equals the value of capital. A high level of Q thus indicates the presence of an investment opportunity. It is argued that Q summarizes all future information that is relevant for a firm when deciding to invest.

[3] Chapman et al. (1996) classify firms as financially unconstrained when a firm's net acquisition of financial assets exceeds its net incurrence of equity and debt liabilities.

[4] The classification schemes based on: payout ratio; asset size; bond rating; commercial paper rating; "KZ index".

[5] KZ re-examine the low- dividend FHP firms and categorize the firms as unconstrained, possibly constrained and constrained.

for Cleary (1999) and Kaplan and Zingales (1997) argument that investment is more sensitive to cash-flow for less constrained firms (influential observations for KZ and negative cash-flow observations for Cleary).

Recent literature has found the ICF relationship to be U-shaped, further adding to the controversy about the interpretation of cash flow sensitivities. Examples of these findings are Povel and Raith (2002), Cleary et al. (2007), Lyandres (2007), within a dynamic framework, and Guarilia (2008) for the U.K. (1993-2003 with 99% firms not traded) finds this relationship to hold for a wide range of industrial sectors.[6]

Finally, another group of literature points that cash-flows might contain information about firm's investment opportunities, meaning that Q should be corrected (e.g. Cummings et al., 2003; Erickson and Whited, 2000; Gomes, 2001; Alti, 2003; and Abel and Eberly, 2004).[7] Nonetheless, Alti (2003) finds that even after Q correction, all firms present sensitivity to cash-flow. In addition, Bhagat et al. (2005) find evidence that financially distressed firms exhibit positive investment-cash flow sensitivities if they operate at a profit, low sensitivity if operate at a loss and invest less than in the previous year and strong negative sensitivity if operate at a loss and invest more than in the previous year, while Chang et al. (2007) find supportive evidence for the case of Australia.

3. STYLIZED RESULTS

Stylized facts on financial constraints are rather hard to find, although one can not state the same with respect to stylized results. A fact is something hard and clear, that raises no possible doubts. On the contrary, a result drives from a robust relationship that seems to exist between variables and might be subject to questioning. With respect to financial constraints, a major problem arises because it is something of "abstract", thus there is no clear measure of this variable. The best one can do is to classify firms as being constrained/unconstrained and the severity of such constraints by using proxies. However, proxies always hide a relation between variables, implying that whichever finding one reaches, it is always a result and never a fact. As a consequence, we will not attempt to distinguish facts from empirical results and will focus on the latter. Keeping these caveats in mind, we were able to pinpoint 10 main results that reflect the topics that suffered major contribution during recent years.

Stylized result 1. Financial constraints are more severe for younger and smaller firms.

The problems with asymmetric information in capital markets are more sever for small and young firms. This will happen either because there is still not much information on these firms available to most potential lenders, or because of "weight" and visibility of such firms. Potential lenders are not able to observe the "quality" of the risk or do not have control over the firm's investment. Under these conditions, smaller and younger firms are expected to be more credit rationed, as shown by Jaffe and Russell (1976) and Petersen and Rajan (1994). In fact, Hyytinen and Vaananen (2006) find evidence that Finnish SMEs face severe adverse selection and moral hazard problems and that the former is empirically more prevalent than the latter. These problems, particularly the former, are found to restrict firms' ability to raise external funds needed to take advantage of investment opportunities.

Although Kadapakkam et al. (1998) and Clearly (1999), for developed countries, find the opposite relationship (i.e. cash flow-investment sensitivity is highest in the large firm size group and smallest in the small firm size group). This result can possibly arise because "larger firms

[6] Agriculture, forestry and mining, manufacturing, construction, retail and wholesale, hotels and restaurants and business and other services; excludes regulated and financial sectors. [] In particular for small firms, cash-flows reveal them the direction to go in presence of growth prospects uncertainty. Bond and Van Reenen (1999) point that ICFS might also indicate other sources of misspecification.

[7] In particular for small firms, cash-flows reveal them the direction to go in presence of growth prospects uncertainty. Bond and Van Reenen (1999) point that ICFS might also indicate other sources of misspecification.

have greater flexibility in timing their investments and have more managerial agency problems" (Kadapakkam et al., 1998). Pratap (2003) advances the hypothesis that adjustment costs explain possible insensitivity of small firms' ICF, as firms do not take major investments before they attain a threshold level of liquidity.

Engeln et al. (1997), classifying firms into financially constrained and unconstrained, find evidence for Germany supporting greater constraints for smaller and younger firms.[8] In addition, using a panel of more than 1,600 small U.S. firms, Carpenter and Petersen (2002) find that the typical firm retains all of its income and raises relatively little external finance and that, for firms that rely mostly on internal finance (90% of their sample), the impact of cash-flow on growth is greater than the unity, meaning that the growth of most small firms is constrained by internal finance. Even if for constrained firms cash-flow is independent of size, then growth will be independent of size, but the variance of growth rates will decline with size as larger firms appear to be less constrained to internal finance. Further supporting evidence is found in Oliveira and Fortunato (2006) who, using an unbalanced panel of 7,653 surviving Portuguese manufacturing firms (1990-2001), also find that smaller firms' growth is more sensitive to cash-flow implying the existence of financing constraints for such firms. Additionally, this finding also applies for younger firms. As to the overall sample, this sensitivity continues to be significant but lower.[9]

Although Audretsch and Elston (2002), for the German case , find that the firms which report higher ICFS are medium sized. They argue that SMEs in Germany appear to benefit from a bank-oriented financial system and an institutional set that prevents smaller firms to face greater liquidity constraints.[10] Their study also confirms that very large firms do not appear to be liquidity constrained. In the case of Italian firms, Becchetti and Trovato (2002) find that while for small firms (<50 employees) the access to external finance is a key limiting determinant of growth, for larger firms (>100 employees) the financial factor appears to be neutral. Bhaduri (2005) also reaches the same conclusions for the Indian case.

Finally, Budina et al. (2000) find evidence suggesting that firms in Bulgaria (1,003 firms, for the period 1993-95) are financially constrained and in particular for smaller firms the constraints are more severe, while Honjo and Harada (2006) for a sample of 6,961 Japanese firms during the period 1995-99, find higher growth CFS for younger SMEs, confirming Ogawa et al. (1996) earlier findings.

Stylized Result 2- Size distribution of firms is skewed due to financial constraints.

A set of studies concerning the evolution of firm size distribution, focus on the idea that, contrary to the stipulated by Gibrat's Law, growth rates and growth volatility are negatively associated with firm size (and age). One of the explanations lies on the financing constraints argument. As a result, one should expect that the presence of financing constraints leads to a skewed distribution of firm size. Cooley and Quadrini (2001), exploring financial market frictions as in Albuquerque and Hopenhayn (1977), find that smaller firms face higher probability of default, take on more debt, issue more shares and pay fewer dividends, and have higher growth rates and volatility. So, they argue, imperfect markets will lead to a skewed size distribution of firms. Using another framework, Cabral and Mata (2003) develop a model of heterogeneous constraints at the entrepreneur level and analyse the evolution of the distribution of firm size.[11] Their findings, obtained using a sample of 514 Portuguese surviving firms in

8 Two classification schemes are used by the authors: self-evaluation and CREDITREFORM (largest credit agency by the time)

9 Small firms are defined as those with less than 50 employees and young firms defined as with less than 10 years. These limits are arguable.

10 Note that the authors divide firms into 4 classes according to their size (<500 employees;500-1300 employees; 1300-5500 employees; >5500 employees) and thus the first class might be seen as small and medium enterprises (SMEs). Higher and most significant sensitivities are found for classes 2 and 3.

11 Note that they only allow for 2 periods and use entrepreneur's age as proxy for financial capacity. They argue that younger firms/entrepreneurs have higher probability of being financially constrained

1984 and 1991, suggest that age has a significant influence upon the size distribution. In particular, the younger firms/entrepreneurs are, the greater is the skewness of the distribution, explained by the financial constraints. Faggiolo and Luzzi (2006) reach the same conclusions using a sample of 14,277 Italian surviving firms during the period 1995-2000. Other studies such as Desai et al. (2003) and Lotti and Santarelli (2004) also support these results.

Stylized Result 3- Start-ups/new entrepreneurs appear to be financially constrained.

The founders of new firms, as being unable to raise the desired capital through their own wealth, have to apply for external finance. In order to finance such projects, banks, due to information asymmetry problems, require collateral as well as minimum equity-debt ratio levels that in their turn are difficult to be achieved by founders. This credit rationing eventually hinder the creation of such firms, possibly firms whose expected returns on investment are higher than interest rates charged on remaining capital required.

Heino (2006) finds that, for the Mexican case, there is evidence that start-ups are financially constrained, as there is a strong positive relationship between probability of entering the market and entrepreneurs using their own personal savings to finance the initial investment. Meanwhile, Blumberg and Letterie (2008), through an entrepreneur perspective in an application to the Dutch reality, find that there are a set of characteristics that banks take into account when providing credit to start-ups. In particular, personal wealth, home ownership, use of own capital in the initial investment and earning capacity, all reduce the probability of a bank credit being denied. They also point out that banks tend to prefer commitments and signals when deciding to lend and that decisions to apply, even with potentially very profitable projects, might not occur due to the fact that entrepreneurs might know in advance that their project will be denied, according to the characteristics that banks value most (financial aspects). In another perspective, Bohacek (2006) finds that business families have high levels of savings that are explained by the incentive to overcome the financing constraint that prevents them from operating their firms at the optimal size. In a recent paper, Nykvist (2008), for Sweden, finds a positive relationship between personal wealth and entrepreneurship.

There is a body of empirical literature that points in the opposite direction. However, such conclusions might be flawed due to correlations between personal wealth and human capital. If human capital is shown to be highly correlated with personal wealth, then including both variables in the model might lead to the loss of explanatory power of one of them. This is what happened with Kim et al. (2006), for the U.S., and Cressy (1996) who uses a UK bank database of 2,000 start-ups, and argues that there is no credit rationing of start- ups. Grilo and Irigoyen (2006) find that the perception of the lack of financial support has no power to explain latent entrepreneurship, and Van Gelderen et al. (2006), for a sample of 21,393 potential Dutch entrepreneurs, finds that one of the factors that affect negatively start-up success is intended start-up capital. The interpretation in this case flows from the fact that is easier to raise smaller amounts to start a business, whilst potential successful entrants who require larger amounts of initial capital and have to resort to external financing, end up compromising their success.

Finally, van Auken, (1999) and Aghion et al. (2007) find financing constraints to work as a barrier to entry, even after controlling for other entry barriers. The reasoning is that capital requirements deter entry of new firms that have limited access to funds, in particular small firms.

Stylized Result 4- Financial constraints are determinant for firm survival.

It is widely accepted that the survival rate of most entrants is low (Geroski, 1995). One of the reasons that lay behind low survival rates of new entrants is financing constraints. Fotopoulos and Louri (2000) find evidence for Greek manufacturing firms that initial financial capital and the ratio of fixed to total assets significantly lowers the probability of a firm dying, while leverage increases the probability of death. For the Portuguese case, Farinha (2005) using a sample of 3,354 firms during the period 1985-98 also finds that the probability of survival is lower for firms that face financing constraints, have smaller initial capital, are more leveraged

and with higher number of credit relationships (as opposite to a stable relationship). She also points that the effects of financing constraints appear to be persistent in time (see also Petrunia, 2007). In a recent paper, Musso and Schiavo (2008) find that for French manufacturing firms, over the period 1996–2004, the greater the financing constraints firms face, the higher the probability that they do not survive and exit the market

With respect to the probability of default, Bottazzi et al. (2007), for a sample of 50,000 Italian firms during 1996-2003, find that the lower it is the credit rating of a firm, the higher is the probability of defaulting,. Also analysing default probabilities Cowling and Mitchell (2003), for the U.K., find that default is related to the cost of capital, with higher severity for smaller firms.

Finally, start-up size is generally accepted to have a positive impact on survival, since smaller entrants face higher probabilities of failure (see for example Audretsch and Mahmood, 1995, for U.S. or Mata and Portugal, 1994, and Farinha, 2005, for Portugal). However, one of the main determinants of start-up size is external financing (see Colombo and Grilli, 2005 for the Italian case and also the Eisfeldt and Rampini, 2007). As a result, financial constraints also have a negative indirect impact on firm survival.

Stylized Result 5- R&D investment is financially constrained.

The existence of financing constraints appears to be particularly severe for firms that decide to invest in R&D because of the risks associated with the investment. As argued before, credit markets will no longer be efficient, generating a wedge between internal and external financing faced by firms as well as a financing hierarchy. This problem is usually associated with Akerlof's (1970) adverse selection in the "market for lemons" (see also Leland and Pyle, 1977, and Myers and Majluf, 1984, for information asymmetries in R&D financing). For example, Hall (1992), Hao and Jaffe (1993) and Himmelberg and Petersen (1994), for U.S. firms, find support for the hypothesis that R&D investment is financially constrained in particular for small firms. Hall et al. (1999) in a comparative study of French, Japanese and U.S. firms also sustain these findings, while Kukuk and Stadler (2001) find that financial constraints adversely affect the timing of innovations for German services firms.

For the Italian case, both Scellato (2007) and Colombo and Grilli (2007) find that most new technology-based firms finance their initial activities through founder's wealth (97.9% of a sample consisting of 386 high-tech Italian firm). Czarnitzki (2006) also finds that, in West Germany case, SME's R&D investment is financially constrained.

When compared with physical capital, R&D appears to be more financially constrained. As pointed by Hall (1992), due to the risky nature of R&D, firms prefer to finance their projects with internally generated funds. Using the same unbalanced panel of roughly 3,000 U.S. firms (41,702 observations from 1959 to 1991), Chiao (2002) tests this hypothesis dividing firms into 2 classes: science based and non-science based. He finds that, in fact, while there is a positive relationship of current debt and current physical investment, the relationship between current debt and current R&D depends on the type of industry where the firm is operating. In particular, for science based industries, current R&D has a negative relationship with debt. Bougheas et al. (2003) also find evidence of financing constraints in R&D investment in Ireland, while they do not find evidence of constraints for investment in physical capital.

Stylized Result 6- Government policy to alleviate financial constraints seems to work.

Despite the inefficiencies that public intervention might cause, subsidies through taxes or in the form of grants and credit guarantees have proved to reduce substantially the severity of financial constraints faced by firms.[12] Becchetti and Trovato (2002), for Italian firms, for example, find evidence that supports a positive impact of subsidies on alleviating firms'

12 As an example of distortions caused by public intervention, de Meza and Webb (1999) points that credit policy encouraging entry might cause excessive entry.

financial constraints, while Li (2002) reaches the same conclusions analysing U.S. Government credit subsidies to poor and capable entrepreneurs. Following Hall's (2002) suggestion that public incentives to R&D might have an alleviating effect on constraints, Czarnitzki (2006) investigates SME's in Germany (1994, 1995 and 1998) finding that, while in West Germany SME's R&D investment is financially constrained, in East Germany (where public subsidies seems to be the driving force behind R&D investment) firms appear not to suffer from external finance constraints.[13] Honjo and Harada (2006), for Japanese firms during the period 1995-1999, also find that public programs to foster SMEs have a high and positive impact in the growth of SMEs, while Klassen et al. (2004) reach the same conclusions for U.S..

With respect to credit guarantees, Cowling and Mitchell (2003), analysing the impact of backed debt contracts on default probabilities for UK firms, find that the policy undertaken by the Government was successful in alleviating a very real capital constraint for the majority of small business, while Kang and Heshmati (2008) observe that a Korean public policy providing credit guarantee to SMEs had significant impact on survival of the latter. Guaranteed firms had clear higher survival rates and better performance.[14]

Stylized Result 7- Close bank relationships alleviate financial constraints.

Close bank relationships facilitate the contact between firms and banks thus reducing the information asymmetries, which means lower financing constraints for firms with close, but few relationships (in particular if such relationships are stable). As Diamond (1991) argues, the risk associated with any particular loan is not neutral with respect to the duration of the relationship. As a result, one can expect differences in financial constraints between market-oriented economies (such as U.S. and U.K.) and bank-oriented ones (Germany for example). In fact, Fohlin (1998), in an historic perspective (1903-1911), finds that German style close bank relationship had a significant impact in reducing financing constraints.

Bond et al. (2003) analyse ICFS for firms in Belgium, France Germany and the U.K. and they find that, while for Belgian firms there is no evidence of the existence of financing constraints, there is strong evidence suggesting that firms in the U.K. face severe financing constraints.[15] For the French and German case, there still exist significant financial constraints but to a lesser extent than the U.K. case. This finding adds to the hypothesis that a bank-oriented financial system (continental case) lightens constrains on (usually small) firms, when compared to a market-oriented system. The same rationale is used by Audretsch and Elston (2002), who find that the firms that report higher ICFS are medium sized.[16] They argue that this results from the fact that SMEs in Germany appear to benefit from a financial system that is bank-oriented and an institutional set that prevents smaller firms to face greater liquidity constraints. Recently, Semenov (2006) also gives supportive evidence of the presence of greater financing constraints in market-oriented economies in a study comparing 11 OECD Countries.

Stylized Result 8- Firms are less financially constrained in Economies with more developed financial markets.

It has been argued that financial constraints arise with imperfect financial markets owing to asymmetric information between firms and finance suppliers, leading to adverse selection and moral hazard problems and so, to a wedge between the costs associated with internal and those

13 With care needed to evaluate the efficiency of R&D investment in East Germany.

14 Backed debt contracts resulted from U.K.'s Small Firms Loan Guarantee Scheme: a policy undertaken by U.K. Government to alleviate financing constraints of small firms.

15 Have high significant cash-flow sensitivities and strongly rejects the Euler's equation null hypothesis of no financing constraints.

16 Note that the authors divide firms into 4 classes according to their size (<500 employees;500-1300 employees; 1300-5500 employees; >5500 employees) and thus the first class might be seen as small and medium enterprises (SMEs). Higher and most significant sensitivities are found for classes 2 and 3.

associated with external finance. As a result, one can expect that firms operating in countries with less developed/integrated financial markets will be more exposed to the constraints as there will be a greater amount of frictions in the market. Oliveira e Fortunato (2006) point out that one of the reasons why Portuguese firms are severely financially constrained, particularly small and young ones, might be due to a relative underdevelopment of capital markets in Portugal when compared to U.S. or U.K. (firms in Portugal rely mostly on banks for external finance).

Hutchinson and Xavier (2006) compare the role of internal finance in Slovenia (a leading transition country with still less developed financial markets) with the well established market economy of Belgium. They find that Slovenian firms' growth is more sensitive to cash-flow than Belgian, with the sensitivity being greater for the case of small firms in either country. Also for Eastern European countries, Budina et al. (2000) for Bulgaria and Konings et al. (2003) for Poland, Czech Republic, Bulgaria and Romania, identify high ICFS possibly indicating the presence of severe financial constraints.[17] Desai et al. (2003), comparing the skewness of distributions between Western, Central and Eastern Europe, as well as particularising for the U.K. case, also support these results. Hartarska and Vega (2006) and Hartarska and Nadolnyak (2008), respectively, for Russia and Bosnia-Herzegovina, find higher ICFS for firms that had no access to microfinance, vital in less financially developed economies.

In addition, for a sample of harmonized firm-level data for 16 industrialized and emerging countries, Aghion et al. (2007) also reinforce the importance of the development of financial markets in allowing the access to external finance for small firms. They also stress the impact that financial constraints do have for entry of small firms, working as an entry barrier. Meanwhile, Islam and Mozumdar (2007), for a sample of 31 countries during the period 1987-1997, confirm that investment is more sensitive to cash-flow for firms in economies with less developed financial markets. Love (2003) also supports this evidence.

On another perspective, De-Haas and Peeters (2006) finds that, for transition economies (10 countries analysed), firms increased their leverage during the transition period in order to reach their optimum target leverage level, despite many firms still prefer internal funding due to information asymmetries in the market for cash. These results imply that, before transition, firms were operating below optimal leverage levels and thus suffered from external finance constraints.

With respect to regional financial development, in a comparison between southern and northern regions of Italy, Sarno (2005) finds that SMEs in the less financially developed southern Italian regions are more severely affected by credit constraints. Guiso et al. (2004) also for the Italian case, analysing the impact of local financial development, confirm that the development of local financial markets has a major impact in relieving financial constraints, in particular for smaller firms that have no access to broader capital markets.

Stylized Result 9- Foreign-owned firms in less financially developed Countries are less financially constrained than domestic ones.

In less financially developed Countries, firms that have access to foreign capital markets are expected to be less financially constrained than those that have to resort to national capital markets, thus facing information asymmetries and severe financial constraints. Ownership (national versus foreign) is believed to create a wedge between firms that have access to outside markets (foreign-owned) and those that do not. Blalock et al. (2008) find evidence for Indonesia suggesting that firm ownership is an important determinant of investment as foreign ownership is associated with firms not being financially constrained. Colombo and Stanca (2006) also find this wedge in budget constraints for firms in Hungary (1989-1999). The introduction of financial reforms in the country leads domestic firms to face stronger constraints with respect to foreign-owned firms. Public firms (particularly small ones), however, were unaffected and remained under soft budget constraints. Hutchinson and Xavier (2006) find that foreign owned firms in Slovenia are less constrained and Ruiz-Vargas (2000) corroborate these results for the

[17] Konings et al. (2003) also find that firms in Poland and Czech Republic experience higher investment- cash flow sensitivities than in Bulgaria and Romania(explained by soft-budget constraints).

case of Puerto Rico. Beck et al. (2006) for more than 10,000 firms, over 80 countries, also find that domestic firms report higher financing obstacles.

Stylized Result 10- Financial liberalization seems to alleviate financing constraints of firms, especially for smaller ones.

In Countries were financial markets are still overregulated, financial liberalization by means of improving the allocation of credit, lightens constraints on firms with potential favourable growth prospects. In particular, it eases the access of small firms to credit, who pre-liberalization had no preferential treatment (eventually given to larger firms under overregulated markets). Gelos and Werner (2002), analysing the Mexican manufacturing sector for the period 1984-94, find that investment was generally financially constrained and that liberalization eased the access of small firms to external funds (note that large firms previously had a preferential treatment). Ghosh (2006) also finds that financial liberalization in India during the 1990s eased financial constraints, in particular for small firms. While, Koo and Maeng (2005) and Laeven (2002) support these findings for Korea, Wang (2003) provides evidence for the Taiwanese case. Finally, Laeven (2003), for a large number of firms in 13 developing countries, finds that financial liberalization affects firms differently- while it eases financing constraints for small firms, it results on the opposite for larger ones.

With respect to securities regulation, MacIntosh (1994) identified legal and institutional barriers to financing innovative companies in the U.S. and Canada, while Chiu (2003) for Canada and Europe and Cohn (1999) for the U.S also find supporting evidence of the negative effects of regulation in aggravating financing constraints.

4. SOME EMPIRICAL NOTES

We shall briefly address some empirical issues, mostly related to data and sampling, that are often raised when analysing financing constraints. Firstly, care is needed when analysing small firms or SMEs due to the heterogeneity of definitions found. Just to illustrate, a small firm might mean a firm with less than 5, 20 or 50 employees and a "small firm" in the U.S. is different from a "small firm" in Europe.

Secondly, one must be aware of the data source and the sample in observation. Most studies focus solely on firms in the manufacturing sector or firms that are publicly traded (both are the easiest samples to obtain). This reduced samples, might generate biases if one expects to draw conclusions for the population. As an example, if one investigates if financing constraints are more severe for smaller firms using traded firms it is reasonable to expect that the observed firms will not be as constrained as if the sample included non-traded firms. This would happen because traded firms have to provide extra information to the market, thus information asymmetries are necessarily lower for these firms.[18]

Thirdly, when analysing start-ups, it is difficult to identify potential entrants and obtaining their financial status information. One only observes firms that indeed overcome entry barriers (stressing the financial barrier). As a result, firms that did not overcome the entry barriers are not taken into account.

Fourthly, several problems emerge in survival analysis: (a) length biased sampling – if using stock sampling, one may not observe firm dynamics in between sample periods, thus survival rates tend to be larger; (b) repeated entry – one must consider "de novo entry" as not independent of previous failures (usually this is not taken into account); (c) death definition – firms may exit the sample even if they did not died (using two missing observations in a row to define death, as it is usually done, probably will not expurgate all other exits).

Fifthly, several problems arise when using ICFS to test for the presence of financing constraints. Despite this discussion has already took place, we must reinforce that using Q

[18] Kim (1999), for Korea, finds that when firms decide to go traded they experience lower ICFS and that the financing constraint relaxation is more noticeable for smaller firms.

theory derived models might entail significant miss measurement errors due to the estimation of Q (see Poterba, 1988, Erickson and Whited, 2000 and Alti, 2003 for a discussion).

Finally, constraint classification is a major issue of financing constraints investigation. Financial constraints are an abstraction, so researchers use proxies such as the followings: (a) dividend payout ratio (FHP); (b) firm self evaluation (Becchetti and Trovato, 2002); (c) cash stocks (Almeida, 2004); (d) degree of leverage; Age; Size (Cabral and Mata, 2003); (e) institutional affiliation (see Chirinko, 1997); (f) credit ratings. Most of them result from previous empirical findings suggesting a strong association between variables despite they can always be subject to questioning.

5. CONCLUDING REMARKS

So, what can we say about financial constraints? This survey clearly shows that financing constraints is a widespread major concern for firms, hindering their ability to carry out their optimal investment and growth trajectories, and that the severity of such constraints depends on institutional and firm specific characteristics, as well as, on the nature of investment projects.

Despite being quite an old topic, financial constraints analysis is still subject to much debate. As it has been shown, the majority of studies rely upon a fragile relationship built to identify the presence of financial constraints (ICFS). Surprisingly, or not, after more than a decade of intensive debate, no definitive conclusions may be drawn. With that caveat in mind, alongside with applications of other extremely useful types of analysis (such as survival analysis or optimal lending contracts), this framework has been applied to a growing body of empirical literature. The core empirical evidence can be summarized in 10 basic results explored in the paper.

Firms' ability to raise external funds appears to be highly dependent of size and age, as information, available to lenders, on these firms is rather scarce due to their small size and/or tender age. In addition, the level of risk associated with the investment projects of such firms is usually not observable, resulting in information asymmetry problems. This credit rationing highly affects firm's investment and growth and reproduces itself, as some firms "are small because they want to be small on efficiency grounds, whereas others are small because they are financially constrained" (Cabral and Mata, 2003), leading to a skewed size distribution of firms. Financing constraints hinder the possibility of potentially profitable firms being successful and grow. In fact, if we think of survival rates, we can observe that these are extremely low in general and that one of the main estimates that explain this fact is financial constraints. However, one must bear in mind that results on survival rates may hide a downward bias as firms that do not overcome entry barriers are not observed. Indeed, financing constraints seem to be particularly severe for start-ups and new entrepreneurs, in many cases to an extent that deters potential entrants, so we may think of financing constraints as a barrier to entry.

Within a Schumpeterian framework, firm dynamics is highly affected by investments in innovation/R&D. However, firms' investment in R&D is severely financially constrained due to the risk associated with innovative projects, which appears to be higher than the risk associated with physical capital investment projects. As a consequence, financing constraints is a pivotal element that shapes the dynamic path of firms. Future research concerning the impact financing constraints have on the dynamics of firms in an evolutionary environment is warranted.

On the other hand, financial constraints are a product of market imperfections, so whatever affects information sets in the financial market will necessarily affect the impact of financing constraints have on other variables. In particular, the development of financial markets, by means of increased market efficiency, seems to result in lighter financial constrains, as information flows faster and in larger amounts. Nonetheless, bank-oriented economies, where firms have few stable and close bank relationships, seem to work better in alleviating constraints than do market-oriented economies. This results from the fact that stable and close relationships greatly reduce the information gap between borrower and lender, to an extent unbeaten by highly efficient financial markets. By contrast, in countries with less developed financial markets, one can expect to find greater financing constraints as well as a wedge between firms

which have access to foreign markets (usually foreign owned firms) and those that do not (domestic firms). Financial markets deregulation seems to have a positive impact in reducing financing constraints, in particular in cases where markets were overregulated or mechanisms of preferential treatment were allowed.

Finally, public policy at the firm level aimed at reducing the severity of financing constraints faced by firms seems to work, in particular intervention through grants, tax policy, as well as credit guarantees. The drawback is that such policies might create inefficiencies and result in welfare loss. In order to study the net effects of such policies, an integrated welfare analysis is needed.

In a last step, a discussion of issues brought up by the way financing constraints' problems are modelled as well as the way data is collected and classified calls for our attention when researching in the field.

REFERENCES

ABEL, ANDREW B. and EBERLY, JANICE C. (2004), "Q Theory without Adjustment Costs & Cash Flow Effects without Financing Constraints," 2004 Meeting Papers. Society for Economic Dynamics.

AGHION, PHILIPPE; FALLY, THIBAULT and SCARPETTA, STEFANO. (2007), "Credit Constraints as a Barrier to the Entry and Post-Entry Growth of Firms," IZA Discussion Papers. Institute for the Study of Labor (IZA).

AKERLOF, GEORGE A. (1970),"The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." The Quarterly Journal of Economics, 84(3), 488 - 500.

ALBUQUERQUE and HOPENHAYN (2004), "Optimal Lending Contracts and Firm Dynamics." The Review of Economic Studies, 71(2), 285- 315.

ALBUQUERQUE, RUI A. and HOPENHAYN, HUGO A. (2000), "Optimal Dynamic Lending Contracts with Imperfect Enforceability," RCER Working Papers. University of Rochester - Center for Economic Research.

ALLAYANNIS, GEORGE and MOZUMDAR, ABON. (2004), "The Impact of Negative Cash Flow and Influential Observations on Investment-Cash Flow Sensitivity Estimates." Journal of Banking & Finance, 28(5), 901 - 30.

ALMEIDA, H. and CAMPELLO, M. (2001), "Financial Constraints and Investment-Cash Flow Sensitivities: New Theoretical Foundations," New York University.

ALMEIDA, HEITOR; CAMPELLO, MURILLO and WEISBACH, MICHAEL S. (2004), "The Cash Flow Sensitivity of Cash." Journal of Finance, 59(4), 1777 - 804.

ALTI, AYDOGAN (2003), "How Sensitive Is Investment to Cash Flow When Financing Is Frictionless?" Journal of Finance, 58(2), 707 - 22.

AUDRETSCH, DAVID B. and ELSTON, JULIE ANN (2002), "Does Firm Size Matter? Evidence on the Impact of Liquidity Constraints on Firm Investment Behavior in Germany." International Journal of Industrial Organization, (20), 1.

AUDRETSCH, DAVID B. and MAHMOOD, T. (1995), "New-Firm Survival: New Results Using a Hazard Function." Review of Economics and Statistics, 77, 97–103.

BECCHETTI, LEONARDO and TROVATO, GIOVANNI (2002), "The Determinants of Growth for Small and Medium Sized Firms: The Role of the Availability of External Finance." Small Business Economics, 19(4), 291- 306.

BECK, THORSTEN; DEMIRGUC-KUNT, ASLI; LAEVEN, LUC and MAKSIMOVIC, VOJISLAV (2006), "The Determinants of Financing Obstacles." Journal of International Money and Finance, 25(6), 932 - 52.

BHAGAT, SANJAI; MOYEN, NATHALIE and SUH, INCHUL (2005), "Investment and Internal Funds of Distressed Firms." Journal of Corporate Finance, 11(3), 449- 72.

BLALOCK, GARRICK; GERTLER, PAUL J. and LEVINE, DAVID I. (2008), "Financial Constraints on Investment in an Emerging Market Crisis." Journal of Monetary Economics, 55(3), 568- 91.

BLUMBERG, BORIS F. and LETTERIE, WILKO A. (2008), "Business Starters and Credit Rationing." Small Business Economics, 30(2), 187 - 200.

BO , HONG; LENSINK, ROBERT and STERKEN, ELMER (2003), "Uncertainty and Financial Constraints." European Finance Review, 7(2), 297-321.

BOHACEK, RADIM (2006), "Financial Constraints and Entrepreneurial Investment." Journal of Monetary Economics, 53(8), 2195 - 212.

BOND, STEPHEN; ELSTON, JULIE Ann; MAIRESSE, JACQUES and MULKAY, BENOÎT (2003), "Financial Factors and Investment in Belgium, France, Germany, and the United Kingdom: A Comparison Using Company Panel Data." The Review of Economics and Statistics, 85(1), 153 - 65.

BOND, STEVE and REENEN, JOHN VAN (1999), "Investment, R&D and Financial Constraints in Britain and Germany," IFS Working Papers. Institute for Fiscal Studies.

BOTTAZZI, GIULIO; GRAZZI, MARCO; SECCHI, ANGELO and TAMAGNI, FEDERICO (2007), "Assessing the Impact of Credit Ratings and Economic Performance on Firm Default," LEM Papers Series. Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy.

BOUGHEAS, SPIROS; GÖRG, HOLGER and STROBL, ERIC (2003), "Is R & D Financially Constrained? Theory and Evidence from Irish Manufacturing." Review of Industrial Organization, 22(2), 159- 74.

BUDINA, NINA; GARRETSEN, HARRY and de JONG, ELKE (2000), "Liquidity Constraints and Investment in Transition Economies - the Case of Bulgaria," Policy Research Working Paper Series. The World Bank.

CABRAL, LUÍS M. B. and MATA, JOSÉ (2003), "Evolution of the Firm Size Distribution: Facts and Theory." American Economic Review, 93, 1075-90.

CARPENTER, ROBERT E.; FAZZARI, STEVEN M. and PETERSEN, BRUCE C (1998), "Financing Constraints and Inventory Investment: A Comparative Study with High-Frequency Panel Data." The Review of Economics and Statistics, 80(4), 513-19.

CARPENTER, ROBERT E. and PETERSEN, BRUCE C (2002), "Is the Growth of Small Firms Constrained by Internal Finance?" The Review of Economics and Statistics, 84(2), 298- 309.

CHANG, XIN; TAN, TEK JUN; WONG, GEORGE and ZHANG, HONGFENG (2007), "Effects of Financial Constraints on Corporate Policies in Australia." Accounting and Finance, 47(1), 85-108.

CHAPMAN, D R; JUNOR, C W and STEGMAN, T R (1996), "Cash Flow Constraints and Firms' Investment Behaviour." Applied Economics, 28(8), 1037-47.

CHIAO, CHAOSHIN (2002), "Relationship between Debt, R&D and Physical Investment, Evidence from Us Firm-Level Data." Applied Financial Economics, 12(2), 105 - 21.

CHIRINKO, ROBERT S (1997), "Finance Constraints, Liquidity, and Investment Spending: Theoretical Restrictions and International Evidence." Journal of the Japanese and International Economies, 11(2), 185 - 207.

CHIU, H. Y. (2003), "Can Uk Small Businesses Obtain Growth Capital in the Public Equity Markets? An Overview of the Shortcomings in Uk and European Securities Regulation and Considerations for Reform." Delaware Journal of Corporate Law, 28(3), 933–77.

CLEARY, SEAN (1999), "The Relationship between Firm Investment and Financial Status." Journal of Finance, 54(2), 673 - 92.

CLEARY, SEAN; POVEL, PAUL E M and RAITH, MICHAEL (2007), "The U-Shaped Investment Curve: Theory and Evidence." Journal of Financial and Quantitative Analysis, 42(1), 1-39.

COHN, S. R. (1999), "Impact of Securities Laws on Developing Companies: Would the Wright Brothers Have Gotten Off the Ground?" Journal of Small and Emerging Business Law, (3), 315–66.

COLOMBO, EMILIO and STANCA, LUCA (2006), "Investment Decisions and the Soft Budget Constraint." The Economics of Transition, 14(1), 171 - 98.

COLOMBO, MASSIMO G. and GRILLI, LUCA (2007), "Funding Gaps? Access to Bank Loans by High-Tech Start-Ups." Small Business Economics, 19(1), 25 - 46.

COLOMBO, MASSIMO G. and GRILLI, LUCA (2005), "Start-up Size: The Role of External Financing." Economics Letters, 88(2), 243 - 50.

COOLEY, THOMAS F. and QUADRINI, VINCENZO (2001), "Financial Markets and Firm Dynamics." American Economic Review, 91(5), 128 - 1310.

COWLING, MARC and MITCHELL, PETER (2003), " Is the Small Firms Loan Guarantee Scheme Hazardous for Banks or Helpful to Small Business?" Small Business Economics, 21(1), 63 - 71.

CUMMINS, J.G.; HASSETT, K.A. and OLINER, S.D. (2003), " Investment Behavior, Observable Expectations, and Internal Funds," Board of Governors of the Federal Reserve System.

CZARNITZKI, DIRK (2006), "Research and Development in Small and Medium-Sized Enterprises: The Role of Financial Constraints and Public Funding." Scottish Journal of Political Economy, 53(3), 335-57.

DASGUPTA, SUDIPTO and SENGUPTA, KUNAL (2007), "Corporate Liquidity, Investment and Financial Constraints: Implications from a Multi-Period Model." Journal of Financial Intermediation, 16(2), 151- 74.

DESAI, MIHIR; GOMPERS, PAUL and LERNER, JOSH (2003), "Institutions, Capital Constraints and Entrepreneurial Firm Dynamics: Evidence from Europe," NBER Working Papers. National Bureau of Economic Research.

DIAMOND, D. (1991), "Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt." Journal of Political Economy, 99(4), 688–721.

EGELN, JÜRGEN; LICHT, GEORG and STEIL, FABIAN (1997), "Firm Foundations and the Role of Financial Constraints." Small Business Economics, 9(2), 137–50.

EISFELDT, ANDREA L. and RAMPINI, ADRIANO A. (2007), "New or Used? Investment with Credit Constraints." Journal of Monetary Economics, 54(8), 2656 - 81.

ENGEL and MIDDENDORF, "Firms' Bank Relationships and Cash-Flow-Investment Sensitivities in Germany."

ERICKSON, TIMOTHY and WHITED, TONI M. (2000), "Measurement Error and the Relationship between Investment and Q." Journal of Political Economy, 108(5), 1027 - 57.

FAGGIOLO and LUZZI (2006), "Do Liquidity Constraints Matter in Explaining Firm Size and Growth? Some Evidence from the Italian Manufacturing Industry." Industrial and Corporate Change, 15(1), 1-39

FARINHA, LUISA (2005), "The Survival of New Firms: Impact of Idiosyncratic and Environmental Factors," F. S. Report, Banco de Portugal.

FAZZARI, STEVEN M.; HUBBARD, R. GLENN and PETERSEN, BRUCE C. (1988), "Finacing Constraints and Corporate Investment." Brookings Papers on Economic Activity, 19(1988-1), 141-206.

FOHLIN, CAROLINE (1998), "Fiduciariand Firm Liquidity Constraints: The Italian Experience with German-Style Universal Banking." Explorations in Economic History, 35(1), 83 - 107.

FOTOPOULOS, GEORGIOS and LOURI, HELEN (2000), "Determinants of Hazard Confronting New Entry: Does Financial Structure Matter?" Review of Industrial Organization, 17(3), 285 - 300.

GELDEREN, MARCO; THURIK, ROY and BOSMA, NIELS (2006), "Success and Risk Factors in the Pre-Startup Phase." Small Business Economics, 26(4), 319 - 35.

GELOS, R. GASTON and WERNER, ALEJANDRO M. (2002), "Financial Liberalization, Credit Constraints, and Collateral: Investment in the Mexican Manufacturing Sector." Journal of Development Economics, 67(1), 1 - 27.

GEROSKI, P. A. (1995), "What Do We Know About Entry?" International Journal of Industrial Organization, 13(4), 421- 40.

GHOSH, SAIBAL (2006), "Did Financial Liberalization Ease Financing Constraints? Evidence from Indian Firm-Level Data." Emerging Markets Review, 7(2), 176 - 90.

GOMES, JOAO F (2001), "Financing Investment." American Economic Review, 91(5), 1263 - 85.

GRILO, ISABEL and IRIGOYEN, JESUS-MARIA (2006), "Entrepreneurship in the Eu: To Wish and Not to Be." Small Business Economics, 26(4), 305 - 18.

GUARIGLIA, ALESSANDRA (2008), "Internal Financial Constraints, External Financial Constraints, and Investment Choice: Evidence from a Panel of Uk Firms." Journal of Banking & Finance, 32(9), 1795 - 809.

GUISO, LUIGI; SAPIENZA, PAOLA and ZINGALES, LUIGI (2004), "Does Local Financial Development Matter?" The Quarterly Journal of Economics, 119(3), 929-69.

HAAS, RALPH and PEETERS, MARGA (2006), "The Dynamic Adjustment Towards Target Capital Structures of Firms in Transition Economies " The Economics of Transition, 14(1), 133 - 69.

HALL, BRONWYN H (2002), "The Financing of Research and Development," NBER Working Papers National Bureau of Economic Research.

HALL, BRONWYN H (1992), "Investment and Research and Development at the Firm Level: Does the Source of Financing Matter?," NBER Working Paper. National Bureau of Economic Research.

HALL, BRONWYN H.; Mairesse, Jacques; Branstetter, Lee and Crepon, Bruno (1999), "Does Cash Flow Cause Investment and R&D?: An Exploration Using Panel Data for French, Japanese, and United States Scientific Firms " IFS Paper.

HAO, K.Y. and JAFFE, A.B. (1993), "Effect of Liquidity on Firms' R&D Spending." Economics of Innovation and New Technology, (2), 275-82.

HARTARSKA, VALENTINA and GONZALEZ-VEGA, CLAUDIO (2006), "What Affects New and Established Firmsâ€™ Expansion? Evidence from Small Firms in Russia." Small Business Economics, 27(2), 195- 206.

HARTARSKA, VALENTINA and NADOLNYAK, DENIS (2008), "An Impact Analysis of Microfinance in Bosnia and Herzegovina," William Davidson Institute Working Papers Series.

HEINO, HEIKKI (2006), "Use of Borrowed Start-up Capital and Micro Enterprises in Mexico: Existence of Liquidity Constraints." Portuguese Economic Journal, 5(1), 1-30.

HIMMELBERG, CHARLES P and PETERSEN, BRUCE C. (1994), "R&D and Internal Finance: A Panel Study of Small Firms in High-Tech Industries." The Review of Economics and Statistics, 76(1), 38- 51.

HONJO, YUJI and HARADA, NOBUYUKI (2006), "Sme Policy, Financial Structure and Firm Growth: Evidence from Japan." Small Business Economics, 27(4), 289 - 300.

HUBBARD, R. GLENN (1998), "Capital-Market Imperfections and Investment." Journal of Economic Literature, 36, 193–225.

HUTCHINSON, JOHN and XAVIER, ANA (2006), "Comparing the Impact of Credit Constraints on the Growth of Smes in a Transition Country with an Established Market Economy." Small Business Economics, 27(2), 169 - 79.

HYYTINEN, ARI and VÄÄNÄNEN, LOTTA (2006), "Where Do Financial Constraints Originate From? An Empirical Analysis of Adverse Selection and Moral Hazard in Capital Markets." Small Business Economics, 27(4), 323- 48.

ISLAM, SAIYID S. and MOZUMDAR, ABON (2007), "Financial Market Development and the Importance of Internal Cash: Evidence from International Data." Journal of Banking & Finance, 31(3), 641 - 58.

JAFFEE, DWIGHT M and RUSSELL, THOMAS (1976), "Imperfect Information, Uncertainty, and Credit Rationing." Quarterly Journal of Economics, 90(4), 651 - 66.

KADAPAKKAM, PALANI-RAJAN; KUMAR, P. C. and RIDDICK, LEIGH A. (1998), "The Impact of Cash Flows and Firm Size on Investment: The International Evidence." Journal of Banking & Finance, 22(3), 293 - 320.

KANG, JAE AND HESHMATI, ALMAS (2008), "Effect of Credit Guarantee Policy on Survival and Performance of Smes in Republic of Korea." Small Business Economics, 31(4), 445 - 62.

KAPLAN, STEVEN N and ZINGALES, LUIGI (1997), "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints." The Quarterly Journal of Economics, 112(1), 169 - 215.

KIM, GRACE (2006), "Do Equally Owned Small Businesses Have Equal Access to Credit?" Small Business Economics, 27(4), 369 - 86.

KIM, JUSEONG (1999), "The Relaxation of Financing Constraints by the Initial Public Offering of Small Manufacturing Firms." Small Business Economics, 12(3), 191- 202.

KIM, PHILLIP; ALDRICH, HOWARD and KEISTER, LISA (2006), "Access (Not) Denied: The Impact of Financial, Human, and cultural Capital on Entrepreneurial Entryin the United States." Small Business Economics, 27(1), 5 - 22.

KLASSEN, KENNETH J.; PITTMAN, JEFFREY A.; REED, MARGARET P. and FORTIN, STEVE (2004), "A Cross-National Comparison of R&D Expenditure Decisions: Tax Incentives and Financial Constraints " Contemporary Accounting Research, 21(3), 639 - 84

KONINGS, JOZEF; RIZOV, MARIAN and VANDENBUSSCHE, HYLKE (2003), "Investment and Financial Constraints in Transition Economies: Micro Evidence from Poland, the Czech Republic, Bulgaria and Romania." Economics Letters, 78(2), 253 - 58.

KOO, JAEWOON and MAENG, KYUNGHEE (2005), "The Effect of Financial Liberalization on Firms' Investments in Korea." Journal of Asian Economics, 16(2), 281 - 97.

KUKUK, MARTIN and STADLER, MANFRED (2001), "Financing Constraints and the Timing of Innovations in the German Services Sector." Empirica, 28(3), 277- 92.

LAEVEN, LUC (2003), "Does Financial Liberalization Reduce Financing Constraints? ." Financial Management, 32(1).

LAEVEN, LUC (2002), "Financial Constraints on Investments and Credit Policy in Korea." Journal of Asian Economics, 13(2), 251-69.

LELAND, HAYNE E. and PYLE, DAVID H. (1977), "Informational Asymmetries, Financial Structure, and Financial Intermediation." Journal of Finance, 32(2), 371 - 87.

LI, WENLI (2002), "Entrepreneurship and Government Subsidies: A General Equilibrium Analysis." Journal of Economic Dynamics and Control, 26(11), 1815 - 44.

LOTTI, FRANCESCA and SANTARELLI, ENRICO (2004), "Industry Dynamics and the Distribution of Firm Sizes: A Nonparametric Approach." Southern Economic Journal, 70(3), 443 - 66.

LOVE, INESSA (2003), "Financial Development and Financing Constraints: International Evidence from the Structural Investment Model." Review of Financial Studies, 16(3), 765-91.

LYANDRES, EVGENY (2007), "Costly External Financing, Investment Timing, and Investment-Cash Flow Sensitivity." Journal of Corporate Finance, 13(5), 959- 80.

MACINTOCH, J.G. (1994), "Legal and Institutional Barriers to Financing Innovative Entreprise in Canada." Queen's at Kingston - School of Policy Studies.

MATA, J. E. and PORTUGAL, A. (1994), "Life Duration of New Firms." The Journal of Industrial Economics, 42, 227-46.

MEZA, DAVID de and WEBB, DAVID (1999), "Wealth, Enterprise and Credit Policy." The Economic Journal, 109, 153-63.

MUSSO, PATRICK and SCHIAVO, STEFANO (2008), "The Impact of Financial Constraints on Firm Survival and Growth." Journal of Evolutionary Economics, 18(2), 135- 49.

MYERS, STEWART C. and MAJLUF, NICHOLAS S. (1984), "Corporate Financing and Investment Decisions When Firms Have Information that Investors Do Not Have." Journal of Financial Economics, 13(2), 187-221.

NYKVIST, JENNY (2008), "Entrepreneurship and Liquidity Constraints: Evidence from Sweden." Scandinavian Journal of Economics, 110(1), 23 - 43.

OGAWA, KAZUO; KITASAKA, SHIN-ICHI; YAMAOKA, HIROSHI and IWATA, YASUHARU (1996), "Borrowing Constraints and the Role of Land Asset in Japanese Corporate Investment Decision." Journal of the Japanese and International Economies, 10(2), 122 - 49.

OLIVEIRA, BLANDINA and FORTUNATO, ADELINO (2006), "Firm Growth and Liquidity Constraints: A Dynamic Analysis " Small Business Economics, 27(2), 139–56.

PETERSEN, MITCHELL A and RAJAN, RAGHURAM G. (1994), "The Benefits of Lending Relationships: Evidence from Small Business Data." Journal of Finance, 49(1), 3-37.

PETERSEN, MITCHELL A. and RAJAN, RAGHURAM G. (1994), "The Effect of Credit Market Competition on Lending Relationships," NBER Working Papers. National Bureau of Economic Research, Inc.

PETRUNIA, ROBERT (2007), "Persistence of Initial Debt in the Long-Term Employment Dynamics of New Firms." Canadian Journal of Economics, 40(3), 861 - 80.

POTERBA, JAMES M. (1988), "Comment on ‘‘Financing Constraints and Corporate Investment,’’." Brookings Papers on EconomicActivity, 200-04.

POVEL, PAUL AND RAITH, MICHAEL (2002), "Optimal Investment under Financial Constraints: The Roles of Internal Funds and Asymmetric Information " AFA 2002 Atlanta Meetings.

PRATAP, SANGEETA (2003), "Do Adjustment Costs Explain Investment-Cash Flow Insensitivity?" Journal of Economic Dynamics and Control, 27(11-12), 1993-2006.

RUIZ-VARGAS, YOLANDA (2000), "Small Business Financing Sources between Immigrants and Natives in Puerto Rico." The Quarterly Review of Economics and Finance, 40(3), 387 - 99.

SARNO, DOMENICO (2005), "Liquidity Constraint on the Production of Firms in Southern Italy." Small Business Economics, 25(2), 133 - 46.

SCELLATO, GIUSEPPE (2007), "Patents, Firm Size and Financial Constraints: An Empirical Analysis for a Panel of Italian Manufacturing Firms." Cambridge Journal of Economics, 31(1), 55 - 76.

SEMENOV, RADISLAV (2006), "Financial Systems, Financing Constraints and Investment: Empirical Analysis of Oecd Countries." Applied Economics, 38(17), 1963 - 74.

STIGLITZ, JOSEPH E AND WEISS, ANDREW (1981), "Credit Rationing in Markets with Imperfect Information." American Economic Review, 71(3), 393-410.

VAN AUKEN, HOWARD E. (1999), "Obstacles to Business Launch." Journal of Developmental Entrepreneurship, 4(2), 175–87.

WANG, HUNG-JEN (2003), \"A Stochastic Frontier Analysis of Financing Constraints on Investment: The Case of Financial Liberalization in Taiwan." Journal of Business & Economic Statistics, 21(3), 406 - 19.

LOS EFECTOS DE LA INVERSIÓN PÚBLICA EN I+D+I SOBRE EL CRECIMIENTO ECONÓMICO PARA UN CONJUNTO DE 15 PAÍSES EUROPEOS

IRENE ALBARRÁN LOZANO
E-mail: ialbarra@est-econ.uc3m.es

Departamento de Estadística
Universidad Carlos III de Madrid

PABLO ALONSO GONZÁLEZ
E-mail: pablo.alonsog@uah.es

Departamento de Estadística, Estructura Económica y O.E.I.
Área de Métodos cuantitativos para la Economía y la Empresa
Universidad de Alcalá

ANTONIO MARTÍNEZ GONZÁLEZ
E-mail: antonio.martinez@urjc.es

Departamento de Economía Aplicada
Universidad Rey Juan Carlos

Resumen

La investigación, desarrollo e innovación absorben importantes recursos económicos y humanos, dirigidos a favorecer la acumulación de capital en la que se basa el moderno crecimiento económico. Muchos autores apoyan la afirmación de que el nivel de inversiones en tecnología realizado por los mercados privados es socialmente sub-óptimo. Para ello, los mercados públicos deberían complementar las inversiones realizadas por las empresas y/o aumentar la financiación de la I+D+i pública. Conocer el impacto de la I+D+i pública sobre el crecimiento es importante, de cara a evaluar si su coste social es justificable. Este trabajo, utilizando datos proporcionados por Eurostat, analiza los efectos que sobre el PIB induce el gasto público en I+D+i para un grupo de países europeos que comparten características económicas similares, estudiándose las posibles correlaciones simultáneas mediante la estimación de las relaciones que se establecen entre las variables con el método SURE.

Palabras clave: I+D+i, crecimiento económico, productividad, competitividad
Área Temática: Economía Industrial.

Abstract

Important economic and human resources are devoted to R&D and innovation, with the aim of contributing to capital accumulation, which is the driving force of modern economic growth. Many authors maintain that the level of investments in R&D carried out by private markets is socially suboptimal. Therefore, public markets should supplement the investments carried out by enterprises and/or increase funding of public R&D and innovation. It is important to know the impact of public R&D and innovation on growth, in order to assess if its social cost is acceptable. This work, using data from Eurostat, analyses the effects on GDP of public expenditure in R&D and innovation for a group of European countries that share similar economic characteristics, considering possible contemporary correlations through the estimation of the relationship among variables with SURE method.

Keywords: R&D, economic growth, productivity, competitiveness
Thematic Area: Industrial Economics.

1. INTRODUCCIÓN

El análisis metodológico en relación al proceso de cambio que sobre el conjunto de la economía induce el fenómeno de la I+D+i, ha permitido apreciar sus efectos positivos sobre los procesos productivos como consecuencia de la mejora de sus rendimientos. De ahí la importancia indudable de poder descifrar los mecanismos por los cuales los avances en materia de investigación, desarrollo tecnológico e innovación inciden sobre el sistema económico en general, y el proceso productivo en particular. Una posible respuesta a esta pregunta reside en el hecho de que el fenómeno de la I+D+i posee una naturaleza económica de carácter intangible, frente a la material de los factores productivos clásicos. Esta particularidad claramente discriminadora permite establecer la condición que identifica a unos factores frente a otros como bienes económicos de naturaleza diferenciada. Mientras que en el caso de la I+D+i su naturaleza económica es "no rival", en el caso del capital físico y humano estaríamos tratando con bienes económicos de naturaleza "rival".

Debido a ello, muchos investigadores suelen apoyar la afirmación de que el nivel socialmente óptimo de inversiones en tecnología debería ser más elevado que el que en realidad se lleva a cabo. Ello implica que los mercados públicos deberían apoyar más ampliamente este tipo de inversiones realizadas por las empresas y/o aumentar la financiación de la I+D+i pública llevada a cabo por universidades e institutos de investigación. Un sólido sistema educativo parece ser uno de los factores clave para fomentar el progreso técnico de un país, razón por la que la investigación universitaria suele recibir un fuerte respaldo en casi todas las políticas públicas. Además, es necesario tener en cuenta que el efecto final de los recursos dedicados a innovación no es igual en toda circunstancia, pues dependerá de factores tales como la inversión previamente realizada, la composición de la misma y la estructura productiva sobre la cual va a incidir ese gasto en innovación. Por ello, puede ser oportuno cuestionarse por el grado de vinculación existente entre el nivel de producto final, medido por el PIB o alguna variable relacionada con él, y el volumen, composición y agentes encargados de ejecutar las inversiones en I+D+i.

Este estudio presenta la metodología empleada y los resultados obtenidos mediante el empleo del sistema de ecuaciones tipo SURE (Seemingly Unrelated Regresión Estimation), con el fin de evaluar si los efectos que producen las inversiones públicas en I+D+i sobre el nivel de renta, medido por el PIB, son similares entre países que comparten una zona geopolítica, con estructuras económicas vinculadas. Ello podría ayudar a identificar, dentro del conjunto de países considerados, la posible existencia de grupos homogéneos en la Unión Europea que permitan explicar la posición relativa de cada uno de ellos en relación al resto. Para ello, en la primera parte del trabajo se ha descrito brevemente el marco teórico que recoge las principales aportaciones realizadas en la literatura empírica relativa a la influencia de las inversiones públicas en I+D+i y el crecimiento económico. A continuación se han revisado las conclusiones obtenidas en las investigaciones realizadas en relación al impacto de las inversiones públicas en I+D+i y sus efectos sobre una elevación del producto, la renta y el bienestar social, en la medida en la que representan un fundamento y el sustento del análisis empírico efectuado en este trabajo. En la segunda parte, se ha desarrollado una investigación de carácter empírico mediante la aplicación del análisis de regresión multiecuacional, para así obtener agrupaciones de países en función de su posición respecto a la media de la relación entre el PIB y el gasto público previsto en I+D+i, que posteriormente es reestimada mediante el uso de relaciones simultáneas tipo SURE.

2. MARCO TEÓRICO

La investigación científica y la innovación utilizan una cantidad importante de recursos económicos y humanos con el fin de alimentar la acumulación de capital intangible en la que se basa el moderno crecimiento económico. Muchos estudios econométricos confirman la influencia positiva de las inversiones en I+D+i sobre el aumento de la productividad de los

factores (Mairesse y Sassenou, 1991; Hall y Mairesse, 1995; Guellec y van Pottelsberghe de la Potterie, 2001), aunque algunos otros muestran una contribución insignificante (Bartelsman, 1990; Lichtenberg y Siegel, 1991; Griliches, 1995; Hall, 1996).

Las políticas industriales de los gobiernos apoyan las inversiones empresariales para recursos destinados a la producción de innovaciones y nuevos descubrimientos técnicos y científicos, con el fin de mejorar la competitividad nacional. De hecho, la investigación produce sus efectos a través de dos formas de innovación: la mejora de la productividad de los factores (Bartelsman y Doms, 2000) y la mejora de la calidad de los productos que aumentan la demanda agregada, como consecuencia de la caída de costes y precios, y por lo tanto el crecimiento del sistema económico. Además, el aumento de la producción científica y de las innovaciones depende de un eficiente sistema nacional de innovación (Lundvall, 1992; Patel y Pavitt, 1994) adecuadamente financiado, y de la interacción de la triple hélice – universidad, gobierno y empresa – (Etzkowitz, 2006). Los gobiernos, por lo tanto, tienen que decidir cuánto y cómo invertir en I+D+i para fomentar la producción de innovaciones y la transferencia tecnológica.

Los excepcionales resultados económicos de Estados Unidos a partir del fin de la segunda guerra mundial hasta la era actual de la globalización, de la internacionalización de los mercados y de la revolución de las TICs, han llamado la atención de los economistas sobre el nivel y las modalidades de las inversiones en investigación y desarrollo que estimulan la innovación tecnológica. De hecho, las modernas teorías sobre el crecimiento económico (Lucas, 1988; Romer, 1990) apoyan la realización de importantes inversiones en investigación y formación, ya que aumentan el crecimiento económico a largo plazo. En particular, la nueva teoría del crecimiento, en la versión de Romer (1990), introduce en el modelo de Solow el cambio tecnológico endógeno como una función del nivel de capital humano.

Muchos han sido los estudios econométricos que han intentado constatar si la inversión pública en I+D+i es un complemento o un sustituto de la inversión privada (Blank y Stigler, 1957), pero a pesar de la extensa literatura científica en la materia, los resultados obtenidos han sido ambiguos. Conocer la relación entre la financiación pública y privada de la I+D+i es importante, puesto que en caso de complementariedad entre los dos tipos de inversiones, existe una justificación para subvencionar las inversiones privadas de I+D+i a costa de la colectividad que paga los impuestos. Si por el contrario, entre las inversiones públicas y privadas existiera un efecto sustitución, entonces resultaría muy difícil justificar dicho gasto a costa de la sociedad. Los que están a favor de la financiación pública de la I+D+i alegan que las ayudas públicas son imprescindibles, ya que, en caso contrario, las empresas dimensionarían a la baja su gasto en I+D+i. Por otro lado, los que se oponen se preguntan si dichas subvenciones se pueden realizar sin suponer al mismo tiempo importantes distorsiones de otros aspectos de la actividad económica -Griffith (2000)-. Así pues, es posible que las inversiones públicas en I+D+i expulsen a las privadas.

Puesto que los objetivos de los gobiernos son los de crear las condiciones para generar crecimiento económico tanto a través de la asignación de los recursos públicos de I+D+i, así como a través de políticas que estimulen la producción de investigación de base y aplicada en los centros de investigación públicos y privados, el análisis de estos aspectos no puede prescindir de una teoría de las elecciones en materia de inversiones públicas y privadas en I+D+i. Metcalfe (1994) identifica dos posibles razones de las inversiones públicas de I+D+i. La primera considera la investigación científica como un producto cultural de consumo que educa a la sociedad. La segunda razón a favor de la financiación pública de la I+D+i, se basa en una visión instrumental, según la cual la ciencia es una inversión que genera, a través del progreso técnico, un retorno compensatorio en términos de riqueza económica nacional y una mejora de la calidad de vida de la sociedad. La tasa social de rendimiento de la I+D+i es igual a la tasa de interés que la sociedad obtiene de dichas inversiones (Griliches, 1958). La financiación pública de la I+D+i puede tener un impacto directo en los resultados económicos de las empresas, y también puede contribuir indirectamente a estimular los gastos privados de I+D+i, generando así un beneficio social en forma de nuevos conocimientos y spillovers. Además, la racionalidad del agente público a la hora de invertir y subvencionar la I+D+i se justifica por la necesidad de corregir el fallo de mercado que se da en la producción del conocimiento científico y

tecnológico, como consecuencia de la apropiabilidad incompleta de los resultados por parte de las empresas que han invertido en investigación (Nelson, 1959; Arrow, 1962). Este planteamiento ha llevado a algunos países a fomentar las inversiones de I+D+i por parte de las empresas, para estimular las innovaciones tecnológicas que repercuten en el sistema económico, generando un superávit tanto para los productores como para los consumidores (David y Hall, 2000).

Las inversiones públicas en investigación pueden llegar a influir sobre las decisiones de las empresas a la hora de financiar la I+D+i privada y su comportamiento en los mercados. De hecho, el gasto público en investigación puede contribuir indirectamente a estimular los gastos privados en I+D+i a través del aumento del conocimiento y de la formación profesional (Klette et al. 2000). Jaffe (1989) y Adams (1990) documentan, a través de análisis econométricos, la existencia de spillovers generados por las inversiones públicas en conocimientos científicos sobre las inversiones privadas en I+D+i. Link y Scott (1998) muestran la influencia que tienen los programas de investigación financiados por los gobiernos sobre las innovaciones tecnológicas.

El debate sobre la complementariedad o sustituibilidad de las inversiones públicas y privadas en I+D+i ha impulsado la realización de muchos estudios tanto a nivel de empresa, (Higgins y Link, 1981; Link, 1982; Toivanen y Niininen, 1998; Wallsten, 1999; Duguet, 2003, Lööf y Heshmati, 2005), de sector, – como es el caso de los estudios de Levin y Reiss (1984), Lichtenberg (1984) – así como a nivel agregado. A pesar de la falta de consenso sobre los resultados, David y Hall (2000), al analizar algunos estudios micro y macroeconómicos, encuentran una preponderancia de complementariedad entre inversiones públicas y privadas en investigación.

Levy y Terleckyj (1983) llevan a cabo el primer estudio sectorial a nivel macro de las relaciones entre financiación pública y privada de la I+D+i. Los resultados más destacados son que las ayudas públicas a proyectos industriales están positiva y significativamente correlacionadas con las inversiones privadas en I+D+i y la productividad de las empresas. Lichtenberg (1987) mantiene que una intensidad más alta del gasto en I+D+i en los centros públicos de investigación no genera un impacto adicional sobre el gasto privado. Robson (1993) encuentra que las inversiones públicas en I+D+i tienen un impacto positivo. A conclusiones parecidas llega Diamond (1998). Kealey (1996) demuestra, a través del análisis de series históricas de inversiones en I+D+i para varios países, cómo inversiones públicas elevadas en investigación con respecto a las privadas reducen de forma más que proporcional el nivel total de las inversiones en I+D+i.

En este contexto nos parece interesante seguir profundizando en las tendencias de las inversiones públicas en I+D+i, puesto que unas eficaces políticas de financiación de la investigación y de gestión de los sistemas nacionales de innovación van a permitir mejorar las capacidades de innovación de las empresas, de las redes, de los sectores y del conjunto de la economía.

3. ANÁLISIS EMPÍRICO

Existen múltiples factores que determinan la evolución de la producción y renta de un país y el gasto en innovación es uno de ellos. A su vez, cabría tener presente que aún cuando se hable de este concepto de forma general, el mismo encierra una realidad con múltiples facetas, pues no todos los distintos usos dados a los gastos son idénticos ni sus consecuencias sobre el producto se manifiestan de manera semejante, ni en intensidad ni en el tiempo que tarda en hacerlo. Adicionalmente, hay que considerar que los resultados de estas inversiones se ven condicionados por la estructura y características del tejido productivo sobre el que se van a ejecutar. Por ello, cabría esperar que gastos semejantes consiguieran resultados diferentes en función de las condiciones imperantes en una u otra economía.

Trabajos como el de Aiginer y Falk (2005), tratan de encontrar una posible relación entre crecimiento del PIB per capita y gasto en I+D+i. La investigación aquí presentada incide en este aspecto desde otro enfoque, pues indaga sobre la posible existencia de relaciones causa-efecto entre gasto público en I+D+i y PIB, a través de la posible identificación de grupos homogéneos de países que responden a pautas de ejecución similares a la hora de invertir desde las instancias públicas parte de su riqueza en desarrollo técnico e innovación.

Este estudio trata de evaluar si los impactos que tienen las inversiones en I+D+i sobre el nivel de renta de un país (medido mediante el PIB) son similares entre naciones que comparten una zona con estructuras económicas que presentan una cierta vinculación y características semejantes. En concreto, se analiza la influencia que tiene el gasto en I+D+i sobre la renta en el conjunto de países formado por la UE-15 excepto Luxemburgo e incluyendo a Noruega. A partir de ahora, el distintivo de cada país incluido en el estudio dentro de los cuadros de resultados serán sus tres primeras letras en mayúsculas -ejemplo, España será ESP-, excepto la Gran Bretaña que se identificará por GBR. Como indicador de la inversión en innovación se han utilizado las cifras correspondientes a GBOARD, siglas de Goverment Budget Appropriations or outlays on Research and Development, que son los importes asignados dentro de los presupuestos de los gobiernos centrales o federales para este fin. Dentro de este concepto también se incluyen las partidas asignadas por los gobiernos provinciales y regionales, siempre que sean significativas. Las cifras incorporan tanto gastos corrientes como de capital y cubren todos los gastos de I+D financiados por los Gobiernos independientemente del agente que los haya ejecutado según el Manual Frascati (OCDE, 2002: 496). Las cifras de GBOARD las recopilan las autoridades nacionales a partir de los presupuestos públicos, los cuales son facilitados posteriormente a Eurostat. Se clasifican en función de objetivos socioeconómicos definidos en las NABS -Nomenclature for the analysis and comparison of scientific programmes and budgets, Eurostat, 1994-. Los objetivos son los siguientes:

Tabla 1: Identificación de las partidas NABS

Rúbrica	Código
Exploración y explotación del medio terrestre	1
Infraestructuras y ordenación del territorio	2
Control y protección el medio ambiente	3
Protección y promoción de la salud humana	4
Producción, distribución y utilización racional de la energía	5
Productividad y tecnologías agrarias	6
Productividad y tecnologías industriales	7
Vida en sociedad (Estructura y relaciones)	8
Exploración y explotación del espacio	9
Investigación financiada por fondos generales de universidades	10
Investigaciones no orientadas	11
Investigaciones no clasificadas	12
Defensa	13

Fuente: Eurostat

Estos datos están recogidos en euros corrientes por lo que se ha procedido a expresarlos en euros constantes de 2005. Para ello se han tomado las tasas de variación del índice de precios de consumo armonizado publicado por Eurostat, para el período incluido en el análisis. Para obtener la tasa media correspondiente al conjunto de los 15 países estudiados se ha procedido a ponderar la tasa de variación del IPC armonizado en cada país por el volumen de su PIB nominal. Esta última variable también se ha tomado de Eurostat (2005 y 2006).

Como se ha indicado anteriormente, el objetivo final del trabajo es la clasificación del conjunto de 15 países en varios grupos que permitan explicar la posición relativa de cada uno de ellos en relación la del resto. Para realizar esta tarea se ha procedido, en primer lugar, a estimar la relación entre el PIB y el gasto previsto en I+D+i, tanto para los 15 países considerados como para el conjunto formado por todos ellos, con cifras expresadas en euros constantes de

2005. A partir de los resultados de las estimaciones y del peso medio que dicho gasto tiene sobre el PIB durante el período considerado (1995-2006) se procede a realizar agrupaciones de países en función de su posición respecto a la media de ambas variables. Con los grupos así formados se procede a reestimar la relación entre el PIB y el gasto previsto en I+D+i mediante el uso de relaciones simultáneas tipo SURE. La tabla 2 recoge los resultados de la estimación individual en logaritmos neperianos, junto con el peso que esta partida tiene sobre el PIB de cada país.

Tabla 2: Resultados de la estimación individual

País	$\hat{\beta}$	test t	I+D+i / PIB
ALE	-0,2365	-1,25	0,81%
AUS	0,7205	7,14	0,64%
BEL	0,5367	8,49	0,57%
DIN	0,6399	5,51	0,73%
ESP	0,3880	35,16	0,65%
FIN	0,8151	7,85	1,02%
FRA	0,8132	3,09	0,99%
GRE	0,7222	6,22	0,28%
HOL	2,0450	6,05	0,77%
IRL	0,6910	8,66	0,34%
ITA	0,4071	7,64	0,64%
NOR	1,0445	8,55	0,74%
POR	0,3550	13,75	0,58%
GBR	1,0271	10,62	0,72%
SUE	0,0680	0,31	0,88%
EUR	0,9230	7,73	0,76%

Fuente: elaboración propia a partir de los datos de Eurostat

Habida cuenta de los resultados medios del conjunto -fila EUR- los grupos formados son los siguientes:

- peso del gasto y elasticidad superiores a la media de ambas variables: Holanda -a partir de ahora, Grupo 1-
- peso del gasto superior a la media pero con una elasticidad inferior a la del promedio: Alemania, Finlandia, Francia y Suecia -a partir de ahora, Grupo 2-
- peso del gasto inferior a la media pero elasticidad por encima del promedio: Noruega y Gran Bretaña -a partir de ahora, Grupo 3-
- peso y elasticidad por debajo de la media: el resto, es decir, Austria, Bélgica, Dinamarca, España, Grecia, Irlanda, Italia y Portugal -a partir de ahora, Grupo 4-.

A partir de esta clasificación se estima la relación por grupos mediante el uso del sistema de ecuaciones tipo SURE (siglas de Seemingly Unrelated Regression Estimation). La diferencia entre este sistema y la regresión tradicional es que en este caso cada ecuación se estima individualmente pero todas ellas se hallan vinculadas entre sí a través de los errores de regresión, ya que se permite que exista correlación contemporánea entre los residuos. Este tipo de estimación permite la agrupación de países según su comportamiento. Mediante esta técnica se estiman sistemas de ecuaciones en donde se supone que la matriz de varianzas y covarianzas se encuentra libre de heterogeneidad y autocorrrelación pero se verifica la existencia de correlación entre los errores, de forma que $\text{cov}(\varepsilon_{it}, \varepsilon_{jt}) = \sigma_{ij} \neq 0$. Este punto es crucial para la utilización de las estimaciones SURE por lo que es preciso contrastar tal hipótesis. Para ello se utilizará el test de independencia de Breusch-Pagan cuya hipótesis nula es precisamente la ausencia de relación entre los residuos de las regresiones estimadas. Obviamente, el primer

grupo no es susceptible de ser estimado mediante este sistema pues está formado únicamente por una nación, por lo que, en este caso, se mantendrá la expresión obtenida mediante la estimación individual.
Los resultados de esta fase de la estimación se recogen en la tabla 3.

Tabla 3: Estimaciones con regresiones SURE

Grupo 2:

País	$\hat{\alpha}$	test t	$\hat{\beta}$	test t	R^2
ALE	15,53	1,35	-0,9471	-,37	0,0937
FIN	6,50	13,61	0,7356	11,13	0,8645
FRA	5,72	4,41	0,8849	6,58	0,5093
SUE	11,03	13,83	0,1866	1,81	0,0220

Test Breusch-Pagan: chi2(6) = 18,489 p-valor = 0,0051

Grupo 3:

País	$\hat{\alpha}$	test t	$\hat{\beta}$	test t	R^2
NOR	4,40	5,19	1,0694	9,08	0,8903
GBR	4,98	5,75	0,9942	10,64	0,9252

Test Breusch-Pagan: chi2(1) = 4,817 p-valor = 0,0377

Grupo 4:

País	$\hat{\alpha}$	test t	$\hat{\beta}$	test t	R^2
AUS	7,23	17,74	0,7016	12,52	0,8488
BEL	8,51	25,50	0,5447	12,00	0,8899
DIN	7,29	19,69	0,6716	13,08	0,7695
ESP	10,22	145,59	0,3906	46,92	0,9928
GRE	7,49	15,89	0,7394	9,61	0,8117
IRL	7,39	21,90	0,7206	12,71	0,8914
ITA	10,36	39,73	0,4130	14,29	0,8670
POR	9,50	65,22	0,3492	15,95	0,9545

Test Breusch-Pagan: chi2(28) = 64,945 p-valor = 0,0001
Fuente: elaboración propia

En los tres grupos se rechaza la independencia de los residuos y, por tanto, es adecuado utilizar las regresiones SURE. A la vista de los resultados cabría preguntarse por la igualdad entre los parámetros estimados en cada grupo. Para realizar estos contrastes de hipótesis se utiliza el test de Wald. Así, dentro del grupo 2 no se puede aceptar la igualdad de parámetros en todos los países pero sí aparecen dos subgrupos con parámetros muy diferentes: Alemania y Suecia por un lado y Finlandia y Francia por otro. Haciendo los correspondientes test de Wald se obtienen los resultados que aparecen en la tabla 4 por lo que se pueden aceptar ambas hipótesis y, por tanto, los parámetros de Alemania y Suecia no se rechaza que sean iguales así como los de Francia y Finlandia.

Tabla 4: Contraste de igualdad entre elasticidades en grupo 2

Países	Número de restricciones	F (p-valor)
ALE = SUE = FRA = FIN	3	26,85 (0,0003)
ALE = SUE	1	3,20 (0,0821)
FRA = FIN	1	0,78 (0,3815)

Fuente: elaboración propia

Por su parte, dentro del grupo 3 la constatación de la igualdad de elasticidades entre Noruega y Gran Bretaña también se acepta, siendo los resultados obtenidos los que se recogen en la tabla 5.

Tabla 5: Contraste de igualdad entre elasticidades en grupo 3

Países	Número de restricciones	F (p-valor)
GBR = NOR	1	0,56 (0,4798)

Fuente: elaboración propia

Finalmente, en el grupo 4 se rechaza que todos los países tiene igual parámetro, contrastándose y no rechazándose la igualdad en las elasticidades en dos grupos de países: Austria, Bélgica, Dinamarca, Grecia e Irlanda por un lado y, por otro, España, Italia y Portugal. Los resultados de los test se recogen en la tabla 6.

Tabla 6: Contraste de igualdad entre elasticidades en grupo 3

Países	Número de restricciones	F (p-valor)
AUS=BEL=DIN=ESP=GRE=IRL=ITA=POR	7	45,67 (0,0000)
AUS=BEL=DIN= GRE=IRL	4	2,17 (0,0856)
ESP=ITA=POR	2	1,98 (0,1449)

Fuente: elaboración propia

4. CONCLUSIONES

En este trabajo se ha analizado la relación entre la inversión en I+D+i y la renta (utilizando el PIB) de los países de EU15 excepto Luxemburgo e incluyendo Noruega durante el período 1995-2006. Tal y como se recoge en los resultados del análisis empírico, se confirma que el impacto de la inversión en innovación sobre el PIB real es heterogéneo dentro del conjunto de países considerados. Existen cuatro grandes grupos distintos y dentro de ellos, a su vez, subgrupos de manera que los 15 países analizados forman seis agrupaciones cuyo comportamiento es diferente. En relación a la o las causas que justifiquen esta situación, cabría pensar en el efecto que tiene el hecho de tener una mayor o menor economía. Sin embargo, no parece adecuado pensar que dicha variable sea determinante a la hora de explicar los resultados obtenidos ya que si observamos el valor del parámetro estimado en los países con mayor economía (Alemania, Francia, Gran Bretaña, Italia y España) se aprecia que cada uno de ellos presenta un valor diferente y que a su vez, cada uno está encuadrado en grupos distintos. Alemania es un caso curioso, pues las estimaciones sugieren que su economía es absolutamente insensible a las modificaciones en gastos públicos de innovación. Sin embargo, es necesario recordar que, durante el periodo considerado y en términos reales, es el único país que ha experimentado un recorte en esta partida (caída del 10% en los once años).

A la vista de los resultados, parece claro que el tipo de estructura productiva incide en el impacto de dichos gastos sobre el PIB. De esta forma se puede entender la enorme sensibilidad de la economía holandesa o la reducida incidencia que se registra en los principales países mediterráneos. El caso de España es altamente significativo, pues nuestro país ha experimentado uno de los crecimientos del PIB más intensos durante el período de tiempo analizado. Sin embargo, es justo mencionar que tal avance se ha llevado a cabo mediante actividades intensivas en mano de obra con escasa cualificación -construcción y hostelería principalmente- por lo que la reducida sensibilidad al volumen de gasto público en I+D+i parece estar justificada.

En consecuencia, los resultados parecen sugerir que si se desea que la inversión en innovación conduzca a los resultados esperados, no basta solamente con realizar mayores desembolsos, sino que es necesario que vayan acompañados de otras medidas que modifiquen las condiciones y la naturaleza de la actividad productiva.

BIBLIOGRAFIA

ADAMS, J.D. (1990), "Fundamental Stocks of Knowledge and Productivity Growth", Journal of Political Economy, vol. 98, n. 4, p. 673702.

AIGINGER, K. Y FALK, M., (2005): "Explaining Differences in Economic Growth among OECD countries", *Empirica* (2005), 32; pp. 1943.

ARROW, K.J. (1962), "Economic Welfare and the Allocation of Resources for Invention", en NELSON, R. (ed.), The Rate and Direction of Inventive Activity: Economic and Social Factors, Princeton University Press, Princeton, 609625.

BARTELSMAN, E.J. (1990), "Federally Sponsored R&D and Productivity Growth", Federal Reserve Economics Discussion Paper, n. 121, Federal Reserve Board of Governors, Washington DC.

BARTELSMAN, E.J., DOMS, M. (2000), "Understanding Productivity: Lessons from Longitudinal Microdata", Journal of Economic Literature, vol. 38, n. 3, p. 569594.

BLANK, D.M., STIGLER, G.J. (1957), The Demand and Supply of Scientific Personnel, National Bureau of Economic Research, New York.

DAVID, P.A., HALL, B.H. (2000), "Heart of Darkness: Modellling PublicPrivate Interactions inside the R&D Black Box", Research Policy, vol. 29, n. 9, p. 11651183.

DIAMOND, A.M. (1998), "Does Federal Funding Crowd Out Private Funding of Science?", Presentation at the American Economics Association Meetings, Chicago.

DUGUET, E. (2003), "Are R&D Subsidies a Substitute or a Complement to Privately Funded R&D? Evidence from France Using Propensities Score Methods for NonExperimental Data", Cahiers de la Maison des Sciences Economiques, Working Paper Series, n. 75, Universidad de París I.

ETZKOWITZ, H. (2006), "The New Visible Hand: an Assisted Linear Model of Science and Innovation policy", Science and Public Policy, vol. 33, n. 5, p. 310320.

GRIFFITH, R. (2000), "How important is business R&D for economic growth and should government subsidize it?", Institute for Fiscal Studies, Briefing Notes 12.

GRILICHES, Z. (1995), "R&D and Productivity: Econometric Results and Measurement Issues", en STONEMAN, P. (ed.), "Handbook of Economics of Innovation and Technological Change, Basil Blackwell, Oxford, p. 5289.

GRILICHES, Z. (1958), "Research Costs and Social Returns: Hybrid Corn and Related Innovations", Journal of Political Economy, vol. 66, n. 5, p. 419431.

GUELLEC, D., VAN POTTELSBERGHE DE LA POTTERIE, B. (2001), "R&D and Productivity Growth: Panel Data Analysis of 16 OECD Countries", OECD Economic Studies, n. 33, p. 111136.

HALL, B.H. (1996), "The Private and Social Returns to Research and Development", en SMITH, B.L.R., BARFIELD, C. (ed.), "Technology, R&D, and the Economy", The Brookings and American Enterprise Institute, Washington DC., p. 289331.

HALL, B.H., MAIRESSE, J. (1995), "Exploring the Relationship between R&D and Productivity in French Manufacturing Firms", Journal Econometrics, vol. 65, n. 1, p. 263293.

HIGGINS, R.S., LINK, A.N. (1981), "Federal Support of Technological Growth in Industry: Some Evidence of Crowding Out", IEEE Transactions on Engineering Management, EM28, 8688.

JAFFE, A.B. (1989), "Real Effects of Academic Research", American Economic Review, vol. 79, n. 5, p. 957970.

KEALEY, T. (1996), "The Economic Laws of Scientific Research", McMillan Press, London.

KLETTE, T.J., MOEN, J., GRILICHES, Z. (2000), "Do Subsidies to Commercial R&D Reduce Market Failures" Microeconomic Evaluation Studies", Research Policy, vol. 29, n. 4, p. 471495.

LEVIN, R.C., REISS, P. (1984), "Test of a Schunpeterian Model of R&D and Market Structure", en GRILIECHES Z. (ed.), "R&D, Patents and Productivity", University of Chicago Press, Chicago.

LICHTENBERG, F.R. (1984), "The Relationship between Federal Contract R&D and Company R&D", American Economic Review Papers and Proceedings, vol. 74, n. 2, p.7378.

LICHTENBERG, F.R., SIEGEL, D. (1991), "The impact of R&D Investment on Productivity. New Evidence Using Linked R&DLRD Data", Economic Inquiry, vol. 29, n. 2, p. 203229.

LINK, A.N. (1982), "An Análisis of the Composition of R&D Spending", Southern Journal of Economics, vol. 49, n. 2, p. 342349.

LINK, A.N., SCOTT, J.T. (1998), "Public Accountability: Evaluating TechnologyBased Institutions," Kluwer Academic, Norwell MA.

LÖÖF, H., HESMATI, A. (2005), "The impact of Public Funds on Private R&D Investment: NEew Evidence from a Firm Level Innovation Study", MIT Discussion Papers, n. 3.

LUCAS, R.E. (1988), "On the Mechanics of Economic Development", Journal of Monetary Economics, vol. 22, n. 1, p. 342.

LUNDVALL, B.A. (1992), "National systems of innovation", Pinter, London.

MAIRESSE, J., SASSENOU, M. (1991), "R&D and Productivity: A Survey of Econometric Studies at the Firm Level", Science Technology and Industry Review, vol. 8, April, p. 945.

METCALFE, J.S. (1994), "Technology Systems and Technology Policy in an Evolutionary Framework", Cambridge Journal of Economics, vol. 19, 2546.

NELSON, R. (1959), "The Simple Economics of Basic Scientific Research", The Journal of Political Economy, vol. 67, n. 3, p. 322.

PATEL, P., PAVITT, K. (1994), "National Innovation Systems: Why They Are Important, and How They Might Be Measured And Compared", Economic Innovation New Technology, vol. 3, n. 1, p. 7795.

ROBSON, M. (1993), "Federal Funding and the Level of Private Expenditure on Basic Research", Southern Economic Journal, vol. 60, n. 1, p. 6371.

ROMER, P.M. (1990), "Endogenous technological change", Journal of Political Economy, vol. 98, n. 5, p. 71102.

TOIVANEN, O., NIININEN, P. (1998), "Investment, R&D, Subsidies and Credit Constraints", Working Paper, Department of Economics MIT and Helsinki School of Economics.

WALLSTEN, S.J. (1999), "Do GovernmentIndustry R&D Programs Increase Private R&D?: The Case of the Small Business Innovation Research Program", Department of Economics Working Paper, Stanford University.

ÁREA II

ECONOMÍA DE LA INFORMACIÓN Y DEL CONOCIMIENTO

ECONOMIA DA INFORMACÃO E DO CONHECIMENTO

Social Capital and Investment in R&D: New Externalities

Tiago Neves Sequeira* Alexandra Ferreira-Lopes†

Abstract

We introduce social capital in an endogenous growth model with physical capital, human capital and varieties, and we compare the market with the efficient solutions. As social capital is not tradable in the market and it favours research networks, it introduces new externalities in this framework. These externalities induce the market to invest less in social capital than a social planner would do and decrease the tendency to underinvestment in R&D. We quantify the distortions in the model. In some conditions, the new distortions are strong enough to overcome the usual result of underinvestment in R&D.

JEL Classification: O15, O31, O41.

Keywords: R&D, Social Capital, Human Capital, Endogenous Growth.

*UBI and INOVA, UNL. Corresponding Author: Tiago Neves Sequeira. Management and Economics Department. Universidade da Beira Interior. Estrada do Sineiro, 6200-209 Covilhã, Portugal. email: sequeira@ubi.pt. Tiago Neves Sequeira acknowledges the financial support from POCI/FCT.

†ISCTE (Lisbon University Institute), Economics Department, ERC-UNIDE, and DINÂMIA. Avenida das Forças Armadas, 1649-026 Lisboa, Portugal, phone: +351 217903901; Fax: +351 217903933; E-mail: alexandra.ferreira.lopes@iscte.pt.

1 Introduction

There are several recent articles that evaluate distortions in the decentralized market economy in models of endogenous economic growth. These models evaluate the differences that arise between the decentralized markets and the social planner choices due to several types of externalities: spillovers in the R&D process, duplications in the R&D process, gains from specialization due to new varieties, and erosion effects from R&D into human capital accumulation. However, social capital has not been considered in these models, despite its increasing interest in the economic growth literature. Social capital can introduce distortions in market allocations mainly due to two features of social capital: the failure of a market for social capital and the impact it can have in R&D due to research networks. The first reason is justified as firms do not pay for social capital when they contract workers; they pay for hours of work and at most for the level of qualifications. This may happen because the features usually included in social capital (confidence, truth, informal networks) are more difficult to evaluate and monitorize than degrees or years of schooling. The second reason is based on the importance of social networks between researchers in R&D productivity. Social networks can help knowledge sharing between researchers in an informal way, simply by proximity at work, "cheap talk" at lunch, etc. An example usually given is the importance of networks of researchers in Silicon Valley. Another can be the proximity of research staff in universities.

Social capital is a sociological concept that has been introduced recently in the economic growth literature. The definition of Putnam[1] refers to this concept as "features of social organization, such as trust, norms, and networks, that can improve the efficiency of society by facilitating coordinated actions". Most of the empirical literature has found a positive influence of social capital on economic growth, although it varies substantially (examples include Knack and Keefer[2]; Temple and Johnson[3], Whiteley[4], and Rupasingha *et al.*[5]). The introduction of social capital in growth models is still very scarce, but a good example is Beugelsdijk and Smulders[6], who also test the model against empirical data using the European Values Survey. Economic agents like to socialize (*bonding*), which they do by loosing consumption, since participation in social networks is time-consuming and reduces time available to work. Hence, higher levels of social capital may decrease economic growth. However, partic-

ipation in community networks (*bridging*) reduces the incentive for rent seeking and cheating, and so through this channel, higher level of social capital produces positive effects on economic growth.

The positive connection between social capital and human capital accumulation was first described in Coleman[7] and in Teachman *et al.*[8] in sociological studies about high school dropouts. Grafton *et al.*[9] test a theoretical growth model against empirical data to explain international country differences in productivity and find a positive impact of people's knowledge connections on productivity. Dinda[10] use an AK-type growth model to study the role of social capital in the production of human capital and in economic growth and compares theoretical results with empirical results finding a positive effect of social capital. Sequeira and Ferreira-Lopes[11] in an endogenous growth model framework also study the interactions between human and social capital and document the decline in social capital found by Putnam[12].

Literature about the connections of social capital, R&D, and economic growth is also very recent, scarce, and empirical. For example, Landry *et al.*[13], De Clercq and Dakhli[14] and Lee *at al.*[15] test empirically if there is indeed a connection and find positive effects of social capital in R&D and in innovation activities, although estimates vary widely.

No previous attempt that we know off brought the positive connection between social capital and R&D to an endogenous growth model. Our main contribution for the literature is to evaluate for the first time the impact of externalities caused by the presence of social capital in an endogenous growth model. Specifically, we want to determine the influence of social capital in the distortions between the decentralized equilibrium and the efficient solution. As already mentioned, social capital can introduce two types of new distortions in the market economy: first, because there is not a market for social capital, the social planner may choose higher levels of social capital than decentralized agents; second, because social capital influence R&D through research networks, which R&D firms do not consider, the social planner may also choose higher levels of social capital than decentralized agents. This is an important issue as opens the possibility of designing policies to enhance accumulation of social capital.

We also want to contribute to the discussion on "Too Much of a Good Thing?" or

on the optimality of R&D investments.[1] Thus, this paper is also inserted in the literature on the macroeconomic efficiency of R&D investments within endogenous growth models without scale effects, whose first contributions were Jones[16] and Jones and Williams[17]. The most common result in the literature tends to indicate that underinvestment in R&D happens in the real world (Romer[18]; Grossman and Helpman[19], and Aghion and Howitt[20] in endogenous growth models with scale effects and also Jones[16] and Jones and Williams[17] in models without scale effects, and Jones and Williams[21] in an empirical article). The exceptions are Stockey[22] and Benassy[23] that, in models with scale effects, discovered that for more general preferences or production, overinvestment in R&D can occur. Most recently, Reis and Sequeira[24] and Strulik[25] showed that overinvestment in R&D can be more plausible than has been thought before.

We build an increasing varieties model with different production sectors. A model with increasing varieties is appropriate as it always predicts more overinvestment in R&D than quality-ladders models or those that combine both increasing variety and quality-ladders. We argue that in this type of model the presence of social capital decreases the scope for underinvestment.

Section two presents the model and sections three and four present, respectively, optimal growth and the decentralized equilibrium. Section 5 compares the shares of human capital allocation in the social planner and in the decentralized equilibrium and discusses distortions in the decentralized equilibrium. In Section 6 we implement a calibration exercise to answer the question how much social capital influences the distortions between the efficient and the decentralized solution. Finally in section 7 we conclude.

2 Model

In this model we combine different types of capital: physical capital, human capital, social capital, and technological capital. Physical capital is used in the production of

[1] When applied to the economics of investment in R&D this expression has been first used by Jones and Williams (2000) as part of the article's title.

the final good. Human capital has different uses: it is employed in the production of differentiated goods, in schools, where it is the main input to new human capital; it is used in the accumulation of social capital, as suggested by previous literature, and is also used in the innovation process. Social capital is used in the production of the final good, in facilitating the accumulation of embodied knowledge (human capital), in facilitating the research networks that increase R&D productivity, and in its own accumulation. In assuming these interactions between different capital types, we based on different bodies of economic literature, revised above. A crucial feature in the model is that there is not a market for social capital. Social capital is produced because it makes families happier. This follows the notion of *bonding* in Beugelsdijk and Smulders (2004). However, firms (both firms in the final good and in the R&D market) benefit from social capital, which follow the notion of *bridging* in the same article. As firms benefit from social capital without paying for it, this carries out externalities with less social capital in the market than in the efficient solution. The distortions caused by social capital act on the opposite direction of gains from specialization and spillovers in the R&D process.

2.1 Production Factors and Final Goods

2.1.1 Capital Accumulation

The accumulation of physical capital (K_P), arises through production that is not consumed, and is subject to depreciation:

$$\dot{K}_P = Y - C - \delta_P K_P \tag{1}$$

where Y denotes production of final goods, C is consumption, and δ_P represents depreciation.

We propose that human capital K_H is produced using human capital allocated to

schooling as well as the total amount of social capital, Ks, according to:

$$\dot{K}_H = \xi H_H + \gamma K_S - \delta_H K_H \tag{2}$$

where H_H are school hours, $\xi > 0$ is a parameter that measures productivity inside schools, $\gamma \geq 0$ measures the sensitivity of human capital accumulation to the stock of social capital, and $\delta_H \geq 0$ is the depreciation of human capital. This expression captures the idea of Coleman[7] and Teachman *et al.*[8] according to which social capital is important to the production of human capital. It also ensures that human and social capital are substitutes in the production of human capital.

Individual human capital can be divided into skills in final good production (H_Y), school attendance (H_H), networking for social capital accumulation (H_S), and doing R&D (H_R). Assuming that the different human capital activities aren't done cumulatively, we have:

$$K_H = H_Y + H_H + H_S + H_R \tag{3}$$

Social capital accumulation requires human capital to be allocated to its production but at each point in time it will also depend on the current stock of social capital, i.e.:

$$\dot{K}_S = \omega H_S + \Omega K_S \tag{4}$$

where ω measures the productivity of human capital in the production of social capital and $\Omega \lesseqgtr 0$ measures the dynamic effect of social capital on its own production. If $\Omega > 0$ existing social networks are strong enough to keep growing without additional human capital.[2] If $\Omega < 0$, on the other hand, there is a net depreciation effect.

[2]Some types of social capital (such as cultural norms or values) are given by the family, which mean that people do not have to make efforts to acquire it. An alternative way of thinking about a positive Ω is that people with stronger social networks find it easier to continue improving networks than people with fewer.

2.1.2 R&D Technology

Technological capital, or new varieties, K_R, is produced in a R&D sector with human capital employed in R&D labs (H_R), by the stock of disembodied knowledge (K_R), and is also influenced by the stock of social capital:

$$\dot{K}_R = \varepsilon H_R^{\upsilon} K_R^{\phi} K_S^{\chi} \tag{5}$$

where $\varepsilon > 0$ measures the productivity of human capital allocated to the production of technological capital, υ measures duplication effects, $0 < \phi < 1$ measures the degree of spillover externalities in R&D across time, as in Jones[16], and $0<\chi < 1$ measures the positive effect of social networks in R&D productivity.[3]

2.1.3 Final Good Production

The final good is a differentiated one, produced with a Cobb-Douglas technology:

$$Y = D^{\beta} K_S^{\sigma} H_Y^{1-\beta-\sigma} K_R^{\eta},\ \beta \text{ and } \sigma < 1,\ \eta > 0 \tag{6}$$

D is an index of intermediate capital goods and is produced using the following Dixit-Stiglitz CES technology:

$$D = K_R \left[\frac{1}{K_R} \int_0^{K_R} x_i^{\alpha} \right]^{\frac{1}{\alpha}} \tag{7}$$

The elasticity of substitution between varieties is measured by $0 < \alpha < 1$. x_i is the intermediate capital good i and is produced in a differentiated goods sector using physical capital: $x_i = K_{P_{x_i}}$.

[3]Theoretically, one could generally assume that χ could assume negative values and values above unity. However, for simplification we assume that it varies between 0 (no effect of researchers networks) and 1 (proportional effect of researchers networks). In fact in the calibration exercise, we reasonably assume that the effect of researchers networks is smaller than the effect of past technological knowledge ($\chi < \phi$). If $\chi = 0$, we would obtain the formulation for R&D technology used in Jones (1995).

This means that (6) can be re-written as:

$$Y = K_R^{\eta} K_P^{\beta} K_S^{\sigma} H_Y^{1-\beta-\sigma} \tag{8}$$

2.2 Consumers

To capture the multi-faceted character of social capital the model for household preferences specifies it, along with consumption, as arguments of the intertemporal utility function:

$$U(C_t, K_{S_t}) = \frac{\tau}{\tau - 1} \int_0^{\infty} \left(C_t K_{St}^{\psi} \right)^{\frac{\tau-1}{\tau}} e^{-\rho t} dt \tag{9}$$

where ψ represents the preference for social capital and ρ is the utility discount rate.[4]

3 Optimal Growth

It is clear that when assets directly provide utility, while simultaneously acting as inputs to the production function, the decentralized equilibrium will in general not maximize aggregate welfare. Thus we must solve a social planner's problem. In this section we derive the conditions associated with the maximization of (9) subject to the production function (6) as well as the transition equations for the different types of capital (1), (2), (4), and (5).

The problem gives rise to the following Hamiltonian function:

$$\begin{aligned} \mathcal{H} \;=\; & \frac{\tau}{\tau - 1} \left(C K_S^{\psi} \right)^{\frac{\tau-1}{\tau}} + \lambda_P \left(K_R^{\eta} K_P^{\beta} K_S^{\sigma} H_Y^{1-\beta-\sigma} - C - \delta_P K_P \right) + \\ & + \lambda_H \left(\xi H_H + \gamma K_S - \delta_H K_H \right) + \lambda_S \left(\omega H_S + \Omega K_S \right) + \lambda_R (\varepsilon H_R^{\nu} K_R^{\phi} K_S^{\chi}) \end{aligned} \tag{10}$$

where the λ_j are the co-state variables for each stock K_j, with $j = P, H, S, R$. Considering choice variables C, H_Y, H_S, and H_R (and substituting H_H for $K_H - H_Y - H_S - H_R$

[4]The t subscripts are dropped in the remaining sections for ease of notation.

using (3)), the first order conditions yield:

$$\frac{\partial U}{\partial C} = \lambda_P \tag{11}$$

$$\lambda_H = \frac{\lambda_P \left(1-\beta-\sigma\right) Y}{\xi H_Y} \tag{12}$$

$$\lambda_H = \frac{\lambda_S \omega}{\xi} \tag{13}$$

$$\lambda_R = \frac{\xi \lambda_H}{\varepsilon \nu H_R^{\nu-1} K_R^{\phi} K_S^{\chi}} \tag{14}$$

as well as:

$$\frac{\dot{\lambda}_P}{\lambda_P} = \rho + \delta_P - \frac{\beta Y}{K_P} \tag{15}$$

$$\frac{\dot{\lambda}_H}{\lambda_H} = \rho + \delta_H - \xi \tag{16}$$

$$\dot{\lambda}_S = \rho\lambda_S - \left(\frac{\partial U}{\partial K_S} + \frac{\lambda_P \sigma Y}{K_S} + \lambda_H \gamma + \lambda_S \Omega + \lambda_R \varepsilon \chi H_R^{\nu} K_R^{\phi} K_S^{\chi-1}\right)$$

$$\dot{\lambda}_R = \rho\lambda_R - \left(\eta \frac{Y}{K_R}\right) \lambda_P - \lambda_R \varepsilon \phi H_R^{\nu} K_R^{\phi-1} K_S^{X} \tag{17}$$

with $\frac{\partial U}{\partial C} = C^{-\frac{1}{\tau}} K_S^{\psi\left(1-\frac{1}{\tau}\right)}$, $\frac{\partial U}{\partial K_S} = \psi C^{\left(1-\frac{1}{\tau}\right)} K_S^{\psi\left(1-\frac{1}{\tau}\right)-1}$ representing the marginal utilities of consumption and social capital respectively.

3.1 Optimal Growth Rates

Growth rates will, by definition, be constant, so equation (1) tells us that K_P, Y, and C all grow at the same rate. Furthermore, K_S and K_H components will also be growing at that same rate, respecting equations (2) to (4).

Denote the growth rate of technological capital as g_{K_R} and the growth rate of human capital as g_{K_H}. From equation (5) we can see that these two growth rates have to respect this relation: $g_{K_H} = \frac{(1-\phi)}{\chi+\nu} g_{K_R}$.

In the steady-state, we can obtain the human capital growth rate as follows. From (12) we find $g_{\lambda_H} = g_{\lambda_P} + g_Y - g_{K_H}$ and using equation (16) we can then replace the previous two equations in $-\frac{1}{\tau} g_Y + \psi \left(1 - \frac{1}{\tau}\right) g_{K_H} = \frac{\dot{\lambda}_P}{\lambda_P}$, which we calculated from (11). Then, using equations (5) and (8) we get:

$$g^*_{K_H} = \frac{\xi - \delta_H - \rho}{\left(\frac{\eta\left(\frac{\chi+\nu}{1-\phi}\right)}{1-\beta} + \psi\right)\left(\frac{1}{\tau} - 1\right) + \frac{1}{\tau}} \tag{18}$$

Using the fact that $g_{K_H} = \frac{(1-\phi)}{\chi+\nu} g_{K_R}$ we solve for the growth rate of technological capital:

$$g^*_{K_R} = \frac{\frac{\chi+\nu}{(1-\phi)}(\xi - \delta_H - \rho)}{\left(\frac{\eta\left(\frac{\nu+}{1-\phi}\right)}{1-\beta} + \psi\right)\left(\frac{1}{\tau} - 1\right) + \frac{1}{\tau}} \tag{19}$$

From (8) and $g_{K_H} = \frac{(1-\phi)}{(\chi+\nu)} g_{K_R}$ we find $g_Y = g_{K_H}\left(\frac{\eta\left(\frac{\chi+\nu}{1-\phi}\right)}{1-\beta} + 1\right)$. By substituting (18) in the previous equality we find:

$$g^*_Y = \frac{(\xi - \delta_H - \rho)\left(\frac{\eta\left(\frac{\chi+\nu}{1-\phi}\right)}{1-\beta} + 1\right)}{\left(\frac{\eta\left(\frac{\chi+\nu}{1-\phi}\right)}{1-\beta} + \psi\right)\left(\frac{1}{\tau} - 1\right) + \frac{1}{\tau}} \tag{20}$$

While the impact of the social capital share (σ) is positive in growth rates, the impact of preference for social capital (ψ) is negative as it has a trade-off with consumption. This has a parallel with the effects of *bonding* and *bridging* in growth rates in the article from Beugelsdijk and Smulders[6]. Optimal growth rates depend on parameters of the model as usual in non-scale models of endogenous growth.

4 Decentralized Equilibrium

In the decentralized equilibrium both consumers and firms make choices that maximize, respectively, their own felicity or profits. Consumers maximize their intertemporal utility function:

$$\frac{\tau}{\tau-1}\int_0^{\infty}\left(C_t K_{St}^{\psi}\right)^{\frac{\tau-1}{\tau}} e^{-\rho t}dt$$

subject to the budget constraint:

$$\dot{a} = (r-\delta_p)a + W_H\left(K_H - H_H - H_S - H_R\right) - C \tag{21}$$

where a represents the family physical assets, r is the return on physical capital, and W_H is the market wage. The market price for the consumption good is normalized to 1. Since it is making an intertemporal choice, the family also takes into account equations (2) and (4), which represent human and social capital accumulation, respectively.[5]

The markets for purchased production factors are assumed to be competitive. However, we assume that the firm cannot buy social capital, as there is, in effect, no market for it. Social capital is treated here as exogenous, although it affects the firm's production. Hence, consumer decisions will carry social capital externalities.

From this problem we know that returns on production are as follows:

$$W_H = \frac{(1-\beta-\sigma)\,Y}{H_Y} \tag{22}$$

$$p_D = \frac{\beta Y}{D}, \tag{23}$$

where p_D represents the price for the index of intermediate capital goods.

Each firm in the intermediate goods sector owns an infinitely-lived patent for selling its variety x_i. Producers of differentiated goods act under monopolistic competition

[5]See Appendix A and B respectively for the FOC and growth rates in the decentralized equilibrium.

in which they sell their own variety of the intermediate capital good x_i and maximize operating profits, π_i:

$$\pi_i = (p_i - r)x_{i,} \tag{24}$$

where p_i denotes the price of intermediate good i and r is the unit cost of x_i. The demand for each intermediate good results from the maximization of profits in the final goods sector. Profit maximization in this sector implies that each firm charges a price of:

$$p_i = p = r/\alpha. \tag{25}$$

With identical technologies and symmetric demand, the quantity supplied is the same for all goods, $x_i = x$. Hence, equation (7) can be written as:

$$D = K_R x. \tag{26}$$

From $p_D D = pxK_R$, together with (23) and (25), we obtain:

$$xK_R = K_P = \frac{\alpha\beta Y}{r}. \tag{27}$$

After insertion of equations (25) and (27) into (24), profits can be rewritten as:

$$\pi = (1 - \alpha)\beta Y/K_R \tag{28}$$

Let υ denote the value of an innovation, defined by:

$$\nu_t = \int_t^\infty e^{-[R(\tau)-R(t)]}\pi(\tau)d\tau, \text{ where } R(\tau) = \int_0^t r(\tau)d\tau. \tag{29}$$

Taking into account the cost of an innovation as determined by eq.(5), free-entry in

R&D implies that,

$$wH_R = \nu\varepsilon H_R^{\upsilon} K_R^{\phi} K_S^{\chi} \text{ if } \dot{K}_R > 0 \quad (H_R > 0); \tag{30}$$

$$wH_R > \nu\varepsilon H_R^{\upsilon} K_R^{\phi} K_S^{\chi} \text{ if } \dot{K}_R = 0 \quad (H_R = 0). \tag{31}$$

Finally, the no-arbitrage condition requires that investing in patents has the same return as investing in bonds:

$$\frac{\dot{\nu}}{\nu} = (r - \delta_P) - \pi/\nu. \tag{32}$$

This fully describes the economy.

5 Optimality of Human Capital Allocations

Using the FOC obtained from the social planner solution (11) to (17) and the equations that describe the evolution of the four capital stocks (1), (2), (4), and (5), it is possible to obtain the shares of human capital allocated to the different sectors in the economy (final good, human capital, social capital, and R&D). The share of human capital allocated to the different sectors are:

$$u_Y^* = \frac{H_Y}{K_H} = \frac{\left[\xi - \delta_H - \frac{\omega\gamma}{\xi} - \Omega\right]\left(\frac{K_S}{K_H}\right)^* - \frac{\omega\chi}{\nu}u_R^*}{\frac{\omega}{(1-\beta-\sigma)}\left[\psi\frac{C}{Y} + \sigma\right]} \tag{33}$$

$$u_S^* = \frac{H_S}{K_H} = \left(\frac{g_{K_S}^* - \Omega}{\omega}\right)\left(\frac{K_S}{K_H}\right)^* \tag{34}$$

$$u_H^* = \frac{H_H}{K_H} = \frac{1}{\xi}\left(g_{K_H}^* + \delta_H\right) - \frac{\gamma}{\xi}\left(\frac{K_S}{K_H}\right)^* \tag{35}$$

$$u_R^* = \frac{H_R}{K_H} = \frac{\frac{\eta\nu}{(1-\beta-\sigma)}g_{K_R}^*}{\xi - \delta_H + g_{K_R}^*\left[(\upsilon - 1 + \chi)\left(\frac{1-\phi}{\chi+\nu}\right)\right]}u_Y^* \tag{36}$$

The equations that were presented in this section provide a basis for a complete analysis of all the relationships between the different capital stocks and also a basis for the comparison with the decentralized equilibrium solution. Using the restriction that $u_Y^* + u_S^* + u_H^* + u_R^* = 1$, we obtain the social to human capital ratio:

$$\left(\frac{K_S}{K_H}\right)^* = \frac{1 - \frac{1}{\xi}\left(g_{K_H}^{DE} + \delta_H\right)}{\Phi^* + \frac{g_{K_S}^{DE} - \Omega}{\omega} - \frac{\gamma}{\xi}} \tag{37}$$

where

$$\Phi^* = \frac{\frac{\left[\xi - \delta_H - \frac{\omega\gamma}{\xi} - \Omega\right]}{\left[\frac{\omega}{(1-\beta-\sigma)}\left[\psi\frac{C}{Y} + \sigma\right]\right]}}{1 + \frac{\frac{\omega\chi}{\nu}}{\frac{\omega}{(1-\beta-\sigma)}\left[\psi\frac{C}{Y} + \sigma\right]} \frac{\frac{\eta\nu}{(1-\beta-\sigma)}}{\xi - \delta_H + g_{K_R}^*\left[(v-1+\chi)\left(\frac{1-\phi}{\chi+\nu}\right)\right]} g_{K_R}^*} \left(1 + \frac{\frac{\eta\nu}{(1-\beta-\sigma)}}{\xi - \delta_H + g_{K_R}^*\left[(v-1+\chi)\left(\frac{1-\phi}{\chi+\nu}\right)\right]} g_{K_R}^*\right). \tag{38}$$

Using the FOC obtained for the decentralized equilibrium solution, equations (45) to (50), the equations that describe the evolution of the four capital stocks (1), (2), (4), and (5), and also equations (22), (28), (30), and (32), it is possible to obtain the shares of human capital allocated to the different sectors in the economy: final good, human capital, social capital, and R&D. The share of human capital allocated to the different sectors in the decentralized equilibrium are:

$$u_Y^{DE} = \frac{H_Y}{K_H} = \frac{\xi - \delta_H - \frac{\omega\gamma}{\xi} - \Omega}{\frac{\omega}{(1-\beta-\sigma)}\left[\psi\frac{C}{Y}\right]} \left(\frac{K_S}{K_H}\right)^{DE} \tag{39}$$

$$u_S^{DE} = \frac{H_S}{K_H} = \left(\frac{g_{K_S}^{DE} - \Omega}{\omega}\right)\left(\frac{K_S}{K_H}\right)^{DE} \tag{40}$$

$$u_H^{DE} = \frac{H_H}{K_H} = \frac{1}{\xi}\left(g_{K_H}^{DE} + \delta_H\right) - \frac{\gamma}{\xi}\left(\frac{K_S}{K_H}\right)^{DE} \tag{41}$$

$$u_R^{DE} = \frac{H_R}{K_H} = \frac{\frac{\beta(1-\alpha)}{(1-\beta-\sigma)} g_{K_R}^{DE}}{\xi - \delta_H + g_{K_R}^{DE}\left[\left(\frac{1-\phi}{\chi+\nu}\right)(\chi + v - 1) + \phi\right]} u_Y^{DE} \tag{42}$$

Using the restriction that $u_Y^{DE} + u_S^{DE} + u_H^{DE} + u_R^{DE} = 1$, we obtain the social to

human capital ratio:

$$\left(\frac{K_S}{K_H}\right)^{DE} = \frac{1-\frac{1}{\xi}\left(g_{K_H}^{DE}+\delta_H\right)}{\Phi^{DE}+\frac{g_{K_S}^{DE}-\Omega}{\omega}-\frac{\gamma}{\xi}} \tag{43}$$

where

$$\Phi^{DE} = \frac{\left[\xi-\delta_H-\frac{\omega\gamma}{\xi}-\Omega\right]}{\left[\frac{\omega}{(1-\beta-\sigma)}\left[\psi\frac{C}{Y}\right]\right]}\left(1+\frac{\frac{\beta(1-\alpha)}{(1-\beta-\sigma)}}{\xi-\delta_H+g_{K_R}^{*}\left[(\upsilon-1+\chi)\left(\frac{1-\phi}{\chi+\nu}\right)+\phi\right]}g_{K_R}^{*}\right). \tag{44}$$

As growth rates and the consumption to output ratio are equal in the social planner and decentralized equilibrium solutions, the differences from the two solutions are spillovers in R&D, duplication effect in R&D, the specialization gains, and the externalities from social capital. As in Alvarez-Palaez and Groth[26], the social gains from specialization (η) compare with the private gain from an innovation ($\beta(1-\alpha)$).[6] From the comparison of (33)-(36) to (39)-(42) taking into account the comparison between (44) and (38), it is possible to write the following proposition.

Proposition 1 *The Decentralized Equilibrium yields a sub-optimal or over-optimal, respectively, social capital to human capital ratio and R&D effort, depending on the opposite effects of the following externalities:*

(i) the social capital externalities (σ and χ) that increases the social to human capital ratio in the social planner solution, increases the human capital allocated to social capital production, and decreases human capital allocated to 'schools', final good, and R&D sectors;

(ii) the spillover externality (ϕ) that decreases human capital allocated to R&D in

[6]See Appendix B for proof.

the market, increases the social to human capital ratio, and then increases allocations to the final good and to social capital production but decreases the allocation to the education sector in the market;

(iii) the duplication externality (υ) that decreases human capital allocated to R&D in the planner's solution, increases the social to human capital ratio, and then increases allocations to the final good and to the social capital production but decreases allocation to the education sector in the planner's solution;

(iv) the social gain from specialization (η) that decreases social to human capital ratio in the social planner's solution, decreases human capital allocated to the final good and to the social capital production, and increases allocation to all the other sectors in the economy.

Thus, relatively high social capital shares in the final good production and R&D technology may contribute to decrease the under-investment in R&D. Spillovers and social gains from specialization acts in favor of underinvestment and duplication and social capital externalities act in favor of overinvestment.

As usual in the studies that intend to evaluate distortions between the social planner and decentralized equilibrium solutions, this evaluation is a quantitative issue. Thus we now implement a calibration exercise to evaluate the distortions.

6 Results and Calibration

6.1 Calibration Procedure

It is not easy to take a model with social capital to data, as research dealing with social capital is still scarce. Some parameters on our model are quite standard in the literature: the intertemporal substitution parameter ($\tau = 0.5$), the intertemporal discount factor ($\rho = 0.02$), the share of physical capital in income ($\beta = 0.36$), the markup ($1/\alpha = 1.33$), and the productivity of R&D ($\epsilon = 0.1$), hence we are not discussing them.[7] For others, there are a range of plausible values: the depreciation rates (δ_K, δ_H), the productivity of human capital accumulation (ξ), the contribution of social capital to economic growth (σ), spillovers (ϕ) and duplication effects (υ). For these values we discuss our options. For other parameters there is greater uncertainty. For γ and ω, we concluded that changes in them are not crucial for the distortion evaluation. Thus we had fixed a value of 0.01 for them. For the preference for social capital (ψ) we have tested different values and concluded that values higher than 0 and less than 1 (i.e., consumers prefer consumption to social capital, which seems reasonable) do not change conclusions on distortions. Thus we have chosen an intermediate value of $\psi = 0.5$. For the spillover and duplication effects we have chosen $\phi = 0.4$ and $\upsilon = 0.5$ as reasonable values pointed out by the literature.[8] For the externality of social capital in the R&D technology we have picked an half of the spillover value, reasonably assuming that the effect of research networks in R&D is much lower than the 'standing on shoulders' effect.

For the depreciation of physical capital, it is clear from the observation of equations in the social planner solution and in decentralized equilibrium that it does not enter in the determination of distortion. Thus, we set $\delta_K = 0$. In fact, in previous articles that simultaneously consider human and physical capital accumulation, a zero depreciation is considered. We consider $\delta_H = 0$ and $\delta_H = 0.05$ in different exercises. For the parameter Ω that can measure a positive effect of social capital in its accumulation

[7] For the markup value, we use a median value from Norrbin[27].

[8] See Reis and Sequeira[24] for a discussion about the value of ϕ in models with human capital accumulation and Pessoa[30] for estimations of ϕ and υ.

or a depreciation of social capital, we use alternatively 0.01 and -0.01. For each of these exercises, we set the steady-state economic growth rate to be equal to 1.85%, which gives us a value for ξ. This procedure yields us values in the range used by human capital literature (e.g. Funke and Strulik[28]). For the impact of social capital in economic growth, we use a lower bound estimate for σ of 0.08 as in Knack and Keefer[2], for whom a 10% increase in trust implied a 0.8% increase in the economic growth rate. We also use a high bound for $\sigma = 1 - \beta - \sigma = 0.32$, suggested by the evidence in Whiteley[4] that points out an effect of social capital as big as the effect of human capital and approximating the evidence on World Bank [29] that pointed out a share of 0.78 to intangible capital, which included both human and social capital. For the consumption-output ratio (C/Y) we use the reasonable value of 0.6.

Table 1 below summarizes the calibration values. We use two sets of calibration values. The first, which we call benchmark, show the distortion caused by social capital in the absence of any other distortion (in this calibration there are not spillovers, duplication effects, specialization gains). In the second (designated by "Reasonable"), we set spillovers to 0.4 (see Reis and Sequeira[24]), duplication to 0.5 (see Pessoa[30]), specialization gains to 0.196 (see Jones and Williams[17]) and the social capital share to a value between 0.08 and 0.32 (see above).

Table 1: Calibration Values

parameter	Benchmark	Reasonable	parameter	Benchmark	Reasonable
τ	2	2	γ, ω	0.01	0.01
β	0.36	0.36	σ	varies	0.08 or 0.32
ρ	0.02	0.02	ϕ	0	0.4
α	0.75	0.75	υ	1	0.5
δ_H	0	0 or 0.05	η	$\beta(1-\alpha)$	0.196
δ_K	0	0	C/Y	0.6	0.6
ξ	depends on g_Y; $[0.05; 0.1]$		χ	varies	0.2
g_Y	0.0185	0.0185	Ω	0.01	0.01 or -0.01

6.2 Distortions from Social Capital

In the next lines we present the differences between the decentralized equilibrium and the optimal solution when there are only distortions from social capital. Thus, we apply benchmark calibration values. In Figure 1 we present the evolution of the K_S/K_H ratio through different values for the share of social capital in production. In Figure 2 we show the evolution of the allocation of human capital to the social capital

sector. In Figure 3, we show the evolution of the allocation of human capital to the R&D sector through different values of the same share. Then we also present three figures with the evolution of the ratio K_S/K_H, the share u_S, and the share u_R through different values for the externality of social capital in the R&D technology.

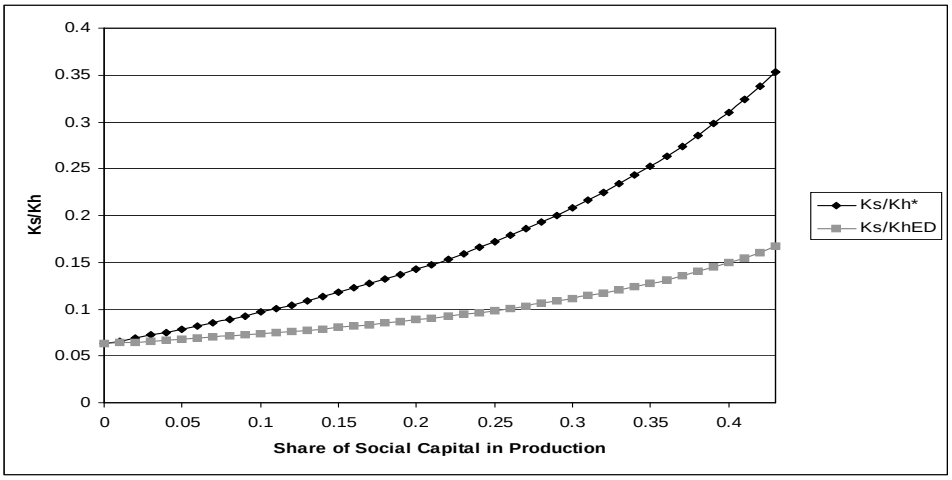


Figure 1 - Comparison between $(K_S/K_H)^{DE}$ and $(K_S/K_H)^*$ for different values of the share of social capital in production

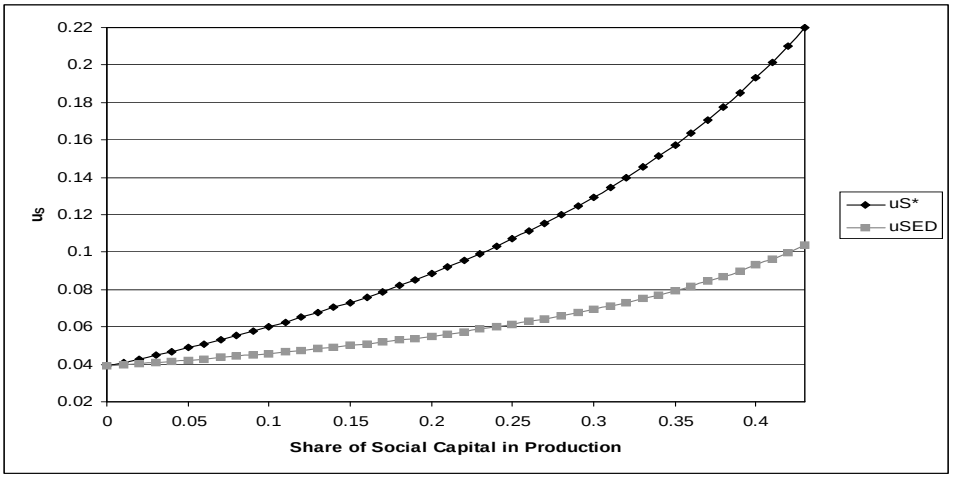


Figure 2 - Comparison between u_S^{DE} and u_S^* for different values of the share of social capital in production

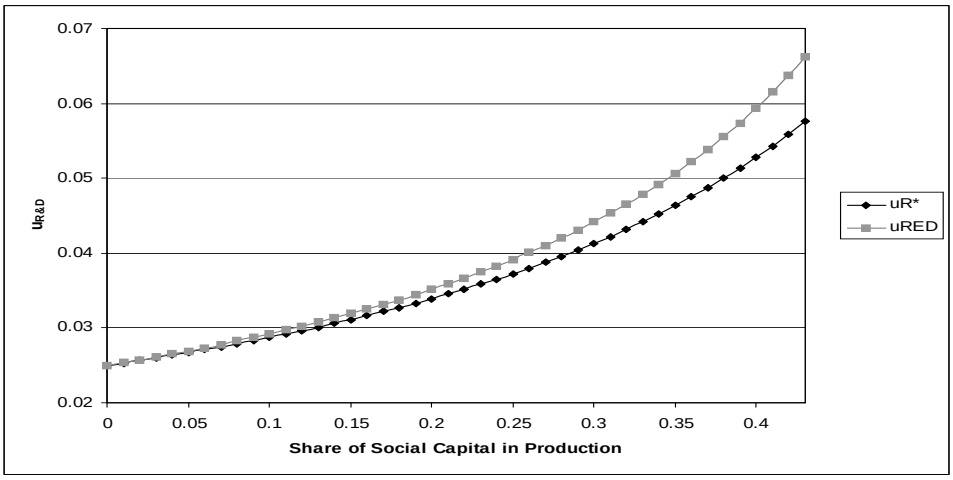


Figure 3 - Comparison between u_R^{DE} and u_R^* for different values of the share of social capital in production

From Figures 1 to 3 we can note that increasing the value of the share of social capital in the final good production increases the distortions in the decentralized economy, increasing the differences between the social desired ratio of social to human capital and the ratio obtained by the decentralized action of different agents (Figure 1). This distortion clearly causes underallocation of human capital to the social capital sector (Figure 2) and overinvestment in R&D as can be seen from Figure 3. While the ratio K_S/K_H rises 7 times until $\sigma = 0.43$ in the optimal solution, it only rises 3 times in the decentralized economy. The difference between the efficient allocation to the R&D sector and the market allocation can rise up to 1%, while the difference between efficient allocation to social capital and the market allocation can go up to 14%.

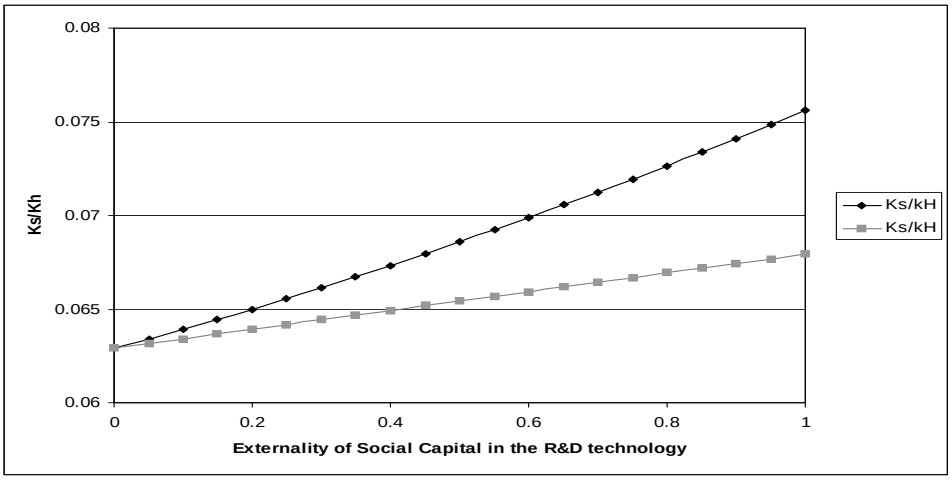


Figure 4 - Comparison between $(K_S/K_H)^{DE}$ and $(K_S/K_H)^*$ for different values of the share of social capital in $R\&D$

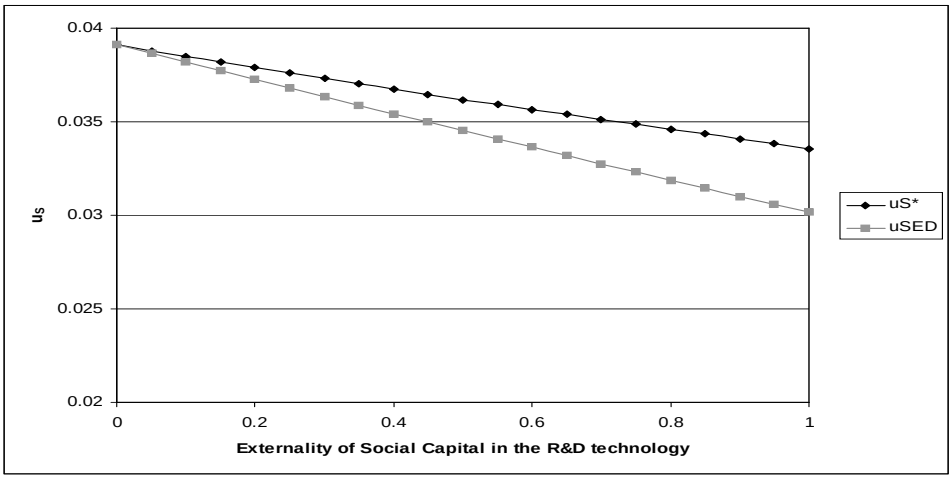


Figure 5 - Comparison between $u_{S^{DE}}$ and u_S^* for different values of the share of social capital in $R\&D$

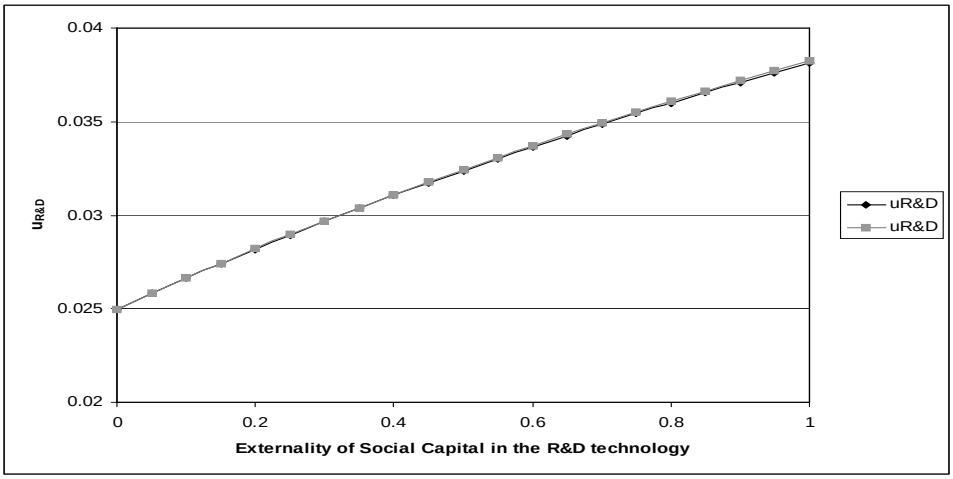


Figure 6 - Comparison between u_R^{DE} and u_R^* for different values of the share of social capital in $R\&D$

From Figures 4 to 6 we can see that increasing the effect of social capital in R&D technology also increases distortions but less than the rise in distortions caused by the final good social capital share. In this case a change from $\chi = 0$ to $\chi = 1$ implies a rise in K_S/K_H by 0.01, the changed caused in u_S is about 0.4% and finally the change implied in u_R is near 0.01%. We have also a tendency for underinvestment in social capital and overinvestment to R&D.

The next step is to evaluate how much decrease in the distortion can social capital impose when all other distortions are present.

6.3 Taking all Distortions Together

In this section, we present results from quantitative exercises in which we apply the calibration values depicted as 'Reasonable' in Table 1. Tables 2 to 4 compare the social planner allocations to the decentralized equilibrium ones. We show three different exercises: the first eliminates the distortions due to social capital and sets other distortions at reasonable levels, given by parameter values discussed above; the second considers a lower limit for the social capital share in production and a reasonable value for the externality of social capital in the R&D technology and keeping other values for parameters at the level used in the first exercise; the third exercise is equal to the second, except for the share of social capital in production that increases to 0.32. The unique difference in Table 3 (from Table 2) is that it uses a depreciation for human capital of $\delta_H = 0.05$. The unique difference in Table 4 (also from Table 2) is to consider a depreciation of social capital ($\Omega = -0.01$). In all these exercises $g_{K_H} = 1.05\%$ and $g_{K_R} = 2.62\%$.

Table 2: Results from Reasonable Calibrations ($\delta_H = 0$, $\Omega = 0.01$)

σ	$\sigma = 0$			$\sigma = 0.08$			$\sigma = 0.32$		
χ	$\chi = 0$			$\chi = 0.2$			$\chi = 0.2$		
	SP	DE	SP/DE	SP	DE	SP/DE	SP	DE	SP/DE
K_S/K_H	0.082	0.084	0.97	0.127	0.097	1.32	0.346	0.163	2.12
u_Y	74.08%	76.10%	0.97	75.16%	76.44%	0.98	73.04%	73.55%	0.99
u_S	0.39%	0.40%	0.97	0.61%	0.46%	1.32	1.65%	0.78%	2.12
u_H	17.81%	17.77%	1.00	16.98%	17.54%	0.97	12.95%	16.32%	0.79
u_R	7.72%	5.72%	1.35	7.26%	5.56%	1.31	12.34%	9.35%	1.32

Table 3: Results from Reasonable Calibrations ($\delta_H = 0.05$, $\Omega = 0.01$)

σ	$\sigma = 0$			$\sigma = 0.08$			$\sigma = 0.32$		
χ	$\chi = 0$			$\chi = 0.2$			$\chi = 0.2$		
	SP	DE	SP/DE	SP	DE	SP/DE	SP	DE	SP/DE
K_S/K_H	0.041	0.043	0.97	0.064	0.048	1.31	0.170	0.081	2.08
u_Y	38.19%	39.22%	0.97	38.55%	39.35%	0.98	36.60%	37.58%	0.97
u_S	0.20%	0.20%	0.97	0.31%	0.23%	1.31	0.81%	0.39%	2.08
u_H	57.63%	57.62%	1.00	57.42%	57.56%	0.99	56.40%	57.25%	0.99
u_R	3.98%	2.95%	1.35	3.72%	2.86%	1.30	6.18%	4.78%	1.29

Table 4: Results from Reasonable Calibrations ($\delta_H = 0.05$, $\Omega = -0.01$)

σ	$\sigma = 0$			$\sigma = 0.08$			$\sigma = 0.32$		
χ	$\chi = 0$			$\chi = 0.2$			$\chi = 0.2$		
	SP	DE	SP/DE	SP	DE	SP/DE	SP	DE	SP/DE
K_S/K_H	0.049	0.050	0.98	0.071	0.056	1.26	0.147	0.086	1.70
u_Y	64.84%	66.38%	0.98	61.55%	65.45%	0.94	45.63%	57.33%	0.80
u_S	9.98%	10.22%	0.98	14.49%	11.51%	1.26	30.03%	17.65%	1.70
u_H	18.43%	18.41%	1.00	18.02%	18.29%	0.99	16.62%	17.74%	0.94
u_R	6.76%	4.99%	1.35	5.94%	4.75%	1.25	7.71%	7.29%	1.05

Without the distortions introduced in this article ($\sigma = 0$, $\chi = 0$), we note that the tendency for underinvestment in R&D is high, as previous literature also predicted (see e.g. Jones and Williams[17]). Human capital allocation in the decentralized economy is almost at the optimal level and there are over-allocations to the final good production and to the social capital sector. It is worth noting that due to the distortions from social gains from specialization, spillovers, and duplication, there is a relatively higher social capital to human capital ratio in the market economy when compared to the social planner choice (0.084 compared with 0.082 in Table 2; 0.043 compared with 0.041 in Table 3, and 0.05 compared with 0.049 in Table 4).

When we consider positive values for σ and χ we note that now the social to human capital ratio is higher in the efficient solution than in the market economy, which is due to the distortions introduced in this article. The absence of a market for social capital is responsible for having relatively lower social capital in the market than in the case in which social welfare would be taken into account. This is reflected in the allocation of human capital to social capital production, which should also be higher than it is in the market economy. Because of that, allocation of human capital to the human capital accumulation sector is above the optimal level and the level of underinvestment in R&D is reduced. In the third exercise, the distortion in the social capital sector is so high that the social planner would allocate to that sector nearly twice the human capital allocated by the market economy (from 0.78% to 1.65% in Table 2; from 0.39% to 0.81% in Table 3, and from 17.65% to 30.03% in Table 4). Thus in these scenarios the social planner would reallocate human capital from final good production and schools to social capital accumulation sectors and to R&D firms. This means that some policies can be designed to enhance the production of social capital. Considering a positive depreciation for human capital as we do in Table 3, introduces almost no differences in distortions from the social planner allocations.

Such a high depreciation in human capital predicts however an implausible share of human capital allocated to the human capital accumulation sector. Nevertheless, when we introduce a depreciation in social capital accumulation, as in Table 4, we note some important differences. The share of human capital allocated to social capital production sector is higher than in previous exercises because human capital allocated to that sector must compensate the depreciation effect, while in previous exercises this did not happen because social capital could grow by itself (exogenously). The most important implication is that underinvestment in R&D is much reduced (from $u_R^*/u_R = 1.35$ to $u_R^*/u_R = 1.05$), opening the possibility to overinvestment in R&D. The threshold level for the share of social capital in production (σ) above which there is overinvestment in R&D is 0.372, which is in the range of plausible values, according to World Bank[29]. The higher the depreciation for social capital, the lower the threshold value for its share in production above which overinvestment in R&D occurs.

7 Conclusion

We motivate this article from the increasing importance of social capital in the economic growth analysis, both in theoretical and empirical perspectives. The interaction between social capital and R&D has been pointed out essentially due to research networks. We build the production side of the model taking into account the interactions between the different types of capital that have been pointed out in previous literature. In particular, we note the importance of the use of human capital in the social capital accumulation and the importance of this last factor in the production of the final good, and also in the discovery of new ideas.

Social capital can introduce distortions in market allocations mainly due to two features of social capital: the failure of a market for social capital and the impact it can have in R&D due to research networks. The first reason is justified as firms do not pay for social capital when contract workers; they pay for hours of work and at most for the level of qualifications. This may happen because the features usually classified as social capital (confidence, truth, networks) are more difficult to evaluate and monitorize than degrees or years of schooling. The second reason is based on

the importance of social networks between researchers in the R&D productivity. This means that social capital is produced because agents like to socialize but it is used in production side in a way that enhance firms productivity without any payment for it.

In the model, we also consider the most important distortions present in previous models: the social benefit from specialization, spillovers, and duplication in R&D. We implement a calibration exercise in order to evaluate the strength of the new distortions from social capital. First we show that both distortions lead to underinvestment in social capital, both when we compare the social to human capital ratio and when we compare allocations of human capital to the social capital accumulation sector. Second, we also show that the presence of these distortions decrease the tendency to underinvestment in R&D. However, quantitatively, these distortions are not strong enough to get overinvestment in R&D when social capital has a positive effect in its own accumulation. The opposite result is obtained when social capital depreciates. In fact, in this case, the social capital externalities introduced in this article are capable of generating overinvestment in R&D. This complements the recent literature (Strulik[25] and Reis and Sequeira[24]) that presented more arguments in favor of overinvestment.

References

[1] Putnam, R. (1993), *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton, Princeton University Press.

[2] Knack, S. and P. Keefer (1997), "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation", *The Quarterly Journal of Economics*, 112, 1251-88.

[3] Temple, J. and P.A. Johnson (1998), "Social Capability and Economic Growth", *The Quarterly Journal of Economics,* 113(3), 965-990.

[4] Whiteley, P. (2000), "Economic Growth and Social Capital", *Political Studies*, 48(3), 443-466.

[5] Rupasingha A., Goetz, S. J., and D. Freshwater (2000), "Social Capital and Economic Growth: A County-Level Analysis", *Journal of Agricultural and Applied Economics*, 32(3), 565-72.

[6] Beugelsdijk, S. and S. Smulders (2004), "Social Capital and Economic Growth", Tilburg University, Department of Economics, *Mimeo*.

[7] Coleman, J. (1988), "Social Capital in the Creation of Human Capital", *American Journal of Sociology*, 94, S95-S120.

[8] Teachman J. D., Paasch, K., and K. Carver (1997), "Social Capital and the Generation of Human Capital", *Social Forces*, 75(4), 1343-1359.

[9] Grafton, R. Q. ; Kompas, T., and P. D. Owen (2007), "Bridging the Barriers: Knowledge Connections, Productivity and Capital Accumulation", *Journal of Productivity Analysis* 28, 219-231.

[10] Dinda, S. (2008), "Social Capital in the Creation of Human Capital and Economic Growth: A Productive Consumption Approach", *The Journal of Socio-Economics*, 37, 2020-2033.

[11] Sequeira, T.N. and A. Ferreira-Lopes (2008) "An Endogenous Growth Model with Human and Social Capital Interactions", *ERC Working Paper 09/08*.

[12] Putnam, R. (2000), *Bowling Alone: The Collapse and Revival of American Community*, New York: Simon and Schuster.

[13] Landry, R. , Amara, N., and M. Lamari (2002), "Does Social Capital Determine Innovation? To What Extent?", *Technological Forecasting and Social Change*, Vol. 69, 681-701.

[14] De Clercq, D. and M. Dakhli (2004), "Human Capital, Social Capital, and Innovation: A Multi-Country Study", *Entrepreneurship and Regional Development*, Vol. 16, Number 2, March, 107-128.

[15] Lee, S.H., Wong, P.K., and C. L. Chong (2005), "Human and Social Capital Explanations for R&D Outcomes", *IEEE Transactions on Engineering Management*, Vol. 52, No. 1, February, 59-68.

[16] Jones, C. I. (1995), "R&D-Based Models of Economic Growth", *Journal of Political Economy*, Vol. 103, Issue 4, August, 759-784.

[17] Jones, C. and J. Williams (2000), "Too Much of a Good Thing? The Economics of Investment in R&D", *Journal of Economic Growth* 5, March, 65-85.

[18] Romer, P. (1990), "Endogenous Technological Change", *Journal of Political Economy* 98, 71-102.

[19] Grossman, G. and E. Helpman (1991), *Innovation and Growth in the Global Economy*, MIT press.

[20] Aghion, P. and P. Howitt (1998), *Endogenous Growth Theory,* MIT press.

[21] Jones, C. and J. Williams (1998), "Measuring the Social Return to R&D", *The Quarterly Journal of Economics*, November, 1119-1135.

[22] Stokey, N. (1995), "R&D and Economic Growth", *Review of Economic Studies* 62, February, 469-489.

[23] Benassy, J. P. (1998), "Is There Always Too Little Research in Endogenous Growth With Expanding Product Variety?", *European Economic Review* 42, 61-69.

[24] Reis, A.B. and T. N. Sequeira (2007), "Human Capital and Overinvestment in R&D", *Scandinavian Journal of Economics* 109(3), 573-591.

[25] Strulik, H. (2007), "Too Much of a Good Thing? The Quantitative Economics of R&D Driven Growth Revisited", *Scandinavian Journal of Economics*, 109(2), 369-386.

[26] Alvarez-Palaez, M. and C. Groth (2005), "Too Little or Too Much R&D?", *European Economic Review* 49, 437-456.

[27] Norrbin, S. (1993), "The Relation Between Price and Marginal Cost in U.S. Industry: A Contradiction", *Journal of Political Economy*, Vol. 101, No. 6, 1149-1164.

[28] Funke, M. and H. Strulik (2000), "On Endogenous Growth with Physical Capital, Human Capital and Product Variety", *European Economic Review* 44, 491-515.

[29] World Bank (2006), "Where is the Wealth of Nations? Measuring Capital for the 21^{st} Century", The World Bank, Washington D.C.

[30] Pessoa, A. (2005), "Ideas' Driven Growth: the OECD Evidence", *Portuguese Economic Journal*, 4(1), 46-67.

A First Order Conditions for the Decentralized Equilibrium

The choice variables for the consumers are C, H_H, and H_S, so the first order conditions for the consumer problem yield:

$$\frac{\partial U}{\partial C} = \lambda_a \tag{45}$$

$$\lambda'_H = \frac{\lambda_a W_H}{\xi} \tag{46}$$

$$\lambda'_S = \frac{\lambda_a W_H}{\omega} \tag{47}$$

as well as:

$$\frac{\dot{\lambda}_a}{\lambda_a} = \rho + \delta_P - r \tag{48}$$

$$\frac{\dot{\lambda}'_H}{\lambda'_H} = \rho + \delta_H - \xi \tag{49}$$

$$\dot{\lambda}'_S = \rho\lambda'_S - \left(\frac{\partial U}{\partial K_S} + \lambda'_H\gamma + \lambda'_S\Omega\right) \tag{50}$$

where λ_a is the co-state variable for the budget constraint, and λ'_H and λ'_S are co-state variables for the stocks of human and social capital, respectively.

B Growth Rates in the Decentralized Equilibrium

In the steady-state, we can obtain the human capital growth rate of the decentralized equilibrium as follows. By using equation(49) and replacing it in $g_{\lambda'_H} = g_{\lambda_a} + g_W$ which we get by (46), we find $\frac{\dot{\lambda}_a}{\lambda_a} = \rho + \delta_H - \xi - g_W$. From (22)we get $g_W = g_Y - g_{K_H}$. Substituting this last equation in the previous one and introducing both in $-\frac{1}{\tau}g_y + \psi\left(1-\frac{1}{\tau}\right)g_{K_H} = \frac{\dot{\lambda}_a}{\lambda a}$ which we find in (45) and using equations (5) and (8) we get:

$$g_H^{DE} = \frac{\xi - \delta_H - \rho}{\left(\frac{\eta\left(\frac{\nu}{1-\phi}\right)}{1-\beta} + \psi\right)\left(\frac{1}{\tau} - 1\right) + \frac{1}{\tau}} \tag{51}$$

Using the fact that $g_{K_H} = \frac{(1-\phi)}{\nu}g_{K_R}$ we solve for the growth rate of technological capital:

$$g_{K_R}^{DE} = \frac{\frac{\nu}{(1-\phi)}(\xi - \delta_H - \rho)}{\left(\frac{\eta\left(\frac{\nu}{1-\phi}\right)}{1-\beta} + \psi\right)\left(\frac{1}{\tau} - 1\right) + \frac{1}{\tau}} \tag{52}$$

From (8) and $g_{K_H} = \frac{(1-\phi)}{\nu}g_{K_R}$ we find $g_Y = g_{K_H}\left(\frac{\eta\left(\frac{\nu}{1-\phi}\right)}{1-\beta} + 1\right)$. By substituting (18) in the previous equality we find:

$$g_Y^{DE} = \frac{(\xi - \delta_H - \rho)\left(\frac{\eta\left(\frac{\nu}{1-\phi}\right)}{1-\beta} + 1\right)}{\left(\frac{\eta\left(\frac{\nu}{1-\phi}\right)}{1-\beta} + \psi\right)\left(\frac{1}{\tau} - 1\right) + \frac{1}{\tau}} \tag{53}$$

These three growth rates are equal to the ones that we found in the social planner problem.

Now we also show that the consumption to output ratio is equal in the decentralized equilibrium and in the social planner solution.

From equation (1) we get:

$$\frac{\dot{K}_P}{K_P} = \frac{Y}{K_P} - \frac{C}{K_P} - \delta_P \tag{54}$$

Also, from this equation we get $g_{K_P} = g_Y = g_C$ because the growth rates have to be constant in steady state. Since we have shown that $(g_{K_P} = g_Y = g_C)^* = (g_{K_P} = g_Y = g_C)^{DE}$ the left hand-side of equation (54) is equal in the social planner and in the decentralized equilibrium and δ_P is a constant.

In the social planner problem by transforming equation (11) we obtain - $\frac{\dot{\lambda}_P}{\lambda_P} = -\frac{1}{\tau}g_C + \psi(1-\frac{1}{\tau})g_{K_S}$. Replacing this last equation in equation(15) we get - $\frac{\beta Y}{K_P} = \rho + \delta_P + \frac{1}{\tau}g_C - \psi(1-\frac{1}{\tau})g_{K_S}$.

Since we have shown that $g^*_{K_P} = g^{DE}_{K_P}$ hence $\left(\frac{Y}{K_P} - \frac{C}{K_P}\right)^* = \left(\frac{Y}{K_P} - \frac{C}{K_P}\right)^{DE}$. Using the fact that $\frac{C}{K_P}$ can be written has $\frac{C}{Y}\frac{Y}{K_P}$ and putting $\frac{Y}{K_P}$ in evidence, we obtain that $\left(\frac{Y}{K_P}\right)^*\left(1-\frac{C}{Y}\right)^* = \left(\frac{Y}{K_P}\right)^{DE}\left(1-\frac{C}{Y}\right)^{DE}$. From (27) we note that $\left(\frac{Y}{K_P}\right)^{DE} = \frac{r}{\alpha\beta}$ and that $r = \alpha P m g_K$, thus $\left(\frac{Y}{K_P}\right)^{DE} = \frac{Pmg_K}{\beta} = \left(\frac{Y}{K_P}\right)^*$. Finally $\left(\frac{C}{Y}\right)^* = \left(\frac{C}{Y}\right)^{DE}$.

COMERCIO ELECTRÓNICO EN LAS PYMES: EL CASO DE SEVILLA

VICENTE FERNÁNDEZ RODRÍGUEZ
E-mail: vicentef@etea.com
SALUD MILLÁN LARA
E-mail: smillan@etea.com
MERCEDES TORRES JIMÉNEZ
E-mail: mtorres@etea.com
Facultad de C. Económicas y Empresariales
Universidad de Córdoba

Resumen

El desarrollo del comercio electrónico afecta con carácter general a todos los sectores productivos pero de manera heterogénea en función del tamaño de las organizaciones. Así la implantación del comercio electrónico ha sido muy rápida en las grandes empresas pero no en el caso de las pymes, como lo demuestran los porcentajes de venta por esta vía en ambos casos. El presente trabajo clasifica las web de las pymes con sede central en Sevilla en diferentes niveles desde el punto de vista del comercio electrónico, tal como se hizo en dos estudios anteriores (en Córdoba y Málaga).
Posteriormente estudia la relación de dicha clasificación con la capacidad de las empresas analizadas para ofrecer comercio electrónico, así como con otros elementos relacionados con la presencia de las empresas en la red.

Palabras Clave: comercio electrónico, pymes, clasificación.
Área temática: Economía de la Información y del Conocimiento.

Abstract

The development of electronic commerce is a reality in all productive sectors in general, but its application varies greatly depending on the company's size. Electronic commerce has been introduced into big companies very quickly, but this has not been the case in SMEs, as has been demonstrated by on line sales volume in the respective cases. In this work the SME webs with headquarters in Seville, Spain, have been classified in different levels according to their respective electronic commerce, as previously done in two research studies in Cordoba and Malaga respectively. After that, the classification obtained was related to the companies' capacity to offer electronic commerce and also to other elements related to their business presence on the net.

Key Words: Electronic commerce, SMEs, classification.
Thematic Area: Information and Knowledge Economy.

1. INTRODUCCIÓN

Siendo prudentes podemos afirmar que ya hace más de un lustro que no existen limitaciones técnicas para que el comercio electrónico sea una herramienta al alcance de cualquier empresa. El desarrollo en capacidad de acceso a contenidos, a seguridad, o a intercambio de información ha posibilitado un crecimiento espectacular del comercio a través de la red, hasta el punto de convertirse en un arma estratégica de primer nivel en el desarrollo comercial de algunas empresas y sectores.

Esa evolución es especialmente destacable en sectores como el financiero o el de negocios de viajes y actividades relacionadas (alojamientos, compra de billetes, alquiler de vehículos o vacaciones) en los que la implantación de las TICs está mucho más desarrollada. Sin embargo

en otros sectores, que a su vez son los que más aportan al PIB de nuestro país, como el comercio minorista (20,5%) o el transporte (9,4%) el nivel de implantación de dichas tecnologías es aún muy reducido, lo que está suponiendo un retraso en la productividad de la economía española (Fundetec, 2008).

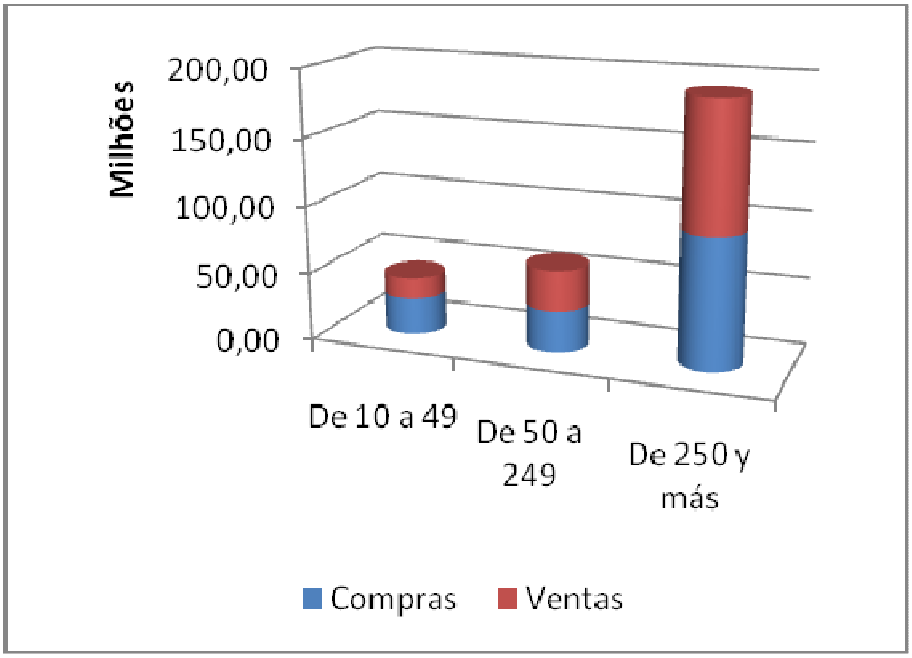
El presente trabajo forma parte del proyecto de investigación sobre "el grado de implantación del comercio electrónico en las pymes andaluzas" que se está llevando a cabo en la facultad de empresariales ETEA (Córdoba).

Los antecedentes de este trabajo son la ponencia presentada por los autores en *Mundo Internet 2004* bajo el título "Implantación del comercio electrónico: el caso de Córdoba", la ponencia de los mismos autores titulada "Web: computing tool, business tool. Malaga firms in a case study" presentada en la conferencia "The Internet & Society 2006" organizada por el Wessex Institute of Technology, y la ponencia presentada en Mundo Internet 2008 bajo el título "Comercio electrónico, ¿privilegio de los grandes?

En esos antecedentes, a través del análisis del comercio electrónico en Córdoba y en Málaga, se pone de manifiesto que el uso del comercio electrónico es más la excepción que la norma en las pequeñas y medianas empresas andaluzas, y cabe esperar que siga siendo así durante algún tiempo. Este hecho se produce no solo en las pymes andaluzas, sino en general en las pymes españolas como se advierte en el gráfico 1.

Es curioso comprobar que en las empresas de menor tamaño (menos de 50 trabajadores) el volumen de ventas por comercio electrónico es inferior al volumen de compras por dicho medio (representando las ventas el 56% de las compras). En el resto de casos ambas variables son prácticamente iguales (97,8 y 99,6%).

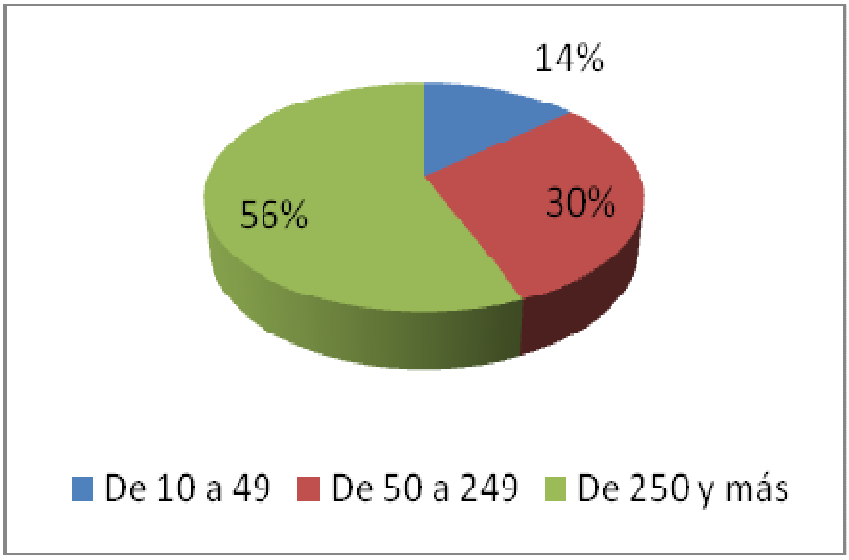
Gráfico 1: Uso del comercio electrónico (compras y ventas en millones de euros) por las empresas españolas según tamaño (número de trabajadores) en el 2007



Fuente: Elaboración propia a partir de la encuesta de uso de TIC y Comercio Electrónico (CE) en las empresas 2007-2008 del INE

Si analizamos la proporción de ventas por comercio electrónico de las empresas españolas sobre el total de sus ventas, según el tamaño de las mismas (número de trabajadores) observamos que en las empresas de menos de 50 trabajadores dicho ratio alcanza el 14%, en las empresas cuyo número de trabajadores osciles entre 50 y 249 alcanza el 30% y en las más grandes el 54% (gráfico 2), lo que corrobora que las variables "tamaño de la empresa" y "desarrollo del comercio electrónico" están directamente relacionadas.

Gráfico 2: Proporción de ventas por comercio electrónico vs total ventas empresas españolas según tamaño (medido por nº de trabajadores) (año 2007)



Fuente: Elaboración propia a partir de la encuesta de uso de TIC y Comercio Electrónico (CE) en las empresas 2007-2008 del INE

En este contexto, la evolución positiva del comercio electrónico está muy ligada al desarrollo de las páginas webs que las empresas hacen, puesto que son las herramientas que permiten las transacciones.

Parece evidente que las empresas que ya han conseguido desarrollar comercio electrónico, continuarán haciéndolo e incorporando todas las actualizaciones y mejoras que les permitan mantener y/o ampliar su servicio y su negocio, pero no es tan evidente la incorporación masiva a este proceso de la pequeña y mediana empresa, tal como lo demuestran los datos de las estadísticas nacionales así como de los estudios que hemos realizado anteriormente, y que éste nuevo estudio, en el que se presentan los resultados que se han obtenido en el análisis de las empresas con sede en Sevilla, corrobora.

2. OBJETIVOS

Los objetivos de este trabajo se centrar en contrastar el grado de cumplimiento de las siguientes hipótesis:

1. El lento desarrollo del comercio electrónico en las pymes en comparación con las grandes empresas es un hecho evidente, pero posiblemente no esté motivado principalmente por dificultades tecnológicas ni económicas; nuestra hipótesis es que la razón fundamental debe buscarse en elementos culturales y de formación.
2. Las pymes disponen de web con dominio propio.
3. Las pymes ofrecen sus productos a través de la web.
4. Las pymes venden a través de la web.

Los niveles de las hipótesis anteriores son distintos en función de las variables que definen las capitales analizadas.

3. METODOLOGÍA

Se ha generado una base de datos con toda la información analizada de empresas con sede central en la ciudad de Sevilla, de tal forma que se facilite, tanto el análisis de esas empresas, como su posterior comparación con los resultados del resto de las capitales andaluzas. Dicha base de datos incluye el almacenamiento de las direcciones URL de las empresas que tienen sede en Sevilla, y a partir de ahí se han analizado los correspondientes sitios webs intentando mantener el contacto con cada empresa a través de su web.

3.1. Empresas analizadas

El estudio incluye solamente las empresas cuyas sedes centrales se encuentran en la capital de la provincia. Partiendo de la totalidad de empresas, se ha realizado una búsqueda para ver cuales de ellas tienen presencia en la red. Aunque no hemos conseguido información de ninguna fuente que nos garantice el análisis del 100% de la población, la cifra total de empresas de las que hemos conseguido algún tipo de referencia, asciende a 10.758, lo que podría suponer una cifra que supere ampliamente el 90% del total de empresas con presencia en la red; esto supone, sin lugar a dudas, que estamos trabajando con una muestra con un altísimo nivel de representatividad.

Para identificar esas empresas, se ha utilizado como fuente la página web de Páginas Amarillas: www.paginasamarillas.com.

A partir del listado de empresas obtenido, la primera fase de selección ha consistido en diferenciar entre aquellas empresas que realmente tienen su propia página web, y las que simplemente aparecen incluidas en la web de Páginas Amarillas.

3.2. Obtención de información.

La obtención de información se ha intentado completar a partir de las propias páginas de las empresas, para lo que intentamos acceder a cada una de ellas.

Aquellas páginas que pudimos cargar, las clasificamos en los niveles que se presentan más adelante, dependiendo de las posibilidades que ofrecían para el desarrollo del comercio electrónico.

Finalmente, intentamos obtener información básica de los productos ofrecidos por las empresas a través de sus páginas web, así como de su utilización para la prestación de diferentes servicios.

3.3. Clasificación de las web por niveles.

La clasificación utilizada incluye seis niveles diferentes:

- Nivel 1: páginas que solamente ofrecen información general de la empresa.
- Nivel 2: Estas páginas incluyen también información de los productos.
- Nivel 3: Permiten establecer contacto directo con la empresa.
- Nivel 4: Además del contacto, la página web admite la posibilidad de realización de pedidos.
- Nivel 5: La web permite concretar y formalizar el proceso de entrega.
- Nivel 6: Ofrece la posibilidad de formalizar el pago también a través de la página.

3.4. Análisis de datos.

Una vez aplicada la clasificación anterior a las empresas estudiadas, se ha realizado un análisis descriptivo en relación con la solicitud de información a los clientes, los medios de pago, las posibilidades de contacto, la presencia en la red, y los servicios ofrecidos por las empresas a través de sus páginas web.

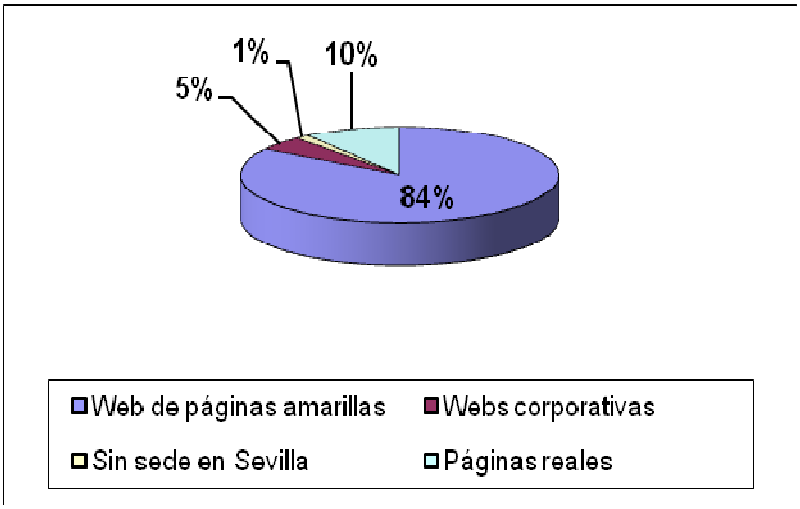
4. RESULTADOS DEL ESTUDIO

4.1 Datos de población y muestreo

Una vez fijados los niveles y características que ha de cubrir una web de comercio electrónico, vamos a analizar qué grado de cumplimiento tienen las webs de las empresas que tienen sede en Sevilla capital.

De las 10.758 empresas analizadas, 9.028 (84%), limitan su presencia a la referencia web que facilita Páginas Amarillas, es decir, estas empresas, aunque no renuncian a su presencia en la red, esa presencia es mínimamente informativa.

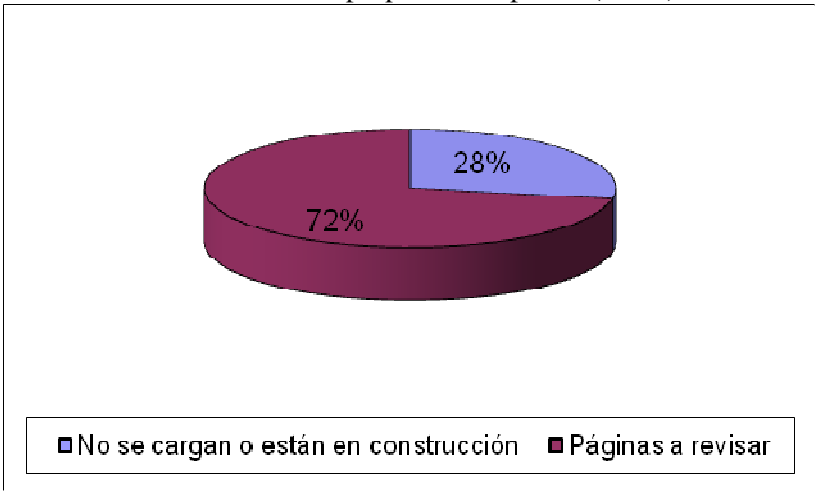
Gráfico 3. Web analizadas



Fuente: Elaboración propia.

Si además eliminamos del análisis las 522 webs corporativas (5%) y las 153 (1%) que corresponden a empresas sin sede en Sevilla, nos quedan finalmente 1.055 empresas que ofrecen una web propia. Puestos a analizar estas páginas nos encontramos que casi el 28% de las mismas (295 páginas), no se cargan, lo que implica que finalmente son 760 empresas aquellas cuyas webs podamos analizar. En éstas nos hemos centrado, que representan el 7% de las páginas detectadas en la búsqueda inicial; de ellas se han analizado una serie de variables que son las que pasamos a describir.

Gráfico 4. Webs propias de empresas (1.055)



Fuente: Elaboración propia.

4.2 Análisis en función del nivel de la web estudiada

Atendiendo a la clasificación de niveles de web descrita en el apartado tres, sabemos que el nivel 1 lo identificamos con sitios webs informativos, el nivel 2 con web que muestran sus productos, el nivel 3 permite el contacto con el cliente, el nivel 4 permite hacer el pedido, el nivel 5 lo tienen aquellas webs que gestionan la entrega del producto y el nivel 6 aquellas que además ofrecen el servicio para gestionar el pago.

El que una empresa alcance un nivel u otro está muy influenciado, entre otros factores, por el tipo de producto que ofrece. Las empresas que ofrecen productos distribuibles a través de la red tienen más facilidad para alcanzar el sexto nivel. Del total de páginas analizadas 34 ofrecen productos virtuales, el 3,2%.

Es lógico pensar que si una web alcanza un nivel, cubre los niveles anteriores, aunque no siempre tiene por qué ser así; por ese motivo hemos analizado cada nivel de forma independiente, calculando su importancia sobre el total de las páginas analizadas.

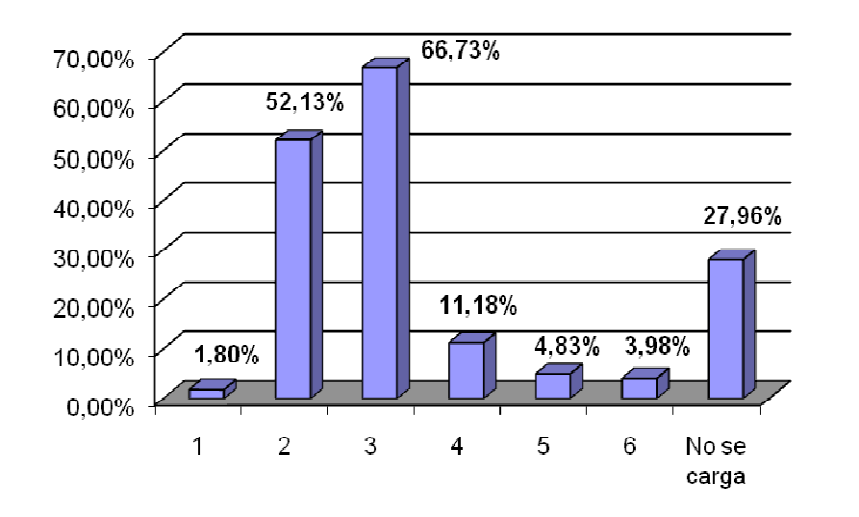
Tabla 1. Distribución de páginas Webs según el nivel máximo alcanzado

NIVEL DE WEB	Sevilla	%
1	19	1,80
2	550	52,13
3	704	66,73
4	118	11,18
5	51	4,83
6	42	3,98
No se carga	**295**	**27,96**
TOTAL	**1.055**	**100,00**

Fuente: Elaboración propia.

Podemos observar que en Sevilla el porcentaje de empresas cuyas webs son solamente informativas es muy bajo (1,8%). El porcentaje más representativo es el que alcanza el nivel de oferta de productos y contacto con el cliente (nivel 3), aunque el nivel 2 también presenta una frecuencia muy alta. A la vista de los datos se puede comprobar cómo la situación se hace más complicada para alcanzar los niveles de transacción de productos y realización del cobro. (Véase gráfico 5 y tabla 1).

Gráfico 5. Nivel de las Webs en Sevilla. Cumplimiento de cada nivel sin considerar si se cumplen o no los anteriores.



Fuente: Elaboración propia.

4.3 Solicitud de información a los clientes

La posibilidad de gestionar datos de clientes la tienen las 206 empresas de Sevilla (19,5%) que solicitan datos vía formulario a sus clientes.

El número de datos solicitados varía de 2 a 35, siendo la media de 14 datos. Como ya hemos comentado, no es conveniente solicitar muchos datos al navegante porque se corre el riesgo de que abandone su opción de compra, pero esta herramienta no se debe despreciar si queremos tener una gestión personalizada de nuestros visitantes.Hay que solicitar sólo los datos necesarios para llevar a cabo la venta, y en próximos contactos ir aumentando los datos solicitados. El objetivo es ir creando una base de datos de clientes para poder gestionar ventas personalizadas, pero esta base se constituirá de forma escalable.

Del total de páginas analizadas, 129 tienen zona privada para clientes, por lo que pueden realizar gestión CRM (12,2%).

4.4 Pago a través de la web

Tal como se muestra en la tabla 2, un total de 42 páginas web permiten gestionar el pago; los medios admitidos en cada caso se recogen en la tabla siguiente:

Tabla 2. Medios de pago

Medio de pago	Nº webs 42	%
Transferencia-Tarjeta-Contra reembolso	12	28,57
Transferencia-Tarjeta	14	33,33
Transferencia-Contra reembolso	4	9,52
Tarjeta-Contra reembolso	1	2,38
Transferencia	4	9,52
Tarjeta	4	9,52
Contra reembolso	3	7,14

Fuente: Elaboración propia.

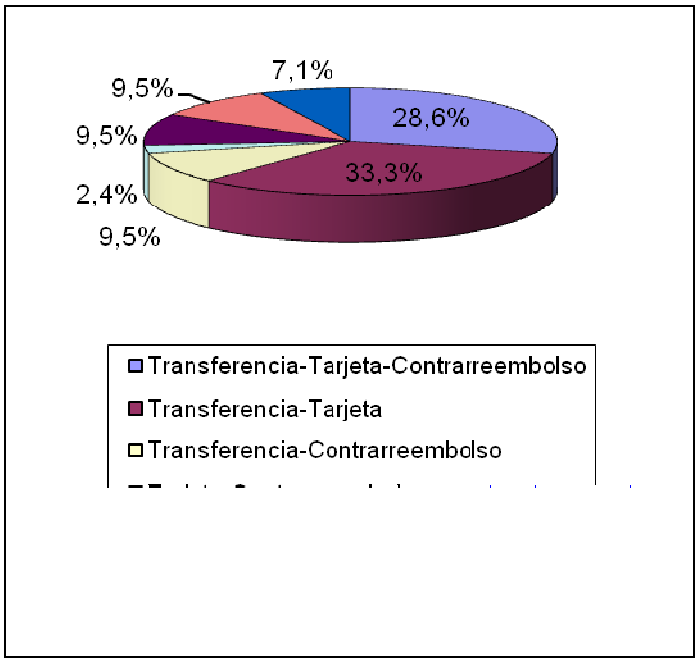
Como cabría esperar, destaca el hecho de que más del 60% de las empresas admiten tanto la transferencia como el pago con tarjeta, siendo la combinación de opciones más habitual. Por separado, la transferencia se admite en más del 80% de los casos, y la tarjeta en casi un 74%. Con alguna de esas dos opciones tenemos casi el 93% de los casos.

La opción de pago mediante contra reembolso es aceptada por algo más del 47% de las empresas.

Las webs sevillanas que entran en página segura son 22, sólo la mitad de las que gestionan el pago a través de la página.

En algunos casos, las empresas ofrecen precios distintos según la forma de pago, siendo el contra reembolso el medio que tiene más riesgo y suele ir acompañado de un precio más alto. En el caso de Sevilla, esto ocurre en 9 de las páginas analizadas.

Gráfico 6. Distribución de los medios de pago en las Webs

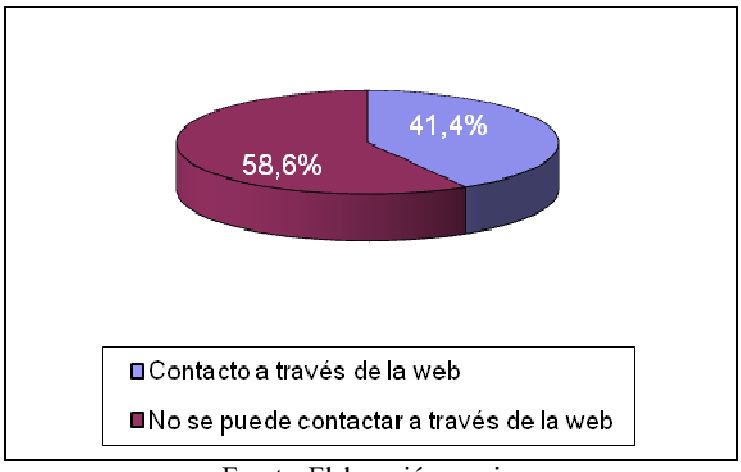


Fuente: Elaboración propia.

4.5 Posibilidad de contacto a través de la web, actualización y control de visitas

De las 1.055 páginas analizadas, sólo hemos conseguido contactar con las empresas en 437 casos (el 41,4%). Ante una consulta genérica se han obtenido muy pocas respuestas y con una tardanza excesiva; este hecho no va en consonancia con la idea de negocio y atención al cliente desde una web.

Gráfico 7. Posibilidad de contacto a través de la web



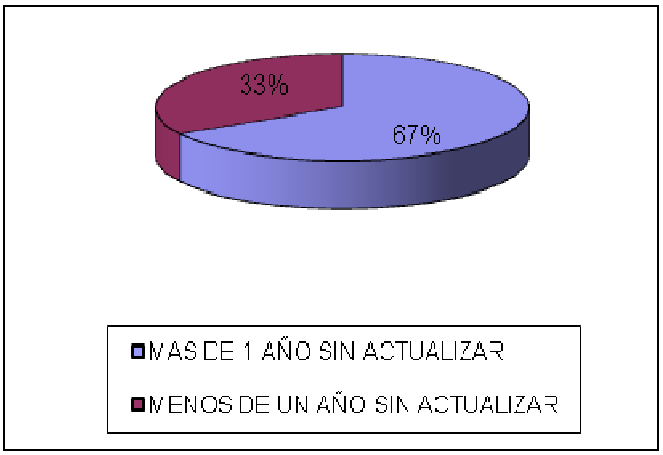
Fuente: Elaboración propia.

Sabemos que el usuario de Internet espera poder contactar rápidamente con la página que le interesa, no sólo para ejercer su opción de compra, sino también para poder hacer alguna consulta, esperando una respuesta casi inmediata.

Una referencia del grado de interés que presenta una web es el número de visitantes de tal forma que cuando este dato no es significativo es preferible omitirlo; en el caso de nuestro estudio, son 34 las páginas que muestran este dato. La media de visitas que muestran esas 34 páginas es de 45.158.

Otro aspecto a tener en cuenta es en qué medida una web se actualiza; no se puede mantener el mismo contenido durante mucho tiempo porque eso puede llevar al usuario a dudar de su operatividad y a no volver a visitar la página. Tan sólo 39 de las páginas analizadas muestran la fecha de actualización, y por eso no fuera un dato suficientemente negativo, sólo un tercio de las que lo hacen, las actualizan con una frecuencia inferior al año. Esto es otro indicador que pone de manifiesto la importancia que las empresas atribuyen a su presencia en la red.

Gráfico 8. Frecuencia de utilización de páginas web



Fuente: Elaboración propia.

En algunos casos, el dato que aparece es que la web lleva hasta 7 años sin actualizar, lo cual consideramos que es mucho más negativo que el hecho de que no apareciese ninguna información al respecto.

4.6 Presencia en otras webs

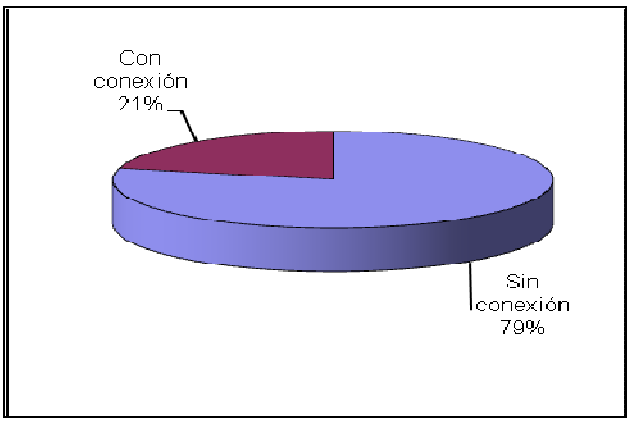
Una web suele albergar publicidad de otras empresas con las que normalmente mantienen hipervínculo, esto es posible sólo para las webs que realmente tienen un elevado número de visitantes.

De las webs analizadas, 159 empresas (el 15,1%) contienen en sus páginas publicidad de otras empresas, con un número de apariciones muy variable; estas apariciones presentan una media de 7,2 eliminando el extremo de más de 2.002.

También es posible que la conexión hacia otras páginas no se haga por motivos publicitarios, sino que se trate de enlaces de interés que la empresa recomienda; eso es lo que ocurre en el 21% de las páginas analizadas, tal como se recoge en el gráfico 7.

La media de conexiones a otras webs de las páginas de las empresas de la capital sevillana es de 8,7.

Gráfico 9. Conexión con otras webs

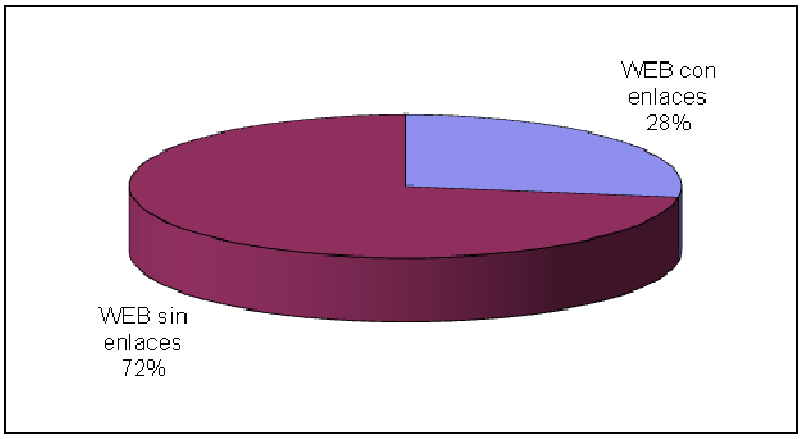


Fuente: Elaboración propia.

La empresa que ofrece menos conexiones ofrece solamente una, y la que más, 651 conexiones.

El análisis de presencia en la red también podría hacerse a través del número enlaces en sentido opuesto, esto es, desde cuantas webs se tiene acceso a nuestra página. En las empresas con sede central en Sevilla, 291 (lo que representa el 27,6% de las analizadas), tienen enlaces desde otras webs, y la media de esos enlaces es de 23.

Gráfico 10. Distribución de páginas a las que se puede acceder desde otras webs



Fuente: Elaboración propia.

4.7. Servicios que prestan las webs

Las empresas a través de sus webs ofrecen servicios que animan a los usuarios a visitar la página, aunque no estén relacionados directamente con su actividad empresarial.

De entre las de empresas sevillanas, 225 páginas (el 21,3%) ofrecen servicios adicionales a los navegantes. La tabla 3 muestra la distribución de dichos servicios.

Tabla 3. Oferta de servicios

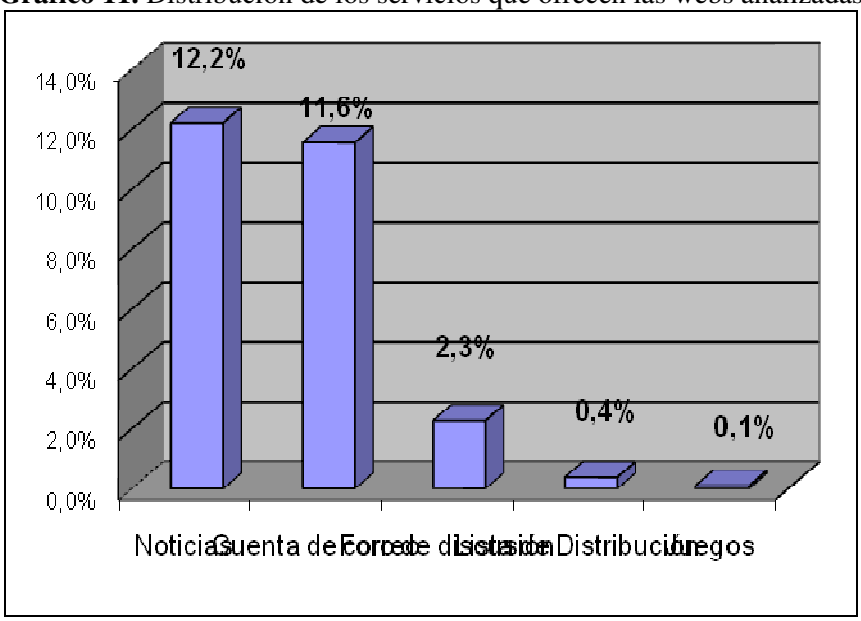
Servicio	Cantidad	%
Noticias	129	12,2
Cuenta de correo	122	11,6
Foro de discusión	24	2,3
Lista de Distribución	4	0,4
Juegos	1	0,1

Fuente: elaboración propia

El servicio más ofrecido en Sevilla es el de noticias, seguido muy de cerca por la cuenta de correo. Ofrecer cuenta de correo, al ser un servicio de los más utilizados, es un método para asegurar que los clientes visiten la página, aunque las empresas se están planteando ofrecer el servicio previo pago.

Son pocas las empresas que ofrecen más de un servicio (sólo el 3%); cuando lo hacen, la combinación más frecuente es noticias y cuenta de correo; esta combinación la ofrecen 19 empresas.

Gráfico 11. Distribución de los servicios que ofrecen las webs analizadas



Fuente: Elaboración propia.

5. CONCLUSIONES

Los aspectos más destacados a los que nos conducen los datos analizados, son los siguientes:

- ✓ Las pymes disponen de web con dominio propio sólo en uno de cada 10 casos. Si además descartamos aquellas que no pudimos cargar a la hora de realizar nuestro análisis, bajamos al 7%.
- ✓ Las pymes ofrecen sus productos a través de la web en la práctica totalidad de las empresas que tienen página propia operativa (704 de 760), pero son mayoría las empresas en las que esa circunstancia no se da.
- ✓ La venta a través de la web es todavía un sueño lejano para la mayoría de las pymes. Otras empresas, y de otros ámbitos, lo están haciendo, las pymes de Sevilla, rotundamente, no.
- ✓ En cuanto a algunos de los parámetros que miden la presencia en la red, y con ello al aprovechamiento potencial del comercio electrónico, las empresas sevillanas sólo el 25% de las páginas analizadas están utilizando en mayor o menor medida este potencial; las tres cuartas partes restantes, no solo no desarrollan y explotan adecuadamente su presencia, sino que posiblemente la imagen que transmiten sea negativa de cara a este canal.

Es posible que la disminución que se está produciendo en el ritmo de crecimiento de las ventas a través de Internet se deba, al menos en parte, a que es más difícil y más lenta la incorporación de las pymes a ese comercio. Nuestra hipótesis es que la razón fundamental debe buscarse en elementos culturales y de formación.

Para esas empresas, y aunque poco a poco van siendo conscientes de la necesidad de tener presencia en la red, por el grado de atención que le prestan a sus páginas, se advierte que hoy por hoy no supone para ellos ni un aspecto de mejora global del negocio, ni –refiriéndonos expresamente a la actividad comercial- un canal alternativo de distribución.

BIBLIOGRAFÍA

Fundetec (2008) "Informe ePyme 2008. Análisis sectorial de implantación de las TIC en la pyme española".

INE (2008) Encuenta de uso de TIC y Comercio Electrónico (CE) en las empresas 2007-2008.

www.paginasamarillas.com.

A IMPORTÂNCIA DAS REDES NO DESEMPENHO INOVADOR: SECTOR DOS MOLDES EM PORTUGAL

FILIPA DIONÍSIO VIEIRA
E-mail: filipadv@dps.uminho.pt
FERNANDO ROMERO
E-mail: dps.uminho.pt
Departamento Produção e Sistemas
Universidade do Minho

Resumo

Com este trabalho de investigação pretendeu determinar-se onde e como oito empresas nacionais do sector dos moldes, todas elas localizadas na região da Marinha Grande, adquirem o conhecimento necessário para o desenvolvimento das suas actividades de inovação. A ênfase deste trabalho foi o estudo das diversas interacções entre os vários elementos do sistema, sejam eles internos ou externos à empresa.
Com a informação recolhida das empresas estudadas, foi possível construir um modelo conceptual de observação das empresas inovadoras do sector dos moldes, relativo ao tipo de relações existentes entre as empresas (clientes, fornecedores, concorrentes) e entre estas e outras instituições (instituições académicas, centro tecnológico, centros de investigação), permitindo uma sistematização e uma generalização na análise da dinâmica inovadora deste sector. O padrão de relacionamento é diferenciado consoante as características da empresa, nomeadamente as diferentes competências específicas, associadas a cada área nuclear, e consoante o seu posicionamento estratégico face à aquisição e criação de conhecimento.

Palavras-Chave: Inovação, Redes, Sector dos Moldes.
Área Temática: Economia da Inovação e Transferência Tecnológica.

Abstract

In this case study the aim was to understand the mechanisms for knowledge acquisition and its transformation in innovation performance by eight Portuguese firms of the moulds industry, all located in the region of Marinha Grande. Emphasis was put on the study of interactions between several external and internal elements of the system.
The information obtained through intensive observation allowed for the creation of an observation model of the moulds industry. This model is based on the type of existing interactions between firms (customers, suppliers, competitors) and between firms and other institutions (academic institutions, technological centres, research institutions). This research work is important because it permits the systematization and generalisation of the innovative dynamics in this industry.
The pattern of relationship depends on the characteristics of the firm, namely its particular abilities which are associated with each nuclear area and according to their strategic position for the knowledge acquisition (i.e. from external sources) and knowledge creation (i.e. own R&D activities).

Key words: Innovation, Networks, Moulds Industry.
Thematic Area: Innovation Economy and Technology Transfer.

1. INTRODUÇÃO

Com base no estudo de oito empresas nacionais do sector dos moldes, todas localizadas na região da Marinha Grande, pretendeu determinar-se onde e como tais empresas adquirem o conhecimento necessário para a realização das suas actividades e desempenho inovadores. Interessa saber se tais conhecimentos foram adquiridos interna e/ou externamente através de redes de inovação e determinar quais os principais elementos que constituem tais redes, bem como qual o seu desempenho, como impulsionadores da inovação. A ênfase deste trabalho está no estudo das interacções entre os vários elementos do sistema, sejam eles internos ou externos à empresa. Segundo Lundvall (1992), o conhecimento é o recurso mais importante na moderna economia e, consequentemente, a aprendizagem é o processo mais importante, sendo este predominantemente interactivo e decorre num enquadramento ou contexto institucional e cultural. A aprendizagem está relacionada com actividades de rotina, em diferentes áreas - na produção, na distribuição ou no consumo – permitindo, assim, a criação de importantes inputs para o processo de inovação.

O estudo efectuado resultou na criação de um modelo de observação do sector dos moldes, relativo ao tipo de relações existentes entre as empresas e entre estas e outras instituições, permitindo uma sistematização e uma generalização na análise da dinâmica inovadora deste sector. O padrão de relacionamento é diferenciado consoante as características da empresa, nomeadamente as diferentes competências específicas, associadas a cada área nuclear, e consoante o seu posicionamento estratégico face à aquisição e criação de conhecimento.

2. REVISÃO BIBLIOGRÁFICA

Este estudo é baseado na abordagem sistémica do processo de inovação por ser particularmente adequado aos objectivos e ao alvo do estudo. Este tipo de abordagem realça a interacção entre agentes, a interdependência e as sinergias. A observação que a caracteriza dá-se ao nível das instituições (empresas, instituições académicas, instituições de investigação, públicas ou privadas, organizações diversas) e não ao nível dos processos individuais, como por exemplo, as abordagens ligadas à gestão do conhecimento (Nonaka et al., 2001), ou as abordagens mais abstractas e formalizadas, que modelam, a um nível de decisão mais agregado ou mais estereotipado, um processo dinâmico, como por exemplo as teorias da difusão da inovação o fazem (Stoneman, 2001; Lissoni e Metcalfe, 1994; David, 1985). O estudo foi também informado por estas teorias, na medida em que se deu a devida atenção aos processos de decisão individuais, embora estas tenham sido posteriormente integradas no nível de decisão do agente institucional (empresa ou outra organização), sendo este o foco específico do estudo. É nesse sentido que a perspectiva de Teece (1998) é utilizada. O modelo, ou a forma de organização sistémica, contextualiza comportamentos de pequenas empresas, ao nível do contributo de outras organizações para a aprendizagem interna, da aprendizagem colectiva, da mobilidade de indivíduos chave, das baixas barreiras de entrada para empresas inovadoras e da flexibilidade operatória. Cada sector, ou sistema, tem, no entanto, dinâmicas próprias de evolução, decorrentes de condicionantes históricos e ambientais, e é nesse sentido que a perspectiva é indicada e pode ser usada como modelo de descrição e interpretação de várias realidades.

É amplamente assumido que a inovação é um factor-chave na competitividade e que o crescimento do *output* e da produtividade dependem do desenvolvimento e da difusão de novas tecnologias (OCDE, 1997). Daí urge a necessidade de as empresas apostarem fortemente na inovação, porque, só assim, conseguirão crescer com mais conhecimento e produzir com mais valor acrescentado, tornando-se mais competitivas.

Contudo, o processo de inovação não é linear, mas sim um processo complexo, interactivo, envolvendo instituições empresariais e não empresariais. O modelo de ligação em cadeia ou "chain-linked model", apresentado por Kline e Rosenberg (1986) é o exemplo de um modelo de inovação que mostra o carácter interactivo do processo de inovação. Na verdade, a inovação não

deve ser vista como o produto de um único actor, mas sim o resultado da interacção entre vários actores, quer institucionais quer organizacionais. E apesar do processo de inovação ser descontínuo e incerto, os sistemas nacionais de inovação ajudam a perceber e explicam porque é que a tecnologia se desenvolve numa determinada direcção.

O estudo de sistemas nacionais de inovação consiste na análise das ligações existentes entre as várias partes que compõem o sistema de inovação, sejam elas as empresas, as instituições académicas ou as instituições de investigação, públicas ou privadas (Nelson, 1993; Lundvall, 1992; Freeman, 1987). Não esquecendo que são as empresas o 'core' destes sistemas, competindo e, simultaneamente, cooperando umas com as outras (Dosi et al., 1988).

Diversos autores (Tidd and Bessant, 2009; Chung and Kim, 2003; Freel, 2003; Beeby and Booth, 2000; Sousa, 1997) referem a importância das interacções entre empresas e o ambiente social e económico onde estas se encontram inseridas, resultando assim em ligações com clientes e fornecedores, com empresas de consultadoria e com infra-estruturas da ciência e tecnologia, salientando a importância dos centros de I&D e das instituições de educação e formação. A optimização destas ligações e das competências associadas resultam no estabelecimento de redes (de inovação e de negócio) que assentam na complementaridade de competências ou de activos, na dependência mútua e em novas formas de organização económica, não primordialmente regidas por vínculos contratuais mas principalmente por relações de confiança e de partilha de riscos e benefícios (Teece et al., 1997).

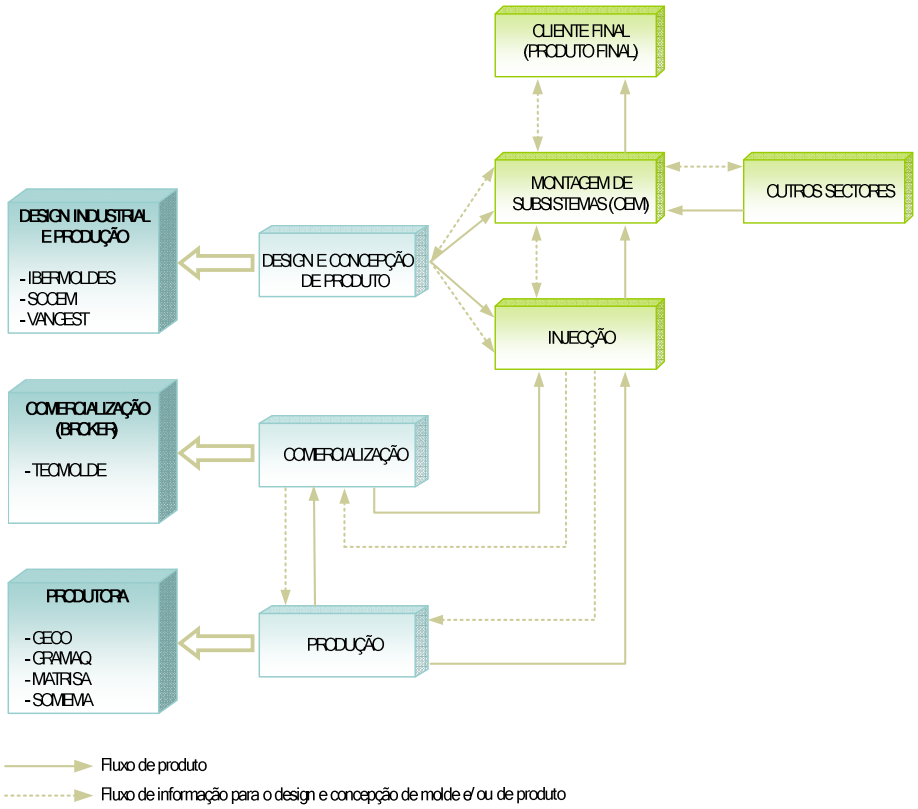
Empresas que apresentam altos níveis de colaboração tecnológica, quer entre empresas, quer com universidades e institutos de investigação públicos, e de aquisição de tecnologia e mobilidade de recursos humanos, conseguem melhorar a sua capacidade inovadora em termos de produtos e processos. A facilidade com que o conhecimento é transferido depende do seu tipo. O conhecimento explícito é formal e sistemático, portanto pode estar facilmente associado a especificações de produtos, fórmulas ou programas de computador. O conhecimento tácito é, em parte, capacidade técnica, mas também tem uma dimensão cognitiva formada por modelos mentais, crenças e perspectivas, que são difíceis de articular. A socialização é a única maneira de transferir conhecimento tácito entre indivíduos, através da observação, imitação e prática (Arias, 1995).

Talvez por isso, as regiões parecem ter importância acentuada para o desenvolvimento de redes e de sistemas de novas tecnologias. As infra-estruturas locais, os recursos humanos especializados, o mercado de trabalho local, os serviços especializados e as relações pessoais são fundamentais e contribuem para o desenvolvimento regional (Muscio, 2006; Chung, 2002; Freeman e Soete, 1997).

3. MODELO DE OBSERVAÇÃO E ANÁLISE DE EMPRESAS INOVADORAS DO SECTOR DOS MOLDES

Apesar de todas as empresas estudadas pertencerem ao mesmo sector da indústria nacional – Sector dos Moldes -, as empresas apresentam competências diferentes, o que permite agrupá-las segundo a sua posição na cadeia de valor do mesmo, criando uma tipologia específica para cada um desses grupos de empresas, como se pode constatar pela análise da figura 1, abaixo apresentada.

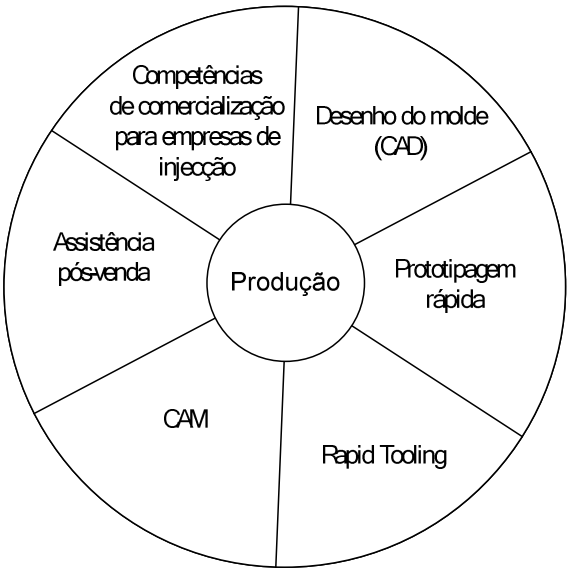
Figura 1: Representação da cadeia de valor do sector dos moldes e sua relação com a tipologia das empresas estudadas



Da observação da figura 1 ressalta: (1) a cadeia de valor do sector assenta em várias actividades, que podem ir desde a produção até à entrega do produto final ao cliente final, evidenciando-se o fluxo de produto e o fluxo de informação para o design e concepção de molde e/ou produto; (2) as áreas nucleares de competências das empresas estudadas que permitiram a criação de uma tipologia de empresas: produção, comercialização e design e concepção de produto e (3) as diferentes empresas de moldes estudadas associadas à tipologia, entretanto criada: GECO, GRAMAQ, MATRISA e SOMEMA (Produtora); TECMOLDE (Comercialização) e IBEROMOLDES, SOCEM e VANGEST (Design Industrial e Produção).

As figuras (2, 3 e 4) que se seguem mostram com mais pormenor as diferentes competências específicas, associadas a cada área nuclear, existentes nos diferentes grupos de empresas estudadas. Estas figuras foram efectuadas com base na perspectiva de Teece (1998) que refere a importância das diferentes competências para as actividades inovadoras das empresas. Estas competências podem existir internamente ou poderão ser adquiridas externamente. As figuras, que representam as três tipologias de empresas, dizem respeito a competências internas das empresas de acordo com as respectivas áreas nucleares de competências.

Figura 2: Competências internas das empresas produtoras



Estas empresas (GECO, GRAMAQ, MATRISA e SOMEMA) são tipicamente empresas produtoras de moldes, possuindo elevado know-how na produção de moldes, sendo as suas competências no desenho do molde (CAD), em prototipagem rápida, no *rapid tooling* e no CAM elevadas, mas com uma limitada capacidade de engenharia e concepção do molde. O desenvolvimento de actividades de I&D, engenharia e design do produto são inexistentes. Apresentam também competências consideráveis na assistência pós-venda e na comercialização para empresas de injecção.

Figura 3: Competências internas da empresa comercial (*Broker*)



A TECMOLDE é uma empresa que não tem componente produtiva e ainda que subcontrate outras empresas exclusivamente produtoras, possui uma elevada capacidade comercial e de integração de outras empresas. Esta empresa tem como principal competência a gestão de redes de subcontratados, quer de produtores de moldes, quer de serviços especializados, o que lhe permite a gestão de grandes projectos, i.e., sistemas de moldes. Também apresenta competências nas áreas de engenharia e concepção do molde e marketing.

Figura 4: Competências internas das empresas com design industrial e produção



Esta categoria de empresas (IBEROMOLDES, SOCEM e VANGEST) posiciona-se num patamar mais elevado da cadeia de valor do sector dos moldes, comparativamente às duas categorias anteriores. As suas competências de engenharia e concepção de moldes, conjuntamente com a parte produtiva, estão bastante desenvolvidas, aliando-se competências emergentes no design e concepção de produtos injectados, o que lhes permite fornecer soluções inovadoras mais completas aos seus clientes, conseguindo entrar nos mercados mais exigentes. As fontes de informação para a inovação são, fundamentalmente, internas e resultam de actividades próprias de I&D. Também apresentam competências na assistência pós-venda, na área do marketing e em serviços especializados. A última competência enunciada leva a concluir que estas empresas apresentam um nível de integração vertical considerável. Devido às suas características já são capazes de comercializar para a OEM (*Original Equipment Manufacturer*), o que representa uma subida significativa, relativamente às empresas nas outras categorias, na cadeia de valor do sector, aproximando-se cada vez mais do cliente final.

A tabela 1, a seguir apresentada, foi construída a partir da síntese dos resultados do estudo realizado e propõe um modelo conceptual de observação das empresas inovadoras do sector dos moldes, mostrando como as características das empresas estudadas, que dependem da sua posição na cadeia de valor do sector em questão, se reflectem em termos das relações existentes com outras empresas, tais como clientes, fornecedores, concorrentes e com instituições, sejam elas, de ensino superior, de investigação públicas e privadas e infra-estruturas tecnológicas. Permite, ainda, tirar ilações sobre as áreas nucleares de competências, a origem das ideias para o desenvolvimento das inovações e o tipo de inovações que são desenvolvidas pelas empresas. Sendo assim, as empresas comportam-se, em termos inovadores, de maneiras distintas, estando fortemente dependentes do padrão de ligações informais que mantêm.

Tabela 1: Modelo de observação do sector dos moldes em Portugal

Tipo de empresa	Área nuclear de competências	Origem das ideias	Tipo de inovação	Dependência dos clientes para o desenvolvimento de inovações	Relação com fornecedores de:			Relação com concorrentes[1]	Relação com instituições académicas	Relação com centros de investigação[2]	Relação com centros de formação
					Materiais[3]	Máquinas e equipamentos	Software				
Produtora - GECO - GRAMAQ - MATRISA - SOMEMA	Produção	Externa	Processo	Forte	Fraca	Forte	Forte	Forte	Fraca ou inexistente	Fraca ou inexistente	Forte
Comercial (*Broker*) - TECMOLDE	Comercialização	Externa	Processo e Produto	Forte	Forte	Forte	Forte	Inexistente	Inexistente	Inexistente	Inexistente
Design Industrial e Produção - IBEROMOLDES - SOCEM - VANGEST	Design e concepção de produto	Interna	Processo e Produto	Fraca	Médio	Forte	Forte	Forte	Forte[4]	Forte	Fraca

[1] Estas relações só ocorrem com alguns concorrentes e esta classificação foi feita com base neles

[2] Refere-se ao CENTIMFE e ao INETI

[3] Refere-se principalmente aos fornecedores de aço

[4] Pode considerar-se forte quando se compara com os restantes grupos, uma vez que nestes as relações existentes são quase nulas, mas continua a evidenciar-se um grande afastamento entre estas duas entidades, devido, acima de tudo, a barreiras culturais que estão a ser transpostas, ainda de forma muito tímida

3.1 Empresas Produtoras

As empresas que são produtoras por excelência têm uma elevada preponderância para desenvolverem inovações de processo, que advêm da sua forte dependência dos clientes em termos de concepção do molde que é, em geral, desenvolvido externamente, pelo cliente.

Neste tipo de empresas, as competências internas necessárias para a geração de ideias para a inovação de produto são escassas. A origem das ideias para a inovação centra-se, principalmente, nos clientes e, pontualmente, nos fornecedores de máquinas e de equipamentos. Na verdade, são as exigências dos clientes que obrigam, por seu turno, ao desenvolvimento do processo produtivo. Os clientes são os grandes impulsionadores da aquisição de novos equipamentos que permitem, através de novas e melhoradas técnicas de produção, a satisfação dos requisitos exigidos.

Como já foi referido, as relações com fornecedores de materiais (nomeadamente os fornecedores de aço) são muito reduzidas, não havendo neles uma fonte de informação significativa para a inovação, como se pode constatar no estudo de caso da GRAMAQ. Contudo, as relações são consistentes com as empresas fornecedoras de máquinas, equipamentos e software.

As visitas a empresas concorrentes estrangeiras, que possuem know-how superior, e que facultam períodos de formação de alguns recursos humanos, sob a iniciativa dos clientes, são frequentes e permitem que algumas destas empresas essencialmente produtoras, se vão actualizando em termos tecnológicos, como acontece com o caso da MATRISA.

Neste tipo de empresas, o relacionamento com centros de investigação é ténue ou mesmo inexistente, principalmente quando se trata de universidades, mas vão mantendo algumas relações com o CENTIMFE, nomeadamente na obtenção de alguma formação.

Estas empresas são típicas empresas subcontratadas, que vão evoluindo tecnologicamente em função das exigências dos seus clientes e do conhecimento transferido pelos fornecedores de máquinas e de equipamentos.

3.2 Empresa Comercial (*BROKER*)

Esta empresa funciona como um *broker*, uma vez que tem um comportamento misto relativamente aos outros dois grupos identificados (Produtora e Design Industrial e Produção). É uma empresa que não possui capacidade produtiva própria, subcontratando toda a sua produção a empresas exclusivamente produtoras (a primeira categoria), bem como àquelas de serviços especializados. Possui competências ao nível da engenharia e concepção do molde e muito conhecimento sobre o mercado, daí apresentar um elevado desempenho comercial. Face ao exposto, as suas características encontram-se entre as duas situações "extremas". No caso da TECMOLDE, como detém competências internas em termos de engenharia e concepção do molde, aproxima-se mais do terceiro grupo.

Este tipo de empresa é fulcral para a sobrevivência de muitas empresas do sector dos moldes, porque assegura duas funções que se revelam essenciais para muitas empresas. A primeira é uma função comercial, de ligação ao mercado externo, que é inexistente em muitas empresas do sector e, através desta função, assegura o escoamento e a exportação de uma parte substancial da produção de muitas empresas. A segunda função prende-se com a gestão de carteiras de encomendas[5] e a potenciação de economias de escala, através da coordenação da actividade produtiva de um conjunto de empresas. Esta vertente advém da incapacidade de resposta da maior parte das empresas a grandes encomendas e da necessidade de distribuir a produção da encomenda por várias empresas, de maneira a cumprir os prazos de entrega. Assim, esta empresa desenvolve capacidades de coordenação e de integração, ao distribuir produção e capacidade por uma rede de empresas produtoras, criando um formato organizacional que é equivalente aos propostos pelos conceitos de empresa "estendida" ou "virtual" (Browne e Zhang, 1999).

[5] A tendência do mercado passa cada vez mais pela procura de sistemas de moldes e não tanto de moldes isolados

3.3 Empresas com Design Industrial e Produção

As empresas deste grupo têm a componente de engenharia e concepção de moldes bastante desenvolvida, conjuntamente com a parte produtiva, além de possuírem competências emergentes no design e na concepção de produtos injectados, sendo, assim, capazes de fornecer soluções inovadoras aos seus clientes. A sua dependência dos clientes para o desenvolvimento de inovações é fraca, pois o seu know-how é suficientemente elevado para fornecer, autonomamente, as soluções que os clientes procuram.

Estas empresas mantêm relações relativamente consistentes com as instituições de educação e de investigação, apesar de considerarem que essas ligações deveriam ser mais frequentes. Recorrem com alguma regularidade às universidades, aos laboratórios de investigação e ao centro tecnológico, funcionando estes como parceiros privilegiados na resolução dos seus problemas técnicos. Como consequência deste comportamento, são empresas em que a realização de inovações de produto se torna mais importante, sem prejudicar a importância das inovações de processo. O know-how entretanto adquirido permite-lhes inovar já em termos de produto, adoptando uma estratégia de diversificação mais acentuada. Como exemplo deste tipo de empresas, é de referir os grupos IBEROMOLDES, SOCEM e VANGEST.

É curioso notar que o comportamento revelado pelo primeiro grupo, em termos de apropriação e criação de conhecimento, tem fortes semelhanças com a categoria intitulada *supplier dominated* proposta por Pavitt (1984), e o terceiro grupo de empresas tem semelhanças assinaláveis com o grupo *specialized suppliers*. O grupo *supplier dominated* incluía os sectores mais tradicionais, em que a actividade inovadora assentava principalmente em conhecimento gerado, externamente, noutros sectores. O grupo *specialized suppliers* incluía os fabricantes de máquinas, cujo know-how assentava, essencialmente, em design e em desenvolvimento próprio. A taxionomia de Pavitt procura determinar e compreender a diversidade comportamental em termos de aquisição e criação de tecnologia, mas estabelece a diferença entre sectores industriais. No caso deste estudo, o sector é só um, mas verificam-se diferenças assinaláveis a nível de empresas individuais. A associação de características e a sua generalização a um sector inteiro pode, assim, esconder uma realidade que parece ser bastante mais complexa e diversificada.

Apesar de as empresas terem comportamentos heterogéneos, que são determinados pela sua especialização, todas elas se relacionam com algumas empresas concorrentes, dependendo essencialmente do grau de confiança existente, e estabelecem com elas fortes ligações informais, desde visitas às suas instalações até ao relacionamento informal entre quadros, o que é uma característica particular deste sector. É importante salientar que esta troca de conhecimento não é feita entre todas as empresas concorrentes, mas apenas com aquelas que assegurem reciprocidade, i.e., quando se estabelece uma transacção de informação e conhecimento que beneficiará ambas as partes envolvidas (Von Hippel, 1988).

De uma maneira geral, pode afirmar-se que a maioria das empresas recorre mais às fontes de informação externas do que às internas, para obtenção de informação no seu processo de aprendizagem, com excepção daquelas que são fortemente inovadoras, dedicando-se não apenas a inovações de processo, mas também de produto. Essas empresas são mais selectivas e exigentes nas suas necessidades externas de conhecimento, como foi visto anteriormente.

No que diz respeito às fontes externas, estas configuram-se, principalmente, nos clientes, em alguns concorrentes e nos fornecedores de máquinas, equipamentos e software e, com menor importância, nos fornecedores de materiais. As instituições de ensino, de investigação e de formação têm um papel marginal, em quase todas as empresas estudadas, com excepção das empresas pertencentes ao terceiro grupo.

A informação disponível de forma generalizada, principalmente, as visitas a feiras e exposições, as mostras de produtos e os encontros são extremamente relevantes para todas as empresas, independentemente da sua posição hierárquica na cadeia de valor do sector em questão.

4 CONCLUSÕES

Este estudo permitiu verificar que o padrão de relacionamento destas empresas de moldes envolve outras empresas, nomeadamente, clientes, fornecedores (de máquinas e de equipamentos, de materiais e de software) e concorrentes, bem como instituições académicas (universidades e institutos superiores) e instituições do sector público (laboratórios, centro tecnológico: CENTIMFE - Centro Tecnológico da Indústria de Moldes, Ferramentas Especiais e Plásticos - e de formação: CENFIM - Centro de Formação Profissional da Indústria Metalúrgica e Metalomecânica); constatar que são fortes e consistentes as ligações que estas empresas têm com outras empresas, sejam elas clientes, fornecedores ou mesmo concorrentes, mas são extremamente fracas e por vezes inexistentes as relações que existem com as referidas instituições académicas e instituições do sector público e confirmar que as empresas mais inovadoras são as que mais se relacionam com as instituições académicas e instituições do sector público.

Com a informação recolhida das empresas estudadas, foi possível construir um modelo que permitisse a caracterização do sector dos moldes, relativamente ao tipo de relações existentes entre as empresas e os diferentes parceiros. O padrão de relacionamento é diferenciado consoante as características da empresa e consoante o seu posicionamento estratégico face à aquisição e criação de conhecimento.

Foi possível concluir que a grande vantagem competitiva deste sector da indústria nacional se deve à sua forte capacidade de se relacionar com outros parceiros, independentemente da sua posição na cadeia de valor. Mas estas ligações são diferenciadas, tanto em termos de parceiros como em termos de intensidade, verificando-se uma correlação entre a capacidade e o tipo de actividade inovadora e as interacções estabelecidas pela empresa.

BIBLIOGRAFIA

ARIAS, J. T. G. (1995), "Do networks really foster innovation?", Management Decision, Vol. 33, nº 9, pp. 52-56

BEEBY, M.; BOOTH, C. (2000), "Networks and inter-organizational learning: A critical review", The Learning Organization, Vol. 7, nº 2, pp.75-88

BROWNE, J.; ZHANG, J. (1999), "Extended and virtual enterprises – Similarities and differences", International Journal of Agile Management Systems, Vol.1, nº 1, pp. 30-36

CHUNG, S.; KIM, G. M. (2003), "Performance effects of partnership between manufacturers and suppliers for new product development: The supplier's standpoint", Research Policy, Vol. 32, nº 4, pp. 587-603

CHUNG, S. (2002), "Building a national innovation system through regional innovation systems", Technovation, Vol. 22, nº 8, pp. 485-491

DAVID, P. A. (1985), "Clio and the economics of QWERTY", American Economic Review Vol.75, nº 2, pp. 332-337

DOSI, G.; FREEMAN, C.; NELSON, R.; SILVERBERG, G.; SOETE, L. (1988), "Technical change and economic theory", Pinter Publishers, London

FREEL, M. S. (2003), "Sectoral patterns of small firm innovation, networking and proximity", Research Policy, Vol. 32, nº 5, pp. 751-770

FREEMAN, C.; SOETE, L. (1997), "The economics of industrial innovation", Pinter, 3rd edition

FREEMAN, C. (1987), "Technology and Economic Performance: Lessons from Japan", Pinter, London

KLINE, S. J.; ROSENBERG, N. (1986), "An overview of innovation", in Landau, R.; Rosenberg, N. (eds.), "The positive sum strategy – Harnessing technology for economic growth", The National Academy Press, Washington

LISSONI, F.; METCALFE, S. (1994), "Diffusion of innovation ancient and modern: A review of the main themes" in: Dodgson, M.; Rothwell, R., "The handbook of industrial innovation", Edward Elgar

LUNDVALL, B.-A. (1992), "National Systems of Innovation – Towards a theory of innovation and interactive learning", Pinter, London

MUSCIO, A. (2006), "From regional innovation systems to local innovation systems: Evidence from Italian industrial districts", European Planning Studies, Vol. 14, nº 6, pp. 773-789

NELSON, R. (1993), "National innovation systems: A comparative analysis", Oxford University Press

NONAKA, I.; TOYAMA, R.; BYOSIÈRE, P. (2001), "A theory of organizational knowledge Creation: Understanding the dynamic process of creating knowledge" in: Dierkes, M.; Antal B.; Child, J.; Nonaka, I. (eds.), "Handbook of organizational learning and knowledge", Oxford University Press, Oxford

OECD (1997), "Oslo Manual"

PAVITT, K. (1984), "Sectoral patterns of technical change: Towards a taxonomy and a theory", Research Policy Vol. 13. nº 6, pp. 343-373

SOUSA, A. (1997), "Sistema Português da Inovação – Imagem para um enquadramento político", FEP, Universidade do Porto, Porto

STONEMAN, P. (2001), "The economics of technological diffusion", Blackwells, Oxford

TEECE, D. J. (1998), "Capturing value from technological innovation: Integration, strategic partnering, and licensing decisions", in: Teece, D. J., "Strategy, technology and public policy", Edward Elgar Publishing

TEECE, D. J.; PISANO G.; SHUEN A. (1997), "Dinamic capabilities and strategic management", Strategic Management Journal, Vol.18, nº 7, pp. 509-533

TIDD J.; BESSANT J. (2009), "Managing innovation – Integrating technological, market and organizational change", John Wiley & Sons, 4th edition

VON HIPPEL, E. (1988), "Sources of innovation", Oxford University Press

INSERCIÓN LABORAL Y ESPECIALIZACIÓN EDUCATIVA EN ESPAÑA

ANTONIO CASQUERO TOMÁS
DOLORES GARCÍA CRESPO
Mª LUCÍA NAVARRO GÓMEZ
E-mail: l_navarro@uma.es
Universidad de Málaga

Abstract

La mayor parte de los análisis del mercado laboral confirman que los jóvenes constituyen uno de los grupos más vulnerables en términos de acceso al empleo. Aunque las elevadas tasas de desempleo juvenil son una característica común a los países de la UE, en España el problema alcanza mayor relevancia cuando se tiene en cuenta su carácter persistente y, además, que se produce en un contexto de elevada demanda educativa.

El objetivo de este trabajo es analizar la duración del primer periodo de desempleo de los jóvenes en España desde el punto de vista de su especialización educativa, en la medida en que éste puede ser un factor explicativo de la misma. Para ello, tras una breve descripción de la evolución reciente del mercado laboral por especialidades educativas según la CNED-2000, se estima el modelo de riesgo proporcional de Cox con el Módulo de Transición de la Educación al Mercado Laboral (INE, 2001), con objeto de examinar los determinantes de la inserción laboral de los individuos de edades comprendidas entre los 16 y 35 años. Los resultados obtenidos indican que en términos de acceso al primer empleo existe un efecto diferencial positivo asociado a la especialización cursada, aunque su influencia difiere según se hayan realizado estos estudios en la universidad o en la formación profesional.

Área temática: Economía de la Información y del Conocimiento.

1. INTRODUCCIÓN

La influencia contrastada del nivel de formación de la población en el proceso de inserción laboral en España (Ahn y Ugidos, 1995; Lassibille, et al., 2001) y la mayor disponibilidad de información a partir de la EPA de 2000 han significado un fuerte impulso del análisis de la transición educación-empleo en nuestro país (Albert, et al., 2003; Corrales y Rodríguez, 2003; Fernández y Jimeno, 2004; Blázquez, 2005; Rahona, 2006). Esta literatura viene a completar la que ha abordado el problema desde el ámbito regional (García-Montalvo, et. al., 1997; García-Espejo, 1998; Congregado y García-Pérez, 2002) o la referida a un nivel educativo concreto (Gil, 1999; Salas, 1999; González-Betancor, 2003).

Aunque las características del *Módulo de Transición de la Educación al Mercado Laboral* (INE, 2001), anexo a la EPA del segundo trimestre de 2000, limita el estudio de la inserción laboral a la década de los noventa al disponer de diez cohortes de individuos, de entre 16 y 35 años, ordenados según el año en el que dieron por finalizada su etapa formativa, sin embargo evita los problemas estadísticos que se derivan del enlace de encuestas por la construcción de pseudopaneles rotatorios (Bover, et al., 1996; Albert, et al., 2000) o aquellas limitaciones más genéricas que resultan del empleo de encuestas en principio no diseñadas para el citado fin (Lassibille, et al., 2001; Aguilar, 2005). Además, esta encuesta es la más actual que existe a nivel nacional sobre la inserción de los titulados de cualquier nivel educativo. Por estos motivos, ella será la que utilicemos para analizar los determinantes del primer periodo de desempleo de los jóvenes, que es el objetivo de este trabajo. Haremos especial hincapié en la contribución de la especialización educativa en la explicación del fenómeno de la inserción, puesto que este factor es muy determinante en el mismo (Fernández y Jimeno, 2004; Blázquez, 2005; Rahona, 2006).

La estructura de este trabajo es la siguiente. El segundo apartado describe el impacto de las especialidades educativas en el mercado laboral mediante la observación de los indicadores de

actividad, empleo y paro. En un tercer apartado se lleva a cabo el análisis causal de la inserción laboral estimando el modelo de riesgo proporcional de Cox, donde se incorporan aspectos personales, familiares y del entorno geográfico de los individuos. Los resultados obtenidos se comentan en el siguiente apartado y se finaliza con las conclusiones.

2. ESPECIALIZACIÓN EDUCATIVA Y MERCADO LABORAL

Teniendo en cuenta que los empresarios seleccionan a sus trabajadores según la educación que éstos poseen, y que, a su vez, éstos invierten en una determinada formación para mejorar su ventaja competitiva en el mercado laboral, vamos a describir en términos generales los indicadores del mercado laboral en función de las distintas especialidades educativas que poseen los individuos, con el objetivo de ver cuáles son las que presentan mayor empleabilidad en el mercado.

Para efectuar este análisis se ha elegido como inicio el año 2000, por ser el correspondiente a la encuesta de Transición que vamos a usar, y que además es el primero del que se dispone de información desagregada por sectores de estudio, según la CNED (INE, 2000), y como final el 2004, por ser el último año cuyos datos son considerados como definitivos. En relación al grupo de población joven seleccionado, es decir los individuos entre 16 y 35 años (ambos inclusive), es evidente que dicho intervalo resulta más amplio que el habitualmente utilizado para definirlo en términos demográficos, pero es acorde con el realizado por la encuesta.

Con los criterios señalados, las tablas 1 y 2 presentan las tasas de actividad, empleo[1] y paro por especialidades de estudio y sexo para los años 2000 y 2004, respectivamente. En ellas se observa que la evolución de la participación laboral de la población ha sido positiva, sobre todo la de los varones, lo que se traduce en mayores niveles de empleo y menores tasas de paro al final del periodo. En cuanto a la diferenciación por sexo, destaca la acusada reducción que experimenta en el periodo la tasa de paro femenina. Este hecho es mucho más apreciable si se tiene en cuenta que esta disminución se produce en un contexto de aumento de la tasa de actividad femenina[2]. Parece que el contrastado impulso de la demanda educativa de las mujeres que se ha producido en España en los últimos años comienza a tener respuesta en términos de mayores niveles de ocupación para este grupo de edad[3].

Tabla 1. Tasas de actividad, empleo y paro por sectores de estudio y sexo, población de 16 a 35 años (%). Año 2000.

Sectores de Estudio	*T. Actividad*			*T. Empleo*			*T. Paro*		
	Total	*Varones*	*Mujeres*	*Total*	*Varones*	*Mujeres*	*Total*	*Varones*	*Mujeres*
Formación básica	57,6	67,2	47,2	46,1	57,6	33,4	20,1	14,2	29,3
Docentes y Educación	81,7	83,6	81,2	61,8	67,1	60,4	24,3	19,7	25,6
Artes y Humanidades	74,1	76,6	72,4	57,4	63,9	52,9	22,5	16,6	27,0
C. Sociales. Admón. y Dº	79,7	82,5	78,3	63,8	70,5	60,5	20,0	14,5	22,7
Ciencias exactas	76,1	76,9	74,9	61,3	64,9	56,7	19,4	15,7	24,3
Ing., industria y const.	82,6	84,1	70,0	74,4	76,5	56,7	9,9	9,0	18,9
Agricultura y ganadería	78,0	81,6	69,6	64,9	71,3	50,0	16,8	12,7	28,2
Medicina y S. Sociales	80,1	76,7	80,9	60,9	62,4	60,5	24,0	18,6	25,2
Servicios	81,0	88,8	78,3	64,8	77,6	60,4	19,9	12,6	22,9
Sin clasificar	41,0	49,8	29,6	33,4	43,3	20,5	18,6	13,0	30,8
Medias	*64,8*	*71,6*	*58,0*	*52,3*	*61,8*	*42,5*	*19,4*	*13,6*	*26,6*
Población	*53.918*	*27.200*	*26.718*	*53.918*	*27.200*	*26.718*	*34.947*	*19.462*	*15.485*

Fuente: Elaboración propia a partir de la EPA 2º trimestre de 2000 (INE)

[1] Las tasas de empleo presentadas se corresponden con las que el INE denomina *tasa específica de empleo* y que define como el cociente entre el número de ocupados en un intervalo de edad considerado y la población en dicho intervalo.

[2] En cualquier caso, estas tasas son aún inferiores que las registradas por los varones.

[3] También este hecho puede estar relacionado con los cambios metodológicos experimentados en la EPA en este periodo.

Por sectores de estudio, la tasa de actividad aumenta en todas las especialidades educativas[4], tanto para los varones como las mujeres, siendo los estudios relacionados con las tecnologías (Ingenierías, industria y construcción) la formación que proporciona una incorporación mayor durante todo el periodo. En el extremo opuesto, las titulaciones conectadas al sector Servicios (transporte, medio ambiente, seguridad) son las que menos crecen en término de actividad.

Tabla 2. Tasas de actividad, empleo y paro por sectores de estudio y sexo, población de 16 a 35 años (%). Año 2004.

Sectores de Estudio	*T. Actividad*			*T. Empleo*			*T. Paro*		
	Total	*Varones*	*Mujeres*	*Total*	*Varones*	*Mujeres*	*Total*	*Varones*	*Mujeres*
Formación básica	61,3	71,7	49,5	51,7	63,3	38,5	15,7	11,8	22,1
Docentes y Educación	85,2	89,6	84,0	73,6	78,0	72,4	13,6	13,0	13,8
Artes y Humanidades	78,5	83,1	75,6	65,3	71,0	61,8	16,7	14,7	18,2
C. Sociales. Admón. y Dº	83,9	89,5	81,3	72,8	81,3	68,8	13,2	9,1	15,4
Ciencias exactas	81,3	83,1	79,1	69,8	72,7	66,2	14,1	12,6	16,3
Ing., industria y const.	90,3	91,5	80,9	83,3	85,3	67,4	7,7	6,8	16,7
Agricultura y ganadería	83,3	86,5	77,0	72,2	78,2	60,6	13,2	9,6	21,3
Medicina y S. Sociales	84,9	89,7	83,8	72,8	82,2	70,7	14,2	8,4	15,6
Servicios	81,6	89,2	78,6	68,6	79,7	64,2	15,9	10,6	18,3
Sin clasificar	36,2	41,7	31,0	27,4	32,9	22,3	24,2	21,1	28,2
Medias	*69,6*	*77,2*	*61,9*	*59,6*	*68,9*	*50,2*	*14,4*	*10,7*	*18,9*
Población	*58.398*	*29.341*	*29.057*	*58.398*	*29.341*	*29.057*	*40.628*	*22.653*	*17.975*

Fuente: Elaboración propia a partir de la EPA 2º trimestre de 2004 (INE)

Por su parte, la evolución de las tasas de empleo por sectores de estudio ayuda a describir el grado de utilización laboral efectiva que proporcionan las diferentes ramas de enseñanza. Así, en el año 2000 y para caso de los varones, los estudios de ingenierías y servicios son los que alcanzan las mayores tasas de empleo (77%), siendo destacable el acusado incremento que registran las tasas de empleo durante el periodo entre los profesionales de Medicina y Servicios Sociales. Esta última situación contrasta fuertemente con los resultados que para dichos estudios médicos se producen en términos de una menor demanda (Fernández, y Jimeno, 2003). En el caso del empleo femenino, el sector de Docentes-Educación, junto con los sectores de Ciencias Sociales y Medicina, es el que se caracteriza al inicio del periodo por un mayor nivel de empleabilidad (60%) y ha experimentado un mayor incremento en el periodo analizado (12 puntos porcentuales). También destaca la mayor importancia que adquiere el empleo femenino relacionado con las Ingenierías (con tasas de 57% y 67% en 2000 y 2004, respectivamente), espacio del mercado laboral tradicionalmente reservado para los varones.

Por último, la evolución de las tasas de paro por especialidades, durante el periodo, señala la mayor disminución del paro para ambos sexos en los sectores Docentes-Educación y Medicina (entre 6 y 12 puntos porcentuales).

3. ESPECIALIZACIÓN EDUCATIVA Y PRIMER EMPLEO

La fuente de datos utilizada para estudiar la inserción de los jóvenes es el Módulo de transición de la educación al mercado laboral (INE, 2001), anexo a la EPA del segundo trimestre de 2000, que genera información específica sobre los procesos de transición en veinte países europeos (Ianelli, 2002).

El referido módulo se dirige a la población comprendida entre 16 y 35 años, que en el transcurso de los diez años anteriores a la encuesta, esto es entre 1991 y 2000, interrumpieron durante más de un año sus estudios, lograron finalizarlos o bien decidieron abandonarlos de

[4] Excepto en el grupo Sin clasificar.

manera definitiva. El enlace del módulo *ad hoc* a la estructura habitual de la EPA amplía la información de aquél con todas las cuestiones personales y familiares procedentes del cuestionario habitual de la EPA, que serán utilizados en el análisis empírico.

3. 1 Análisis descriptivo

Por lo general, en el estudio de la inserción laboral la variable de interés es el tiempo transcurrido desde que el joven sale del sistema educativo hasta que encuentra un primer empleo significativo[5]. En este sentido, y con el fin de describir el comportamiento en el mercado laboral de los individuos de acuerdo a sus especialidades educativas, la tabla 3 refleja un indicador de inserción rápida al empleo por sectores de estudios; es decir, el porcentaje de individuos que encuentra un empleo significativo durante el primer año tras su salida del sistema educativo. Igualmente se recoge la duración media del primer periodo de desempleo. Ambos aspectos se observan para todos los niveles educativos, así como para los individuos con estudios terminados de Formación Profesional de 2º grado y de universidad, al ser ambos niveles los que otorgan verdaderas especialidades profesionales.

En la tabla 3 se observa que los jóvenes tardan, por término medio, 40,36 meses[6] en acceder al primer empleo significativo, independientemente del nivel de estudios realizados. Esta duración se reduce en ocho meses si han realizado estudios universitarios o de formación profesional. Por otra parte, durante el primer año de búsqueda desde la salida del sistema educativo sólo el 21,4% de los jóvenes encuentran un empleo, elevándose hasta el 26% y el 28% para los titulados de formación profesional y los universitarios, respectivamente.

Además, estas bajas tasas de inserción laboral no son homogéneas ni por especialidades ni por niveles de estudios. Así, excepto Ingenierías y Ciencias Exactas, que se sitúan por encima del nivel medio de inserción durante el primer año (por ejemplo, el 42% y 29% en el caso de la universidad), el resto de sectores educativos muestran un comportamiento prácticamente inverso según sean de tipo profesional o universitario. Algo similar ocurre cuando se analiza la duración del primer periodo de desempleo, destacando como casos extremos la situación favorable de los estudios universitarios de Ingenierías (8 meses menos que la duración media de su grupo) frente a Agricultura y Ganadería en el nivel de estudios profesionales (9 meses más que la media de su grupo).

[5] Según el INE un empleo de al menos 6 meses de duración y 20 horas semanales de trabajo. En el contexto de la población juvenil, esta definición de empleo puede resultar restrictiva, puesto que esta población mayoritariamente accede al empleo mediante contratos en prácticas, de aprendizaje, para obra, servicio o por temporadas, es decir fórmulas contractuales que en buena medida aquí no se consideran empleos (significativos).

[6] En los datos ofrecidos por el INE para España (INE, 2001) la duración media de desempleo estimada es de 28,6 meses, si bien el resultado no es comparable puesto que el INE, a diferencia de esta investigación, censura a los individuos que finalmente no encuentran un empleo significativo.

Tabla 3. Porcentaje de individuos de 16 a 35 años con un empleo significativo en el primer año y duración en meses del primer periodo de desempleo, por sectores de estudio*

Sectores de estudio	*Todos los niveles educativos*		*FP(1)*		*Universidad(2)*	
	Encuentran un empleo en el primer año	Duración del desempleo	Encuentran un empleo en el primer año	Duración del desempleo	Encuentran un empleo en el primer año	Duración del desempleo
Formación básica(3)	14,9	48,56	0,01	54,01	0,01	25,37
Docentes y Educación	20,7	37,71	33,3	30,03	20,5	36,95
Artes y Humanidades	25,2	34,78	27,5	33,20	23,7	36,08
C. Sociales. Admón. y Dº	24,3	34,73	23,6	37,23	26,2	31,66
Ciencias exactas	27,8	32,42	28,8	33,.65	28,8	31,04
Ing., industria y const.	31,5	32,43	29,0	32,54	41,6	24,26
Agricultura y ganadería	34,6	30,40	21,1	41,65	37,8	29,57
Medicina y S. Sociales	24,2	32,69	17,6	34,66	27,3	31,57
Servicios	25,1	32,49	30,4	31,74	26,7	28,17
Sin clasificar	16,1	45,74	23,1	42,24	0,0	27,45
Medias	*21,2*	*40,36*	*25,8*	*32,69*	*27,9*	*32,49*
Nº de individuos	*12.300*	*12.300*	*2.012*	*2.012*	*3.321*	*3.321*

Fuente: Elaboración propia a partir del Módulo EPA 2º Trimestre de 2000 (INE).
*Individuos que salieron del sistema educativo entre 1991 y 1999.
(1) Se refiere sólo a la FP de segundo grado.
(2) La Universidad recoge los ciclos corto y largo.
(3) Los resultados para este sector de estudios no son interpretables desde el punto de vista de la especialización, ya que casi el 100% de los individuos que se incluyen en él carecen de ella.

3. 2 Análisis explicativo

La variable dependiente en el análisis es el número de meses trascurridos desde que el joven interrumpe, abandona o finaliza sus estudios hasta que obtiene su primer empleo significativo. Con objeto de incorporar variables explicativas en el análisis y evaluar en qué medida influyen en la tasa de salida del desempleo, se estima el modelo de riesgo proporcional de Cox (1972), cuyos resultados se presentan en la tabla 4.

Los valores obtenidos tras la estimación del modelo de riesgo proporcional de Cox muestran, en primer lugar, que los varones tienen una menor duración en el primer periodo de desempleo. En cuanto a las otras características personales, destaca que las variables educativas muestran su efecto positivo para la inserción laboral a través de la especialización cursada en el sistema de enseñanza, aunque difiere su influencia según se realice o no en la universidad dicha especialización.

Tabla 4. Estimación del modelo de riesgo proporcional de Cox[a]

Características	*Coef.*	*D. E.*	*Exp(β)[b]*
Varón	0,221*	0,039	1,24
Formación Profesional (FP)	-0,055	0,047	
Sectores educativos			
Docentes y Educación	-0,189**	0,098	0,82
Artes y Humanidades	-0,247*	0,095	0,78
C. Sociales (Admón. y Derecho)	0,048	0,065	
Ingenierías, Industria y Const.	0,225*	0,082	1,25
Agricultura y Ganadería	0,109	0,139	
Medicina y S. Sociales	-0,056	0,087	
Servicios	0,004	0,126	
Sin clasificar	0,349	0,415	
*Sectores educativos * FP*			
Docentes y Educación * FP	0,201	0,291	
Artes y Humanidades * FP	0,312*	0,312	1,36
C. Sociales (Admón. y Dº)* FP	-0,134*	0,060	0,87
Ingenierías, Ind. y Const.* FP	-0,173*	0,079	0,84
Agricultura y Ganadería * FP	-0,346	0,298	0,70
Medicina y S. Sociales * FP	-0,208*	0,123	0,81
Servicios * FP	-0,149	0,161	
Sin clasificar * FP	-0,457	0,470	
Características familiares			
Padre con estudios medios	-0,030	0,049	
Padre con estudios superiores	0,108*	0,049	1,11
Madre con estudios medios	0,230*	0,057	1,25
Madre con estudios superiores	0,191*	0,052	1,21
Padre parado	-0,255*	0,108	0,77
Padre inactivo	-0,098*	0,042	0,90
Madre parada	-0,473*	0,084	0,62
Madre inactiva	-0,030	0,038	
Tasa paro prov. juvenil > media nacional	-0,024*	0,004	0,977
Regiones			
Noroeste	0,006	0,067	
Noreste	0,206*	0,079	1,22
Madrid	-0,073	0,083	
Centro	0,128*	0,065	1,13
Este	0,191	0,083	
Canarias	0,171**	0,094	1,18
Año Salida sistema educativo			
1992	0,088	0,078	0,97
1993	0,168*	0,079	1,18
1994	0,364*	0,078	1,43
1995	0,639*	0,078	1,89
1996	0,848*	0,079	2,33
1997	1,118*	0,080	3,05
1998	1,348*	0,084	3,85
1999	1,771*	0,090	5,87
Razón de verosimilitudes			*1432*
N			*5333*

Nota: *significativo al 5%; ** significativo al 10%.

(a) El individuo de referencia es una mujer del sector de ciencias exactas, cuyos padres están ocupados, no tienen estudios o sólo cuentan con estudios primarios, reside en una comunidad autónoma de la zona sur y abandonó el sistema educativo en 1991.

(b) Se ha calculado Exp (β) para los coeficientes significativamente distintos de cero.

Así, respecto a un joven del sector de Ciencias Exactas[7], sólo las especialidades de Ingenierías –tanto universitarias como de FP- y los titulados en formación profesional en Artes y Humanidades presentan una transición más rápida hacia el empleo, cuando estas últimas especialidades en la universidad tiene un efecto negativo, como se espera *a priori*. Quizás la explicación al hecho aparentemente paradójico de Artes y Humanidades se encuentre en la mayor homogeneidad de dichos estudios en FP. Los universitarios del sector Docentes-Educación, y los titulados en FP en las áreas de Ciencias Sociales y Medicina tienen un acceso más lento al empleo que la media. Para el resto de titulaciones no se aprecia un comportamiento diferenciado.

Por otro lado, respecto a las características familiares, tal y como se obtienen en otros trabajos (Casquero y García, 2006) el nivel de estudios del padre y de la madre continúa siendo un factor importante a la hora de explicar el acceso al empleo de los hijos, al igual que la actividad laboral de los progenitores. En concreto, un mayor capital humano de los padres incrementa la probabilidad de acceso al primer empleo significativo entre un 11% y un 25%, mientras que una situación de paro o inactividad lo reduce entre un 10% y un 38%.

En cuanto a las variables que recogen la región de residencia del joven, se observa un efecto diferencial en las zonas Noreste, Centro, Este y Canarias, donde se accede antes al empleo que en la zona Sur, probablemente debido a su mayor desarrollo económico o turístico no estacional[8]. El coeficiente relativo a la tasa de paro provincial, por su parte, muestra un signo negativo, indicando que si el individuo se encuentra en una provincia con una tasa de desempleo juvenil superior a la media, su probabilidad de acceso al empleo se reduce en un 3%. Por último, las variables ficticias correspondientes el año de salida del individuo del sistema educativo no permiten ser interpretadas con sentido cíclico, pues reflejan las duraciones diferentes en el mercado laboral de los jóvenes y se introducen en la estimación para controlar esta estructura de la muestra.

4. CONCLUSIONES

El objetivo de este trabajo es poner de manifiesto que, desde el punto de vista del acceso al primer empleo, no es tan importante que el joven haya cursado estudios universitarios o de formación profesional como el sector educativo en el que se ha especializado. Así, destacan las Ingenierías en ambos niveles educativos y Artes-Humanidades en formación profesional como las vías de acceso más rápidas al empleo, seguidas de las de Ciencias Exactas que, en definitiva, refleja la gran demanda de los estudios de Informática que están englobados bajo esta última disciplina.

Las características familiares también muestran un papel determinante en el proceso de transición del sistema educativo al mercado de trabajo. Así, un mayor nivel de estudios de los padres favorece el acceso al empleo, mientras que una situación de inactividad o paro lo retrasa. Por último, el mercado local en el cual el joven inicia la búsqueda de empleo afecta a su inserción, confirmando, al igual que en el caso anterior, los resultados de otros estudios previos.

5. REFERENCIAS BIBLIOGRÁFICAS

AGUILAR, M. I. (2005): *La inserción laboral de los jóvenes en España. Un enfoque microeconométrico.* Estudios y Monografías de Economía, Editorial Thomson-Civitas.

AHN, N y UGIDOS, A (1995): "Desempleo juvenil en España: ¿qué determina su duración?". *Ekonomiaz* nº 31-32, pp. 247-265.

[7] Se ha escogido este sector de estudios como referencia porque su duración media de acceso al empleo coincide con la media muestral.

[8] El efecto positivo de la región Centro puede quizás deberse a que engloba municipios que actúan como ciudades dormitorio de Madrid.

ALBERT, C., JUAREZ, J., SACHEZ, R, y TOHARIA, L. (2000): "La transición de la escuela al trabajo en España: años noventa" *Papeles de Economía española*, nº 86, pp.42-58.

ALBERT, C., DAVIA, M. A, y TOHARIA, L (2003): "The find or not to find significant job: the case of Spain", *Actas de las V Jornadas de Economía Laboral*, Universidad Rovira i Virgili, Reus, 9-11 de Julio.

BLÁZQUEZ, M. T. (2005): "Youth labour market integration in Spain: Search, time, job duration and skill mismatch". *Spanish Economic Review,* 7, pp. 191-208.

BOVER, O., ARELLANO, M. y BENTOLILA, S. (1996): "Duración del desempleo, duración de las prestaciones y ciclo económico", *Estudios Económicos* nº 57*,* Servicio de Estudios del Banco de España, Madrid.

CASQUERO, A. y GARCÍA CRESPO, D. (2006). "La inserción laboral de la formación profesional en España". XV Reunión de AEDE, Granada.

CONGREGADO, E. y GARCÍA PÉREZ, J. L (2002): "El problema del desempleo en la economía andaluza (1990-2001): análisis de la transición desde la educación al mercado laboral". *Documento de Trabajo E2002/17, Serie Economía*. Fundación Centro de Estudios Andaluces (CENTRA), Sevilla.

CORRALES, H y RODRIGUEZ, B. (2003): "La transición del sistema educativo al merado laboral. Análisis de los factores determinantes del primer desempleo". *Actas de las V Jornadas de Economía Laboral*, Universidad Rovira i Virgili, Reus, 9-11 de julio.

COX, D. (1972): "Regression models and life-tables (with discussion)", *Journal Royal Statistic Society*, B.34, pp. 187-220.

FERNÁNDEZ, C. y JIMENO, J. F. (2004): "Educación e inserción laboral", *Papeles de Economía española*, nº 100, pp.112-122.

GARCÍA ESPEJO, M. I. (1998): *Recursos formativos e inserción laboral de los jóvenes.* Colección Monografía. Centro de Investigaciones Sociológicas. Madrid.

GARCÍA MONTALVO, J. PEIRÓ, J. M y SORO, A. (2003): *Capital humano. Observatorio laboral de los jóvenes: 1996-2002*. Fundación Bancaja-IVIE. Valencia.

GIL, J. A (1999): *Los titulados universitarios y el mercado de trabajo. Un estudio del desajuste educativo*. La Caja de Canarias. Las Palmas de Gran Canaria.

GONZÁLEZ BENTANCOR, S. (2003): Inserción Laboral, Desajuste Educativo y Trayectorias Laborales de los Titulados en Formación Profesional Específica en la isla de Gran Canaria (1997-2000). Tesis doctoral. Universidad de Las Palmas de Gran Canaria.

IANELLI, C. (2002): *Evaluation and analyses of the LFS 2000 ad hoc module data on scholl-to-work transitions. Report on data quality and cross-country comparability,* Centre for Educational Sociology, Universty of Edinburg.

INE (2001): *Encuesta de Población Activa. Módulo de transición de la educación al mercado laboral*. Instituto Nacional de Estadística, Segundo trimestre de 2000. Madrid.

LASSIBILLE, G., NAVARRO, Mª. L., AGUILAR, M. I. y DE LA O, C. (2001): "Youth transition from school to work in Spain", *Economics of Education Review*, vol. 20, nº 2, pp.139-149.

RAHONA, M .(2006): "¿La posesión de un título universitario facilita el acceso de los jóvenes al primer empleo?. Una aproximación para el caso español". *Revista del Ministerio de Trabajo y Asuntos sociales*, *Economía y Sociología,* nº 61, pp. 105-121.

SALAS, M. (1999): Inversión en capital humano: demanda, búsqueda de empleo y rendimientos privados de la educación superior. Tesis doctoral. Universidad de Granada.

A INTERACÇÃO ENTRE A GESTÃO ESTRATÉGICA E A GESTÃO DO CONHECIMENTO

ANTÓNIO ANACLETO VIEGAS FERREIRA
E-mail: barquiconta@sapo.pt
Instituto Politécnico de Tomar
JOÃO JOSÉ MATOS FERREIRA
E-mail: jjmf@ubi.pt
Universidade da Beira Interior
JORGE MANUEL MARQUES SIMÕES
E-mail: jorgesimoes@ipt.pt
PEDRO AUGUSTO NOGUEIRA MARQUES
E-mail: pedromarques@ipt.pt
Instituto Politécnico de Tomar

Resumo

Este artigo discute a interacção entre a gestão estratégica e a gestão do conhecimento (GC). Assume que o conhecimento é uma capacidade de actuar, em que o conhecimento não se pode descrever com palavras já que é essencialmente tácito, e ao mesmo tempo dinâmico e estático, sendo a gestão do conhecimento um processo colectivo e interactivo que envolve as seguintes actividades: geração do conhecimento; codificação do conhecimento; e transferência do conhecimento. É relevante estudar as várias actividades da gestão do conhecimento pois, a gestão do conhecimento é um processo sistemático, articulado e intencional, apoiado na criação, na codificação e na transferência do conhecimento, com o objectivo de atingir a excelência organizacional, permitindo assim atingir uma vantagem competitiva sustentável. Assim, neste artigo iremos investigar as várias actividades da GC, em termos empíricos, para aprofundar os conceitos envolvidos.

Palavras Chave: gestão estratégica, gestão do conhecimento, actividades do conhecimento.
Área Temática: Economia da Informação e do Conhecimento.

Abstract

This article discusses the interaction between strategic management and knowledge management. Assumes that knowledge is an ability to act, on that knowledge can not be describe with words since it is essentially tacit, and at the same time dynamic and static, being the knowledge management a collective and interactive process that involves the following activities: generation of knowledge; codification of knowledge and transfer of knowledge. It is important to study the various activities of knowledge management because the management of knowledge is a systematic process, and intentionally articulated, supported in the creation, in the codification, and transfer of knowledge, in order to achieve organizational excellence, aiming to a sustainable competitive advantage. Thus, in this article we develop the various activities of the GC in empirical terms, to deepen the concepts involved.

Key Words: strategy management, knowledge management, knowledge activities.
Thematic Area: Information and Knowledge Economy.

1. INTRODUÇÃO

Na literatura sobre a gestão estratégica, o conhecimento baseado nas organizações altera o foco dos recursos do conhecimento e propõe que o conhecimento é o recurso mais importante para criar uma vantagem competitiva sustentável. Mas, nem todas as actividades da gestão do conhecimento foram provadas como influências positivas para a performance empresarial ou

como resultado de uma vantagem competitiva (Greiner *et al.,* 2007). Inclusivamente, Learned *et al.* (1969), já defendia, na altura, que na formulação de vantagens estratégicas, as competências distintivas, os recursos fortes e as oportunidades é que determinam a sustentabilidade das vantagens competitivas, sendo, segundo Dinis (2004), o sucesso de uma empresa dependente da criação de vantagem competitiva que, por sua vez, resulta da capacidade que a empresa tem de promover inovações. Para obter sucesso no século XXI, as empresas necessitam de apostar no conhecimento constantemente (Winkelen *et al*., 2008), principalmente nos países mais desenvolvidos, onde fabricar os produtos *standards* usando métodos standards não sustenta vantagens competitivas.

Nesse seguimento o que se entende por gestão do conhecimento e quais as actividades que compõem o mesmo? Assim, neste artigo iremos aprofundar as várias actividades da GC, em termos empíricos, para aprofundar os conceitos envolvidos, sendo a gestão do conhecimento entendida como um processo colectivo, de natureza interactiva, que pressupõe uma partilha de informação e atitudes nas várias fases do desenvolvimento do conhecimento (Simões, 2008). A gestão do conhecimento, vista como um conjunto de processos que governa a criação, a disseminação e a utilização do conhecimento para atingir plenamente os objectivos da organização, é uma nova área na confluência entre a tecnologia da informação e da gestão, um novo campo entre a estratégia, a cultura e os sistemas de informação de uma organização. Com o enfoque da gestão do conhecimento começa-se a rever a empresa, as suas estratégias, a sua estrutura e a sua cultura. Segundo Loureiro (2003, p. 72-73), para transformar o conhecimento em activo organizacional com valor, "o conhecimento, a experiência e a perícia devem ser formalizados, distribuídos e aplicados". A gestão do conhecimento é considerada a parte fundamental da estratégia para usar conhecimento na criação de vantagens competitivas sustentáveis no actual ambiente de negócio (Bloodgood e Salisbury, 2001; Loureiro, 2003; Sher e Lee, 2004; Grant 2006; Gray e Meister, 2006; Greiner *et al.,* 2007).

Para Boisot (2002: 65) "a maioria dos desafios postados pela efectiva gestão de recursos do conhecimento não é particularmente nova. Eles têm, na realidade, estado connosco desde a revolução científica do século XVII, se não mesmo antes". As sociedades instruídas do século XVII, a Académia Lincei em Roma (fundada em 1603), a Royal Society em Londres (fundada em 1660), e a Académie des Sciences em Paris (fundada em 1666), estavam todas preocupadas com a rotinização da descoberta e estavam todas prontas para a disseminação do conhecimento útil (Boisot, 2002). Segundo Sprat (1662), referido por Boisot (2002: 65), os cientistas europeus da época estavam a lutar com o mesmo tipo de questões que confrontam a gestão do conhecimento hoje: como gerar conhecimento que é válido e ao mesmo tempo esperançosamente útil, como o compartilhar, e como o manter em contacto com um outro, bem como também para o actualizar.

Assim, a gestão do conhecimento será um "processo sistemático, articulado e intencional, apoiado na criação, codificação, disseminação e apropriação de conhecimentos, com o propósito de atingir a excelência organizacional. Na prática, a gestão do conhecimento consiste na identificação e no mapeamento dos activos de conhecimento da organização, divulgando e gerando novos conhecimentos para a vantagem competitiva e partilhando as melhores práticas e tecnologias que impulsionarão estes processos" (Serrano e Fialho, 2003: 126). Logo, as actividades principais para a gestão do conhecimento, de acordo com Silva e Neves (2003) e Franco e Mariano (2007), prendem-se com a geração (criação ou aquisição), codificação (ou armazenamento) e transferência do conhecimento. Porém, existem autores que defendem que os níveis de conhecimento das organizações são influenciados por quatro processos: criação do conhecimento, absorção de conhecimento público pertinente (externo), conhecimento transferido dentro da organização e disseminação no ambiente externo da organização do conhecimento pertencente à mesma. Cada um destes processos afecta o nível de conhecimento global da organização (Matusik, 2002: 606).

Porém é de salientar que, na gestão do conhecimento, as qualidades do gestor, de acordo com Sveiby (2000), só se adquirem mediante a prática. Ninguém pode chegar a mestre sem ter passado primeiro pela fase de aprendiz. Este tipo de aprendizagem, através da prática, existe em todas as profissões e as regras são mais ou menos universais. A tradição estende-se para além

dos organismos e das nacionalidades, permite que os profissionais falem uns com os outros e deste modo consigam fazer circular eficazmente o conhecimento. Mas, "ainda que a tradição seja mais lenta que a informação na transmissão dos acontecimentos, é um meio muito eficaz para a transmissão de competências já que utiliza todos os sentidos, inclusive o da intuição. Muitas vezes, basta trabalhar junto a alguém que sabe mais do que nós para aprender; a transferência efectua-se automática e inconscientemente" (Sveiby, 2000: 92).

"Como mercados para bens e serviços, o mercado do conhecimento tem compradores e vendedores que negociam para alcançar um preço mutuamente satisfatório para a troca. Tem corretores que reúnem os compradores e vendedores e até intermediários que usam o mercado do conhecimento para criar as suas bases de poder internas" (Davenport e Prusak, 1998; citados por Empson, 1999: 68).

O objectivo deste artigo será esclarecer as várias fases da gestão do conhecimento, de forma empírica, fazendo uma abordagem prévia das estratégias competitivas, para no futuro ser efectuado uma investigação relacionado ambos os conceitos nas instituições de ensino superior militar. Este assunto é relevante, dado a gestão do conhecimento ter sido vista como um assunto operacional e não como um assunto estratégico. Assim, neste trabalho, iremos abordar os seguintes assuntos: as estratégias competitivas, na secção 2.1; a geração do conhecimento, na secção 2.2; a codificação do conhecimento, na secção 2.3; e a transferência do conhecimento, na secção 2.4. O trabalho terminará, na secção 3, com uma proposta do modelo conceptual e, na secção 4, com uma conclusão.

2. REVISÃO DA LITERATURA

2.1. Estratégias Competitivas

A função do general do exército (Ghemawat, 2000; Freire, 2006) significa em grego *strategos*, sendo esta a origem do termo estratégia. O termo anteriormente mencionado, estratégia, esteve sempre presente na gestão desde a origem da civilização, surgindo apenas investigações científicas sobre este termo, durante o século XX, mais concretamente na década de sessenta com Chandler (1962). Este autor considerava que a estratégia é a determinação dos fins e dos objectivos de longo prazo de uma empresa e os programas de acção e a alocação de recursos para atingir as metas traçadas. Evoluindo, em termos de conceito, Glueck (1976), assumiu que estratégia era um plano integrado, unido e compreensivo desenhado para assegurar que os objectivos básicos da organização eram atingidos (Bracker, 1986). Em 1979, Mintzberg assumiu que a estratégia era uma força mediática entre a organização e o ambiente, sendo as decisões estratégicas da empresa suportadas nas expectativas das alterações do citado ambiente (Bracker, 1986). Porter, em 1996, considerava que a estratégia significa fazer escolhas de posicionamento competitivo. Actualmente, com os autores Ocasio e Joseph (2008), estes entendem a estratégia como uma estrutura, implícita ou explícita, a qual direcciona uma organização na selecção das suas acções.

Segundo Hofer e Schendell (1978) e Digman (1990), todas as estratégias deveriam conter 4 componentes:

(1) O âmbito ou domínio de acção pelo qual a organização pretende alcançar os seus objectivos;
(2) As competências e os recursos que a organização irá usar para alcançar os objectivos;
(3) As vantagens que a organização espera alcançar sobre os seus competidores através da utilização eficaz dos seus recursos e competências (vantagens competitivas);
(4) As sinergias que resultarão da forma como a organização utilizará as suas competências e recursos.

Neste seguimento, Porter (1980) apresentou as tipologias de estratégias genéricas, visando dessa forma ser uma alternativa de orientação estratégica para atingir vantagens competitivas sustentáveis: liderança de custos, diferenciação do produto e enfoque. Segundo Barbosa (2006), a selecção de entre uma das estratégias de Porter tem implicações directas na escolha da

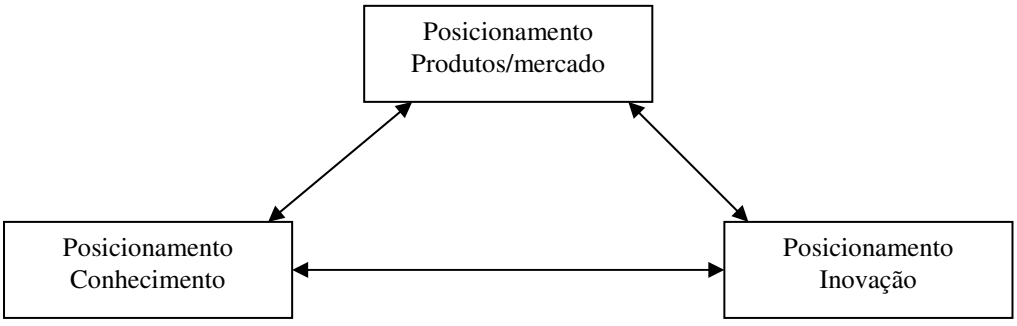
estratégia tecnológica (conhecimento), sendo as "vantagens competitivas", segundo Freire (2006), designada por: uma empresa possui uma "vantagem competitiva" quando a sua taxa de rentabilidade a longo prazo é superior à média da indústria num dado mercado ou segmento de mercado. Uma empresa poderá possuir várias vantagens sobre outra empresa tais como um sistema de produção superior, nível inferior de custos salariais ou uma capacidade para fornecer um serviço superior ao cliente, mas as vantagens competitivas são aquelas às quais os clientes atribuem valor (Fahy e Smithee, 1999). É esta a visão do conceito adoptada na investigação.

Inclusivamente, os autores Prahalad e Hamel (1990), Barney (1991) e Grant (1991) realçaram que os recursos, por si só, são insuficientes para a obtenção de uma vantagem competitiva ou para um crescimento sustentável. Segundo estes investigadores, é necessário que os recursos sejam transformados em capacidades, capazes de influenciar positivamente o crescimento. Logo, será necessário conhecimento para o efectuar e a organização deverá possuir uma forma, modelo, de gerir as diversas actividades do conhecimento inerentes para alcançar uma vantagem competitiva ou um crescimento sustentável. Assim, Mcdonough III *et al.* (2008) referiram que uma estratégia eficiente é constituída por três componentes chaves (ver figura 1): O posicionamento dos produtos/mercado; O conhecimento; A inovação.

Por outro lado, vários investigadores têm definido a estratégia como um processo: (i) continuo de experiência, investigação, pesquisa e exploração que resulta em novos produtos, processos e formas de organização (Lundvall, 1992); (ii) não linear, evolucionário, complexo e interactivo de aprendizagem e de relacionamentos entre a empresa e o seu meio envolvente (Silva, 2003). Para este investigador os resultados deste processo designa-se por capacidade empresarial (inovação/conhecimento). A capacidade inovadora/conhecimento, segundo Silva (2003), compreende a criação e lançamento de novos produtos ou processos e as melhorias tecnologicamente significativas introduzidas nos produtos ou processos, bem como abarca novas formas de negócio, de organização de trabalho, de gestão das empresas e de relacionamentos internos e externos. Em 1979, Freeman, sugere uma nova abordagem da inovação/conhecimento com a visão interactiva do processo da inovação. Nesta nova abordagem os principais factores impulsionadores da inovação/conhecimento residem nas oportunidades científicas e tecnológicas, combinadas com as necessidades económicas que emergem do mercado e da sociedade (Silva, 2003). Segundo Freeman (1979), existe um conjunto de mecanismos e relações interactivas que suportam os processos de inovação/conhecimento, com contribuições de conhecimentos científicos e tecnológicos, experiências de produção, utilização de equipamentos complexos e políticas governamentais.

Consequentemente, quando o meio envolvente se altera, as empresas necessitam de rever continuadamente o posicionamento deste alinhamento. Para Mcdonough III *et al.* (2008), a competição das empresas não se baseia, apenas, nos produtos e serviços oferecidos mas, também, nos seus conhecimentos e na forma como inova, sendo estes aspectos o que representa a posição competitiva das empresas, que deve ser avaliada tendo em conta as suas capacidades. Como as capacidades estão inevitavelmente sujeitas à imitação, à substituição por inovação ou a tornarem-se obsoletas pelas alterações do mercado, as empresas devem ser capazes de actualizar as suas capacidades se quiserem manter as suas vantagens competitivas.

Figura 1 – Alinhamento estratégico para obter vantagens competitivas



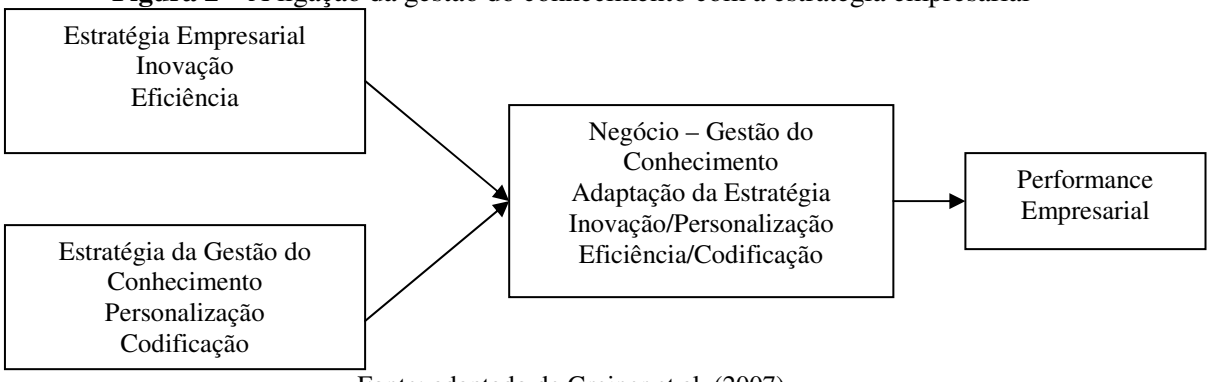
Fonte: adaptado de McDonough III *et al.* (2008: 56)

(1) Segundo Tallman (2006), o posicionamento dos produtos/mercado somente fornece vantagens competitivas temporárias, assim como podem ser copiadas se forem bem sucedidas. Sendo assim a organização deve alinhar o posicionamento dos produtos/mercados com o posicionamento da inovação (Mcdonough III *et al.,* 2008).

(2) Actualmente o conhecimento tem sido explicitadamente entendido como tendo valor e por isso um recurso que deve ser gerido. Assim, Winkelen *et al.* (2008) referiram que uma gestão eficiente dos conhecimentos suporta a capacidade das empresas para serem inovadoras e consequentemente trazer produtos novos ou melhorados para o mercado. A importância da partilha e transferência dos conhecimentos nas empresas como estimulante na criação de novos conhecimentos foi descrita por Nonaka em 1991. No entanto, a gestão do conhecimento tem sido visto como um assunto operacional e não como um assunto estratégico. Para Zack (1999), a ligação entre conhecimento e estratégia raramente tem sido efectuada de forma explícita e as diferenças entre os conhecimentos que a organização possui e os que necessita para competir com sucesso cria um *gap* de conhecimento estratégico que deve ser eliminado. Assim, Greiner *et al.* (2007), demonstram, esquematicamente, a ligação da gestão do conhecimento com a estratégia empresarial, ver figura 2.

(3) A inovação normalmente sugere o desenvolvimento de novos produtos, mas as empresas podem, também, competir baseando-se na inovação em diversas áreas, reflectindo os seus pontos-chaves estratégicos, incluindo a liderança tecnológica, novos e melhores serviços, baixos preços, melhor execução operacional e melhor entendimento dos clientes e mercado (McDonought III *et al.,* 2008). O posicionamento da inovação nas empresas especifica o quanto se irá focalizar no desenvolvimento de inovações externas que advém das experiências directas dos clientes, por exemplo através dos produtos e serviços, ou inovações internas usadas somente na organização como novos processos e procedimentos. O posicionamento da inovação também reflecte a extensão da inovação, desde um aumento refinado às mudanças radicais, e o grau de novidade para os clientes/mercado que se deseja incorporar nas inovações externas. No entanto, na análise do posicionamento da inovação nas empresas há que ter em conta os vários factores, apresentados na literatura, que influencia e/ou limita a inovação empresarial. Tendo em conta que o processo de inovação não é linear e que depende de múltiplos factores, Todtling (1991) elaborou um esquema que interrelaciona os factores que influenciam a inovação empresarial: ambiente económico e características do trabalho em rede, características internas, factores locais e política pública de inovação. Na sua investigação, Silva (2003) considerou os seguintes factores: esforço tecnológico, capacidade tecnológica, dimensão empresarial, sector de actividade, mercado, relacionamentos externos, região e financiamento Público à inovação.

Figura 2 – A ligação da gestão do conhecimento com a estratégia empresarial



Fonte: adaptado de Greiner et al. (2007)

De forma sintética, a competição não se deve basear somente no que a empresa produz, ou nos serviços fornecidos, mas também nos seus conhecimentos e na forma como se inova (McDonough III *et al.,* 2008). Nesta perspectiva, uma estratégia eficiente tem que integrar três componentes chaves: posicionamento produto/mercado, conhecimento e inovação. Neste

seguimento, nas secções seguintes iremos esclarecer as diversas actividades que constituem a gestão do conhecimento.

2.2. Geração do Conhecimento

"Numa economia onde a única certeza é a incerteza, a única fonte segura de vantagem competitiva é o conhecimento" (Nonaka, 1991: 96). Quando os mercados se alteram, as tecnologias proliferam, os competidores multiplicam-se e os produtos ficam obsoletos quase de um dia para o outro, as empresas de sucesso são as que constantemente criam novo conhecimento, que o disseminam amplamente em toda a organização e depressa o inserem em novas tecnologias e produtos. Estas actividades de criação, disseminação e inserção do conhecimento definem uma organização, esclarecida por Nonaka (1991: 96) como uma organização "criadora de conhecimento" cujo negócio exclusivo é a inovação contínua. E ainda, apesar dos termos de "capacidade intelectual" e "capital intelectual", poucos gestores agarram a verdadeira natureza da organização "criadora de conhecimento". A razão, segundo Nonaka (1991), é que eles mal entendem o que é o conhecimento e o que as empresas têm que fazer para explorá-lo.

Neste seguimento, Davenport e Prusak (1998), consideram cinco modos de geração do conhecimento nas organizações: aquisição, recursos dedicados, fusão, adaptação e redes de conhecimento (knowledge networking):

- Aquisição: o conhecimento de que uma organização necessita, não tem de ser criado de novo, basta que seja novo na organização. Logo, uma das melhores formas, de uma organização gerar conhecimento é adquirir outra organização ou contratar pessoas que possuam esse conhecimento;
- Outro modo de geração de conhecimento por aquisição consiste na contratação temporária;
- Recursos dedicados: trata-se de estabelecer grupos ou unidades organizacionais com o fim expresso da criação de conhecimento;
- Fusões: a fusão de organizações e, por conseguinte, de grupos de pessoas com conhecimentos, experiências e culturas diferentes pode criar um clima propício à geração de conhecimento;
- Adaptação: a necessidade de adaptação às alterações profundas na envolvente em que a organização desenvolve a sua actividade leva à criação de conhecimento;
- Redes de conhecimento: em todas as organizações existem grupos de pessoas com interesses comuns, que comunicam frequentemente por meios geralmente informais, que criam e partilham conhecimento, mas que não constituem qualquer estrutura formal na organização e, geralmente, até desempenham as suas funções em unidades organizativas com missões muito distintas.

Como demonstra Nonaka e Takeuchi (1995), Grant (2006) e Harlow (2008), a criação de conhecimento consiste na capacidade de uma empresa, como um todo, criar conhecimento, disseminá-lo através da organização e incorporá-lo em produtos, serviços e sistemas. Segundo os mesmos autores, a chave da criação do conhecimento nas organizações reside na distinção e relação entre o conhecimento tácito e explícito. Mesmo Sveiby (2000: 92) reafirma que a ideia, segundo a qual o conhecimento na realidade nasce de uma interacção entre os dois tipos de conhecimento – explícito e tácito – foi idealizada por Nonaka e Takeuchi (1995). Por conhecimento tácito entendem "o conhecimento do corpo, que é subjectivo, prático e analógico. Por conhecimento explícito, o do espírito, que é objectivo, teórico e numérico. À interacção de ambos os conhecimentos denominam de processo de conversão do conhecimento" (Sveiby, 2000: 93). Nonaka e Takeuchi (1995) também explicam como, no decurso dos anos 80, as empresas japonesas utilizaram quatro processos de conversão do conhecimento para conceber produtos novos e inovadores: a socialização, a externalização, a internalização e a combinação.

Tabela 1 – Os quatro processos de conversão do conhecimento

	Conhecimento Tácito	Conhecimento Explícito
Conhecimento Tácito	Socialização	Externalização
Conhecimento Explícito	Internalização	Combinação

Fonte: Nonaka e Takeuchi (1995)

Assim, Nonaka (1991: 98) indicou que a "distinção entre conhecimento tácito e explícito sugestiona quatro padrões básicos para criar conhecimento em qualquer organização:

1. De Tácito para Tácito. Às vezes, um indivíduo compartilha conhecimento tácito directamente com outro. Por exemplo, quando Ikuko Tanaka aprende com o padeiro do Hotel Internacional de Osaka, ela aprende as habilidades tácitas dele por observação, imitação, e prática. Eles tornam-se parte da sua própria base de conhecimento tácito. Posto de outro modo, ela é "socializada" na arte. Mas em si próprio, a socialização é uma forma bastante limitada da criação de conhecimento. Na verdade, o aprendiz aprende as habilidades do mestre. Mas nem o aprendiz nem o mestre ganham alguma perspicácia sistemática no conhecimento da sua arte. Porque o conhecimento deles nunca fica explícito, não pode ser facilmente transportado pela organização como um todo.

2. De Explícito para Explícito. Um indivíduo também pode combinar pedaços discretos de conhecimento explícito num novo. Mas quando o conhecimento tácito e explícito interage, como no exemplo de Matsushita, acontece algo poderoso. É precisamente esta troca entre conhecimento tácito e explícito que as companhias japonesas são especialmente boas a desenvolver.

3. De Tácito para Explícito. Quando Ikuko Tanaka puder articular as fundações do conhecimento tácito dela de fazer pão, ela o converte em conhecimento explícito, permitindo assim ser compartilhado com o seu projecto de desenvolvimento da equipa.

4. De Explícito para Tácito. O que é mais, o novo conhecimento explícito é compartilhado ao longo duma organização, outros empregados começam a interiorizá-lo, isto é, eles usam-no para alargar, estender, e reformular o seu próprio conhecimento tácito.

Tabela 2 – Os factores que constituem o processo de conversão do conhecimento nas organizações

Factor	Descrição
Socialização: do Tácito para Tácito	
Acumulação do conhecimento tácito	Os gestores recolhem informação de *sites* de vendas e de produção, partilham experiências com fornecedores e clientes e entram em diálogo com os competidores.
Recolha de informação social extra-firma	Os gestores questionam-se sobre o exterior das suas organizações, reunindo ideias para a estratégia orgaÿÿzacional rÿÿÿÿtaÿÿesÿÿa vida social diária, interagindo com peritos externos e reunindo informalmente com os competidores.
Recolha de informação social intra-firma	Os gestores encontram novas estratégias e oportunidades de mercado, questionando-se no interior da organização.
Transferência do conhecimento tácito	Os gestores criam um ambiente de trabalho que permita aos semelhantes observarem demonstrações de e a prática dominada pelos artesãos, visando o entendimento das técnicas que o seus trabalhos exigem.
Externalização: do Tácito para Explícito	
	Os gestores facilitam o diálogo essencial e criativo, o uso de metáforas para apressarem a concepção da criação e a inclusão de *designers* industriais em equipas de projectos
Combinação: do Explícito para Explícito	
Aquisição e integração	Os gestores planeiam estratégias e operações, "afogando-se" em literatura publicada, simulações computorizadas e prevendo, visando, reunir dados internos e externos.
Sintetização e processamento	Os gestores criam manuais, documentos e bases de dados para os produtos e serviços e reúnem figuras de gestão, informação técnica, ou ambas, através da organização.

Tabela 2 (cont.)

Factor	Descrição
Disseminação	Os gestores planeiam e fazem apresentações visando transmitir os novos conceitos criados.
Internalização: do Explícito para Tácito	
Experiência pessoal, aquisição do conhecimento do mundo real	Os gestores procuram e partilham novos valores e pensamentos, partilham e tentam entender as visões e valores da gestão, comunicando com colegas da organização
Simulação e experimentação: aquisição do conhecimento do mundo virtual	Os gestores facilitam a prototização e o *benchmarking* e nutrem um espírito de desafio dentro da organização. Os gestores formam equipas como modelo, conduzem experiências e partilham resultados com todo o departamento.

Fonte: adaptado de Nonaka *et al.* (2001: 496)

De forma mais sintética, segundo Silva e Neves (2003: 188-189):

- "De Tácito para Tácito: Socialização – processo de criar conhecimento tácito comum a partir da partilha de experiências.
- De Tácito para Explícito: Externalização – processo de articular conhecimento tácito em conceitos explícitos. Geralmente esta articulação é efectuada através de metáforas, analogias, conceitos, hipóteses ou modelos.
- De Explícito para Explícito: Combinação – processo de agregar conhecimentos explícitos, novos ou já existentes, num sistema de conhecimento como, por exemplo, um conjunto de especificações para um novo produto ou serviço.
- De Explícito para Tácito: Internalização – processo de incorporar conhecimento explícito em tácito. Está geralmente relacionado com aprender fazendo".

Nonaka (1991: 98) realça que este processo vai conduzir, novamente, ao começar da espiral de conhecimento, mas a um nível mais alto. "A externalização (convertendo conhecimento tácito em conhecimento explícito) e a internalização (usando aquele conhecimento explícito para estender a própria base de conhecimento tácita da pessoa) são os passos críticos nesta espiral de conhecimento.

A razão é que ambos requerem o envolvimento activo do próprio – quer dizer, compromisso pessoal. A decisão de Ikuko Tanaka para aprendiz de um padeiro mestre é um exemplo deste compromisso. Semelhantemente, quando o *controller* articula o conhecimento tácito dele e o incorpora numa nova inovação, a identidade pessoal dele é de certo modo directamente envolvida, não sendo apenas quando ele "mastigar" os números de um plano financeiro convencional.

Assim, como o conhecimento tácito inclui modelos mentais e convicções para além da experiência, enquanto movendo de tácito para o explícito, realmente é um processo de articular a visão da pessoa do mundo – o que é e o que deveria ser. Quando os empregados inventarem um novo conhecimento, eles também estão a reinventarem-se, até a empresa e mesmo o mundo.

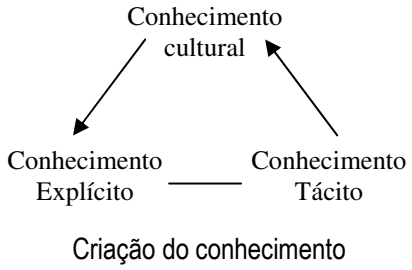
Quando os gestores se aperceberem destas ideias, eles entenderão que as ferramentas apropriadas para gerir as organizações "criadoras de conhecimento" são muito diferentes das encontradas na maioria das empresas ocidentais".

Mas, o que é a espiral do conhecimento? O novo conhecimento sempre começou com o indivíduo, logo o conhecimento pessoal de um indivíduo é transformado, como um todo, em conhecimento organizacional valioso para a organização. Nonaka (1991: 98) refere que tornar o conhecimento pessoal disponível a outros é a actividade central das organizações "criadoras de conhecimento". Nonaka e Takeuchi (1995) e Lai e Chu (2002) indicaram que, uma organização, apenas por si só, não pode criar conhecimento, sendo a base para a criação de conhecimento organizacional o conhecimento tácito dos indivíduos. Assim, a espiral do conhecimento será um processo em que o conhecimento tácito é amplificado através dos quatro processos de conversão e elevado até ao nível organizacional.

Note-se que também o conhecimento organizacional é internalizado pelos indivíduos. Segundo Silva e Neves (2003: 190) "neste processo dá-se também uma deslocação e conversão do conhecimento ao longo dos níveis das entidades envolvidas nos modos de conversão do conhecimento (individuo, grupo, organização ou inter-organizações) ".

Por outro lado, Choo (2002: 81) defendeu que uma organização tem três tipos de conhecimento: conhecimento tácito baseado nas perícias e experiências dos indivíduos; conhecimento explícito baseado em artefactos, regras e rotinas; e conhecimento cultural baseado nas suposições e convicções usadas pelos sócios para referirem o valor e o significado para uma informação nova ou conhecimento. Assim, segundo este autor, a criação do conhecimento é precipitada pelo reconhecimento de falhas no conhecimento existente na organização. Tais falhas de conhecimento podiam impedir o resolver de um problema, o desenvolvimento de um produto novo ou dificultar a detecção de uma oportunidade. As organizações criam então conhecimento novo, convertendo conhecimento tácito em conhecimento explícito, integrando e combinando conhecimento e adquirindo ou transferindo conhecimento para além dos limites.

Figura 3 – Criação do conhecimento

Modelo	Processo	Dinâmica
Criação do conhecimento	Falhas no conhecimento → Conversão do conhecimento, integração, transferência → Conhecimento novo - Conversão do conhecimento - Integração do conhecimento - Transferência do conhecimento	

Fonte: adaptado de Choo (2002: 81)

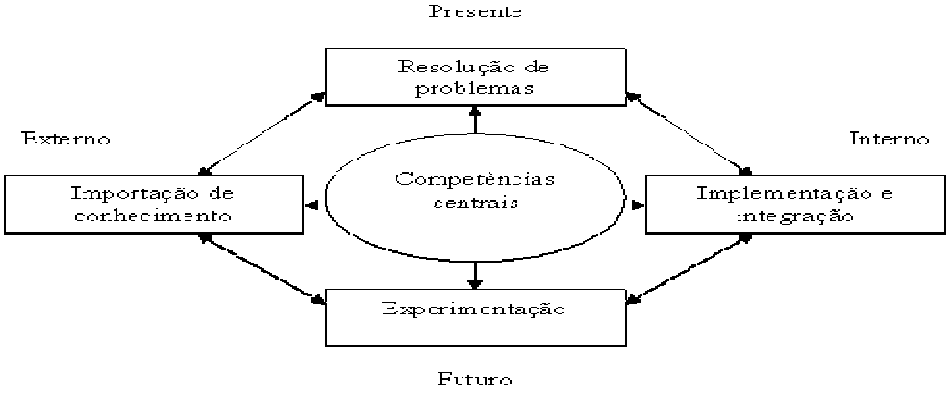
Na conversão do conhecimento, a organização cria, de forma continuada, um novo conhecimento convertendo o conhecimento pessoal, tácito de indivíduos que desenvolvem perspicácia criativa para a partilha, em conhecimento explícito, através do qual a organização desenvolve novos produtos e inova.

O conhecimento tácito é compartilhado e externalizado através de diálogos que usam metáforas e analogias. São criados novos conceitos e estes são justificados e avaliados de acordo com o seu ajuste para com a intenção da organização. (…) Finalmente, os conceitos que foram criados, justificados e modelados são movidos para outros níveis da organização para gerarem novos ciclos de criação do conhecimento (Choo, 2002: 81).

Ainda, Silva e Neves (2003: 190) "considera quatro actividades primárias na criação do conhecimento numa organização:

- Resolução criativa de problemas;
- Implementação e integração de novas metodologias e ferramentas;
- Experimentação formal e informal;
- Importação de conhecimento do exterior".

Figura 4 – Actividades criadoras de conhecimento



Fonte: Leonard-Barton (1995), adaptado por Silva e Neves (2003: 191)

Será a interacção entre estas quatro actividades de criação de conhecimento e as capacidades centrais da organização que determinam as características duma organização no que respeita à sua capacidade de inovação.

Assim, de forma conclusiva, Nonaka (1991: 101) refere que "entender a criação de conhecimento como um processo de tornar o conhecimento tácito explícito – um assunto de metáforas, analogias e modelos – tem implicações directas para como uma empresa projecta a sua organização e define papéis administrativos e responsabilidades dentro dela. Isto é o "como" das organizações "criadoras de conhecimento", as estruturas e práticas que traduzem a visão de uma organização em tecnologias e produtos inovadores". Para ele, Nonaka (1991: 101-102), "o princípio fundamental do *design* organizacional nas empresas japonesas que eu estudei é a redundância – o sobrepondo consciente de informações de empresas, actividades empresariais, e responsabilidades administrativas. Para os gestores ocidentais, o termo "redundância", com as suas conotações de duplicação desnecessária e desperdício, pode soar indesejado. E ainda, construir uma organização redundante é o primeiro passo para administrar as organizações "criadoras de conhecimento"".

2.3. Codificação do Conhecimento

Para Silva e Neves (2003: 191-192) "a codificação consiste na representação do conhecimento de modo a que este possa ser acedido ou transferido. Trata-se de tornar o conhecimento organizacional acessível àqueles que dele necessitam, e no momento em que dele necessitam".

Silva e Neves (2003: 192), considera que "o conhecimento que reside na mente das pessoas necessita de ser convertido em conhecimento que possa ser armazenado e transformado em inovação".

"O objectivo da codificação é pôr o conhecimento organizacional numa forma que o torna acessível para os que precisam dele. Transforma-se o conhecimento, literalmente, num código (não necessariamente um código de computador) para o tornar organizado, explícito, portátil e fácil de entender quanto possível" (Davenport e Prusak, 1998: 68).

Davenport e Prusak (1998), Grant (2006), Franco e Mariano (2007) e Harlow (2008) afirmam mesmo que a codificação nas organizações, semelhantemente, converte o conhecimento em formatos aplicáveis e acessíveis. Os gestores do conhecimento e usuários podem categorizar o conhecimento, descrevê-lo, mapeá-lo e modelá-lo, simulá-lo e embuti-lo em regras. Cada uma destas aproximações tem o seu próprio conjunto específico de valores e limitações e eles podem ser aplicados isoladamente ou em combinação. Obviamente, as novas tecnologias têm um papel importante na codificação do conhecimento e fazem os prospectos para as actividades que crescem de uma forma promissora.

O conhecimento codificado é sistematizado em conhecimento explícito que pode ser facilmente transferido se o remetente e o receptor tiverem o mesmo, ou semelhante, vocabulário e modelos cognitivos no domínio do conhecimento que é transferido. Procedimentos "de como fazer" são um exemplo típico do conhecimento codificado. Na maioria dos contextos, os receptores são capazes de absorver conhecimento codificado, mas mesmo conhecimento codificado altamente explícito não é provável que seja transferido se o equilíbrio dos factores motivacionais não favorecer a sua transferência. Por exemplo, o possuidor do conhecimento pode não o desejar compartilhar porque a retenção do conhecimento lhe confere uma vantagem competitiva. Ou o potencial utilizador do conhecimento pode ter um preconceito de "não-inventar-aqui" que trabalha contra a adopção de melhores práticas (Huber, 2001).

Assim, quais serão os princípios básicos da codificação do conhecimento? De acordo com Davenport e Prusak (1998: 68) "a dificuldade primária encontrada no trabalho da codificação é a pergunta de como classificar o conhecimento sem perder as suas propriedades distintivas e se transformar em informações menos vibrantes ou dados. Noutras palavras, alguma estrutura para

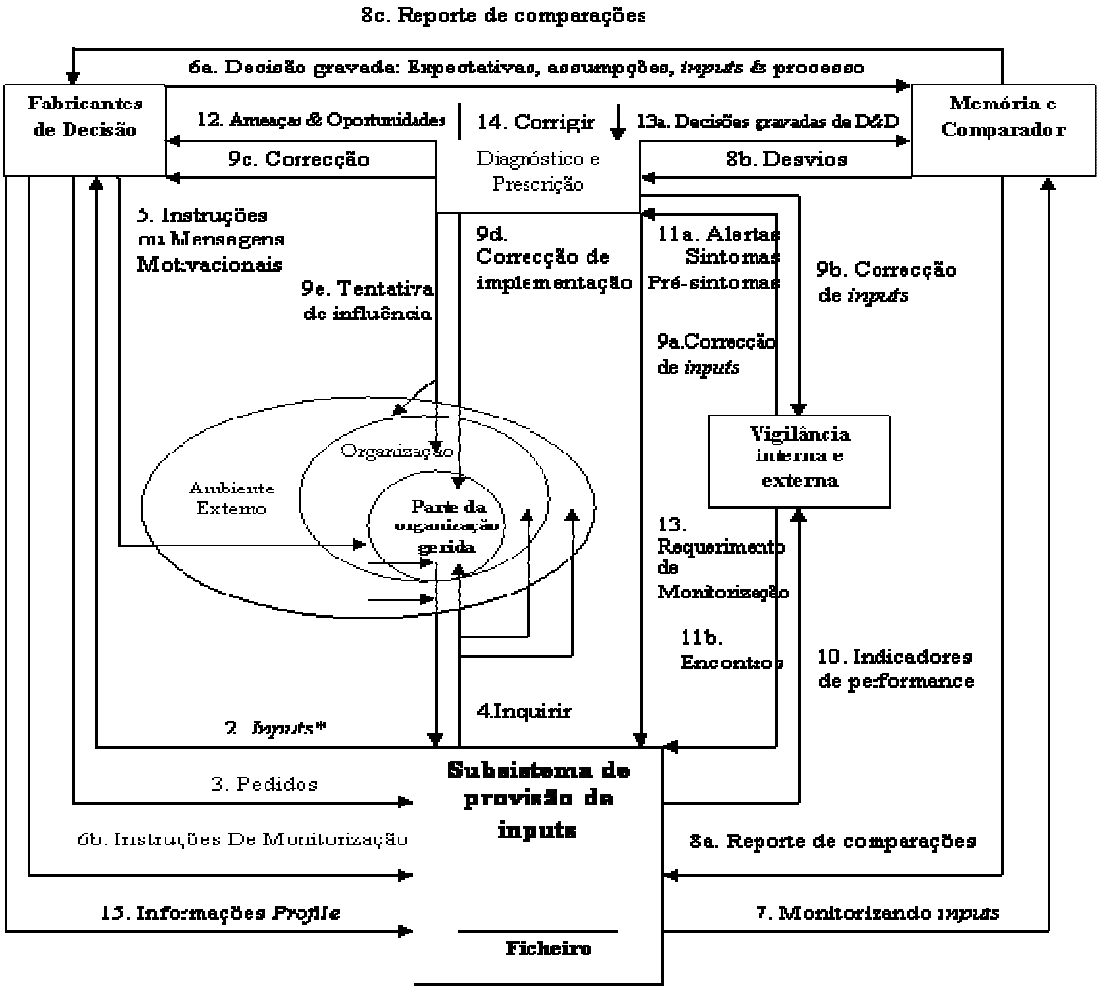
o conhecimento é necessária, mas muitas vão arrasá-lo. As organizações que querem classificar o conhecimento com sucesso deveriam-se lembrar então dos seguintes quatro princípios".

1. Os gestores têm que decidir que metas empresariais o conhecimento codificado servirá (por exemplo, empresas cuja intenção estratégica envolve colocarem-se mais próximas do cliente, podem escolher codificar o conhecimento do cliente);
2. Os gestores devem ser capazes de identificar o conhecimento que existe em várias formas, apropriado para alcançar as suas metas;
3. Os gestores de conhecimento têm que avaliar o conhecimento por utilidade e conveniência para a codificação;
4. Os codificadores têm que identificar um meio apropriado para a codificação e distribuição".

Porém, quando se fala em codificação do conhecimento, é usual também encontrar o termo de conhecimento integrativo, mas porquê e qual o seu conceito? Helfat e Raubitschek (2002) definem o conhecimento integrativo como o "conhecimento que integra, ou conhecimento de como integrar, actividades diferentes, capacidades e produtos em uma ou mais cadeias verticais. O conhecimento integrativo habilita as organizações a coordenar actividades dentro de uma cadeia vertical ou por cadeias verticais, obter a avaliação do mercado dos clientes sobre produtos e obter avaliações de dentro de qualquer uma das cadeias verticais ou dos mercados externos, relativos a tecnologia. A natureza da coordenação dentro e por cadeias verticais depende em parte dos tipos de conhecimento que devem ser coordenados, como tácito contra conhecimento codificado" (Helfat e Raubitschek, 2000: 319). Por exemplo, Monteverde (1995), referido por Helfat e Raubitschek (2000: 319), recorre ao termo de "diálogo técnico destruturado" que envolve o "incodificável destruturado, geralmente verbal, e, frequentemente, a comunicação cara-a-cara exigida pela gestão integrada do projecto" de *design* do produto e a sua fabricação.

Nesta situação, a coordenação do conhecimento tácito que reside em múltiplas fases de uma cadeia vertical requer mecanismos organizacionais firmemente "arrumados". Porém, a codificação de conhecimento classificado, não impede a necessidade de conhecimento integrativo. Como exemplo, Flaherty (1996), referido por Helfat e Raubitschek (2000: 319-320), menciona o "*just-in-time* na manufacturação e distribuição de bens para as vendas finais, que requer uma estreita coordenação entre os fornecedores, os fabricantes e os distribuidores". Embora a informação que flúi entre as várias fases da cadeia de provisão seja largamente classificada, a complexidade de coordenar estes fluxos de informação pode requerer mecanismos organizacionais integrativos e conhecimento. Mas, constata-se que a informação que percorre a organização cada vez se encontra mais dispersa e em maior quantidade, inclusive mesmo Ackoff (1967), referido por Barabba *et al.* (2002: 363), defendeu que os gestores sofrem mais de uma superabundância de informação irrelevante do que de uma escassez de informação pertinente. Então o que deve ser feito para filtrar toda essa informação, que muitas vezes também se encontra codificada dentro da organização? Ackoff (1967), referido por Barabba *et al.* (2002: 363), sugeriu, então, que um sistema de apoio de gestão deveria filtrar as mensagens entradas por relevância e as condensar para minimizar o tempo exigido para adquirir o seu conteúdo.

Figura 5 – Sistemas de suporte de aprendizagem e adaptação



Fonte: adaptado de Barabba *et al.* (2002: 364)

Assim, como os dados devem ser processados para serem convertidos em informação, conhecimento, compreensão, ou sabedoria, o processamento de dados é uma parte necessária do subsistema de provisão de *inputs*. Mas o que é este sistema? Barabba *et al.* (2002: 363), defendem que os inputs (2) são transmitidos aos "fabricantes de decisão" em relação aos seus pedidos (3). Quando os "fabricantes de decisão" receberem os *inputs*, eles nem sempre os vão achar úteis. Eles podem-nos achar ilegíveis ou incompreensíveis, podem duvidar da sua validez ou podem questionar a sua perfeição. Então, o recibo dos inputs (2) conduz frequentemente a pedidos (3) adicionais. Tais pedidos requerem duas capacidades adicionais do subsistema de provisão de inputs. Este subsistema deve poder gerar dados novos – quer dizer, investigar (4) a parte da organização gerida e o seu ambiente de forma que os inputs (2) adicionais requeridos possam também ser obtidos. Também tem que ter a habilidade para reutilizar dados, informação, conhecimento ou compreender os previamente recebidos ou gerados. Isto significa que deve poder armazenar dados numa forma recuperável. Uma facilidade de armazenamento de dados será num "arquivo". É uma parte do subsistema de provisão de *inputs*.

Uma vez processados, os dados novos ou velhos, para providenciar a informação "acreditada" para responder ao pedido recebido dos "fabricantes de decisão", é transmitida a eles. Este ciclo de pedido-cumprimento ou pode continuar até os "fabricantes de decisão" terem todos os *inputs* que querem ou passar o tempo e ter que se fazer uma decisão com tudo o que eles têm. Em alguns casos, eles podem acreditar que não é provável que o tempo e o custo de investigação adicional sejam justificados pela melhoria ou aumento dos *inputs* que eles

acreditam que são possíveis. O valor esperado da "informação-para-sabedoria" (chamada informação +) deveria ser calculado antes de qualquer *input* pedido e usado. Deveria ser feita uma análise sobre se a quantia máxima que o "fabricante de decisão" está disposto a pagar pelo *input* é justificada pelo valor esperado do seu uso.

O *output* de uma decisão para fazer algo é uma mensagem que é uma instrução ou uma mensagem motivacional e é endossada aos indivíduos que fazem parte da organização, de quem a responsabilidade é para levar a cabo as instruções ou de quem a motivação é o objectivo. Uma instrução será uma mensagem para outros ou para si mesmo que é pretendido que aumente ou mantenha a eficiência da organização. Uma mensagem motivacional é a pessoa pretender afectar os valores de uma organização ou alguns dos seus elementos (internos ou externos). Uma decisão, claro que, pode ser para não fazer nada como também para fazer algo. Quando uma decisão para não fazer nada é tomada, nenhuma instrução pode ser requerida, mas um registo de decisão (6a) é.

Toda a decisão tem um de dois únicos propósitos possíveis: fazer algo acontecer que caso contrário não irá acontecer ou impedir algo de acontecer que caso contrário acontecerá. Além disso, existe sempre um tempo no qual o efeito da decisão é esperado. Então, controlar uma decisão, os seus efeitos esperados e o tempo das realizações esperadas deveria ser feito de forma explícita e registado. Tudo isso é igualmente verdadeiro em decisões que envolvem a implementação de uma decisão. Por exemplo, se a decisão tomada foi construir uma nova fábrica, há expectativas de quando, aproximadamente, deveria ser completada, o que deveria valer e assim por diante.

Deveriam ser registadas decisões de implementação separadamente e deveriam ser localizadas. Além de registar os efeitos esperados e quando os esperam, para cada decisão, deveria ser mantido um registo das informações e suposições nas quais as expectativas são baseadas e o processo pelo qual a decisão foi alcançada, por quem e quando. Tudo isso deveria ser registado no registo de decisão (6a), deveria ser colocado e armazenado numa memória inactiva e num "*comparador*". Porque as recordações humanas são distorcidas, modificam o seu conteúdo com o passar do tempo, especialmente previsões e expectativas, logo é importante que a memória empregue seja completamente inactiva. O armazenamento inactivo de *inputs* será a única coisa que um computador pode fazer, que um humano não pode.

Uma versão do registo de decisão (6a), monitorizando as instruções (6b), deveria ser enviado ao subsistema de provisão de inputs que tem a responsabilidade de conferir a validez das expectativas, suposições e a informação+ usada, tomada pela decisão e sua implementação. Quando obtida a informação sobre a validade dos efeitos esperados, deveriam ser enviadas as suposições pertinentes e a informação+ usada, à memória e ao *comparador*, bem como monitorizando os *inputs* (7).

Então, usando a informação sobre a decisão registada (6a) armazenada na memória e os *inputs* monitorizados (7), deveria ser feita uma comparação dos efeitos actuais e dos esperados e as suposições e ocorrências pertinentes. Quando o *comparador* não encontra nenhuma diferença significante entre as expectativas/assumpções e o desempenho actualmente observado e informado nos *inputs* monitorizados (7), nada mais necessita ser feito do que dar entrada de um relatório de comparações (8a), na memória, para futura referência e para o "fabricante de decisão" (8c). Este registo preserva o que é conhecido ou acreditado. Então, deveria ser armazenado de uma forma facilmente recuperável, por exemplo, pelo uso de palavras-chaves. Porém, se são encontradas diferenças significantes, elas são reportadas como divergências (8b) para a função de diagnóstico e prescrição. Tais divergências indicam que ocorreu algo errado. Um diagnóstico é exigido para determinar o que tem, e o que deveria ser feito. O propósito do diagnóstico é encontrar o que é responsável pelas divergências e prescrever acções correctivas. Por outras palavras, a função diagnóstico consiste em explicar o engano e, então, produzir entendimento disto.

Assim, este será um exemplo de um sistema de apoio de gestão que irá filtrar as mensagens entradas por relevância e as condensar para minimizar o tempo exigido para adquirir o seu conteúdo, logo codificação. Mas e a transferência do conhecimento codificado, como se efectuará?

2.4. Transferência do Conhecimento

A transferência do conhecimento segundo os autores Davenport e Prusak (1998), Lai e Chu (2002), Franco e Mariano (2007) e Simões (2008) é permanente e espontânea nas organizações, passiva e não previamente planeada, apesar de dividida e localizada. O desafio da gestão do conhecimento será ter presente a necessidade da partilha e aceitação do conhecimento e promover essa mesma partilha, desenvolvendo estratégias que encorajem a transferência do conhecimento, de forma espontânea. Mas, segundo diversos autores, onde se destacam Wong e Aspinwall (2004), Tirpak (2005) e Grant (2006), continuam a considerar a transferência do conhecimento como uma área apetecível para a informática e para os sistemas de informação, o que é correcto, em parte, na medida em que diversa tecnologias podem facilitar a transferência do conhecimento (Simões, 2008). Note-se, inclusive, que "nenhuma tecnologia isoladamente satisfaz todos os critérios pedidos por um sistema de gestão do conhecimento porque a gestão do conhecimento não é só uma tecnologia. É um campo multidisciplinar que inclui aspectos da ciência da informação, comunicação interpessoal, aprendizagem organizacional, ciência cognitiva, motivação, instrução, divulgação e análise de processos do negócio" (Loureiro, 2003: 112).

Destaque-se que, segundo Silva e Neves (2003), a comunicação pessoal terá mais resultados para a transferência de informação complexa do que qualquer género de tecnologia. Webber (1993), referido por Silva e Neves (2003: 193), realça que "as conversas são o modo como os trabalhadores do conhecimento descobrem aquilo que sabem, partilham-no com os seus colegas e nesse sentido criam novo conhecimento para a organização".

Neste sentido, o sucesso da transferência do conhecimento, de acordo com Silva e Neves (2003: 194) "é determinado pelos valores, normas e padrões de comportamento que incorporam a cultura organizacional mais do que pelas ferramentas proporcionadas pela tecnologia, embora estas sejam essenciais, em particular no caso de organizações grandes e complexas".

Assim, quais as estratégias para a transferência de conhecimento? "Espontânea, a transferência de conhecimento destruturada é vital para o sucesso de uma empresa. Embora o termo "gestão de conhecimento" insinua transferência formalizada, um dos seus elementos essenciais é desenvolver estratégias específicas para encorajar as tais trocas espontâneas. Isto é particularmente necessário para organizações cujo papel primário é criar conhecimento" (Davenport e Prusak, 1998: 89).

Porém, Davenport e Prusak (1998: 96) referem que existem muitos factores culturais que inibem a transferência de conhecimento, aos quais apelidam de "fricções" de inibidores porque, segundo os autores, reduzem a velocidade ou previnem a transferência, sendo também provavelmente, para corroer algum do conhecimento conforme ele fluí pela organização. São as fricções mais comuns, tabela 3, e os modos de as superar.

Tabela 3 – As fricções mais comuns e os modos de as superar

Fricção	Soluções Possíveis
Falta de confiança	Construir relações e ganhar confiança através de encontros cara-a-cara
Diferentes culturas, vocabulários e modelos de referência	Criar uma base sólida através da educação, discussão, trabalho de equipa e rotação de postos de trabalho
Falta de tempo e de local de reuniões	Estabelecer períodos de tempo e lugares para a transferência de conhecimento: conferências, palestras e relatórios
Status e prémios são atribuídos aos detentores do conhecimento	Avaliar a performance e providenciar incentivos aos que partilham o conhecimento
Falta de capacidade de absorção	Educar os empregados para a flexibilidade, providenciar tempo para a aprendizagem e contratar com a esperança de ideias
Crer que o conhecimento é apenas destinado a determinados grupos	Encorajar uma aproximação não hierárquica ao conhecimento e tornar a qualidade de ideias mais importante do que a sua proveniência
Intolerância para com os erros	Aceitar erros de criatividade e não perder *status* devido a não saber tudo

Fonte: Davenport e Prusak (1998: 97)

Logo, a transferência de conhecimento efectiva, segundo Davenport e Prusak (1998), é de longe mais fácil quando os participantes falam os mesmos, ou semelhantes, idiomas (através de idioma não se quer dizer apenas inglês ou espanhol mas também "engenheiro mecânico" ou "vendedor"), ou, como diz Thomas Allen, citado por Davenport e Prusak (1998: 98-99), "quando há "porteiros" para traduzir entre culturas e sistemas de valor". Mas, por vezes, "a transferência de conhecimento só pode trabalhar se as várias partes da organização são reunidas fisicamente" (Davenport e Prusak, 1998: 99). Davenport e Prusak (1998: 100) afirmam mesmo que "as pessoas que compartilham a mesma cultura de trabalho podem comunicar melhor e podem transferir conhecimento, de forma mais eficiente, do que as pessoas que não compartilham". As pessoas julgam a informação e o conhecimento que recebem, numa medida significante, com base em quem lhes confere essa informação e conhecimento. As organizações que ignoram este facto irão, provavelmente, ficar desapontadas com os resultados dos projectos de transferência de conhecimento. "A reputação é uma procuração para o valor, que nós usamos para avaliar o fluxo de informação que vem a nós. Nós não temos tempo para olhar cuidadosamente para tudo, assim nós seleccionamos o que pensamos valer a pena, baseados na reputação do emissor" (Davenport e Prusak, 1998: 101).

Assim, a transferência do conhecimento é composta por duas acções, de acordo com Davenport e Prusak (1998: 101-102), sendo as mesmas a transmissão e a absorção. "Tornando o conhecimento meramente disponível, não é nenhuma transferência. O acesso é necessário, mas não é nenhum meio suficiente para assegurar que aquele conhecimento será usado. O objectivo da transferência de conhecimento é melhorar a habilidade de uma organização para fazer coisas e então aumentar o seu valor. Até mesmo a transmissão e absorção, juntas, não têm nenhum valor útil se o novo conhecimento não conduzir a alguma mudança no comportamento, ou ao desenvolvimento de alguma ideia nova que conduza a um comportamento novo. É bastante comum para alguém, entender e absorver novo conhecimento mas não o pôr em uso por uma variedade de razões. A mais importante é não respeitar ou confiar na fonte do conhecimento, sendo as restantes o orgulho, teimosia, falta de tempo, falta de oportunidade e um medo de correr riscos (numa organização que castiga enganos)". Inclusivamente, Zahra e George (2002) Melkas e Harmaakorpi (2008) reafirmam que a transferência do conhecimento se baseia na transmissão e absorção do conhecimento, assumindo que a absorção possui 2 tipos: a potencial capacidade de absorção (aquisição e assimilação) e a real capacidade de absorção (transformação e exploração) em adquirir e assimilar o conhecimento.

Todos os factores, até agora mencionados, afectam o sucesso e a eficiência da transferência do conhecimento nas organizações. Eles vão influenciar a "velocidade" da transferência, quer dizer, a velocidade com que o conhecimento se move através de uma organização. Mas, com que rapidez e amplitude, é disseminado? Com que rapidez as pessoas que necessitam do conhecimento se dão conta dessa sua necessidade, e adquirem o seu acesso? Computadores e redes, claro que, superam a forma de aumentar a velocidade do conhecimento, porém surge a noção de viscosidade, mas do que se trata? A viscosidade, de acordo com Davenport e Prusak (1998: 102-103), "recorre à riqueza (ou densidade) do conhecimento transferido (…) a viscosidade é influenciada por vários factores, especialmente o método de transferência. O conhecimento transferido por meio de uma aprendizagem longa, ou por meio de uma relação de mentoring é provável que tenha uma viscosidade alta: o receptor ganhará uma tremenda quantidade de conhecimento detalhado com o passar do tempo. O conhecimento vindo de um banco de dados ou adquirido da leitura de um artigo, será muito mais objectivo. Obviamente, a velocidade e a viscosidade são preocupações importantes para os gestores do conhecimento, na determinação da eficiência de como a sua empresa usa o seu conhecimento capital. Com que velocidade coloca o conhecimento onde possa gerar valor e quantos dos activos de conhecimento estão a adquirir o que de facto eles necessitam? Porque a aprendizagem genuína é um tal empenho, profundamente humano, e porque a absorção, não só mas também, a aceitação do conhecimento novo envolve tantas pessoas e factores fisiológicos, a velocidade e a viscosidade encontram-se frequentemente em conflito. O que aumenta a velocidade pode

reduzir a viscosidade. A maioria dos esforços da transferência de conhecimento afecta um acordo entre estes dois factores".

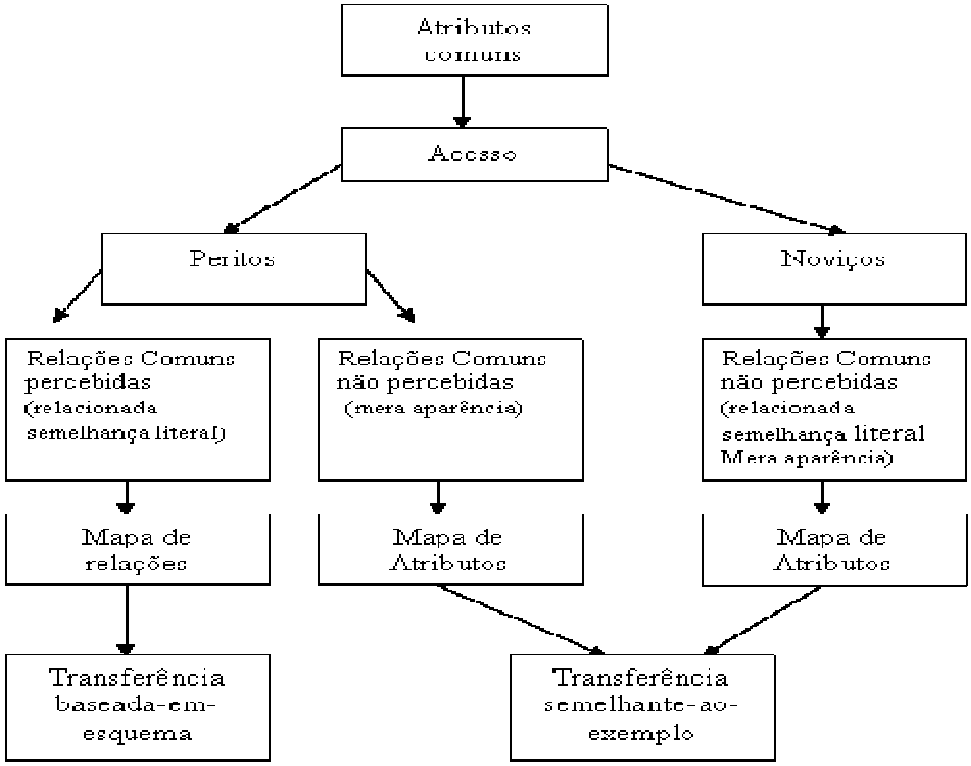
Neste seguimento, Davenport e Prusak mencionam a aprendizagem, dita genuína, mas como é que esta surge relacionada com a transferência do conhecimento? De acordo com Paxton e John (1997: 267) "a aprendizagem efectua-se por meio de um processo que transfere o conhecimento da base mais conhecida para o objectivo menos conhecido. A força motriz por detrás da aprendizagem através da analogia será a ideia que os domínios relacionados em alguns cumprimentos serão aspectos, que provavelmente estarão também relacionados noutros". Assim, fala-se das fases do processo de aprendizagem analógico, mas como se processam?

As fases da transferência mencionam que o processo de aprendizagem analógico (figura 4) é actualmente composto por numerosas fases, cada uma com o seu próprio conjunto de assuntos e influências (Anderson e Thompson, 1989; Gentner, 1989; Holyoak e Thagard, 1989; Reeves e Weisberg, 1994; Winston, 1980; referidos por Paxton e John, 1997: 267). Estas fases incluem o (a) acesso ao domínio básico, o (b) traçar dos elementos do objectivo sobre a base, (c) transferir conhecimento da base para o objectivo, e (d) induzir um esquema. A fase do acesso do processo de aprendizagem analógico está preocupada com a habilidade de um objectivo para recobrar a representação mental do "estudioso" de um domínio básico. O objectivo da fase de acesso é activar a representação mental do "estudioso" de um domínio básico, de forma que possa ser considerado como uma fonte potencial de informação sobre o domínio designado (Paxton e John, 1997: 267). O assunto central da fase do mapeamento do processo de aprendizagem analógico será se é, ou não, activado o conhecimento nas referidas fases de transferências da situação familiar à situação moderna. A meta do mapeamento será alinhar os domínios básicos e designados de tal forma que, o conhecimento associado com a base possa ser transferido para o objectivo. Isto é conseguido através de um processo que constrói correspondências um-para-um, entre os elementos das representações do estudioso dos domínios básicos e designados (Gentner 1983, 1989; Holyoak, 1984; Holyoak e Thagard, 1989; referidos por Paxton e John, 1997: 267).

Tendo-se notado a potencial relevância de uma base previamente adquirida de conhecimento a uma situação moderna (acesso) e tendo-se executada uma cartografia dos elementos dos dois domínios (mapeamento), a próxima fase na aprendizagem analógica será a transferência actual da informação. É nesta fase que a aprendizagem acontece, com conhecimento do domínio básico que se move para o objectivo. Sob a decisão da transferência do conhecimento está a convicção de que o domínio conhecido para ser semelhante em certos cumprimentos (como estabelecido na fase de cartografia) é provável para ser, como bem semelhante em outros cumprimentos (Paxton e John, 1997: 267-268).

Finalmente, em muitos casos, uma estrutura de conhecimento mais abstracta, como o esquema, pode ser criada como um subproduto do processo de aprendizagem analógico. Como muita pesquisa revelou (Catrambone e Holyoak 1989; Gick e Holyoak 1980, 1983; Spencer e Weisberg 1986; referidos por Paxton e John, 1997: 268) esta estrutura de conhecimento abstracta pode ser usada como uma base na futura aprendizagem analógica.

Figura 6 – Um modelo de consumidor que aprende por analogia



Fonte: adaptado de Paxton e John (1997: 270)

Assim, Matusik (2002: 606) sintetiza que os níveis de conhecimento das organizações são influenciados por quatro processos: criação do conhecimento, absorção de conhecimento público pertinente (externo), conhecimento transferido dentro da organização e disseminação no ambiente externo da organização do conhecimento pertencente à mesma. Cada um destes processos irá afectar o nível de conhecimento global da organização. Mas quais são os níveis de conhecimento das organizações? As organizações possuem "tesouros" de conhecimentos privados e públicos. O conhecimento privado é único para a organização, considerando que o conhecimento público reside no domínio público. O conhecimento privado – ou específico – da organização pode ser uma fonte de vantagem competitiva. É um exemplo de um recurso que é único, valioso, raro e imperfeitamente imitável. O conhecimento privado inclui itens tais como as rotinas sem igual de uma organização, processos, documentação ou segredos do negócio. O conhecimento público consiste num conhecimento sem igual para qualquer uma das organizações. Ele reside no ambiente externo e é, na essência, um bem público. O conhecimento público inclui itens tais como as melhores práticas industriais e profissionais, Total Quality Management (TQM), Inventário Just – in – Time, são todos exemplos de "melhores práticas" no domínio público.

Como se efectua a transferência de conhecimento dentro da organização? A dispersão de conhecimento dentro de uma organização afecta quem pode fazer uso produtivo do conhecimento da organização. (...) As características específicas do próprio conhecimento afectam a habilidade para transferir conhecimento dentro da organização. Por exemplo Zander e Kogut (1995), referidos por Matusik (2002: 608), defenderam que o grau de codifiabilidade e de ensinamento afectam a habilidade da organização para transferir capacidades dentro da empresa mas por locais geográficos distantes. Outro foco de estudos, em situações de transferência específicas dentro de uma organização, examina o contexto e os indivíduos envolvidos numa transferência intencional. A falta de relação directa ou comunicação por subunidade inibe a partilha de conhecimento (Galbraith, 1973; Lawrence e Lorsch, 1967; referidos por Matusik, 2002: 608). Também, a viscosidade da informação a ser transferida e as características dos

indivíduos envolvidos afectam as transferências intencionais ao nível do projecto (Szulanski, 1996; referido por Matusik, 2002: 608). Hansen (1999), referido por Matusik (2002: 608), toma uma perspectiva de redes-sociais olhando para as transferências de conhecimento intencionais por subunidades e considera que os laços fracos das redes estão positivamente relacionados com a velocidade da transferência quando o conhecimento a ser transferido não seja complexo. Porém, laços fracos afectam negativamente a velocidade da transferência quando o conhecimento é complexo. Assim, a complexidade do conhecimento, bem como também o contexto de rede, são considerações importantes. Huber (1991), referido por Matusik (2002: 608), também notou que o corpo de pesquisa na distribuição de informação dentro das organizações afecta a aprendizagem dentro de organizações.

Como se efectua a preservação/disseminação de conhecimento privado? Interessantemente, poucos estudos especificamente consideram a disseminação de conhecimento pertencente à organização fora da organização. Argumentos clássicos na visão recurso-baseada da organização articulam que um recurso, como conhecimento, é uma fonte potencial de vantagem competitiva se é raro, imitável, não substituível e valioso (Barney 1991; referido por Matusik, 2002: 609). Certamente, as condições de singularidade, inimitabilidade e, sugestões de não-substituição que a disseminação de um recurso como o conhecimento, pode afectar negativamente a vantagem competitiva. Porém, só alguns estudos especificamente examinam este processo. March discutiu a disseminação de conhecimento na simulação dele dos efeitos na reviravolta no conhecimento e aprendizagem (March, 1991; referido por Matusik, 2002: 609). Ele assumiu que uma reviravolta moderada pode afectar positivamente a criação de conhecimento; há alguma perda de conhecimento como também algum ganho com a reviravolta. Embora a disseminação de conhecimento seja reconhecida teoricamente como importante, houve relativamente poucos exames empíricos dos mecanismos que a podem inibir ou apressar. Sintetizando, quando examinados conjuntamente, estes quatro processos possibilitam uma visão compreensiva do conhecimento da organização. O nível de conhecimento da organização em qualquer ponto do tempo é relativo das suas acções e fluxos de conhecimento, dentro e pelos limites da organização (Matusik, 2002: 609).

3. MODELO CONCEPTUAL PROPOSTO

Como vimos ao longo da revisão da literatura e nos pontos anteriores, a aposta na gestão do conhecimento pode criar vantagens competitivas sustentáveis. Inclusive, vários autores afirmaram que a gestão do conhecimento é considerada a parte fundamental da estratégia para usar conhecimento na criação de vantagens competitivas sustentáveis no actual ambiente de negócio (Bloodgood e Salisbury, 2001; Loureiro, 2003; Sher e Lee, 2004; Grant 2006; Gray e Meister, 2006; Greiner *et al.,* 2007). Com base nestas considerações, elaborou-se o seguinte modelo (Figura 7), que revela a interacção entre as diversas actividades da gestão do conhecimento e as intuições que serão alvo de uma investigação futura.

Como se pode constatar pela figura 7, uma estratégia competitiva deve abarcar três actividades (geração do conhecimento, codificação do conhecimento e transferência do conhecimento) interligados entre si e estes devem estar em sintonia com a estratégia da organização. Se houver uma interligação entre estes quatro aspectos a organização irá obter vantagens competitivas e desempenhos superiores.

Figura 7 – Modelo das actividades de gestão do conhecimento nas organizações



Fonte: elaboração própria

4. CONCLUSÃO

Na literatura sobre a gestão estratégica, o conhecimento baseado nas organizações altera o foco dos recursos do conhecimento e propõe que o conhecimento é o recurso mais importante para criar uma vantagem competitiva sustentável. Mas, nem todas as actividades da gestão do conhecimento foram provadas como influências positivas para a performance empresarial ou como resultado de uma vantagem competitiva (Greiner *et al.,* 2007). Inclusivamente, Learned *et al.* (1969), já defendia, na altura, que na formulação de vantagens estratégicas, as competências distintivas, os recursos fortes e as oportunidades é que determinam a sustentabilidade das vantagens competitivas, sendo, segundo Dinis (2004), o sucesso de uma empresa dependente da criação de vantagem competitiva que, por sua vez, resulta da capacidade que a empresa tem de promover inovações. Para obter sucesso no século XXI, as empresas necessitam de apostar no conhecimento constantemente (Winkelen *et al*., 2008), principalmente nos países mais desenvolvidos, onde fabricar os produtos *standards* usando métodos standards não sustenta vantagens competitivas. Assim, a gestão do conhecimento estratégico entende-se como um processo colectivo, de natureza interactiva, que pressupõe uma partilha de informação e atitudes nas várias fases do desenvolvimento do conhecimento (Simões, 2008). A gestão do conhecimento, vista como um conjunto de processos que governa a criação, a disseminação e a utilização do conhecimento para atingir plenamente os objectivos da organização, é uma nova área na confluência entre a tecnologia da informação e da gestão, um novo campo entre a estratégia, a cultura e os sistemas de informação de uma organização. Esta forma de gestão será constituída pelas seguintes actividades: a criação, a codificação e a transferência do conhecimento.

A criação de conhecimento consiste, assim, na capacidade de uma empresa, como um todo, criar conhecimento, disseminá-lo através da organização e incorporá-lo em produtos, serviços e sistemas (Nonaka e Takeuchi, 1995). Inclusivamente, Lu e Etzkowitz (2008) mencionam que o desenvolvimento das capacidades de inovação por meio da criação de conhecimento domina a agenda política na China.

Segundo os mesmos autores, Nonaka e Takeuchi, a chave da criação do conhecimento nas organizações reside na distinção e relação entre o conhecimento tácito e explícito, o que sugestiona quatro padrões básicos para criar conhecimento em qualquer organização.

Por sua vez, o objectivo da codificação será colocar o conhecimento da organização numa forma que o torne acessível para os que precisam dele, facilitando o seu acesso. Transforma-se o conhecimento, literalmente, num código para, desta forma, o tornar organizado, explícito, portátil e fácil de entender.

Por fim, a transferência do conhecimento é permanente e espontânea nas organizações, passiva e não está previamente planeada, apesar de dividida e localizada.

Assim, de forma conclusiva, o êxito da transferência do conhecimento, segundo Silva e Neves (2003: 194) "é determinado pelos valores, normas e padrões de comportamento que incorporam a cultura organizacional mais do que pelas ferramentas proporcionadas pela tecnologia, embora estas sejam essenciais, em particular no caso de organizações grandes e complexas".

Como limitação aponta-se que este artigo apenas efectua uma revisão de literatura sobre a temática da estratégia relacionada com a gestão do conhecimento, sem efectuar uma aplicação prática da mesma (uma análise qualitativa ou quantitativa).

Tendo por base estas limitações, sugere-se como futuras investigações: verificando-se que não existem artigos explícitos sobre a gestão do conhecimento, relacionado com a estratégia, abordada na instituições de ensino superior militares, estão assim criadas as condições para um estudo empírico dessas mesmas actividades junto das instituições anteriormente mencionadas.

5. BIBLIOGRAFIA

BARABBA, VINCENT P.; POURDEHNAD, JOHN; ACKOFF, RUSSELL L. (2002), "Above and Beyond Knowledge Management", In: Choo, Chun Wei; Bontis, Nick (eds.), *The strategic management of intellectual capital and organizational knowledge,* Oxford University Press, New York, pp. 359-369.

BARBOSA, M. (2006), "A capacidade inovadora empresarial como factor criador de vantagens competitivas nas empresas da Indústria Transformadora da Região da Beira Interior", Dissertação de Mestrado em Gestão, Universidade da Beira Interior.

BARNEY, J. (1991), "Firm resources and sustained competitive advantage", J*ournal of Management*, Vol. 17, nº. 3, pp. 99-120.

BLOODGOOD, J.M.; SALISBURY, W.D. (2001), "Understanding the influence of organizacional change strategies on information technology and knowledge management strategies", *Decision Support Systems,* Vol. 31, nº 1, pp. 55-69.

BOISOT, MAX (2002), "The creation and sharing of knowledge", In: Choo, Chun Wei; Bontis, Nick (eds.), "The strategic management of intellectual capital and organizational knowledge", Oxford University Press, New York, pp. 64-77.

BRACKER, JEFFREY (1986), "The Historical Development of the Strategic Management Concept", *Academy of Management*, Vol. 5, pp. 219-224.

CHANDLER, A. D. (1962), "Strategy and Structure", MIT Press, Cambridge.

CHOO, CHUN WEI (1998), "The Knowing Organization: How organizations use information to construct meaning, create knowledge, and make decisions", *http://choo.fis.utoronto.ca/FIS/KO/default.html (9* de Novembro de 2004).

CHOO, CHUN WEI (2002), "Sensemaking, Knowledge Creation, and Decision Making", In: Choo, Chun Wei; Bontis, Nick (eds.), "The strategic management of intellectual capital and organizational knowledge", Oxford University Press, New York, pp. 79-88.

CHOO, CHUN WEI (2003), "Gestão de Informação para a Organização Inteligente – a arte de explorar o meio ambiente", Editorial Caminho, Lisboa.

CHOO, CHUN WEI; BONTIS, NICK (eds.) (2002), "The strategic management of intellectual capital and organizational knowledge", Oxford University Press, New York.

CUNHA, MIGUEL PINA; REGO, ARMÉNIO; CUNHA, RITA CAMPOS; CARDOSO, CARLOS CABRAL (2003), "Manual de Comportamento Organizacional e Gestão", 1ª Edição, Editora RH, Lisboa.

DAVENPORT, THOMAS H.; JARVENPAA, SIRKKA L.; BEERS, MICHAEL C. (1996), "Improving knowledge work processes", *Sloan Management Review* (Verão), pp. 53-65.

DAVENPORT, THOMAS H.; LAURENCE, PRUSAK (1998), "Working Knowledge – How organizations manage what they know", 1ª Edição, Harvard Business School Press.

DAVENPORT, THOMAS H.; GLASER, JOHN (2002), "Just-in-Time Delivery Knowledge Management", *Harvard Business Review* (Julho), pp. 107-111.

DIERKES, MEINOLF; BERTHOIN ANTAL, ARIANE; CHILD, JOHN; NONAKA, IKUJIRO (eds) (2001), "Handbook of organizational learning & knowledge", Oxford University Press, New York.

DIGMAN, L. (1990), "Strategic Management: concepts, decisions, cases", 2nd Ed., Homewood, Illinois: Irwin.

DINIS, A. (2004), "Empresarialidade em meio rural e periféricos: um modelo multidimensional de análise", Tese de Doutoramento em Gestão, Universidade da Beira Interior.

FAHY, J.; SMITHEE, A. (1999), "Strategic Marketing and the Resource based View of the Firm", *Academy of Marketing Science Review*, p.1.

FRANCO, MASSIMO; MARIANO, STEFANIA (2007), "Information Technology Repositories and Knowledge Management Processes", *Journal of Information and Knowledge Management Systems,* Vol. 37, n. º 4, pp. 440-451.

FREIRE, A. (2006), "Estratégia: sucesso em Portugal", Editorial Verbo, 8ª Edição, Lisboa.

GHEMAWAT, P. (2000), "Competition and Business Strategy in Historical Perspective", *http://papers.ssrn.com/paper.taf?abstract_id=264528*, acedido em 01 de Dezembro de 2008.

GRANT, R. M. (1991), "The resource-based theory of competitive advantage: implications for strategy formulation", *California Management Review*, Vol. 33, nº. 3, pp. 114-135.

GRANT, ROBERT (2006), "The Kowledge-Based View of the Firm", In: Faukner, D.; Campbell, A. (2006), "Strategy: A Strategy Overview and Competitive Strategy", Oxford, University Press, pp. 203-227.

GRAY, P. H.; MEISTER, D. B. (2006), "Knowledge Sourcing Methods", *Information and Management,* Vol..43, nº 2, pp.142-156.

GREINER, MARTINA E.; BOHMANN, TILO; KRCMAR, HELMUT (2007), "A Strategy For Knowledge Management", *Journal of Knowledge Management,* Vol. 11, nº 6, pp 3-15.

HELFAT, CONSTANCE E.; RAUBITSCHEK, RUTH S. (2002). "Product Sequencing – Coevolution of knowledge, capabilities, and products", In: Choo, Chun Wei; Bontis, Nick (eds.), "The strategic management of intellectual capital and organizational knowledge", Oxford University Press, New York, pp. 317-337.

HARLOW, HAROLD (2008), "The Effect of Tacit Knowledge on Firm Performance", *Journal of Knowledge Management* Vol. 12, nº 1, pp. 148-163.

HOFER, C e SCHENDELL, D (1978), "Strategy formulation: Analytical Concepts", St Paul, MN: West.

HUBER, GEORGE P. (1999), "Facilitating project team learning and contributions to organizational knowledge", *Creativity and Innovation Management***,** Vol. 8, n.º 2 (Junho), pp. 70-76.

HUBER, GEORGE P. (2001), "Transfer of knowledge in knowledge management systems: unexplored issues and suggested studies", *European Journal of Information Systems,* Vol. 10, Issue 2 (Junho).

LAI, H.; CHU, T. H. (2002), "Knowledge Management: a review of industrial cases", *Journal of Computer Information Systems,* Vol. 42, nº 5, pp. 26-39.

LEARNED, E.; CHRISTENSEN, R.C.; ANDREWS, K.R.; GUTH, W.D. (1969), "Business Policy: Text and Cases", Homewood, Ill, Irwin.

LOUREIRO, JOAQUIM LUÍS (2003), "Gestão do Conhecimento", 1ª Edição, Editora Centro Atlântico, Lisboa.

LU, LUCY; ETZKOWITZ, HENRY (2008), "Strategic Challenges for Creating Knowledge – Based Innovation in China", *Journal of Technology Management,* Vol. 3, nº 1, pp. 5-11.

LUNDVALL, B. (1992), "National Systems of Innovations: towards a theory of innovation and interactive learning", Printer, London.

MATUSIK, SHARON F. (2002), "Managing Public and Private Firm Knowledge within the context of Flexible Firm Boundaries", In: Choo, Chun Wei; Bontis, Nick (eds.), "The strategic management of intellectual capital and organizational knowledge", Oxford University Press, New York, pp. 605-617.

MCDONOUGH III, E.; ZACK, M.; BERDROW, I.(2008), "Integrating Innovation Style and Knowledge into Strategy"; *MIT Sloan Management Review*, Fall, pp. 53-58.

MELKAS, HELINA; HARMAAKORPI, VESA (2008), "Data, Information and Knowledge in Regional Innovation Networks", *European Journal of Innovation Management,* Vol. 11, nº 1, pp. 103-124.

NONAKA, IKUJIRO (1991), "The knowledge – creating company", *Harvard Business Review,* (Novembro - Dezembro), pp. 96-104.

NONAKA, IKUJIRO (1994), "A dynamic theory of organizational knowledge creation", In: Choo, Chun Wei; Bontis, Nick (eds.), "The strategic management of intellectual capital and organizational knowledge", Oxford University Press, New York, pp. 437-462.

NONAKA, IKUJIRO; TAKEUCHI, HIROTAKA (1995), **"**The Knowledge – Creating Company: How Japanese Companies create the Dinamics of Innovation", Oxford University Press, New York.

NONAKA, IKUJIRO; KONNO, NOBORU (1998), "The Concept of "Ba": building a foundation for knowledge creation", *California Management Review,* Vol. 40, n.º 3 (Primavera), pp. 40-54.

NONAKA, IKUJIRO; TOYAMA, RYOKO; BYOSIÈRE, PHILIPPE (2001), "A theory of organizational knowledge creation: understanding the dynamic process of creating knowledge", In: Dierkes, Meinolf; Berthoin Antal, Ariane; Child, John; Nonaka, Ikujiro (eds), "Handbook of organizational learning & knowledge", Oxford University Press, New York, pp. 490-517.

OCASIO, W e JOSEPH, J. (2008), "Rise and Fall-or transformation? The evolution of strategic planning at the General Electric Company, 1940-2006", *Long Range Planning*, Vol. 41, pp. 248-272.

PAXTON, JENNIFER GREGAN; JOHN, DEBORAH ROEDDER (1997), "Consumer learning by analogy: a model of internal knowledge transfer", *Journal of Consumer Research,* Vol. 24 (Dezembro), pp. 266-284.

PORTER, M. (1980), "Competitive Strategy", The Free Press, New York

PORTER, M. (1996), "What is strategy?", *Harvard Business Review*, November-December, pp. 61-78.

PRAHALAD, C.K.; HAMEL, G. (1990), "The core competence of the corporations", *Harvard Business Review, May-June*, pp.79-91.

SERRANO, ANTÓNIO; FIALHO, CÂNDIDO (2003), "Gestão do Conhecimento – o novo paradigma das organizações", FCA – Editora de Informática, Lisboa.

SILVA, RICARDO VIDIGAL DA; NEVES, ANA (2003), "Gestão de Empresas na Era do Conhecimento", Edições Sílabo, Lisboa.

SHER, J. P.; LEE, C. V. (2004), "Information Technology as a Facilitator for Enhancing Dynamic Capabilities through Knowledge Managements", *Information and Management,* Vol. 41, nº 8, pp.933-945.

TIRPAK, T.M. (2005), "Five Steps to Effective Knowledge Managements", *Research Technology Management,* Vol. 48, nº 3, pp. 15-16.

SIMÕES, JORGE (2008), "Transferência do Conhecimento no Ensino Superior Público em Portugal", *Revista Universo Contábil,* Vol. 4, nº1, pp. 95 – 113.

SVEIBY, KARL ERIK (2000), "Capital intelectual – La nueva riqueza de las empresas, Como medir y gestionar los activos intangibles para crear valor", Maxima Laurent du Mesnil Éditeur, Paris.

TALLMAN, S. (2006) in Faukner, D.; Campbell, A (eds), "A strategy of overview and competitive strategy", Oxford, university press.

TODTLING, F. (1991), "Spatial differentiation of innovation-locational and structural factors" in Bergman, E., Maier, G; Todling, F. (eds), "Regions Reconsidered, economics networks, innovation and local development in industrialized countries", Ed. Villiers House.

WINKELEN, C., FORUM, H.; JORDAN, W. (2008), " Building the Capability to be Innovative"; *KM Review*, Vol. 11, nº. 3, pp. 8-13.

WONG, K.Y.; ASPINWALL, E. (2004), "Knowledge Management Implementation Framework: a Review", *Knowledge and Process Management,* Vol. 11, nº 2, pp 93-104.

ZACK, M. (1999), "Developing a knowledge strategy ", *Califórnia Management Review*; Vol. 41, nº. 3, pp. 125-145.

ZAHRA, A. Z.; GEORGE, G. (2002), "Absorptive Capacity: a review, reconceptualization, and extension", *Academy of Management Review,* Vol. 27, nº 2, pp 185-203.

PATRONES DE DIFUSIÓN INTERNACIONAL Y SECTORIAL DE LA TECNOLOGÍA ESPAÑOLA

ANA MARÍA FERNÁNDEZ PÉREZ
E-mail: ana.fernandez@uca.es

MARÍA DOLORES LEÓN RODRÍGUEZ
Facultad de Ciencias Económicas y Empresariales
Departamento de Economía General
Universidad de Cádiz, España

Resumen

El presente trabajo analiza el proceso de difusión tecnológica considerando su dimensión internacional y sectorial y utilizando para ello datos relativos a patentes y citas de patentes concedidas a inventores españoles por parte de la Oficina de Patentes de los Estados Unidos. Los resultados muestran que los inventores españoles se nutren principalmente de conocimiento procedente de patentes estadounidenses, japonesas y alemanas. En cuanto al destino geográfico del conocimiento español éste fluye igualmente hacia EE. UU., Japón y Alemania mayoritariamente. En cuanto a los flujos sectoriales de conocimiento el análisis realizado pone de manifiesto una escasa variabilidad tecnológica entre las diferentes categorías o clases, tanto en procedencia como en destino. ABSTRACT: This paper analyzes the process of technology diffusion considering its international and sectorial dimension, thereby using data on patents and citations of patents granted to Spanish inventors by the U.S. Patent Trade Office. The results show that Spanish inventors are fed mainly from knowledge of U.S., Japanese and German patents. As for the geographical destination of Spanish knowledge it also flows mostly to U.S., Japan and Germany. In terms of knowledge flows on sectorial analysis shows a low variability between the different technology categories or classes, both in origin and destination.

Palabras Clave: Difusión del conocimiento / patentes / citas de patentes / dimensión internacional y sectorial.
Área temática: Economía de la Información y del Conocimiento.

1. INTRODUCCIÓN

Muchos países han elegido estimular el conocimiento y el desarrollo tecnológico como medio para promover el crecimiento económico y la mejora del bienestar social. Pero la cuestión básica para la mejora de la competitividad de los países se encuentra no solo en la generación de conocimiento tecnológico sino también en el entendimiento de los mecanismos de difusión que tienen lugar en los distintos ámbitos económicos y que provocan flujos de conocimiento en general desde los agentes económicos generadores hacia terceros agentes. Es crucial, por tanto, entender la dimensión geográfica y sectorial de la difusión del nuevo conocimiento creado.

El objetivo de este trabajo es contribuir a la literatura empírica existente sobre el proceso de generación y difusión del conocimiento tecnológico, utilizando los documentos de patentes como una medida de la generación de nuevo conocimiento y las citas de patentes como instrumento para determinar los flujos tecnológicos que tienen lugar entre países y entre sectores económicos. Para ello analizamos un conjunto de datos extraídos del National Bureau

of Economic Research (NBER) relativos a patentes y citas de patentes concedidas por la Oficina de Patentes y Marcas de los EE. UU. (USPTO) a inventores españoles.

Este estudio aporta varios elementos novedosos. En primer lugar no tenemos constancia de investigaciones empíricas previas desarrolladas para el caso de los flujos de conocimiento tecnológicos de las patentes generadas por inventores españoles. En segundo lugar, presenta una revisión sobre cómo se difunde el conocimiento tecnológico en las patentes españolas y hacia dónde se dirige dicho conocimiento. Además constituye una novedad el intentar identificar la difusión del conocimiento tecnológico desde un punto de vista internacional y sectorial. Nuestras principales cuestiones de investigación están relacionadas con la difusión geográfica y sectorial del conocimiento tecnológico tratando de dar respuesta a las dos cuestiones siguientes: ¿De dónde procede desde un punto de vista geográfico y sectorial el conocimiento tecnológico utilizado para desarrollar las patentes de inventores españoles? ¿Qué destino geográfico y sectorial tiene el conocimiento generado en patentes de inventores españoles?.

Nuestra metodología empírica utiliza un análisis de datos relativos a patentes y citas de patentes en orden a contestar las cuestiones planteadas en la investigación, haciendo para ello el supuesto, al igual que muchos otros trabajos, de que éstas representan una unidad de medida del conocimiento tecnológico que es utilizado y, por tanto, considerado de utilidad, para el desarrollo de la tecnología patentada.

Este trabajo se ha organizado como sigue. A continuación se revisan los argumentos teóricos existentes en la literatura reciente en torno a la generación de conocimiento tecnológico y su difusión para posibilitar un mayor crecimiento económico en los países y las regiones aportando también evidencia empírica relativa a la dimensión espacial del proceso de difusión del conocimiento tecnológico. En el tercer epígrafe se analizan los aspectos metodológicos relativos a la utilización de las patentes y citas de patentes como indicador de los flujos de conocimiento, justificando su uso a través de una amplia revisión de estudios previos. En el epígrafe cuarto se realiza un análisis sobre la dimensión internacional y sectorial del proceso de difusión tecnológica. Por último se exponen las principales conclusiones obtenidas.

2. DIFUSIÓN TECNOLÓGICA Y CRECIMIENTO ECONÓMICO

La razón básica para estudiar la transferencia de conocimiento en general y de conocimiento tecnológico en particular se encuentra en que éste contribuye a mejorar las actividades innovadoras de las empresas y, de ahí, a generar posibilidades de crecimiento económico en los países y en las regiones. La interrelación entre crecimiento económico y cambio técnico es claramente aceptada entre académicos, políticos y analistas en el ámbito regional, nacional e internacional y viene dada por un proceso de generación y difusión del conocimiento que desemboca en el desarrollo de innovaciones en los procesos productivos y en los productos terminados. Por ello en los últimos años los economistas han dedicado gran atención a los estudios relativos a la tecnología en general y al conocimiento y su transmisión en particular como uno de los factores determinantes del crecimiento económico y del progreso. La literatura existente en torno a la relación entre la tecnología y el crecimiento económico es extensa, y abarca desde el modelo básico de crecimiento neoclásico (Solow, 1956) hasta las modernas teorías del crecimiento endógeno (Romer, 1986, 1990; Lucas, 1988; Grossman y Helpman, 1991).

Si bien los procesos de invención e innovación son precondiciones necesarias para el desarrollo de una nueva tecnología, es su proceso de difusión el que determinará su alcance y puesta en uso productivo, lo que condicionará en última instancia el nivel de dinamismo tecnológico en la empresa, industria o economía. Desde que se publicara el trabajo pionero de Schumpeter (1934) han sido muy abundantes las investigaciones que han tratado de modelizar los procesos de difusión tecnológica. Si bien inicialmente se prestó más atención a los procesos de invención e innovación en sí, dicha percepción comenzó a cambiar a partir de los años 50, alcanzando las investigaciones relativas a la difusión tecnológica un lugar prominente en la literatura científica (Silverberg, 1990; Freeman, 1994), pudiéndose caracterizar la mayoría de

los análisis teóricos acerca del proceso de difusión tecnológica en torno a los modelos de equilibrio neoclásicos o la perspectiva relativa a los modelos de desequilibrio evolucionista[1].

La premisa básica para entender la difusión tecnológica que puede tener lugar entre diversos agentes económicos se encuentra en la primera generación de teorías del crecimiento endógeno (Romer, 1990; Grossman y Helpman, 1991; Aghion y Howitt, 1992, 1998) que enfatizan la naturaleza no rival de la tecnología y sus posibilidades de difusión (externalidad positiva) como atributos que la distinguen del resto de factores productivos tradicionales -capital físico y capital humano- permitiendo que los beneficios del progreso tecnológico sean compartidos. En este sentido, el conocimiento difiere de la concepción habitual del resto de factores productivos de modo que cuando un agente económico (empresa) utiliza determinado conocimiento éste no puede ser apropiado de forma exclusiva, ni puede, por tanto, impedir su uso por parte de terceros agentes (en ausencia al menos de protección legal).

Las fuentes extranjeras de tecnología son, por otra parte, un importante factor que contribuye al crecimiento de la productividad de los países especialmente en aquellos menos desarrollados que dedican una muy reducida proporción de gasto propio a I+D, por lo que el grado de difusión de las tecnologías generadas en países más avanzados suponen una cuestión clave para el crecimiento de su productividad. En este sentido, las investigaciones de Xu y Chiang (2005) sobre los efectos de la difusión tecnológica internacional en una muestra de países que diferencian en tres grupos de renta concluye que todos los países independientemente de sus niveles de renta se benefician de la difusión de tecnología extranjera, aunque hay diferencias en los canales de transmisión. Igualmente Canepa y Stoneman (2004) aprecian diferencias internacionales en la velocidad y en el alcance del proceso de difusión a través de distintas tecnologías. Esta materia ha sido ampliamente estudiada en diversos trabajos en los que se sugieren varios canales por los cuales la tecnología puede ser transmitida a otros países y que se detallan a continuación.

En primer lugar, la transferencia tecnológica se facilita a través de los contactos comerciales. La tecnología incorporada en los bienes intermedios y de capital a través de las importaciones directas de dichos bienes suponen un canal principal de transmisión. Los exportadores, por su parte, pueden venir influenciados por los contactos con clientes extranjeros que demandan nuevos productos y nuevas soluciones técnicas. Sin embargo, al considerar la influencia ejercida por las actividades exportadoras éstas no muestran efectos empíricos relevantes como canal de difusión tecnológica para las ganancias de productividad (Keller, 2004).

En segundo lugar, las inversiones directas extranjeras realizadas por parte de empresas multinacionales suponen otro canal para la transmisión de tecnología, razón por la cual dichas inversiones son favorecidas desde instancias políticas en los países receptores. Hay una extensa literatura sobre los mecanismos a través de los que las inversiones directas extranjeras intensifican los flujos de tecnología a través de las fronteras internacionales (Keller y Yeaple, 2003; Aitken y Harrison, 1999; Xu, 2000; Laursen y Meliciani, 2002). Así, las subsidiarias de empresas multinacionales instaladas en terceros países utilizan conocimientos generados en el país de origen tanto al importar bienes intermedios que utilizan en su proceso productivo como al utilizar el conocimiento generado por la I+D desarrollada en dicho país de origen. Adicionalmente la instalación de empresas multinacionales en terceros países conlleva igualmente el traslado físico de personal, normalmente cualificado, que podrá ejercer actividades de colaboración y cooperación con empresas locales con la consiguiente transmisión de conocimientos (tanto a través de canales formales como informales). Son, por tanto, mecanismos de difusión derivados de los efectos demostración (por observación o imitación de

[1] En el trabajo de Sarkar (1998) se puede consultar una extensa revisión tanto de la perspectiva neoclásica como de las teorías evolucionistas que modelizan los procesos de difusión tecnológica. En él se examina la literatura existente sobre la caracterización del proceso de difusión tecnológica evaluándose las perspectivas que modelizan las distintas alternativas teóricas con la evidencia empírica de los procesos de difusión y tratando de proporcionar una amplia revisión de ambas perspectivas dentro de un marco unificado de análisis. La investigación esencialmente concluye que ambas alternativas no son mutuamente excluyentes como habitualmente se han considerado sino que, por el contrario, pueden solaparse y complementarse para proporcionar un mejor entendimiento del proceso de difusión tecnológica entre las industrias y los países.

los productos o las prácticas desarrolladas por las empresas extranjeras), de la movilidad de los trabajadores y de las relaciones entre suministradores y clientes extranjeros.

En tercer lugar, la I+D desarrollada en países avanzados puede originar y generar conocimientos e ideas que podrán ser utilizados en el proceso productivo de empresas distintas a las que desarrollaron inicialmente dichas investigaciones, las cuales pueden además estar localizadas en el mismo país o en terceros países, dando lugar en este caso a una transmisión internacional de conocimientos. Esta posibilidad de generación y transmisión de la tecnología se puede manifestar de manera expresa en el desarrollo de formas codificadas de conocimiento como las patentes, cuyo uso a través de sus citas facilitan la difusión del conocimiento[2] (Eaton y Kortum, 1996; Xu y Chiang, 2005) y pueden ser descritas como canales de difusión tecnológica activa por diferencia al comercio de bienes y las inversiones extranjeras que se consideran formas de difusión pasiva (Abreu et al., 2004).

Varios trabajos analizan empíricamente el impacto de la transferencia de tecnología extranjera sobre el crecimiento de la productividad a través de sus canales de transmisión. El trabajo de Savvides y Zachariadis (2005) contribuye a la literatura existente sobre las externalidades positivas de la difusión de la tecnología extranjera para el crecimiento de la productividad doméstica, analizando si la intensidad de la I+D extranjera, las importaciones de bienes de capital y las inversiones directas extranjeras son determinantes significativos del crecimiento del factor de productividad total y del valor añadido en el sector manufacturero. Sus descubrimientos avalan el impacto positivo de estos factores, principalmente de la intensidad de la I+D extranjera. El trabajo de Keller (2002) evalúa empíricamente cómo el conocimiento tecnológico es transmitido a otras industrias (tanto domésticas como extranjeras) a través del comercio de bienes intermedios. Sus resultados sugieren que la inversión en I+D de la propia industria es la principal fuente de incrementos de la productividad, siguiéndole en importancia la I+D en otras industrias domésticas y la I+D extranjera. En un trabajo posterior Keller y Yeaple (2003) analizando la difusión tecnológica vía importaciones e inversiones directas extranjeras sugieren que ambas conducen a ganancias de productividad, aunque de mayor cuantía en el caso de éstas últimas. Xu y Chiang (2005) descubren que las patentes extranjeras generan difusión tecnológica principalmente en los países de rentas medias y bajas, pero menos en los de renta alta. Por el contrario la tecnología extranjera incorporada en los bienes de capital importados tienen un efecto significativo en los países de renta media y alta. Varios trabajos más valoran la influencia de estos factores como determinantes de la difusión tecnológica (Eaton y Kortum, 2001; Coe y Helpman, 1995; Xu y Wang, 1999; Xu, 2000; Globerman et al., 2000; Markusen, 2002).

3. ASPECTOS METODOLÓGICOS RELATIVOS A LA UTILIZACIÓN DE LAS CITAS DE PATENTES

Uno de los principales obstáculos o inconvenientes de los estudios sobre tecnología y difusión tecnológica es la identificación y medida de la difusión de la I+D, es decir, los efectos positivos o externalidades positivas que una industria o empresa recibe de la actividad de I+D desarrollada por otra. Esto se ha solucionado recientemente, con mayor o menor acierto y asumiendo sus bondades y limitaciones utilizando la información contenida en los documentos de patentes y citas de patentes.

Una patente es un documento proporcionado por autoridades legales que reconoce un monopolio temporal para la explotación comercial de una innovación[3]. Los inventos patentados incorporan porciones de conocimiento tecnológico generado en organismos e instituciones públicas y privadas de distinta naturaleza jurídica (empresas, universidades, centros de investigación, etc.) que sirven de base tecnológica para generar el conocimiento que se desea

[2] Sería el caso igualmente de las citas en artículos científicos.

[3] Las solicitudes de patentes muestran la capacidad para producir innovaciones, pudiéndoseles considerar un indicador de la fortaleza innovadora de un país. Las patentes solicitadas son, por tanto, un indicador ampliamente utilizado como criterio para medir la actividad innovadora de los países, no solo en cantidad sino también en calidad.

patentar. Dichas porciones de conocimiento son las citas a otras patentes y a literatura científica que describen la parte de otros conocimientos que han sido utilizados como apoyo para la invención en cuestión y sobre los cuales no se pueden hacer reivindicaciones. Es lo que se conoce como el "estado de la técnica" previo a la invención patentable. Las solicitudes de nuevas patentes incluyen, por tanto, referencias a patentes previas que son esenciales para el desarrollo de la nueva innovación. Estas citas de patentes son añadidas a los documentos de patentes cuando el conocimiento intrínseco en la patente citada es relevante para el conocimiento patentado por la patente que cita. Por tanto el hecho de que una patente B cite a la patente A puede ser indicativo de un flujo de conocimiento desde la patente A hacia la patente B (Hall et al., 2001).

Si bien tradicionalmente numerosos trabajos han usado datos de patentes como variable para valorar los resultados de la innovación (Pavitt, 1988; Griliches, 1990), recientemente se ha generalizado el uso de la información contenida en las citas de patentes. Al no existir ninguna otra medida de la influencia que ejercen unas invenciones sobre otras las citas de patentes se han convertido en una atractiva variable para el estudio de los flujos de conocimiento. La información procedente de las citas contenidas en los documentos de patentes (citas hechas y citas recibidas respectivamente) resulta muy valiosa para el investigador como medida de los flujos de conocimiento tecnológico. Las citas recogidas en un documento de patente constituyen las citas hechas y pueden ser usadas para trazar o determinar la influencia de invenciones pasadas a través del tiempo y del espacio, proporcionando valiosa información acerca de los flujos de conocimiento que se producen. Estos datos han sido utilizados para investigar modelos de difusión del conocimiento tanto sectorial como espacial. A continuación hacemos una revisión de los principales trabajos desarrollados y que utilizan la información proporcionada por las citas de patentes como medida de los flujos tecnológicos.

Es pionero el trabajo de Jaffe et al. (1993) que usando datos de citas compara la localización geográfica del inventor (primer inventor) con el origen geográfico de la patente citada para medir el alcance de la difusión tecnológica dentro de los EE. UU. En la misma línea Jaffe y Trajtenberg (1996) utilizan citas de patentes para comparar la magnitud de los flujos tecnológicos a través de los países. MacGarvie (2005) utiliza las citas para medir la difusión internacional del conocimiento tecnológico, descubriendo que dicha difusión se intensifica por la proximidad física y por compartir un idioma común. Globerman et al. (2000) examinan los canales para la difusión de conocimiento extranjero hacia Suecia distinguiendo entre empresas multinacionales y pequeñas y medianas empresas locales. Sus resultados muestran que ambos tipos de empresas hacen más referencias a países con un gran stock de patentes y próximos a Suecia. MacGarvie (2006) descubre, utilizando citas de patentes de empresas francesas, que las invenciones de exportadores e importadores tienen más probabilidad de estar influenciadas por tecnología extranjera que las invenciones de empresas que no se comprometen en actividades de comercio internacional, aportando evidencia sobre el modo en que las empresas que desarrollan actividad internacional aprenden de la tecnología extranjera.

Los flujos de conocimiento que se dan entre sectores aportan también evidencia en torno a lo que podemos denominar localización o proximidad tecnológica, entendiendo con ello la difusión sectorial del conocimiento que tiene lugar a través de las citas de patentes (Jaffe y Trajtenberg, 1996; Lee et al., 2007; Laursen y Meliciani, 2002). Algunos trabajos muestran, en general, que una patente cita con mayor probabilidad a otra patente de su misma clase tecnológica. Duguet y Macgarvie (2005) revelan en su investigación que las citas, además de estar geográficamente concentradas lo están también dentro de las mismas clases tecnológicas. Igualmente Jaffe y Trajtenberg (1999) encuentran que las patentes dentro de la misma clase tecnológica son citadas con mayor probabilidad que las patentes de distinta clase tecnológica. También Macgarvie (2005) confirma esta cuestión. Gay et al. (2005), por su parte, concluyen que una patente que es frecuentemente citada es más citada por patentes de otros campos tecnológicos, lo que indica que la patente citada tiene una difusión sectorial más amplia y por tanto una importancia tecnológica mayor. Maurseth (2005) demuestra en la misma línea que las patentes que reciben citas de campos tecnológicos diversos tienen una tasa de supervivencia

mayor que otras patentes. Por el contrario, las patentes que reciben citas dentro del mismo campo tecnológico decaen antes en su uso.

4.- DIFUSIÓN TECNOLÓGICA DEL CONOCIMIENTO

El análisis de la difusión tecnológica del conocimiento que se realiza a continuación utiliza los datos extraídos de la base de datos NBER[4] de patentes estadounidenses, de las que se han extraído las patentes cuyo primer inventor es español[5]. El estudio se realiza desde dos puntos de vista simultáneamente, citas hechas y citas recibidas, lo que posibilita analizar tanto la procedencia como el destino del conocimiento tecnológico utilizado por las patentes de inventores españoles. Según muestra la Tabla 1 el conjunto de datos utilizado para el análisis de las citas hechas comprende un total de 2065 patentes cuyo año de concesión se encuentra entre 1985 y 1999. Este conjunto de patentes realizó en total 11025 citas de las que se ha podido extraer información[6], lo que supone un total de 5,34 citas hechas por término medio por cada patente. Por su parte, el conjunto de datos correspondiente al análisis de las citas recibidas comprende un total de 2269 patentes cuyo año de concesión se encuentra entre 1963 y 1999 que reciben un total de 9306 citas[7], lo que supone un total de 4,10 citas recibidas por término medio por cada patente.

La Tabla 1 nos muestra además la distribución del total de patentes según el número de citas hechas y recibidas. Si atendemos a las referencias citadas por cada patente podemos observar que la mayor parte, un 49,78 por ciento de las patentes, realizan hasta un máximo de 5 citas. El 35 por ciento de las patentes realizan entre 6 y 10 citas, seguidas por un 11,23 por ciento de las patentes con un número de citas comprendido entre 11 y 15. El porcentaje de patentes que realizan más citas es poco relevante. En cuanto a las citas recibidas éstas son menores. El 77,17% de las patentes reciben entre 1 y 5 citas seguidas por un 16% de las patentes que reciben hasta un máximo de 10 citas. El dato acumulado correspondiente al número de patentes que reciben más de 10 citas es únicamente del 6,87%.

TABLA 1. Distribución de patentes según el número de citas

	CITAS HECHAS	CITAS RECIBIDAS
Nº Patentes	2065	2269
Nº Citas	11025	9306
Nº medio de citas por patente	5.34	4.10
% de patentes según número de citas:		
Entre 1 y 5 citas	49.78	77.17
Entre 6 y 10 citas	35.01	15.95
Entre 11 y 15 citas	11.23	4.05
Entre 16 y 20 citas	2.37	1.59
Más de 20 citas	1.60	1.23
TOTAL	100	100

Para la realización de un análisis estrictamente riguroso de los datos extraídos de patentes y de citas de patentes que nos lleve al conocimiento de los flujos tecnológicos que tienen lugar tanto desde un punto de vista espacial como sectorial hubiera sido deseable eliminar del

[4] La estructura de esta fuente de datos es descrita en profundidad en Hall et al. (2001).

[5] La base de datos de citas de patentes de NBER asigna la nacionalidad de cada patente geográficamente atendiendo al domicilio del primer inventor. En este trabajo, ante la imposibilidad de disponer de información adicional asumimos este mismo criterio.

[6] El número de citas realizadas por dicho conjunto de patentes es mayor, concretamente 12635, pero no se dispone de información sobre dichas citas por corresponder a referencias anteriores a 1975. En el conjunto de datos utilizado se ha optado por un margen de 10 años, utilizando por tanto las patentes correspondientes al período 1985-1999 para garantizar la disponibilidad de información relativa a un cierto número de citas hechas.

[7] En este caso la información disponible sobre las citas recibidas corresponde a la fecha de cierre de la base de datos, por lo que es necesario tener en cuenta que las patentes más recientes disponen razonablemente de un número menor de citas recibidas que las patentes más antiguas, lo que introduce ciertamente desviación en el análisis resultante que deberá ser tenido en cuenta en la interpretación de los resultados.

conjunto de datos disponibles las autocitas realizadas. Éstas son citas hechas o recibidas por patentes cuyo solicitante es la misma empresa o institución que la patente en cuestión, por lo que no se puede considerar que en la utilización de tal conocimiento haya existido difusión o flujo tecnológico. En el conjunto de datos que venimos utilizando no siempre está disponible la información relativa a todos los solicitantes de las citas hechas o recibidas, por lo que no ha sido posible eliminar dichas autocitas.

4.1. Difusión Internacional

En la Tabla 2 se recoge la información relativa a la procedencia geográfica del conocimiento utilizado por las patentes de inventores españoles a través de las citas hechas. Como podemos observar, el mayor porcentaje de citas hechas por el conjunto de patentes considerado corresponde a Estados Unidos, con un 57,37% de las referencias hechas, seguidas a gran distancia por las citas hechas a patentes de Japón y Alemania, con valores del 10,56 y 9,40% respectivamente. Considerando conjuntamente, como información complementaria, el porcentaje de citas a países del entorno[8] este valor asciende al 21,67%. Por su parte, las citas hechas a España suponen un 3,15%. Si consideramos en lugar del número de citas, el número de patentes según la procedencia geográfica de las citas la información resulta igualmente concluyente. Más del 84% de las patentes citan a otras patentes estadounidenses, seguidas de un 31,09% de las patentes que utilizan conocimientos de Japón y de casi el 30% de las patentes que utilizan conocimientos de Alemania. Considerando el conjunto de países del entorno este porcentaje de eleva hasta el 80,34% de las patentes, que hacen referencia a otras patentes de países cercanos.

TABLA 2. Procedencia geográfica del conocimiento: citas hechas

	NÚMERO DE CITAS		NÚMERO DE PATENTES	
	Valor absoluto	Porcentaje	Valor absoluto	Porcentaje
ESTADOS UNIDOS	6325	57.37	1743	84.41
JAPÓN	1164	10.56	642	31.09
CANADA	216	1.96	172	8.33
ESPAÑA	347	3.15	269	11.03
ALEMANIA	1036	9.40	609	29.49
FRANCIA	423	3.84	318	15.40
GRAN BRETAÑA	362	3.28	284	13.75
ITALIA	227	2.06	179	8.67
SUIZA	200	1.81	158	7.65
SUECIA	141	1.21	111	5.38
PAÍSES ENTORNO	2389	21.67	1659	80.34
OTROS PAÍSES	584	5.30	493	23.87
TOTAL	11025	100	2065 (*)	-

(*) El dato total no coincide con la suma algebraica ya que una misma patente cita a varias patentes de diversos países.

Los datos de la Tabla 3 ponen de manifiesto el destino geográfico del conocimiento de patentes de inventores españoles que es utilizado por otras patentes. El mayor porcentaje de citas recibidas corresponde, como era de esperar, a los Estados Unidos, con un 57,47%, seguido a gran distancia por las citas recibidas de patentes japonesas y alemanas. Al igual que se manifestó al analizar la procedencia del conocimiento, si se considera el conjunto de países del entorno el porcentaje de citas recibidas es ciertamente significativo, suponiendo un total del 18,92% de las citas recibidas. Atendiendo al número de patentes, y considerando que cada patente recibe citas de varias patentes de diversos países las conclusiones son similares, confirmando el destino geográfico del conocimiento hacia Estados Unidos, seguido de Japón y de Alemania, aunque al considerar el conjunto de países del entorno éstos son el principal destino geográfico tras los Estados Unidos.

[8] Como países del entorno se han considerado aquellos países próximos cuyo stock de patentes citadas es más importante (Alemania, Francia, Gran Bretaña, Italia, Suiza y Suecia). En el resto de países europeos el número de citas resulta irrelevante.

TABLA 3. Destino geográfico del conocimiento: citas recibidas				
	NÚMERO DE CITAS		NÚMERO DE PATENTES	
	Valor absoluto	Porcentaje	Valor absoluto	Porcentaje
ESTADOS UNIDOS	5348	57.47	1707	75.23
JAPÓN	865	9.30	515	22.70
CANADA	216	2.32	165	7.27
ESPAÑA	423	4.55	318	14.01
ALEMANIA	716	7.69	479	21.11
FRANCIA	328	3.52	255	11.24
GRAN BRETAÑA	224	2.41	185	8.15
ITALIA	227	2.44	175	7.71
SUIZA	181	1.94	146	6.43
SUECIA	85	0.91	69	3.04
PAÍSES ENTORNO	1761	18.92	1309	57.69
OTROS PAÍSES	693	7.45	546	24.06
TOTAL	9306	100	2269 (*)	-
(*) El dato total no coincide con la suma algebraica ya que una misma patente recibe cita de varias patentes de diversos países.				

4.2. Difusión Sectorial

El análisis de la difusión sectorial del conocimiento de las patentes de inventores españoles lleva a considerar la posibilidad de que una patente de un sector determinado utilice conocimientos procedentes de otra patente del mismo sector o de un sector diferente. Del mismo modo el destino del conocimiento aportado por las patentes españolas consideradas puede ser el mismo sector de la patente en cuestión u otro sector distinto. Es lo que se puede denominar variabilidad tecnológica del conocimiento[9].

Atendiendo a la procedencia tecnológica del conocimiento los datos aportados por la Tabla 4 muestran que, en general, en el 80,27% de las citas no hay variabilidad tecnológica en la categoría correspondiente, es decir, la patente citada y la patente en cuestión pertenecen a la misma categoría tecnológica, mientras que en el 19,73% de las citas hay variabilidad tecnológica, por lo que la patente en cuestión utiliza conocimientos tecnológicos de una patente perteneciente a una categoría tecnológica diferente. En función de la categoría considerada los datos oscilan más o menos a favor de la variabilidad tecnológica (las categorías de mayor variabilidad son, por este orden, las categorías 2, 1 y 3 correspondientes a *Informática y Comunicaciones*, *Química* y *Farmacia y Medicina*). Si se considera la variabilidad tecnológica por subcategorías estos porcentajes oscilan notablemente (Tabla 5) como es lógico esperar.

TABLA 4. Variabilidad tecnológica por categorías (citas hechas)		
	VARIABILIDAD TECNOLÓGICA (%)	NO VARIABILIDAD TECNOLÓGICA (%)
CAT 1	23.83	76.17
CAT 2	25.11	74.89
CAT 3	23.65	76.35
CAT 4	19.28	80.72
CAT 5	17.97	82.03
CAT 6	16.67	83.33
TOTAL	19.73	80.27
Cat 1: Química; Cat 2: Informática y Comunicaciones; Cat 3: Medicina y Farmacia; Cat 4: Eléctrica y Electrónica; Cat 5: Mecánica; Cat 6: Otras		

[9] En el Anexo I figura la relación detallada de todas las categorías y subcategorías tecnológicas consideradas.

El análisis de la procedencia del conocimiento se completa con los datos aportados por la Tabla 6 relativa a la procedencia tecnológica de cada cita hecha. En ella se puede consultar la procedencia tecnológica de los conocimientos utilizados en cada categoría, observándose que cada una se nutre principalmente de conocimiento tecnológico de su propia categoría y únicamente en porcentaje residual utiliza conocimientos aportados por patentes de otras categorías tecnológicas. Así, al margen de los conocimientos aportados por la categoría propia, las patentes de la categoría *Química* se nutre también de conocimientos de la categoría *Otros*. La categoría *Informática y Comunicaciones* se nutre principalmente de *Eléctrica y Electrónica*. La categoría *Medicina y Farmacia* utiliza conocimientos aportados por la categoría *Química*. La categoría *Eléctrica y Electrónica* principalmente se nutre de *Mecánica*. Por su parte la categoría *Mecánica* usa conocimientos de la categoría *Otros* y, esta última cita principalmente patentes de la categoría *Mecánica*.

TABLA 5. Variabilidad tecnológica por subcategorías (citas hechas)

SUBCAT	VARIABILIDAD TECNOLÓGICA (%)	NO VARIABILIDAD TECNOLÓGICA (%)	SUBCAT	VARIABILIDAD TECNOLÓGICA (%)	NO VARIABILIDAD TECNOLÓGICA (%)
11	57.85	42.15	45	35.83	64.17
12	61.49	38.51	46	50.00	50.00
13	30.00	70.00	49	36.42	63.58
14	38.17	61.83	51	30.50	69.50
15	26.81	73.19	52	36.60	63.40
19	26.71	73.29	53	30.15	69.85
21	27.19	72.81	54	37.18	62.82
22	42.03	57.97	55	31.95	68.05
23	35.59	64.41	59	22.94	77.06
24	66.67	33.33	61	21.43	78.57
31	38.55	61.45	62	6.73	93.27
32	14.63	85.37	63	19.39	80.61
33	37.62	62.38	64	18.18	81.82
39	12.31	87.69	65	25.93	74.07
41	23.40	76.60	66	36.62	63.38
42	22.22	77.78	67	42.77	57.23
43	31.94	68.06	68	19.46	80.54
44	28.21	71.79	69	35.08	64.92

Subcategorías: Anexo I

TABLA 6. Procedencia tecnológica del conocimiento (citas hechas)

	CAT 1	CAT 2	CAT 3	CAT 4	CAT 5	CAT 6	TOTAL
CAT 1	76.17	0.21	6.85	1.69	6.64	8.43	100
CAT 2	0.61	74.89	0.45	14.22	7.87	1.97	100
CAT 3	16.02	0.08	76.35	1.58	0.91	5.06	100
CAT 4	2.12	3.92	0.21	80.72	9.22	3.81	100
CAT 5	4.50	0.53	0.53	2.64	82.03	9.77	100
CAT 6	3.48	0.35	1.09	2.59	9.15	83.33	100
TOTAL	17.33	5.12	10.05	9.72	26.55	31.23	100

Cat 1: Química; Cat 2: Informática y Comunicaciones; Cat 3: Medicina y Farmacia; Cat 4: Eléctrica y Electrónica; Cat 5: Mecánica; Cat 6: Otras

En cuanto al destino que tiene el conocimiento generado en las patentes de inventores españoles, los datos aportados por la Tabla 7 muestran igualmente un porcentaje de variabilidad tecnológica en las citas recibidas del 19,90 por ciento, siendo por tanto más del 80 por ciento el número de citas recibidas por patentes pertenecientes a la misma categoría tecnológica de la patente en cuestión. Son en este caso las categorías 1 y 4, correspondientes a *Qímica* y *Eléctrica*

y Electrónica las que presentan los mayores porcentajes de variabilidad tecnológica y, por tanto, el mayor aprovechamiento de los conocimientos tecnológicos de la patente hacia otros sectores. Observando los datos por subcategorías tecnológicas que muestran la Tabla 8 se aprecian mayores oscilaciones en los porcentajes de variabilidad.

TABLA 7. Variabilidad tecnológica por categorías (citas recibidas)

	VARIABILIDAD TECNOLÓGICA (%)	NO VARIABILIDAD TECNOLÓGICA (%)
CAT 1	32.77	67.23
CAT 2	15.94	84.06
CAT 3	14.30	85.70
CAT 4	24.40	75.60
CAT 5	18.29	81.71
CAT 6	16.97	83.03
TOTAL	19.90	80.10

Cat 1: Química; Cat 2: Informática y Comunicaciones; Cat 3: Medicina y Farmacia; Cat 4: Eléctrica y Electrónica; Cat 5: Mecánica; Cat 6: Otras

Los datos de la Tabla 9 confirman el destino del conocimiento tecnológico de las patentes de inventores españoles. Igual que sucede al analizar la procedencia del conocimiento, el destino mayoritario del conocimiento generado por las patentes se dirige a otras patentes de la misma categoría tecnológica, quedando el uso de dicho conocimiento en otras categorías limitado a porcentajes de citas bastante reducidos. En particular es la categoría *Química* la que destina mayor parte de sus conocimientos a otras categorías, principalmente a la categoría *Medicina y Farmacia*. La categoría *Informática y Comunicaciones* por su parte destina conocimientos a la categoría *Eléctrica y Electrónica*. La categoría *Medicina y Farmacia* fundamentalmente a la categoría *Química*. La categoría *Eléctrica y Electrónica* a la categoría *Mecánica*. La categoría *Mecánica*, por su parte recibe citas principalmente de la categoría *Otras* y, por último, la categoría *Otras* de la categoría *Mecánica*.

TABLA 8. Variabilidad tecnológica por subcategorías (citas recibidas)

SUBCAT	VARIABILIDAD TECNOLÓGICA (%)	NO VARIABILIDAD TECNOLÓGICA (%)	SUBCAT	VARIABILIDAD TECNOLÓGICA (%)	NO VARIABILIDAD TECNOLÓGICA (%)
11	32.67	67.33	45	39.85	60.15
12	68.52	31.48	46	77.78	22.22
13	28.57	71.43	49	39.82	60.18
14	60.69	39.31	51	35.54	64.46
15	52.17	47.83	52	33.89	66.11
19	29.17	70.83	53	26.47	73.53
21	31.28	68.72	54	37.67	62.32
22	20.86	79.14	55	25.99	74.01
23	66.67	33.33	59	32.99	67.01
24	25.00	75.00	61	18.00	82.00
31	29.35	70.65	62	14.08	85.92
32	13.03	86.97	63	18.60	81.40
33	23.75	76.25	64	20.24	79.76
39	21.30	78.70	65	25.32	74.68
41	26.11	73.89	66	17.04	82.96
42	26.67	73.33	67	53.28	46.72
43	26.83	73.17	68	28.45	71.55
44	34.15	65.85	69	38.91	61.09

Subcategorías: Anexo I

TABLA 9. Destino tecnológico del conocimiento (citas recibidas)							
	CAT 1	CAT 2	CAT 3	CAT 4	CAT 5	CAT 6	TOTAL
CAT 1	67.23	0.71	12.55	2.13	9.01	8.37	100
CAT 2	1.03	84.06	0.26	6.94	4.88	2.83	100
CAT 3	8.84	0.08	85.70	0.48	1.93	2.97	100
CAT 4	2.28	5.90	0.27	75.60	10.59	5.36	100
CAT 5	4.77	1.00	0.71	3.90	81.71	8.92	100
CAT 6	3.93	0.74	2.29	1.55	8.47	83.03	100
TOTAL	14.14	4.61	14.35	8.01	26.67	32.23	100
Cat 1: Química; Cat 2: Informática y Comunicaciones; Cat 3: Medicina y Farmacia; Cat 4: Eléctrica y Electrónica; Cat 5: Mecánica; Cat 6: Otras							

Consideradas de forma conjunta, la información contenida en las Tablas 6 y 9 permiten establecer el canal de procedencia-destino del conocimiento tecnológico generado en una patente. Así, la Tabla 10 muestra, por un lado, que cada categoría se nutre mayoritariamente de conocimiento generado por patentes de su propia categoría e igualmente destina sus conocimientos a otras patentes de la misma categoría tecnológica y, por otro, que hay una correspondencia clara en todas las categorías, a excepción de la primera, entre la procedencia y el destino de los flujos de conocimiento.

TABLA 10. Flujos de conocimiento tecnológico sectorial		
	PRINCIPALES CATEGORIAS DE PROCEDENCIA DE LAS CITAS HECHAS	PRINCIPALES CATEGORÍAS DE DESTINO DE LAS CITAS RECIBIDAS
CAT 1	1-6	1-3
CAT 2	2-4	2-4
CAT 3	3-1	3-1
CAT 4	4-5	4-5
CAT 5	5-6	5-6
CAT 6	6-5	6-5
Cat 1: Química; Cat 2: Informática y Comunicaciones; Cat 3: Medicina y Farmacia; Cat 4: Eléctrica y Electrónica; Cat 5: Mecánica; Cat 6: Otras		

5. CONCLUSIONES

En los últimos años las políticas económicas se han dirigido a estimular el conocimiento y el desarrollo tecnológico como medio para promover el crecimiento económico de los países y las regiones. La cuestión básica se encuentra en el entendimiento de los mecanismos de difusión que tienen lugar y que provocan estos flujos. El objetivo del presente trabajo es contribuir al conocimiento de los mecanismos que subyacen en los procesos de difusión tecnológica, ya que si bien la invención e innovación son precondiciones necesarias para el desarrollo de una nueva tecnología, es su proceso de difusión el que determinará su alcance y puesta en uso productivo. Todos los países, independientemente de sus niveles de renta, se pueden beneficiar de la difusión de tecnología extranjera pudiendo haber, sin embargo, diferencias en sus canales de transmisión. Estos son principalmente los contactos comerciales, las inversiones directas extranjeras y la I+D desarrollada en países avanzados que puede generar conocimientos e ideas que se transmitan a través de formas codificadas de conocimiento. El uso de la información contenida en los documentos de patentes nos permite, asumiendo sus bondades y limitaciones metodológicas, hacer un primer sondeo que puede ser un complemento a otros procedimientos para el análisis de la difusión tecnológica y utilizar estos datos como variable para el estudio de los flujos de conocimiento y de sus modelos de difusión sectorial y espacial.

El análisis de datos de patentes extraídos del National Bureau of Economic Research (NBER) se realiza desde dos puntos de vista simultáneamente, citas hechas y citas recibidas, lo que posibilita la obtención de información tanto de procedencia como de destino del

conocimiento tecnológico utilizado por las patentes de inventores españoles. Del conjunto de citas de las que se ha podido extraer información se desprende que el número de citas hechas y recibidas por término medio por cada patente es 5. En este sentido, cerca del 50 por ciento de las patentes realizan entre 1 y 5 citas, siendo este porcentaje muy superior, más del 77 por ciento, en el caso de las citas recibidas. En cuanto a la difusión espacial de la tecnología patentada por inventores españoles, ésta se nutre principalmente de conocimiento procedente de EE. UU., seguido a gran distancia de Japón y Alemania, si bien al considerar conjuntamente los países de nuestro entorno este porcentaje es ciertamente significativo y superado únicamente por EE. UU. Los datos analizados permiten confirmar los mismos países como principal destino geográfico del conocimiento generado. El análisis de la difusión sectorial del conocimiento, por su parte, lleva a considerar la posibilidad de que una patente utilice conocimientos procedentes del mismo sector o de un sector diferente. Los datos permiten concluir tanto en procedencia como en destino del conocimiento, a través de sus citas hechas y recibidas respectivamente, una escasa variabilidad tecnológica en el conjunto de categorías mostrando porcentajes superiores al 80% correspondiente a los casos en los que la patente citada y la patente en cuestión pertenecen a la misma clase tecnológica. Igualmente el destino mayoritario del conocimiento generado por las patentes se dirige a otras patentes de la misma categoría. Considerada de forma conjunta esta información nos permite establecer, para terminar, el canal de procedencia-destino del conocimiento tecnológico generado. Cada categoría se nutre mayoritariamente de conocimiento de patentes de su propia categoría e igualmente destina sus conocimientos a otras patentes de la misma categoría tecnológica, existiendo una correspondencia clara en dichas categorías entre la procedencia y el destino de los flujos de conocimiento.

AGRADECIMIENTOS:

Los autores agradecen el apoyo financiero recibido de la Junta de Andalucía (P06-SEJ-02087 y P08-SEJ-03981) para la realización de este trabajo.

BIBLIOGRAFÍA

ABREU, M., DE GROOT, H. L., FLORAX, R. J. (2004), "Spatial Patterns of Technology Diffusion", Discussion Paper (Tinbergen Institute) TI 2004-079/3 pp. 1-31.

AGHION, P., HOWITT, P. (1992), "A Model of Growth Through Creative Destruction", Econometrica 60, pp. 323-351.

AGHION, P., HOWITT, P. (1998), "Endogenous Growth", MIT Press, Cambridge.

AITKEN, B. J., HARRISON, A. E. (1999), "Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela", American Economic Review 89, pp. 605-618.

CANEPA, A., STONEMAN, P. (2004), "Comparative International Diffusión: Patterns, Determinants and Policies", Economics of Innovation & New Technology 13 (3), pp. 279-298.

COE, D. T., HELPMAN, E. (1995), "International R&D Spillovers", European Economic Review 39 (5), pp. 859-887.

DUGUET, E., MACGARVIE, M. (2005), "How Well Do Patent Citations Measure Flows of Technology? Evidence from French Innovation Surveys", Economics of Innovation & New Technology 14 (5), pp. 375-393.

EATON, J., KORTUM, S. (1996), "Trade in Ideas: Patenting and Productivity in the OECD". Journal of International Economics 40, pp. 251-278.

EATON, J., KORTUM, S. (2001), "Trade in Capital Goods". European Economic Review 45(7), pp. 1195-1235.

FREEMAN, C. (1994), "The Economics of Technical Change", Cambridge Journal of Economics 18, pp. 463-514.

GAy, C., LE BAS, C., PATEL, P., TOUACH, K. (2005), "The Determinants of Patent Citations: an Empirical Analysis of French and British Patents in the US", Economics of Innovation & New Technology 14 (5), pp. 339-350.

GLOBERMAN, S., KOKKO, A., SJÖHOLM, F. (2000), "International Technology Diffusion: Evidence from Swedish Patent Data", Kyklos 53 (1), pp. 17-38.

GRILICHES, Z. (1990), "Patent Statistics as Economic Indicators: A Survey". Journal of Economic Literature 28 (4), pp. 1661-1707.

GROSSMAN, G. M., HELPMAN, E. (1991), "Trade, Knowledge Spillovers, and Growth", European Economic Review 35 (2-3), pp. 517-526.

HALL, B., JAFFE, A. B., TRAJTENBERG, M. (2001), "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools", NBER Working Paper 8498.

Jaffe, A., Trajtenberg, M., HENDERSON, R. (1993), "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", Quarterly Journal of Economics 79 (3), pp. 577-598.

JAFFE, A., TRAJTENBERG, M. (1996), "Flows of Knowledge from Universities and Federal Labs: Modelling the Flow of Patent Citations Over Time and Across Institutional and Geographic Boundaries", NBER Working Paper 5712.

JAFFE, A., TRAJTENBERG, M. (1999), "International Knowledge Flows: Evidence from Patent Citations", Economics of Innovation & New Technology 8, pp. 105-136.

KELLER, W. (2002), "Trade and the Transmission of Technology", Journal of economic Growth 7, pp. 5-24.

KELLER, W. (2004), "International Technology Diffusion", Journal of Economic Literature 42, pp. 752-782.

KELLER, W., YEAPLE, S. R. R. (2003), "Multinational Enterprises, International Trade, and Productivity Growth Firm-Level. Evidence from the United States", IMF Working Paper 03/248.

LAURSEN, K., MELICIANI, V. (2002), "The Relative Importance of International vis-à-vis National Technlogical Spillovers for Market Share Dynamics", Industrial and Corporate Change 11 (4), pp. 875-894.

LEE, Y. G., LEE, J. D., Song, Y. I., Lee, S. Y. (2007), "An In-depth Empirical Analysis of Patent Citation Counts Using Zero-inflated Count Data Model: the Case of KIST", Scientometrics 70 (1), pp. 27-39.

LUCAS, J. R. E. (1988), "On the Mechanics of Economic Development", Journal of Monetary Economics 22 (1), pp. 3-42.

MACGARVIE, M. (2005), "The Determinants of International Knowledge Diffusion as Measured by Patent Citations", Economics Letters 87 (1), pp. 121-126.

MACGARVIE, M. (2006), "Do Firms Learn from International Trade?", The Review of Economics and Statistics 88 (1), pp. 46-60.

MARKUSEN, J. (2002), "Multinational Firms and the Theory of International Trade", MIT Press, Cambridge.

Maurseth, P. B. (2005), "Lovely but Dangerous: the Impact of Patent Citations on Patent Renewal", Economics of Innovation & New Technology 14 (5), pp. 351-374.

PAVITT, K. (1988) "Uses and Abuses of Patent statistics", in Raan, A. F. (ed.) Handbook of Quantitative Studies of Science and Technology, Elsevier, Amsterdam, pp. 509-536.

ROMER, P. (1986), "Increasing Returns and Long-Run Growth", Journal of Political Economy 94, pp. 1002-1037.

ROMER, P. (1990), "Endogenous Technological Change", Journal of Political Economy 98 (5), pp. 71-102.

SARKAR, J. (1998), "Technological Diffusion: Alternative Theories and Historical Evidence", Journal of Economic Surveys 12 (2), pp. 131-176.

SAVVIDES, A., ZACHARIADIS, M. (2005), "International Technology Diffusion and the Growth of TFP in the Manufacturing Sector of Developing Economies", Review of Development Economics 9 (4), pp. 482-501.

SCHUMPETER, J. (1934), "The Theory of Economic Development", Harvard University Press, Cambridge.

SILVERBERG, G., (1990), "Adoption and Diffusion of Technology as a Collective Evolutionary Process", in Freeman, C., Soete, L. (eds.), New Explorations in the Economics of Technological Change, Pinter, London.

SOLOW, R. M. (1956), "A Contribution to the Theory of Economic Growth", Quarterly Journal of Economics 70, pp. 65-94.

XU, B., WANG, J. (1999), "Capital Goods Trade and R&D Spillovers in the OECD", Canadian Journal of Economics 32, pp. 1258-1274.

XU, B. (2000), "Multinational Enterprises, Technology Diffusion and Host Country Productivity Growth", Journal of Development Economics 62, pp. 477-493.

XU, B., CHIANG, E. P. (2005), "Trade, Patents and International Technology Diffusion", Journal of International Trade and Economic Development 14 (1), pp. 115-135.

ANEXO

RELACIÓN DE CATEGORÍAS Y SUBCATEGORÍAS TECNOLÓGICAS (NBER)			
CAT 1	Chemical	SUBCAT 11	Agriculture, Foods, Textiles
		SUBCAT 12	Coating
		SUBCAT 13	Gas
		SUBCAT 14	Organic Compounds
		SUBCAT 15	Resins
		SUBCAT 19	Miscellaneous-Chemical
CAT 2	Computers & Communications	SUBCAT 21	Communications
		SUBCAT 22	Computer Hardware & Software
		SUBCAT 23	Computer Peripherials
		SUBCAT 24	Information Storage
CAT 3	Drugs & Medical	SUBCAT 31	Drugs
		SUBCAT 32	Surgery & Medical Inst.
		SUBCAT 33	Biotechnology
		SUBCAT 39	Miscellaneous-Drugs&Med
CAT 4	Electrical & Electronics	SUBCAT 41	Electrical Devices
		SUBCAT 42	Electrical Lighting
		SUBCAT 43	Measuring & testing
		SUBCAT 44	Nuclear & X-rays
		SUBCAT 45	Power Systems
		SUBCAT 46	Semiconductor Devices
		SUBCAT 49	Miscellaneous-Elec
CAT 5	Mechanical	SUBCAT 51	Mat Proc & Handling
		SUBCAT 52	Metal Working
		SUBCAT 53	Motors & Engines + Parts
		SUBCAT 54	Optics
		SUBCAT 55	Transportation
		SUBCAT 59	Miscellaneous-Mechanical
CAT 6	Others	SUBCAT 61	Agricultura, Husbandry, Food
		SUBCAT 62	Amusement Devices
		SUBCAT 63	Apparel & Textile
		SUBCAT 64	Herat Working & Wells
		SUBCAT 65	Furniture, House Fixtures
		SUBCAT 66	Heating
		SUBCAT 67	Pipes & Joints
		SUBCAT 68	Receptacles
		SUBCAT 68	Miscellaneous-Others

Evaluating the Effect of Maternal Time on Child Development Using the Generalized Propensity Score

Pedro Carneiro
UCL, IFS and CEMMAP

Margarida Rodrigues
Faculdade de Economia, UNL
m.margarida.r@fe.unl.pt

Resumo
O tempo que as mães passam com as crianças não é considerado como um input nos trabalhos empíricos que estudam os determinantes do desenvolvimento cognitivo das crianças. Este trabalho fá-lo usando uma base de dados única com informação sobre o uso do tempo diário das crianças e implementando uma extensão da metodologia de avaliação de tratamentos, aplicada a tratamentos contínuos. Neste trabalho conclui-se que, independentemente da idade, as crianças que passam mais tempo com as mães tendem a apresentar melhores resultados nos testes cognitivos, mas este efeito é mais forte para as crianças mais novas. Quando a amostra é dividida de acordo com a raça das crianças conclui-se que as crianças negras mais novas tendem a apresentar piores resultados quando passam mais do que 5 horas com as mães durante um dia. Este efeito negativo não é extensível para as crianças brancas, que tendem a beneficiar do tempo que passam com as mães e beneficiam até mais tarde na sua vida.

Palavras-chave: Tempo maternal, Testes cognitivos e não cognitivos, avaliação de tratamentos.

Abstract
Mainly due to data unavailability, time with the mother is usually not included as an input in empirical papers trying to find the determinants of child's achievement. We attempt to do so in this paper by using unique data that collects a child's time-use diary, cognitive and non cognitive test scores. We implement the methodology developed recently on the program evaluation of continuous treatments. We find that more time with mothers leads both young and old children to perform better in cognitive tests, but the effect for the former group is more pronounced. Once we divide our sample according to race we find that young black children tend to perform worse if they spend more than 5 hours with the mother in a day. This negative effect is not present for white children, who tend to perform better if the time spent with mothers increases. Also white children benefit from this effect until later ages.
Keywords: Maternal time, Cognitive and non-cognitive outcomes, Treatment evaluation.

1 Introduction

There is evidence from psychology, neuropsychology and other sciences that the time the child spends with the primary caregiver (usually the mother) is essential for her cognitive, social and emotional development. Neuropsychological research on attachment theory finds that the interaction between the primary caregiver and infant stimulates brain development (Shore (1997)). In particular it stimulates the orbifrontal cortex maturation, which controls emotions and behavior, having a significant and permanent impact on the development of social and emotional skills: self esteem, self confidence, optimism, discipline, endurance, motivation and emotional security. All these are important components of human capital if we think of human capital as anything that affects productivity. In fact, economic literature acknowledges the importance of noncognitive skills: Heckman and Carneiro (2003) present evidence that noncognitive outcomes determine potential productivity and socioeconomic success later in life. Heckman et al. (2006) demonstrate that cognitive and non-cognitive skills are equally important in explaining a variety of aspects of social and economic life. Carneiro et al. (2003) and Cunha et al. (2006) find that cognitive and non-cognitive skills are strongly influenced by parental behavior and abilities, suggesting that economic success is already partly determined in early childhood.

Several papers in the economic literature attempt to determine the determinants of child's output from cognitive and non-cognitive test scores. There are not many papers that include the time mother and child spend together as an explanatory variable mainly due to data unavailability. The best some papers do is to relate mother employment (in intensive and extensive margin) to child outcomes, i.e. they assume that the mother's non-working time is entirely spent with the child (e.g. Bernal and Keane (2005), James-Burdumy (2005), Blau and Grossberg (1992), Ruhm (2004)). The above methodology, despite being very interesting, allows uniquely to assess the impact of working (or working hours) on test scores while it does not allow to make any considerations about the impact of the actual time they spend together. We can think of two cases in which the former analysis would lead to wrong conclusions: i) mothers that don't work or do work very few hours but spend very few of their free time with children; and ii) mothers that do work or work for longer hours but that spend almost their entire free time with their children. So, if one attempts to relate mother-child time with child outcomes it is important to have a direct measure of the actual time they spend together. The type of data needed to do such an analysis is very hard to find because we must have simultaneously child outcomes and an accurate measure of time. The usual time diary surveys do not have the former while other databases do not have the later. In this paper we use a unique data set from the 1997 PSID - Child Development Supplement made with the purpose of providing researchers with a comprehensive database of children and families that would enable to study the dynamic process of early human capital formation. Up to 2 children per PSID family aged 0-12 years during the calendar year of 1997 were selected for interview. The interview included a time use diary with 24-hour detailed accounting of time use for one randomly selected weekday and weekend, with the type, duration and location of activities and social context of activities (detailed information about whom participated in the activity and who else was there but not directly engaged). It was also collected aptitude and achievement test scores (reading and math and memory) and psychological, emotional and social well being (behavior problem index, positive behavior) as well as a bunch of family context variables. So this seems the perfect data set to assess the impact of mother-child time on child's test scores, as we have the two necessary variables and a bunch of variables to control for. We focus on three outcomes: letter word identification, applied problems and behavior problem index. Our sample has 1497 children, which we then divide into two six sub-samples according to their age and race.

In order to assess the causal effect of the time spent with mothers in the cognitive and non cognitive achievement we use the recent development in program evaluation made by Hirano and Imbens (2004). The latter extends the usual binary treatment case to a continuous treatment,

which suits particularly well our needs since the treatment variable we consider, time, is continuous. We are able then to use the generalized propensity scores and estimate dose response functions, i.e. the response function of each outcome to each level of the treatment variable.

The remainder of the paper is organized as follows. Next section describes the use of the generalized propensity score and methodology of estimating a dose response function to evaluate a continuous treatment. Section 3 presents the data used in this paper. The fourth section presents our empirical results and section 5 concludes.

2 The Generalized Propensity Score

In recent years the research in program evaluation has made comprehensive use of matching methods. The standard case considers a binary treatment. Rosenbaum and Rubin (1983) provided the key result that has made the matching such an attractive method: rather than conditioning on the full set of covariates, conditioning on the propensity score, i.e. on the probability of receiving the treatment given the covariates, is sufficient to balance treatment and comparison groups.

More recently, the literature has extended propensity score methods to the cases of multi-valued treatments (Imbens(2000) and Lechner (1999)) and continuous treatments (Hirano and Imbens (2004). The approach of the latter paper is particularly suitable for our purpose because it enables to estimate the entire dose response function of our continuous treatment- time. Therefore we follow closely Hirano and Imbens's (2004) methodology and implementation, that are summarized below.

2.1 Methodology

We have a random sample of units $i = 1, ..., N$ and for each we observe a set of potential outputs $Y_i(t)$ from a treatment t. In the usual binary case the treatment set is $\check{T} = \{0, 1\}$ whereas in the continuos case $\check{T}$ is an interval $[t_0, t_1]$. For each sample unit i we observe a vector of covariates X_i, the level of treatment actually received $T_i \in [t_0, t_1]$ and the outcome $Y_i(T_i)$. Our objective is to estimate the average dose response function $\mu(t) = E[Y(t)]$. In the remainder of the section we ignore subscript i.

The key assumption of Hirano and Imbens (2004) generalizes the unconfoundedness assumption for binary treatments to the continuos case. The weak unconfoundedness assumption is the following:

$$Y(t) \perp T \mid X \ , \forall t \in \check{T} \tag{1}$$

and it is named so because it only requires independence to hold for each level of treatment t rather than the joint independence of all potential outcomes.

Call $r(t, x) = f_{T|X}(t|x)$, i.e. $r(t, x)$ is the conditional density of the treatment given the covariates. The generalized propensity score (GPS) is defined as:

$$R = r(T, X). \tag{2}$$

The GPS has a balancing property similar to that of the standard propensity score, as within the strata with the same value of $r(t, x)$ the probability of $T = t$ does not depend on the value of X. In other words, GPS has the following property:

$$X \ \perp \ 1\{T = t\} \mid r(t, x). \tag{3}$$

Hirano and Imbens (2004) highlight that this property does not require unconfoundedness. However, when combined with unconfoundedness, it implies that assignment to treatment is un-

confounded given the GPS:

$$Y(t) \perp T \mid X \ , \forall t \in \check{T} \quad \Rightarrow \quad Y(t) \perp T \mid r(T,X) \ , \forall t \in \check{T}. \tag{4}$$

Given this result it is possible to use the GPS to remove the bias associated with differences in covariates in two steps:

1. Estimate the conditional expectation of the outcome as a function of two variables- treatment T and GPS R:
$$\beta(t,r) = E\left[Y \mid T = t, R = r\right] \tag{5}$$

2. Estimate the dose response function (DRF) at each particular level of the treatment. This is implemented by averaging the conditional expectation over the GPS at that particular level of treatment:
$$\mu(t) = E\left[\beta(t, r(t,X))\right] \tag{6}$$
Notice that we do not average over the GPS $R = r(T,X)$ but instead we average over the score evaluated at the treatment level of interest $r(t,X)$.

It should be stressed that the regression function $\beta(t,r)$ does not a causal interpretation but $\mu(t)$ corresponds to the value of the DRF for treatment level t, which compared to another treatment level t' does have a causal interpretation.

2.2 Implementation

In the practical implementation of the methodology outlined in the previous section we use a normal distribution for the treatment given the covariates:

$$T_i \mid X_i \ \tilde{} \ N(\beta_0 + \beta_1' X_i, \sigma^2). \tag{7}$$

1. In the first stage we estimate the parameters β_0, β_1 and σ^2 by OLS. Then the estimated GPS is:
$$\widehat{R}_i = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(T_i - \widehat{\beta}_0 - \widehat{\beta}_1' X_i)^2}{2\widehat{\sigma}^2}\right] \tag{8}$$

2. In the second stage we model the conditional expectation of Y_i given T_i and R_i as a flexible function of its two arguments. In the application used by Hirano and Imbens (2004) a quadratic approximation is used:
$$E[Y_i|T_i, R_i] = \alpha_0 + \alpha_1 T_i + \alpha_2 T_i^2 + \alpha_3 R_i + \alpha_4 R_i^2 + \alpha_5 T_i R_i \tag{9}$$
The parameters are estimated by OLS using the estimated GPS $\widehat{R}_i$.

3. In the third stage, and given the estimated parameters in the second stage, the average potential outcome at treatment level t is estimated:
$$\widehat{E[Y(t)]} = \frac{1}{N}\sum_{i=1}^{N}\left[\widehat{\alpha}_0 + \widehat{\alpha}_1.t + \widehat{\alpha}_2.t^2 + \widehat{\alpha}_3\widehat{r}(t,X_i) + \widehat{\alpha}_4\widehat{r}(t,X_i)^2 + \widehat{\alpha}_5.t.\widehat{r}(t,X_i)\right] \tag{10}$$
We should do this for every treatment level we are interested in order to obtain the entire dose-response function. It is convenient to use bootstrap methods to form standard errors and confidence intervals.

3 Data

In this paper we use the Child Development Supplement (CDS) of the Panel of Income Dynamics (PSID). The PSID is a nationally representative sample of U.S. families that collects data on income, work and consumption. PSID families have been interviewed since 1968 and sample members are followed as they split off into new households. In 1997, with the purpose of providing researchers with a comprehensive database of children and families that would enable to study the dynamic process of early human capital formation, it was developed the PSID Child Development Supplement. Up to 2 children per PSID family aged 0-12 years during the calendar year of 1997 were selected. Then these children were followed and re-interviewed in 2002/03 with ages 5-18 years. The interview included a time use diary with 24-hour detailed accounting of time use for one randomly selected weekday and weekend, with the type, duration and location of activities and social context of activities (detailed information about whom participated in the activity and who else was there but not directly engaged). It was also collected aptitude and achievement test scores (Reading and math and memory) and psychological, emotional and social well being (behavior problem index, positive behavior) as well as a bunch of family context variables. So this seems the perfect data set to assess the impact of mother-child time on child's test scores.

For each individual in our sample we compute the total time they spend with mothers in a week day [1]. This will be our treatment variable. This variable englobes the time that the mother is participating in the activity with the child or not, i.e. the mother might just be around while the child does an activity by herself. We are particular interested in three outcomes: Letter-Word Identification (LW), Applied Problems (AP) and Behavior Problem Index (BPI). So we have information both on cognitive and non-cognitive skills. Table 1 in the appendix describes all the variables we use in our analysis.

In 1997, 3563 children were interviewed. We will only consider the 2478 children that completed the time diary questionnaire (both in week and weekend days), that live with the mother and for whom their mother was reported to be the primary caregiver. Our final sample consists of 1497 children that have available at least one of our three outcome variables and for whom none of the several covariates we consider (child's, mother's and family's characteristics) is missing. In the forthcoming analysis we always divide children in two age groups: the first group is composed by younger children (from 3 to 6 years old) and is named *Young* and the second group is composed by older children (from 7 to 12 years old) and is named *Old*. We believe it is important to make this distinction because the time needs for one and other group seem to be very different and can therefore affect determinantely the impact of the treatment. Furthermore we distinguish between white and black children because the majority of their covariates is significantly different from each other. We see this by computing the difference in means and checking whether these are significantly different from zero. We conclude that, both for young and old children, the means of the following covariates are significantly different between races: birth weight, age breast feeding stopped, mother's age at birth, marital status at birth, mother's education, dummy for both parents living with the child, dummy of other adults living in house, dummy of living at own house, father's education, 5 year average of total income, mother's fixed effect and state's average weekly wage.

So, summarizing we have 6 different samples: young children with 561 observations, old children with 936 observations, young white children with 341 observations, young black children with 220 observations, old white children with 540 observations and old black children with 396 observations. We implement the methodology presented in the previous section separately for each of them. Table 2 presents the summary statistics for every variable we use and for each of the six samples.

On average mothers with young children spend more hours in a week day with children than

[1] For now we focus on a week day as these are the days in which a time constraint is more relevant. Later on we intend to use also the weekend time diary and see whether there is any time compensation during the weekend.

mothers with old children. In this table we present the standard outcomes but in all our analysis we normalize them to have mean 0 and standard deviation 1 within each sample. Young children preform worse than old children in LW, AP and better in BPI. Analyzing the columns corresponding to the division of children according to age and race (columns (3) to (12)) we conclude that white children preform much better in all outcomes considered. Also, white mothers spend mor than 1 hour per week day with their children with their children irrespectively of being young or old. This could lead us to think that indeed time with mothers determines the outcomes as we observe an association between higher treatments and higher outcomes. However, notice that there exists selection bias, i.e. we observe that the characteristics of black and white samples are very different and significant as explained above. Furthermore the black samples seem to be disadvantaged in comparison with their white counterparts in almost every characteristics we can observe: white children weighted more at birth, were breast fed longer, had an older mother at birth, have more educated mothers and fathers, are more likely to be part of a biparental family, have less siblings, are more likely to own their house, belong to richer families, and have abler mothers. So indeed white children seem to have a privileged background.

Table 2- Summary Statistics

	Young		Old		Young Black		Young White		Old Black		Old White	
Treatment	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Time	5.99	3.23	4.29	2.50	5.28	2.91	6.46	3.45	4.11	2.46	4.42	2.52
Outcomes (standard)												
LW	99.2	15.6	106.7	19.1	95.4	15.7	101.6	15.1	99.3	16.5	112.5	18.9
AP	101.9	18.2	109.2	16.8	94.1	18.6	107.1	15.9	101.1	14.1	115.5	16.0
BPI	7.49	4.83	7.96	5.95	7.20	4.68	7.67	4.93	8.04	6.36	7.91	5.63
Covariates												
Ageatpcg	60.6	13.3	121.8	22.1	61.2	12.9	60.2	13.5	123.1	22.5	120.8	21.7
RaceWhite	0.61	0.49	0.58	0.49	0	0	1	0	0	0	1	0
RaceBlack	0.39	0.49	0.42	0.49	1	0	0	0	1	0	0	0
Chgender	0.56	0.49	0.49	0.50	0.56	0.49	0.55	0.49	0.53	0.49	0.47	0.49
Birthorder	1.89	1.12	2.08	1.12	1.98	1.31	1.83	0.97	2.29	1.27	1.93	0.96
Birthweight	6.88	1.39	6.94	1.38	6.51	1.47	7.13	1.28	6.59	1.44	7.19	1.27
Ageatbreast	3.23	5.97	3.10	5.31	1.30	3.79	4.47	6.74	1.11	3.25	4.56	6.00
AgeBirthM	27.2	5.95	27.3	5.18	25.1	5.96	28.5	5.56	26.3	5.32	28.1	4.94
MaritalBirth	0.68	0.46	0.73	0.44	0.39	0.49	0.87	0.33	0.45	0.49	0.93	0.25
EducationM	13.2	2.02	13.1	2.14	12.5	1.77	13.6	2.04	12.4	1.87	13.5	2.19
ParentLive	0.66	0.47	0.66	0.47	0.38	0.48	0.85	0.36	0.42	0.49	0.84	0.37
ChildrenFU	2.22	1.03	2.42	1.02	2.26	1.21	2.19	0.90	2.52	1.21	2.35	0.83
AgeYoungest	3.41	1.58	7.27	3.14	3.49	1.52	3.354	1.62	7.19	3.41	7.32	2.93
OtherAdults	0.77	0.49	0.86	0.59	0.53	0.58	0.93	0.34	0.71	0.73	0.97	0.44
OwnHouse	0.56	0.49	0.66	0.47	0.31	0.46	0.72	0.45	0.47	0.49	0.80	0.4n
EducationF	9.06	6.68	8.86	6.67	4.93	6.45	11.7	5.34	5.21	6.30	11.5	5.57
Avg(Total Income)	10.6	0.97	10.8	0.91	10.1	0.96	11.0	0.76	10.3	0.93	11.2	0.69
Mother's FE	-0.77	2.10	0.47	1.99	-1.36	2.02	-0.39	2.06	0.15	1.96	0.71	1.98
AvgWeeklyWage	561	76.9	567	86.3	543	64.7	573	81.8	554	86.9	576	84.5
Observations	516		936		220		341		396		540	

4 Results

4.1 GPS, Covariate Balance and Common Support

We first estimate the conditional distribution of the treatment by estimating equation (7) by OLS. Table 3, in the appendix, presents the results of this estimation for all the samples. We then compute the GPS according to equation (8).

As in the binary treatment case it is important to impose common support. This guarantees that for the same propensity score there are observations from the different treatment groups and therefore they can compared. Given the GPS estimate we divide the sample in 3 groups according to the 33th (p33) and 66th (p66) percentile of the treatment variable T_i. So:

$$\begin{aligned} i &\in G1 \text{ if } T_i \in [0, p33] \\ i &\in G2 \text{ if } T_i \in (p33, p66] \\ i &\in G3 \text{ if } T_i \in (p66, p33] \end{aligned}$$

Then we evaluate the GPS for each observation in each group, which yields the variables:

$$\begin{aligned} \text{if } i &\in G1 \Rightarrow R_i^1 = \widehat{R}_i; \text{ if } i \notin G1 \Rightarrow R_i^1 = NaN \\ \text{if } i &\in G2 \Rightarrow R_i^2 = \widehat{R}_i; \text{ if } i \notin G2 \Rightarrow R_i^2 = NaN \\ \text{if } i &\in G3 \Rightarrow R_i^3 = \widehat{R}_i; \text{ if } i \notin G3 \Rightarrow R_i^3 = NaN. \end{aligned}$$

We assume that the control group for each group is formed by the other two groups combined. For instance the control group of G1 is formed by individuals in the other two groups combined (G2 and G3). Then we eliminate observations in each group that have a GPS higher than the maximum or lower than the minimum GPS observed in the control group, i.e.:

$$\begin{aligned} \text{eliminate } i &\in G1 \text{ iff } \left[R_i^1 < \min\left\{\min(R^2), \min(R^3)\right\} \text{ or } R_i^1 > \max\left\{\max(R^2), \max(R^3)\right\}\right] \\ \text{eliminate } j &\in G2 \text{ iff } \left[R_j^2 < \min\left\{\min(R^1), \min(R^3)\right\} \text{ or } R_j^2 > \max\left\{\max(R^1), \max(R^3)\right\}\right] \\ \text{eliminate } k &\in G3 \text{ iff } \left[R_k^3 < \min\left\{\min(R^1), \min(R^2)\right\} \text{ or } R_k^3 > \max\left\{\max(R^1), \max(R^2)\right\}\right]. \end{aligned}$$

Part A of table 4 present the percentage of observations eliminated for each sample as well as the maximum percentage of observations eliminated in our 1000 bootstrap samples.

It is also important to evaluate how well the adjustment for the GPS works in balancing the covariates, i.e. if the specification of equation (7) is adequate. Whereas in the binary case the typical approach is to compare the covariate means for the treated and control groups, before and after matching, testing for covariate balance is more difficult in the continuos case.

Here we implement exactly what is proposed by Hirano and Imbens (2004). We use the 3 groups defined in point 2 above: G1, G2 and G3.

1. **Unadjusted**: For each of the covariates we investigate the balance by testing whether the mean in one of the three treatment groups is different from the mean in the other two treatment groups combined. Take for instance variable X and suppose we want to test the balance of this covariate in G1 vs the other two groups. We compute the mean and variance of X for observations belonging to G1, $\overline{X}_1$ and $Var(X_1)$. Next we compute the mean and variance of X for the observations belonging to G2 *and* G3, $.\overline{X_{2\&3}}$ and $Var(X_{2\&3})$. Then we compute the T-test statistic corresponding to null that these means are indeed the same and compare them to the 10% and 5% critical values. We reject the null, i.e. the covariates are unbalanced, if the t-statistic is higher than the critical value.

2. **Adjusted for the GPS**: As an example consider we want to test the balance of the covariate age in the sample of young children[2]. To implement the GPS-adjusted version of this statistic we discretize both the level of treatment and the GPS. The steps are the following:

   - Given the groups formed in point 2, we find the median treatment in each group and evaluate the GPS at this median treatment. For instance the median treatment in group 1 is 3.17 and we get $r(3.17, X_i)$. Basically what we will test is whether:

$$X_i \perp 1\{0 \leq T_i < 4.17\} \mid r(3.17, X_i).$$

   - Next we block on the score $r(3.17, X_i)$. We make five blocks defined by the quintiles of $r(3.17, X_i)$ in the group with $1\{0 \leq T_i < 4.17\}$: $[0, 0.0066]$, $(0.0066, 0.0380]$, $(0.0380, 0.0819]$, $(0.0819, 0.1912]$ and $(0.1912, 0.3242]$.
   - In each of these five blocks we compute the difference in means of age with respect to individuals that have a GPS such that they belong to that block, but have a treatment level different from the one being evaluated. I.e., if we are testing for group 1, we will compute the following difference in means:

$$\begin{aligned} mean(age_i \mid i \in & \text{ G1 \& } i \in \text{block 1 of G1}) - \\ -mean(age_j \mid j \notin & \text{ G1 \& } j \in \text{block 1 of G1}) \end{aligned}$$

   - For example:
     - the first of these five blocks, with $r(3.17, X_i) \in [0, 0.0066]$, has a total of 174 observations: 42 with $T \in [0, 4.17]$ and 132 with $T \notin [0, 4.17]$. Testing for equality of the mean of age in the first versus the other treatment groups in this GPS group give a mean difference of -0.057 and a standard deviation of 1.72.
     - the second of these five blocks, with $r(3.17, X_i) \in (0.0066, 0.0380]$, has a total of 147 observations: 42 with $T \in [0, 4.17]$ and 105 with $T \notin [0, 4.17]$. Testing for equality of the mean of age in the first versus the other treatment groups in this GPS group give a mean difference of 2.86 and a standard deviation of 2.32.
     - the third of these five blocks, with $r(3.17, X_i) \in (0.0380, 0.0819]$, has a total of 84 observations: 41 with $T \in [0, 4.17]$ and 43 with $T \notin [0, 4.17]$. Testing for equality of the mean of age in the first versus the other treatment groups in this GPS group give a mean difference of 0.95 and a standard deviation of 2.48.
     - the fourth of these five blocks, with $r(3.17, X_i) \in (0., 0819.1912]$, has a total of 89 observations: 42 with $T \in [0, 4.17]$ and 47 with $T \notin [0, 4.17]$. Testing for equality of the mean of age in the first versus the other treatment groups in this GPS group give a mean difference of -2.69 and a standard deviation of 2.13.
     - the fifth of these five blocks, with $r(3.3, X_i) \in (0.1912, 0.3242]$, has a total of 58 observations: 42 with $T \in [0, 4.17]$ and 16 with $T \notin [0, 4.17]$. Testing for equality of the mean of age in the first versus the other treatment groups in this GPS group give a mean difference of -2.47 and a standard deviation of 2.15.
   - Combining these five difference in means, weighted by the total number of observations in each block leads to a mean difference of 0.70 and a standard error of 3.91 and thus to a t-statistic of 0.18. So in this example we do not reject the null of equality of means between individuals with the same GPS but in different treatment groups. This means that, when adjusted for the GPS, the covariate age is balanced between groups.

[2]For this sample the 33^{th} and 66^{th} percentiles of treatment variable are 4.17 and 6.75 hours per week day, respectively.

- This procedure needs to be repeated for each treatment group and for each covariate, therefore we will have 57 t-statistics for the young and old samples and 51 t-statistics for the samples in which we divide both for age and race.

Part B of table 4 shows the number of t-statistics explained before that are higher than 1.645 or 1.96.

Table 4- Results on common support and covariate balance

	Part A		Part B			
			t.st.$\geq$ 1.645		t.st.$\geq$ 1.96	
	(1)	(2)	Unadj	Adj.	Unadj.	Adj.
Young	0.016	0.060	20	0	19	0
Old	0.038	0.081	11	0	2	0
Young Black	0.045	0.100	18	0	14	0
Young White	0.009	0.102	9	0	6	0
Old Black	0.035	0.119	8	0	3	0
Old White	0.046	0.144	8	0	6	0

If adjustment for the GPS properly balances the covariates, we expect that the differences in means to not be statistically different from zero. The values in the table show that the adjustment for the GPS improves the balance as, before the adjustment, and depending on the sample considered, there is an important number of covariates that present t-statictis higher than 1.645 which indicates a clear unbalanced covariates distribution. After the adjustment, the number of t-statistics higher than 1.645 or 1.96 is zero for all samples meaning that adjusting for the GPS is successful in balancing the covariates.

4.2 Estimating and Plotting the Dose Response Functions

The final step is to estimate the GPS-adjusted dose response function. Table 5 contains the estimation results for the dose response function, i.e. it presents the estimated coefficients of equation (9). Again notice that these coefficients do not have any interpretation but we use these to estimate equation (10). For young children we evaluate this equation for levels of treatment between 0 and 12 hours per week day and for old children we evaluate that equation for levels of treatment between 0 and 8 hours per week day.

4.2.1 Age Divided Samples

Figures 1 and 2 present the dose response functions of LW, AP and BPI for young and old children, respectively. Notice that these figures show the results of equation (10), i.e. the expected average outcome for each level of treatment we have considered. For young children, the LW and AP d.r.f. are clearly increasing in the treatment variable, meaning that spending more hours per week day with the mothers lead young children to perform better in these tests. For old children, it seems that the LW and AP d.r.f. slightly increase in the treatment variable but much less pronoucedly than those of young children. Furthermore, for higher levels of treatment, the d.r.f. seems to decrease. The BPI d.r.f. is more flat both for young and old children, suggesting that spending more time with the mother does not have an impact on the BPI score.

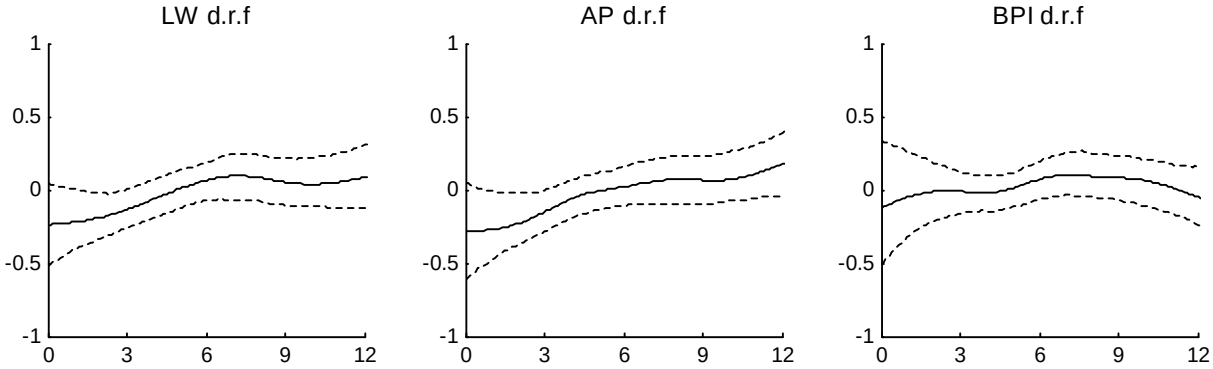


Figure 1- Dose response functions for young children

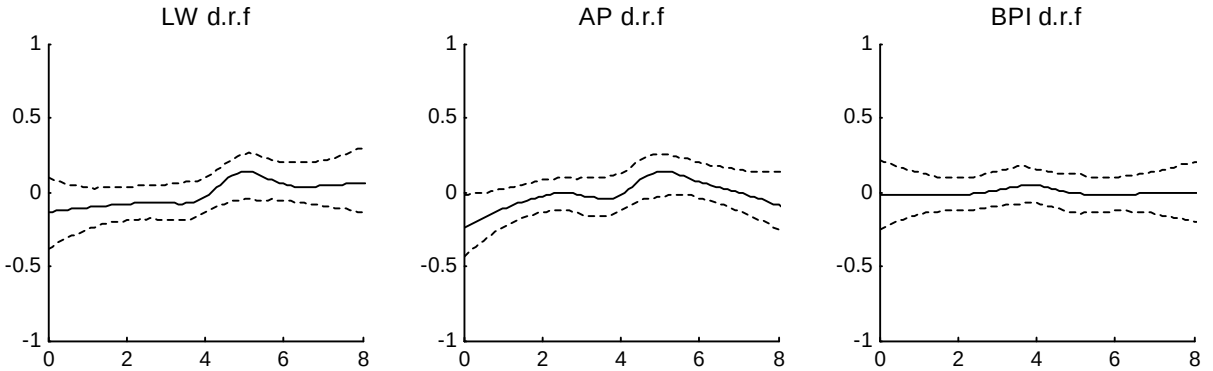


Figure 2- Dose response functions for old children

Notice that these figures allow us to compare the estimated d.r.f. for the different levels of treatment but do not give any information on whether the differences are significant. For instance, for young children, we can tell that spending 12 hours per day with the mother is certainly better for LW score than spending any time at all. However we can not tell whether it is significantly better. To give some insight about the significance of these differences we evaluate the d.r.f. in 5 treatment levels, compute the difference between them and test whether these differences are indeed significant. This information is presented in table 6.

Table 6 - Dose response differences

	Young children					Old children				
	T	0	3	6	9	T	0	2	4	6
LW	3	0.10	-	-	-	2	0.05	-	-	-
	6	0.30*	0.20***	-	-	4	0.10	0.05	-	-
	9	0.29	0.19**	-0.01	-	6	0.18	0.13*	0.08	-
	12	0.32*	0.22	0.02	0.03	8	0.19	0.14	0.09	0.01
AP	3	0.13	-	-	-	2	0.21**	-	-	-
	6	0.30	0.17**	-	-	4	0.22*	0.01	-	-
	9	0.35*	0.22**	0.05	-	6	0.30**	0.09	0.08	-
	12	0.46***	0.33***	0.16	0.11	8	0.15	-0.06	-0.07	-0.16*
BPI	3	0.11	-	-	-	2	0.01	-	-	-
	6	0.19	0.08	-	-	4	0.07	0.06	-	-
	9	0.21	0.09	0.01	-	6	0.01	0.00	-0.06	-
	12	0.08	-0.04	-0.12	-0.13	8	0.01	0.01	-0.05	0.00

For young children we evaluate treatments 0, 3, 6, 9 and 12 and for old children treatments 0, 2, 4, 6 and 8. In each cell of the table we present the difference between the d.r.f. of the treatment

in the row and in the column. For instance in the first cell (top left cell) we have 0.10 meaning that we expect to have a difference of 0.10 in the LW scores of a child that spends 3 hours with the mother and a child that spends no time at all with the mother. However this difference is not significantly different from zero.

Almost all the values for LW and AP are positive for young children, which indeed confirms that the d.r.f. is increasing in the treatment variable. Several values are indeed significant and it is worthwhile to mention the magnitude of these differences as they reach almost half a standard deviation. Notice that the figures are much smaller for old children than those of young children, in particular for LW. For AP it is significantly better to spend more time with the mother up until the 6 hours per day. In both samples the BPI values are very small and none is significant, which again suggests that the treatment does not have an impact on children behavior.

4.2.2 Age and Race Divided Samples

Figures 3 and 4 present the d.r.f. for young black and white children, respectively.

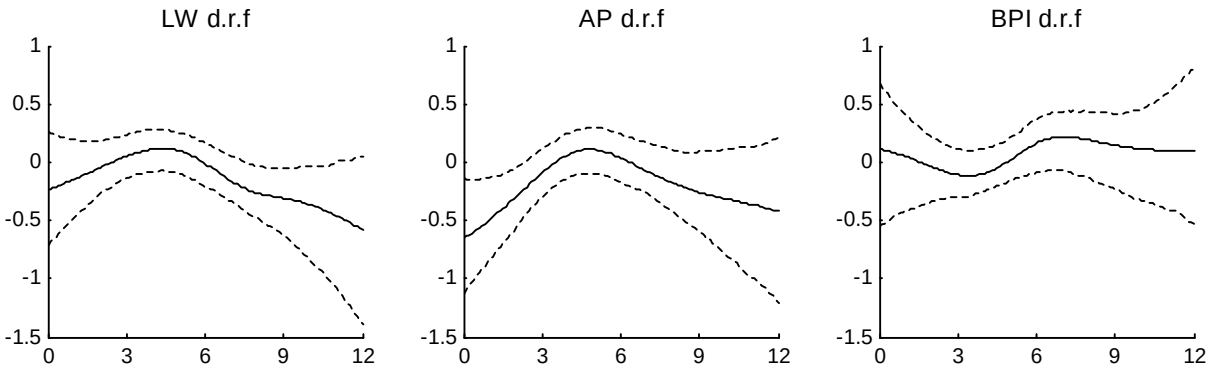


Figure 3- Dose response functions for young black children

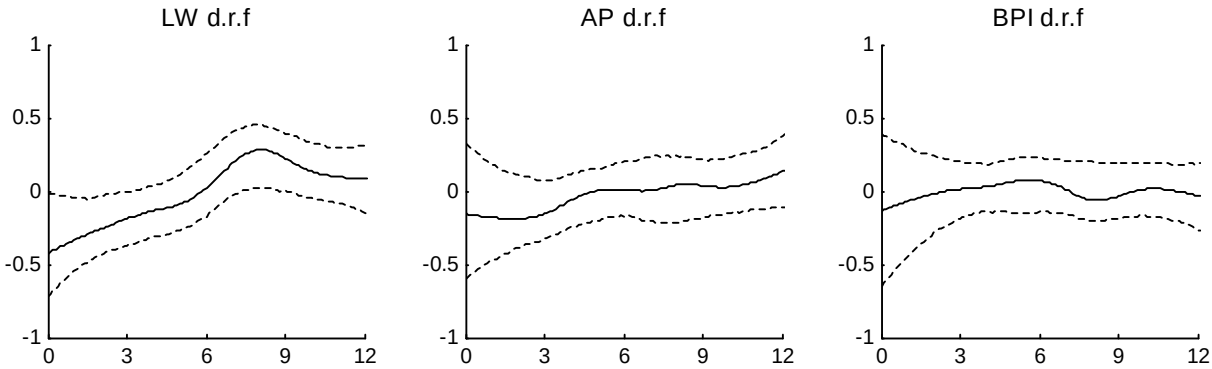


Figure 4- Dose response functions for young white children

Interestingly the d.r.f. present very differently patterns when we take each race separately. For young white children the LW and AP d.r.f. seem to be increasing in the treatment. Even though the LW d.r.f. seems to decrease for higher levels of treatment, overall there is no doubt that it presents an upward trend. This is confirmed by table 7: almost all the values for the white sample are positive and are particularly significant for LW. The black sample present very different d.r.f.. The LW and AP d.r.f. increase for smaller levels of treatment but then present a sharp decrease for high levels of treatment, presenting a peak around the 5 hours per week day. Notice that we expect black children to perform better in LW, and almost the same in AP, if they spend no time with their mothers or if they spend 12 hours per day with them. These findings are supported by the values in table 7: all the significant values in LW are negative and for AP the peak seems

indeed significant. The BPI d.r.f. is rather flat for young white children and for black children there is no clear pattern.

Table 7 - Dose response differences for young children

		Black Young				Young White			
	T	0	3	6	9	0	3	6	9
LW	3	0.29	-	-	-	0.23	-	-	-
	6	0.23	-0.06	-	-	0.44*	0.21*	-	-
	9	-0.07	-0.36**	-0.30*	-	0.64***	0.41***	0.20	-
	12	-0.34	-0.66*	-0.57	-0.27	0.51***	0.27*	0.07	-0.14
AP	3	0.56**	-	-	-	-0.00	-	-	-
	6	0.69***	0.12	-	-	0.16	0.06	-	-
	9	0.39	-0.17	-0.29*	-	0.19	0.19	0.03	-
	12	0.23	-0.33	-0.45	-0.16	0.29	0.29*	0.12	0.09
BPI	3	-0.22	-	-	-	0.15	-	-	-
	6	0.05	0.27*	-	-	0.21	0.06	-	-
	9	0.03	0.25	-0.02	-	0.09	-0.05	-0.11	-
	12	-0.01	0.21	-0.06	-0.04	0.10	-0.05	-0.10	0.00

Figures 5 and 6 present the d.r.f. for black and white old children, respectively.

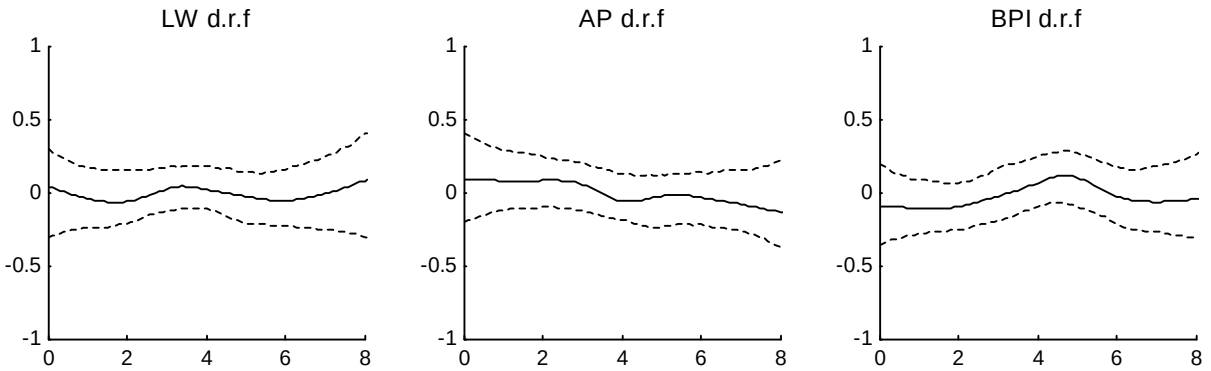


Figure 5- Dose response functions for old black children

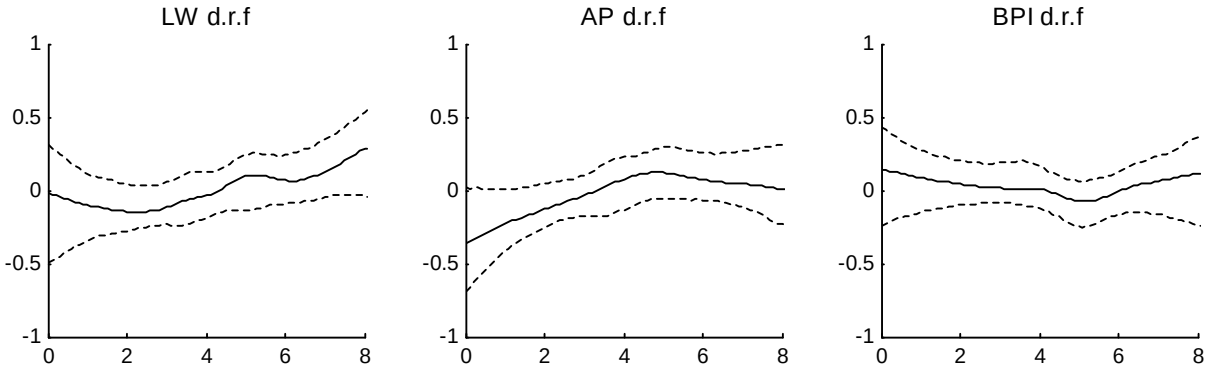


Figure 6- Dose response functions for old white children

For black children the d.r.f. seem flat suggesting that increasing spending more time with the mother may not have a significant effect on child's performance in LW, AP and BPI. Indeed, we can see in table 8 that all the values are insignificant. For white children the LW and AP d.r.f.

are overall positively inclined, even though for LW there is a slight decrease for low levels of the treatment and for AP there is a slight decrease for high levels of treatment. However in table 8 all the significant values for the white sample are positive suggesting that spending more time with the mother increases significantly the performance in LW and AP tests. There are no significant values for the white sample when the BPI d.r.f. is considered.

Table 8 - Dose response differences for old children

		Old Black				Old White			
	T	0	2	4	6	0	2	4	6
LW	2	-0.10	-	-	-	-0.12	-	-	-
	4	-0.02	0.08	-	-	-0.02	0.11	-	-
	6	-0.09	0.01	0.08	-	0.09	0.21*	0.10	-
	8	0.04	0.14	0.06	0.14	0.30	0.43***	0.32*	0.21
AP	2	-0.00	-	-	-	0.23	-	-	-
	4	-0.14	-0.14	-	-	0.43**	0.20*	-	-
	6	-0.12	-0.12	0.02	-	0.43*	0.20**	-0.00	-
	8	-0.22	-0.22	-0.08	-0.09	0.37*	0.14	-0.07	-0.06
BPI	2	-0.00	-	-	-	-0.10	-	-	-
	4	0.16	0.16	-	-	-0.14	-0.04	-	-
	6	0.07	0.07	-0.09	-	-0.15	-0.05	-0.01	-
	8	0.05	0.05	-0.11	-0.02	-0.03	0.07	0.11	0.12

4.3 Sensitivity Analysis

We perform a sensitivity analysis of our estimation by considering higher orders of approximation when estimating (9) and (10). In particular we use a cubic approximation and a 4^{th} degree approximation. Equation (10) for these cases is respectively:

$$\widehat{E[Y(t)]} = \frac{1}{N}\sum_{i=1}^{N}\left[\begin{array}{c}\widehat{\alpha}_0+\widehat{\alpha}_1.t+\widehat{\alpha}_2.t^2+\widehat{\alpha}_3\widehat{r}(t,X_i)+\widehat{\alpha}_4\widehat{r}(t,X_i)^2+\widehat{\alpha}_5.t.\widehat{r}(t,X_i)+\\ +\widehat{\alpha}_6t^3+\widehat{\alpha}_7\widehat{r}(t,X_i)^3+\widehat{\alpha}_8t^2\widehat{r}(t,X_i)+\widehat{\alpha}_9\widehat{r}(t,X_i)^2.t\end{array}\right]$$

$$\widehat{E[Y(t)]} = \frac{1}{N}\sum_{i=1}^{N}\left[\begin{array}{c}\widehat{\alpha}_0+\widehat{\alpha}_1.t+\widehat{\alpha}_2.t^2+\widehat{\alpha}_3\widehat{r}(t,X_i)+\widehat{\alpha}_4\widehat{r}(t,X_i)^2+\widehat{\alpha}_5.t.\widehat{r}(t,X_i)+\\ +\widehat{\alpha}_6t^3+\widehat{\alpha}_7\widehat{r}(t,X_i)^3+\widehat{\alpha}_8t^2\widehat{r}(t,X_i)+\widehat{\alpha}_9\widehat{r}(t,X_i)^2.t+\alpha_{10}t^4\\ +\alpha_{11}\widehat{r}(t,X_i)^4+\alpha_{12}t^3\widehat{r}(t,X_i)+\alpha_{13}\widehat{r}(t,X_i)^3t+\alpha_{14}t^2\widehat{r}(t,X_i)^2\end{array}\right]$$

In figures 7 to 12 (in the appendix) we plot the d.r.f. for the three approximations considered for each of the six samples. Even though there are some differences in detail, the six figures show that in general the d.r.f present the same trends and shapes under the different specifications.

4.4 Falsification exercise

In the appendix.

5 Conclusion

In this paper we use unique data from the PSID that collects a diary on children's time use and test scores. This data allows to measure the time that children spend with mothers during a week day and to assess the importance that this time has on child's cognitive and non-cognitive test scores. We use the treatment evaluation methodology applied for continuous treatments, as our

treatment variable, time, is continuous. This methodology eliminates the selection bias problem and ensure that we find a causal effect between time with mothers and test scores.

We conclude that more time with mothers leads both young and old children to perform better in cognitive tests, and that this effect is more pronounced for younger children. Once we divide our sample according to race and age we find that young black children tend to perform worse in those tests if they spend more than 5 hours in a day with mothers. This negative effect is not present for young white children who perform significantly better as time with mothers increases. For old black children we find no relation between time with mothers and test scores, while for old white children the positive effect of time with mothers still persists. So we conclude that white children always benefit from time spent with mothers in a week day and benefit until later ages. For all the age and race samples we consider we find no effect between time spent with mothers and non-cognitive test scores.

References

[1] Bernal, R. and Keane, M.

[2] Blau, F. and Grossberg, A. (1992), Maternal Labor Supply and Children's Cognitive Development, *The Review of Economic and Statistics*, vol. 74, n^{o}3, pp. 474-481

[3] Carneiro, P. K. Hansen and J. Heckman (2003), Estimating Distributions of Counterfactuals with an Application to the Returns to Schooling and Measurement of the Effect of Uncertainty on Schooling Choice, *International Economic Review*, vol. 44(2), pp. 361-422

[4] Cunha, F., J. Heckman, L. Lochner and D. Masterov (2006), Interpreting the Evidence on Life Cycle Skill Formation", in *Handbook of the Economics of Education*, ed. by E. Hanushek, and F. Welch. Amsterdam: North-Holland

[5] Heckman, J. and P. Carneiro (2003), Human capital policy, IZA discussion paper 821

[6] Heckman, J., J. Stixrud and S. Urzua (2006), The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior, *Journal of Labor Economics*, vol.24(3), pp. 411-482

[7] Hirano, K. and Imbens, G. (2994), The Propensity Score with Continuous Treatments, in A. Gelman and X. Meng (eds) *Applied Bayesian Modeling and Causal Inference from Incomplete Data Perspectives*

[8] Imbens, G. (2000), The Role of the Propensity Score in Estimating Dose-Response Functions", *Biometrika*, vol.87, pp. 706-710

[9] James-Burdumy, S. (2005), The Effect of Maternal Labor Force Participation on Child Development, *Journal of Labor Economics*, vol.23, n^{o}1

[10] Lechner, M. (1999), Identification and estimation of causal effects of multiple treatments under the conditional independence assumption, IZA, discussion paper n^{o}91

[11] Ruhm, C. (2994), Parental Employment and Child Cognitive Development, *The Journal of Human Resources*, vol. 39, n^{o}1, pp. 155-192

[12] Rosenbaum, P. and Rubin, D. (1983), The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika*, vol. 70, pp- 41-55

[13] Shore, R. (1997), "Rethinking the Brain: New Insights Into Early Development", New York: Families and Work Institute

A Tables

Table 1- Variable names and definitions

Variable Name	Definition
Outcomes	
LW	Letter- Word Identification
AP	Applied Problems
BPI	Behavior Problem Index
Treatment	
Time	Hours child spend with mother in a week day
Covariates	
Ageatpcg	Child's age at caregiver interview (months)
RaceWhite	Race - White (=1 if yes)
RaceBlack	Race - Black (=1 if yes)
Chgender	Child's gender (=1 if male)
Birthorder	Birth order from mother
Birthweight	Birth weight
Ageatbreast	Age breastfeeding stopped (months)
AgeBirthM	Mother's age at birth
MaritalBirth	Mother's marital status at birth (=1 if married)
EducationM	Mother's years of education
ParentLive	Both parents living (=1 if yes)
ChildrenFU	Number children in household
AgeYoungest	Age youngest child
OtherAdults	Number of other adults (excluding mother)
OwnHouse	Family own house (=1 if yes)
EducationF	Spouse years of education
Avg(Total Income)	5 year average of total income in previous year (log, real terms)
Mother's FE[1]	Mother's fixed effect
AvgWeeklyWage	State's average weekly wage

[1]- We compute the Mother's FE as follows. We have information on the wage the mother receives every 2 yeras from 1990 until 2003, so approximately 7 years before after the CDS data collection point. Furthermore we can compute the age of the mother at each of these data points. We build a panel with the wage, the age and the age squared and run a fixed effect model ($xtreg\ wage\ age\ age^2,\ fe$).

Table 3- Coefficients of the regression (7)

	Young		Old		Young Black		Young White		Old Black		Old White	
	β	S.D.	β	S.D.	β	S.D.	β	S.D.	β	S.D.	β	S.D.
Ageatpcg	−0.05	0.01***	−0.00	0.00	−0.06	0.02***	−0.04	0.02**	−0.01	0.01	−0.00	0.01
RaceWhite	8.60	2.91***	5.09	1.89***	–	–	–	–	–	–	–	–
RaceBlack	7.62	2.82***	5.01	1.85***	–	–	–	–	–	–	–	–
Chgender	−0.07	0.26	−0.14	0.16	−0.18	0.38	0.03	0.37	−0.19	0.25	−0.14	0.22
Birthorder	−0.18	0.26	−0.13	0.12	−0.20	0.35	−0.04	0.39	−0.20	0.16	−0.01	0.18
Birthweight	−0.01	0.09	−0.04	0.06	0.05	0.13	−0.02	0.14	−0.01	0.09	−0.03	0.09
Ageatbreast	0.05	0.02***	0.02	0.02	0.05	0.05	0.06	0.03**	0.03	0.04	0.01	0.02
AgeBirthM	0.07	0.04	0.05	0.03**	0.01	0.06	0.09	0.06	0.00	0.04	0.10	0.04***
MaritalBirth	0.11	0.42	0.21	0.25	0.75	0.57	0.02	0.64	0.49	0.31	−0.13	0.46
EducationM	0.02	0.09	0.04	0.05	0.10	0.13	−0.01	0.13	−0.01	0.08	0.06	0.07
ParentLive	−0.34	1.19	−0.24	0.70	−0.44	2.15	0.48	1.57	−0.71	1.34	0.69	0.92
ChildrenFU	−0.03	0.27	−0.08	0.12	−0.01	0.37	−0.20	0.39	−0.10	0.15	−0.11	0.19
AgeYoungest	0.02	0.13	0.00	0.04	0.02	0.19	−0.06	0.19	0.00	0.05	−0.01	0.06
OtherAdults	−0.23	0.43	0.17	0.20	−0.08	0.51	0.18	0.73	0.59	0.25***	−0.46	0.33
OwnHouse	0.29	0.35	−0.14	0.22	0.30	0.50	0.25	0.48	0.01	0.31	−0.26	0.31
EducationF	0.06	0.08	0.04	0.05	0.01	0.16	0.04	0.11	0.04	0.10	0.02	0.06
Avg(Total Income)	−0.33	0.23	−0.29	0.15*	−0.77	0.33**	−0.07	0.34	−0.46	0.24*	−0.19	0.21
Mother's FE	−0.43	0.12***	−0.19	0.07***	−0.35	0.20*	−0.43	0.15***	−0.11	0.12	−0.24	0.08***
AvgWeeklyWage	0.00	0.00	0.00	0.00***	0.00	0.00*	0.00	0.00	0.00	0.00***	0.00	0.00
Intercept	0.00	0.00	0.00	0.00	11.6	4.55***	5.53	3.98	7.89	3.17***	3.24	2.52
Observations	516		936		220		341		396		540	

Table 5a- Estimated dose response function

	Young			Old		
Outcome	LW	AP	BPI	LW	AP	BPI
Intercept	−0.24	−0.28	−0.12	−0.13	−0.23	−0.02
(*s.e.*)	0.18	0.18	0.17	0.12	0.12**	0.11
Time	0.01	−0.00	0.09	0.03	0.13	0.00
(*s.e.*)	0.07	0.07	0.07	0.07	0.07**	0.07
$Time^2$	0.00	0.00	−0.01	−0.00	−0.01	−0.00
(*s.e.*)	0.01	0.01	0.01	0.01	0.01	0.01
GPS	1.57	3.87	−3.75	−0.98	−1.55	0.58
(*s.e.*)	2.02	2.03	1.95	0.95	0.95	0.94
GPS^2	−4.83	−10.6	8.38	−0.77	−0.06	−0.07
(*s.e.*)	5.49	5.51*	5.33	0.81	0.82	0.80
$Time * GPS$	0.08	−0.03	0.19	0.36	0.35	−0.11
(*s.e.*)	0.19	0.19*	0.18	0.18**	0.18*	0.18
N	491	487	529	827	825	879

Table 5b- Estimated dose response function

	Young Black			Old Black		
Outcome	LW	AP	BPI	LW	AP	BPI
Intercept	−0.25	−0.68	0.14	0.04	0.08	−0.09
(*s.e.*)	0.13	0.33**	0.31	0.16	0.16	0.15
Time	0.08	0.11	−0.01	−0.09	0.00	−0.02
(*s.e.*)	0.14	0.14	0.14	0.01	0.01	0.01
$Time^2$	−0.01	−0.01	0.00	0.01	−0.00	0.00
(*s.e.*)	0.01	0.01	0.01	0.01	0.01	0.01
GPS	0.66	5.66	−5.60	1.82	0.66	0.66
(*s.e.*)	3.52	3.55	3.50	1.47	1.48	1.45
GPS^2	9.97	−6.23	6.74	−1.69	−1.39	−2.18
(*s.e.*)	10.2	10.2	10.1	1.93	1.94	1.91
$Time * GPS$	−0.44	−0.41	0.73	−0.15	−0.04	0.22
(*s.e.*)	0.34	0.34	0.33**	0.25	0.25	0.25
N	188	188	200	363	361	370

Table 5c- Estimated dose response function (White)

	Young White			Old White		
Outcome	LW	AP	BPI	LW	AP	BPI
Intercept	−0.42	−0.15	−0.13	−0.02	−0.36	0.14
(*s.e.*)	0.23*	0.23	0.22	0.19	0.19*	0.18
Time	0.09	−0.03	0.07	−0.09	0.14	0.14
(*s.e.*)	0.09	0.09	0.09	0.11	0.11	0.10
$Time^2$	−0.00	0.00	0.00	0.02	−0.01	0.01
(*s.e.*)	0.00	0.01	0.01	0.01	0.01	0.01
GPS	−0.99	4.43	−0.98	0.53	0.12	0.29
(*s.e.*)	2.47	2.54	2.41	1.36	1.36	1.27
GPS^2	−6.61	−10.2	8.25	−1.71	−0.23	0.81
(*s.e.*)	5.66	5.79*	5.53	1.14	1.15	1.10
$Time * GPS$	0.53	−0.11	−0.27	0.17	0.04	−0.19
(*s.e.*)	0.25**	0.26*	0.24	0.25	0.25	0.24
N	299	295	325	461	461	505

B Sensitivity Analysis Figures

C Falsification Exercise

To validate our identification assumptions we do a falsification exercise. We take a variable that intuitively should not be affected by our treatment and use it as the outcome variable in steps 2 and 3 presented in the section implementation. We use the years of education of the child's grandfather (GrandF) and grandmother (GrandM) for this exercise. Our estimated dose response functions should be rather flat for our procedure to be valid as this would mean that child's grandfather and grandmother education is not affected by the time the mother spends with the child as one would expect.

Similar to the text we present the dose response functions and the tables.

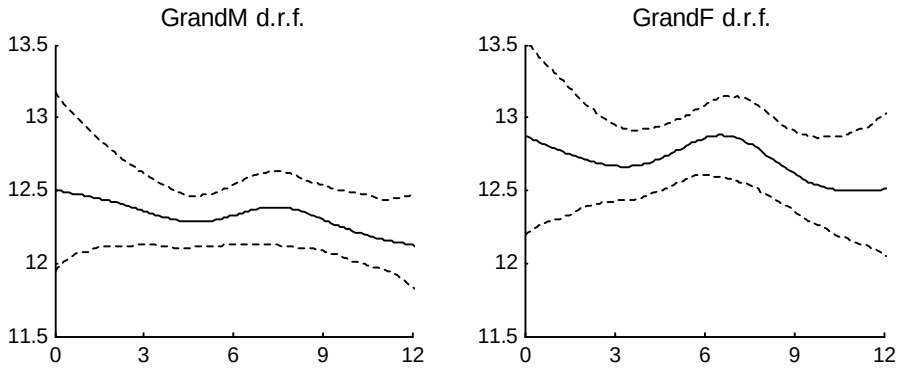


Figure 13- Sensitivity analysis dose response functions for young children

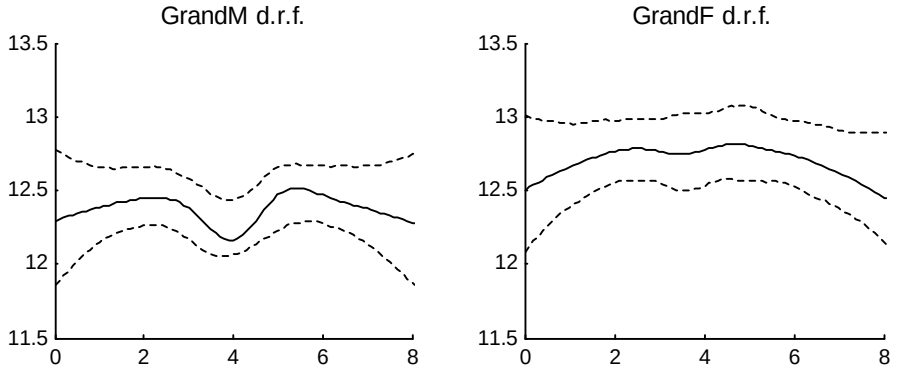


Figure 14- Sensitivity analysis dose response functions for old children

Table 9 - Dose response differences for falsification exercise

		Young children					Old children			
	T	0	3	6	9	T	0	2	4	6
GrandM	3	-0.14	-	-	-	2	0.15	-	-	-
	6	-0.17	-0.03	-	-	4	-0.13	-0.28**	-	-
	9	-0.20	-0.06	-0.03	-	6	0.17	0.02	0.30**	-
	12	-0.37	-0.23	-0.20	-0.17	8	-0.01	-0.17	0.11	-0.19
GrandF	3	-0.21	-	-	-	2	0.26	-	-	-
	6	-0.02	0.19	-	-	4	0.27	0.01	-	-
	9	-0.26	-0.06	-0.25*	-	6	0.23	-0.03	-0.04	-
	12	-0.37	-0.16	-0.35	-0.10	8	-0.05	-0.31	-0.32	-0.28*

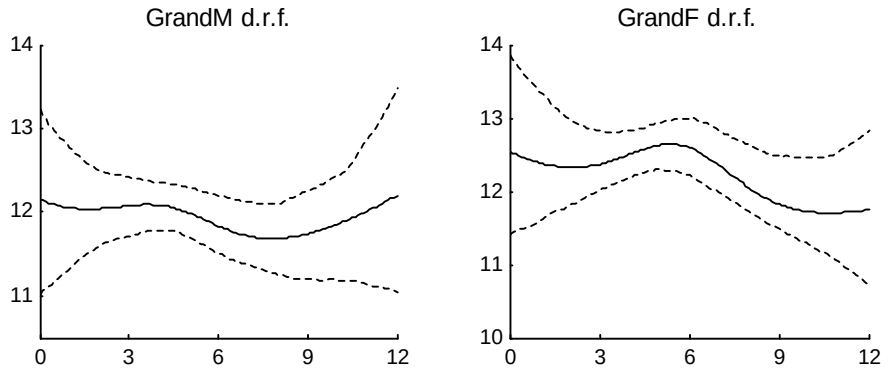


Figure 15- Sensitivity analysis dose response functions for young black children

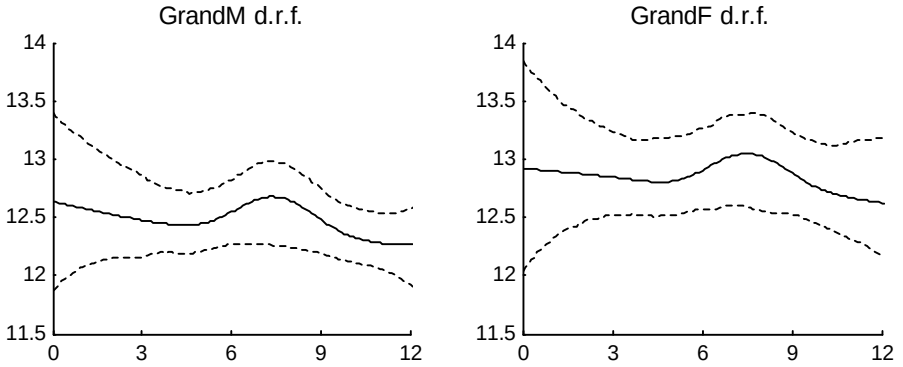


Figure 16- Sensitivity analysis dose response functions for young white children

Table10 - Dose response differences for falsification exercise

		Black Young				Young White			
	T	0	3	6	9	0	3	6	9
LW	3	-0.07	-	-	-	-0.16	-	-	-
	6	-0.31	-0.24	-	-	-0.09	0.07	-	-
	9	-0.40	-0.33	-0.09	-	-0.16	0.00	-0.07	-
	12	0.05	0.12	0.36	0.45	-0.37	-0.21	-0.28	-0.21
AP	3	-0.17	-	-	-	-0.07	-	-	-
	6	0.06	0.23	-	-	-0.02	0.06	-	-
	9	-0.71	-0.55*	-0.77***	-	-0.04	0.03	-0.02	-
	12	-0.79	-0.62	-0.85	-0.07	-0.29	-0.22	-0.28	-0.25

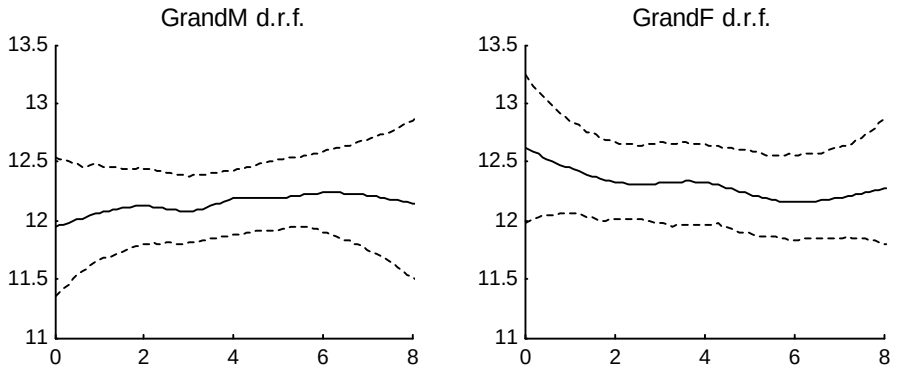


Figure 17- Sensitivity analysis dose response functions for old black children

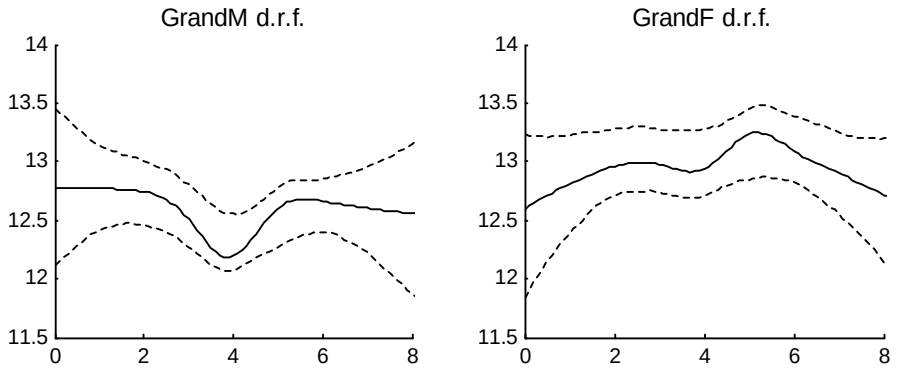


Figure 18- Sensitivity analysis dose response functions for old white children

Table 8 - Dose response differences for falsification exercise

		Old Black				Old White			
	T	0	2	4	6	0	2	4	6
LW	2	0.18	-	-	-	-0.04	-	-	-
	4	0.24	0.06	-	-	-0.59	-0.55***	-	-
	6	0.29	0.11	0.05	-	-0.12	-0.08	0.47***	-
	8	0.20	0.23	-0.04	-0.09	-0.23	-0.19	0.36	-0.11
AP	2	-0.29	-	-	-	0.36	-	-	-
	4	-0.29	-0.01	-	-	0.33	-0.02	-	-
	6	-0.46	-0.17	-0.16	-	0.49	0.13	0.16	-
	8	-0.34	-0.05	-0.05	0.11	0.11	-0.24	-0.22	-0.39

ÁREA III

ECONOMÍA DE LA INNOVACIÓN Y TRANSFERENCIA TECNOLÓGICA

ECONOMIA DA INOVACÃO E TRANSFERÊNCIA TECNOLÓGICA

INOVAÇÃO NOS SERVIÇOS: UM IMPERATIVO ESTRATÉGICO

ESTER AMORIM NOGUEIRA FERNANDES
E-mail: esteramorim@ipg.pt
Instituto Politécnico da Guarda

Resumo

O crescimento económico das últimas décadas está fortemente relacionado com o desenvolvimento do sector dos serviços. A sociedade actual é dominada por este sector, tendo-se tornado o sector mais dinâmico da economia. Tradicionalmente existia a ideia de que o sector de serviços era um sector que desenvolvia pouca actividade inovadora e que consumia inovações produzidas pelo sector industrial. Contudo, nos últimos anos existe a convicção de que o sector de serviços tem um papel importante no uso e na generalização das inovações. Daí, haver um interesse crescente pelo estudo da inovação no sector de serviços e na sua importância como motor de crescimento económico e chave para a competitividade.
Este trabalho pretende definir o que é inovação no sector dos serviços, compreender como as empresas inovam e a importância da inovação como factor estratégico da economia.

Palavras-chave: Inovação, Serviços, Inovação tecnológica.
Área temática: Economia da Inovação e Transferência de Tecnologia.

Abstract

The economic growth society is currently dominated of the last decades is strongly related to the development of the services sector. Society is currently dominated by this sector, having itself become the most dynamic sector of the economy. Traditionally there was idea of that the services sector was a sector that developed a small amount of innovative activity and that it consumed innovations produced by the industrial sector. However, in recent years the belief that the services sector has an important role in the use and the generalization of the innovations has become prevalent. Thus there is an increasing interest for the study of the innovation in the services sector and for its importance as a motor of economic growth and a key for competitiveness.
The objective of this study is to define what innovation is in the services sector, in order to understand how companies innovate and the importance of the innovation as a strategically factor of the economy.

Key words: Innovation, Services, technological Innovation
Thematic Area: Innovation and Technological Transfer Economics.

1. INOVAÇÃO E DESENVOLVIMENTO TECNOLÓGICO

O actual contexto caracteriza-se por mudanças aceleradas nos mercados, nas tecnologias e nas formas organizacionais. Com isso, a capacidade de gerar e absorver inovações vem sendo considerada crucial para que uma organização se torne efectivamente competitiva.

O fenómeno da Revolução Industrial que fragmentou o tecido da sociedade e o teceu novamente num padrão diferente, segundo Davidow & Malone (1992), repete-se actualmente. A complexidade e dinamicidade do ambiente provocam constantes alterações no quotidiano e nas expectativas: o que era inovador ontem tornou-se obsoleto hoje. Com o detalhe que estes "ontem" e "hoje" estão cada vez mais próximos, ou seja, as mudanças ocorrem num ritmo cada vez mais frenético.

Este novo ambiente conduziu a um novo tipo de mercado muito mais aberto, competitivo e com maiores desafios. As organizações precisaram de acompanhar outra área muito dinâmica: a

evolução tecnológica. Surge assim, a partir da década de 80, o que viria a ser uma revolução tecnológica internacional, as designadas indústrias tecnológicas, com base nas Tecnologias de Informação e Comunicação.

Apesar de muitos considerarem, actualmente, que o processo de globalização e a disseminação das Tecnologias de Informação e Comunicação, permitem a fácil transferência de conhecimento, observa-se que, ao contrário desta tese, apenas informações e alguns conhecimentos podem ser facilmente transferíveis. Elementos cruciais do conhecimento, implícitos nas práticas de pesquisa, desenvolvimento e produção, não são facilmente transferíveis espacialmente, pois estão enraizados em pessoas, organizações e locais específicos. Somente os que detêm este tipo de conhecimento, podem ser capazes de se adaptar às velozes mudanças que ocorrem nos mercados e nas tecnologias e gerar inovações em produtos, processos e em formas organizacionais. Desta forma, um dos limites mais importantes à geração de inovação por parte das organizações, países e regiões é o não compartilhar destes conhecimentos que permanecem específicos e não transferíveis.

As novas tecnologias e a necessidade do desenvolvimento da superioridade tecnológica de forma a ser-se competitivo têm feito com que a inovação e a capacidade de inovar adquiram um valor muito mais importante, tornando-se num objectivo central entre organizações e países. A estreita aliança entre a ciência e tecnologia faz com que os Governos, Gestores e as Indústrias disponham hoje de processos tecnológicos, cuja finalidade é conseguir resultados práticos e concretos, as designadas inovações tecnológicas.

No ambiente concorrencial actual, a inovação é essencial para conferir a flexibilidade, a agilidade e a produtividade indispensáveis para a competitividade sustentada e para a modernização permanente. Crescer é um imperativo para as organizações, como tal elas necessitam encontrar soluções para os problemas ou necessidades manifestadas pelos clientes, identificar internamente formas de produção mais eficientes, que minimizem recursos, inventar novos produtos ou serviços. Desta forma, o crescimento das organizações está dependente da sua capacidade criativa e de mobilização de recursos para o encontro de soluções que potenciem a sua estratégia de redução de custos e de diferenciação.

Embora a inovação possa verificar-se em qualquer área de actividade, ela está mais acentuadamente vinculada à dimensão tecnológica. A informação é cada vez mais um instrumento chave para o sucesso, a inovação assume-se como sendo uma condição necessária para o desenvolvimento estratégico organizacional.

A inovação é o modo de mudar uma organização, quer como resposta a mudanças nos micro e macro ambientes quer como uma forma de exercer domínio e influenciar o ambiente em que a mesma se encontra inserida. Contudo, nas organizações a inovação está sujeita a influência individuais, organizacionais e ambientais.

Herriot & Pemberton (1995) referem que, dadas as condições ambientais, as novas prioridades das organizações estão acima e além das constantes preocupações com a eficiência e a produtividade. Elas são hoje a necessidade de inovar e a necessidade de aprender. Mas como a inovação se baseia, normalmente, em conhecimento existente, as organizações e as pessoas só conseguirão inovar se estiverem, receptivas a encetar novos processos de aprendizagem de um modo contínuo.

Assim, é necessário considerar que uma organização não inova sozinha, pois as fontes de informações, conhecimentos e inovação podem-se localizar tanto dentro, como fora dela. O processo de inovação é, portanto, um processo interactivo, realizado com a contribuição de variados agentes económicos e sociais que possuem diferentes tipos de informações e conhecimentos. Esta interacção dá-se em vários níveis, entre os diversos departamentos de uma mesma organização, entre organizações distintas e com outras instituições, nomeadamente as de ensino e de pesquisa.

As informações e conhecimentos passaram, mais recentemente, a ser consideradas importantes para que as organizações se convençam a gerar inovações e enfrentar mudanças, tendo em vista que a solução da maioria dos problemas tecnológicos implica o uso de conhecimento de vários tipos.

Analogamente, e tendo em conta o papel cada vez mais importante desempenhado pelo sector dos serviços na actividade económica, o conhecimento assume uma importância fundamental e estratégica para a inovação. No sector de serviços, o capital humano é muito importante quando a inovação não depende da tecnologia, e são neste caso as pessoas e as suas habilidades que ajudam a conduzir a inovação (OCDE, 2000a).

Presentemente é reconhecido que a inovação é factor chave da competitividade, do desenvolvimento económico e da criação de emprego, e que as novas ideias, os novos métodos e os novos produtos são motores de crescimento económico (Caraça e Martins, 1999). É este o desafio que se coloca, torna-se imperativo que as organizações se tornem mais competitivas e sejam mais eficazes nos seus processos de inovação, isto é, minimizem o tempo de resposta ao mercado, acedam às tecnologias, ao saber-fazer e às ideias, e possuam maior flexibilidade organizacional.

Neste âmbito, a inovação assenta na capacidade da organização para reconhecer oportunidades de mercado ou um serviço diferente, nas suas capacidades internas para responder de forma inovadora e na respectiva base de conhecimento. A procura de mercado, as condições de mercado e as atitudes do cliente/utente influenciam fortemente o comportamento inovador das organizações.

Pode-se, então afirmar, que o processo de inovação consiste num processo interactivo de aprendizagem, realizado através da recolha de informações e gerado pelo conhecimento de diferentes actores económicos e sociais.

Focalizando a criação de conhecimento, Guimarães (1998) define o processo de inovação *"como uma síntese de conhecimentos diversos, integrados à base privada de conhecimento de uma organização através de um processo de aprendizagem"*. Este processo procura a oportunidade tecnológica e a apropriação de benefícios gerados pelas actividades inovativas, actuando simultaneamente. A autora salienta a importância do contexto social, de como uma interacção social pode gerar conhecimento e como o curso do desenvolvimento tecnológico é moldado por essas interacções.

Numa economia cada vez mais baseada no conhecimento, a inovação é fundamental devido aos seus efeitos profundos no trajecto da mudança social e económica, (Stuart e Sorensen, 2000). Nunca em nenhuma época, como hoje, os fenómenos que têm a ver directamente com a inovação tiveram tão grande importância para o envolvimento das nações na moldagem do seu futuro.

De acordo com Rodrigues (1999), os elementos constitutivos de qualquer inovação, considerada em si mesma emergem de três áreas, que, de algum modo, são interdependentes: *"os utilizadores, a tecnologia e o mercado"*. Caraça e Martins (1999) referem que, actualmente no contexto da organização, a inovação poderá consistir, pois, em coisas tão diversas como: "*a renovação e alargamento dos serviços; a introdução de alterações na gestão, na organização do trabalho e nas condições e habilitações dos trabalhadores; na qualidade dos serviços prestados e satisfação dos clientes/utentes",* ou seja, a aposta na melhoria das capacidades das organizações ao nível de todas as suas áreas funcionais, tanto ao nível da geração e introdução de novas tecnologias, como ao nível da mudança organizacional (reforço e revitalização constante da base de conhecimentos da organização) e do desenvolvimento da sua intervenção nos mercados.

De acordo com Mendonça (1998) e Fernandes (2000), um dos pressupostos da inovação é a mudança. A chave para que uma organização se mantenha competitiva é a sua adaptabilidade, ou seja, a capacidade para melhorar o que faz e como faz, reagindo às alterações de mercado ou, melhor ainda, provocando-as.

2. O PROCESSO DE INOVAÇÃO NO SECTOR DOS SERVIÇOS

Os estudos sobre o fenómeno inovatório têm sido dominados pela visão industrialista. Tradicionalmente existia a ideia de que o sector de serviços era um sector que desenvolvia pouca actividade inovadora e que consumia inovações produzidas pelo sector industrial.

Contudo, nos últimos anos há cada vez mais a convicção de que o sector de serviços tem um papel importante no uso e na generalização das inovações, embora a inovação neste sector tenha características específicas. Por isso, tem existido um interesse crescente pelo estudo da inovação no sector de serviços e na sua importância como motor de crescimento económico e chave para a competitividade.

É inequívoco que o crescimento económico das últimas décadas está fortemente relacionado com o desenvolvimento do sector dos serviços. A sociedade actual é dominada por este sector, tendo-se tornado o sector mais dinâmico da economia, com o comércio, os transportes e comunicações, o turismo e os serviços financeiros a apresentar taxas de crescimento positivas.

De facto o sector dos serviços tem actualmente um peso determinante nas economias ocidentais. De acordo com os dados da OCDE, este é o sector que mais contribui para o PIB e cujo peso relativo para a criação de emprego é maior nos países mais desenvolvidos.

A maioria dos estudos existentes sobre a inovação nos serviços prevalecem o papel da informação e do conhecimento, sendo de notar que existem empresas do sector muito inovadoras, que através da inovação procuram melhorar a eficiência e a qualidade no processo de produção de serviços e desenvolver novos conceitos de serviços, baseadas nas oportunidades criadas através do incremento na utilização das novas tecnologias de informação e de comunicação. Exemplos tais como: serviços intensivos em escala ou serviços dependentes das redes de informação (serviços financeiros, serviços imobiliários ou de seguros) e nos chamados "serviços de negócios intensivo em conhecimento" - *Knowledge Intensive Business Sectors* (KIBS)[1] (serviços de consultoria e desenho, serviços de software, etc.).

As empresas de serviços não só dependem intensamente da tecnologia como também canalizam recursos e se organizam para a execução de actividades de investigação e desenvolvimento (I&D). O desenvolvimento tecnológico nos serviços recorre muito ao *stock* de conhecimento disponível para a indústria transformadora, particularmente, no que diz respeito aos sistemas de informação computorizados.

Neste sentido a inovação torna-se bastante mais difícil de estudar nos serviços. Pois, se no caso dos produtos, a invenção pode ser mais facilmente transformada em patentes que podem originar um produto novo com alguma protecção face à imitação. No caso dos serviços, a sua natureza e características dificultam a invenção, pois são facilmente imitáveis, o que se traduz numa maior prevalência de inovações de carácter incremental. Por outro lado a sua heterogeneidade[2], em termos de dimensão de empresas, graus de sofisticação, propriedade pública ou privada, entre outros, leva-nos a considerar que a inovação no sector dos serviços ainda que possa ser caracterizada por alguns traços comuns comporta realidades diferenciadas consoante a classe de serviços que estivermos a analisar.

2. 1. Conceito de Serviço

A análise do processo de inovação no sector dos serviços e, em termos mais latos, a própria delimitação do que é serviço, impõe uma breve reflexão sobre o próprio conceito de serviço e as principais características dos serviços. O Quadro 1 inclui algumas definições de serviço.

[1] KIBS são empresas que dependem do conhecimento ou experiência profissional sobre uma disciplina específica (técnica) e que fornecem produtos e serviços intermédios com frequência, baseados no conhecimento.

[2] Algumas tentativas de clusterização do sector dividem-no em 9 grupos: comércio por grosso, comércio a retalho, hotéis e restaurantes, transportes e comunicações, banca e seguros, agências imobiliárias, serviços de I&D e de tecnologia de informação, serviços para empresas e serviços pessoais (Hollenstein, 2000:17).

Quadro 1 – Definições de Serviço

	AUTORES
"(...) acto ou desempenho que uma parte pode oferecer a outra e que seja, essencialmente, intangível e não resultando na propriedade de coisa alguma e podendo a sua produção estar, ou não, ligado à existência de um produto físico (...)" a definição apresenta, não obstante a sua abrangência, um conteúdo conceptual simultaneamente consistente e não limitativo.	*Kotler,* 1994
"uma actividade ou uma série de actividades de natureza mais ou menos intangível – que normalmente, mas não necessariamente, acontece durante as interacções entre cliente e empregados de serviço e/ou recursos físicos ou bens e/ou sistemas do fornecedor de serviços – que é fornecida como solução ao(s) problema(s) do(s) cliente(s)".	*Gronroos,* 1993
"os serviços proporcionam ajuda, utilidade ou cuidados, experiência, informação ou outro conteúdo intelectual e a maioria do valor é intangível em lugar de residir em qualquer produto físico"	*Department of Industry, Science and Resources*, 1999

O conceito de serviço, embora interiorizado pela maior parte das pessoas, está longe de merecer uma definição consensual, pois deparamo-nos com diferentes interpretações quanto à sua definição o que dificulta o estudo de forma cientifica, do âmbito da inovação nos serviços[3].

É um conceito ambíguo e heterogéneo, já que abarca diversas actividades económicas, todas aquelas que não se incluem no sector primário – fundamentalmente dedicado à obtenção de matérias-primas do meio ambiente – nem no sector secundário – cuja actividade principal é a transformação das matérias-primas em produtos.

Assim a sua definição é mais facilmente inteligível quando é contraposta à de produto. Desta forma, tentaremos ilustrar o conceito de serviço fazendo um paralelismo entre produto e serviço. Um produto é um objecto concreto, enquanto que o serviço é uma representação abstracta. Um produto é tangível, apropriável e transferível, enquanto que num processo de criação do serviço, além deste não poder ser apropriado, a sua prestação é concomitante ao consumo.

O sector dos serviços é, portanto, uma parte muita diversificada da economia que abarca desde sectores intensivos em tecnologia e conhecimentos como serviços informáticos ou de negócios até serviços pouco tecnológicos e pouco qualificados como a maior parte dos serviços pessoais.

A caracterização dos serviços designada por Rust e Oliver (1994) estabelece num quadro o número de componentes principais que compõem um serviço (Quadro 2).

[3] Apenas no final do século XX aparecem as primeiras referências à inovação nos serviços, e, ainda, hoje a EU e a OCDE procura indicadores que meçam inovação nos serviços.

Quadro 2 – Componentes de um Serviço

COMPONENTES	
Produto físico	Componente tangível de um serviço com o qual o cliente/utente é confrontado. Podem aqui considerar-se os impressos a preencher, pedido de informação, etc.
Ambiente de serviço	Em termos gerais, corresponde ao meio envolvente em que decorre o seu fornecimento. Neste sentido, é frequente encontrarmos uma distinção entre ambiente externo e interno. O primeiro incide sobre o conceito mais intuitivo de meio físico, o segundo incide sobre a cultura organizacional subjacente à prestação do serviço.
Projecto do serviço	Diz respeito à sua concepção, ou seja, à forma através da qual se supõe que o serviço seja fornecido. Serão aqui contemplados aspectos tão díspares como a concepção de um produto financeiro ou a especificação de tempos previstos para atendimento de chamadas telefónicas.
Fornecimento do serviço	Tem como finalidade caracterizar de que forma o serviço é efectivamente prestado. Qualquer organização justifica a sua existência trocando com o ambiente externo, bens que a sociedade valoriza.

Fonte: Rust e Olivier (1994): 15

Com base nestes componentes comuns dos serviços, um serviço pode ser todas as actividades que não têm por objecto final a produção de objectos materiais, mas sim o serviço prestado directamente ao cliente, ou seja, um serviço pode ser uma prestação totalmente imaterial ou, incorporar bens materiais com a finalidade de satisfazer necessidades públicas. Esta qualidade híbrida torna-se um pouco mais clara se compararmos as diferenças que distinguem o produto do serviço (Quadro 3).

Quadro 3 – Distinção entre Produto e Serviço

PRODUTO	SERVIÇO
Tangíveis	Intangíveis
Homogeneidade	Heterogeneidade
Produção e distribuição separadas do consumo	Produção, distribuição e consumo constituem um processo simultâneo
Uma coisa	Uma actividade ou processo
O núcleo do seu valor produz-se na fábrica	O núcleo do seu valor produz-se na interacção prestador / utilizador
Os clientes não participam no processo de produção	Os clientes participam na sua produção
Pode ser armazenado	Não pode ser armazenado
Pode transferir-se a propriedade	Não é transferível

Fonte: Gronroos, Ch. (1990)

2. 2. Características dos Serviços

Segundo Kotler (1987), não se pode separar produtos dos serviços. No entanto, Lendrevie *et al.* (1998) acrescentam que os serviços apresentam algumas características particulares, que originam uma abordagem algo de diferente, da que tradicionalmente é considerada para os

produtos. Importante pois, conhecer as características intrínsecas dos serviços que, apesar de não se aplicarem a todo o tipo de serviços, dão indicações acerca das diferenças que existem entre eles e os produtos e as consequências dessas diferenças. As características dos serviços sobre as quais existe consenso são a intangibilidade, a inseparabilidade, a heterogeneidade e a perecibilidade (Zeithaml *et al.* 1990; Murray 1991; Moreira 2000).

A *intangibilidade* está associada ao facto de os serviços não terem nenhuma substância física, não poderem ser armazenados, e se traduza numa impossibilidade de transacção ou transferência de propriedade. Isto não significa, a inexistência de componentes tangíveis associados ao processo de fornecimento do serviço enquanto globalmente considerado. O que fica de posse do consumidor é o resultado da prestação do serviço, ou seja, por exemplo, o transporte, a obtenção da carta de condução, etc.

A *inseparabilidade* pretende reflectir que como prestações que são, os serviços são produzidos e consumidos simultaneamente e no mesmo lugar, não sendo, como tal, susceptíveis de serem consideradas de forma separada, ou seja, prestador e consumidor interagem directamente no momento e local da transacção.

A *heterogeneidade* expressa a ideia de que um serviço pode variar em padrão ou qualidade de um prestador para outro ou de uma ocasião para outra, decorrente da profunda interacção prestador/consumidor. Na generalidade os serviços são menos estandardizados e menos uniformes que os produtos, dependendo muito dos elementos que interactuam. Um consumidor de serviços contribui para a qualidade desse serviço, através do seu bom ou mau humor, pelo seu grau de exigência, pela sua experiência, pela sua competência, pela forma como se envolve no acto de produção e consumo do serviço.

A *perecibilidade* consubstancia-se no desaparecimento do serviço no acto de consumo, ou seja, como os serviços são produzidos e consumidos em simultâneo, não existe a oportunidade de serem produzidos com antecedência e armazenados, pelo que a capacidade produtiva não utilizada se perde.

Estas características de um serviço determinam alguns factores com influência relevante para a prestação organizativa e acarretam determinadas implicações para as empresas de serviços e consequentemente, para os seus clientes. Zeithaml *et al.* (1990), opinam que, dadas as características dos serviços, os consumidores têm dificuldade em proceder à sua avaliação, sendo a qualidade percepcionada em função do processo utilizado para prestar o serviço e não um certificado de qualidade. Por outro lado, a percepção da qualidade irá resultar do encontro das expectativas do consumidor com o desempenho de quem presta o serviço. A qualidade depende muito das pessoas que prestam o serviço e dos seus métodos de trabalho.

Do que ficou expresso, evidencia-se de forma clara que, existe uma grande dificuldade de adequação, às empresas de serviços, dos métodos quantitativos utilizados no desenvolvimento de formas de obtenção de qualidade em empresas industriais. Nestas são abordados os processos de concepção e de produção, tendo como base as especificações e suas tolerâncias, sendo feita posteriormente a medição para verificação da sua conformidade. Relativamente às empresas de serviço é muito mais difícil, onde cada prestação é um caso único, tornando-se a sua padronização difícil, uma vez que existem muitas variáveis aleatórias.

2. 3. Qualidade dos Serviços

Peritos de qualidade, como Crosby (1980), definem qualidade como a *"conformidade com as exigências"*. Neste caso, segundo Asubonteng e*t al.* (1996), é assumido que as variáveis da qualidade estão dependentes das suas exigências, uma vez que Crosby define qualidade como *"ciclo sem fim de melhoria contínuas"*, ou seja, a qualidade não é um destino, mas antes uma caminhada.

Em geral, o consumidor determina o que num produto ou num serviço representa qualidade. O consumidor determina quando o nível de qualidade é alto ou baixo e fá-lo usando toda a informação disponível, resultante da sua experiência, expectativas ou por raciocínio dedutivo.
Porém, no caso dos serviços da Administração Pública, o consumidor está limitado pela quantidade e qualidade da informação disponível. A qualidade dos serviços da Administração

Pública visa a satisfação de necessidades explícitas ou implícitas. Estas necessidades apresentam três perspectivas:

- A do cidadão/utilizador que espera que lhe sejam prestados os melhores e mais adequados serviços em tempo útil;
- A dos financiadores e gestores que esperam que estes cuidados sejam prestados com a máxima eficiência, efectividade, equidade e eficácia;
- A dos prestadores que esperam que exista uma gestão eficaz de recursos que lhes propicie a necessária aplicação dos seus conhecimentos às situações com que se deparam no dia a dia.

Esta perspectiva de *feedback* dos consumidores tem ganho importância a partir da última década, em consequência da evolução tecnológica. Esta atitude de *feedback* constante em relação aos serviços, permite determinar o que é qualidade para o utente, identificar serviços necessários, mas não disponíveis, identificar serviços a serem melhorados, determinar as necessidades fundamentais dos utentes no serviço e visualizar o valor da melhoria contínua dos serviços.

Gronroos (1993) divide a avaliação da qualidade do serviço pelo cliente em dimensões técnicas e funcionais. Estas dimensões de uma forma simplificada, significam que a qualidade é avaliada não só pela eficiência, mas pela forma como o cliente percebe o serviço. O autor aponta seis critérios para a avaliação de serviços:

1. *Profissionalismo e Habilidades*: conhecimentos e habilidades envolvendo a equipa, na solução de problemas.
2. *Atitudes e Comportamento*: forma como os funcionários se preocupam e se interessam na resolução de problemas.
3. *Facilidade de Acesso e Flexibilidade*: capacidade do sistema em facilitar ou ajustar-se às necessidades dos clientes.
4. *Confiabilidade e Honestidade*: cumprimento do serviço prometido aos clientes.
5. *Recuperação*: acções tomadas para correcção e imprevisibilidades ou erros no processo.
6. *Reputação e Credibilidade*: imagem do prestador do serviço perante os clientes.

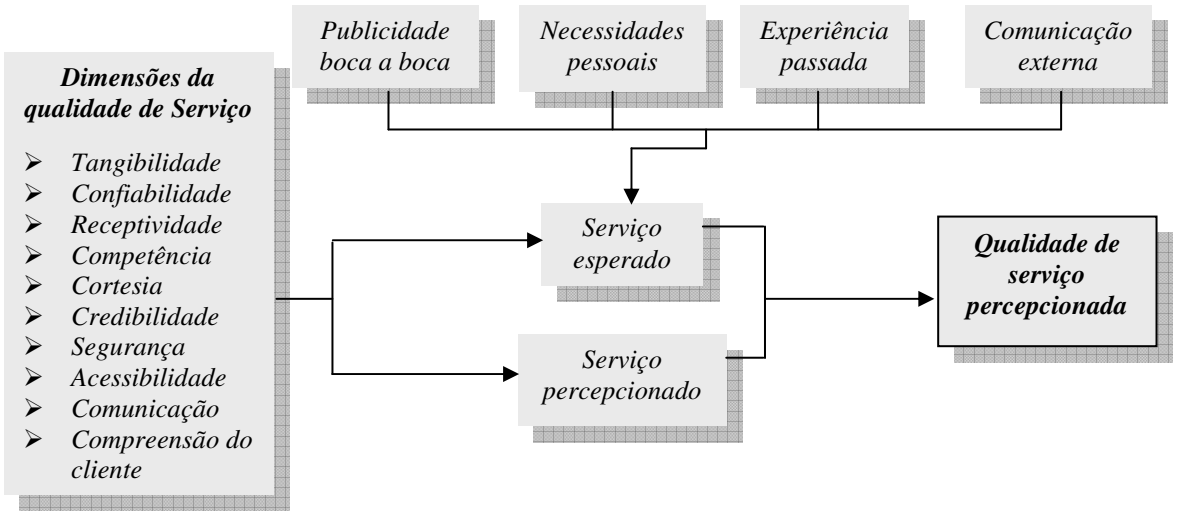
Segundo Berry (1996), a qualidade é definida pelo cliente, que avalia o desempenho segundo dois padrões: o que desejam (serviço desejado) e o que aceitam (serviço adequado). Para melhorar a qualidade nos serviços, o autor sugere a *"tomada de impressões"* sistemáticas sobre o desempenho do serviço na óptica do cliente, assim como as expectativas do serviço. Barros (2002) chama essa *"tomada de impressões"* de conceito integrado de qualidade.

Parasuraman, Zeithaml e Berry citados em Davis *et al.* (2001) identificaram as dez dimensões genéricas que contribuem para o nível da qualidade do serviço que uma organização proporciona aos clientes:

- *Tangibilidade*: são as evidências físicas do serviço.
- *Confiabilidade*: relaciona-se à consistência do desempenho.
- *Receptividade*: refere-se à disposição e/ou prontidão dos funcionários para prestar um serviço.
- *Competência*: relacionada às habilidades e ao conhecimento dos trabalhadores para realizar o serviço apropriadamente.
- *Cortesia*: refere-se à gentileza, ao respeito, à consideração e ao comportamento amigável do pessoal que está em contacto com o cliente.
- *Credibilidade:* refere-se às características de confiança, possibilidade de acreditar e honestidade no trabalhador que presta o serviço.
- *Segurança*: refere-se à isenção de qualquer perigo, risco ou problema.
- *Acesso*: refere-se à acessibilidade e à facilidade de contacto.
- *Comunicação*: disponibilização de informações sobre o serviço em andamento ou sobre os procedimentos.
- *Compreensão a respeito do cliente*: refere-se ao esforço feito pelo trabalhador que realiza o serviço para entender as necessidades específicas de cada cliente.

A figura 1 mostra como as expectativas derivam de várias fontes:

Figura 1 – Qualidade Percebida do Serviço



Fonte: Fitzsimmons & Fitzsimmons (2000): 250

De acordo com Gianezi & Corrêa, (1994: 196), a qualidade em serviços pode ser definida como o grau em que as expectativas do cliente são atendidas ou excedidas pela percepção do serviço prestado. Os mesmos autores apontam alguns factores que indicam a baixa qualidade em empresas de serviços:

- Pouca formação e motivação dos funcionários, considerados como mão-de-obra temporária;
- Excessiva ênfase no corte de custos e aumento de produtividade quantitativa;
- Falta de maior exigência por parte dos clientes;
- Pouca concorrência;
- Dificuldade em padronizar os serviços;
- Falta de tempo para inspecções e correcções, já que o serviço é prestado e consumido instantaneamente;
- O serviço geralmente é intangível, sendo difícil de medir e controlar.

Os serviços, como resultado de actos sociais e do contacto directo entre o cliente e os representantes da organização, determinam que a qualidade percebida seja atribuída exactamente no preciso momento em que o prestador do serviço e o cliente se confrontam Porém convém salientar a importância da participação do cliente no processo de produção do serviço como um atributo fundamental dos serviços, o *"cliente é quase sempre mais do que um simples cliente"* (Normann, R. 1990: 7-8). É este atributo que distingue a distribuição de bens e a de serviços e constitui uma referência obrigatória para a concepção e aplicação de novas tecnologias e formas organizacionais.

2. 4. Formas de Inovação no Sector de Serviços

Autores como Hipp *et al.* (2000); Coombs e Miles (2000), referem que cada vez há mais certeza de que o sector dos serviços joga um papel importante no uso e na generalização das inovações, ainda que a actividade inovadora deste sector tenha características específicas já mencionadas anteriormente.

Estudos da OCDE (2000b) salientam que o sector de serviços está a ser cada vez mais inovador. Observa-se um rápido crescimento nas actividades de I&D nos serviços, com um aumento da percentagem de participação dos serviços nos gastos totais em I&D. Além do mais, as indústrias de serviços, como os serviços financeiros e de negócios, são grandes utilizadores

das novas tecnologia e de bens intensivos em I&D, principalmente Tecnologias de Informação e Comunicação.

Uma das particularidades do sector dos serviços é ser um importante incentivo na preparação e formação dos Recursos Humanos, factor competitivo que joga um papel chave nas estratégias inovadoras das empresas.

Estudos recentes (OCDE, 2000c; Miles *et al.*, 1995) também demonstram que os serviços beneficiaram em grande medida do uso das tecnologias incorporadas nos novos bens de capital.

Sucede que, a inovação nos serviços apresenta características próprias. Por isso, é importante recordar os quatro aspectos típicos das indústrias de serviços identificados na literatura (Evangelista e Sirilli, 1995) que têm implicações no momento de definir e analisar a inovação nos serviços:

- A interacção estreita entre produção e consumo.
- O conteúdo intensivo em informação das actividades de serviços e de produção.
- O papel fundamental dos Recursos Humanos como factor básico de competitividade.
- A importância dos factores organizativos na *performance* das empresas.

A interacção estreita entre produção e consumo leva a que uma grande parte das actividades de inovação em serviços esteja orientada na adaptação/*customization*[4] dos serviços às necessidades dos clientes ou utilizadores.

O conteúdo intensivo em informação confere uma grande importância na generalização e uso das Tecnologias de Informação e Comunicação nas actividades de inovação das organizações de serviços. O papel fundamental do factor humano na organização e provimento dos serviços associa-se a grandes incentivos em Recursos Humanos e, por isso, as actividades de formação das organizações devem ser consideradas como fonte de desenvolvimento das capacidades tecnológicas das organizações de serviços. A importância do factor organizativo no sector de serviços requer uma revisão do conceito tradicional de inovação para incluir as mudanças organizativas.

Tendo em conta estes aspectos característicos dos serviços, Evangelista e Sirilli (1995) e o relatório de Eurostat (1995) aconselham uma modificação dos conceitos e definições de inovação desenvolvidos inicialmente para o sector de indústria transformadora para adapta-los ao sector de serviços. O Quadro 4 inclui algumas definições de inovação dos serviços.

Quadro 4 – Definições de Inovação nos Serviços

	AUTORES
"Produzir um serviço (…) é arranjar solução para um problema, esta não envolve como actividade principal o fornecimento de um bem. Traduz-se na disponibilização de um conjunto de competências (humanas, tecnológicas e organizacionais) ao serviço do cliente para arranjar uma solução".	*Gadrey,* 1995
"Inovação nos serviços é principalmente pequenos ajustamentos nos procedimentos de carácter incremental e raramente radical. O tempo necessário ao desenvolvimento da inovação nos serviços é geralmente curto (…). O processo de inovação nos serviços é normalmente muito prático".	*Sundbo and Gallouj,* 1999
"Inovação nos serviços é por definição multidimencional. Quando comparado com a industria, por exemplo, a inovação nos serviços caracteriza-se por uma maior ênfase na dimensão organizacional (novos conceitos de serviços, novos interfaces com os clientes, novos sistemas de distribuição) comparativamente às opções tecnológicas".	*Van Ark et al,* 2003

Para estes autores, as inovações no sector dos serviços compreendem novos serviços e novas formas de produzir ou fornecer serviços, assim como mudanças significativas nos serviços já existentes ou na sua distribuição e entrega. As inovações serão o resultado de incentivos substanciais no novo conhecimento, com ou sem tecnologia. As inovações de processo são

[4] Concepção de produtos ou serviços à medida do consumidor.

formas novas ou melhoradas de forma significativa de produzir ou fornecer serviços para o exterior. Portanto, as inovações de processo podem ser inovações nos processos de produção dos serviços ou nos processos de distribuição e entrega.

Contudo, esta nova classificação esquece-se das inovações organizativas e cada vez há mais certeza de que a actividade inovadora em serviços é desta natureza (Gallouj, 1998; Sundbo e Gallouj, 1998a; Miles, 1994). As maiorias das inovações em serviços são não tecnológicas e implicam mudanças pequenas e incrementais nos processos e procedimentos, de forma que não costumam requerer muita I&D.

Sundbo e Gallouj categorizaram a inovação nos serviços em 5 tipos de acordo com as suas motivações e natureza da organização que presta o serviço (Quadro 5).

Quadro 5 – Formas de Inovação nos Serviços

FORMAS DE INOVAÇÃO	"*CUSTOMISATION*" CARACTERÍSTICAS	FORNECEDOR DE SERVIÇO	EXEMPLOS
Inovação "por medida"	Os produtos são feitos para clientes específicos, atendendo aos seus pedidos e necessidades.	Pequenas empresas Serviços operacionais	Projectos de Arquitectura ou Engenharia
Inovação Ad Hoc	Implementação de um serviço que requer uma adaptação a um problema em cooperação com o cliente de uma forma interactiva e no contexto de um cliente especifico.	Serviços Conhecimento intensivos Fornecimento de serviços por "*experts*" em determinada matéria	Consultadoria, Advocacia, Solicitadores
Inovação baseada em recombinações ou Arquitectural	Novos produtos são obtidos através da dissociação ou de novas combinações de elementos dos serviços *standardizados*.	Serviços Financeiros Serviços Operacionais	Crédito ao consumo, Contas poupança reformados
Diferenciação ou Inovação adjunta	Novos produtos são obtidos através da adição de serviços periféricos novos ou melhorados.	Comércio	Venda do automóvel + Serviço de crédito Assistência pós-venda
Inovação baseada na Distribuição	Quando o modo de entrega ou de interacção com o cliente é alterado.	Serviços Financeiros Serviços Operacionais	*e-commerce* *e-banking*

Fonte: Adaptado de Sundbo e Gallouj (1998a)

Distinguem, ainda, quatro tipos de inovações em serviços, tendo em conta a importância das inovações organizativas (Sundbo e Gallouj (1998a):

1. *Inovações de produto*. Implicam a apresentação de serviços novos ou melhorados aos clientes. Por exemplo, os bancos oferecem nos últimos anos, novas formas de contas bancárias; as empresas de software introduzem novos produtos melhorando estes de forma regular.
2. *Inovações de processo*. São novidades ou melhorias dos processos de produção e provisões do serviço. Estas inovações podem-se dividir em duas categorias: inovações nos processos de produção ou inovações nos processos de distribuição e entrega (provisões).
3. *Inovações de mercado*. São novos comportamentos no mercado, como por exemplo, encontrar um novo segmento de mercado.
4. *Inovações organizacionais*. São novas formas de organização ou de gestão nas empresas. Estão muitas vezes relacionadas com as inovações tecnológicas.

As inovações organizacionais contribuem de forma importante na melhoria da produtividade e qualidade dos serviços. Alguns tipos de mudanças organizativas podem estimular inovações para enfrentar novas ameaças e dificuldades: por exemplo, a competência em serviços de

telecomunicações leva a novas necessidades de como estabelecer, aceder e manter os directórios. Outro tipo de inovações organizacionais é a procura da garantia de qualidade.
Normann (1991) identifica três tipos de inovações organizacionais nos serviços:

1. *Novas combinações de serviços.* Implicam uma estratégia de diversificação, mas também podem haver economias de alcance associadas às novas combinações (por exemplo, os dados sobre clientes podem ser reutilizados). A esta forma de inovação podem associar-se características como a intangibilidade e simultaneidade.
2. C*liente como co-produto*. Focalizada estritamente em actividades de auto-serviço. A fronteira entre a actividade do produtor e o consumidor não é fixa. Põe em realce a importância do cliente para antecipar a experiência dos fornecedores de serviços e identificar novas vias para a sua aplicação.
3. *Organização de serviço multi-unidade*. Em muitos casos é difícil definir uma fronteira entre a produção e consumo de serviços. Por conseguinte, na maioria dos casos, o crescimento da organização conduzirá à reprodução do sistema de administração de serviço noutro local, perto de um novo mercado (criação de unidades novas). Exemplo da Hertz pela reprodução do sistema de administração de serviços e da OPG que desenvolveu uma fórmula de franquia para farmácias.

As pesquisas sobre a inovação (Evangelista e Savona, 1998; Pattison *et al.*, 1995) sugerem que poucas empresas de serviços realizam um só tipo de inovação, já que geralmente, se praticam ao mesmo tempo inovações de produto, de processo e organizacional. As inovações em serviços têm uma grande dependência de *software* e frequentemente implicam inovações organizativas e tecnológicas complementares. O problema é que as inovações organizativas são mais difíceis de medir que as inovações tecnológicas, apesar de que as primeiras são as inovações mais frequentes nos serviços (Miles, 1995; Marklund, 1998).

O aparecimento das Tecnologias de Informação e Comunicação, que facilitam o trabalho virtual, é apresentado como fonte de crescimento da inovação de serviços.

Evans e Wurter (2000) sublinharam que o tradicional quadro de serviços que são oferecidos como uma norma a um mercado mais vasto ou mais especializado e personalizado, destinado aos de maior poder de compra está inundado de *bits* pelas oportunidades das tecnologias da *WEB*. Sendo possível oferecer simultaneamente riqueza e acessibilidade, e criar novos mercados. De acordo, com Tidd *et al* (2003), o desafio da Internet não compreende apenas as grandes empresas de serviços e as empresas de vendas a retalho, pode implicar a sobrevivência de muitas pequenas e médias empresas, exemplificando com uma empresa do sector de turismo. Uma pequena Agência de viagens que passa a disponibilizar via *Net* os seus pacotes turísticos, podem do o consumidor reservá-los no conforto do seu lar a mais baixos preços. O processo de inovação nos serviços funciona em rede, entre o sistema interno e o ambiente externo. A organização interna da empresa pode ser mais favorecedora do intraempreendedorismo, ou seja, do aparecimento e desenvolvimento de novas ideias no seio das empresas que a podem diferenciar no mercado gerando melhores resultados.

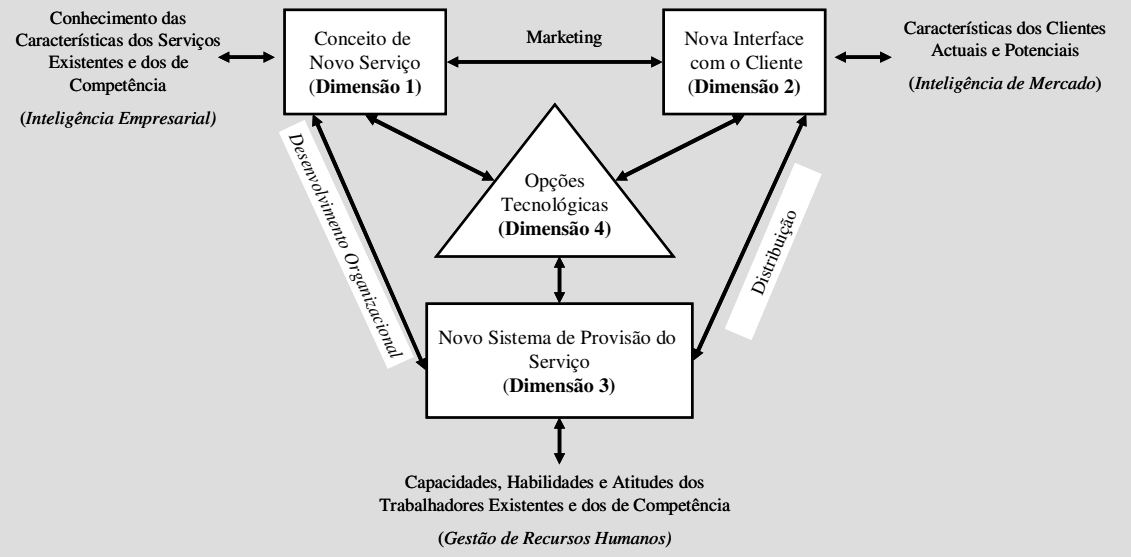
As redes de inovação envolvem empresas do sector dos serviços e industrial em trabalho conjunto. Pelo que, qualquer estudo do sector serviços não se poderá dissociar dos outros sectores.

2. 5. Modelos de Inovação no Sector dos Serviços

A inovação nos serviços raramente se limita a uma mudança das características do serviço em si mesmo, pelo contrário implica frequentemente novas formas de distribuição do produto, de interacção com o cliente, de controlo de qualidade, etc. Na prática a maioria das inovações parecem ser uma combinação de mudanças e de melhores serviços anteriormente existentes.

Bilderbeek *et al.* (1998) apresentam um modelo de quatro dimensões de inovação nos serviços que serve de ferramenta para descrever a actividade inovadora do sector de serviços. A figura 2, mostra como as inovações se situam em qualquer das quatro dimensões da actividade do sector.

Figura 2 - Modelo Dimensional da Inovação nos Serviços



Fonte: Bilderbeek *et al.* 1998.

Dimensão 1 - Conceito de novo serviço. As inovações de serviços podem ser visíveis, sobretudo quando implicam fornecimento do produto, mas geralmente são uma nova ideia o conceito de como organizar uma solução a um problema. Ou seja, o fundamental é que a sua aplicação seja novidade dentro de um mercado particular. Por exemplo, a introdução de um serviço de transporte porta a porta é um novo conceito de transporte que atende às necessidades dos utilizadores. A sua introdução supõe um novo conceito de transporte no sector de transporte público.

Dimensão 2 - Nova interface com o cliente. Um elemento importante das inovações em serviços é o *design* da interface entre o prestador do serviço e os clientes. No sector de serviços é cada vez mais frequente a oferta e comercialização dos produtos ser orientada de forma específica a grupos de consumidores atendendo em particular às suas necessidades e exigências. Por isso, cada vez é mais importante a comunicação entre os prestadores do serviço e os clientes, gerando assim uma grande área para a introdução de inovações.

Estas inovações nas relações entre prestadores e os clientes geralmente implicam tecnologias de informação. A introdução de centros de chamadas ou de informação ao cliente (que podem ser assistidos mediante computador ou Internet) são exemplo de mudanças na forma de abordar e atender os clientes.

Dimensão 3 - Novo sistema de fornecimento (distribuição e entrega) do serviço e organização. Esta dimensão centra-se na relação entre prestador e os clientes. Refere-se aos arranjos organizativos internos que se fazem para permitir aos trabalhadores realizar melhor o seu trabalho e oferecer os produtos de serviços de forma adequada. Um exemplo é a introdução do comércio electrónico nas compras, que supõe uma mudança na forma de relacionamento entre prestador e cliente e uma mudança na forma de fornecer o serviço.

Dimensão 4 - Opções tecnológicas. Como referem Bilderbeek *et al.* (1998), a inovação em serviços é possível sem inovação tecnológica, quer dizer, a tecnologia não é sempre uma dimensão. Segundo Djellal e Gallouj (1999), há serviços em que a tecnologia é um factor externo na produção das características do serviço uma vez que este poderia fornecer-se sem ela. Porém, em muitas inovações em serviços a tecnologia joga um papel importante. Na prática há uma gama de relações entre tecnologias e inovação de serviço. Não obstante, o conhecimento das opções tecnológicas disponíveis, o grau de disponibilidade das tecnologias necessário ao conhecimento essencial para que possam procurar as suas necessidades de tecnologia varia de uma empresa de serviços para outra. Actualmente as tecnologias mais comuns em serviços são as tecnologias de informação.

Este facto deve-se em grande medida ao carácter intangível dos serviços e ao papel fundamental que a informação e o conhecimento jogam em muitas destas actividades.

Em geral qualquer inovação em serviços implica uma combinação das quatro dimensões mencionadas. Um serviço totalmente novo requer o desenvolvimento de um novo sistema de fornecimento do produtor, provocando mudanças na organização do trabalho e na relação prestador-clientes, gerando novos usos das Tecnologias de Informação e Comunicação.

Vários autores como Pavitt (1984); Soete e Miozzo (1989) consideram que os sectores da economia têm distintos modelos de inovação (produto e processo) e que desenvolvem tecnologias e inovações por várias razões. Assim, Pavitt (1984) classifica os modelos de inovação dos sectores em quatro categorias:

- *Sectores dominados pelos fornecedores*. São aqueles em que as empresas ou indústrias do sector não desenvolvem directamente uma grande actividade inovadora. Esta surge incorporada nas matérias-primas, na maquinaria e equipamento que adquirem aos seus fornecedores. Entre as empresas *"dominadas pelos fornecedores"* a competências tende a basear-se nas habilidades da mão-de-obra e preços, do que em vantagens tecnológicas. Encontram-se dentro deste grupo o sector têxtil, o sector de pronto-a-vestir, o sector da madeira e couro, etc.
- *Sectores produtores em grande escala*. Nestes sectores a inovação surge fundamentalmente pela necessidade de reduzir custos, já que produzem produtos destinados a grandes mercados com um certo grau de estandardização. Dentro deste grupo destacam-se as indústrias de automatização, de alimentação e bebidas, de combustível e refinação de petróleo, indústria naval, indústria tabaqueira, etc.
- *Sectores de fornecimentos especializados*. São empresas que realizam desenvolvimentos próprios e mantêm-se ligadas ao processo inovador dos seus fornecedores e dos seus clientes. É o caso das máquinas de calcular e computadores.
- *Sectores de base científica.* São sectores em que a inovação assenta essencialmente num intenso nível de investigação própria, como é exemplo da indústria farmacêutica, as telecomunicações, etc.

Dentro desta classificação, Pavitt (1984) descreve os serviços (de mercado) como empresas *"dominadas pelos fornecedores*", em lugar de serem produtores de novas tecnologias.

Posteriormente Pavitt *et al.* (1989) acrescentam uma nova categoria na classificação sectorial: os sectores *"intensivos em informação"*. Esta nova categoria destaca-se pela importância das Tecnologias de Informação e Comunicação numa trajectória de inovação baseada no processamento de informação e no fornecimento de produtos relacionados. Para estes autores, nesta nova categoria incluíam-se os serviços financeiros e a venda a retalho.

Ao mesmo tempo, Soete e Miozzo (1989) identificam de acordo com os distintos estilos inovadores das actividades de serviços quatro tipos de empresas de serviços:

- *Sectores dominados pelos fornecedores*. Aqui estariam incluídos os serviços públicos e sociais, como a educação, a saúde e administração pública, e os serviços pessoais (reparações, hotéis, bares, restaurantes, comercio, etc.). São serviços dominados pelos fornecedores de equipas e sistemas técnicos, de onde as empresas não participam de forma significativa na produção das tecnologias de processo que utilizam. São sectores, portanto, que dependem das tecnologias desenvolvidas por outros sectores.
- *Serviços intensivos em produções, intensivas em escala e serviços de redes*. São serviços onde existe uma grande divisão de trabalho através da simplificação e coordenação das actividades de produção e/ou fornecimento, e a substituição de mão-de-obra qualificada por máquinas. Nestes sectores dão-se inovações que se baseiam na introdução de tecnologias de informação para reduzir custos. Dentro deste grupo identificam-se dois tipos de serviços:
  1. *Serviços de redes*. Trata-se de sectores que dependem de redes de Tecnologias de Informação e Comunicação, como, por exemplo, banca, seguros, e serviços de telecomunicações. O desenvolvimento das Tecnologias de Informação e Comunicação tem permitido melhorar a precisão e qualidade dos serviços oferecidos pelos fornecedores; têm facilitado a personalização dos serviços e

desempenham um papel fundamental no estabelecimento de padrões em muitas actividades de serviços.

2. *Serviços intensivos em escala.* Estes são, serviços dependentes das redes físicas (por exemplo, serviços de transporte e turismo, comercio de distribuição); são serviços que dependem fortemente das tecnologias hardware desenvolvidas no sector da indústria transformadora, pelo que desempenham um papel fundamental na definição e especificação de inovações, influenciando assim os fornecedores de novas tecnologias.

- *Fornecedores especializados em tecnologia e sectores de base científica*. São sectores que incluem serviços de negócios especializados, serviços de software, serviços de laboratórios e desenho. São sectores muito dinâmicos, de onde a maior fonte de tecnologia é a actividade inovadora dos serviços em si mesmos que costuma ser fruto da investigação, desenvolvimento e das actividades de software das próprias empresas.

Soete e Miozzo (1989) afirmam que um mesmo sector, como é o caso das comunicações, pode encontrar-se em mais do que uma categoria, já que é um sector intensivo em escala e de base científica, mas também é um sector de fornecimento especializado.

3. MEDIR A INOVAÇÃO NOS SERVIÇOS: INDICADORES DE INOVAÇÃO

Os governos e os políticos estão, em geral, cada vez mais interessados em promover a inovação para estimular o crescimento económico e o emprego. Há uma necessidade crescente de medir e avaliar a inovação e a mudança tecnológica para haver um maior conhecimento sobre as forças que conduzem à inovação e as suas consequências socio-económicas (Kleinknecht, 2000).

Contudo, alguns autores afirmam que a caracterização, a nível quantitativo, da inovação torna-se num processo bastante complexo devido à dificuldade em determinar que tipos de dados são relevantes. *"O produto gasto em I&D, ou número de investigadores ou cientistas por cada mil habitantes fazem parte do cervo estatístico padrão que é utilizado para caracterizar e comparar países entre si"* (Conceição e Ávila, 2001: 3).

Defendem ainda, que este tipo de amostragem de processo de inovação não é suficiente, salientando que o número de investigadores ou cientistas representa apenas recursos afectos às actividades ligadas ao processo de inovação. Outros indicadores referem-se ao número de submissão ou concessão de patentes pedidas como defendem alguns autores como (Griliches, 2000) ou (Jaffe *et al.*, 2000).

Mas, o processo de avaliação de inovação torna-se tão complicado que o mesmo pode acontecer em pequenas organizações ou pessoas que não são investigadores ou cientistas e que abordam a inovação como uma nova descoberta, uma necessidade de conhecimento e sobrevivência sem essa estar directamente ligada a centros de I&D. Por outro lado, é ainda de realçar que é necessário determinar se as inovações são realizadas apenas para benefício da organização ou se correspondem a inovações mais generalistas que englobam não só o interior da organização mas também inovações do contexto exterior. Outro aspecto essencial para a quantificação da inovação prende-se com o facto da mesma não ser limitada a organizações ou a pessoas mas também a serviços.

A maioria da literatura tem-se centrado, para medir a inovação, em dois indicadores de inovação tecnológica: O gasto em I&D e o número de patentes.

Como refere Martínez-Ros (1998), a primeira medida é um *input* do processo de produção das inovações, enquanto que a segunda proporciona informação sobre o *output* do processo de inovação tecnológica.

Tradicionalmente, no sector da indústria transformadora são os dados sobre as actividades de I&D das organizações que se utilizam como principal indicador dos investimentos em processos de inovação tecnológica. Porém, segundo Marklund (1998) no caso do sector de serviços a informação disponível sobre I&D tem algumas limitações na análise de actividades de inovação:

- I&D é apenas uma parte dos processos de inovação, tanto em termos de custo como em termos de processos de gerar conhecimentos.
- As pesquisas de I&D não cobrem todos os sectores de serviços e os sectores que são abrangidos não o são de forma adequada.
- As pesquisas centram-se nas grandes empresas, em detrimento das PME's.

Como a maioria das inovações em serviços estão ligadas a mudanças nos processos, na organização e nos mercados, o gasto em I&D só abrange uma pequena parte do esforço inovador das organizações de serviços (Pilat, 2000). Além do mais, como a inovação nos serviços não é necessariamente tecnológica pode realizar-se, sem I&D prévio (Djellal e Gallouj, 2000). Ou seja, como I&D é só um *input* do processo de inovação, especialmente nos serviços, e os *inputs* podem ser utilizados mais ou menos eficientemente, seria recomendável dispor de indicadores de *output* do processo de inovação. Para além de existirem outros *inputs* que incluiriam análises de mercado, formação de empregados, desenho de produto, etc.

Quanto ao segundo indicador, o número de patentes, só serve para medir a inovação tecnológica no sector da indústria transformadora, esquecendo-se da inovação organizacional, uma vez que no sector dos serviços não é aplicável porque as inovações são dificilmente patenteáveis, ainda que nalguns serviços se utilizem os *copyrights* com êxito. Em qualquer caso, a importância da informação como componente de muitos serviços faz com que, em geral, seja muito difícil proteger as inovações. Vários autores (Levin *et al.,* 1987; Brouwer e Kleinknecht, 1999) demonstraram que muitas vezes as organizações não consideram a protecção de patentes como o meio mais apropriado para beneficiar das inovações. Outras formas de protecção da propriedade intelectual, como os direitos de autor ou de marcas, costumam ser mais utilizadas nos serviços. Não obstante, factores como a vantagem temporal sobre os concorrentes, o manter o secretismo sobre a inovação, ou reter o pessoal qualificado na organização tendem a ser mais valorizados. A boa imagem da organização serve geralmente como protecção à cópia da inovação.

Por outro lado, a qualificação, formação e habilidades pessoais dos empregados de uma organização são um factor primordial que tornam possíveis as actividades de inovação. Portanto, podia-se utilizar como indicador da distribuição e dos fluxos de recursos para as actividades de inovação, a informação sobre os recursos humanos das organizações de serviços.

Outra fonte que facilita o desenvolvimento de actividades de inovação nas organizações é as capacidades organizativas das mesmas. Marklund (1998, 2000) sugere incluir nos estudos sobre inovação indicadores de organizações flexíveis e mudanças organizativas. Para Marklund (2000), as mudanças na estrutura organizativa das organizações podem ser interpretados frequentemente como uma inovação em si mesma. Contudo, hoje em dia os estudos actuais sobre inovação não costumam incluir indicadores sobre mudanças organizativas.

Por último, tradicionalmente mede-se o impacto ou eficiência económica e a *performance* inovadora do sector dos serviços em termos de produtividade. Contudo, as medidas de produtividade são em geral problemáticas, principalmente no sector dos serviços. Em muitos serviços as medidas do produto ou *output* são de duvidosa qualidade, principalmente pela falta de dados ou pela dificuldade de definir o *output* (Pilat, 2000). Por isso, o conceito de produtividade deve ser utilizado com cautela ou adaptar a sua interpretação ao sector dos serviços.

BIBLIOGRAFIA

ASUBONTENG, P; MCCLEARY, K. & MUNCHUS, G. (1996) –"The evolution of quality in the US health care industry: an old wine in a new bottle", *International Journal of Health Care Quality Assuranc***e,** N.º 9, pp. 11 - 19.

BARROS, Claudius D´Artagnan C. (2002) – *Excelência em Serviços – Uma Questão de Sobrevivência no Mercado*, 2.ª Edição, Rio de Janeiro, Editora Qualitymark.

BERRY, Leonard L. (1996) – *Serviços de Satisfação Máxima – Guia Prático de Acção*, Rio de Janeiro. Editora Campus.

BILDERBEEK, R., P. den HERTOG, G. MARKLUND e I. MILES (1998) - "Services in Innovation: Knowledge Intensive Business Services (KIBS) as Co-Producers of Innovation". SI14S Synthesis Paper nº 3.SI14S Project, STEP Group.

BROUWER, E. y A.H. KLEINKNECHT (1999) – "Innovative Output and a Firm's Propensity to Patent", *Research Policy*, pp.27.

CARAÇA, J. e MARTINS, M. (1999) – "Inovação, Desenvolvimento Empresarial e Bem - Estar Social", *Economia & Prospectiva*, n.º 10, Jul-Set, pp. 31-39.

CROSBY, B. P. (1980) - *Quality is free: The Art of Making Quality Certain*, New York, Mentor Book.

DAVIDOW, William H., MALONE, Michael S.(1992) – *A Corporação Virtual – Estruturação e Revitalização da Corporação para o Século 21*, São Paulo, Livraria Pioneira Edÿÿora.

DAVIS, Mark M.; AQUILANO,ÿÿÿÿholas J.; CHASE, Richard B. (2001) - *Fundamentos da Administração da Produção*, 3:ª Edição, Porto Alegre, Bookman.

Department of Industry, Science and Resources (1999) - The Australian Service Sector Review 2000, Canberra.

DJELLAL, F. y F. GALLOUJ (1999) – "Services and the Search for Relevant Innovation Indicators: A Review of National and International Surveys", *Science and Public Policy* 26/4, pp. 218-232.

DJELLAL, F. y F. GALLOUJ (2000) – "Innovation surveys for service industries: a review, ponencia presentada a la conferencia" *Innovation and Enterprise Creation: statistics and indicators*, Niza, Francia.

EUROSTAT (1995) - *Report of the Eurostat Pilot Project to Investigate the Possibilities to Measure Innovation in the Service Sectors*, Luxemburgo.

EVANS, P; WURSTER, T. (2000) – *Blown to Bits: How the economics of information transforms strategy*, Harvard Business school Press, Boston.

EVANGELISTA, R. y M. SAVONA (1998). *Patterns of Innovation in Services: The Results of the Italian Innovation Survey*. Ponencia presentada al 7th Annual RESER Conference, Berlin.

EVANGELISTA, R. y G. SIRILLI (1995) – "Measuring innovation in services", *Research Evaluation*, 5/3, pp. 207-215.

FERNANDES, J. (2000) – "Competitividade e Inovação", *Economia & Prospectiva*, n.º 10, Jul-Set, pp. 81-90.

FITZSIMMONS, James A.; FITZSIMMONS, Mona J. (2000) - *Administração de Serviços – Operações, Estratégia e Tecnologia da Informação*, 2.ª Edição, Bookman, Porto Alegre.

GADREY, J. (1995) – "Relations contrats et conventios de services", in Gadrey J., de Bandt J., *Relations de Service, Marchés de services*, CNRS, Paris

GALLOUJ, F. (1998) – "Innovation in Reverse Services and the Reverse Product Cycle", SI4S Topical Paper nº 5, STEP Group.

GIANEZI, Irineu G. N. & CORRÊA, Henrique Luiz (1994) - *Fundamentos da Administração da Produção*, Editora Atlas, São Paulo.

GRONROOS, Christian (1990) – *Service Management and Marketing: managing the moments of truth in service competition*, Lexington: Free Press, Lexington Books.

GRONROOS, Christian (1993) – *Marketing – Gerenciamento e Serviços: A Competição por Serviços na Hora da Verdade*, Editora Campus, Rio de Janeiro.

GUIMARÃES, Maria Cristina Soares (1998) – *Tecnologia como Conhecimento: o Público e o Privado*, Convénio CNPq/IBICT – UFRJ, Janeiro.

HERRIOT, Peter, PEMBERTON, Carole (1995) – *Competitive Advantage Through Diversity: Organizational Learning from difference*, London: Sage publications.

HIPP, C., B.S. TETHER e I. MILES (2000) – "The Incidence and Effects of Innovation in Services: Evidence from Germany", *International Journal of Innovation Management*, 4/4, pp. 417- 453.

HOLLENSTEIN, Heinz (2000) – *Innovation Modes in the Swiss Service Sector. A Cluster Analysis based on firm-level data*, 3rd Workshop of the focus group on Innovative firms and Networks, OECD, Project on National Innovation Systems (Phase III), Rome, October, pp. 17.

KLEINKNECHT, A. (2000) – "Indicators of Manufacturing and Service Innovation: Their Strengths and Weaknesses" en *Innovation Systems in the Service Economy: Measurement and Case Study Analysis*, Economics of Science, Technology and Innovation, vol. 18, J. S. Metcalfe e I. Miles (editores), pp. 169-186. Kluwer Academic Publisher, USA.

KOTLER, Philip (1994) – *Marketing – para organizações que não visam o lucro*, Editora Atlas, Brasil.

KOTHER, Philip (1987) – *Marketing*, Atlas, S. Paulo.

LENDREVIE, Jacques; DIONÍSIO, Pedro; RODRIGUES, Vicente (1998) - *Mercator - Teoria e Prática do Marketing*, Publicações D. Quixote, 7ª ed., Lisboa.

LEVIN, R.C., A.K. KLEVORICK, R.H. NELSON y S. WINTERS (1987) – "Appropriating the Returns from Industrial Research and Development", *Brookings Papers on Economic Activity*, 3, pp. 783-831.

MARKLUND, G. (1998) - *Need for New Measures of Innovation in Services*, SI4S Topical Paper nº 9, STEP Group.

MARKLUND, G. (2000) – "Indicators of Innovation Activities in Services" en *Services and theKnowledge-Based Economy*, M. Boden e I. Miles (editores). Science, Technology and the International Political Economy Series, Series editor: J. de la Mothe, Continuum, Londres, Gran Bretaña, pp. 86-108.

MARTÍNEZ-ROS, E. (1998) - *Innovation and Rents: Three Empirical Essays using Data from Spanish Manufacturing Firms*, Tesis Doctoral, Universidad Autónoma de Barcelona, España.

MILES, I., N. KASTRINOS, R. BILDERBEEK, P. DEN HERTOG, K. FLANAGAN y W. HUNTIK (1995) - *Knowledge-intensive business services: their role as users, carriers and sources of innovation*, Informe para el EC DG XIII Sprint EIMS Programme, Luxemburgo.

MILES, I. (1995) – "Services Innovation, Statistical and Conceptual Issues". Working Group on Innovation and Technology Policy, OCDE (Doc. DSTI/EAS/STP/NESTI (95)12, Paris, Francia.

MOREIRA, Daniel A. (2000) - *Administração da Produção e Operações*, São Paulo, Editora Pioneira, 5.ª Edição.

MURRAY, K. (1991) – "A test of services marketing theory: consumer information acquisition activities", *Journal of Marketing*, Vol. 55, January, pp. 10 - 25.

NORMANN, R. (1991) – *Service Management, Strategy and Leadership in Service Business*, New York, John wiley & Sons.

OCDE (2000a) – "Promoting Innovation and Growth in Services", *Science, Technology and Industry Outlook 2000,* OCDE, Paris.

OCDE (2000b) – "Innovation and Growth", *Science, Technology and Industry Outlook 2000*, OCDE, Paris.

OCDE (2000c) – "A New Economy?. The Changing Role of Innovation and Information Technology in Growth", *Information Society*, OCDE, Paris.

PATTISON, W., J. OVINGTON y E. FINLAY (1995) - *Innovation in selected Australian industrie*s, ponencia presentada al 10th meeting of the Voorburg Group on service statistics, Australian Bureau Statistics.

PAVITT, K. (1984) – "Sectoral Patterns of Technological Change: Towards a Taxonomy and a Theory", *Research Policy* 13, pp. 343-373.

PAVITT, K., M.; ROBSON y J. TOWNSEND (1989) – "Technological Accumulation, Diversification and Organisation in UK Companies, 1945-1983", *Management Science*, 35.1, pp. 81-99.

PILAT, D. (2000) - *Innovation and Productivity in Services: State of the Art*, Ponencia presentada al OCDE/Australia workshop on "Innovation and Productivity in Services", Sydney, Australia.

RODRIGUES, E. (1999) – *A Inovação como um Sistema quase Biológico...,* Dissertação apresentada na Universidade Católica Portuguesa com vista à obtenção do grau de Mestre em Estudos Europeus, Lisboa.

RUST, R. T.; OLIVER, R. L. (1994) - *Service Quality: Directons in theory and pratice*, Sage Publications, pp.1 - 20.

SOETE, L. y M. MIOZZO (1989) – "Trade and Development in Services: A Technological Perspective", Working Paper nº 89-031, MERIT, Maastricht.

SUNDBO, Jon; GALLOUJ, Faiz (1998a) - *Innovation in Services*. SI14S Project Synthesis, WP, pp. 3-4.

SUNDBO, Jon; GALLOUJ, Faiz (1998b) - *Innovation as a loosely coupled system in Services*. SI14S, topical Paper 04.

SUNDBO, Jon; GALLOUJ, Faiz (1999) - *Innovation in Services in seven European countries, 99:1, Synthesis report for Europeasn Commission*, DGXII, TSER-SI4S.

TIDD, J.; BESSANT, J.; PAVITT, K. (2003) – *Gestão da Inovação. Integração das Mudanças Tecnológicas de Mercado e Organizacionais*, Monitor.

VAN ARK, Bart, BROESMA, Lourens; HERTOG, den Pim (2003) – "Services Innovation, Performance Policy"**:** *A Review*, synthesis Report in the Project Structural Information Provision on Information in services, June.

ZEITHLAML, V.; PARASURAMAN, A.; BERRY, L. (1990) - *Delivering Quality Service Balancing customer perception and expectations*, Free Press, New York.

PADRÕES SECTORIAIS DE INOVAÇÃO

ANA ABRUNHOSA
E-mail: ampa@fe.uc.pt
CARLOS CARREIRA
E-mail: ccarreir@fe.uc.pt
RITA MARTINS
E-mail: rvmartin@fe.uc.pt
Faculdade de Economia
Universidade de Coimbra

Resumo

A relação entre dimensão e inovação não é consensual, é sensível a vários factores como o tipo de sector considerado e a base de conhecimento/regime tecnológico subjacente. O objectivo do presente trabalho consiste em analisar, para o caso português, a relação entre intensidade de inovação e dimensão empresarial. Começámos, numa primeira etapa, por explicar a intensidade de inovação (medida através da intensidade de I&D e da intensidade da inovação total) através de variáveis explicativas ao nível da empresa e ao nível do sector, socorrendo-nos de informação recolhida no âmbito do CIS 4 e de estatísticas da OCDE, respectivamente. Numa segunda etapa, o objectivo foi enfatizar a importância dos regimes de conhecimento para a referida relação. Finalmente, numa terceira etapa, colocou-se a tónica na influência dos regimes tecnológicos na relação acima referida. Em todas estas etapas, verificámos que a relação entre intensidade de inovação e dimensão empresarial é negativa e estatisticamente significativa.

Palavras-chave: inovação, dimensão empresarial, regime tecnológico.
Área temática: Economia da Inovação e Transferência de Tecnologia.

Abstract

The relation between dimension and innovation is not consensual, it is sensible to some factors as the type of sector and the underlying base of knowledge and technological regimen. The objective of the present work is to analyze, for the Portuguese case, the relation between innovation intensity and firms' dimension. We started by taking into consideration, besides, firm-specific variables, collected from the 4th Community Innovation Survey, and from industry specific variables, obtained from OECD database. Next, we emphasized the impact of the knowledge regimes into the abovementioned relation. Finally, we focused on the influence of the technological regimes on the relation between innovation and firm dimension. The main conclusion obtained is that there is a statistically significant negative d relation between firm size and innovation intensity.

Key words: innovative intensity, firm size, technological regimen.
Thematic Area: Innovation and Technological Transfer Economics.

1. INTRODUÇÃO

No contexto actual, os factores de competitividade não-preço assumem uma importância crescente, onde a diversidade e a qualidade dos produtos se tornou fundamental. Neste processo a capacidade de inovação das empresas é fundamental, não só para manterem os mercados onde actualmente operam, mas para poderem entrar em novos mercados.

Porque é que as empresas apresentam capacidades de inovação diferentes entre si? Para Schumpeter (1934) as empresas de maior dimensão apresentavam maiores capacidade de inovação. Ainda hoje persiste a ideia de que as grandes empresas desenvolvem mais actividades de inovação do que as pequenas empresas. Esta percepção deve-se, por um lado, ao facto de existirem proporcionalmente menos empresas de pequena dimensão a dedicarem-se a actividades de inovação comparativamente às grandes empresas, e, por outro, ao facto das grandes empresas disporem em termos absolutos de mais recursos passíveis de serem afectos a este tipo de actividade.

A discussão e os estudos em torno desta temática têm sido múltiplos e com resultados nem sempre coincidentes.

O objectivo do presente trabalho é analisar, para o caso português, e ao nível da empresa, se existe uma associação positiva entre intensidade de inovação e dimensão das empresas. Começámos, numa primeira etapa, por explicar a intensidade de inovação (medida através da intensidade de I&D e da intensidade da inovação total) através de variáveis explicativas ao nível da empresa e ao nível do sector, socorrendo-nos de informação recolhida no âmbito do CIS 4 e de estatísticas da OCDE, respectivamente. Numa segunda etapa, o objectivo foi enfatizar a importância dos regimes de conhecimento para a referida relação. Finalmente, numa terceira etapa, colocou-se a tónica na influência dos regimes tecnológicos na relação acima referida.

Para as variáveis ao nível da empresa recorreu-se à base de dados resultante do 4º Inquérito Comunitário à Inovação feito às empresas portuguesas. As variáveis caracterizadoras dos sectores de actividade foram obtidas a partir da base de dados da OCDE.

A estrutura do trabalho é a seguinte. Após a presente Introdução, no segundo Capítulo é feito um breve enquadramento teórico do tema em análise. O terceiro Capítulo é dedicado à exposição da metodologia, dados e das opções seguidas para dar resposta ao objectivo desta investigação. O quarto Capítulo é dedicado à análise e discussão dos resultados, dedicando uma secção a cada uma das etapas definidas anteriormente. Finalmente, no quinto e último Capítulo são apresentadas as principais conclusões.

Este estudo visa contribuir para uma discussão que está longe de estar encerrada. Para além disso, em Portugal, esta temática está pouco explorada, pelo que, a este nível, se pretende dar um tributo empírico para a discussão. Nesta fase, este estudo está numa fase ainda de exploração dos resultados, pretendendo-se, de futuro, aprofundar as razões de alguns dos resultados obtidos.

2. REVISÃO DA LITERATURA

Schumpeter (1942) enfatizou a ideia de que as grandes empresas apresentam uma importância crescente na promoção de mudanças tecnológicas nas modernas economias industrializadas. Desde então, tem-se vindo a discutir/estudar a existência de uma associação positiva entre dimensão empresarial e capacidade de inovação. No entanto, não existe consenso na literatura empírica sobre se são realmente as grandes empresas as que têm mais actividades de inovação. A relação entre dimensão e inovação é sensível a vários factores específicos ao sector de actividade onde as empresas desenvolvem as suas actividades de inovação, como por exemplo, a dimensão do sector, a concentração de mercado, a taxa de rotação e a base de conhecimento subjacente, entre outros. Neste contexto, há uma corrente de autores que consideram que o regime tecnológico do sector de actividade onde a empresa opera pode influenciar a relação entre inovação e dimensão (Mazzucato, 2000; Acs *et al.* 1990, entre outros). Um regime tecnológico reflecte as propriedades do processo de inovação específicas ao

sector de actividade, bem como a natureza do conhecimento científico e tecnológico em que se baseia o processo de inovação nas empresas. Um regime tecnológico é definido a) pelas oportunidades tecnológicas e condições da sua apropriação, b) pela acumulação de conhecimento resultante da aprendizagem e c) pela base de conhecimento subjacente ao processo de inovação (Dosi e Orsenigo, 1988; Dosi *et al*., 1995).

Nelson e Winter (1982) identificam dois regimes tecnológicos (schumpeterianos), o rotineiro e o empreendedor. No regime rotineiro, "as empresas acumulam a sua base de conhecimentos tecnológicos a partir de fontes internas – daí ser também designado por *regime cumulativo*. O conhecimento é cumulativamente construído através de processos de aprendizagem internos às empresas. Como a acumulação de conhecimentos pelas empresas estabelecidas gera elevadas barreiras (tecnológicas) à entrada, este regime potencia a inovação das (grandes) empresas em actividade, limitando as oportunidades dos agentes empresariais exteriores à indústria" (Carreira, 2004: 23). O regime empreendedor, "caracteriza-se por um conhecimento tecnológico baseado na ciência e, consequentemente, de base alargada e universal, geralmente com origem em programas de investigação públicos (por exemplo, centros de investigação ou universidades). Cada empresa acumula assim a sua base de conhecimentos tecnológicos a partir de fontes externas – este regime é por isso também designado de *não-cumulativo* (em relação à base de conhecimentos da empresa). O carácter universal da base de conhecimentos potencia a entrada de novas empresas inovadoras, desfavorecendo a actividade inovadora das empresas em actividade. Este tipo de indústria é assim caracterizado pela existência de uma elevada taxa de entrada e de um grande número de pequenas empresas" (*idem*, *ibidem*).

Em regimes empreendedores as empresas de pequena dimensão tendem a ser mais inovadoras e nos regimes rotineiros são as grandes empresas que demonstram maior capacidade de inovação (Acs *et. al*, 1990).

Assim sendo, para se analisar a relação entre dimensão e intensidade de inovação deve ter-se em conta o regime tecnológico em que a empresa se insere, ou seja, as características específicas dos sectores de actividade onde as empresas operam, bem como os regimes de conhecimento subjacentes.

Marsili (2001) procurou analisar o efeito dos regimes de conhecimento na relação entre inovação e dimensão, considerando as diferenças na natureza das bases de conhecimento subjacentes aos processos inovadores. A base de conhecimento de uma empresa pode ser caracterizada em termos da combinação específica de "corporate competencies in different technological fields, defined by fairly stable, universal and operational systems of classification" (Granstrand, Patel e Pavitt, 1997: 9).

Marsili (2001: pág 93 e segs.) faz uma classificação dos regimes de conhecimento em 5 grupos (veja-se Tabela 1 do Anexo), sendo eles: regime baseado na ciência, processos fundamentais, sistema de conhecimento complexo, regime baseado em engenharia de produto e por último regimes de processos contínuos.

O regime baseado na ciência distingue-se pelas elevadas oportunidades tecnológicas, elevadas barreiras tecnológicas à entrada em termos de conhecimento e de escala e pela persistência de inovação. As empresas são caracterizadas por um baixo grau de diversidade de taxas e direcções das inovações e da base do conhecimento. O contributo da investigação académica é bastante importante e directo. As actividades de inovação são mais viradas para a inovação de produto.

O regime de processos fundamentais apresenta características semelhantes ao anterior, mas com relativamente baixos níveis de oportunidades tecnológicas e inputs provenientes da investigação académica e de outras instituições públicas. Neste regime a inovação é fundamentalmente de processo.

O regime sistema de conhecimento tecnológico é caracterizado por um nível médio-alto de oportunidades, barreiras à entrada em conhecimento e de escala e persistência na inovação. Uma característica que distingue este regime é o elevado grau de diferenciação das competências desenvolvidas pelas empresas, especialmente nas tecnologias das primeiras etapas de produção

e também por fontes externas de conhecimentos, incluindo uma importante, no entanto indirecta, contribuição de investigação académica.

O regime baseado em engenharia de produto caracteriza-se por grandes oportunidades tecnológicas, baixas barreiras à entrada e pela não persistência na inovação. As empresas são muito heterogéneas nas suas taxas e direcções das inovações.

O regime baseado em processos contínuos caracteriza-se pela baixa oportunidade, baixas barreiras à entrada e por baixa persistência na inovação.

A base de conhecimento das empresas reflecte a sua capacidade para explorar diferentes áreas de oportunidades tecnológicas em mercados de produtos e em tecnologias de produção relacionados. Áreas de elevadas oportunidades tecnológicas para a inovação estão associadas a sectores com elevada intensidade de I&D, elevadas barreiras à entrada no conhecimento e elevadas vantagens para a inovação associadas à escala, elevada persistência da inovação e ligações directas com a investigação académica. Contudo, elevadas oportunidades de inovação também podem coexistir com reduzidas barreiras tecnológicas. Este é o caso dos sectores pertencentes ao regime engenharia do produto,

Com o intuito de analisar as relações entre intensidade de inovação e dimensão empresarial, a maioria dos estudos (Evangelista *et al.*, 1997; Evangelista, 1996, entre outros) utiliza como unidade de análise a empresa. Em geral, são realizadas análises seccionais. Em alguns estudos empíricos os recursos financeiros e humanos aplicados em actividades de I&D foram utilizados como *proxy* da capacidade de inovação para analisar se a hipótese de que a intensidade de inovação cresce com a dimensão do sector de actividade, se verifica. Utilização essa que veio a ser criticada por não reflectir as actividades tecnológicas que são levadas a cabo pelas pequenas e médias empresas, e que são bastante importantes no âmbito de actividades de inovação (Archibugi *et al*., 1995). Se considerarmos que a inovação corresponde à produção de conhecimento novo ou à recombinação de conhecimento já existente para desenvolver e criar produtos e/ou processos, novos ou significativamente melhorados, com valor para o mercado, verificamos que nem toda a I&D dá, de facto, origem a inovação. Por outro lado, verifica-se que, em muitas empresas, o processo de inovação não resulta necessariamente das actividades de I&D, ou seja, da produção de conhecimento novo, resultando, antes, da recombinação de conhecimento já existente (Abrunhosa, 2006). Assim, para além da I&D, é importante considerar outras actividades promotoras de inovação, como as actividades de marketing, aquisição conhecimentos, entre outras.

3. METODOLOGIA E DADOS

3.1. Introdução

O objectivo do presente trabalho é responder à seguinte questão: *Que relação existe entre inovação e dimensão empresarial?*

Dada a complexidade de factores que podem condicionar esta relação, procurámos responder à questão em três etapas. Na primeira delas procurámos analisar a referida relação, integrando informação sobre os sectores onde as empresas operam. Na segunda etapa considerámos os regimes de conhecimento a que cada empresa pertence. Finalmente, numa terceira etapa tivemos em consideração o respectivo regime tecnológico.

Os dados utilizados no presente estudo provêm da recolha obtida a partir do 4º Inquérito Comunitário à Inovação (Community Innovation Survey 4 - CIS 4) administrado pelo Observatório da Ciência e do Ensino Superior (OCES) e referente ao período de 2002 a 2004.

Este inquérito recolhe informação sobre as actividades relacionadas com a inovação de produtos e processos e, pela primeira vez, com as actividades de inovação organizacional e de marketing desenvolvidas em 2002, 2003 e 2004. A maioria das questões refere-se a bens e serviços, bem como à implementação de processos e de métodos de logística e de distribuição, novos ou significativamente melhorados. No que se refere à dimensão empresarial é recolhida informação para 2004.

A base inicial contém informação relativa a 4809 empresas nacionais, de 37 sectores de actividade económica.

O CIS 4 considera como actividades orientadas para a inovação *a investigação & desenvolvimento (I&D) intramuros e extramuros*; a *aquisição de maquinaria, equipamento e software*; a *aquisição de outros conhecimentos externos*; e *actividades de marketing*.

A *I&D intramuros* diz respeito a «trabalhos criativos empreendidos dentro da empresa com vista a aumentar o conhecimento, bem como a utilização desse conhecimento em novas aplicações, tais como produtos (bens/serviços) ou processos novos ou significativamente melhorados» (GPEARI/MCTES, 2004: 9). A *I&D extramuros* respeita às actividades atrás mencionadas mas «executadas por outras empresas (incluindo outras empresas do grupo), ou por instituições de I&D públicas ou privadas e adquiridas pela empresa» (GPEARI/MCTES, 2004: 9). A *aquisição de maquinaria, equipamento e software* compreende a «aquisição de maquinaria avançada, equipamento, hardware e software especificamente para produzir bens ou serviços novos ou significativamente melhorados ou implementar processos novos ou significativamente melhorados» (GPEARI/MCTES, 2004: 9). A *aquisição de outros conhecimentos externos* compreende a «compra ou licenciamento dos direitos de patentes e/ou invenções não patenteadas, "*know-how*" ou outras formas de conhecimento, a outras empresas ou instituições» (GPEARI/MCTES, 2004: 9). O *marketing* refere-se a «actividades orientadas para a introdução no mercado dos bens e serviços novos ou significativamente melhorados, incluindo estudos de mercado, testes de mercado e publicidade de lançamento» (GPEARI/MCTES, 2004: 9).

No CIS 4 questionam-se as empresas se, durante o período de 2002 a 2004, estiveram envolvidas em alguma das actividades orientadas para a inovação anteriormente referidas, e, para além disso, qual o montante estimado de despesa, para o ano 2004, nas seguintes quatro actividades de inovação: *I&D intramuros*, *I&D extramuros*, *aquisição de maquinaria, equipamento e software* e *aquisição de outros conhecimentos externos*. O somatório destes valores corresponde à nossa variável *inovação total*.

Face ao objectivo da investigação, foi aplicado um primeiro filtro à base inicial tendo-se excluído as empresas que não declararam ter realizado despesa em pelo menos uma das actividades de inovação atrás referidas. Após este procedimento, a base ficou reduzida a 1609 empresas.

Uma ideia comum é a de que as grandes empresas desenvolvem mais actividades de inovação do que as pequenas empresas. Esta percepção deve-se, por um lado, ao facto de existirem proporcionalmente menos empresas de pequena dimensão a dedicarem-se a actividades de inovação comparativamente às grandes empresas, e, por outro, ao facto das grandes empresas disporem em termos absolutos de mais recursos passíveis de serem afectos a este tipo de actividade.

Os dados recolhidos via CIS 4 confirmam a percepção anteriormente referida, conforme se pode verificar pelos valores inscritos na tabela 3.1.

Tabela 3.1. Empresas com actividades de inovação, por dimensão, em Portugal (2002-2004)

Class size	Number of firms	Innovative activities (share of firms)	
(number of employees)	in the population	With (%)	Without (%)
10-49	18.000	36	64
50-249	4.590	56	44
250 or more	757	63	37

Fonte: Extraído de Gonçalves *et al.* (2006: 5).

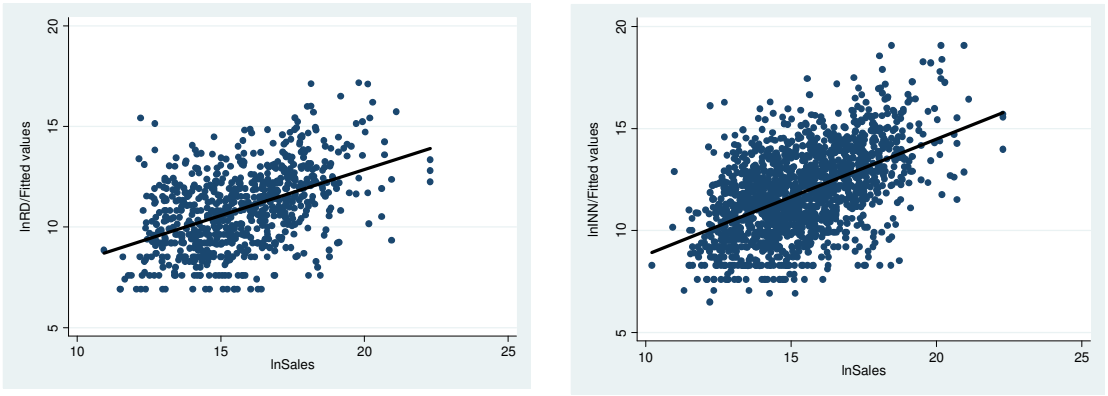
De facto, entre as empresas de pequena dimensão, apenas 36% realizam actividades de inovação, ao passo que, para o universo das grandes empresas, o mesmo indicador sobe para 63%.

No que se refere à relação entre despesas realizadas em actividades de inovação, em termos absolutos, e dimensão, para o universo das 1609 empresas que declararam desenvolver

actividades de inovação, a Figura 3.1 também confirma a ideia comum de que há uma relação directa entre dimensão e inovação.

Figura 3.1. Relação entre despesas de I&D e volume de negócios e entre despesas em inovação total e volume de negócios, por empresa (valores logaritmizados)

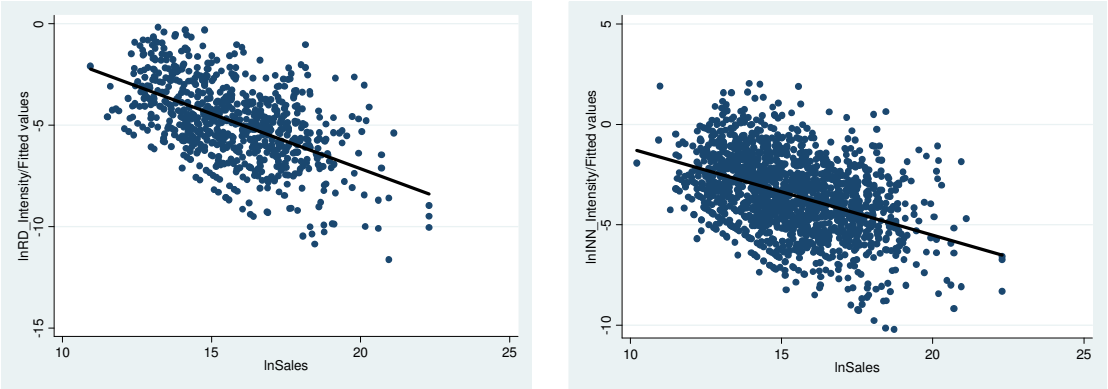
(a) Log of R&D *(a) Log of innovation*



Contudo, quando em vez das despesas em valor absoluto consideramos a intensidade de despesas de I&D (despesas de I&D / volume de negócios) e a intensidade de inovação total (despesas totais de inovação / volume de negócios) a relação de cada uma destas variáveis com a variável dimensão considerada (volume de negócios), parece revelar uma tendência negativa (conforme ilustrado pela Figura 3.2).

Figura 3.2. Relação entre intensidade de I&D e entre intensidade de inovação e volume de negócios por empresa (valores logaritmizados)

(a) Log of R&D intensity *(a) Log of innovation intensity*



A forte dispersão em torno da tendência quer no gráfico (*a*) quer em (*b*), da Figura 3.2, indicia a importância de outros factores na definição do padrão de inovação das empresas.

A Tabela 3.2 apresenta um sumário de estatísticas descritivas das variáveis utilizadas, ao nível da empresa e do sector, bem como a respectiva descrição.

Tabela 3.2. Lista de variáveis e estatísticas descritivas

Variable	Description	No. Obs.	Mean	Standard deviation
(a) At firm-level				
R&D	Total of R&D expenditures (10^3 €)	1602	192,7	1386,2
Innovation	Total of expenditures in R&D, licenses, patents, machinery and equipments (10^3 €)	1602	1568,1	10200,0
R&D intensity	R&D expenditures to sales ratio	1602	0,018	0,068
Innovation intensity	Innovation expenditures to sales ratio	1602	0,161	0,518
Size	Total of sales (10^6 €)	1602	37,2	229,0
Marketing (dummy)	1: marketing investments; 0: otherwise	1602	0,440	0,497
Equipment (dummy)	1: machinery and equipments investments; 0: otherwise	1602	0,880	0,325
Services (dummy)	1: services sector; 0: otherwise	1602	0,422	0,494
Regime (dummies)				
RC1	1: knowledge regime 1; 0: otherwise	1602	0,091	0,288
RC2	1: knowledge regime 2; 0: otherwise	1602	0,069	0,253
RC3	1: knowledge regime 3; 0: otherwise	1602	0,029	0,169
RC4	1: knowledge regime 4; 0: otherwise	1602	0,137	0,344
RC5	1: knowledge regime 5; 0: otherwise	1602	0,673	0,469
REmp	1: technological regime empr; 0: otherwise	1602	0,228	0,420
RRot	1: technological regime rot; 0: otherwise	1602	0,772	0,420
(b) At industry-level				
Size	Total of production (10^6 €)	1569	7800,0	6270,0
Concentration	C5 concentration ratio	1569	0,212	0,200
Churn rate	Sum of entry and exit rate	1601	0,263	0,989
Export share	Exports to production ratio	959	0,330	0,269

Source: CIS4 and OECD database (http://stats.oecd.org/wbos/Index.aspx)

O número máximo de observações para cada variável é de 1602, resultante da exclusão de 7 empresas da base, cuja intensidade de I&D apresentou valores superiores à unidade, pelo que foram consideradas outliers.

Os dados ao nível do sector foram recolhidos da base de dados da OECD (http://stats.oecd.org/wbos/Index.aspx), que não possui informação para todas as NACE consideradas no inquérito CIS 4. Por este motivo o número de observações ao nível do sector é inferior.

A informação sobre a variável *marketing* foi recolhida sob a forma de variável binária, com 1 para o caso em que a empresa declara ter realizado despesas de marketing e 0 para o caso contrário. Por este facto, apenas foi possível considerá-la como variável dummy.

Já no caso da dummy *equipamento*, a sua inclusão pretende corrigir algum efeito de sobreavaliação da variável *inovação total*, que decorre da tendência para se considerar como inovação qualquer aquisição de maquinaria, equipamento e software.

Considerou-se ainda uma variável binária que assume o valor 1 no caso da empresa pertencer ao sector dos serviços. Ou seja, assume-se a hipótese da existência de um comportamento diferenciado entre as empresas do sector de serviços e de outros sectores.

Ao nível do sector considerámos três potenciais tipos de efeitos da envolvente: dimensão, dinâmica de concorrência e grau de abertura. No que respeita à dimensão do sector, utilizou-se como *proxy* a variável produção total da NACE a que a empresa pertence. Para a dinâmica de concorrência, considerámos dois tipos de variáveis: a *ratio* de concentração (C5) e a taxa de rotação das empresas. A primeira variável corresponde à soma das cinco maiores quotas de mercado. A segunda é a soma da taxa de entrada com a taxa de saída em cada sector. Em mercados muito concentrados a concorrência é menor, pelo que a pressão das empresas para inovar pode ser também menor. Apesar disto, se a taxa de rotação das empresas for elevada, a concorrência potencial é maior, pelo que poderá haver uma maior pressão para inovar. Finalmente, considerámos como *proxy* para o grau de abertura, a variável quota de exportações.

Na Tabela 3.3. apresentam-se as correlações entre as diversas variáveis.

Tabela 3.3. Matriz das correlações entre as variáveis

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]
[1] R&D_Int	1,000																
[2] Inov_Int	0,468	1,000															
[3] Firm Size	-0,025	-0,036	1,000														
[4] Dummy Markt	0,070	0,010	0,054	1,000													
[5] Dummy Equipment	0,008	0,102	-0,011	-0,047	1,000												
[6] Ind_SIZE	-0,001	-0,010	0,006	0,014	0,060	1,000											
[7] C5	0,062	-0,015	0,222	0,133	-0,136	-0,264	1,000										
[8] Churn_Rate	0,069	0,008	-0,079	-0,066	0,008	0,068	-0,227	1,000									
[9] Export Share	0,034	-0,011	-0,030	-0,029	-0,091	-0,757	0,270	0,042	1,000								
[10] Dummy Servces	0,052	0,033	-0,052	-0,032	0,047	0,753	-0,302	0,502	-0,636	1,000							
[11] RC1	0,084	-0,015	0,009	0,092	-0,093	-0,274	0,365	0,375	0,230	-0,130	1,000						
[12] RC2	0,012	-0,050	0,065	0,130	-0,133	0,072	0,549	-0,497	-0,014	-0,136	-0,066	1,000					
[13] RC3	0,040	0,004	-0,007	-0,045	-0,032	-0,163	0,060	0,020	0,464	-0,117	-0,057	-0,060	1,000				
[14] RC4	-0,042	0,039	-0,035	0,036	-0,044	-0,387	-0,118	-0,026	0,144	-0,263	-0,128	-0,134	-0,115	1,000			
[15] RC5	-0,030	-0,002	-0,005	-0,120	0,163	0,492	-0,384	0,083	-0,432	0,403	-0,322	-0,337	-0,291	-0,652	1,000		
[16] RRot	-0,007	-0,028	0,027	-0,082	0,090	0,501	-0,087	-0,177	-0,255	0,310	-0,419	0,158	0,136	-0,847	0,769	1,000	
[17] REmpr	0,007	0,028	-0,027	0,082	-0,090	-0,501	0,087	0,177	0,255	-0,310	0,419	-0,158	-0,136	0,847	-0,769	-1,000	1,000

Da informação apresentada destaca-se a correlação negativa entre a dimensão empresarial e intensidade de inovação, conforme referido anteriormente a propósito da tendência evidenciada na Figura 3.2.

Numa primeira etapa analisámos as determinantes da intensidade de I&D e de inovação, considerando, para além da variável interna dimensão das empresas, um conjunto de variáveis, atrás referidas, que caracterizam a envolvente.

Com o objectivo de averiguar a natureza da associação entre intensidade de inovação e dimensão empresarial para o caso português, foram estimada as seguintes equações (1) e (2) pelo Método dos Mínimos Quadrados (OLS), em que todas as variáveis foram logaritmizadas, à excepção das variáveis *dummy*.

Atendendo a que nestas formulações a constante não tem um significado económico, optámos pela não inclusão da mesma em todas as regressões.

$$\log R\&DInt_i = \beta_1 \log Firmsize_i + \beta_2 \log Indsize_i + \beta_3 \log CR5_i + \beta_4 \log ChurnRate_i + \beta_5 \log ExpSh_i + \beta_6 Dservices_i + \varepsilon_i \quad (1)$$

$$\log InovInt_i = \beta_1 \log Firmsize_i + \beta_2 DMark_i + \beta_3 DEquip_i + \beta_4 \log Indsize_i + \beta_5 \log CR5_i + \beta_6 \log ChurnRate_i + \quad (2)$$
$$+ \beta_7 \log ExpSh_i + \beta_8 Dservices_i + \mu_i$$

sendo ε e μ os usuais termos de erro e o índice *i* relativo à empresa.

A partir desta fase, e uma vez que se passa a trabalhar com grupos com características relativas a sectores de que fazem parte as empresas deixa de se usar, naturalmente, as variáveis explicativas ao nível do sector (produção total da NACE; a *ratio* de concentração (C5); a taxa de rotação das empresas; e a quota de exportações).

Numa segunda etapa analisámos as determinantes da intensidade de I&D e de inovação, considerando, para além da variável interna dimensão das empresas, os regimes de conhecimento a que cada empresa pertence. Para tal, previamente à estimação adoptou-se a metodologia desenvolvida por Marsili (2001), tendo-se classificado os sectores de actividade de acordo com o seu regime de conhecimento, da forma apresentada na Tabela 3.4.

Embora esta classificação seja feita para sectores industriais, aplicámo-la, com adaptações, ao sector dos serviços considerando factores possíveis de serem aplicados, como por exemplo, barreiras tecnológicas à entrada, persistência na inovação e fontes de conhecimentos externos.

Tabela 3.4: Regimes de conhecimento para os vários sectores de actividade[♣]

Regimes de Conhecimento	*Designação*	*Código NACE (a dois dígitos)*
Baseado nas ciências	RC1	NACE_32; NACE_33; NACE_64; NACE_72_73
Processos fundamentais	RC2	NACE_11_14; NACE_23_24; NACE_37
Sistemas complexos	RC3	NACE_34; NACE_35;
Engenharia de produto	RC4	NACE_22; NACE_25; NACE_26; NACE_29; NACE_30_31; NACE_65; NACE_67
Processos contínuos	RC5	NACE_15_16; NACE_17; NACE_18; NACE_19; NACE_20; NACE_21; NACE_27; NACE_28; NACE_36; NACE_40; NACE_41; NACE_45; NACE_51; NACE_52; NACE_55; NACE_60; NACE_61; NACE_62; NACE_63; NACE_66; NACE_74

♣ A designação detalhada dos correspondentes sectores de actividade, encontra-se na Tabela 2 em Anexo.

Cada regime de conhecimento corresponde a uma *dummy*, *RC*, pela mesma ordem que é apresentada na Tabela 3.4, nos modelos apresentados nas equações 3 e 4.

$$\log R\,\&\,DInt_i = \beta_1 \log Firmsize_i + \beta_2 RC_{1i} + \beta_3 RC_{2i} + \beta_4 RC_{3i} + \beta_5 RC_{4i} + \beta_5 RC_{5i} + \delta_i \quad (3)$$

$$\log InovInt_i = \beta_1 \log Firmsize_i + \beta_2 DMark_i + \beta_3 DEquip_i + \beta_4 RC_{1i} + \beta_5 RC_{2i} + \beta_6 RC_{3i} + \beta_7 RC_{4i} + \beta_8 RC_{5i} + \eta_i \quad (4)$$

sendo δ e η os usuais termos de erro e o índice *i* relativo à empresa.

Após as estimações com as dummies relativas aos regimes de conhecimento, procedemos à análise do mesmo fenómeno mas regime de conhecimento a regime de conhecimento. Ou seja, estimaram-se 5 de cada uma equações seguintes, uma por cada regime de conhecimento.

$$\log R\,\&\,DInt_i = \beta_1 \log Firmsize_i + \delta_i \quad (3.1)$$

$$\log InovInt_i = \beta_1 \log Firmsize_i + \beta_2 DMark_i + \beta_3 DEquip_i + \eta_i \quad (4.1)$$

Finalmente, numa terceira etapa procurámos estudar a natureza da associação entre intensidade de I&D e intensidade de inovação e dimensão das empresas, considerando os regimes tecnológicos do sector a que a empresa pertence.

Primeiramente foi feita uma estimação considerando todas as empresas da base, distinguindo-as consoante se considere que pertencem ao regime tecnológico empreendedor ou ao regime tecnológico rotineiro.

Consideraram-se empresas do regime empreendedor as empresas que fazem parte dos regimes de conhecimento 1 (baseado na ciência) ou 4 (baseado na engenharia de produto) e que, simultaneamente, se integrem em sectores com elevada mobilidade e baixa concentração (variáveis relativas à dinâmica da concorrência). A justificação subjacente a esta opção de trabalho prende-se com a ideia de que do regime empreendedor fazem parte muitas pequenas empresas que entram e saem do sector com alguma facilidade. Assim, impôs-se, para além da condição relativa ao regime de conhecimento, características específicas do sector, ou seja, que as empresas pertencessem a sectores cujas taxas de rotação fossem iguais ou superiores a 15,1[1] ou a sectores cujo índice CR5 fosse inferior a 2/3.

As empresas consideradas como fazendo parte do regime rotineiro, são todas as restantes. Ou seja, são as empresas que se classificaram nos regimes de conhecimento 2, 3 e 5 e, para além

[1] Valor correspondente ao decil inferior da distribuição da variável taxa de rotação.

destas, as que pertencendo aos regimes de conhecimento 1 e 4 se integram em sectores cujas variáveis taxa de rotação e índice de concentração CR5, assumem valores inferiores a 15,1 e maiores ou iguais a 2/3, respectivamente.

Nas primeiras regressões desta etapa, cada regime tecnológico corresponde, assim, a uma *dummy*, em que *REmpr* toma o valor 1 caso se considere que a empresa se insere no regime tecnológico empreendedor e 0 em caso contrário, e com *RRot* a assumir o valor 1 para as empresas consideradas pertencentes ao regime tecnológico rotineiro.

Os modelos estudados encontraram-se representados pelas equações 5 e 6.

$$\log R\,\&\,DInt_i = \beta_1 \log Firmsize_i + \beta_2 REmpr_{1i} + \beta_3 RRot_{2i} + \gamma_i \quad (5)$$

$$\log InovInt_i = \beta_1 \log Firmsize_i + \beta_2 DMark_i + \beta_3 DEquip_i + \beta_4 REmpr_{1i} + \beta_3 RRot_{2i} + \lambda_i \quad (6)$$

sendo γ e λ os usuais termos de erro e o índice *i* relativo à empresa.

Num momento posterior da análise, e com a intenção de melhor explorar a relação entre dimensão empresarial e intensidade de inovação, procedemos a uma partição da base, separando as empresas consideradas pertencentes ao regime empreendedor (366 empresas) das empresas que se integrem no regime rotineiro (1236 empresas). A partir daqui repetiram-se as regressões para cada regime tecnológico em separado.

As equações (7) e (8) foram estimadas, portanto, duas vezes cada, uma para o grupo de empresas empreendedoras e uma para o grupo de empresas rotineiras, com o intuito de discernir eventuais diferenças na relação entre dimensão e intensidade de inovação, entre os grupos referidos.

$$\log R\,\&\,DInt_i = \beta_1 \log Firmsize_i + \gamma_i \quad (5.1)$$

$$\log InovInt_i = \beta_1 \log Firmsize_i + \beta_2 DMark_i + \beta_3 DEquip_i + \lambda_i \quad (6.1)$$

4. ANÁLISE E DISCUSSÃO DE RESULTADOS

4. 1. Introdução

Nesta secção analisamos os resultados das estimações enunciadas anteriormente e de acordo com as etapas descritas na metodologia.

4. 2. Influência da dinâmica competitiva do sector na relação entre inovação e dimensão empresarial

Uma vez que se verificou, a partir do teste de White, que os modelos expressos pelas equações (1) e (2) eram heterocedásticos, isto é, que a variância do termo de erro não era constante, procedeu-se à correcção da heterocedasticidade a partir da metodologia sugerida por Davidson and MacKinnon (1993)[2].

Os resultados da estimação das equações (1) e (2) pelo Método dos Mínimos Quadrados encontram-se na Tabela 4.1.

As variáveis foram introduzidas na forma de logaritmo, excepto as variáveis binárias, pelo que a leitura dos coeficientes poderá ser feita directamente sob a forma de elasticidade. Foram feitos testes de significância conjunta e individual das variáveis.

[2] Este tipo de procedimento foi sempre seguido em qualquer das estimações realizadas.

Tabela 4.1. Intensidade de I&D e de inovação e características do sector I

Variables	(1) R&D intensity	(2) Innovation intensity
Firm's size	-0.447*** (0.046)	-0.379*** (0.032)
Marketing (dummy)		0.450*** (0.110)
Equipment (dummy)		1.160*** (0.162)
Industry size	0.167*** (0.039)	0.103*** (0.030)
Industry concentration	0.299*** (0.098)	0.232*** (0.074)
Churn rate	0.435**(0.192)	0.377** (0.155)
Industry export share	0.215** (0.094)	0.001 (0.071)
Sector of services (dummy)	1.826* (0.944)	-0.294 (0.686)
Number of observations	453	959
F test	699.05***	515.25***
R-squared	0.903	0.812

Notes: Independent and dependent variables are all measured in logarithmic form, except dummy variables. Robust standard errors are given in parenthesis. ***, **, and * denote statistical significance at the .01, .05, and .10 levels, respectively.

De uma forma geral, todas as variáveis apresentam significância estatística e o R^2 revela uma boa performance nos dois modelos.

Um aumento de 1% na dimensão das empresas leva a uma redução da intensidade de I&D e da intensidade de inovação de 0,45% e 0,38%, respectivamente, *ceteris paribus*.

Este resultado contraria a hipótese convencional de Schumpeter, segundo a qual grandes escalas produtivas influenciam positivamente a intensidade de inovação.

No entanto, está em sintonia com outros estudos empíricos que vieram suscitar dúvidas a este propósito. Soete (1979), por exemplo, concluiu que a intensidade de I&D das empresas aumenta com a dimensão da empresa, até um certo patamar, a partir do qual a relação se inverte. O mesmo autor destaca ainda que este tipo de relação pode variar de intensidade e de sinal, consoante o tipo de sector.

Para além destes aspectos há que sublinhar que o nosso estudo foi conduzido com dados ao nível da empresa mas que, em simultâneo, considerou o efeito de factores que caracterizam o sector em que a mesma se insere.

De igual modo, Cohen, Levin e Mowery (1987) analisam a relação entre dimensão empresarial e intensidade de inovação, ao nível da empresa, considerando efeitos das condições do sector de actividade, ao nível da oportunidade tecnológica, apropriação do conhecimento e procura de mercado. Todavia, os seus resultados mostraram que a intensidade de I&D não era estatisticamente influenciada pela sua dimensão.

Conforme era expectável, quer a dimensão sectorial, quer a dinâmica competitiva (concentração de mercado e taxa de rotação) revelam uma relação directa e estatisticamente significativa com a intensidade de inovação, embora essa relação seja maior no primeiro modelo.

Em relação à variável *dummy* marketing, verifica-se uma relação directa e estatisticamente significativa com a intensidade de inovação, o que significa que se as empresas realizam despesas de marketing associadas a actividades de inovação então a sua intensidade de inovação tende a ser maior. Este resultado está em consonância com os estudos empíricos realizados neste área.

A dummy *equipamento*, encontra-se directamente correlacionada com a intensidade de inovação, sendo essa relação estatisticamente significativa, o que nos dá uma ideia da sobreavaliação da variável *inovação total*, nos caso das empresas que declara realizar este tipo de despesa. Como foi referido anteriormente, este tipo de situação decorre da tendência para ser considerado como inovação qualquer aquisição de maquinaria, equipamento e software.

As variáveis taxa de abertura e a dummy relativa ao sector dos serviços são estatisticamente significativas, sendo essa relação positiva com a intensidade de I&D. No entanto, e contrariamente ao que seria de esperar, as relações não são estatisticamente significativas com a variável intensidade de inovação.

4. 3. Influência do regime de conhecimento do sector na relação entre inovação e dimensão empresarial

Na presente secção pretende-se, analisar os resultados da segunda etapa do nosso estudo.

Os resultados da estimação das equações (3) e (4), pelo Método dos Mínimos Quadrados e após a correcção da heterocedasticidade conforme descrito anteriormente, encontram-se na Tabela 4.2.

Tabela 4.2. Intensidade de I&D e de inovação e regimes de conhecimento

Variables	(3) R&D intensity	(4) Innovation intensity
Firm's size	-0.465*** (0.036)	-0.392*** (0.023)
Marketing (dummy)		0.267*** (0.088)
Equipment (dummy)		1.018*** (0.133)
Knowledge regime 1	3.427*** (0.531)	2.162*** (0.391)
Knowledge regime 2	2.483*** (0.619)	1.472*** (0.426)
Knowledge regime 3	2.653*** (0.667)	1.727*** (0.475)
Knowledge regime 4	2.667*** (0.592)	1.718*** (0.400)
Knowledge regime 5	2.148*** (0.570)	1.312*** (0.387)
Number of observations	730	1602
F test	1095.28***	886.74***
R-squared	0.9023	0.8192

Notes: Independent and dependent variables are all measured in logarithmic form, except dummy variables. Robust standard errors are given in parenthesis. ***, **, and * denote statistical significance at the .01, .05, and .10 levels, respectively.

Todas as variáveis apresentam significância estatística nos dois modelos, mantendo-se a relação negativa entre dimensão empresarial e intensidade de I&D e intensidade de inovação.

É de destacar que as empresas que pertencem ao regime de conhecimento 1, investem mais 1,3% das vendas em I&D do que aquelas que pertencem ao regime de conhecimento 5, não sendo esta diferença tão acentuada quando a variável de inovação em explicação é a intensidade de inovação.

Após esta análise global, procedeu-se à estimação das 5 equações de cada um dos tipos (3.1) e (4.1), uma para cada regime de conhecimento, pelo Método dos Mínimos Quadrados, cujos resultados se encontram na Tabela 4.3.

Tabela 4.3. Intensidade de I&D e de inovação por regimes de conhecimento

	RC 1 R&D intensity	RC 2 R&D intensity	RC 3 R&D intensity	RC 4 R&D intensity	RC 5 R&D intensity
Firm's size	-0.241***(0.013)	-0.312***(0.015)	-0.302***(0.024)	-0.299***(0.010)	-0.332***(0.005)
Number of observations	101	55	28	98	447
F test	330***	413***	155.170***	885.300***	4561.500***
R-squared	0.805	0.893	0.864	0.901	0.912
	Inov intensity	Inov intensity	Inov intensity	Inov intensity	Inov intensity
Firm's size	-0,229***(0,027)	-0,278***(0,020)	-0,317***(0,059)	-0,345***(0,020)	-0,319***(0,010)
Mark dummy	0,541* (0,541)	0,452(0,365)	0,716(0,572)	0,538**(0,233)	0,093 (0,107)
Equip dummy	0,495 (0,495)	0,503*(0,301)	1,481(0,962)	2,038***(0,309)	1,292***(0,153)
Number of observations	146	110	47	220	1078
F test	115,9***	217,5***	42,800***	288,460***	1756,41***
R-squared	0,734	0,812	0,776	0,802	0,833

Notes: Independent and dependent variables are all measured in logarithmic form, except dummy variables. Robust standard errors are given in parenthesis. ***, **, and * denote statistical significance at the .01, .05, and .10 levels, respectively.

Um primeiro comentário que a tabela nos merece refere-se ao facto de a grande maioria das empresas pertencer ao regime de conhecimento 5. Em termos de número de empresas seguem-se os regimes de conhecimento 4, depois o 1, a seguir o 2 e, finalmente, o 3.

Da análise da Tabela verificamos que no regime de conhecimento 1, um aumento de 1% do volume de negócios conduz a uma redução nas despesas de I&D de 0,24%. Se considerarmos o regime de conhecimento 5 verificamos, por seu lado, que um aumento de 1% do volume de negócios conduz a uma redução nas despesas de I&D de 0,33%. Regista-se, assim, que, *ceteris paribus*, para um mesmo aumento hipotético no volume de negócios, há uma diminuição de cerca de mais 9 p.p. nas despesas de *I&D* no regime de conhecimento 5 do que no regime de conhecimento 1. Isto significa que as empresas pertencentes ao regime de conhecimento 1 canalizam uma maior percentagem do seu volume de negócios para I&D do que as empresas pertencentes aos restantes regimes de conhecimento.

Este tipo de leitura e a própria dimensão das relações aplica-se também para a variável intensidade de inovação total.

4. 4. Influência do regime tecnológico do sector na relação entre inovação e dimensão empresarial

Na presente secção analisam-se os efeitos dos regimes tecnológicos em que se integram as empresas, na relação entre intensidade de inovação e dimensão empresarial. Isto significa que, nesta etapa, considera-se simultaneamente os efeitos das características dos sectores e dos regimes de conhecimento a referida relação.

Os resultados da estimação das equações do tipo (5), (6) e (5.1), (6.1), pelo Método dos Mínimos Quadrados, encontram-se nas Tabelas 4.4. e 4.5, respectivamente.

Tabela 4.4. Intensidade de I&D e de inovação e regimes tecnológicos

Variables	(5) R&D intensity	(6) Innovation intensity
Firm's size	-0.489*** (0.034)	-0.397*** (0.024)
Marketing (dummy)		0.278*** (0.088)
Equipment (dummy)		0.981*** (0,135)
Technological regime Empr	3.419*** (0.521)	2.001*** (0.401)
Technological regime Rot	2.588*** (0.541)	1.452*** (0.403)
Number of observations	730	1602
F test	2119.99***	1396.98***
R-squared	0.9007	0.8182

Tabela 4.5. Intensidade de I&D e de inovação por regime tecnológico

(5.1)	**REmpr** R&D intensity	**RRot** R&D intensity
Firm's size	-0.271***(0.008)	-0,328***(0.005)
Number of observations	199	531
F test	1030.02***	5065.99***
R-squared	0.8519	0.9075
(6.1)	Inov intensity	Inov intensity
Firm's size	-0,294*** (0,017)	--0,312*** (0,009)
Mark dummy	0,584*** (0,182)	0,151 (0,101)
Equip dummy	1,317*** (0,251)	1,169*** (0,139)
Number of observations	366	1236
F test	370,65***	1944,91***
R-squared	0,768	0,828

Notes: Independent and dependent variables are all measured in logarithmic form, except dummy variables. Robust standard errors are given in parenthesis. ***, **, and * denote statistical significance at the .01, .05, and .10 levels, respectively.

Numa primeira análise verificamos que quase todas as variáveis são estatisticamente significativas.

A Tabela 4.4 diz-nos que o facto de as empresas pertencerem ao regime empreendedor, significa que a sua intensidade de inovação é maior do que se pertencerem ao regime rotineiro. De facto, o regime empreendedor acresce cerca de 1 p.p. à intensidade de I&D e 0.6 p.p. à intensidade de inovação.

Na Tabela 4.5 verificamos que, *ceteris paribus*, no regime empreendedor, um aumento de 1% no volume de negócios, conduz a uma diminuição na intensidade de inovação (I&D e inovação total) inferior à registada no regime rotineiro.

Isto poderá denunciar que no regimes empreendedor as empresas tendem a realizar um maior esforço de inovação, canalizando, portanto, uma maior percentagem dos seus resultados para actividades de inovação. Ou seja, a relação entre intensidade de inovação e dimensão empresarial é negativa, mas, em média, no regime empreendedor as empresas, grandes ou pequenas, realizam um maior esforço de inovação do que as empresas do regime rotineiro.

5. CONCLUSÃO

O objectivo do presente trabalho foi analisar a relação entre intensidade de inovação e dimensão empresarial. Começámos, numa primeira etapa, por explicar a intensidade de inovação (medida através da intensidade de I&D e da intensidade da inovação total) através de variáveis explicativas ao nível da empresa e ao nível do sector, socorrendo-nos de informação recolhida no âmbito do CIS 4 e de estatísticas da OCDE, respectivamente. Numa segunda etapa, o objectivo foi enfatizar a importância dos regimes de conhecimento para a referida relação. Finalmente, numa terceira etapa, colocou-se a tónica na influência dos regimes tecnológicos na relação acima referida.

Nas sucessivas etapas da nossa análise, verificámos que a relação entre intensidade de inovação e dimensão empresarial é negativa e estatisticamente significativa.

As grandes empresas possuem mais recursos para em termos absolutos realizarem maiores volumes de despesas em actividades de inovação, pelo que seria de esperar que a relação fosse positiva. por outro lado. Contudo, pelo facto de a variável explicada ser uma variável (I&D e inovação total) relativizada pelo volume de negócios, a relação revelou ser negativa, ou seja, são as empresas de menor dimensão que realizam um maior esforço de inovação.

Quando considerámos o efeito dos regimes de conhecimento, concluímos que o facto de as empresas fazerem parte do regime de conhecimento baseado nas ciências parece revelar um maior esforço em termos de intensidade de inovação.

Quando considerámos o efeito dos regimes tecnológicos, verificámos que é no regime empreendedor que as empresas tendem a realizar um maior esforço de inovação, canalizando, portanto, uma maior percentagem dos seus resultados para actividades de inovação. Isto poderá ter que ver com o facto de no primeiro regime as empresas adquirem externamente o conhecimento de que necessitam, e no segundo o processo de aprendizagem ser interno e cumulativo.

REFERÊNCIAS BIBLIOGRÁFICAS

ABRUNHOSA, A. (2006) – "Como competir em Sectores Tradicionais? A Inovação Organizacional como Fonte de Vantagem Competitiva", Dissertação de Doutoramento em Economia, FEUC.

ACS, ZOLTAN J. e AUDRETSCH, DAVID B. (1991), "Innovation and Small Firms", MIT Press, Cambridge.

ARCHIBUGI, D., EVANGELISTA, R. e SIMONETTI, R. (1995), "Concentration, firm size and innovation: evidence from innovation costs", *Technovation,* 15 (3), , pp. 153-163.

CARREIRA, C. (2004) – "Dinâmica Industrial e Crescimento da Produtividade - Uma Análise Microeconómica do Papel da Aprendizagem e da Reafectação dos Recursos no Crescimento Industrial", Dissertação de Doutoramento em Economia, FEUC.

COHEN, LEVIN e MOWERY (1987), "Firm Size and R&D Intensity: a Re-examination", Journal of Industrial Economics, 35(4), 543-565.

DAVIDSON, R., e J. G. MACKINNON (1993), "Estimation and Inference in Econometrics", New York: Oxford University Press.

DOSI, G., O. MARSILI, L. ORSENIGO e R. SALVATORE (1995), "Learning, Market Selection and the Evolution of Industrial Structures", Small Business Economics, 7(6), 411-436.

DOSI, G. e L. ORSENIGO (1988), "Industrial Structure and Technological Change" in Heertje, A. (ed.), *Innovation, Technology and Finance*, Oxford: Basil Blackwell, pp. 14-37.

EVANGELISTA, R. (1996), *"Embodied and disembodied innovative activities: evidence from the Italian Innovation Survey", in* OECD, *Innovation, Patents and Technological Strategies,* OECD, Paris, pp. 139-157.

EVANGELISTA, R., PERANI, G., RAPITI, F. e ARCHIBUGI, D. (1997), "Nature and Impact of Innovation in Manufacturing Industry: some evidence from the Italian innovation survey", Research Policy 26, pp. 521 – 536.

GRANSTRAND, O., P. PATEL e K. PAVITT (1997), "Multi-technology Corporations: Why They Have "Distributed" Rather than "Distintive Core" Competencies", *California Management Review*, 39 (4), 8-25.

GONÇALVES e VIEIRA A.; ALMEIDA, F.; ALMEIDA I.; FERREIRA J.; AMADO N. (2006), CIS – Inquérito Comunitário à Inovação – Quadros/Síntese dos resultados, Observatório da Ciência e do Ensino Superior.

GPEARI/MCTES (2004) Gabinete de Planeamento, Estratégia, Avaliação e Relações Internacionais/Ministério da Ciência, Tecnologia e Ensino Superior, *Community Innovation Survey*.

MARSILLI, O. (2001), "The Anatomy and Evolution of Industries – Technological Change and Industrial Dynamics", Edward Elgar Publishing, Inc., Massachusetts.

MAZZUCATO, M. (2000), "Firm Size, Innovation and Market Structure – The Evolution of Industry Concentration and Instability", Edward Elgar Publishing, Inc., Massachusetts.

SCHUMPETER, J.A. (1942), *Capitalism, Socialism and Democracy*, Harper, New York.

SOETE, L. G. (1979), "Firm Size and Inventive Activity. The Evidence Reconsidered", *European Economic Review*, 12 (4), 319-340.

Anexo

Tabela 1: Regimes de conhecimento para os vários sectores de actividade

Regimes	Sectores de actividade
Baseado nas ciências	Medicamentos
	Engenharia biomédica
	Computadores
	Telecomunicações
	Aparelhos eléctricos
	Instrumentos (fotografia e fotocopiadoras)
Processos fundamentais	Química
	Minas e petróleo
Sistemas complexos	Veículos a motor
	Aviões
Engenharia de produto	Maquinaria não eléctrica
	Instrumentos (controlos mecânicos, eléctricos e instrumentos mecânicos)
	Produtos metálicos fabricados
	Borracha e produtos plásticos
	Outro fabrico
Processos contínuos	Processo metalúrgico (metais básicos, materiais de construção)
	Processos Químicos (tecidos, papel e algodão)
	Comida e bebida (comida, bebida e tabaco)

Fonte: Adaptado de Marsili (2001:95)

Tabela 2: Sector de actividade económica e respectivo regime tecnológico

Sector actividade económica	Regime tecnológico
Extracção de petróleo bruto, gás natural e actividades dos serviços relacionadas, excepto a prospecção, indústrias extractivas, com excepção da extracção de produtos energéticos	2
Indústrias alimentares, das bebidas e do tabaco	5
Fabricação de têxteis	5
Indústria do vestuário, preparação, tingimento e fabricação de artigos de pele com pêlo	5
Curtimenta e acabamento de peles sem pêlo, fabricação de artigos de viagem, marroquinaria, artigos de correeiro, seleiro e calçado	5
Indústrias da madeira e da cortiça e suas obras, excepto mobiliário, fabricação de obras de cestaria e de espartaria	5
Fabricação de pasta, de papel e cartão e seus artigos	5
Edição, impressão e reprodução de suportes de informação gravados	4
Fabricação de coque, produtos petrolíferos refinados e tratamento de combustível nuclear e fabricação de produtos químicos	2
Fabricação de artigos de borracha e de matérias plásticas	4
Fabricação de outros produtos minerais não metálicos	4
Indústrias metalúrgicas de base	5
Fabricação de produtos metálicos, excepto máquinas e equipamento	5
Fabricação de máquinas e de equipamento, n.e.	4
Fabricação de máquinas de escritório e de equipamento para o tratamento automático da informação e fabricação de máquinas e aparelhos eléctricos, n.e.	4
Fabricação de equipamento e de aparelhos de rádio, televisão e comunicação	1
Fabricação de aparelhos e instrumentos médico-cirúrgicos, ortopédicos, de precisão, de óptica e de relojoaria	1
Fabricação de veículos automóveis, reboques e semi-reboques	3
Fabricação de outro material de transporte	3
Fabricação de mobiliário, outras indústrias transformadas, n.e.	5
Reciclagem	2
Produção e distribuição de electricidade, de gás, de vapor e água quente	5
Captação, tratamento e distribuição de água	5
Construção	5
Comércio por grosso e agentes do comércio, excepto de veículos automóveis e de motociclos	5
Comércio a retalho excepto de veículos automóveis, motociclos e combustíveis para veículos, reparação de bens pessoais e domésticos	5
Alojamento e restauração, restaurantes e similares	5
Transportes terrestres, transportes por oleodutos ou gasodutos	5
Transportes por água	5
Transportes aéreos	5
Actividades anexas e auxiliares dos transportes, agências de viagens e de turismo e de outras actividades de apoio turístico	5
Correios e telecomunicações	1
Intermediação financeira, excepto seguros e fundos de pensões	4
Seguros, fundos de pensões e outras actividades complementares de segurança social	5
Actividades auxiliares de intermediação financeira	4
Actividades informáticas e conexas e investigação e desenvolvimento	1
Outras actividades de serviços prestados principalmente às empresas	5

Barreiras à Capacidade Inovadora das Empresas Portuguesas: Evidências Empíricas ao Nível da Inovação no Produto

ALEXANDRA SOFIA CARAMELO NUNES
E-mail: alex.nunes@sapo.pt
MARIA JOSÉ AGUILAR MADEIRA SILVA
E-mail: msilva@ubi.pt
Universidade da Beira Interior

Resumo

A presente investigação visa identificar as barreiras à inovação que influenciam o processo inovador nas empresas portuguesas e, consequentemente, a sua capacidade inovadora ao nível da inovação no produto. Apresenta um modelo conceptual assente nas abordagens de referência sobre inovação empresarial de referência, nomeadamente a abordagem sistémica da inovação e a abordagem das redes e das relações inter-organizacionais. No sentido de testar as hipóteses formuladas utilizam-se dados secundários facultados pelo "Observatório da Ciência e do Ensino Superior" (OCES), pertencentes ao 4º Inquérito Comunitário à Inovação – *Community Innovation Survey* (CIS 4). Este questionário foi implementado sob a supervisão do EUROSTAT. O método utilizado no teste é o modelo de regressão logística. Ao nível dos resultados destaca-se, como barreira principal, o factor de âmbito económico: falta de financiamento de fontes externas. No entanto, os resultados evidenciam outras barreiras que condicionam a actividade e o desempenho inovador das empresas ao nível da inovação no produto.

Palavras-chave: Barreiras à Inovação, Redes, Capacidade Inovadora Empresarial, CIS.
Área temática: Economia da Inovação e Transferência de Tecnologia.

Abstract

This paper aims to identify and analyze the barriers to innovation that influence the innovation capability of Portuguese firms. The focus is on the study of entrepreneurial innovation capability, specifically the technological innovations made by the firm in relation to product innovation. The conceptual model that is proposed makes use of the system innovation approach, as well as the networks and inter-organisational relationships approach. The database is extracted from the Community Innovation Survey 4 (CIS 4). According to the data granted by the "Observatório da Ciência e do Ensino Superior" (OCES), that was conducted by EUROSTAT. In order to identify the significant determinants of entrepreneurial innovative capability, a logistic regression is performed. The most frequent barrier in almost all the dimensions of the innovation capability is the economic factor: lack of external financial sources. The analysis suggests several barriers to innovation have influence on the innovation capability of Portuguese firms.

Key words: Barriers to Innovation, Networks, Entrepreneurial Innovative Capability, CIS.
Thematic Area: Innovation and Technological Transfer Economics.

1. INTRODUÇÃO

Hoje em dia, é possível verificar que a inovação assume uma importância relevante, com impacto em todos os sectores e níveis da sociedade, sendo vista como o factor chave da competitividade das empresas e mesmo das nações, num contexto de globalização. É pois aceite que a procura de vantagens competitivas sustentáveis passa a depender cada vez mais da capacidade de inovação da empresa (Silva, 2003). Desta forma, torna-se necessário um conhecimento mais aprofundado do processo de inovação, incidindo principalmente, no estudo dos factores que impulsionam e limitam a inovação empresarial.

Em Portugal, o interesse por esta temática, por parte da Administração Pública e das empresas, despertou nos finais da década de 80, início de 90, através do desenvolvimento de um conjunto de iniciativas destinadas ao fortalecimento das capacidades de inovação e dos mecanismos de difusão tecnológica. Mais recentemente, surgiram vários apoios e incentivos no âmbito do QREN – Quadro de Referência Estratégico Nacional - Portugal 2007-2013, QREN (2007), assumindo-se como grande ponto estratégico a qualificação dos portugueses, valorizando o conhecimento, a ciência, a tecnologia e a inovação, bem como a promoção de níveis elevados e sustentados de desenvolvimento económico e sociocultural e de qualificação territorial, ou do aumento da eficiência e qualidade das instituições públicas.

É certo que existem já, em termos académicos, alguns estudos sobre as barreiras à capacidade inovadora das empresas portuguesas, (Simões, 1996; Barrau, 2000; Conceição e Ávila, 2001; Ussman *et al*, 2001; Silva, 2003; Bóia, 2003; Faria, 2004; Marques e Monteiro-Barata, 2005; Barbosa, 2006; Marques e Monteiro-Barata, 2006; Silva e Leitão, 2007; Natário *et al*, 2007; Silva *et al*, 2008), no entanto, revela-se que se trata de uma temática bastante premente e em constante actualização, de acordo com o desenvolvimento de novos inquéritos CIS. Desta forma, o principal contributo deste artigo passa pela identificação e análise dos principais factores limitadores da capacidade inovadora empresarial, a nível de inovação no produto, incidindo na identificação da sua natureza, origem e importância, promovendo uma melhor compreensão dos mesmos e abrindo caminho para estudos futuros, na descoberta das melhores práticas para ultrapassar essas barreiras.

Tomando como referência as abordagens actuais sobre a temática, destacam-se a abordagem sistémica da inovação e a abordagem das redes e relações interorganizacionais, como as mais adequadas ao estudo dos factores limitadores da capacidade inovadora empresarial (Silva, 2003). Desta forma, tendo em conta este quadro conceptual, pretende-se com esta investigação elaborar um suporte empírico que permita identificar e analisar os factores que condicionam a actividade e o desempenho inovador das empresas portuguesas, ao nível da inovação no produto.

O trabalho estrutura-se da seguinte forma, no ponto dois, com base na literatura relevante, formulam-se as hipóteses que se pretendem testar empiricamente no modelo estatístico. No ponto três define-se a amostra e, posteriormente, descrevem-se e caracterizam-se os dados utilizados no estudo empírico e apresenta-se o modelo de regressão logística para a inovação no produto. No ponto quatro efectua-se a análise dos resultados, e as principais conclusões surgem no último ponto.

2. REVISÃO DA LITERATURA

Ao longo dos últimos anos e numa era de globalização, a inovação tem despertado um grande interesse, ao nível académico. O vasto reconhecimento de que as empresas necessitam de inovar para sobreviverem, sob pena de desaparecerem, assim como a dificuldade em estimular e sustentar a inovação e o espírito empreendedor nas mesmas, têm gerado inúmeras investigações. Em grande parte dos estudos sobre inovação empresarial, tende-se a associar a noção de inovação às actividades de I&D, ou à ideia de tecnologia material, no entanto, esta não se restringe apenas a estas dimensões (Silva, 2003).

Tendo em conta vários autores (Kaufmann e Tödtling, 2001; Silva, 2003; Foxon e Pearson, 2008), a *inovação empresarial* define-se como um processo não linear, evolucionário, dinâmico, complexo e interactivo de aprendizagem e de relacionamentos entre a empresa e o seu meio envolvente.

Desta forma, a Inovação não é algo que as empresas fazem uma única vez e esquecem, ou que suceda por acidente, é sim uma capacidade que necessita de ser desenvolvida e praticada frequentemente. Esta capacidade inovadora, de acordo com Silva (2003), define-se como o resultado do processo de Inovação empresarial, integrando as diversas componentes resultantes desse mesmo processo: inovação no produto, inovação no processo e inovação organizacional, às quais o Manual de Oslo (OCDE, 2005), acresce a inovação de marketing.

Neste artigo decidiu-se apenas analisar a componente inovação no produto, levada a cabo pelas empresas portuguesas. De acordo com o CIS 4 (2005:3), a inovação de produto "é a introdução no mercado de um bem ou serviço novo ou significativamente melhorado relativamente às suas capacidades iniciais, tais como a melhoria no software, "mais amigável", novos componentes ou subsistemas".

Antes de mais, é necessário verificar que uma invenção não se torna inovação até ao momento em que esta entra no processo de produção e de operações de marketing e sendo depois difundida pelo mercado (Garcia e Calantone, 2002). Uma inovação não necessita de ser desenvolvida pela própria empresa, podendo ser adquirida de outras empresas ou instituições através do processo de difusão. Difusão é pois, a forma como as inovações se dispersam, quer seja, através dos canais de mercado ou não (OCDE, 2005).

Entre as décadas de 50 e 80, verificou-se uma grande preocupação à volta desta temática. Nas últimas duas décadas, a perspectiva sistémica da inovação e a perspectiva das redes de inovação, têm ganho uma grande importância, ao acrescentar ao estudo outros factores que influenciam o processo de inovação (Silva, 2003). Qualquer uma destas perspectivas defende que as empresas não inovam isoladas, logo procuram estabelecer relacionamentos com outras empresas ou organizações, com o objectivo de aceder a recursos, informações e conhecimentos de que necessitam para sobreviver a um ambiente bastante competitivo. Ambas as abordagens realçam, também, que a capacidade inovadora varia de empresa para empresa, bem como é determinada por um vasto e complexo número de factores, tanto impulsionadores como limitadores do processo de inovação empresarial (Silva *et al*, 2008).

A perspectiva sistémica da inovação desenvolveu-se em meados dos anos 80, graças às contribuições de vários autores, Nelson e Winter (1982), Freeman (1987, 1988), Lundvall (1985, 1988, 1992), Nelson e Rosenberg (1993), Tödtling (1995), Edquist (1997), Cooke *et al* (1997), Braczyk *et al* (1998), Mytelka (2000) e Kaufmann e Tödtling (2001), que acrescentaram vários factores, tais como a organização institucional, a cultura e a história, à análise da inovação (Silva, 2003). De acordo com estes autores, a abordagem sistémica apresenta a capacidade inovadora como o resultado de um processo de interacção das empresas com o seu meio envolvente, realçando as sinergias de aprendizagem inerentes ao sistema económico e o estímulo das instituições de suporte à inovação (Silva *et al*, 2008). Esta abordagem pretende, desta forma, abarcar um amplo conjunto dos determinantes de inovação no contexto nacional, regional ou sectorial.

Por seu lado, a abordagem das redes e das relações inter-organizacionais analisa as ligações estabelecidas entre as empresas e outras organizações, tendo em vista a melhoria das suas capacidades inovadoras (Magee, 1992; UNCTAD, 1998; Silva, 2003; Freire, 2006), realçando, portanto, o papel da cooperação, da complementaridade e da dependência entre os diversos actores económicos (Barbosa, 2006). Nas duas últimas décadas, esta abordagem tem registado um crescimento significativo em termos de artigos publicados, novas perspectivas teóricas e resultados de investigações empíricas (Silva, 2003) a destacar a abordagem do Cluster Industrial (Porter, 1994, 1998; UNCTAD, 1998; Hsien-Che e Shyu, 2005) que dá uma grande ênfase à pressão concorrencial exercida pelo meio envolvente sobre as empresas (Silva *et al*, 2008); a abordagem do Distrito Industrial (Becattini, 1990, UNCTAD, 1998; Muscio, 2006) onde se destacam os relacionamentos funcionais entre os vários actores do sistema (Silva, 2003); as Redes Industriais (Hakansson, 1987; Axelsson e Easton, 1992; Hakansson e Johanson, 1992),

que realçam o papel dos actores, actividades e recursos (Silva *et al*, 2008), focando-se a importância dos relacionamentos externos no âmbito da inovação (Silva, 2003); e a abordagem dos Recursos e Capacidades que reconhece que os recursos e capacidades existentes no interior da empresa são imprescindíveis para a obtenção de vantagens competitivas sustentáveis (Pfeffer e Salancik, 1978; Prahalad e Hamel, 1990; Silva, 2003).

2.1. Barreiras à Inovação

É pois, perceptível que a capacidade inovadora varia de empresa para empresa, sendo determinada por um número complexo e vasto de factores, internos e externos, tanto estimulantes como restritivos, que promovem um impacto significativo no processo de concepção, implementação e de difusão da inovação (Hadjimanolis, 1999; Conceição e Ávila, 2001; Bóia, 2003; Silva, 2003; Silva *et al*, 2008). Neste contexto, definem-se as barreiras à inovação como factores internos ou externos à empresa que diminuem a propensão da mesma para inovar, isto é diminuem a sua capacidade de introduzir um produto novo ou um significativamente melhorado.

Após análise do processo de inovação, verifica-se que as barreiras à actividade inovadora apercebidas pelas empresas são bastantes. Como a percepção das mesmas depende de diferentes características ou determinantes empresariais (Tourigny e Le, 2004), assim, as que irão ser estudadas nesta investigação, resultam das definidas pelo Inquérito à inovação mais actual em Portugal, o CIS 4 (2005:10) e para o qual dispomos de dados. Desta forma, o CIS 4 apresenta as barreiras à inovação divididas em quatro tipos de factores: (1) económicos; (2) de conhecimento; (3) de mercado e (4) razões para não inovar, como se pode verificar no Quadro 1.

Quadro 1. Barreiras à inovação de acordo com o CIS 4

Factores de Impedimento / Barreiras à Inovação	
Factores Económicos	Insuficiência de Capitais Próprios ou do Grupo a que pertence
	Falta de Financiamento de Fontes Externas
	Custos com a Inovação demasiado elevados
Factores de Conhecimento	Falta de Pessoal Qualificado
	Falta de Informação sobre Tecnologia
	Falta de Informação sobre os Mercados
	Dificuldade em encontrar Parceiros para Cooperação em Projectos e Inovação
Factores de Mercado	Mercado dominado por Empresas estabelecidas
	Incerteza na procura / mercado para os Bens ou Serviços novos
Razões para não Inovar	Desnecessário por já existirem inovações anteriores
	Desnecessário pela inexistência de procura / mercado para Inovações

Fonte: CIS 4 (2005:10)

Os factores económicos são considerados como barreiras quando os gerentes se apercebem da insuficiência de capitais, da falta de financiamento de fontes externas ou os custos de inovação são demasiado elevados. Hadjimanolis (1999) no seu estudo releva o papel dos problemas de financiamento no desenvolvimento de novos produtos, bem como o papel da burocracia, como uma das barreiras mais importantes ao processo inovador. Utilizando os dados do CIS 2, foi também possível verificar que, os custos elevados e a falta de financiamento são os impedimentos mais frequentemente apontados como causas de atrasos e do não início de projectos de inovação, em empresas industriais (Silva, 2003; Silva *et al*, 2008). No que diz

respeito aos projectos cancelados, para além dos custos elevados, a percepção de riscos excessivos e as regulamentações e normas surgem como os mais importantes (Conceição e Ávila, 2001). Também para Barbosa (2006) os elevados custos de investimento que implica o processo inovador, são a principal barreira ao desenvolvimento de processos de inovação tecnológica.

Os factores de conhecimento estão, por um lado, associados ao nível de qualificação do pessoal presente no seio de cada organização e, por outro, à quantidade e qualidade das fontes de informação a que a empresa consegue aceder, no meio em que se insere. A falta de pessoal qualificado é abordada por vários autores, Hoffman *et al* (1998), Romijn e Albaladejo (2002), Conceição e Ávila (2001), Silva (2003); Bóia (2003) e Silva *et al* (2008) verificaram que as empresas possuidoras de maior nível de pessoal qualificado revelam maiores propensões para inovar. Barrau (2000) chega mesmo a considerar a falta de qualificação e formação técnicoprofissional como um das barreiras mais críticas ao desenvolvimento, à cooperação, à endogeneização associada aos processos de criação e adopção de tecnologias e, consequentemente, à inovação.

Vários estudos desenvolvidos em diversos países evidenciam, também a importância dos relacionamentos externos, caracterizados por trocas de informação ou difusão de conhecimento, como estimulantes do processo de inovação empresarial (Kaufmann e Tödtling, 2001; Romijn e Albaladejo, 2002). Em Portugal, os resultados da análise ao CIS 2 efectuada por Conceição e Ávila (2001) e Silva (2003) e ao CIS 3 efectuada por Bóia (2003) demonstram que, tanto na indústria como nos serviços, as empresas integradas em grupos inovam substancialmente mais do que as que actuam independentemente. Silva (2003), Barbosa (2006), Nieto e Santamaria (2007) e Silva e Leitão (2009) verificaram que os relacionamentos externos estabelecidos com os parceiros de negócio e com as universidades influenciam a empresa a desenvolver avanços inovadores a nível de inovação no produto.

No que respeita aos factores de mercado, estes são visionados como entraves à inovação quando as empresas se apercebem que o mercado é já dominado por outras empresas ou se verifica uma grande incerteza na procura / mercado para os bens ou serviços novos. Silva *et al* (2008) concluíram que a falta de receptividade dos clientes a novos produtos, leva as empresas a demonstrarem menor propensão para inovar. Assim, se a empresa percepciona que o mercado não está interessado nos novos produtos, não tem incentivos para inovar e, desta forma, esta percepção actua como um obstáculo à inovação. Por último, as razões para não inovar surgem neste trabalho, como variáveis exploratórias em análise, pois constatou-se que os vários estudos não consideraram as razões para não inovar como barreiras à inovação.

2.2. Definição das Hipóteses de Investigação

A cada um dos factores e barreiras do modelo foi associada uma hipótese teórica, com a finalidade de obter conhecimentos acerca das influências que essas mesmas barreiras exercem sobre a capacidade inovadora empresarial, mais concretamente sobre a inovação no produto, tal como apresentadas no Quadro 2.

Quadro 2. Hipóteses de Investigação

Hipóteses de Investigação	
Factores Económicos	H_1: *A Insuficiência de Capitais Próprios ou do Grupo a que pertence está negativamente relacionada com a propensão da empresa para inovar no produto.*
	H_2: *A Falta de Financiamento de Fontes Externas está negativamente relacionada com a propensão da empresa para inovar no produto.*
	H_3: *Os Custos com a Inovação demasiado elevados estão negativamente relacionados com a propensão da empresa para inovar no produto.*
Factores de Conhecimento	H_4: *A Falta de Pessoal Qualificado está negativamente relacionada com a propensão da empresa para inovar no produto.*
	H_5: *A Falta de Informação sobre Tecnologia está negativamente relacionada com a propensão da empresa para inovar no produto.*
	H_6: *A Falta de Informação sobre os Mercados está negativamente relacionada com a propensão da empresa para inovar no produto.*
	H_7: *A dificuldade em encontrar Parceiros para Cooperação em Projectos e Inovação está negativamente relacionada com a propensão da empresa para inovar no produto.*
Factores de Mercado	H_8: *O Mercado dominado por Empresas estabelecidas está negativamente relacionado com a propensão da empresa para inovar no produto.*
	H_9: *A incerteza na procura / mercado para os Bens ou Serviços novos está negativamente relacionada com a propensão da empresa para inovar no produto.*
Razões para Não Inovar	H_{10}: *Ser desnecessário por já existirem inovações anteriores está negativamente relacionado com a propensão da empresa para inovar no produto.*
	H_{11}: *Ser desnecessário pela inexistência de procura / mercado para Inovações está negativamente relacionado com a propensão da empresa para inovar no produto.*

Segundo o Observatório da Criação de Empresas (IAPMEI, 2007), Portugal é constituído predominantemente por micro, pequenas e médias empresas. Desta forma, a realidade portuguesa é apresentada como um laboratório ideal para testar as hipóteses, com o objectivo de gerar conhecimento e propor orientações que guiem as entidades públicas e privadas na formulação de medidas que visem a promoção da capacidade inovadora empresarial e a superação das barreiras à inovação (Silva *et al*, 2008).

3. METODOLOGIA DE INVESTIGAÇÃO

Neste ponto procedeu-se à escolha da população objecto de estudo, que neste trabalho de investigação, são as empresas industriais, de construção e de serviços portuguesas. Definiu-se o método de recolha dos dados e por fim, escolheu-se o método de análise de dados: o modelo de regressão logística.

3.1. Apresentação dos Dados: População e Amostra

Os dados a usar, foram recolhidos através de inquérito por questionário, denominado de 4º Inquérito Comunitário à Inovação (CIS 4 - *Community Innovation Survey 4*) entre Junho e Novembro de 2005. Em Portugal, o inquérito foi conduzido pelo OCES, com a colaboração do INE – Instituto Nacional de Estatística, sendo referente às actividades de inovação das empresas portuguesas no período de 2002 a 2004.

A população contempla todas as empresas industriais, de serviços e construção com pelo menos cinco colaboradores. A nível de classificação de actividades económicas (CAE), a população foi subdividida pelas secções C à K da CAE – Rev. 2.1. (2003). A amostra inicial foi construída pelo INE, de acordo com as especificações metodológicas do EUROSTAT, sendo extraída de uma população de 23.348 empresas registadas no FUE – Ficheiro de Unidades

Estatísticas do INE. Portanto, desta população, extraiu-se uma amostra inicial de 7.370 empresas, que mais tarde foi corrigida para 6.482 empresas. Desta amostra, 4.815 empresas responderam ao questionário, correspondendo, portanto a uma taxa de resposta de 74,3%. As empresas da amostra foram consideradas como inovadoras no produto, se durante o período de 2002 a 2004, introduziram algum bem ou serviço, novo ou significativamente melhorado.

3.2. Descrição e Caracterização dos Dados

Como, neste trabalho, a capacidade inovadora é medida a partir da informação recolhida, ao nível da inovação no produto, sendo considerada como a variável dependente. Esta dimensão apresenta-se como uma variável dicotómica baseada em dados binários: assume o valor 0 para empresas que não inovaram e o valor 1 para aquelas que inovaram. No que concerne às variáveis independentes, estas são representadas pelas barreiras à inovação. Para esta investigação, apenas se teve em conta se o factor funcionava ou não como barreira, assumindo-se também, como uma variável dicotómica baseada em dados binários: assume o valor 1 quando é percepcionada como uma barreira e o valor 0 quando o não é.

3.3. Método: Modelo de Regressão Logística

A partir da revisão teórica da literatura efectuada e do modelo conceptual proposto, verificou-se que a capacidade inovadora empresarial é um fenómeno complexo influenciado por um vasto conjunto de factores. Como se torna necessário explorar as relações entre esses factores e a capacidade inovadora, mais especificamente pretende-se estudar a relação estatística de uma variável dependente em relação a mais do que uma variável explicativa, decidiu-se utilizar o Modelo de Regressão Logística (*Logit Model*). Na dimensão de análise da capacidade inovadora do produto, corresponde um modelo de regressão em que a variável dependente ou resposta é dicotómica, um dos modelos adequados de aplicação é o modelo de regressão Logit (Manso, 1998).

Este modelo foi o mais utilizado nos estudos empíricos realizados (Kaufmann e Tödtling, 2001; Conceição e Heitor, 2001; Silva, 2003; Bóia, 2003), apresentando-se assim, como a técnica analítica apropriada para os modelos conceptuais propostos, uma vez que estes incluem uma variável dependente categórica (binária ou dicotómica) e várias variáveis independentes também categóricas, tal como se apresenta seguidamente:

$$IP = \beta_0 + \beta_1 C_1 + \beta_2 C_2 + \beta_3 C_3 + \beta_4 C_4 + \beta_5 C_5 + \beta_6 C_6 + \beta_7 C_7 + \beta_8 C_8 + \beta_9 C_9 + \beta_{10} C_{10} + \\ + \beta_{11} C_{11} + \varepsilon_i$$

onde: IP = Inovação Produto ε_i = Resíduo β = Coeficientes e C_i = Barreiras

4. RESULTADOS: APRESENTAÇÃO E DISCUSSÃO

No período de 2002 a 2004, 44,8% das empresas da amostra inovaram, no entanto, 55,2% das empresas não desenvolveram qualquer tipo de inovação. É de referir que, 407 empresas não concretizaram os projectos de inovação, dado que os mesmos não foram bem sucedidos, levando, consequentemente, ao seu abandono. Por seu lado, 483 empresas vêem os seus projectos de inovação bastante atrasados. É de destacar que nestes dados encontram-se empresas que desenvolveram mais do que um tipo de inovação.

No modelo de inovação no produto as empresas da amostra foram consideradas como inovadoras no produto, se durante o período de 2002 a 2004, introduziram algum bem ou serviço, novo ou significativamente melhorado. Do total da amostra, isto é, 4.815 empresas, apenas 1.275 (26,5%) inovaram no produto durante o referido período, destacando-se, consequentemente, 3.540 (73,5%), que o não fizeram.

Aplicando o modelo de regressão logística à inovação no produto, obtiveram-se os seguintes resultados:

Quadro 3. Resultados da regressão logística para o modelo de inovação no produto

Modelo 1	Modelo Inicial		Modelo Final		
	Estimativa Coeficiente (B)	Sig.	Estimativa Coeficiente (B)	Sig.	Exp (B)
Insuficiência Capitais Próprios ou do Grupo	-0,017	0,704			
Falta Financiamento Fontes Externas	-0,108	0,012	-0,103	0,005	0,902
Custos com Inovação demasiado elevados	0,137	0,002	0,151	0,000	1,163
Falta de pessoal qualificado	0,026	0,560			
Falta informação sobre tecnologia	0,088	0,064			
Falta informação sobre mercados	-0,037	0,410			
Dificuldade em encontrar parceiros	-0,002	0,959			
Mercado dominado por empresas estabelecidas	0,000	0,999			
Incerteza na procura / mercado	0,152	0,000	0,166	0,000	1,180
Existência inovações anteriores	-0,038	0,408			
Inexistência de procura / mercado	-0,098	0,038	-0,115	0,000	0,892
SIZE		0,000			
SIZE_1	-1,040	0,000	-1,039	0,000	0,354
SIZE_2	-1,049	0,000	-1,042	0,000	0,353
SIZE_3	-0,571	0,000	-0,560	0,000	0,571
SIZE_4	-0,370	0,020	-0,369	0,020	0,691
SIZE_5	-0,518	0,141	-0,505	0,149	0,604
SIZE_6	-0,766	0,007	-0,769	0,007	0,463
SIZE_7	-0,893	0,001	-0,903	0,001	0,405
IT		0,000		0,000	
IT_1	-0,466	0,000	-0,469	0,000	0,626
IT_2	0,208	0,059	0,208	0,058	1,232
FU	1,755	0,000	1,755	0,000	5,786
Constante	-0,353	0,014	-0,329	0,021	0,719
Qualidade do Ajuste do modelo					
Casos correctamente predictos	76,2%		76,2%		
Qui-quadrado	483,263	0,000	476,404	0,000	
Log Likelihood	5.082,410		5.089,269		
Total de Casos	4.814		4.814		

O modelo inicial evidencia os resultados das relações entre a capacidade inovadora empresarial, ao nível da inovação no produto e as barreiras à inovação. Como se pode verificar, algumas variáveis associadas às barreiras não são estatisticamente significativas ao nível de 5%, logo as hipóteses H_1, H_4, H_5, H_6, H_7, H_8 e H_{10} não foram testadas empiricamente. Desta forma, procedeu-se, de seguida, à estimação do modelo, sem considerar essas variáveis, do qual resultou o modelo final. Como estatística de teste foi usada a estatística de Wald, promovendo um estudo da significância ao nível de 5%, podendo-se constatar que todas as barreiras são significativas a este nível, bem como a capacidade preditiva do modelo ou a percentagem de casos correctamente classificados se situa nos 76,2%. A estatística de teste do qui-quadrado tem o valor de 476,404 com valor de prova inferior ao nível de significância de 0,05. A estatística da

log verosimilhança, com o valor 5.089,269, também corrobora a significância global do modelo comparativamente ao modelo nulo.

Da análise do quadro, é possível verificar que das barreiras à inovação consideradas, existem quatro variáveis estatisticamente significativas: Falta de financiamento de fontes externas, Custos com a inovação demasiado elevados, Incerteza na procura / mercado para os bens ou serviços novos e Desnecessário pela inexistência de procura/mercado para inovações.

No que respeita à Falta de financiamento de fontes externas, verifica-se que tem um efeito significativo e negativo sobre a inovação no produto, levando a que se possa rejeitar a hipótese nula de inexistência de relação entre as variáveis, corroborando portanto a hipótese H_2. Assim, as empresas que enfrentam escassez de fontes de financiamento têm menor probabilidade de inovar no produto. Este resultado encontra-se de acordo com estudos anteriores (Hadjimanolis, 1999; Silva, 2003; Silva *et al*, 2008), que revelam a falta de fontes de financiamento, como uma das barreiras mais importantes à capacidade inovadora das empresas.

Surpreendentemente, a variável Custos com a inovação demasiado elevados, tem um efeito significativo mas positivo, surgindo não como uma barreira, mas como uma variável impulsionadora da inovação no produto. Desconhecendo-se a razão deste resultado, constitui uma limitação desta investigação, para ser analisada em estudos posteriores. Desta forma, pelo estudo rejeita-se H_3.

A barreira Incerteza na procura / mercado para os bens ou serviços novos apresenta um resultado significativo e positivo, evidenciando que quanto maior for a incerteza na procura / mercado maior é a probabilidade para inovar. Esta situação poderá explicar-se tendo em conta que o factor incerteza, apercebido pelos empresários, ao não saberem o que podem esperar, poderá levá-los a tender para o desenvolvimento da inovação como para a sua restrição, neste caso, para a maioria das empresas da amostra, esta variável funciona como um factor impulsionador, levando-os a arriscar na busca de rendibilidade. Desta forma, rejeita-se H_9.

Os resultados do modelo demonstram que a barreira Desnecessário pela inexistência de procura/mercado para inovações tem um efeito significativo na propensão para inovar. Deste modo, rejeita-se a hipótese nula de inexistência de relação entre as variáveis, permitindo confirmar a hipótese H_{11}. Esta informação confirma que a inexistência de procura/mercado para inovações inibe fortemente as empresas de inovar. Os resultados finais são apresentados no quadro seguinte:

Quadro 4. Resultados das hipóteses do modelo de inovação no produto

Hipóteses	Variáveis	Resultado
H_1	Insuficiência Capitais Próprios ou do Grupo	Não confirmada
H_2	Falta de financiamento de fontes externas	Confirmada
H_3	Custos com inovação demasiado elevados	Rejeitada
H_4	Falta de pessoal qualificado	Não confirmada
H_5	Falta informação sobre tecnologia	Não confirmada
H_6	Falta informação sobre mercados	Não confirmada
H_7	Dificuldade em encontrar parceiros	Não confirmada
H_8	Mercado dominado por empresas estabelecidas	Não confirmada
H_9	Incerteza na procura / mercado	Rejeitada
H_{10}	Existência inovações anteriores	Não confirmada
H_{11}	Inexistência de procura / mercado	Confirmada

5. CONCLUSÕES

De acordo com o trabalho teórico e empírico, foi possível verificar que a temática das barreiras à capacidade inovadora se apresenta de importância crucial no processo inovador empresarial e sobre o qual ainda existe muito para investigar. A formulação do modelo conceptual teórico apoiou-se na revisão da literatura que demonstrou que a capacidade inovadora empresarial é influenciada por um vasto conjunto de barreiras. Nesta investigação, decidiu-se utilizar as barreiras definidas pelo CIS 4 e verificar até que ponto os empresários as

percepcionam como factores limitadores das actividades de inovação. É certo que as barreiras não se limitam somente a estas, mais factores podem ser acrescentados ao estudo.

Pelos resultados obtidos, é possível verificar que o factor económico: falta de financiamento de fontes externas, tem uma acção bastante limitadora sobre este tipo de inovação. É, também, de destacar que as variáveis obtidas, através deste estudo, como barreiras à capacidade inovadora empresarial, no âmbito da inovação no produto, se podem englobar em factores externos à empresa. Bem como, nenhum factor de conhecimento foi destacado como significativo, de forma a representar uma barreira percepcionada pelos empresários inquiridos. Estranhamente, a variável custos com a inovação demasiado elevados não se apresenta como barreira, mas sim como factor impulsionador da actividade inovadora, pois o seu coeficiente é sempre positivo. Trata-se de uma investigação que poderá ser explorada em estudos posteriores.

As principais limitações deste estudo passaram pela restrição dos dados obtidos, não tendo sido facultados, na sua totalidade os resultantes do CIS 4, como por exemplo, a subdivisão da inovação por regiões. Por outro lado, não foi sido possível elaborar uma comparação de resultados, com os CIS anteriores, de modo a avaliar tendências evolutivas no âmbito das barreiras à inovação.

Futuras investigações no âmbito desta temática poderão assentar na incorporação de outras variáveis, não consideradas neste estudo, de forma a alargar o horizonte de análise. Outra proposta assenta em comparar os resultados de todos os CIS efectuados até ao momento, como forma de melhor compreender o fenómeno, evolução e suas tendências. Não descurando o CIS 2006 que se encontra em desenvolvimento. Seria também recomendável, proceder a uma comparação entre os vários países europeus que responderam ao CIS.

Após a análise das barreiras à inovação, um outro trabalho poderá passar pela definição de estratégias que permitam alcançar a melhor forma de ultrapassar essas dificuldades sentidas pelas empresas, aquando da necessidade de inovar. Assim, em termos de implicações do presente estudo, propõem-se actuações futuras ao nível das empresas que visem a superação de barreiras à inovação. Estas actuações poderão traduzir-se em medidas que possam proporcionar recursos, informação e conhecimentos necessários ao processo inovador das empresas, principalmente, das pequenas e médias empresas que se caracterizam pelos recursos escassos e fortes limitações financeiras. Assim, estas medidas devem comportar a oferta de um conjunto de instrumentos e incentivos para financiar e premiar, respectivamente, as actividades de inovação e a aquisição de novas competências empreendedoras e de inovação.

REFERÊNCIAS BIBLIOGRÁFICAS

AXELSSON, B. e EASTON, G. (1992), "Industrial Networks: A New View of Reality", London, Routledge.

BARBOSA, M. (2006), "A capacidade inovadora empresarial como factor criador de vantagens competitivas nas empresas da indústria transformadora da região da Beira Interior", Dissertação de Mestrado em Gestão, Universidade da Beira Interior, Covilhã.

BARRAU, I. (2000), "Inovação e Competitividade: O perfil inovador das empresas industriais da Beira Interior", Dissertação de Mestrado em ciências Económicas, Universidade da Beira Interior, Covilhã.

BECATTINI, G (1990), "The Marshallian Industrial District as a Socio-Economic Notion", in "Industrial Districts and Inter-Firm Co-operation in Italy", Pyke F, Becattini G, and SENGENBERGER W (eds), International Institute for Labour Studies, Geneva, pp. 37-51.

BÓIA, M. (2003), "Determinants of Innovation in Portugal. Designing, Implementing and Analyzing Evidence from the Third Community Innovation Survey", Dissertação de Mestrado em Engenharia e Gestão de Tecnologia, Universidade Técnica de Lisboa, Instituto Superior Técnico, Lisboa

Braczyk, H-J., Cooke, P. e HEIDENREICH, M. (1998), "Regional Innovation Systems", London and Bristol, Penn, UCL Press.

CAE – Rev. 2 (1993), "Classificação das Actividades Económicas – Rev. 2". Ministério do Planeamento e da Administração do Território, Diário da Republica nº 112, Iª Série, Decreto-Lei nº 182 de 14 de Maio de 1993, pp. 2539-2560.

CIS 4 (2005), "Inquérito Comunitário à Inovação", Observatório da Ciência e do Ensino Superior, Lisboa, Ministério da Ciência e da Tecnologia.

CONCEIÇÃO, P. e ÁVILA, P. (2001), "A Inovação em Portugal - II Inquérito Comunitário às Actividades de Inovação", Oeiras, Celta Editora.

COOKE, P., URANGA, M. e ETXEBARRIA, G. (1997), "Regional Innovation Systems: Institutional and Organizational Dimensions", *Research Policy*, Vol. 26, nº 4-5, pp.475-491.

EDQUIST, C. (1997), "Systems of Innovation Approaches - Their Emergence and Characteristics", em "Systems of Innovation: Technologies, Institutions and Organizations", EDQUIST C. (Ed.) Chapter One, London, Pinter, pp. 1-35.

FARIA, PEDRO (2004), "Innovation and Productivity: What can we learn from the CIS III Results for Portugal?" Dissertação de Mestrado em Engenharia e Gestão de Tecnologia, Universidade Técnica de Lisboa, Instituto Superior Técnico, Lisboa.

FOXON, T. e PEARSON, P. (2008), "Overcoming barriers to innovation and diffusion of cleaner technologies: some features of a sustainable innovation policy regime", *Journal of Cleaner Production* 16S1, S148-S161.

FREEMAN, C. (1987), "Technology policy and economic performance: lessons from Japan", London: Pinter. Publishers.

FREEMAN, C. (1988), "Japan: A New National System of Innovation?" em "Technical Change and Economic Theory", Dosi, G., Freeman, C., Nelson, R., Silverberg, G. e Soete, L. (Eds) London, Pinter.

FREIRE, A. (2006), "Inovação - Novos Produtos, Serviços e Negócios para Portugal", 2ª Reimpressão, Lisboa, Verbo.

GARCIA, R. e CALANTONE, R. (2002), "A critical look at technological innovation typology and innovativeness terminology: a literature review" *The Journal of Product Innovation Management,* Vol. 19, pp. 110-132.

HADJIMANOLIS, A. (1999), "Barriers to innovation for SMEs in a small less developed country (Cyprus)", *Technovation,* Vol. 19, pp. 561-570.

HAKANSSON, H. (1987), "Industrial Technology Development – A Network Approach", London: Croom Helm.

HAKANSSON, H. e JOHANSON, J. (1992), "A Model of Industrial Networks", em "Industrial Networks - A New View of Reality", Axelsson, B. e Easton, G. (Eds), Routledge, pp. 28-36.

HOFFMAN, K., PAREJO, M., BESSANT, J. e PERREN, L.. (1998), "Small Firms, R&D, Technology and Innovation in the UK: A Literature Review", *Technovation,* Vol. 18, nº.1, pp. 39-55.

HSIEN-CHE, L, e SHYU J.Z. (2005), "A comparison of innovation capacity at science parks across the Taiwan Strait: the case of Zhangjiang High-Tech Park and Hsinchu Sciencebased Industrial Park", *Technovation,* Vol. 25, pp. 805–813.

IAPMEI (2007), "Observatório da Criação de Empresas - Resultados do Inquérito 2006"; Direcção de Planeamento e Estudos, Lisboa.

KAUFMANN, A. e TÖDTLING, F. (2001), "Science - industry interaction in the process of innovation: the importance of boundary - crossing between systems" *Research Policy* , Vol. 30, pp. 791-804.

LUNDVALL, B.A. (1985), "Product Innovation and User-Producer Interaction", *Industrial Research*, Series Nº 31, Aalborg, Aalborg University Press.

LUNDVALL, B.A. (1988), "Innovation as an Interactive Process: From User-Producer Interaction to the National System of Innovation", em "Technical Change and Economic Theory", Dosi, G., Freeman, C., Nelson, R., Silverberg, G. e Soete, L. (Eds) Chapter 17, London, Pinter, pp. 349-269.

LUNDVALL, B.A. (1992), "National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning", London, Pinter.

MAGEE, J.F. (1992), "Strategic Alliances: Overcoming Barriers to Success*", Technology,* Vol. 81, pp. 56-61.

MARQUES, C. e MONTEIRO-BARATA, J. (2005), Determinantes de inovação e desempenho económico - financeiro: Ensaio para a Indústria Transformadora Portuguesa. Em *4th International Meeting of the Iberoamerican Academy of Management, Lisboa, Portugal, 8 a 11 de Dezembro, 2005.* Faculdade de Economia - Universidade Nova de Lisboa.

MARQUES, C. e MONTEIRO-BARATA, J. (2006), "Determinants of the Innovation Process", *Management Research,* Vol. 4, nº.2, pp. 113-126.

MANSO, J.R. (1998), "Curso de Econometria", Universidade da Beira Interior, Covilhã.

MUSCIO, A. (2006), "Patterns of Innovation in Industrial Districts: An Empirical Analysis", *Industry and Innovation*, Vol.13, nº.3, pp. 291-312.

MYTELKA, L. (2000), "Local systems of innovation in a globalized world economy", *Industry and Innovation,* Vol. 7, nº. 1, pp. 15-32.

NATÁRIO, M., COUTO, J., TIAGO, M. e BRAGA, A. (2007), "Determinantes da Capacidade Nacional de Inovação: Uma Análise à Realidade Europeia" em *Proceedings XVII Jornadas Hispano-Lusas de*

Gestão, Logroño, Espanha, 8 e 9 de Fevereiro, 2007. Ayala CALVO, J.C. y grupo de investigación FEDRA, Universidade de la Rioja, pp. 1650-1662.

NELSON, R. R. e ROSENBERG, N. (1993), "Technical Innovation and National Systems", em "National Systems Of Innovation: A Comparative Analysis", Nelson, R.R. (Ed.), Oxford University Press, pp. 3-21.

NELSON, R.R. e WINTER, S.G. (1982), "An Evolutionary Theory of Economic Change", Cambridge, Belknap Press.

NIETO, M. e SANTAMARIA, L. (2007), "The importance of diverse collaborative networks for the novelty of product innovation", *Technovation,* Vol. 27, pp. 367-377.

OCDE (2005), "OSLO Manual: Proposed Guidelines for Collecting and Interpreting Technological Innovation Data", Paris, OCDE.

PFEFFER, J. e SALANCIK, G. (1978): "The External Control of Organizations: A Resource Dependence Perspective", Harper & Row, New York.

PORTER, M. (1994), "Construir as Vantagens Competitivas de Portugal", 1ª Edição do Forúm para a Competitividade, Lisboa, Cedintec.

PORTER, M.E. (1998), "Clusters and The New Economics of Competition", Harvard Business Review, Nov-Dec., pp. 77-90.

PRAHALAD, C.K. e HAMEL, G. (1990): "The Core Competence of the Corporation", *Harvard Business Review*, May-June, pp. 79-91.

QREN (2007), "Quadro De Referência Estratégico Nacional - Portugal 2007-2013", Lisboa, Observatório do QCA II Ministério do Ambiente, do Ordenamento do Território e do Desenvolvimento Regional.

ROMIJN, H. e ALBALADEJO, M. (2002), "Determinants of innovation capability in small electronics and software firms in southeast England", *Research Policy,* Vol. 31, pp. 1053-1067.

SILVA, M. (2003) "Capacidade Inovadora Empresarial - Estudo dos factores impulsionadores e limitadores nas empresas industriais portuguesas", Tese de Doutoramento em Gestão, Universidade da Beira Interior, Covilhã.

SILVA, M.J. e LEITÃO, J. (2007), "What Determines the Entrepreneurial Innovative Capability of Portuguese Industrial Firms?" *Munich Personal RePEc Archive* Nº 5216.

SILVA, M.J. e LEITÃO, J. (2009), "Reinforcing the Entrepreneurial innovation capacity of industrial SMES with networks", em "Public Policies for Fostering Entrepreneurship: An European perspective", (Eds) Leitão, J. e Baptista, R., International Studies in Entrepreneurship, Vol. 22, Springer1.

SILVA, M.J., LEITÃO, M. e RAPOSO, M. (2008), "Barriers to innovation faced by manufacturing firms in Portugal: how to overcome it for fostering business excellence?" *International Journal of Business Excellence,* Vol. 1, nº 1-2, pp. 92-105.

Simões, V.C. (1996) "Inovação e Gestão em PME Industriais Portugueses", Relatório Técnico, Lisboa, GEP-CEDE.

TODTLING, F. (1995), "The Innovation Process and Local Environment", em "The Industrial Enterprise and its Environment: Spatial Perspectives", Conti, S. et al. (Ed.) U.K. Avebury, Aldershot, pp. 171-193.

TOURIGNY, D. e LE, C.D. (2004), "Impediments to innovation faced by Canadian manufacturing firms", *Economics of Innovation and New Technology,* Vol. 13, nº 3, pp.217 - 250.

UNCTAD (1998), "Promoting and Sustaining SMEs Clusters and Networks for Development", Geneve, United Nations, pp. 24.

USSMAN, A., ALMEIDA, A., FERREIRA, J., MENDES, L. e FRANCO M. (2001) "SME's and Innovation: Perceived barriers and behavioural patterns". *The International Journal of Entrepreneurship and Innovation,* Vol. 2, nº 2, pp.111-118.

ÁREA IV

ECONOMÍA DE LA ENERGÍA,
DEL AMBIENTE, DE LOS TRANSPORTES
Y DE LOS RECURSOS NATURALES

ECONOMIA DA ENERGIA,
DO AMBIENTE, DOS TRANSPORTES
E DOS RECURSOS NATURAIS

ECONOMIC TOOLS TO DESIGN EFFICIENT INTEGRATED WILDFIRE SUPPRESSION MANAGEMENT STRATEGIES

MARIA ISABEL DE DEUS MENDES
E-mail: midm@iseg.utl.pt
SOCIUS/ISEG
Technical University of Lisbon

Resumo

O cenário dos fogos florestais nos países do Sul da Europa e particularmente em Portugal tem piorado nos últimos anos e teme-se que a situação se agrave no futuro com o impacte das alterações climáticas previstas para a região. Este agravamento, conjugado com a existência de severas restrições financeiras e orçamentais e com a necessidade de afectar recursos escassos a outras actividades, sugere a premência de se adoptarem estratégias de combate aos fogos, economicamente eficientes. Neste artigo a teoria microeconómica do produtor é utiilizada para elaborar um modelo teórico de decisão, que permita escolher as estratégias de supressão mais produtivas e baratas. As soluções do modelo poderão ser operacionalizadas na prática se forem conjugadas com a utilização de técnicas de programação linear, bases de dados GIS e programas electrónicos de simulação de fogos florestais.

Palavras-chave: fogos florestais, supressão, planeamento eficiente, teoria do produtor.
Área Temática: Economia da Energia, do Ambiente, dos Transportes e dos Recursos Naturais.

Abstract

Within a backdrop scenario of resource scarcity, huge budget constraints, and existing alternative public choices with equivalent priority to that of wildfires prevention and suppression, new suppression wildfire management strategies seem to be necessary, to get efficient answers to questions arising from the increasing complexity of economic and technical Mediterranean forest fire risk management requirements. In this paper, we use traditional microeconomic theory of the producer to design a theoretical model enabling the choice of efficient wildfire suppression strategies. Along with linear programming techniques, GIS based data, and computer simulator programs, the solutions of the model can be applied and be used to improve the wildfire suppression process.

Key-Words: forest fires, combat, efficient planning, theory of the producer.
Thematic Area: Economics of Energy, Environment, Transports, and Natural Resources.

1. INTRODUCTION

Wildfires (a wildfire is a large destructive fire that rapidly spreads out of control, happening most frequently in the summer, when the brush is dry and flames can move unchecked through a wooded area; the fire often begins unnoticed and spreads quickly, lighting brush, trees and homes and it may be started by a campfire that was not doused properly, a tossed cigarette, burning debris, lightning or arson or other causes) has become in the last decades one of the biggest environmental problems Mediterranean forest is faced with. These occurrences are characterized by high levels of destruction of build, natural, and human capital, imposing to Mediterranean society a great financial burden. The nature and virulence of these events and related consequences are forcing the Mediterranean administrations to affect greater amounts of financial, human and material resources to the design and implementation of forest fire

prevention and suppression strategies. However, the results seem very far from the objectives traced within the frame of the fire management strategies. Forest wildfires continued to grow in number and extension, and so did the physical and monetary damages, particularly in the first 8 years of the current century (DGRF 2007). In spite of the increasing efforts on behalf of the administrations to meliorate qualitatively the prevention and combat measures undertaken, the wildfire scenario in Mediterranean regions has been worsen and one fears that the situation will be aggravated particularly within the framework of the present climacteric changes. Following this state of the art, one concludes that the methodology commonly used in the design and implementation of wildfire prevention and fighting strategies, seems no longer to be the more adequate. Equally it seems not reasonable or socially efficient, to simply continue to adopt a wildfire prevention and suppression management based on the *ad hoc* accruing of financial and other human and material resources, each time the backdrop wildfire scenarios turns worst. Firstly, because wildfires have a strong component of uncertainty which is measured in terms of the probability of ignition and the probability of existing conditions to the development of the fire. Secondly, because there is another important component of risk associated to forest fires: the heavy damage and level of destruction wildfire enhances (include forest ecosystem destruction and/or damage which implies the loss or damage of the stock of natural capital and of all direct and indirect benefits produced by this type of ecosystems, the destruction and damage over properties and infrastructures, as well as human losses and human health damages, psychological damages supported by the populations, and so on (Weiner 2001, Ramachandran1998). And finally, because the climacteric and social economic conditions that partially explain this type of occurrences in the present will persist and even aggravate in the future. Within a backdrop scenario of resource scarcity, huge budget constraints, and other socio-economic choices with equivalent priority to that of wildfire prevention and combat, new management wildfire strategies are necessary to respond efficiently, both from the economic and technical point of view, to the increasing complexity requirements of Mediterranean forest fires.

Till very recently, prevention and combat strategies have been traditionally designed by technicians coming from other scientific areas, different from economics (Riedout and Ziesler 2004). The absence of economic thought within these areas of management has been evidence, particularly till the last century. While prevention and suppression attempt to reduce the risk associated with forest fires, only two economic models have been around since the 1920's and generally designated as the Cost Plus Net Value Change (C+NVC) Model of Wildfire Economics (Sparhawk 1925 and Donovan and Rideout 2003, the latter an improvement on the former), focusing on minimizing suppression costs and fire damage as a trade off against prevention costs even if until 2006 none were empirically estimated (Lankoande 2006). This shortage of economic analysis (Rodriguez and Silva 2007, Riera 2005) can be partially explained by two major arguments. Firstly because of the existence of some lack of communication between different scientific areas of knowledge, which explains that those technicians with scientific formation different from economics, that usually design prevention and combat strategies, are not able to apply or even simply to understand how and where economic tools can be used to meliorate the forest fire combat. Secondly, economic approach to these matters is particularly data consuming, as a large bundle of variables in number and different type and nature is generally involved in the management process. However problems of scarcity and budget constraints one referred to earlier, points clearly to the direction of the increasing necessity of introducing economic thought into the process, in order to obtain more productive and efficient management of forest fires. To respond to these necessities, economic scientists have been developing empirical studies in order to use economic tools in the design and implementation of efficient and costless prevention and combat management forest fire strategies (Kline 2004, Riera and Mogas 2004, Rodriguez and Silva 2004, Loomis et al 2003, Prestemon et al 2001, Cleaves et al 2000).

In this paper, one wants to contribute to clarify how economic tools can be used to improve wildfire management planning in order to reduce fire risk and damage within a context of financial scarcity and probable raising of forest fire risk. We'll begin by defining and clarifying

the concept of wildfire risk's economic cost as decreasing risk is the main objective of forest fire prevention and suppression strategies (section 2). It will be shown that the concept has two components, one associated with a high level of uncertainty, and the other associated with the level of destruction and/or damage. Further, in section 3, we'll describe how the producer theory can be used, along with linear programming techniques, to design a theoretical model, whose solutions enable to estimate the more productive and the cheaper wildfire suppression strategies, per different wildfire scenarios. To apply the solutions of the model, sufficient data must be gathered in order to measure efficiently the variables involved by using geographic information systems – GIS – along with computer simulator programs (section 4). Finally Conclusions will be drawn.

2. HAZARD AND FURY: "LES BÊTES NOIRS" OF A FIRE SUPPRESSION STRATEGY PLANNER

As we outlined above, wildfire have three main characteristics that strongly affect and condition any type of management exercise. Firstly, these events are largely destructive and rapidly spread out of control. Secondly they characterize by high levels of destruction imposing a great financial burden to society, due not only to damages on properties and build infra-structures, but on human health and ecosystem related damage as well. And thirdly, wildfires have a strong component of uncertainty which is commonly measured in terms of the probability of ignition and the probability of existing conditions to the development of the fire.

These three characteristics turn any attempt to manage fire suppression strategies in an effective way, into a very rough exercise. To run efficiently the combat of a wildfire it is necessary to predict and detect locally its occurrence, to predict the way and the intensity of the fire spreading, and to predict as well the associated damages. The monetary value of potential damages, when a fire occurs, configures a very important type of information for administrations in order to help them to decide to what extent it will be socially and financially relevant to spend scarce resources in fire combats. Then, it is necessary to choose the better combination of resources to be used in the fire combat, which must be the most productive and the costless one.

Wildfire occurrences are uncertain and depend of two factors: the probability of existing ignitions and the existence of previous conditions that favours fire development. Ignitions may have a human or a non-human origin, although humans are the more common. In Portugal 98.8% of the ignitions are caused by humans, of which 35.4% are non-intentional, 27.4% are unknown, and 26.8% are caused by human negligence (DGRF 2006). It is difficult to estimate the probability of ignition occurrences because the set of possible causes is very large. Some authors like Bachmann and Algöwer (1998) defend that this type of probability will rise with the degree of geographic spreading of human activities: the greater the number and intensity of human activities, the greater the probability of the occurrence of ignitions. The previous conditions that favour fire development are basically related with the quantity and the quality of the existent fuel on the terrain. To less fuel corresponds a smaller opportunity to fire development.

The monetary costs associated to the risk of physical damages is clearly related with the probability of something be destroyed by a fire (the so called probability impact) and the monetary value under risk. Accordingly to Bachmann and Algöwer (1999), the impact probability depends on the way the fire spreads geographically. The fire spreading speed depends, in his turn, on the existent fire suppession and prevention strategies and on the behaviour of the fire itself. The only way to decrease the number and the intensity of wildfires and to reduce the monetary damages as well, -which is equivalent to diminish the wildfire risk - is to implement and/or improve prevention and combat actions. The main goal of prevention and combat actions is, so, that of minimizing the number of occurrences and the monetary value of damage and losses (Bennetton et al 1997). The main risk factors affecting the fire intensity and the level of damages associated to it, include local geographic and climacteric

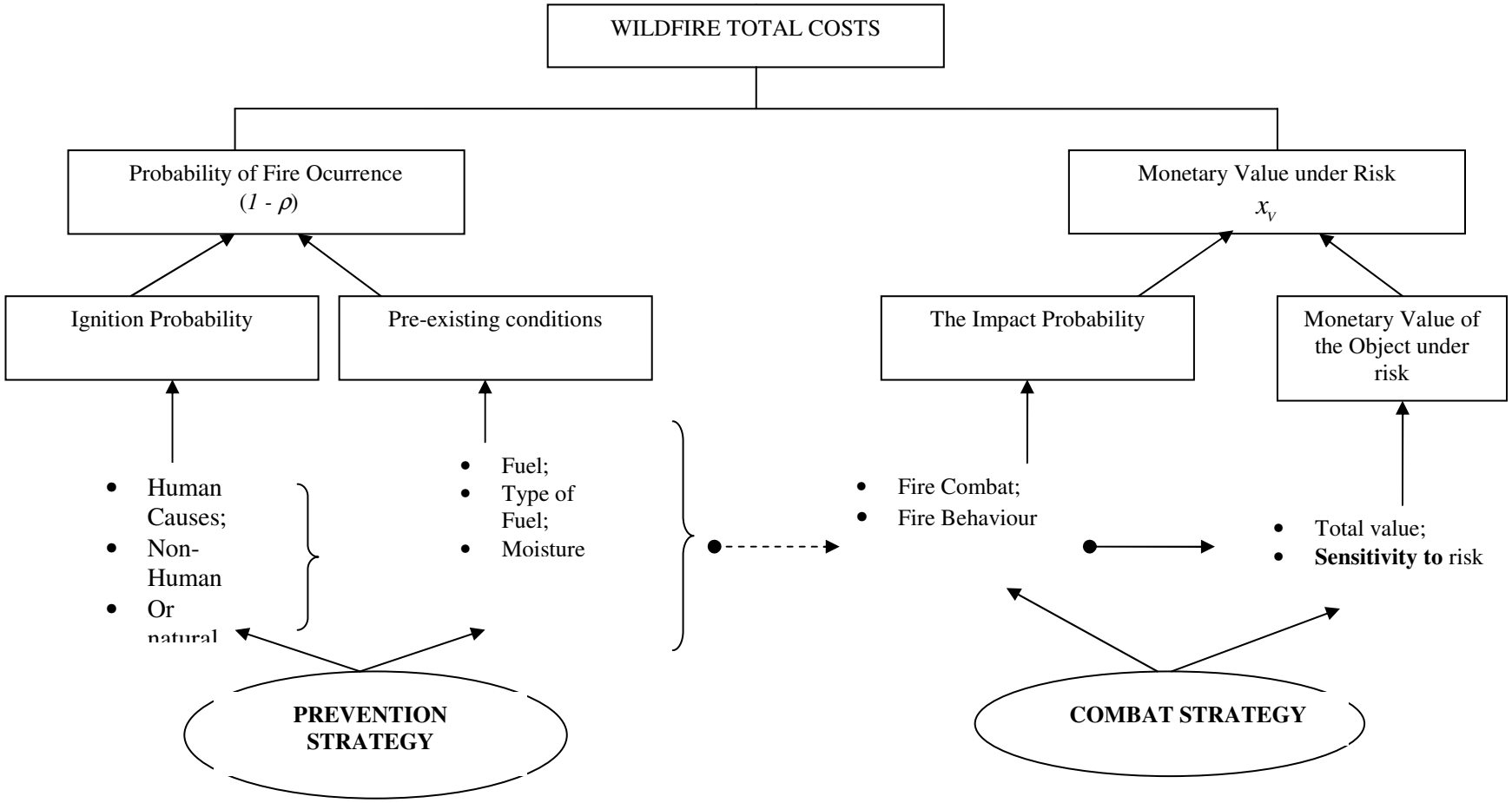
characteristics, the nature and the quantity of fuel on the terrain, and the efficiency of the already existent strategy of prevention and combat (Mercer and Prestemon 2005; Prestemon et al 2001). It is evident that among all these, fuel and the prevention and combat strategy are the risk factors over which it is easier and costless to intervene to decrease the risk of wildfire.

In *figure 1* are schematically represented the fire risk factors and the actions that have to be implemented to reduce the risk. It seems to be evident to the reader that fire risk expected value is a notion that incorporates two levels of analysis. The first is the monetary values under risk, their geographical dispersion and the assessment of the existing risk factors that can condition the fire occurrence probability. The second level of analysis is the design and the implementation of more productive and costless strategies to prevent and fight the wildfires, in order to reduce the level of both wildfire ignition and spreading probabilities. Therefore, if we take into account: *i)* there is a great number of prevention and suppression actions with different types of productivity and different total operational costs too; *ii)* the set of factors affecting the number and the intensity of wildfires is large and very heterogeneous; and *iii)* society has to take efficient decisions under severe financial, human and environmental restrictions, it is therefore reasonable to conclude that:

- It is necessary to design several prevention/suppression action scenarios to deal with the imperative of reduce wildfire risk and associated losses and damages;
- It is necessary to adopt economic methodologies to choose the more productive and costless fire combat strategy.

Till very recently, prevention and combat strategies have been traditionally designed by technicians coming from other scientific areas, different from economics (Riedout and Ziesler 2004). The absence of economic thought within these areas of management has been evidence particularly till the last century (Rodriguez and Silva 2007, Riera 2005), and such a situation can be partially explained by two major arguments Firstly because of the existence of some lack of communication between different scientific areas of knowledge, which explains that technicians that ordinarily design prevention and combat strategies, possessing scientific formation different from economics, are not able to apply or even simply to understand how and where economic tools can be used to improve the forest fire combat. Secondly, because economic approach to these matters is particularly data consuming, as a large bundle of variables in number, of different type and nature, is generally involved. However, problems such as scarcity and budget constraints outlined above, point clearly to the direction of the increasing necessity of introducing economic thought into the decision process, in order to obtain more productive and efficient management of forest fires. To respond to these necessities, economic scientists, together with others with distinct scientific formations, have being developing theoretical and empirical studies in order to introduce economic decision in the process of designing and implementing efficient and costless prevention and suppression management forest fire strategies (Kline 2004, Riera and Mogas 2004, Rodriguez and Silva 2004, Loomis *et al* 2003, Prestemon *et al* 2001, Cleaves *et al* 2000).

Figure 1 Fire Risk Factors and Fire Risk Integrated Management Actions



3. THE FIRE SUPPRESSION PRODUCTION DECISION PROBLEM

Following the microeconomic Producer Theory, wildfire forest suppression can be red as a production activity, where the output can be measured in physical units per unit of time (extinguish x hectares of burning hectares during a time period y or the designing of a controlled fire line of x meters during a time period y). The level of output depends of the type and number of material and human inputs used to the accomplishment of the task, and the way these inputs are combined. That is, the suppression output is a function of the wildfire forest combat technology used. The main planning fire combat's problem is how to choose the adequate bundle of fire fighting inputs (meaning the adequate technology) that must be at the same time the costless and the more technologically efficient. Theoretically this is the optimal bundle and is compatible with an economic Pareto equilibrium, i.e., it maximizes social well being, and there is no other bundle capable of improving even more general well being, without turning someone else worst.

The Wildfire Suppression Production Function

Let $\boldsymbol{q}$ be the output of the wildfire suppression production function measured, for instance, in terms of number of extinct burning hectares per time period. Let $\left(\boldsymbol{r}_i,\boldsymbol{r}_j\right)$ be the bundle of different *r* quantities of fire fighting inputs of type *i* ($\boldsymbol{r}_i$), and *r* quantities of fire fighting inputs of type *j* ($\boldsymbol{r}_j$). The fire combat planners combine the inputs per type and quantities to produce or to achieve the desired output. This process can be described by precise engineering formulas where the quantity and type of inputs combined with one another will be exactly specified. The end product, the output, can then be expressed as a function of all the inputs used in the fire combat. Such a function is the analytical form of the fire suppression production function and is expressed by an equation, whose specific mathematical form will be one that better describes the engineering process of fire combat.

One of the basic assumptions that economists make about production functions, is that they are technologically efficient. This means that, when inputs (firemen, helicopters, firemen's cars, tools, and others) are combined to fight a fire of several hectares along a certain period of time, a particular combination of inputs is said to be technologically efficient to achieve a certain suppression output if it is not possible to get the same output using less of an input or no more of any other input.

The general form of the production function of fire combat is given by:

$$\boldsymbol{q}=\boldsymbol{q}\left(\boldsymbol{r}_i,\boldsymbol{r}_j\right) \qquad \textbf{(1)}$$

where *(1)* represents a fire suppression technology, defined as different types and quantities of fire fighting inputs $\left(\boldsymbol{r}_i,\boldsymbol{r}_j\right)$ combined in a specific mathematical form (Cobb-Douglas Production Function, or CES Production Function for instance, or other specific mathematical type) to which several fire fighting outputs are associated: to raising quantities of input used to combat the fire, corresponds raising levels of output.

The first derivative of the suppression function in order of both inputs is positive - $\partial \boldsymbol{q}\left(\boldsymbol{r}_i,\boldsymbol{r}_j\right) \Big/ \partial \boldsymbol{r}_i,\boldsymbol{r}_j > 0$ - whose value is a measure of the Marginal Productivity of each input ($\boldsymbol{MgP}_{r_i,r_j}$). The $\boldsymbol{MgP}_{r_i,r_j}$ is the increment of number of hectares of fire extinguished associated to the unitary increment of the input *i,j* used in the fire combat and is a positive value, meaning that the number of extinguished burning hectares will raise with the number of fire combat inputs used. The second derivative of the production function is a negative value - $\partial^2 \boldsymbol{q}\left(\boldsymbol{r}_i,\boldsymbol{r}_j\right) \Big/ \partial \boldsymbol{r}_i^2,\boldsymbol{r}_j^2 < 0$ - which means that $\boldsymbol{MgP}_{r_i,r_j}$ will decreasingly rise as more fire fighting

inputs are added to combat a specific fire. These two properties of a production function are generally referred, within the production theory, to the law of diminishing returns. Accordingly to it, the first units of fire combat inputs will have a higher productivity, measured in number of hectares extinguished; however, there is a specific degree of concentration of fire combat inputs on the terrain which, if exceeded, will generate diminishing increases of the fire fighting results instead of increasing or constant ones. This is equivalent to say that fire combat inputs have diminishing productivities. If the fire combat's input prices remain constant, the existence of diminishing returns signifies that fire combat total costs will increase exponentially while the fire costs saved as a result of the fire output combat will be lesser, if a "more forest fire/more fire combat inputs" strategy is adopted. As a consequence, society will support a raising financial burden associated with the inefficiency of an expensive fire combat policy.

If one divides the level of output given by the production function in *(1)* by the total quantity of each type of fire fighting input, the planner will get a measure of the average productivity of the inputs used, i.e., the planner will get the median productivity of each input - $\boldsymbol{MeP}_{r_i,r_j}$ - calculated by equations *(2)* and *(3)* respectively for the inputs of type *i* and *j*:

$$MeP_{r_i} = \frac{q\left(r_i, r_j\right)}{r_i} \quad (2) \; ; \; MeP_{r_j} = \frac{q\left(r_i, r_j\right)}{r_j} \quad (3)$$

The concept of fire suppression production function enables the planner to compare each of all the existing fire fighting technologies to the others, by using their technological efficiency as the main criterion of ranking, accordingly to the level of fire fighting outputs and the productivity of each of the inputs used.

If the planner wants to know, what and how much of fire fighting inputs may be used to get a specific level of fire fighting output, the isoquant is the answer. An isoquant is a line that represents the bundle of all input combinations $\left(\boldsymbol{r}_i, \boldsymbol{r}_j\right)$ that can be used to perform a specific level of fire fighting output $\bar{\boldsymbol{q}}$ one wants to achieve. If the planner wants to know what and how much of fire fighting inputs have to be used to achieve several pre-setting levels of fire fighting outputs, then isoquant map is the answer. The isoquant map is defined as the set of isoquants representing different previously defined levels of fire fighting outputs. Analytically, one can get the general formula of the isoquant map by setting equation ***(1)*** equal to some cardinal – $\boldsymbol{k}$ - that represents some level of fire suppression output one wants to perform and then solving for the input $\boldsymbol{r}_j$, like in equation ***(2)***:

$$\boldsymbol{q} = \boldsymbol{q}\left(\boldsymbol{r}_i, \boldsymbol{r}_j\right) = \boldsymbol{k} \Leftrightarrow \boldsymbol{r}_j = \boldsymbol{f}\left(\boldsymbol{r}_i, \boldsymbol{k}\right) \qquad (2)$$

In short terms, the map of isoquants is a common method used by economists to rank different type of fire fighting technologies accordingly to the level of outputs they generate. If the ranking process follows the three properties mentioned below:

i) all the possible input combinations $\left(\boldsymbol{r}_i, \boldsymbol{r}_j\right)$ may be ranked;

ii) all the combinations are transitive, that is if combination A is preferred to combination B which is preferred to C's, then A is preferred to C;

iii) combinations using more inputs are preferred to those using less inputs, because output will be smaller,

Then, isoquants are said to be well behaved, which geometrically means that they can be represented by convex lines, as those shown in *figure 2*:

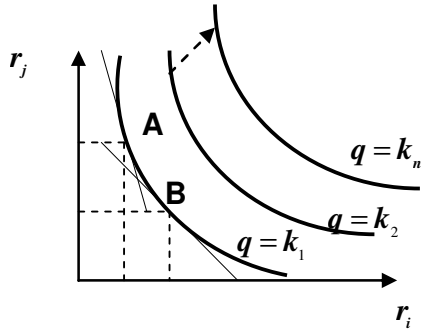


Figure 2 Fire Combat Isoquant Map

In *figure 2*, each isoquant represents a set of all combinations of fire combat input quantities that can be used, to achieve the same level of fire extinction, represented by some q equal to a constant k, where $k_i = k_1, k_2, ..., k_n$. For instance, the level of fire extinction k_1 ($q = k_1$) can be performed by using the combination of fire combat inputs *A* or *B*, or even any other, as long as it belongs to the isoquant $q = k_1$; k_n is the *n* highest level of output.

The property concerning the convexity to the origin of the isoquants means the existence of decreasing marginal rates of technical substitution of one input by the other (MRTS), while keeping constant the level of output. Geometrically, this concept is defined as the negative of the slope of an isoquant, and analytically, as the ratio of the marginal product of the fire fighting inputs as given in *(3)*:

$$MRTS = \frac{MgP_{r_i}}{MgP_{r_j}} = \frac{\dfrac{\partial q(r_i, r_j)}{\partial r_i}}{\dfrac{\partial q(r_i, r_j)}{\partial r_j}} \qquad (3)$$

Decreasing MRTS economically means that as the quantity of input i - r_i - , increases all along an isoquant, the marginal product of r_i must decline relative to the marginal product of r_j, in order to maintain constant the level of output.

By using the concept of fire fighting production function, the planner will also be able to compare the rhythm of variations of the quantities of all the inputs used in a fire combat with the correspondent rhythm of variations of the outputs. This particular relationship is designated by returns to scale, accordingly to the terminology of the production function theory, and is a long run concept where all inputs like the number of firemen or the number of fire combat vehicles, vary (while in the short run some inputs like firemen's quarters are fixed). As output expands in the long run, fire fighting production functions may exhibit the property of homogeneity. There are three types of homogeneous production functions. If doubling all the inputs exactly doubles, less than doubles, or more than doubles the output, we say the production function exhibits constant, decreasing, or increasing returns to scale - in other words, multiplying all the inputs by the same positive constant, multiplies output by the same constant, an inferior constant, or a greater constant. Analytically, the degree of homogeneity of the production function is known when one multiplies all the inputs of the production function in *(1)* by a positive constant λ such that:

$$q(\lambda r_i, \lambda r_j) = \lambda^{\beta} q(r_i, r_j)$$

The fire fighting production function has constant, increasing, or decreasing returns to scale if $\beta = 1$, $\beta > 1$, or $\beta < 1$, respectively.

The Costs of wildfire suppression

Economists can estimate fire fighting production functions trough the observation of choices of the quantities of inputs, per type of input, used to combat wildfires on the terrain, and assuming that, to carry out a certain fire fighting output, planners always choose combinations of fire fighting inputs that are both technologically efficient and cost minimizing. The assumption that input's choices are cost minimizing for a given set of fire fighting outputs is referred to as an assumption of economic efficiency in contrast with the assumption of technological efficiency outlined above. Any input combination used to achieve a certain level of output is economically efficient, if it is not possible to produce that level of output at a lower cost, given the prevailing input prices. Assuming the input prices are given (they are set in competitive markets and the producers do not get in the way of the process, which means producers are price takers) and setting the level of output, the wildfire combat planner can use a wildfire combat production function to find the least-cost way of doing it.

To do so, firstly the planner has to rank the costs society has to support with wildfire combat. There are two main bundles of costs to be considered. The planner has to buy or rent the inputs necessary to combat a specific fire during a certain period of time, and the operational cost will simply be equal to the sum of all input's expenditures. This, is the first bundle of fire combat costs. If the planner wants to extinguish the fire in a small period of time, more inputs have to be allocated, the higher will be the total operational cost of fire combat, and the less will be the number of hectares of forest burnt. Alternatively, if the planner wants to extinguish the fire in a greater period of time, less input will be allocated, total operational cost with fire combat will be lower, but more hectares of forest will be burned.

The costs related with the hectares of forest meanwhile burned during the fire fighting combat, is the second bundle of fire combat costs the planner has to consider when assessing the overall costs of each fire combat strategy.

It seems evident that there is a clear trade-off between the costs of the inputs used in the fire combat and the costs supported by society related with the hectares of forest burned during the fire combat period. The more inputs are used, the more expensive the fire fighting strategy will be from the operational cost point of view, but the value of the losses in natural and built capital avoided will be higher. On the contrary, if planners, for the sake of budget restrictions, reduce the inputs used to fight the fire, the operational costs of the combat strategy will be less but the losses due to the fire will obviously increase. This means that the use of more inputs to fight a fire is not necessarily synonymous of raising costs to society, because the costs will be automatically compensated by the savings in terms of avoided damages.

Formally, let SRTC and LRTC be the short run and the long run total costs of a fire combat strategy respectively, VC the variable costs associated with the fire combat inputs whose value depends on the output q, FC the fixed costs that do not depend of the output q. The total cost of some fire fighting strategy used to achieve a certain output q is given by *(4)* and *(5)* for the short and the long run respectively:

$$SRTC(q) = VC(q) + FC = p_i r_i(q) + p_i \bar{r}_j + c(1-q) \quad (4)$$

$$LRTC(q) = p_i r_i(q) + p_i r_j(q) + c(1-q) \quad (5)$$

where $\bar{r}_j$ is the input that does not vary in the short run, p_i and p_j are the prices of the inputs, c is the value of damages per each hectare of forest burnt, and $(1-q)$ is the hectares of forest lost during the fire fighting.

By dividing *(4)* and *(5)* by q, the average cost (*ASRC(q)* and *ALRC(q)* respectively for the short and the long run respectively) of each additional hectare of fire extinguished, per fire, will be known. This information is given by equations *(6)* and *(7)* for the short and the long run respectively.

$$ASRC(q) = \frac{SRTC(q)}{q} = \frac{VC(q) + FC}{q} = \frac{p_i r_i(q) + p_i \bar{r}_j + c(1-q)}{q} \quad (6)$$

$$ALRC(q) = \frac{LRTC(q)}{q} = \frac{p_i r_i(q) + p_i r_j(q) + c(1-q)}{q} \quad (7)$$

It is also possible to know the total cost of each additional hectare of fire extinguished per period of time, or the marginal cost of q (*MSRC(q)* and *MLRC(q)* respectively, for the short and the long run), by taking the derivative of *(4)* and *(5)* in order to q. The results are given in *(8)* and *(9)*.

$$MSRC(q) = \frac{\partial\left[p_i r_i(q) + p_i \bar{r}_j + c(1-q)\right]}{\partial q} = \frac{\partial VC(q)}{\partial q} - c \quad (8)$$

$$MLRC(q) = \frac{\partial[LRTC(q)]}{q} = \frac{\partial\left[p_i r_i(q) + p_i r_j(q) + c(1-q)\right]}{q} = \frac{p_i \partial\left[r_i(q)\right]}{\partial q} + \frac{p_j \partial\left[r_j(q)\right]}{\partial q} - c \quad (9)$$

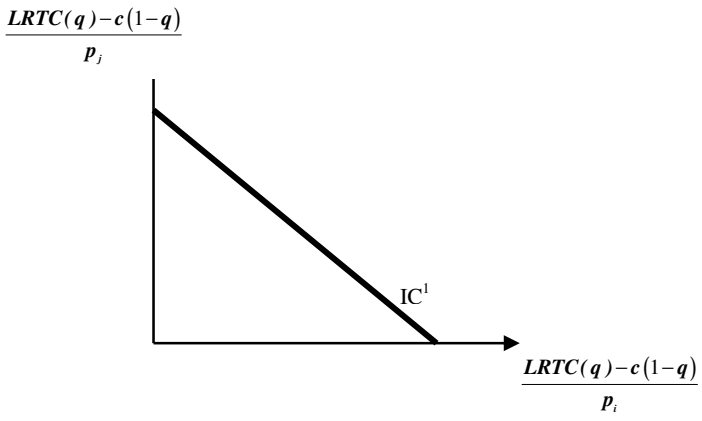
By *(8)* and *(9)* one concludes that the additional cost of one more hectare of fire extinguished per period of time has two components: one positive related with the cost of the additional units of input fire combat, and the other, negative, related with the damages saved by extinguishing one more hectare of fire.

Holding *TC* and the input prices fixed and varying r_i and r_j to satisfy the equation *(5)* tells the planner all the combinations of inputs that can be hired for a certain total cost of fire suppression given the prices of the inputs, p_i and p_j. All the input combinations with se same cost are given by a linear equation called the isocost line. Analytically, the isocost line is obtained by resolving equation *(5)* in order to r_j as given in *(10)*:

$$r_j(q) = \frac{LRTC(q) - c(1-q)}{p_j} - \frac{p_i}{p_j} r_i(q) \quad (10)$$

The line *(10)* is a linear function and can be represented in a $r_i - r_j$ graph as illustrated in Figure 3.

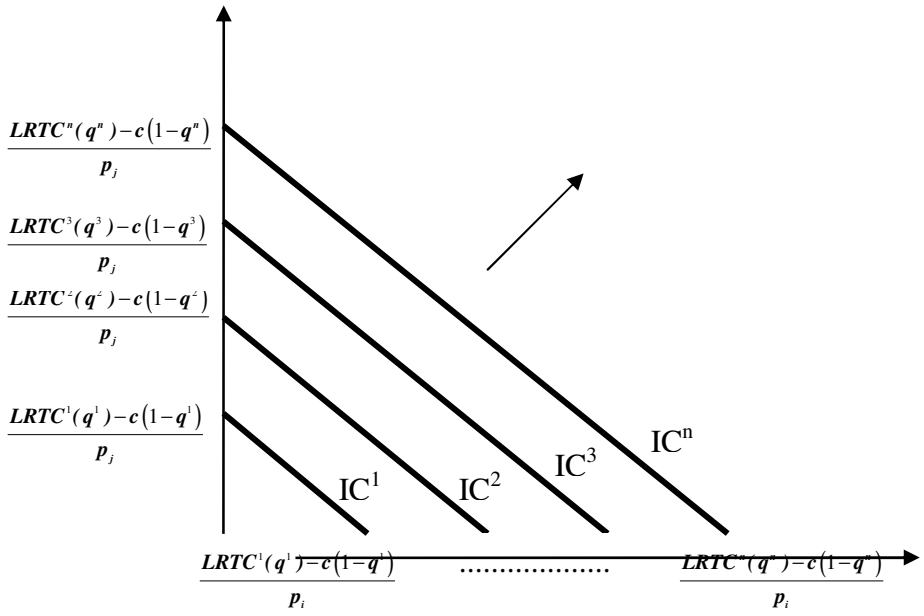
Figura 3 An Isocost Line



The slope of this line is the negative of the ratio of the input prices, $-p_i/p_j$ - and gives the opportunity cost of using one more unit of input r_i measured in terms of the decreasing of the use of the other input r_j, in order to maintain the same level of *LRTC(q)*. The intercept is

$\frac{LRTC(q)-c(1-q)}{p_j}$. If *LRTC (q)* assumes different values, it is possible to estimate several isocost lines (IC^1, IC^2, …, IC^n) as given in *figure 4*. Each of the *IC* lines represents combinations of inputs of fire combat with the same cost.

Figura 4 A Map of Isocost Lines



The arrow in *figure 4* indicates the trajectory of dislocation of the *IC* lines and of the total costs of combating wildfires, as more inputs are allocated. IC^n represents the more expensive input combinations with the same total cost $LRTC^n$, while IC^1 represents the costless combinations corresponding to the cost $LRTC^1$.

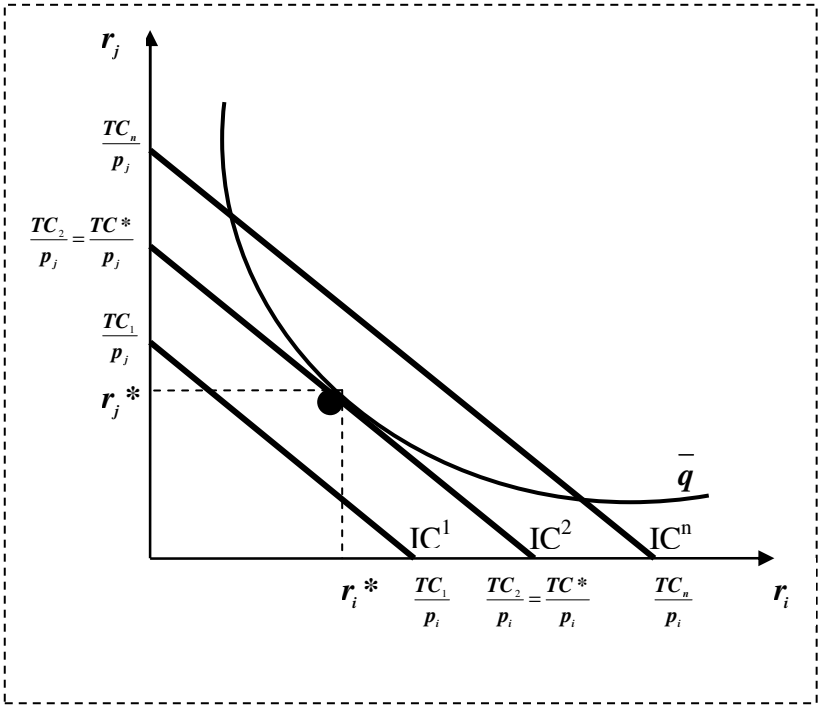
Choosing the more efficient combination of fire suppression inputs

The problem of knowing which is, from the economic point of view, the efficient combination of fire fighting inputs to combat fires, is similar to the problem of choosing the input combination that is more technologically efficient (the more productive), and the costless as well.

As outlined above, wildfire suppression functions represent sets of technologically efficient bundles of fire combat inputs; cost functions represent the total costs society has to support with the use of those sets of bundles, for every level of fire fighting performance. To choose the more efficient input combination is similar to choose the more technologically efficient input bundle and the costless as well, to achieve a predetermined level of fire fighting output. This is the so called cost minimization problem. If we assume that the planner has a fire fighting suppression function that can be described by isoquant lines convex to the origin as those outlined in *figure 1*, and that fire fighting input market prices are parameters in the planner's decision problem, the minimization cost problem for each level of fire fighting output is resolved every time the fire combat planner adjusts its fire combat inputs mix until the technologically determined marginal rate of technical substitution equals the predetermined price ratio for those inputs.

The cost-minimizing fire combat input combination $\left(r_i^*, r_j^*\right)$ for a predetermined fire combat output $\bar{q}$ is geometrically founded at the point of tangency between the isoquant $\bar{q}$ and the isocost line at *TC** (figure 5). This tangency condition can be interpreted as follows:

Figure 5 Cost Minimization



$$slope\ of\ \ isoquant = \frac{dr_j}{dr_i} = -\frac{p_i}{p_j} = slope\ \ of\ \ iso cos t\ \ line$$

Or

$$MRST = -\frac{dr_j}{dr_i} = \frac{p_i}{p_j} = \frac{MgP_i}{MgP_j} \qquad (11)$$

Equation *(11)* summarizes the cost-minimization condition, and can be viewed as the rule of thumb for optimal choice of fire fighting inputs, both from the technological and the economic points of view, in order to achieve a predetermined fire combat result.

Analytically, the optimal bundle choice $\left(r_i^*, r_j^*\right)$ is the solution of the minimizing cost problem outlined in *(12)*:

$$\min_{\left(r_i^*, r_j^*\right)} \quad p_i r_i + p_j r_j + c\left(1-q\right) \qquad (12)$$
$$subject\ to \quad q\left(r_i, r_j\right) = \bar{q}$$

where $q\left(r_i, r_j\right)$ is the fire fighting production function, specified by the analytical expression that better describes the engineering process of a fire combat; $\bar{q}$ is the level of fire combat output to be achieved and is predetermined; and $p_i r_i + p_j r_j + c\left(1-q\right)$ is the cost of doing it. We wish to minimize the cost of producing $\bar{q}$, given the prices of the fire combat inputs $\bar{p}_i$ and $\bar{p}_j$.

By using optimal programming techniques, the summarized first-order conditions for this problem are:

$$\begin{cases} \dfrac{MgP_i}{MgP_j} = \dfrac{p_i}{p_j} & (13) \\ q(r_i, r_j) = \bar{q} & (14) \end{cases}$$

where *(13)* is the cost-minimization condition given in *(11)* and *(14)* is the technological restriction or the fire fighting production function. Solving the system above in order to the objective variables (r_i, r_j) gives *(15):*

$$\begin{aligned} r_i^* &= f_i(p_i, p_j, \bar{q}) \\ r_j^* &= f_j(p_i, p_j, \bar{q}) \end{aligned} \qquad (15)$$

The equations *(15)* are the solution of the cost-minimization problem defined by *(12)*, and they represent the optimal choice of r_i and r_j to perform $\bar{q}$, both from the technical and the economical points of view. The general form of the functions given by *(15)*, named input conditioned demand functions, means that the optimal quantity of suppression inputs (r_i^*, r_j^*) depends of the input prices, and of the suppression strategy, *i.e.* the suppression technology (the technology).

Substituting *(15)* in the objective function *(12)*, we estimate the minimum cost of $\bar{q}$, given the input prices and the technological restrictions to combat the fire. If we want to know the minimum cost of achieving different levels of $\bar{q}$, we only have to change those levels in *(12)* and repeat the calculus.

4. APPLYING THE FIRE SUPPRESSION PRODUCTION DECISION PROBLEM: THE USE OF GIS AND SIMULATOR PROGRAMS

In practical terms, some countries have being evaluating their wildfire combat programmes by using methodologies with a theoretical structure similar to the outlined above: they are based on marginal analysis, linear programming, and the estimation of functions. In practice, these tools reveal to be very useful to implement the task of planning efficiently the wildfire combat strategy, particularly within a backdrop scenario of accruing hazard, rising costs, and budget restrictions.

There is, however, a magna problem concerning the application of these methodologies: they are very data dependent. Not only extensive lists of different variables measured in physical as well as in monetary terms are necessary, but also simulations of wildfire behaviour are an imperative to know the impacts that each of the fire combat inputs (evaluated either in isolation or in combination with others) has on the development of a specific wildfire. For instance, to estimate the fire suppression production function in order to assess the more productive input fire fighting techniques, within a predetermined period of time, and for a wildfire with specific characteristics, we have to simulate the wildfire behaviour with and without intervention, to measure and quantify in physical terms the quantity of inputs used, and their productivity in terms of fire extinction as well. To estimate the cost functions of fire suppression, not only data concerning the quantity of input resources used during the combat are necessary, but also the prices of the inputs and the monetary evaluation of the economic damages associated with wildfires are obligatory as well. Besides the data needs we just outlined, there is another important problem concerning specifically the monetary evaluation process of wildfire associated damages in terms of losses of natural, build, and human capital which constitutes, by itself, a difficult and very exigent task; this process is quite demandable in terms of data bases and in terms of the use of sophisticated technical evaluation methodologies

(see more details in Mendes *forthcoming*). Rodriguez y Silva (2007) made an inventory of some of the data bases that, in his opinion, are necessary:

- Identification of the geographical unit under analysis: natural protected area, hydro graphic basin, forest stain, territorial administrative unit, region, country, etc;
- Data sufficient enough to characterize the geographical unit under analysis in geo-physical, climacteric and socio-economic terms, including inventories of all the natural, build, and human units of capital that it contains;
- Analysis of the wildfire behaviour at the ignition phase, in the absence of intervention;
- Cartography of all type of existing fuels and of the forest and types of forest as well;
- Number of fires per year and per levels of intensity and duration;
- Wildfire speed propagation;
- Dimension in hectares of the surfaces affected at the moment of the ignitions;
- Historical of the incidents in geographical terms (by using GIS for instance);
- Historical of the climacteric conditions in median terms, per geographic units;
- Assessment of the potential losses and damages measured in physical and monetary terms, with and without fire combat;
- Data to plan the efficient use of resources in the fire combat: number and types of resources and productivity of the resources per types of fuel models;
- Data for the assessment and characterization of already existing fire system's detection and fire combat system;
- Data to evaluate the times of arrival of the resources of combat after the fire detection phase;
- Fire frequencies per type of intensity;
- Comparison between the speed of development of the fire control line and the speed of development of the fire perimeter;
- Operational costs associated with the combat interventions, per hectare;
- Number of fires with a development far beyond the expected;

By reading this list of information, one confirms immediately that the appliance of economic techniques to improve the fire suppression planning decision and management process is particularly demandable in terms of data. But a closer look at the list, highlighted the idea that data exclusively needed by economists, and by no other technicians, are those concerning the quantification of variables measured in monetary terms, particularly monetary costs of the resources used in combat and the monetary value of the damages and losses avoided with the intervention. All the others are data needs common both to economists and engineers. If this mix of physical and monetary data needed to assess fire risks, fire behaviour, and the impacts of combat strategies, exists, economists will then be able to use their economic models to make efficient choices. In short, the active and effective participation of economists to improve the process of fire suppression strategy planning is exclusively dependent on existing data; and these data can only be gathered by other technicians, but economists. This means that new wildfire combat integrated management strategies must be compulsorily an interdisciplinary exercise, which is the contrary of the planning fire combat processes that are more commonly put into practice (particularly in Portugal).

The main question one has to set now is how to assess the data. The answer can be found by analyzing the empirical experiences that have been put into practice by some countries. In the last few years some of the countries more exposed to this kind of natural disasters, and very aware of the huge monetary costs they provoke, have being developing and applying new integrated management practices based in models that use economic decision techniques together with others, used to simulate wildfire behaviour. Such models simulate the effects of the combat strategies per type of fire and per period of time of the extinguish phase. By combining several techniques like computer simulation programmes and GIS techniques, together with economic evaluation of losses and damages, these integrated models provide the sufficient data bases that allows the choice of the more productive and costless fire combat resources in efficient terms, the integration of cost-benefit analysis, the modelling of the

probability of the occurrences, and the evolution of line fires as well, with and without intervention.

Examples of these models include the Fire Protection Analysis (FPA) de 2003 applied in USA, the Level of Protection Analysis System (LEOPARDS) used in Canada, the *Sistema Nacional para el Manejo de Incêndios Florestais* (SINAMI) plus the program of economic analysis the ECOSINAMI used in Spain (these two were concluded at the end of 2006). Other examples of this integrated management models include the Fire Management Decision Support Systems (FFMDSS), KITRAL, California Fire Economics Simulation Model (FPPS/CFES), FirePro, ARMS/ADFF, ARCAR41, Rideout Model and Wildfire Initial Response Assessment System (WIRAS). All of these integrated models of planning use economic analysis methodologies to assess the value of the resources used to combat wildfires as the value of the losses and damages as well, to minimize the intervention costs and to choose the more productive combat fire strategies.

From the Portuguese point of view and in the absence of this type of integrated models, it is interesting to know in more detail the SINAMI and the ECOSINAMI models used in Spain because they were elaborated to resolve the specific problems of planning efficiently the wildfires combat concerning Mediterranean forest stains, which is a need common to the two Peninsula countries. Rodriguez y Silva (2007) describes shortly the SINAMI model and its potentialities. The model was elaborated to put into practice integrated management of Mediterranean forest stains, and combines the economic analysis of the operational costs of wildfire prevention and combat measures and related productivity, with the economic valuation of wildfire associated damages and losses. Accordingly to Rodriguez *op. cit* the SINAMI allows: *i)* the identification of the more efficient sets of fire combat strategies (that is the fire production function, to use the theoretical lexical of *section 3*), per type of fire and per type of period of time; *ii)* the simulation and measure of the productivity of human and non-human combat resources, used within the context of different scenarios of fire spreading designated as Fuel Models: to frame the scenarios and to define the Fuel Models, a GIS was applied to collect the necessary physical data to characterize the geographical unity under analysis in meteorological and topographic terms, as well as the type of local vegetation and its moisture level: to assess the productivity of the resources used in the fire combats per fuel model, computer simulations has been used; *iii)* allows to choose the costless and the more productive resources combination to built a controlled fire perimeter on one certain period of time that minimizes the economic fire damages and losses.

Some examples of data obtained from the SINAMI model are presented in the following tables. *Table 1* shows the productivity of different type of resources used in wildfire combat, per fuel model type, measured in meters per minute. For instance, and following Rodiguez y Silva (*op. cit.*) the fuel model nº5 corresponds to the more common fuel combustion model in the geographical region under analysis, and is characterized by dense and creeping bushes. With these data it is possible to measure the monetary operational costs of the combat strategies, per type of fuel model, and assess the total cost functions of fire combat (see *section 3*). *Table 2* shows the number of hectares burnt and the perimeter of the controlled fire line, per each period of time, for the fuel model nº 5. These numbers were set by considering the more productive combination of combat resources chosen after the productivity results given in *Table 1*. Combining the information of Tables 1 and 2 it is possible to set the more productive combination of combat resources, to extinguish a fire in a region characterized by a fuel model nº 5, per different periods of time. The results are in *Table 3*. With these data we could obtain the isoquanta map (see *section 3*) for the fuel model nº 5. Note that if we add to the operational costs of some combination of combat resources, the monetary costs of the hectares burnt during the combat, we obtain the total cost of fire combat (see *section 2*), per type of combat resources combination and period of time.

If we combine all these information accordingly to the algorithm described in the *section 3*, the planner will be able to choose the more efficient (that is the more productive and the costless) combination of resources to combat a fire in a particular local under specific fuel characteristics, per period of time.

Table 1 Productivity of Fire Combat Resources

Fuel Model	Firemen Brigades m/min	Extinguishing Groups m/min	Vehicles m/min	Bell 212 m/min	Tractor D6 m/min	Anfib Vehicle m/min	ACT802 m/min	Bell412 m/min
1	15	11	25	95	65	120	100	95
2	11	8	20	95	65	120	100	95
3	10	6	20	85	45	110	90	85
4	5	4	10	25	25	40	30	25
5	8	6	18	45	45	60	50	45
6	7	5	10	30	40	50	45	30
7	7	5	9	20	35	35	25	20
8	7	5	8	50	15	90	70	50
9	6	5	8	50	15	90	70	50
10	6	4	7	45	15	80	65	45
11	5	3	7	35	12	45	35	35

Font: Rodriguez y Silva (2007)

Table 2 Number of Hectares Burnt and Perimeter of the Controlled Fire Line per Period of Time (Fuel Model nº5)

Period of Time (hours)	Burnt Area (hectares)	Perimeter of the controlled fire line (meters)
2	92.3	4213.4
4	369.3	12 640.0
6	830.8	12 640.0
8	1477.0	16 853.5

Font: Rodriguez y Silva (2007)

Table 3 Combinations of Combat Resources (fuel model nº 5) Considering Four Time Periods

Period of Time	2 hours	4 hours	6 hours	8 hours
Type of Resources Combination	**1**	**1**	**3**	**4**
Machinery	-	1	1	1
Specialist Group(7 persons)	1	1	-	1
Specialist Brigade (12 persons)	-	-	1	1
Water Vehicles	1	-	1	1
Semi-heavy helicopter Bell412 (transport and extinction)	-	-	-	1
Backing helicopter of transport and extinction	1	-	-	-
Semi-heavy load plane (AT-802)	-	-	1	-
Hydroplane (CL215T)	-	-	-	1
Operational Costs € (dislocation + extinction)	2539.24	1291.68	9789.30	22778.51

Font: Rodriguez y Silva (2007)

5. CONCLUSION

Wildfires are characterized by high levels of destruction of build, natural, and human capital, imposing to Mediterranean society a great financial burden due not only to damages on properties and build infra-structures, but human wealth and ecosystem related damages too.

Wildfire occurrences are uncertain and they depend of two factors: the probability of existing ignitions and the existence of previous conditions that favors fire rushing. The only way to diminish the number and the intensity of wildfires and to reduce the monetary damages is to implement prevention and combat actions.

Till very recently, prevention and combat strategies have been traditionally designed by technicians coming from other scientific areas, different from economics. The absence of economic thought within these areas of management has been evidence, particularly till the last century. However problems of scarcity and budget constraints are pointing clearly to the direction of the increasing necessity of introducing economic thought into the process, in order to get more productive and efficient management of forest fires. To comply with these needs, economic scientists have being developing recently theoretical and empirical studies that use economic tools to design and implement efficient and costless prevention and suppression management forest fire strategies. In this paper we describe how the microeconomic producer theory can be used along with linear programming techniques in order to choose the better and cheaper way to combat wildfires, per each wildfire type. We describe how economists, by using the theory of the producer, are able to design a theoretical model to choose more efficient (meaning the more productive and the costless) strategies to combat wildfires, within a backdrop scenario of high uncertainty and financial restrictions. Further, we also demonstrate how this theoretical model and solutions may be fully put into practice, if adequate data bases exist. However, data are dependent of other technicians different from economists, which mean that if administrations want to dispose of a more adequate operational methodology to minimize the risk of wildfires and the suppression costs, they will have to incentive the constitution of multidisciplinary research teams.

This paper is one of the scientific outputs of the project *Fogos Florestais: Factores de Risco Associados a Danos Económicos e Ambientais* (*Forest Fires: Environmental and Economic Risk Factors*), coordinated by the Technical University of Lisbon, and financed by *Banco Santander Totta* and *Fundo Florestal Permanente*. Financial support received from the Fundação para a Ciência e Tecnologia (FCT) under FCT/POCTI, partially funded by FEDER is also gratefully appreciated.

REFERENCES

BACHMANN A. AND ALLGÖWER, B., THE NEED FOR A CONSISTENT WILDFIRE RISK TERMINOLOGY, 2000. In: Neuenschwander, L.F., Ryan, K.C., Golberg, G.E., Greer, J.D. (Eds.), Proceedings of the Joint Fire Science Conference and Workshop Crossing the Millennium: Integrating Spatial Technologies and Ecological Principles for a New Age in Fire Management, Boise Idaho, Jun 15 – 17 1999, University of Idaho, 67-77.

BACHMANN A. AND ALLGÖWER, B. (1998). "Framework for Wildfire Risk Analysis", in Viegas, X. (ed). "III International Conference on Forest Fire Research", 16 – 29 November, Universidade de Coimbra (ADAI), Luso.

BENNETTON J., CASHIN, P., JONES, D. AND SOLIGO, J. (1997). "An Economic Evaluation of Bushfire Prevention and Supression", *Research Paper nº 598,* The University of Melborne, Department of Economics, Victoria. Cleaves, D.A., Martinez, J., and Haines, T.K. 2000. *Influences on Prescribed Burning Activity and Costs in the National Forest System*. USDA Forest Service, Southern Research Station.

COOL, T. 2002. Proper Definitions for Uncertainty and Risk. *EconWPA Paper 9902002*.

DGRF 2006. *Incêndios Florestais: Relatório de 2005*. Direcção-Geral dos Recursos Florestais: Lisboa.

DGRF 2007. *Estratégia Nacional para as Florestas*. Direcção-Geral dos Recursos Florestais:Lisboa.

DONOVAN, G.H. AND RIDEOUT, D.B. 2003. A Reformulation of the Cost Plus Net Value Change (C+NVC) Model of Wildfire Economics. *Forest Science*, **49**(2): pp 318-323.

KLINE, J.D. 2004. *Issues in Evaluating the Costs and Benefits of Fuel Treatments to Reduce Wildfire in the Nation's Forests.* US Department of Agriculture, Forest Service, Pacific Northwest Research Station.

LANKOANDE, M. AND YODER, J., 2006. An Econometric Model of Wildfire Suppression Productivity. *Working Paper Series WP 2006-10*, School of Economic Sciences, Washington State University.

LOOMIS J., WOHLGEMUTH, P., GONZÁLEZ-CABÁN, A., AND ENGLISH D., 2003. Economic Benefits of Reducing Fire-Related Sediment in Southwestern Fire-Prone Ecosystems. *Water Resources Research,* **39**(9): p. 1260.

MENDES, I. (forthcoming). *A Avaliação Económica dos Fogos Florestais (The Economic Evaluation of Forest Fires), in Fogos Florestais: Factores de Risco Associados a Danos Económicos e Ambientais* (*Forest Fires: Risk Factors and Economic and Environmental Damages*). Technical University of Lisbon: Lisbon.

MERCER, D. R, AND PRESTEMON, J. P. 2005. Comparing Production Function Models for Wildfire Risk analysis in the Wildland – Urban Interface. *Forest Policy and Economics* **7**: pp 782 – 795.

PRESTEMON, J.P, MERCER, D.E., PYE, J.M, BUTRY, D.T, HOLMES, T.P., ABT, K.L. 2001. *Economically Optimal Wildfire Intervention Regimes*. American Agricultural Economics Association Annual Meeting: Chicago http://www.srs.fs.usda.gov/econ/pubs/jpp012.htm.

RAMACHANDRAN, G. 1998. *The Economics of Fire Protection*. E & FN Spon: London.

RIDEOUT D.B. AND ZIESLER P.S. 2004. The Great Myths of Wildland Fire Management. Presented at the *2nd Symposium on the Fire Economics and Policy: a Global View*. 19-22 April 2004: Cordoba.

RIERA, P. 2005. Forest Fire Valuation and Assessment: A Survey. Paper presented to the international conference *The Multifunctional Role of Forest-Policies, Methods and Case-Studies*, 28-30 Abril: Padova.

RIERA, P. AND MOGAS, J. 2004. Evaluation of Risk Reduction in Forest Fires in a Mediterranean Region. *Forest Policy and Economics* **6**: pp 521 – 528.

RODRIGUEZ Y SILVA, F. 2007. Coste y Eficiencia en las Operaciones de Extinción de Incendios Forestales, Fundamentos y Herramientas para su Estudio Y Análisis. 4th Internacional Wildland FIRE Conference, 13-17 May: Sevilla.

RODRIGUEZ Y SILVA, F. 2004. Economic Analysis in the Control of Fuel Loading in Mediterranean Forest Ecosystems. Prescribed Burning, an Alternative to Mechanical Methods. Paper presented at the 2nd Symposium on Fire Economics and Policy: a Global View, 19-22 April 2004: Cordoba.

SPARHAWK, W.N. 1925. The Use of Liability Ratings in Planning Forest Fire Protection. *Journal of Agricultural Research* **30**(8); pp 693-792.

WEINER, M. 2001. *The Economic Costs of Fire*. Home Office Research Study 229: London. http://www.homeoffice.gov.uk/rds/pdfs/hors229.pdf.

A HYBRID TOP-DOWN/BOTTOM-UP MODEL FOR ENERGY POLICY ANALYSIS IN A SMALL OPEN ECONOMY - THE PORTUGUESE CASE

SARA A. PROENÇA
E-mail: sproenca@esac.pt
Polytechnic Institute of Coimbra
MIGUEL ST. AUBYN
E-mail: mstaubyn@iseg.utl.pt
Technical University of Lisbon

Resumo

As questões energéticas/climáticas são, actualmente, reconhecidas como prioritárias na agenda política mundial, tanto pelos decisores políticos como pela comunidade científica.
O objectivo deste trabalho é contribuir para o desenvolvimento de uma metodologia consistente para a análise da política energética, capaz de apoiar os decisores políticos na concepção e promoção de instrumentos eficientes de combate às alterações climáticas, reduzindo assim a incerteza. É proposto um modelo híbrido top-down/bottom-up formulado como um problema de complementaridade mista (MCP), enquanto ferramenta de avaliação de impactes económicos de diferentes opções de política energética numa pequena economia aberta. O formato MCP permite integrar as duas abordagens de modelação num modelo único, introduzindo detalhe tecnológico nos modelos de equilíbrio geral.
A implementação do modelo é ilustrada através de uma aplicação empírica ao caso Português para avaliar os impactes económicos da actual política energética Portuguesa de promoção de fontes de energia renováveis na produção de electricidade.

Área Temática: Economia da Energia, do Ambiente, dos Transportes e dos Recursos Naturais.

Abstract

Energy/climate change is currently acknowledged by policy makers and the scientific community as one of the major global issues of the worldwide policy agenda.
This paper aims to contribute to the development of a consistent methodology for energy policy analysis, to assist, with decrease uncertainty, policy makers on the design and promotion of cost-effective instruments for climate change mitigation. For this purpose, a hybrid top-down/bottom-up model formulated as a mixed complementarity problem (MCP) will be proposed as a reliable modelling tool to assess the economic effects of different energy policy options in a small open economy. The MCP format permits to join the two modelling forms in a single integrated model, introducing technological detail in general equilibrium models.
The practical suitability of the model is illustrated by an empirical application to the Portuguese case to evaluate the economic impacts of the current Portuguese energy policy promoting electricity from renewable energy sources.

Key words: Economy-energy modelling, top-down and bottom-up, mixed complementarity problems, energy policy analysis.
Thematic Area: Economics of Energy, Environment, Transports, and Natural Resources.

1. INTRODUCTION

Energy and climate issues are today at the top of the political and economic agenda worldwide. In order to achieve energy policy goals of sustainability (low-carbon), security of supply and competitiveness multiple policy instruments are being adopted. Impact assessment of policy instruments plays an important role in the decision-making process. It is particularly relevant for policy makers on the design of cost-effective policy instruments. Methodologically, the impact assessment of energy policy measures have been made in the literature by two modelling approaches: bottom-up (or partial equilibrium) engineering-based models and top-down (mainly computable general equilibrium (CGE) models) economy models. However, both models have specific weaknesses which originate often divergent results, leading to uncertainty in impact assessment and policy design. Top-down models do not include technological detail, overestimating economic adjustments of the energy system, and tend to ignore technological changes induced by prices adjustments. These models help policy makers assess economy-wide policy instruments such as carbon taxes, but are ineffective in assessing the role of technological evolution to achieve a low-carbon economy. Bottom-up models, on the other hand, rely heavily on cost-effectiveness of energy technologies but neglect the macro-economic feedbacks induced by different energy-climate policy instruments, resulting in cost underestimation to achieve environmental goals. Aware of the shortcomings of each model type, recent studies have argued for the need to integrate the two modelling paradigms, combining technological explicitness of bottom-up models with the economic comprehensiveness of top-down models, i.e. developing hybrid models. Three main hybrid approaches have been developed: i) "soft-link" between two independent top-down and bottom-up models through an interactive process among inputs and outputs; ii) link one model to a reduced form of the other, usually link a bottom-up model to a simplified CGE model; iii) Mixed Complementarity Problem (MCP) format to join the two modelling forms in a single integrated model, introducing technological detail in general equilibrium models (e.g., Böhringer and Rutherford, 2008).

In this context, the main purpose of this study is to contribute to the development of a reliable hybrid modelling tool for applied energy policy analysis in a small open economy, in order to assist policy makers on the design and impact assessment of alternative cost-effective energy policy instruments, thereby reducing the uncertainty. For this purpose, a static, multi-sectoral, open, hybrid applied general equilibrium model formulated in the mixed complementarity format will be proposed. The practical suitability of our hybrid MCP model is illustrated by an empirical application for the Portuguese case-study to evaluate the economic effects of a prominent policy measure on the Portuguese energy-climate policy agenda: promoting electricity from renewable energy sources (RES-E).

The remainder of the paper is organized as follows. Section 2 introduces the analytical framework which underlies the merging of hybrid top-down/bottom-up models formulated as a mixed complementarity problem. Section 3 describes in detail the algebraic formulation of the hybrid open MCP model. Section 4 discusses an illustrative empirical application of the model. Section 5 concludes.

2. ANALYTICAL FRAMEWORK

This section introduces the theoretical concepts that underlie the merging of hybrid top-down/bottom-up models formulated as a Mixed Complementarity Problem (MCP). The purpose is to illustrate how the complementarity format can be used in general equilibrium modelling for the hybrid representation of economy-wide production possibilities with a detailed technology description in the energy sectors and top-down functional forms in the other production sectors.

2.1 General Economic Equilibrium Models

We assume a competitive Arrow-Debreu (1954) economy with n commodities (goods and factors), m production activities (sectors), and h households (including government). Following the Mathisen´s formulation (1985), the endogenous variables of the general economic equilibrium can be classified into three classes:

y represents a non-negative m-vector of activity levels of constant-returns-to-scale (CRTS) production sectors ($j = \{1,...,m\}$),

p is a non-negative n-vector of prices for all goods and factors ($i = \{1,...,n\}$), and

M is a non-negative k-vector of income levels ($h = \{1,...,k\}$).

A competitive equilibrium in these decision variables satisfies a system of three classes of equilibrium conditions: zero profit condition (quantity condition for all sectors), market clearance condition (price condition for all tradables), and income balance condition (values condition for all households). These three classes of conditions correspond to the general first-order necessary conditions of the producers´ profit maximizing and consumers´ cost minimizing problems, being the core of the general equilibrium problems formulation.

Zero Profit Condition

The first class of equilibrium conditions is based on producers' profit maximization and implies that, in equilibrium, the marginal cost of a commodity cannot be lower than its market price, i.e. no production activity makes a positive profit:

$$-\prod_j(p) = C_j(p) - R_j(p) \geq 0 \quad \forall_j \tag{1}$$

where,

$\prod_j(p)$ represents the unit-profit function for CRTS production sector j, which is calculated as the difference between unit revenue and unit cost functions[1].

In equilibrium, any activity which earns negative unit profits is inactive (i.e. no production activity takes place), hence:

$$y_j \prod_j(p) = 0 \quad \forall_j \tag{2}$$

Market Clearance Condition

The market clearance condition implies that, for equilibrium prices and activity levels, the supply of any commodity must balance or exceed the demand by consumers:

$$\sum_j y_j \frac{\partial \prod_j(p)}{\partial p_i} + \sum_h w_{ih} \geq \sum_h d_{ih}(p, M_h) \quad \forall_i \tag{3}$$

where,

$\frac{\partial \prod_j(p)}{\partial p_i}$ denotes (by Hotelling´s lemma)[2] the compensated supply of good i per unit operation of sector j,

w_{ih} is the initial endowment of household h with commodity i, and

---

[1] $R_j(p) \equiv \max\left\{\sum_i p_i \frac{\partial \Pi_j}{\partial p_i} \middle| g_j(.) = 1\right\}$; $C_j(p) \equiv \min\left\{\sum_i p_i \frac{\partial \Pi_j}{\partial p_i} \middle| f_j(.) = 1\right\}$.

The functions f_j and g_j represents feasible input and output combinations of production in sector j.

[2] Hotelling´s lemma: $\frac{\partial \Pi(p)}{\partial p_i} = y_i(p)$, for all $i = 1,...,n$, assuming derivative exists and $p_i > 0$.

$d_{ih}(p, M_h)$ is the utility maximizing demand function for good i by household h [3].

Any commodity in excess supply has an equilibrium price of zero, hence:

$$p_i\left[\sum_j y_j \frac{\partial \prod_j(p)}{\partial p_i} + \sum_h w_{ih} - \sum_h d_{ih}(p, M_h)\right] = 0 \quad \forall_i \tag{4}$$

Income Balance Condition

The third class of equilibrium conditions, budget constraint for households, requires that expenditure for each household h does not exceed disposable income:

$$M_h = \sum_i p_i w_{ih} \geq \sum_i p_i d_{ih}(p, M_h) \quad \forall_h \tag{5}$$

Given that underlying utility functions exhibit non-satiation Walras´ law holds[4] and then households are on their budget line, hence:

$$M_h\left(\sum_i p_i w_{ih} - \sum_i p_i d_{ih}(p, M_h)\right) = 0 \quad \forall_h \tag{6}$$

As demonstrated by Mathiesen (1985), in equilibrium any production activity operated at a positive intensity makes zero profit and any production activity which earns a negative net return is inactive. Similarly, any commodity with positive price has a balance between aggregate supply and demand and any commodity in excess supply has an equilibrium price of zero. And, for each household disposable income equals expenditure. This means that the economic equilibrium features complementarity between equilibrium variables and equilibrium conditions: an activity level to a zero profit condition, a commodity price to a market clearance condition, and a consumer income variable to a budget constraint. An equilibrium allocation determines quantities, prices and income levels. Thus, complementarity is a feature rather than a condition for equilibrium in the competitive model. This characteristic of the market equilibrium motivates the formulation of economic CGE models in a mixed complementarity format.

2.2 General Equilibrium and the MCP Format

A Mixed Complementarity Problem (MCP) consists of a combination of binding relationships (=) and weak inequalities ($\geq$) and complementarity slackness between system variables and system conditions. Following Rutherford (2002), the MCP format is defined as:

$$\begin{aligned} &\textit{Given} && F: R^n \to R^n, \quad l, u \in R^n \\ &\textit{Find} && z, w, v \in R^n \\ &\textit{subject to} && F(z) - w + v = 0 \\ & && l \leq z \leq u, \quad w \geq 0, \quad v \geq 0 \\ & && w^T(z - l) = 0, \quad v^T(u - z) = 0 \end{aligned} \tag{7}$$

in which $-\infty \leq l \leq u \leq +\infty$, u upper bound and l lower bound, z decision variable, v, w slack variables[5].

The MCP format keeps general equilibrium conditions in its most general form by stating the equilibrium problem as (Böhringer, 1998):

---

[3] $d_{ih}(p, M_h) \equiv \arg\max\left\{U_h(x) \middle| \sum_i p_i x_i = M_h\right\}$, in which U_h denotes the utility function for household h.

[4] As pointed out in Mas-Colell *et al.* (1995): "Walras´ law should be understood broadly: the consumer's budget may be an intertemporal one allowing for savings today to be used for purchases tomorrow. What Walras´law says is that the consumer fully expends his resources over his lifetime."

[5] The expression $x^T y = 0$ means $x_i y_i = 0$, for all $i = 1, \ldots, n.$ The variables x_i and y_i are called a complementary pair and are said to be complements to each other.

$$\begin{aligned} &\textit{Given} && f: R^n \to R^n \\ &\textit{Find} && z \in R^n \\ &\textit{subject to} && f(z) \geq 0,\ z \geq 0,\ z^T f(z) = 0 \end{aligned} \tag{8}$$

which is a special case of MCP format (a nonlinear complementarity problem) by setting $l = 0,\ u = +\infty,\ and\ F(z) = f(z)$.

Substituting z by $[y, p, M]$ and $f(z)$ by $\left[\prod_j(p), \left(\sum_j y_j \frac{\partial \prod_j(p)}{\partial p_i} + \sum_h w_{ih} + \sum_h d_{ih}\right), \left(\sum_i p_i w_{ih} - \sum_i p_i d_{ih}\right)\right]$ we notice that the complementarity format accommodates equilibrium conditions (1), (3), and (5) as well as the complementarity features (2) and (4) of a general equilibrium problem, showing equivalence of the economic equilibrium problem with the complementarity problem.

The MCP format is particularly useful for mathematical programs which cannot be solved as optimization problems (e.g., problems that include several consumers, taxes, or other market distortions). Then the problem can be turned into a mixed complementarity problem and solved as a system of non-linear equations. However, only if the complementarity problem is "integrable" (see Takayma and Judge, 1971), the solution corresponds to the first-order conditions for a (primal or dual) programming problem. Given integrability, the optimization problem can be interpreted as a market equilibrium problem where prices and quantities are defined using duality theory. In this case, a system of weak inequalities and complementary slackness conditions replace the minimization operator. In particular, the MCP formulation of general equilibrium problems permits integration of bottom-up activity analysis in which alternative technologies may produce one or more products subject to system constraints such as carbon emission limits or capacity constraints into a top-down CGE framework, enabling the merging of hybrid models.

3. THE HYBRID MODEL STRUCTURE

Based on the modelling framework described above, this section outlines the specification of a generic hybrid top-down/bottom-up model for energy policy analysis in a small open economy. Following Böhringer and Rutherford (2008), we present a detailed algebraic description of a single integrated model to combine technological detail of a bottom-up energy system with a top-down general equilibrium framework, designed to assess the medium-run effects of different energy policy measures. It is a static, multi-sectoral, applied CGE model for a small open economy formulated in the mixed complementarity format as a non-linear system of inequalities. The functional forms are written in calibrated share form which reduces the efforts as well as sources of errors in the calibration of free function parameters, since it is based on value shares that can be directly read from benchmark prices and quantities without inverting demand functions (Rutherford, 1995).

Following the structure of the Arrow-Debreu model presented in the previous section, the decision variables of our model are defined as follows[6]:

Activity Variables

Y_i denotes the production of non-energy good i,

E_i refers to energy aggregate input in sector i,

FF_i represents the production of fossil-fuel $i \in \{coal, oil, gas\}$,

ELE_t denotes the production of electricity by technology t,

[6] The notation follows closely Böhringer and Rutherford (2008). $\prod_i^x$ represents the profit function of sector i where x is the name assigned to the associated activity.

A_i represents the Armington aggregate good i,
U is the utility (consumption and leisure),
C refers to aggregate household final consumption, and
M_i is the world import aggregate of good i.

Price Variables

p_i is the output price of good i produced for the domestic market,
p_i^X is the output price of good i produced for the exported market,
p_i^A is the price of Armington aggregate good i,
p_i^E is the price of energy aggregate in sector i,
w is the labour price,
r_K is the capital price,
r_i is the rent to natural resource i,
p_{ELE} is the electricity price,
p_U is the utility price index,
p_C is the price of the household aggregate consumption,
p_i^M is the price of world import aggregate for good i, and
μ_t is the shadow price on generating capacity.

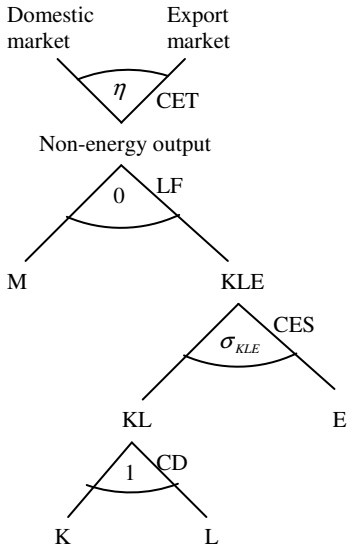
Income Variable

M is the income of the representative household.

3.1 Production

In each production sector a representative firm minimizes costs of producing output subject to nested constant-elasticity-of-substitution (CES) production functions, which reflect the substitution possibilities in domestic production between capital (K), labour (L), energy goods (E), and non-energy intermediate inputs (M). Each intermediate input M represents a composite of domestic and imported varieties (Armington assumption). Fig. 1 provides a diagrammatic overview of the nesting structure of non-energy goods (Y) production function. Production of non-energy goods (commodities other than fossil fuels and electricity) to the domestic and export market is described by an aggregate production function which characterizes the technology through transformation possibilities on the output side and substitution possibilities on the input side. On the output side, production is split between goods produced for the domestic market and goods produced for the exported market according to a constant-elasticity-of-transformation (CET) function. On the input side, non-energy intermediate inputs (used in fixed coefficients among themselves) are combined in fixed proportions (Leontief production structure) with an aggregate of capital, labour and energy, at the top level. At the second level, a CES function describes the substitution possibilities between the energy aggregate and the value-added aggregate. At last, at the third level, labour is combined with capital, trading off at a constant elasticity of substitution (Cobb-Douglas-CD).

Fig. 1: Nesting structure of non-energy production function



The unit-profit function of non-energy output can be written as:

$$\Pi_i^Y = \left(\theta_i^X \left(p_i^X\right)^{1-\eta} + \left(1-\theta_i^X\right)p_i^{1-\eta}\right)^{\frac{1}{1-\eta}} - \sum_{j\notin E}\theta_{ji}\,p_j^A - \theta_i^{KLE}\left[\theta_i^E\left(p_i^E\right)^{1-\sigma_{KLE}} + \left(1-\theta_i^E\right)\left(w^{\theta_i^L} r_K^{1-\theta_i^L}\right)^{1-\sigma_{KLE}}\right]^{\frac{1}{1-\sigma_{KLE}}} = 0 \quad \forall_{i,j} \notin E \tag{9}$$

where,

θ_i^X is the value share of ROW exports in sector i,

θ_{ji} is the cost share of non-energy intermediate input j in sector i,

θ_i^{KLE} is the cost share of KLE aggregate in sector i,

θ_i^E is the cost share of energy in the KLE aggregate of sector i,

θ_i^L is the labour cost share in sector i,

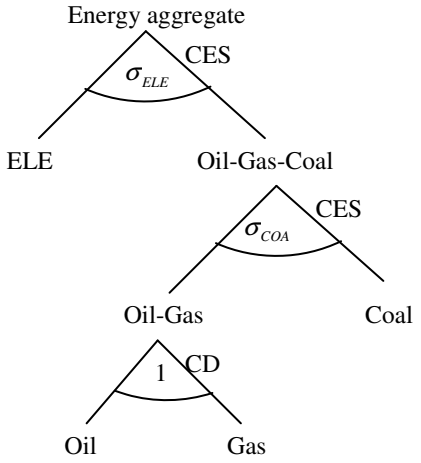
η is the elasticity of transformation between production for the domestic market and production for the exported market,

σ_{KLE} is the elasticity of substitution between the energy aggregate and the value-added in non-energy production, and

Y is the associated complementary variable.

The energy aggregate (E), in turn, is a nested CES composite of electricity and primary energy inputs, as illustrated in Fig. 2. The fossil-fuel composite is defined as a CES function of coal and a CD aggregate of liquid fuels (LQ) oil and natural gas.

Fig. 2: Nested energy aggregate production structure



The unit-profit function for energy aggregate is:

$$\Pi_i^E = p_i^E - \left\{ \theta_i^{ELE} \left(p_{ELE}^A \right)^{1-\sigma_{ELE}} + \left(1-\theta_i^{ELE}\right) \left[\theta_i^{COA} \left(p_{COA}^A \right)^{1-\sigma_{COA}} + \left(1-\theta_i^{COA}\right) \left(\prod_{j \in LQ} \left(p_j^A \right)^{\beta_{ji}} \right)^{1-\sigma_{COA}} \right]^{\frac{1-\sigma_{ELE}}{1-\sigma_{COA}}} \right\}^{\frac{1}{1-\sigma_{ELE}}} = 0 \qquad (10)$$

where,

θ_i^{ELE} is the cost share of electricity in energy demand by sector i,

θ_i^{COA} is the cost share of coal in fossil fuel demand by sector i,

β_{ji} is the share of liquid fossil fuel j in liquid fossil fuel demand by sector i,

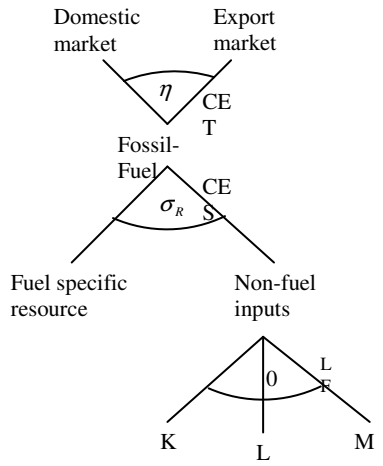
σ_{ELE} is the elasticity of substitution between electric and non-electric energy goods (fossil fuel aggregate) in production,

σ_{COA} is the elasticity of substitution between coal and liquid fossil fuel composite in production, and

E is the associated complementary variable.

In the primary fossil-fuels production (FF: coal, oil, and natural gas), all non-fuel inputs (capital, labour, and intermediate inputs) are aggregated in fixed proportions at the lower nest. At the top level, this aggregate trades off with the specific fossil-fuel resource at a CES function. The latter is calibrated in consistency with exogenously given price elasticity of fossil fuel supplies. The structure of fossil-fuel production function is represented in Fig. 3.

Fig. 3: Nested fossil-fuel production structure



The unit-profit function for fossil-fuel production is:

$$\prod_i^{FF} = \left(\theta_i^X \left(p_i^X\right)^{1-\eta} + \left(1-\theta_i^X\right) p_i^{1-\eta}\right)^{\frac{1}{1-\eta}} - \left[\theta_i^R r_i^{1-\sigma_{R,i}} + \left(1-\theta_i^R\right)\left(\theta_{Li}^{FF} w + \theta_{Ki}^{FF} r_K + \sum_j \theta_{ji}^{FF} p_j^A\right)^{1-\sigma_{R,i}}\right]^{\frac{1}{1-\sigma_{R,i}}} = 0 \quad \forall_i \in FF \tag{11}$$

where,

θ_i^R is the cost share of fossil-fuel natural resources in sector i,

θ_{Li}^{FF} is the labour cost share in the non-fuel composite input in sector i,

θ_{Ki}^{FF} is the capital cost share in the non-fuel composite input in sector i,

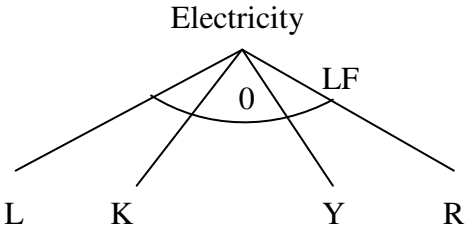
θ_{ji}^{FF} is the cost share of good j in the non-fuel composite input in sector i,

$\sigma_{R,i}$ is the elasticity of substitution between fossil-fuel natural resources and non-fuel inputs in fossil fuel production, and

FF is the associated complementary variable.

As Böhringer and Rutherford (2008), we integrate bottom-up activity analysis into a top-down general equilibrium framework through the detailed technological representation of the electricity production sector (ELE). Therefore, we represent electricity production possibilities by Leontief technologies which are active or inactive in equilibrium depending on their profitability. Fig. 4 provides an overview of the nested electricity production structure.

Fig. 4: Nested electricity production structure



The unit-profit function of power generation technology t is:

$$\prod_t^{ELE} = p_{ELE} - \left(a_t^L w + a_t^K r_K + \sum_{j \notin EG} a_{jt} p_j^A + \sum_i a_{it} p_i^A + \mu_t\right) = 0 \quad \forall_i \in R \ \forall_j \notin EG \tag{12}$$

where,

a_t^L is the labour input coefficient in electricity production by technology t,

a_t^K is the capital input coefficient in electricity production by technology t,

a_{jt} is the non-energy good input coefficient in electricity production by technology t,

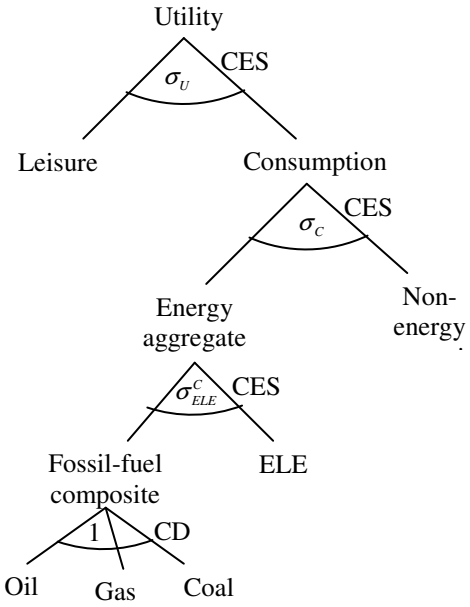
a_{it} is the natural resource (fossil-fuels and renewable resources) input coefficient in electricity production by technology t, and

ELE is the associated complementary variable.

3.2 Households

We consider a representative consumer agent[7] that maximizes utility subject to a budget constraint given by the income level. The CES utility function is shown in Fig. 5. At the top level, final consumption demand is combined with leisure at a CES function to form an aggregate utility good (i.e., households decide over current versus future consumption). Final consumption demand of the representative agent, in turn, is given by a composite good with a three-level nested CES technology. At the top level, the energy aggregate trades off with non-energy composite good at a CES function. Substitution patterns within the non-energy consumption bundle are reflected via Cobb-Douglas functions with an Armington aggregation of imports and domestic commodities. The fossil-fuel composite is combined at the second level with electricity subject to a CES function. Fossil fuels enter in the third level non-electric energy nest at a unitary elasticity of substitution.

Fig. 5: Utility nested function of representative consumer



The unit-profit function for aggregate utility is:

$$\prod^U = p_U - \left[\theta_U w^{1-\sigma_U} + (1-\theta_U)(p_C)^{1-\sigma_U}\right]^{\frac{1}{1-\sigma_U}} \quad (13)$$

where,

θ_U is the leisure value share of aggregate utility,

σ_U is the compensated elasticity of substitution between leisure and final consumption, and

U is the associated complementary variable.

The unit-profit function for household final consumption demand is:

[7] A population of identical households (including government).

$$\Pi^C = p_C - \left\{ \theta_C \left[\prod_{i \notin EG} \left(p_i^A \right)^{\gamma_i} \right]^{1-\sigma_C} + \left(1-\theta_C\right) \left[\theta_{ELE}^C \left(p_{ELE} \right)^{1-\sigma_{ELE}^C} + \left(1-\theta_{ELE}^C\right) \left(\prod_{j \in FF} \left(p_j^A \right)^{\varphi_j} \right)^{1-\sigma_{ELE}^C} \right]^{\frac{1-\sigma_C}{1-\sigma_{ELE}^C}} \right\}^{\frac{1}{1-\sigma_C}} = 0 \quad \forall_i \notin EG, \forall_j \in FF \tag{14}$$

where,

θ_C is the cost share of non-energy composite in aggregate household consumption,

θ_{ELE}^C is the cost share of electricity in household energy aggregate demand,

γ_i is the cost share of non-energy good i in non-energy household demand,

φ_i is the cost share of fossil fuel i in non-electric household energy demand,

σ_C is the elasticity of substitution between energy and non-energy goods in household consumption ,

σ_{ELE}^C is the elasticity of substitution between electric and non-electric energy in household energy consumption, and

C is the associated complementary variable.

In our small open economy, the representative consumer agent is endowed with primary factors labour (time) and capital, and natural resources. Thus, total income of the representative household consists of factor payments and scarcity rents on capacity constraints. The income balance of the representative household is then written:

$$M = w\,\overline{L} + r_K\,\overline{K} + \sum_i r_i \overline{R}_i + \sum_t \mu_t \overline{ELE}_t + \overline{B} \tag{15}$$

where,

$\overline{L}$ is the exogenous aggregate endowment with time,

$\overline{K}$ is the exogenous aggregate capital endowment,

$\overline{R}_i$ is the exogenous endowment with natural resource i,

$\overline{ELE}_t$ is the generation capacity for technology t, and

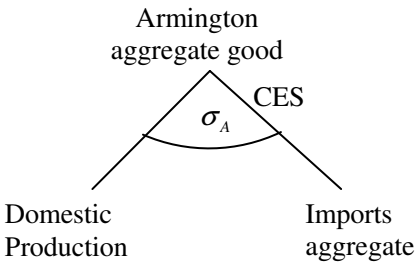
$\overline{B}$ is the balance of payment surplus (deficit).

3.3 International Trade

International trade is modelled assuming two common assumptions in the literature: i) the small open economy assumption, meaning that export and import prices in foreign currency are not affected by the behaviour of the domestic market (i.e., the domestic market is too small to influence world prices, being assumed how price-taker in relation to the ROW), and that the world market can satisfy all the importing and exporting needs of the domestic economy[8]; ii) the Armington assumption (Armington, 1969) of international product differentiation (in the sense that imported and domestically produced goods of the same type are imperfect substitutes) for imports and, symmetrically, the constant-elasticity-of-transformation (CET) supply function for exports, meaning that domestically produced goods may be supplied either to domestic market and export market. The Armington assumption of product heterogeneity means that each product variety used on the domestic market in intermediate and final demand corresponds to a combination of domestic production and imports from different regions with a CES composite function, as illustrated in Fig. 6.

[8] See Shoven and Whalley (1992).

Fig. 6: Nested Armington good



The unit-profit function for Armington aggregate good is:

$$\prod_i^A = p_i^A - \left[\left(\theta_i^A \, p_i^{1-\sigma_A} + \left(1-\theta_i^A\right) \left(p_i^M\right)^{1-\sigma_A} \right)^{\frac{1}{1-\sigma_A}} \right] = 0 \qquad (16)$$

where:

θ_i^A	is the cost share of domestic variety i in Armington aggregate good,
σ_A	is the Armington substitution elasticity between domestic and imported varieties of the same good, and
A	is the associated complementary variable.

Finally, notice that there is an imposed balance of payments constraint to ensure trade. The rest-of-the-world (ROW) closure requires that the value of imports to the ROW is equal to the value of exports from the ROW after including a constant benchmark trade surplus or deficit (B). A small open economy is assumed to be price-taker with respect to world market prices (world prices are considered to be exogenous), and hence trade with ROW is included by perfectly elastic ROW import-supply and export-demand functions.

4. APPLICATION OF THE MODEL

The purpose of this section is to illustrate the practical suitability of the hybrid MCP model described above with a simple numerical example based on empirical data for the Portugal case-study for the base-year 2005. We apply our "prototype" model to evaluate the economic effects of a prominent policy measure on the Portuguese energy-climate policy agenda: promoting electricity from renewable energy sources (RES-E).

The sample model describes a small open economy with five production sectors – coal, natural gas (Gas), oil, electricity (Ele), and a macro-good (Y) which represents the rest of economy – whose outputs are demanded by intermediate production, exports and one representative final consumer (Ra). Primary factors include labour (L), capital (K, sector-specific), natural fossil fuel resources (coal, gas, and oil) and natural renewable resources (wind, sun, trees, and water). The electricity production sector covers a set of seven electricity generation representative technologies (coal, gas, oil, hydro, wind, solar, and biomass). The numerical model is implemented as a system of non-linear inequalities using GAMS/MPSGE[9] mathematical programming language (Rosenthal, 2008; Rutherford, 1995, 1999) and solved by using the PATH solver (Dirkse and Ferris, 1995).

[9] General Algebraic Modeling System (GAMS)/Mathematical Programming System for General Equilibrium (MPSGE).

4.1. Benchmark Data

The calibration approach is adopted in parameter specification of the model[10], which is calibrated to a single base-year equilibrium such that the model base solution exactly reproduces the values of the adjusted data. This replication characterizes the initial allocation of resources, the so-called benchmark equilibrium. Benchmark prices and quantities, jointly with exogenous elasticities determine the free parameters of the functional forms required to model calibration. The macroeconomic data comes in the form of an aggregate national accounting matrix (NAM), constructed from the integrated system of input-output tables for Portugal for 2005 (DPP, 2008). The final NAM[11], shown in Table 1, provides the benchmark data base for the top-down model calibration for Portugal in the base-year 2005. However, to calibrate a hybrid model we have to disaggregate the column in the NAM representing the electric power sector into specific electricity generation technologies (bottom-up engineering data).

Table 1. Benchmark Accounting Matrix for Portugal, 10^6 Euros (NAM-2005)

	Y	Coal	Gas	Oil	Ele	Ra	Exports	Imports	Total Sum
Y	138 499		-178	-356	-976	-147 941	-36 083	47 036	0
Coal	-7				-271	-19		297	0
Gas	-350		1 303	-14	-515	-424			0
Oil	-3 487		-734	4 817	-430	-5 645	-1 117	6 597	0
Ele	-2 984		-5	-3	4 573	-1 837	-93	349	0
Labour	-74 653		-90	-70	-546	75 358			0
Capital	-52 327		-143	4	-1 802	54 267			0
Rent			-147	-4 372		4 519			0
Trf	-4 691		-6	-6	-34	4 766	-28		0
Trade balance						16 956	37 321	-54 278	0
Total Sum	0	0	0	0	0	0	0	0	0

Note: The row Trf represents benchmark transfers between economic agents.

Table 2 and 3 summarize the engineering bottom-up characteristics for the set of electricity generation technologies active in the benchmark equilibrium (coal, gas, oil, and hydro). These data are computed from a range of sources namely the International Energy Agency (IEA) statistics, the TIMES_PT model[12], and the Directorate General for Energy and Geology (DGEG). Given the output share and the cost structure of benchmark active technologies the top-down electric power sector in the NAM is disaggregate into individual electricity supply technologies in a way that is consistent with these technologies´ bottom-up engineering characteristics.

[10] See e.g., Mansur and Whalley (1984) and Devarajan *et al.* (1994) for more discussion on the calibration approach.

[11] The benchmark NAM is accounted in monetary values (price time quantity) from the national accounts. Then it is not straightforward to distinguish between prices and physical quantities. As usual in the literature, we adopt the Harberger convention, which consists in normalizing all benchmark prices to unity, whereas the amount of production and consumption are set equal to the monetary values in the base accounting matrix (the benchmark values are treated as quantities). Following this convention an Arrow-Debreu economy only depends upon relative prices.

[12] TIMES_PT is a linear optimisation bottom-up technology model, which results from the implementation for Portugal of the model generator TIMES developed by ETSAP (Energy Technology Systems Analysis Programme) of the International Energy Agency. TIMES are the acronym for The Integrated MARKAL-EFOM system. The implementation of the TIMES model for Portugal is being done within the international research project NEEDS (New Energy Externalities Development for Sustainability). The Portuguese research team is responsible for the base-year information, for the validation of technologies information, and for calibration and validation of the national model.

Table 2. Technology Specification Data

Nr.	Technology	Electricity generation GWh	Technology share %	Physical capacity MWh
1	Coal	15 226	33	1 776
3	Gas	13 606	29	2 166
2	Oil	8 791	19	1 909
4	Hydro	8 952	19	6 464
	Total	46 575	100	12 315

Note: The tiny amounts of other renewable technologies (biomass, waste, wind, solar, and geothermal) are accounted for within hydro.

Table 3. Cost Structure of Benchmark Technologies

	Coal	Gas	Oil	Hydro
Ele	1 495	1 336	863	879
Y	-369	-247	-130	-265
Coal	-271			
Gas		-515		
Oil			-430	
Labour	-199	-134	-70	-143
Capital	-656	-440	-232	-471

The elasticity values, as usual in the calibration of CGE models, are taken from review of the relevant literature. Table 4 provides a summary of key elasticities underlying our policy simulations.

Table 4. Summary of Key Elasticities

Production	
Capital-labour vs. Energy aggregate in macro-good production	0.6
Electricity vs. Primary energy inputs (fossil fuel aggregate) in macro-good production	0.3
Fossil fuel inputs in macro-good production	2
Fossil fuel resources vs. Other non-fuel inputs in fossil fuel production	
Coal	0.5
Gas	0.25
Oil	0.25
Consumption	
Consumption vs. Leisure	1.4
Energy vs. Non-energy goods	0.5
Electricity vs. Primary energy inputs (fossil fuel composite)	0.5
Trade	
Domestic production vs. Exported goods (Armington elasticity of transformation)	2
Domestic production vs. Imported goods (Armington elasticity of substitution)	2

Source: Drawn up by authors for this study based on relevant literature - see e.g., GTAP-E (Truong *et al.*, 2007); Kemfert and Welsch (2000); Hertel, 1997; Böhringer and Andreas (2008).

4.2. Policy simulation and results

In the policy simulation, we measure the economic effects of Portugal RES-E promotion against a business-as-usual (*BaU*) scenario without political support for renewable energy. According to the Directive 2001/77/EC the RES-E target share to be achieved by Portugal in 2010 is 39% of gross electricity consumption. In January 2007 the Portuguese Government raised voluntarily this target share to 45% in 2010 - we have assumed the RES-E target share in terms of total electricity production. Within this policy background, in the counterfactual policy scenario we impose a cumulative quantity constraint on the share of electricity produced from

renewable energy sources at a level which assures compliance with the Portugal RES-E target under the EU obligations. This quantity boundary is related with a complementary uniform endogenous subsidy on renewable electricity generation technologies inactive in the base year (wind, solar, and biomass). We assume that the subsidy is financed by the representative household.

In the *BaU* scenario (year 2005) Portugal renewable energy (i.e. hydro) share in overall power generation amounts to roughly 19%. In our policy simulation, we increase this share gradually in five steps by 20 percentage points up to the Portugal RES-E target share of 39% in 2010.

Simulation results show that the uniform subsidization of renewable power technologies leads to an increase in RES-E deployment, leading to a diversification of energy power supply as can be seen in Figs. 7 and 8. There is a shift from high-carbon fossil fuels technologies such as oil toward carbon-free and high-efficiency power generation technologies such as wind and, consequently, a reduction in CO_2 emissions. The electricity generation from oil decreases more than 50%, which results in a drop of the oil share from by 19% under *BaU* to roughly 8% under the RES-E target share of 39%. The share of wind technology in total electricity generation, in turn, increases from 0% in *BaU* scenario to roughly 23%. Notice, however, that the support scheme for RES-E is not high enough for solar and biomass renewable technologies to break even. Further, the overall level of electricity supply increases by 17% compared with the *BaU* scenario level, leading to a reduction of the electricity imports with positive impacts in terms of security of energy supply.

Fig. 7. Impact of RES-E promotion on electricity generation technology mix

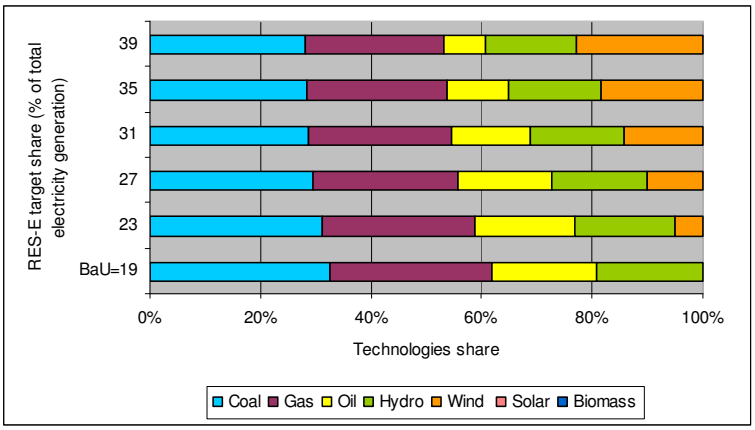


Fig. 8. Impact of RES-E promotion on electricity supply

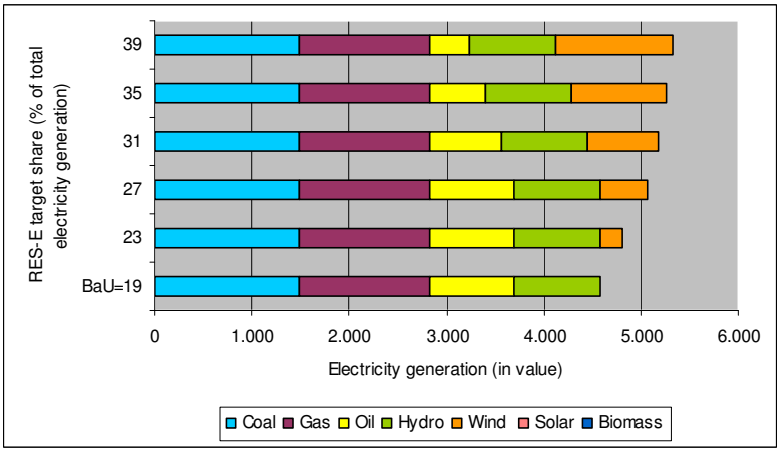
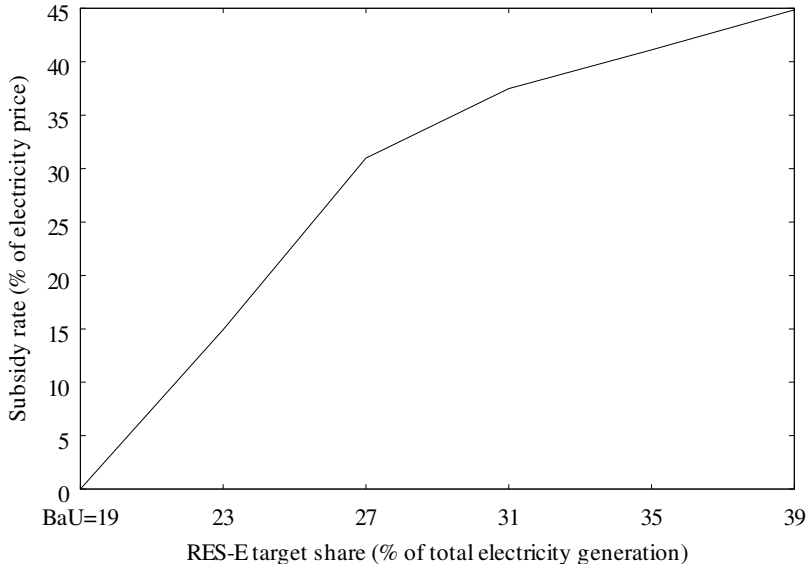


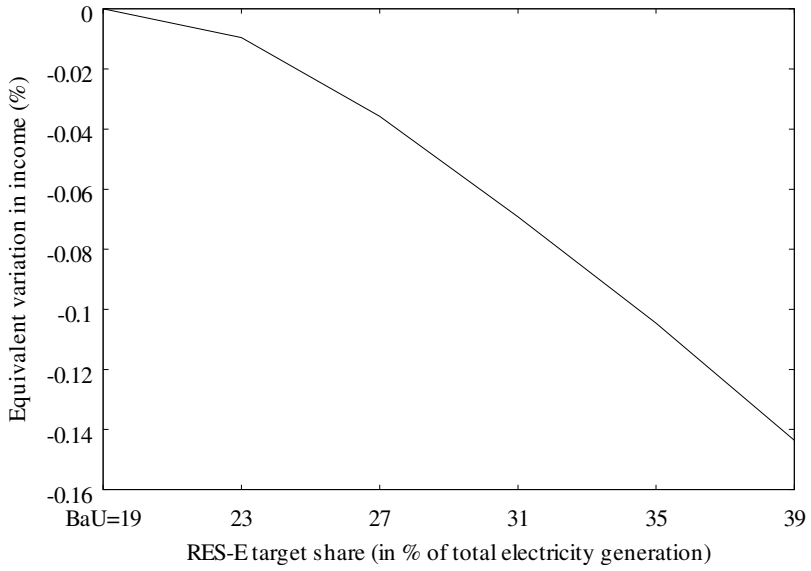
Fig. 9 illustrates the required subsidy rate to achieve rising RES-E shares from de *BaU* scenario to the Portugal RES-E target share of 39% in 2010, where the subsidy rate is around 45% of electricity price.

Fig. 9. Subsidy rates of RES-E promotion



The welfare impacts of the RES-E promotion policy are measured in terms of Hicksian equivalent variation (HEV) in income, which captures changes in the economic agent's income induced by the endogenous subsidy. As shown in Fig. 10, the welfare effects are negative. The extent of welfare losses is closely related to the RES-E share value, wherein a share of 39% leads to an income loss of roughly -0.14%. Notice, however, that such induced income losses should be compared with the welfare gains from positive environmental externalities, e.g. from lower CO_2 emissions. Yet, no such effects are captured by the model.

Fig. 10. Welfare effects of RES-E promotion



5. CONCLUSIONS

Conventional top-down and bottom-up models have inherent weaknesses which limited their usefulness to police-makers in energy policy analysis. Accordingly, policy modellers have explored the development of a new generation of hybrid energy-environment-economy models which contains both technological foundation of bottom-up models and the economic richness of top-down general equilibrium models.

In this paper, we have motivated the development of a hybrid top-down/bottom-up modelling tool for energy policy analysis in a small open economy, formulated as a mixed complementarity problem. The MCP format permits to join the two modelling forms in a single integrated model, introducing technological detail in general equilibrium models. A static, multi-sectoral, applied hybrid general equilibrium MCP model is presented as a reliable modelling tool to assess the economic effects of alternative energy policy measures. The economic costs induced by energy-climate policies can be substantially reduced if an assessment is made of the cost-effectiveness of alternative policy instruments and technological options. This impact assessment can be properly made by using this type of modelling approach, which can be applied to a wide variety of policy analysis and countries.

The practical suitability of the model is illustrated by a simple numerical application for the Portuguese case-study, analysing the economic effects of Portugal RES-E promotion. Our simulation results show that the uniform subsidization of renewable power technologies require relatively modest adjustment costs to the Portuguese economy.

REFERENCES

ARMINGTON, P. (1969), "A Theory of Demand for Products Distinguished by Place of Production", *IMF Staff Papers*, Vol. 16, nº 1, pp. 159–178.

ARROW, K., DEBREU, G. (1954), "The existence of an equilibrium for a competitive economy", *Econometrica*, Vol. 22, pp. 265-269.

BÖHRINGER, C., RUTHERFORD, T. (2008), "Combining bottom-up and top-down", *Energy Economics*, Vol. 30, nº 2, pp. 574-596.

BÖHRINGER, C. (1998), "The synthesis of bottom-up and top-down in energy policy modelling", *Energy Economics*, Vol. 20, nº 3, pp. 233–248.

BÖHRINGER, C., LÖSCHEL, A., WELSCH, H. (2008), "Environmental Taxation and Induced Structural Change in an Open Economy: The Role of Market Structure," *German Economic Review*, Vol. 9, nº 1, pp.17-40.

DEVARAJAN, S., LEWIS, J., ROBINSON, S. (1994), "Getting the Model Right: The General Equilibrium Approach to Adjustment Policy", Mimeo.

DGEG (2008), "Estatísticas e Preços", available at [http://www.dgge.pt].

DIRECTIVE 2001/77/EC of the European Parliament and of the Council of 27 September 2001 on the promotion of electricity produced from renewable energy sources in the internal electricity market, Official Journal of the European Communities, 27.10.2001, L 283/33.

DIRKSE, S., FERRIS, M. (1995), "The PATH Solver: a non-monotone stabilization scheme for mixed complementarity problems", *Optimization Methods and Software*, Vol. 5, pp. 123–156.

DPP (2008), "Integrated System of Input-Output Tables for Portugal, 2005", available at [http://www.dpp.pt].

HERTEL, T. (1997), "Global Trade Analysis: Modeling and Applications", Cambridge University Press.

IEA (2008), "Energy Statistics for 2005", available at [http://www.iea.org].

KEMFERT, C., Welsch, H. (2000), "Energy-Capital-Labour substitution and the economic effects of CO_2 abatement: Evidence for Germany", *Journal of Policy Modelling*, Vol. 22, nº 6, pp. 641-660.

MANSUR, A., WHALLEY, J. (1984), "Numerical specification of applied general equilibrium models: estimation, calibration, and data", in Scarf and Shoven (eds.), Applied general equilibrium analysis, Cambridge University Press.

MAS-COLELL, A., WHINSTON, M., GREEN, J. (1995), Microeconomic Theory, Oxford University Press, Inc., Madison Avenue, New York.

MATHIESEN, L. (1985) "Computation of economic equilibrium by a sequence of linear complementarity problems", in Manne, A. (ed.), Economic Equilibrium — Model Formulation and Solution, Vol. 23, pp. 144–162.

ROSENTHAL, R. (2008), "GAMS – A User´s Guide", available at [http://www.gams.com].

RUTHERFORD, T. (2002), "Mixed Complementarity Programming with GAMS", Lecture Notes for Econ 6433, Springer Semester, available at [http://pascal.iseg.utl.pt/~depeco/summerschool2008/P1-jedc1999.pdf].

RUTHERFORD, T. (1995), "Extension of GAMS for complementarity problems arising in applied economic analysis", *Journal of Economic Dynamics and Control*, Vol. 19, nº 8, pp. 1299-1324.

RUTHERFORD, T. (1999), "Applied General Equilibrium Modelling with MPSGE as a GAMS Subsystem: An Overview of the Modelling Framework and Syntax", *Computational Economics*, Vol. 14, pp. 1-46.

SHOVEN, J., WHALLEY, J. (1992), "Applying General Equilibrium, Cambridge University Press, Cambridge.

TAKAYMA, T., JUDGE, G. (1971), "Spatial and Temporal Price and Allocation Models", North-Holland, Amsterd.

TRUONG, T., KEMFERT, C., BURNIAUNX, J-M. (2007), "GTAP-E: An Energy-Environmental Version of the GTAP Model with Emissions Trading", DIW Discussion Paper 668, February, available at [http://www.diw.de/documents/publikationen/73/55787/dp668.pdf].

"CRIME AND PUNISHMENT" – AN ECONOMIC ANALYSIS OF ILLEGAL FISHING

MANUEL PACHECO COELHO
E-mail: coelho@iseg.utl.pt
SOCIUS/ISEG
JOSÉ ANTÓNIO BONITO FILIPE
E-mail: jose.filipe@iscte.pt
MANUEL ALBERTO M. FERREIRA
E-mail: manuel.ferreira@iscte.pt
ISCTE

Resumo

O reforço da fiscalização sobre o cumprimento da Lei imposto pelas instituições públicas é uma temática, teórica e empírica, muito relevante na área das Ciências Sociais. Montesquieu, Beccaria and Bentham referiram inicialmente este assunto no Séc. XVIII. A Literatura Económica olvidou-o até BECKER (1968), "Crime and Punishment: An Economic Approach". Na Economia das Pescas, o problema pode ser visto como uma externalidade quando os direitos exclusivos de propriedade estão ausentes. Isto depende, entre outros factores, dos custos de definir e impor a exclusividade. Considerações de eficiência não impõem, *de per si*, a escolha do regime correcto de direitos de propriedade. Na "Propriedade Comum" o realinhar dos direitos pode ter custos proibitivos. O modelo proposto conjuga o Modelo Básico de Gordon/Schaefer com a Teoria do "Crime and Punishment" de Becker. As conclusões permitem discutir o desenho e a reforma do regime de monitorização e controlo da Política Comum das Pescas.

Palavras-chave: Pesca, Fiscalização, Política Comum das Pescas.
Área Temática: Economía de la Energía, del Ambiente, de los Transportes y de los Recursos Naturales.

Abstract

Public enforcement of law is an obvious important theoretical and empirical subject for Social Sciences. First literature dates from eighteen century: Montesquieu, Beccaria and Bentham. After the sophisticated analysis of Bentham, enforcement subject "lay essentially dormant in economic scholarship", until BECKER (1968) article, "Crime and Punishment: An Economic Approach". In Fisheries Economics, this can be seen as an externality arising when exclusive property rights are absent, and that absence depends, particularly, on the costs of defining and enforcing exclusivity. Efficiency considerations don't dictate, only by themselves, the choice of a certain property rights regime. In "common property" the re-alignment of the property rights can have a very high or even prohibitive cost. This model combines standard Economics of Fisheries analysis (Gordon/Schaefer model) with Becker's Theory of "Crime and Punishment". The conclusions are used to discuss the design and reform of the control and monitoring regime of the Common Fisheries Policy.

Key Words: Fisheries, Enforcement, Common Fisheries Policy.
Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

INTRODUCTION

Public enforcement of law, that is, the use of public agents to detect and sanction violators of legal rules, is an obvious important theoretical and empirical subject for Social Sciences. First literature on the subject of law enforcement dates from eighteen century: Montesquieu, Beccaria and Bentham. Curiously, after the sophisticated analysis of Bentham, the subject of enforcement

"lay essentially dormant in economic scholarship" (Polinsky and Shavell (2000)), until the influential article of Becker, 1968, "Crime and Punishment: An Economic Approach".

In the context of Fisheries Economics, the problem can be seen as an externality arising when exclusive property rights are absent (Cheung (1970), and that absence, among other things, depends on the costs of defining and enforcing exclusivity.

Differences in the costs of information, negotiation and enforcement, associated to the privatisation of the property rights, implicate differences in the choice and evolution of the alternative contracts (Alchian and Demsetz (1973); Cheung (1968)).

Efficiency considerations, only by themselves, don't dictate the choice of a certain property rights regime. In some systems of property rights (as it is the case of "common property") the re-alignment of the property rights can have a very high or even prohibitive cost. The establishment and enforcement of a system of rights depends, of course, on efficiency considerations, but also on the individual preferences and the ethical, political and social realities in a community. These include the lack of means (or other insufficiencies) of the administration to control and enforce the execution of legal rules - Demsetz (1967).

By definition, anything that is an infringement of the law is illegal. Illegal fishing therefore covers a wide range of behaviours, which can take place at three levels for members of the European Union: national, Community and international. Any violation of national laws or EU regulations, or failure to comply with the recommendations of international bodies, especially those of Regional Fisheries Organisations, constitutes an infringement.

Illegal fishing has always existed, but, in recent decades, there has been a sharp rise in violating activities, due to technical progress: motorization, freezing techniques, improved gear, new forms of stocks detection and information. This process was majored by the evolution of the Law of the Sea - a "creeping jurisdiction" process which seems to have given an end to the principle of open access.

Obviously it is impossible to quantify or qualify infringements. They are known to take place at all levels and take different forms at different times; some violations are detected but many remain unnoticed. Infringements take the traditional forms of fishing over the quota or using non-permitted mesh-size, but are also in situations of non-permitted by-catches or transhipment, even in the fake world of convenience flags. Illegal fishing occurs at all stages of fishing activity.

A large number of offenders are fishermen motivated by various interests, the fundamental being the lure of short term profit. But fishermen are not the only ones involved. Fraud can take place along the entire channel. National administrations sometimes bear part of the blame. Every state is responsible for enforcing the existing rules and monitoring activities (policing its territory, conducting controls and penalising offenders). Its inefficacy in controlling activities is the reason of a lot of enforcement problems.

Most of the literature on fisheries management implicitly assumes law can be perfectly and costless enforced. Even when such costs and imperfections are recognised, they are not incorporated in the analysis to show how management and regulatory policies are affected by their presence. This paper explores this issue with a formal model of fisheries law enforcement to show how fishing firms behave and fisheries policies are affected by costly, imperfect enforcement of fisheries law. This model combines standard Economics of Fisheries analysis (Gordon/Schaefer model) - Clark and Munro (1975), Clark (1985) - with the Theory of "Crime and Punishment" of Becker (see Sutinen and Andersen (1985) and, for an extensive review of the literature on fisheries regulation enforcement, see Nostbakken (2008) and Sumaila *et al* (2006)). The conclusions of the model are used to discuss the design and reform of the control and monitoring regime of the Common Fisheries Policy.

1. "CRIME AND PUNISHMENT" – THEORETICAL ANALYSIS

Despite the enormous volume of literature on Fisheries Economics, the issue of enforcement has always been "the neglected element in fishery management" (Sutinen and Hennessey (1986)).

We can explore this issue with a formal model of fisheries law enforcement showing how fishing firms behave and fisheries policies are affected by costly, imperfect enforcement of fisheries regulation. This model combines standard Gordon /Schaefer fisheries model with the theory of crime and punishment of Becker.

The fundamental problem in fisheries management is to obviate the tendency towards overexploitation of the resources under open access. Regulation methods used to curb this tendency of overfishing and overcapacity includes gear restrictions, area and seasonal closures, TACs, ITQs, limiting entry and other forms of reducing fishing effort.

Let's assume that, whatever means are applied to reduce catch rates, any catch level above the level of the permitted quota for a certain fishing, q*, is illegal. If we suppose a system of individual non-transferable quotas, the amount of the individual firm catch above its quota (q_i – q_i*) is illegal.

If detected and convicted, a penalty fee is imposed on the firm in an amount given by f,

$$f = f(q_i - q_i^*),$$

where $f > 0$, if $q_i > q_i^*$; and f =0, otherwise;

and $\frac{\partial f}{\partial q} \geq 0$; $\frac{\partial^2 f}{\partial q^2} \geq 0$; $\forall\ q_i > q_i^*$.

For hypothesis the function *f(.)* is continuous and differentiable for all $q_i^* > q_i$. This penalty fee has a finite upper bound and each firm is assumed to face the same penalty fee schedule.

An individual firm's profit before penalty is given by

$$\Pi^i(q_i,x) = pq_i - c^i(q_i,x),$$

where *p* denotes the price of fish, *x* is the size of fish stock and *c(.)* is the cost function. Let's assume that firms are price takers.

In an imperfect law enforcement regime not every violator is detected and convicted. Let the probability of detection and conviction be given by θ, and, to simplify, let' s assume that all firms face the same probability.

If detected and convicted of a violation, a firm's profit will be $\Pi^i(q_i,x) - f(q_i - q_i^*)$; if not, $\Pi^i(q_i,x)$.

So, expected profits are

1) $\theta[\Pi^i(q_i,x) - f(q_i - q_i^*)] + (1-\theta)\,\Pi^i(q_i,x)$

Assuming that firms are risk neutral and maximising expected profits, each q_i is determined by the first order condition (subscripts other than i denote partial derivatives)

2) $\Pi_q^{\ i}(q_i,x) \geq \theta\, f_q(q_i - q_i^*)$.

The solution to 2) for one form of the marginal penalty schedule, f_q, has a clear economic meaning. The model sustains a rule of optimal behaviour for a rational ("homo economicus") operator:

For a given stock size (x), the firm sets its catch rate at a level in excess of its quota, where marginal profits equal the expected marginal penalty. If there were no penalty for fishing beyond legal quota, or if there were no probability of being detected and convicted ($f = 0$ or $\theta = 0$) the firm would set its catch at the open access catch rate. If the expected marginal penalty schedule lies above the marginal profit schedule for all q_i above the legal quota, the firm's "optimum" catch equals its quota. Firms with no quota have an expected net gain for entering, illegally, in the fishery, if their expected marginal penalty schedule begins below their marginal profit schedule.

This approach reveals the importance of empirical studies (see Sutinen and Gauvin (1989)) trying to estimate the factors that ensure compliance with the regulation. These studies give important basis for public authority decision about the actions to be implemented.

Stigler (1970) argues that public authorities have four basic means to improve compliance:

- minimise the chances that violations will go undetected,
- maximise the probability that sanctions will follow the detection of violations,
- speed up the process from time to detection to assignment of sanction,
- make the sanctions large.

There is dispute among experts about the best alternatives. Some scholars have argued that the probability of being detected is more important than the size or magnitude of the sanction, while others argue that making the charging time follow as closely as possible to the detection of illegal behaviour is the most important factor in enhancing compliance. Others, also, put in evidence the level of expenditure oriented to monitoring activities (Tietenberg (2003)).

2. THE MONITORING AND CONTROL REGIME OF THE COMMON FISHERIES POLICY

The main objective of the Common Fisheries Policy (CFP) is to provide sustainable exploitation of fish resources. In order to ensure the achievement of this objective, Community rules shall be applied, in an effective and uniform manner.

The effectiveness of the Common Fisheries Policy depends on the compliance of the various operators concerned with the CFP rules. Member states are responsible for ensuring the correct application of the CFP rules on their territory and in the waters under their jurisdiction. They must also ensure that all vessels flying their flags comply with these rules wherever they operate.

To ensure the equity and fairness of control and monitoring through the Community, Commission inspectors oversee the activities of the national enforcement services and report to the Commission. In addition, periodically, member states must transmit information to the Commission on various aspects of their enforcement activities.

It is up to the member states to ensure the effective implementation of regulation. Of course, inspection and surveillance of fisheries and related activities implicate establishing administrative and technical structures. So, it is of Community interest that the member states have structures (human and technical resources) that are adequate for the control purposes and allow them to meet the requirements of European Union regulation.

The Commission has implemented, since 1978, a series of provisions for financial support towards member states expenditure. This reflects the interest of the European Commission for the control improvement in the fishery activities. The first Council Decision was approved in 1978. A decade later, in 1987, a new Decision went on to grant financial contribution to all member states. The objective was to promote the development on the monitoring and control system capacity and capability within specific areas of control and surveillance.

In the 90s, the Community introduced a more complete aid scheme to provide financial support to the member states for the establishment of control structures, and, in particular, to promote certain forms of control which are particularly effective, such as systems to monitor fishing activities at a distance. Council Decision 89/631 established a framework for Community financial assistance. Assistance Programs were developed by the Commission since then to promote fairness among member states and an efficient control of fishing activities.

On the basis of the experience gained and in conformity with the conclusions drawn by the ex-post evaluation reports, it can be said that, in general, funding has been directed to the most deficient states in terms of EU average means for control purposes. This has addressed the imbalance in member states control capacity, but it has not completely solved it and there was a further scope for additional expenditure.

In this context, and considering the evaluation of the two first decades of CFP made in the "Green Paper"/2002, the 2003 Reform has brought an important evolution, particularly in the control regime of the European Fisheries.

The fundamental guidelines of this reform were the following:

A new framework and a more complete body of regulation were introduced. This new Regulation stressed the need to further improve control of fishing activities in order to fight, by any means, against illegal and undeclared fishing within and outside Community waters. It identified remote control technologies as a tool to better achieve the control objectives under CFP and extended the obligation for remote monitoring by means of vessel monitoring system - VMS (that became the real "jewel of the crown" of the control regime).

With ten new countries joining the Community, the monitoring structures of some of these countries needed upgrading rapidly. For these reasons, and for the sake of greater clarity, it was proposed to redefine the measures eligible for financial support. Those measures included, among others, purchase of computer equipment and installation of IT networks; purchase of equipment devices to locate fishing vessels and equipment for electronic transmission of information; programs of training and exchange of inspectors; purchase and modernisation of boats and aircraft for inspections at sea. Commission's contribution and procedural provisions on applying for aid and payment/ reimbursement have been amended to make them clearer and more precise. The fundamental objective of CFP reform was to enable the Commission and member states to tackle weaknesses in enforcement. Augmenting the expenditure in enforcement was the central element of the new regime. Other possible measures against Member States include the taking of preventive measures, suspension of financial assistance and deductions from future fishing opportunities.

At the same time, the Council decided to step up cooperation among the member states and coordination of their inspection activities by establishing a common inspection structure, using the structures, equipment and communications networks provided by the member states. The coordination of theses actions should rest on a European supra-national Agency with important powers in terms of enforcement. This European Control Agency was created in 2005 and has its official seat in Vigo, Spain. The reform also provided for increased powers for the Commission inspectors because it had been found, through the annual reports on infringement of the CFP rules, that the member states prosecuted infringements with varying degrees of diligence and severity**.** This meant that awareness of the importance of penalising non-responsible behaviour should be increased and that it required promotion of a culture of protection of natural resources and inspection, so that effective penalties could be applied. The role of the Commission inspectors, in this sense, was fundamental.

In the context of the reform, the Commission was committed to increase transparency in the information related to the compliance of member states with their enforcement obligations. This is why it has presented a Communication on compliance with the CFP rules which included, for the very first time, a scoreboard which should be updated on an annual basis.

The "scoreboard of compliance" is a pure example of this new environment of transparency and trust. This scoreboard represents a clear and easily accessible source of information about member states compliance. Among other things, it reports on catches taken by their fleets, the capacity and fishing effort of these fleets, and national monitoring and inspection activities. In addition, it sets out the infringement procedures initiated by the Commission, in respect of member states that have failed to comply with certain CFP rules. The scoreboard seeks also to serve as a tool for comparing data about regulation enforcement by member states. Using a number of key indicators, it shows in summary form, the general level of compliance with those rules. The scoreboard contains details, for each of the above subjects, on the legal basis of the member states obligations, their nature and the intervals at which they must be met along with information concerning compliance. It also contains information about actions taken by the European Commission, firstly, under its powers to check the monitoring activities by the competent national authorities and, secondly, in its role as guardian of compliance with Community law.

3. DISCUSSION

Applying theoretical analysis to the guidelines of CFP design and reform, it is suggested the following:

Implementing Community policies in Member States is never easy, especially when myopic individual interests do not match long term collective interests. This is the case of fisheries. Fishermen do not have a greater propensity to altruism than the rest of the society; so, they are little inclined to refrain catches (for the sake of a clear conscience) if they think their competitors are less scrupulous (European Comission/ DGF). That is, without a clear and effective policy of control and enforcement, the Commission is certain that the "Tragedy of the Commons" will result and that overfishing and overcapacity will occur (Hardin (1968), Filipe *et al* (2008), Filipe (2007), Filipe *et al* (2007). The reform of CFP insists in the philosophy of intervention of its early days, when the Commission put the problem of control in terms of ethical reasons: "It's the only way to assure that the sacrifices of some member states in the recovery of the stocks are not in vain because of the irresponsible action of others" (European Commission, 1976).

According to Becker, individuals rationally decide whether or not engage in criminal activities by comparing the expected returns to crime with the legitimate business. His main thesis is that crime is less attractive if the government increases the probability and severity of punishment (Garoupa (2000). The analysis of the Commission proposals seems to give a special attention to the increase of the probability of detection as a means to deter criminal behaviour and increase compliance with regulation. The introduction of severe penalties is not a priority (Coelho *et al* (2008)). Of course they are considered and an important effort is made to define and make clear the legal procedures to penalise the violators. However, the severity of penalties is not in the centre of the European Fisheries Policy. The Commission believes that the financial support will guarantee the indispensable means of surveillance and control to the member states and this will increase the deterrence capacity of control in member states, in a uniform way, and, also, the transparency and trust between partners. But the Commission also knows that legal administrations, in the member states, have significant differences and that judicial machinery has a great inertia. The capacity and efficiency of member states justice is not just a question of financial means devoted to his mission. It has also cultural and historical roots. It's virtually impossible to put all the member states in a uniform position in terms of speed and severity in the application of penalties.

The enforcement issue points out another advantage of private property rights based management - they are self-enforcing. This may be an important step to proceed to the discussion of regulatory instruments. First, if enforcement costs are significant, the more common forms of regulation (direct controls as TACs, mesh size or areas / seasons closures, for example) should require further re-evaluation. Usually they are detracted because they are not economically efficient. But is also commonly recognised that costs of enforcement are weaker in these cases. Second, the analysis of Individual Transferable Quotas (ITQs) reveals the equivalence between ITQs and taxes. But, with the consideration of enforcement costs, this may not hold. In any case, the reduced costs of enforcement support this tool. As the fishermen are given almost private property rights of resource use, this means that some kind of auto-regulation is guaranteed. In theory, this engages fishermen in compliance with the regulation and diminishes enforcement costs. However, in practice, the implementation of ITQs system is confronted with a lot of problems, including illegal behaviour (Copes (1986) and Coelho and Lopes (2000).

CONCLUSION

The European Common Fisheries Policy has proved that there are some important measures that have been implemented to create a more transparent control regime in the European Fisheries.

Besides the efforts made in the sequence of 2003 Reform, several problems, in the CFP control and monitoring regime, still remain. This implies a new reform that gives the member states a more trustable framework and allows to (re)discover the original philosophy of intervention of the Commission that saw the issue of the fisheries enforcement as an ethical question (equating the stocks recovery and the development of sustainable fisheries with a fair distribution of costs and benefits among member states).

In our opinion, this new Reform should be accomplished with a more intense debate on the issue of the regulation tools of the Common regime of conservation and management of fisheries resources. This implies a cost/benefit analysis that confronts the different instruments with their advantages/disadvantages. And this should include, of course, the issue of monitoring and enforcement costs. By now, it isn't, in fact, possible to predict the Commission position about the introduction of management instruments as ITQs, which could introduce economic efficiency in European fisheries and, at the same time, introduce forms of auto-regulation in the fisheries that could reduce, in theory, the enforcement costs.

REFERENCES

ALCHIAN, AND DEMSETZ, (1973), "The Property Rights Paradigm", *Journal of Economic History,* Vol. 33 , pp. 16-27.

BECKER(1968), "Crime and Punishment: An Economic Approach", *Journal of Political Economy,* Vol.86, pp. 169-217.

CHEUNG (1970), "The Structure of a Contract and the Theory of a Non-Exclusive Resource", Journal of Law and Economics, Vol. 13, pp. 49-70.

CHEUNG (1968), "Private Property Rights and Sharecropping", *Journal of Political Economy*, Vol. 76, pp. 1107-1122.

CLARK (1985), Bioeconomic Modelling and Fisheries Management, John Wiley & Sons.

CLARK (1980), "Towards a Predictive Model for the Economic Regulation of Commercial Fisheries", *Canadian Journal of Fisheries and Aquatic Sciences*, Vol.37, pp. 1111-1129.

CLARK AND G. MUNRO (1975), "The Economics of Fishing and Modern Capital Theory: A Simplified Approach", *Journal of Environmental Economics and Management*, Vol. 2, pp. 92-106.

COELHO, FILIPE, FERREIRA AND PEDRO (2008), "Illegal Fishing: An Economic Analysis", *APLIMAT, Journal of Applied Mathematics*, Volume 1, Number 2, pp. 167-173.

COELHO AND LOPES (2000), The Common Fisheries Policy and the Feasibility of ITQs, in Hatcher e Robinson (eds), The Definition and Allocation of Use Rights in European Fisheries, CEMARE/University of Portsmouth.

COPES (1986), "A Critical Review of the Individual Quota as a Device in Fisheries Management", *Land Economics*, Vol. 62, pp. 278-291.

DEMSETZ (1967), "Toward a Theory of Property Rights", *American Economic Review*, Vol. 57, pp. 347-359.

FILIPE, FERREIRA, COELHO AND PEDRO (2008), "Anti-Commons: How tragedies happen. Some cases and the evidences on Fisheries", *China - USA Business Review*, Volume 7, Number 11, pp. 9-13.

FILIPE (2007), The Drama of Fishing Commons: Cournot-Nash Model and Cooperation. Working Paper ISEG/DE. http://www.iseg.utl.pt/departamentos/economia/wp/wp0302007de.pdf. 2007.

FILIPE, FERREIRA AND COELHO (2007), O Drama dos Recursos Comuns nas Sociedades Actuais: à procura de soluções para os Ecossistemas em perigo. Edições Sílabo.

GAROUPA (2000), "Optimal Law Enforcement and Criminal Organization", Seminário do Departamento de Economia Nº 8/2000, ISEG/UTL. 2000

HARDIN (1968), "The Tragedy of the Commons", *Science*, 162, pp.1243-1248.

NOSTBAKKEN (2008), "Fisheries Law Enforcement: A Survey of the Economic Literature", *Marine Policy*, Vol. 32, pp. 293-300.

POLINSKY AND SHAVELL (2000), "The Economic Theory of Public Enforcement of Law", *Journal of Economic Literature*, Vol. 38, pp. 45-76.

STIGLER (1970), The Optimum Enforcement of Laws", *Journal of Political Economy*, Vol. 78, pp. 526-536.

SUMAILA, ALDER AND KEITH (2006), "Global Scope and Economics of Illegal Fishing", *Marine Policy*, Vol. 30, pp. 696-703

SUTINEN AND ANDERSEN (1985), "The Economics of Fisheries Law Enforcement", *Land Economics*, Vol. 61, pp. 387-397.

SUTINEN AND HENNESSEY (1986), "Enforcement: The Neglected Element in Fishery Management", in Miles, Pealy and Stokes (eds), Natural Resource Policy and Management: Essays in Honor of James Crutchfield.

SUTINEN AND GAUVIN (1989), "An Econometric Study of Regulatory Enforcement and Compliance in the Commercial Inshore Lobster Fishery of Massachusetts", in Neher (ed), Rights Based Fishing, Kluwer Academic Publishers.

TIETENBERG (2003), Environmental and Natural Resource Economics, sixth edition, Addison Wesley, 2003.

ELECTRICITY INDUSTRIAL ORGANIZATION: WHAT ABOUT THE STRATEGIC BEHAVIOR OF HYDRO AND THERMAL OPERATORS?

AHLEM DAKHLAOUI
E-mail: Ahlem.Dakhlaoui@ept.rnu.tn
Laboratoire d'Economie et de Gestion Industrielle
École Polytechnique de Tunisie

Abstract

In this paper, we develop a two-period model where we analyze and compare a hydro/thermal electrical system under different industrial organization: monopoly, Cournot competition and collusion; under storage constraint, water availability constraint and thermal turbine capacity constraint. First, we prove that the technological complementarity have an important role in the satisfaction of electricity demand in the different industrial organization. Second, by the analytical resolution, we show that inter temporal private monopoly water transfer from off-peak season to peak season is not high as the same transfer under a public monopoly and therefore this increases the power price. Under Cournot competition, an increase in the peak season implies a water transfers from off peak season to the peak season. The results of collusion depend on availability of water resource and the demand parameters. The results of collusion depend on availability of water resource and the demand parameters.

Keywords: electricity market, industrial organization, monopoly, competition, hydroelectric plants and thermal plants.
Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

I. INTRODUCTION

The market structure of the electricity sector has been characterized by natural monopolies. In the last decade, the market liberalization processes suggests that the electricity production is done by private generators which use different technologies: wind, hydroelectric, solar, thermal, etc. The main target of liberalization is to achieve more cost efficient production, lower electricity prices, better resources allocation and more supply security. The economic literature, there are little number of papers which analyze mixed hydro thermal system problems in a liberalized industry such as Scott et Read (1996) [1], Bushnell (1998) [2], Crampes and Moreaux (2001) [3] and Dakhlaoui and Moreaux (2004) [4]. Ambec and Doucet (2002) [5] compare the centralized and the decentralized industry of the hydraulic system under water turbine capacity constraint and water abundance constraint. They show that the exercise of the capacity of the market can reduce the results envisaged of the deregulation of this industry. In this paper, we use Ambec and Doucet (2002) framework to develop a two-period model where we analyze and compare a hydro/thermal electrical system under different industrial organization: monopoly, Cournot competition and collusion; under storage constraint, water availability constraint and thermal turbine capacity constraint. First, we prove that the technological complementarity have an important role in the satisfaction of electricity demand in the different industrial organization. Second, by the analytical resolution, we show that inter temporal private monopoly water transfer from off-peak season to peak season is not high as the same transfer under a public monopoly and therefore this increases the power price. Under

Cournot competition, an increase in the peak season implies a water transfers from peak season to the peak season. The results of collusion depend on availability of water resource and the demand parameters.

The rest of the paper proceeds as follow. In the section 2, we present the model. Then, in section 3, we deal with public monopoly and private monopoly. Section 4, we analyze the production problem under Cournot competition and collusion of hydro and thermal operators. We conclude in section 5.

II. THE MODEL

We consider two-period model that analyze the production of electricity with hydro (*H*) and thermal plants (*T*). The inverse demand of electricity during period t for any quantity Q_t : $P_t = a_t - b_t Q_t$ with $a_t > 0$ and $b_t > 0$, $t=1,2$. Let is β the discount factor. The hydroelectric firm uses natural inflows of water in order to produce Q_t^H unities of electricity during the period *t*. e_t is the exogenous volume of water supplied in the hydrographic basin during period *t*. It is assumed that one unit of water yields α units of electricity $(\alpha > 0)$. The total production of H during the two periods is :

$$Q_1^H + Q_2^H = \overline{Q}^H = \alpha(e_1 + e_2) \qquad (1)$$

Water available during period *1* can be used to produce electricity during the first period, or can be stored in reservoir for use in the second period. Hence, it is able to produce at no cost during the first period any quantity Q_1^H such that :

$$Q_1^H \leq \alpha e_1 \qquad (2)$$

The volume of water stored in reservoir during the period 1 is used in its entirety to produce electricity in the second period. This volume is bounded by the reservoir capacity, denoted $\overline{s}$. In terms of first period, this storage constraint is written:

$$Q_1^H \geq \alpha(e_1 - \overline{s}) \qquad (3)$$

A production plan is any vector, $Q^H = (Q_1^H, Q_2^H) \in \Re_+ \times \Re_+$, verifying the constraints.

The firm *T* produces power gaze or fuel input. The production of Q_t^T unities of electricity requires an assuming quadratic cost, which is $C_t = \frac{1}{2} c_t (Q_t^T)^2$ with $c_t > 0$ for $t=1,2$. During each period t, the *T* plant is subject to capacity production constraint which is :

$$Q_t^T \leq \overline{Q}^T, \forall t = 1,2 \qquad (4)$$

A production plan of *T* is $Q^T = (Q_1^T, Q_2^T) \in \Re_+ \times \Re_+$, verifying the capacity constraint.

III. MONOPOLISTIC STRUCTURE

A. Public monopoly

The optimal monopoly production plan is a solution of the following problem:

$$Max_{Q_1^{H_T}, \{Q_t^{T_H}\}_{t=1,2}} \int_0^{Q_1} P_1(x)dx - C_1(Q_1^T) + \beta \left[\int_0^{Q_2} P_2(x) - C_2(Q_2^T) \right]$$

With respect to (1), (2), (3) and (4).

Denoting by $\overline{\lambda}^{H_T}$, $\underline{\lambda}^{H_T}$ and $\lambda_t^{T_H}$ the Lagrange multipliers associated respectively to the water availability constraint at the off-peak season, the storage constraint and the maximum

capacity of the plant *T*.

The first order conditions yield that the monopoly equalizes the marginal benefit net of the one unity of water release for the *H* plant at the season *1* with the actualized marginal benefit comes from the storage of water until the peak period:

$$P_1(Q_1) - \overline{\lambda}^{H_T} = \beta P_2(Q_2) - \underline{\lambda}^{H_T}$$

From the first order condition, we have:

$$C_1^{'}(Q_1^{T_H}) + \lambda_1^{T_H} - \beta C_2^{'}(Q_2^{T_H}) - \lambda_2^{T_H} = \overline{\lambda}^{H_T} - \underline{\lambda}^{H_T}$$

1) If the production maximal capacity is bounded during the peak season

Suppose that during the off-peak season, the demand is not high enough to reach the maximal capacity and that in the second period, the constraint (4). In that case, $\lambda_1^{T_H} = 0$ and $\lambda_2^{T_H} > 0$. So, the first order conditions relative to thermal plants are :

$$P_1(Q_1) = C_1^{'}(Q_1^{T_H})$$

$$\beta P_2(Q_2) = \beta C_2^{'}(Q_2^{T_H}) + \lambda_2^{T_H}$$

The condition (2) is reduced as such:

$$C_1^{'}(Q_1^{T_H}) - \beta C_2^{'}(Q_2^{T_H}) - \lambda_2^{T_H} = \overline{\lambda}^{H_T} - \underline{\lambda}^{H_T}$$

The maximal capacity of the thermal plant is reached in the peak season. To satisfy the peak demand, the use of this plant in the off-peak season will intensify as will the storage of inflows to be sure to latter satisfy the peak demand.

Consequently, as $\lambda_1^{T_H} = 0$, to balance the equation :

- $\overline{\lambda}^{H_T}$ increases, so more water is stored because the marginal cost of the water's immediate exploitation increases.
- $\underline{\lambda}^{H_T}$ decreases, so the marginal cost of the water storage decreases.

2) If the constraints are never saturated

Proposition 1: *The optimal mixed system operating at the public monopoly case is given by*:

$Q_1^{H_T} = \frac{a_1 - \beta a_2}{b(1+\beta)} + \frac{\beta}{1+\beta}\overline{Q}^H$, $Q_1^{T_H} = \beta\frac{a_1 + a_2 - b\overline{Q}^H}{(b+c)(1+\beta)}$, $Q_2^{H_T} = \frac{\beta a_2 - a_1}{b(1+\beta)} + \frac{\overline{Q}^H}{1+\beta}$ and $Q_2^{T_H} = \frac{a_1 + a_2 - b\overline{Q}^H}{(b+c)(1+\beta)}$.

The *optimal* production by the *T* plant, in two periods depends on the water turbine capacity of *H* plant, the demand characteristics and the production cost of the plant *T*. So, the presence of the plant H makes dynamic the problem of the thermal plant. Consequently, any increase in *H* turbine capacity equal to $\Delta\overline{Q}^H$, reduces the plant *T* production during the off-peak season by a quantity equal to $\frac{b}{(b+c)(1+\beta)}\Delta\overline{Q}^H$. A free technology (*H*) supplementary exploitation of $\frac{\beta}{1+\beta}\Delta\overline{Q}^H$, compensates for the decision to reduce expensive technology exploitation (thermal). The substitution of the expensive technology by the free one reduces the price by a quantity equal to $\frac{\beta bc}{(1+\beta)(b+c)}\Delta\overline{Q}^H$. To satisfy a supplementary demand during the peak season, the producer uses both technologies. The producer reduces indeed his water use during the off-peak

season of $\frac{\beta}{b(1+\beta)}\Delta a_2$, in order to increase his plant *T* production by a quantity equal to $\frac{\beta}{(b+c)(1+\beta)}\Delta a_2$. The technological complementarity allows the storage of the water to face any increase of the peak season demand. In this season, the producer uses indeed all the inter temporal supplementary transfer of water ($\frac{\beta}{b(1+\beta)}\Delta a_2$) and increases the *T* production by a quantity equal to $\frac{1}{(b+c)(1+\beta)}\Delta a_2$ in order to satisfy the peak supplementary demand. We notice that the producer increases his use of the thermal plant in different proportions ($\Delta Q_1^{T_H} < \Delta Q_2^{T_H}$).

The technological complementarity allows the substitution of the expensive technology by the free one.

B. Private monopoly

The optimal production plan maximizes :

$$Max_{Q_1^{H_T},\{Q_t^{TH}\}_{t=1,2}} \Pi^{HT} = P_1.(Q_1^{H_T} + Q_1^{T_H}) - \frac{c}{2}(Q_1^{T_H})^2 + \beta\left[P_2.(Q_2^{H_T} + Q_2^{T_H}) - \frac{c}{2}(Q_2^{T_H})^2\right]$$

with respect to (2), (3) and (4).

The first order conditions with respect to *H* and *T* plant imply: $\Pi_1^{HT'} - \overline{\lambda}^{H_T} = \beta\Pi_2^{HT'} - \underline{\lambda}^{H_T}$, $P_1^{'}.(Q_1^{H_T} + Q_1^{T_H}) + P_1 - c.Q_1^{T_H} = \lambda_1^{T_H}$ and $P_2^{'}.(Q_2^{H_T} + Q_2^{T_H}) + P_2 - c.Q_2^{T_H} = \frac{\lambda_2^{T_H}}{\beta}$.

If the constraints are never bounded, we have the following result.

Proposition 2: *The private monopoly equilibrium is given by*: $Q_1^{H_T} = \frac{a_1 - \beta a_2}{2b(1+\beta)} + \frac{\beta}{1+\beta}\overline{Q}^H$, $Q_1^{T_H} = \beta\frac{(a_1+a_2) - 2b\overline{Q}^H}{(1+\beta)(2b+c)}$, $Q_2^{H_T} = \frac{\beta a_2 - a_1}{2b(1+\beta)} + \frac{\overline{Q}^H}{1+\beta}$ and $Q_2^{T_H} = \frac{a_1 + a_2 - 2b\overline{Q}^H}{(2b+c)(1+\beta)}$.

To satisfy a supplementary peak season demand, the producer reduces his *H* production at the season *1* by $\frac{\beta}{2b(1+\beta)}\Delta a_2$ in order to store water resources which increase the production of *T* by $\frac{\beta}{(1+\beta)(2b+c)}\Delta a_2$ units that compensates partially for the decrease of the plant *H* use. The water transfer face the demand growth by a supplementary production via the *H* plant ($\frac{\beta}{2b(1+\beta)}\Delta a_2$) and the T plant ($\frac{\Delta a_2}{(1+\beta)(2b+c)}$).

We find that the water transfer under a private monopoly is lower than the one under the public monopoly. This behavior guarantees a high price to the private monopoly during the peak season. Any increase of turbinage capacity equal to $\Delta\overline{Q}^H$ increases his usage during the off-peak

season and the peak season by respectively, $\frac{\beta}{1+\beta}\Delta\overline{Q}^H$ and $\frac{1}{1+\beta}\Delta\overline{Q}^H$. On the other hand, with technological complementarity, the producer reduces his usage of T plant in the off-peak and the peak season by respectively, a quantity equal to, $\frac{2\beta b}{(1+\beta)(2b+c)}\Delta\overline{Q}^H$ and $\frac{2b}{(1+\beta)(2b+c)}\Delta\overline{Q}^H$.

Between the public monopoly and the private one, we notice a disproportion on the substitution of T by H during the two periods. The presence of the hydroelectric plant makes the exploitation problem of the thermal one dynamic, under the private and the public monopoly. In both cases, the technological complementarity played an important role in the efficiency of the electricity supply. The private monopoly tends to transfer less water than the public one

IV. COMPETITION AND COLLUSION

We consider a decentralized electrical industry where first the two operators compete. Second, we analyze the collusion case.

A. Cournot competition

The *H* production $q^{H*} = (q_1^{H*}, q_2^{H*})$ maximizes the inter temporal profit under the *H* constraints (2) and (3) :

$$Max_{q_1^H} \Pi^H = P_1(q_1^H, q_1^T).q_1^H + \beta P_2(q_2^H, q_2^T).q_2^H$$

The first order conditions yields that the producer seeks to equalize marginal revenues to marginal costs:

$$P_1^{'}q_1^H + P_1 - \beta(P_2^{'}q_2^H + P_2) = \overline{\lambda}^H - \underline{\lambda}^H$$

The *T* operator production $q^{T*} = (q_1^{T*}, q_2^{T*})$ that maximizes the inter temporal profit with respect to the T capacity constraint satisfy the first order condition:

$$\beta(P_2 + P_2^{'}q_2^T - C_2^{'}(q_2^T)) = \lambda_2^T$$

The Nash equilibrium strategies of the *T* and *H* producers, respectively q^{T*} and q^{H*}. In order to resolve the equilibrium we distinguish two different cases :

Case 1: If one of the H plant's constraints is bounded:

We distinguish two cases.

- The inflows of water are scarce: The producer exhausts the inflows of the off-peak season. The first order condition is reduced to the following equality: $\Pi_1^{H'} - \overline{\lambda}^H = \beta.\Pi_2^{H'}$ and the *H* production is $(\alpha e_1, \alpha e_2)$.

- The inflows of water are abundant: The *H* producer chooses to store water for a future use. The stock reaches the maximum capacity of storage $\overline{s}$. The first order condition is reduced to : $\Pi_1^{H'} = \beta\Pi_2^{H'} - \underline{\lambda}^H$. Besides, the inter temporal transfer of water implies the vector of production levels $(\alpha(e_1 - \overline{s}), \alpha(e_2 + \overline{s}))$.

Case2: If the constraint is never bounded:

Proposition 3: The equilibrium Cournot game is given

$$q_1^{H*} = \frac{(a_1 - \beta a_2)(b+c)}{b(1+\beta)(3b+2c)} + \frac{\beta \bar{q}^{H}}{(1+\beta)} \quad , \quad q_1^{T*} = \frac{a_1\left[b(2+3\beta)+c(1+2\beta)\right]+a_2\beta(b+c)}{(2b+c)(1+\beta)(3b+2c)} - \frac{\beta b \bar{q}^{H}}{(2b+c)(1+\beta)}$$

$$q_2^{H*} = \frac{(\beta a_2 - a_1)(b+c)}{b(1+\beta)(3b+2c)} + \frac{\bar{q}^{H}}{(1+\beta)} \quad \text{and} \quad q_2^{T*} = \frac{a_2\left[b(3+2\beta)+c(2+\beta)\right]+a_1(b+c)}{(2b+c)(1+\beta)(3b+2c)} - \frac{b \bar{q}^{H}}{(2b+c)(1+\beta)}$$

To face any increase in the peak season's demand equal to Δa_2 unities of electricity, the hydroelectric producer store during the off-peak season a quantity equal to $\frac{\beta(b+c)}{b(1+\beta)(3b+2c)}\Delta a_2$ of water inflows, while the *T* production increase in the same season by $\frac{\beta(b+c)}{(2b+c)(1+\beta)(3b+2c)}\Delta a_2$ unities of electricity.

This transfer of water is completely used by *H* producer during the peak season. However, during this season the T producer increases his production by a quantity equal to $\frac{b(3+2\beta)+c(2+\beta)}{(2b+c)(1+\beta)(3b+2c)}\Delta a_2$.

At this stage, we can assess that technological heterogeneousness allows more supply security.

B. Collusion

We assume that both firms, *H* and *T*, produce quantities that maximized the joint profit. $q^{Tm} = (q_1^{Tm}, q_2^{Tm})$ and $q^{Hm} = (q_1^{Hm}, q_2^{Hm})$ would be the solutions of the maximization problem with regard to the variables of decision; q_1^{Hm}, q_2^{Hm}, q_1^{Tm} and q_2^{Tm} :

$$Max_{q_1^H, \{q_t^T\}_{t=1,2}} \Pi^{jo\,\text{int}} = \Pi_1^H(P_1(q_1^H, q_1^T), q_1^H) + \beta\Pi_2^H(P_2(q_2^H, q_2^T), q_2^H) +$$
$$\Pi_1^T(P_1, q_1^T, C(q_1^T)) + \beta\Pi_2^T(P_2, q_2^T, C(q_2^T))$$

With respect to (2), (3) and (4).

The *H* producer must equalize the marginal joint profit of the immediate exploitation of one unity of water in the off-peak season to actualized marginal joint profit water storage until the peak season:

$$\frac{\partial \Pi_1^{jo\,\text{int}}}{\partial q_1^H} - \bar{\lambda}^H = \beta \frac{\partial \Pi_2^{jo\,\text{int}}}{\partial q_1^H} - \underline{\lambda}^H$$

The thermal producer equalizes the marginal joint profit to zero during the off-peak season. During the peak-season, he equalizes the actualized marginal profit and the marginal cost of the capacity constraint bounding:

$$\frac{\partial \Pi_1^{jo\,\text{int}}}{\partial q_1^T} = 0 \quad \textit{and} \quad \frac{\partial \Pi_2^{jo\,\text{int}}}{\partial q_2^T} = \frac{\lambda_2^T}{\beta}$$

If the constraints are never bonded, the first order conditions implies the following functions :

$$\begin{cases} \hat{q}_1^H = \dfrac{a_1 - \beta a_2}{2b(1+\beta)} + \dfrac{\beta}{1+\beta}\bar{q}^{H} - \dfrac{1}{1+\beta}q_1^T + \dfrac{\beta}{1+\beta}q_2^T \\ \hat{q}_2^H = \dfrac{\beta a_2 - a_1}{2b(1+\beta)} + \dfrac{1}{1+\beta}\bar{q}^{H} + \dfrac{1}{1+\beta}q_1^T - \dfrac{\beta}{1+\beta}q_2^T \end{cases}$$

$$\begin{cases} \hat{q}_1^T = \dfrac{a_1 - 2bq_1^H}{2b+c} \\ \hat{q}_2^T = \dfrac{a_2 - 2bq_2^H}{2b+c} \end{cases}$$

At the collusion case, *T* quantities fluctuate only according to quantities of the same season of the competitor. Whereas, hydro producer chooses his level of production according to both thermal quantities. We notice also that, correlation between correlation between competitors quantities is more intense than in competition ($\dfrac{\partial \hat{q}_t^T}{\partial \hat{q}_t^H} = 2\dfrac{\partial q_t^{T*}}{\partial q_t^H}$ *and* $\dfrac{\partial \hat{q}_t^H}{\partial \hat{q}_t^T} = 2\dfrac{\partial q_t^{H*}}{\partial q_t^T}$ **).**

Proposition 4: *If competitors are indifferent between the two seasons (* $\beta = 1$ *), the solution of collusion program is:*

$\hat{q}_1^H = \dfrac{a_1 - a_2}{4b} + \dfrac{\bar{q}^{-H}}{2}$, $\hat{q}_1^T = \dfrac{a_1 + a_2}{2(2b+c)} - \dfrac{b}{2b+c}\bar{q}^{-H}$, $\hat{q}_2^H = \dfrac{\bar{q}^{-H}}{2} - \dfrac{a_1 - a_2}{4b}$ and $\hat{q}_2^T = \dfrac{a_1 + a_2}{2(2b+c)} - \dfrac{b}{2b+c}\bar{q}^{-H}$.

Prices are given by: $\hat{P}_1 = \dfrac{a_1(4b+3c) + ca_2 - 2bc\bar{q}^{-H}}{4(2b+c)}$ and $\hat{P}_2 = \dfrac{a_2(4b+3c) + ca_1 - 2bc\bar{q}^{-H}}{4(2b+c)}$.

The collusion is possible only if the maximum capacity of the plant *H* belongs to the interval $\left[\dfrac{a_2 - a_1}{2b}, \dfrac{a_2 + a_1}{2b}\right]$.

The hydroelectric firm produces more during the peak season. Whereas, the thermal firm produces similar quantities during the two seasons ($\hat{q}_2^H > \hat{q}_1^H$ and $\hat{q}_1^T = \hat{q}_2^T$). The maximum capacity affects (negatively) the price. An increase in the maximum capacity implies a more important decrease of the price in competition. Thus, in collusion, the price is less elastic according to the maximum capacity.

V. CONCLUSION

In this paper, we prove that the technological complementarity has an important role in the satisfaction of electricity demand in the different industrial organizations. By the analytical resolution, we show that inter temporal private monopoly water transfer from off peak to peak season is not as high as this same transfer under a public monopole and therefore this increases the power price. Under Cournot competition, an increase in the peak season implies a water transfer from off peak season to the peak season. The results of collusion depend on the availability of water resources and the demand parameters. We remark that in contrast to the competition case, collusion's quantities of *H* are not dependable on *T*'s costs. Thus, if a variation in the production cost of thermal energy has no effect on hydroelectric quantities, then and under these hypotheses, it is very likely that the competition is eased by a collusive agreement.

VI. REFERENCES

SCOTT T. J. et READ E. G.(1996) : Modelling hydro reservoir operation in deregulated electricity market, International Transactions in Operational Research, vol 3 n° 3, pp 243-254.

BUSHNELL . J . B.(1998) : Water and power: Hydroelectric resources in the Era of competition in the western U.S, POWER working paper PWP-056 r. University of California energy institute, July.www.ucei.berkeley.edu/ucei.

CRAMPES C. et MOREAUX M. (2001) : Water resources and power generation, International Journal of Industrial Organization, 19, pp975-997.

DAKHLAOUI A. and M. MOREAUX (2004): « The mixed hydrothermal system operating under

regulated and deregulated electrical Industry », International Network for Economic Research Studies: Competition policy in energy market, n°10, pp 77-94.

AMBEC A. et DOUCET J. A. (2002): Competition along a river: Decentralizing hydropower production, mémo GREEN université Laval, December.

DYNAMIC GAMES IN THE WHOLESALE ELECTRICITY MARKET

Ahlem Dakhlaoui*

LEGI

Abstract: *In this paper, we analyze infinite discrete-time games between hydraulic and thermal power operators in the wholesale electricity market. Two types of games are considered: Cournot closed-loop game and Stackelberg closed-loop game. We consider a deregulated electrical industry where certain demand is satisfied by hydraulic and thermal technologies. The hydraulic operator decides the production in each season of each period that maximizes the sum of expected profit from power generation with respect to the stochastic dynamic constraint on the water stored in the dam, the environmental constraint and the non-negative output constraint. In contrast, the thermal plant is operated with quadratic cost function, with respect to the capacity production constraint and the non negativity output constraint. This paper is devoted to the numerical computations of equilibrium strategies and value function in each kind of games. We show that under imperfect competition, the hydraulic operator has a strategic storage of water in the peak season. Then, we quantify the strategic inter annual and intra annual water transfer in the both games and we compare the numerical results. Under Cournot closed-loop game, we show that the traditional principle of least-cost operation is inverted at the binding capacity constraint of thermal operator. Finally, under Stackelberg closed-loop game, we show that thermal operator can restrict the hydraulic output without compensation. The technical complementarities and Stackelberg competition may distorted the traditional "merit order" operating principal.*

Key Words: water resource, wholesale electricity market, thermal system, inflows uncertainty, Cournot closed-loop game and Stackelberg closed-loop game.

JEL Classification: Q4, Q25, C73, C61.

---

* Laboratoire d'Economie et de gestion industrielle, école polytechnique de Tunisie. E-mail : Ahlem.Dakhlaoui@ept.rnu.tn

DYNAMIC GAMES IN THE WHOLESALE ELECTRICITY MARKET

1-INTRODUCTION

In the last decades, the electricity industry set out the very important structural reform by inserting a competition in generation of high voltage. The transmission and distribution of electricity have been considered by most economists as "natural monopoly". These changes have split the single state owned firm into several private owned firms that compete in a single market: the wholesale electricity market. The economic analysis has been turned toward some problems like as congestion of transmission electric power network, concentration measures, horizontal market power, electricity spot market, bilateral transaction, nodal pricing in a simple and a complex network, etc; and which has been made in a static framework compatible with thermal power generation system. Thus, there are few analyses concerning the restructuring of the electricity industry with heterogeneous technologies. This paper takes part in the new research field concerning the mixed hydrothermal system operating problems under deregulated industry. The objectives of the paper are also to compute numerical equilibrium strategies in two kinds of games and to find the conditions under which the imperfect competition breeds a distortion to traditional least-cost rule used in the mixed system operating.

In several countries, the liberalization process had led to different asymmetric duopoly. The first asymmetric case concerned the British electricity supply industry which has been transferred to two successors companies: National power and Gen power. In this case, Green and Newbery (1992) have shown that "... in an asymmetric case, less output would be sold at higher price, and industry operating costs will be further raised for any level of output since the stations will no longer operate in "merit order". The second asymmetric case is spreading across countries owned with significant hydraulic resources like as Norway, New Zealand, Western United States, etc. In the economic literature, there are few studies on the strategic behavior of hydraulic and thermal operators after the reform of the electricity industry. In the New Zealand case, Scott T. J and E. G. Read (1996) have developed a dual dynamic programming approach in order to characterize an optimal hydraulic schedule for a strategic

firm that controls all the storage hydro capacity in a market with other Cournot producer that controls the thermal generation. They show that «... there is relatively little loss in coordination efficiency if, but only if, there is a high level of contracting and / or a high effective elasticity". The same problem has been analyzed in J. Bushnell (1998) model which developed a sub-game perfect equilibrium of a multi-period Cournot game between strategic producer who controls both hydro and thermal technologies. In the latter, the author has shown that hydraulic releases are directly done in contradictory to the principle of least-cost production since this decision appears as profitable for certain firms. Also, C. Crampes and M. Moreaux (2000) have analyzed the strategic behavior of thermal and hydraulic operators in closed-loop game and open-loop game with two periods. The authors have shown how the presence of the hydroelectric station changes the optimal as well as the market equilibrium outputs of the thermal station. They found that the thermal plant facing hydro plants has to be managed as if it were dynamically connected. The interrelation between heterogeneous technologies can imply other results. Indeed, the technological complementation may incite, under some conditions, operators to behave strategically in the electricity market and in the inverse to the traditional "merit order" operation.

None of theses models take into account the uncertainty in hydraulic resource evolution. The objective of this paper is to develop a multi-periods game between thermal and hydraulic operators in an uncertain framework. We based our analyses on the C. Crampes and M. Moreaux model in order to develop a more realistic asymmetric case. Indeed, the main characteristic of the thermal plants is the possibility to supply certain but an expensive output. Regarding hydraulic plant, its output is linked to a sub-renewal resource with intrinsic uncertain profile. Before the inflows realization, hydraulic power is provisory an exhaustible resource and any additional release reduces the residual water quantity stored in the dam. But, once the inflows are realized, any quantity of water devoted for power generation can be totally or partially renewed by the natural process. We think that this kind of uncertainty and the technical complementation may be in favor of strategic behavior particularly in higher priced peak market. Consequently, this may distort the traditional "merit order" rule and which is against the target of the structural change in the electricity industry.

The rest of the paper proceeds as follow. In the section 2, we present the model. Then, in section 3, we deal with a Cournot closed-loop game between thermal and hydroelectric operators in the wholesale electricity market. Section 4 is devoted to the numerical analysis of Stackelberg closed-loop game between the two operators. Next, we compare the numerical results of imperfect competition. We conclude in section 5.

THE MODEL

Consider a multi-periods model where electric power is generated by two technologies: thermal technology (T) and hydraulic technology (H). The total output is used to satisfy a known demand. Since electric power demand is characterized by fluctuation, then we divide time into infinite years indexed by $t = 0,1,...$; and the year is divided into two seasons ($j = \{h,l\}$, the peak season (h) and the off-peak season (l). Let $P_j\left(q_{jt}\right)$ be the inverse demand function in the season j. It is assumed to be a linear demand function: $P_j\left(q_{jt}\right) = a_j - q_{jt}^H - q_{jt}^T$, where q_{jt}^H is the electricity demand satisfied by the hydroelectric technology, q_{jt}^T is the electricity demand satisfied by the thermal technology and a_j is a positive constant representing the demand characteristic in each season. From year to year, we take the assumption of the stationary of electrical demand function. Since in each season the output of two plants is homogenous then total electric power generation is written as the some of thermal plant output, q_{jt}^T and hydroelectric plant output q_{jt}^H : for a period t and a season j, $q_{jt} = q_{jt}^H + q_{jt}^T$.

The hydraulic plant uses as input a sub-renewable resource stored in a dam. In every year t, the total output must satisfy the following technical constraint: $S_{t+1} = S_t \quad q_t^{hH} \quad q_t^{lH} \quad f_{t+1}$

be quadratic: $-\frac{1}{2}\left(S^* - S_{t+1}\right)^2$. This assumption is compatible also with a mobile dam[2] and a fixed dam. First, if the dam is mobile, then S^* represents the maximum capacity of the fixed part which to simplify is confounded with maximum alert level. Consequently, any current stock S_t beyond the level S^* represents the supplementary water resource stored by closing the gates in order to avoid flood. Second, if the dam is a fixed one, then the penalty cost paid by producer is due to bad release policy not compatible with flooding period. This case is observed in July1996 in Kénogami dam located in sanguinary region in Quebec. During this period, the planner observed an abundant in flows in three days which has affected the water level. The water level attends 166,08 m which exceeds the maximum capacity of the dam equals to 165,67 m. Since the planner does not adopt a release policy compatible with this flooding period, then this over taking of maximal capacity cause a flood with enormous damages. Consequently, the quadratic form penalty cost compatible with the area flooded around the dam. Now, if the expected storage at the end of a period t is less then the level S^*, then it causes a negative externality on other activity around the dam such as navigation, agriculture...etc.

The technical characteristics of the thermal plant are the following. At each season j during the period t, the generation of q_{jt}^T units of energy from thermal plant needs a total cost equals to $C\left(q_{jt}^T\right)$. We assume that the thermal operating cost function is quadratic on $\left[0, \bar{q}^T\right]$: For all $q_{jt}^T \in \left[0, \bar{q}^T\right]$ and for all season $j \in \{h,l\}: C\left(q_{jt}^T\right) = \frac{1}{2} c_j \left(q_{jt}^T\right)^2$. Where c_j is a positive constant for all season $j \in \{h,l\}$ and $\bar{q}^T$ is a known and a positive constant representing the installed capacity. By this hypothesis we neglect the depreciation problem linked to thermal plant operating.

We consider a wholesale electricity market with duopolistic structure. We assume that each operator controls one process. The new industrial structure is represented as following:

---

[2] More information about mobile dam and a fixed dam are available on line at http://retd.edf.fr/futur/publications/disponible.htm.

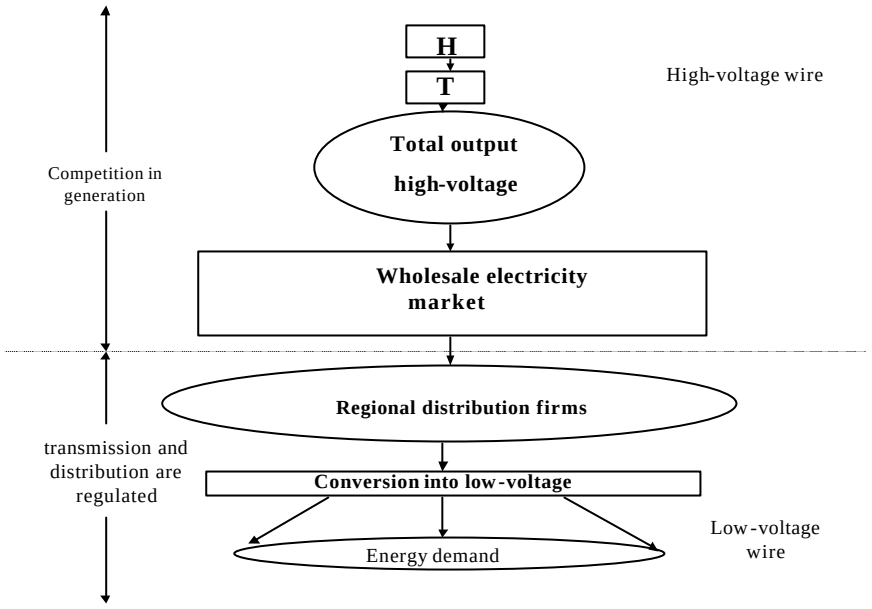


Fig 1: The new structure of electricity industry.

COURNOT CLOSED-LOOP GAME

In this section, we assume that the two operators compete in Cournot closed-loop game in the electricity market.

The thermal operator chooses the output in the peak and off-peak season of every period that maximizes his profit and where the hydroelectric operator output q_{jt}^{H}, for $j \in \{l,h\}$ is taken as given. The thermal operator solves the following problem:

$$\max_{\{q_{lt}^{T},q_{ht}^{T}\}} P_h\left(q_{ht}^{T}+q_{ht}^{H}\right)q_{ht}^{T}-C_h\left(q_{ht}^{T}\right)+P_l\left(q_{lt}^{T}+q_{lt}^{H}\right)q_{lt}^{T}-C_l\left(q_{lt}^{T}\right) \qquad (1)$$

With respect to the capacity constraint and the non negativity constraint:

$$0 \le q_{jt}^{T} \le \bar{q}^{T} \quad \text{, for all } \; j \in \{h,l\} \; \text{ and all } \; t = 0,1,... \qquad (2)$$

The thermal operator problem is a static one. The first order conditions are:

$$\frac{\partial P_j}{\partial q_{jt}^{T}}\left(q_{jt}^{H}+q_{jt}^{T}\right)q_{jt}^{T}+P_j\left(q_{jt}^{H}+q_{jt}^{T}\right)-\frac{dC_j}{dq_{jt}^{T}}\left(q_{jt}^{T}\right)-m_{jt}^{T}+q_{jt}^{T}=0 \qquad (3)$$

$$m_{jt}^{T}\left(\bar{q}^{T}-q_{jt}^{T}\right)=0 \;\; , \;\; q_{jt}^{T}q_{jt}^{T}=0 \qquad (4)$$

$$m_{jt}^{T} \ge 0 \;\; , \;\; q_{jt}^{T} \ge 0 \; , \; q_{jt}^{T} \ge 0 \qquad (5)$$

Where m_{jt}^{T} and q_{jt}^{T} are the multipliers associated respectively to the capacity constraint and the non-negativity constraint.

The first order conditions give the reaction function of the thermal operator in each season of the period t:

$$R_{jt}^{T}\left(q_{jt}^{H}\right)=\frac{a_j-q_{jt}^{H}}{2+c_j} \qquad (6)$$

Let $R_{jt}^{T}\left(q_{jt}^{H}\right)$ be the reaction function of thermal operator in season j of period t. Given the assumption on the inverse demand function and the cost function, then the reaction function is a strictly decreasing one: $\frac{dR_{jt}^{T}}{dq_{jt}^{H}}=-\frac{1}{\left(2+c_j\right)^2}<0$.

The hydraulic operator maximizes the sum of expected profit from power generation under the stochastic dynamic constraint on the water stored in the dam, the environmental constraint and the non-negativity output constraint. The hydraulic operator solves the following problem, where S_t, q_{jt}^{T} are taken as given, for all $j\in\{h,l\}$:

$$\max_{\left\{q_{lt}^{H},q_{ht}^{H}\right\}_{t=0}^{+\infty}} E_0\left\{\sum_{t=0}^{+\infty}P_h\left(q_{ht}^{T}+q_{ht}^{H}\right)q_{ht}^{H}+P_l\left(q_{lt}^{T}+q_{lt}^{H}\right)q_{lt}^{H}-\frac{1}{2}\left(S^{*}-S_{t+1}\right)^2\right\}$$

With respect to the dynamic constraint:

$$S_{t+1}=S_t-q_t^{hH}-q_t^{lH}+f_{t+1}, \quad \text{for all } t=0,1,...$$

$$S_0 \text{ is given.}$$

The closed-loop strategy of hydraulic operator satisfies the Bellman equation:

$$J^{c}\left(S_t\right)=\max_{\left\{q_{lt}^{H},q_{ht}^{H}\right\}_{t=0}^{+\infty}}\sum_{t=0}^{+\infty}P_h\left(q_{ht}^{T}+q_{ht}^{H}\right)q_{ht}^{H}+P_l\left(q_{lt}^{T}+q_{lt}^{H}\right)q_t^{H}+E_t\left\{J^{c}\left(S_{t+1}\right)-\frac{1}{2}\left(S^{*}-S_{t+1}\right)^2\right\}$$

Where $E_t\left(\cdot\right)$ is the mean operator conditional to the information at period t which includes S_t and $J^{c}\left(\cdot\right)$ is the current value function associated to the stochastic dynamic program. It is assumed to be quadratic with unknown coefficients b^c and B^c:

$$J^{c}\left(S_t\right)=b^{c}S_t-\frac{1}{2}B^{c}\left(S_t\right)^2$$

In the stochastic dynamic programming terminology, this solution is called the closed-loop strategy. Under the above hypothesis, the equilibrium strategy exists and it is unique[3].

---

[3] See Blume L. Easley D. and M. O'Harra (1982).

The first order condition in the season j at period t is:

$$\frac{\partial P_j}{\partial q_{jt}^H}\left(q_{jt}^H + q_{jt}^T\right)q_{jt}^H + P_j\left(q_{jt}^H + q_{jt}^T\right) = V^c{}_{t+1}$$

$$P^j\left(\frac{1}{h_{jt}}+1\right) = V_{t+1}^c$$

Where $V_{t+1}^c = E_t\left(J'^c\right) - E_t\left\{S_{t+1} - S^*\right\}$ represents the net marginal value of the hydraulic resources in stock under the decentralized industry. Since J^c is a concave function then V_{t+1}^c is an increasing function in q_t^{jH} and a decreasing function in S_t[4]. We denote h_{jt} the demand elasticity in season j of period t. We assume that h_{jt} is negative and $h_{ht} < h_{lt}$.

The combination of the two first order conditions implies

$$q_{jt}^H = -\frac{1}{3+B^c}q_{jt}^T + \frac{1+B^c}{3+B^c}S_t - \frac{1+B^c}{3+B^c}q_{j't}^H + \frac{a_{jt} - b^c - S^* + \left(1+B^c\right)\overline{f}}{3+B^c}$$

For all $j \neq j'$ and $(j, j') \in \{h,l\}\times\{h,l\}$; where $\overline{f}$ represents the inflows mean.

This equation gives the hydroelectric output in the peak season as a function of the thermal operator output in the same season, the current state of the hydraulic storage and the water release in the off-peak season at the period t.

The resolution of Bellman equation by undetermined coefficient method gives the equilibrium strategy of each player in period t[5]. The results are presented in the following proposition. The resolution of this equations system determines the hydraulic operator strategy in each season as a function of the current stock: $q_t^{Hc*}(S_t) = \left[q_{ht}^{Hc*}(S_t), q_{lt}^{Hc*}(S_t)\right]$. Next, we replace this strategy in the reaction function of thermal operator; we find that this latter is written as a function of the current stock of hydraulic operator $q_{jt}^{Tc} = R_{jt}\left[q_{jt}^{Hc}(S_t)\right]$. The equilibrium strategy of thermal operator in the period t is written as: $q_t^{Tc*}(S_t) = \left[q_{ht}^{Tc*}(S_t), q_{lt}^{Tc*}(S_t)\right]$.

---

4 $\frac{\partial V_{t+1}^c}{\partial S_t} = E_t\left\{\frac{d^2J^c}{S_{t+1}^2}\right\} - 1 < 0$ and $\frac{\partial V_{t+1}^c}{\partial q_t^{jH}} = -\frac{\partial V_{t+1}^c}{\partial S_t} > 0$.

Proposition 1

In a closed-loop game, the thermal and the hydroelectric operators output in the peak season and in the off-peak season is given by the following equations:

$$q_{ht}^{Hc*} = \frac{25\left(1+B^c\right)}{97+52B^c} S_t + \frac{4\left(8+3B^c\right)a_{ht} - 10\left(1+B^c\right)a_{lt} - 25\left[\left(1+B^c\right)\overline{f} - S^* - b^c\right]}{97+52B^c}$$

$$q_{ht}^{Tc*} = -\frac{5\left(1+B^c\right)}{97+52B^c} S_t + \frac{\left(13+8B^c\right)a_{ht} + 2\left(1+B^c\right)a_{lt} + 5\left[\left(1+B^c\right)\overline{f} - S^* - b^c\right]}{97+52B^c}$$

$$q_{lt}^{Hc*} = \frac{27\left(1+B^c\right)}{97+52B^c} S_t + \frac{2\left(14+5B^c\right)a_{lt} - 12\left(1+B^c\right)a_{ht} - 27\left[\left(1+B^c\right)\overline{f} - S^* - b^c\right]}{97+52B^c}$$

$$q_{lt}^{Tc*} = -\frac{9\left(1+B^c\right)}{97+52B^c} S_t + \frac{\left(23+14B^c\right)a_{lt} + 4\left(1+B^c\right)a_{ht} + 9\left[\left(1+B^c\right)\overline{f} - S^* - b^c\right]}{97+52B^c}$$

$$B^c = 2.0527 \text{ and } b^c = 2.96\times10^{-4}\overline{f} - S^* - 1.29a_{ht} - 0.37a_{lt} \blacksquare$$

We remark that both thermal operator strategy and hydraulic operator strategy depend on current stock S_t, on the inflows mean $\overline{f}$, on the storage level S^* and on the electric demand in the peak season and off-peak season. The comparative static in the regulated industry (m) [6] and the Cournot competition case (c) is given by the following table:

t	*Systems*	q_{ht}^{H*}	q_{ht}^{T*}	q_{lt}^{H*}	q_{lt}^{T*}
S_t	c	0,374	-0,075	0,405	-0,134
	m	0,322	-0,08	0,48	-0,24
a_{ht}	c	0,118	0,169	-0,349	0,115
	m	0,6	0,1	-0,165	0,3
a_{lt}	c	-0,195	0,038	0,187	0,269
	m	-0,11	0,22	0,4	0,3
$\overline{f}$	c	-1,9· 10? ³	-0,03	-0,4· 10? ³	-3,05.10? 5
	m	0,4	-0,1	0,6	-0,3
S^*	c	-1,098	0,221	-1,193	0,397
	m	-0,32	0,08	-0,48	0,24

Table 3: Comparative static in the Cournot competition case.

[5] See Basar T. and Olsder G. J (1995).

[6] See Dakhlaoui A. and M. Moreaux (2004).

From this table, we remark that any increase in water resource encourages the hydraulic operator to make an intra-annual transfer from peak season to off-peak season. We compare the Cournot solution with the first best solution; we conclude that the intra-annual transfer sense is maintained also with the structural change in the electricity industry. There is a change only in the volume of transfer. Indeed, the intra-annual transfer is appraised at $0{,}158\,\Delta S_t$ under the regulated industry and it is appraised only at $0{,}031\,\Delta S_t$ under Cournot competition. This show that, for a given period, the use of supplementary resource in the peak season under deregulated system is greater than those under regulated system. Thus, the storage effect ($SE_t = \Delta S_t - \Delta q_t^H$)[7] of competition solution is greater than the storage effect of the first best solution[8]. In addition, the intra-annual transfer affects the thermal operator strategy. Indeed, this latter cannot decrease his output only with a quantity equals to $0{,}075\,\Delta S_t$ units in the peak season and with a quantity equal to $0{,}134\,\Delta S_t$ units in the off-peak season. Consequently, the substitution of expensive technology output with costless. Technology output under competition is less than the substitution under regulated system[9]. Regarding the satisfaction of any increase in the peak demand, both the thermal operator and the hydraulic operator increase their output in different proportions. The hydraulic producer prefers to make an intra-annual transfer from off-peak season to peak season. Indeed, the release increase in the peak season and it is appraised at $0{,}118\,\Delta a_{ht}$ units, but the release decrease in the off-peak season is appraised at $0{,}349\,\Delta a_{ht}$ units. For the thermal operator, he increases the output in both seasons in order to let possible the storage of water resource in the off-peak season. We remark that the hydraulic operator does not use all the intra-annual transfer to satisfy the supplementary demand because the storage in the off-peak season is greater than the release in the peak season. Thus, there is storage of water resource appraised at $0{,}231\,\Delta a_{ht}$ units that will be used for the satisfaction of electricity demand for the following periods.

---

[7] Storage effect is defined as the interaction between the increase in inter annual transfers comes from preceding periods (ΔS_t) and the increase in the release from the stock at the end of period t .

[8] If the renewable effect is nil, then $SE^c = 0.221\Delta S_t$ and $SE^m = 0.198\Delta S_t$.

[9] The substitution of thermal output by hydraulic output is equals to $0{,}299\,\Delta S_t$ under competition and it is equal to $0{,}322\,\Delta S_t$ under regulated industry.

The numerical approximation lets possible both the quantification of water transfer and the substitution between technologies. We conclude the structural reform in the electricity industry not only changes solution greatness, but also it saves the least cost operating principle if the thermal operator does not hurt a binding capacity constraint problem.

Now, we assume that the thermal operator has a binding capacity constraint problem in the peak-season. Implicitly, we consider a large capacity constraint which is impossible to be bound in the off-peak season. In this case, the thermal operator cannot produce only $\bar{q}^{T}$. The hydraulic operator must satisfy the residual demand: $q_{ht} - \bar{q}^{T}$, where q_{ht} is the total demand in the peak-season of period t. The hydraulic operator can choose between Cournot equilibrium output or to exert a market power on the residual demand with an output $\widetilde{q}_{ht}^{T}$ less than q_{ht}^{Tc}. The strategic behavior of hydraulic operator leads to an increase in energy price in the peak-season. This behavior affects the strategy of two players in the off-peak season. Thus, the output decreasing is in favor of intra-annual transfer from peak season to off-peak season. Besides the transfer under constrained Cournot competition is greater than the transfer under unconstrained Cournot competition, it is not used to increase the off-peak season supply and consequently to substitute the thermal energy. The hydraulic operator tends to storage much water resource in order to push the thermal operator to produce more than the Cournot equilibrium output. This strategic behavior is, then, in favor of an important storage effect. This hydraulic storage is done in a strategic motivation in the peak-season of following periods. At the binding capacity constraint, the peak-season thermal plant operating is greater than those of hydraulic plant even with abundant water resource. Consequently, we can conclude that the binding capacity constraint is among the conditions under which the operating system rule is inverted in a deregulated industry.

STACKELBERG CLOSED-LOOP GAME

We analyze an other kind of quantity game in the wholesale electricity market such as the Stackelberg game between hydroelectric and thermal operator. We assume that the thermal operator constitute a dominant firm in the market. At each period the thermal operator first decides its production also in the peak and the off-peak season. Then, the hydroelectric operator decides its output with respect to the current stock of water, the stochastic futur storage in the dam and the output of the thermal operator.

Given the thermal strategy, the hydroelectric player decides the energy output that maximizes the expected sum of its annual profit with respect to the stochastic dynamic of the water stock stored in the dam.

The value function $J_t^s(S_t)$, evaluated at the period t must satisfy the Bellman equation:

$$J_t^s(S_t) = \underset{\{q_{ht}^H, q_{lt}^H\}_{t=0}^{+\infty}}{Max} \sum_{t=0}^{+\infty} P_h\left(q_{ht}^T + q_{ht}^H\right)q_{ht}^H + P_l\left(q_{lt}^T + q_{lt}^H\right)q_{lt}^H + E_t\left\{J_{t+1}^s\left(S_{t+1}\right) - \frac{1}{2}\left(S^* - S_{t+1}\right)^2\right\}$$

Where $E_t(\cdot)$ is the mean operator conditional to the information at period t which include S_t. At each season j of the period t, the hydroelectric operator decides the quantity produced given the thermal operator strategy and the current stock of water. The reaction functions of the hydroelectric operator are solutions of theses equations:

At the peak season of the period t:

$$\frac{\partial P_h}{\partial q_{ht}^H}\left(q_{ht}^T + q_{ht}^H\right)q_{ht}^H + P_h\left(q_{ht}^T + q_{ht}^H\right) - E_t\left\{S^* - S_{t+1}\right\} - E_t\left\{\frac{dJ_{t+1}^s}{dS_{t+1}}\left(S_{t+1}\right)\right\} = 0$$

At the off-peak season of the period t :

$$\frac{\partial P_l}{\partial q_{lt}^H}\left(q_{lt}^T + q_{lt}^H\right)q_{lt}^H + P_l\left(q_{lt}^T + q_{lt}^H\right) - E_t\left\{S^* - S_{t+1}\right\} - E_t\left\{\frac{dJ_{t+1}^s}{dS_{t+1}}\left(S_{t+1}\right)\right\} = 0$$

Theses equations give the peak season hydroelectric production as a function of the water release at the off-peak season, the thermal output at the peak season and the stock of water available at the period t. The same resolution at the of-peak season:

$$q_{ht}^H = Q_{ht}^H\left(q_{lt}^T, q_{ht}^T, S_t\right)$$

$$q_{lt}^H = Q_{lt}^H\left(q_{lt}^T, q_{ht}^T, S_t\right)$$

The thermal profit of season j at the period t is equals to the total receipt of energy purchase comes from thermal plant mines the total cost of production:

$$p_{jt}^T\left(q_{jt}^H, q_{jt}^T\right) = P_j\left(q_{jt}^H + q_{jt}^T\right)q_{jt}^T - C_j\left(q_{jt}^T\right)$$

Given the hydroelectric operator strategy at current period, $q_{ht}^H = Q_{ht}^H\left(q_{lt}^T, q_{ht}^T, S_t\right)$ and $q_{lt}^H = Q_{lt}^H\left(q_{lt}^T, q_{ht}^T, S_t\right)$, the thermal operator decides to supply energy that maximizes the profit with respect to the capacity constraint and the non-negativity output constraint. The reaction function of the thermal operator is a solution this problem:

$$\underset{q_{ht}^T, q_{lt}^T}{Max} p_t = P_h\left(q_{ht}^H + q_{ht}^T\right)q_{ht}^T - C_h\left(q_{ht}^T\right) + P_l\left(q_{lt}^H + q_{lt}^T\right)q_{lt}^T - C_l\left(q_{lt}^T\right)$$

With respect to the constraints:

$$q_{ht}^H = Q_{ht}^H\left(q_{lt}^T, q_{ht}^T, S_t\right)$$

$$q_{lt}^H = Q_{lt}^H\left(q_{lt}^T, q_{ht}^T, S_t\right)$$

The thermal operator problem is a static one. The Lagrange function is written as:

$$J^s = \underset{q_{ht}^T, q_{lt}^T}{Max} P_h\left(Q_{ht}^H\left(q_{lt}^T, q_{ht}^T, S_t\right) + q_{ht}^T\right)q_{ht}^T - C_h\left(q_{ht}^T\right) + q_{ht}^{Ts}q_{ht}^T + P_l\left(Q_{lt}^H\left(q_{lt}^T, q_{ht}^T, S_t\right) + q_{lt}^T\right)q_{lt}^T - C_l\left(q_{lt}^T\right) + q_{lt}^{Ts}q_{lt}^T$$

Where q_{jt}^{Ts} is the multiplicator of non negativity output constraint of the thermal plant in period j at the period t, for all $j \in \{h, l\}$. To simplify, we neglect the turbinate capacity of thermal plant.

The reaction function of the thermal operator is characterized by the Kuhn-Tucker conditions. For the season $j \in \{h, l\}$ at period t:

$$\frac{\partial P_j}{\partial q_{jt}^T} \cdot q_{jt}^T + \frac{\partial Q_{jt}^H}{\partial q_{jt}^T} \cdot \frac{\partial P_j}{\partial q_{jt}^H} + P_j - \frac{dC_h}{dq_{jt}^T} + q_{ht}^T = 0, q_{jt}^T q_{jt}^T = 0,\ q_{jt}^T \geq 0 \text{ and } q_{jt}^T \geq 0.$$

If the multiplicator of non-negativity output constraint is nil, the thermal operator strategy is a solution of following system: $\frac{dC_h}{dq_{jt}^T} = \frac{\partial P_j}{\partial q_{jt}^T} \cdot q_{jt}^T + \frac{\partial Q_{jt}^H}{\partial q_{jt}^T} \cdot \frac{\partial P_j}{\partial q_{jt}^H} + P_j$

The resolution of the equations system gives the thermal operator strategy at the two seasons as a function of the current water stock of the thermal operator: $q_{jt}^{Ts} = Q_{jt}^T\left(S_t\right)$, for all $j \in \{h, l\}$.

We replace the thermal operator strategy at the first order conditions of the hydroelectric operator. In the season j at the period t:

$$\frac{\partial P_j}{\partial q_{jt}^H}\left(Q_{jt}^T\left(S_t\right) + q_{jt}^H\right)q_{jt}^H + P_j\left(Q_{jt}^T\left(S_t\right) + q_{jt}^H\right) = E_t\left\{S^* - S_{t+1}\right\} + E_t\left\{\frac{dJ_{t+1}^s}{dS_{t+1}}\left(S_{t+1}\right)\right\}$$

The resolution of theses equations system gives the hydroelectric operator strategy at the equilibrium as a function of the current water stock: $q_{jt}^{Hs} = Q_{jt}^H\left(S_t\right)$.

We replace this strategy in the Bellman equation; we can find the value function associated to the hydroelectric operator in the Stackelberg competition with the thermal operator:

$$J^s(S_t) = P_h\left[q_{ht}^T(S_t) + q_{ht}^H(S_t)\right]q_{ht}^H(S_t) + P_l\left(q_{lt}^T(S_t) + q_{lt}^H(S_t)\right)q_{lt}^H(S_t)$$
$$+E_t\left\{J^s\left(S_t - q_{lt}^H(S_t) - q_{lt}^H(S_t) + f_{t+1}\right) - \frac{1}{2}\left(S^* + q_{lt}^H(S_t) + q_{lt}^H(S_t) - f_{t+1}\right)^2\right\}$$

NUMERICAL APPROXIMATION OF EQUILIBRIUM: LINEAR-QUADRATIC CASE

We consider a linear inverse demand function at the electricity market: $P_j\left(q_{jt}^H + q_{jt}^T\right) = a_j - q_{jt}^H - q_{jt}^T$. Let's $J^s(S_t)$ the value function associated to the hydroelectric operator problem under Stackelberg competition between the two producers. We assume that the value function is quadratic: $J^s(S_t) = b^s S_t - \frac{1}{2}B^s S_t^2$, where b^s and B^s are unknown coefficients.

The hydroelectric operator strategy at the two seasons verifies the Bellman equation:

$$J_t^s(S_t) = \left(a_h - q_{ht}^H - q_{ht}^T\right)q_{ht}^H + \left(a_l - q_{lt}^H - q_{lt}^T\right)q_{lt}^H - \frac{1}{2}E_t\left\{\left(S^* - S_t + q_{ht}^H + q_{lt}^H - f_{t+1}\right)^2\right\}$$
$$+E_t\left\{b^s\left(S_t - q_{ht}^H - q_{lt}^H + f_{t+1}\right)\right\} - \frac{1}{2}B^s E_t\left\{\left(S_t - q_{ht}^H - q_{lt}^H + f_{t+1}\right)\right\}$$
$$-\frac{1}{2}B^s E_t\left\{\left(S_t - q_{ht}^H - q_{lt}^H + f_{t+1}\right)^2\right\}$$

Given $J_t^s(S_t)$, the first order conditions are given by the following equations:

$$J_t^s(S_t) = \left(a_h - q_{ht}^H - q_{ht}^T\right)q_{ht}^H + \left(a_l - q_{lt}^H - q_{lt}^T\right)q_{lt}^H - \frac{1}{2}E_t\left\{\left(S^* - S_t + q_{ht}^H + q_{lt}^H - f_{t+1}\right)^2\right\}$$
$$+E_t\left\{b^s\left(S_t - q_{ht}^H - q_{lt}^H + f_{t+1}\right)\right\} - \frac{1}{2}B^s E_t\left\{\left(S_t - q_{ht}^H - q_{lt}^H + f_{t+1}\right)^2\right\}$$

The combination of the two latter equations gives the hydroelectric operator output at the peak season as a function of the thermal operator output at the peak and the off-peak season:

$$q_{ht}^H = L_1^H q_{ht}^T + L_2^H S_t + L_3^H q_{lt}^T + L_0^{hH}$$

Where

$$L_1^h = -\frac{3 + B^s}{4\left(2 + B^s\right)}, L_2^H = \frac{1 + B^s}{2\left(2 + B^s\right)}, L_3^H = -\frac{1 + B^s}{4\left(2 + B^s\right)} \text{ and}$$

$$L_0^{hH} = \frac{\left(3 + B^s\right)a_{ht} - \left(1 + B^s\right)a_{lt} + 2\left[\left(1 + B^s\right)\overline{f} - S^* - b^s\right]}{4\left(2 + B^s\right)}.$$

For the hydroelectric output at the off-peak season:

$$q_{lt}^{H} = L_1^H q_{lt}^T + L_2^H S_t + L_3^H q_{ht}^T + L_0^{lH}$$

Where $L_0^{lH} = \dfrac{\left(3+B^s\right)a_{lt} - \left(1+B^s\right)a_{ht} + 2\left[\left(1+B^s\right)\overline{f} - S^* - b^s\right]}{4\left(2+B^s\right)}$.

The thermal operator decides the output given the hydroelectric operator strategy. The thermal operator problem is written as:

$$\underset{q_{lt}^T, q_{ht}^T}{Max} p_t = \left(a_{ht} - L_1^H q_{ht}^T - L_2^H S_t - L_3^H q_{lt}^T - L_0^{hH} - q_{ht}^T\right) q_{ht}^T - \frac{1}{2} c_h \left(q_{ht}^T\right)^2$$
$$+ \left(a_{lt} - L_1^H q_{lt}^T - L_2^H S_t - L_3^H q_{ht}^T - L_0^{lH} - q_{lt}^T\right) q_{lt}^T - \frac{1}{2} c_l \left(q_{lt}^T\right)^2$$

The first order conditions give the thermal operator output as a function also of his production in the other season and the current water stock: $q_{ht}^T = q_1^T q_{lt}^T + q_2^T S_t + q_0^{hT}$ and $q_{lt}^T = q_1^T q_{lt}^T + q_2^T S_t + q_0^{lT}$.

Where $q_1^T = \dfrac{1+B^s}{5+3B^s+2\left(2+B^s\right)c_h}, q_2^T = -\dfrac{1+B^s}{5+3B^s+2\left(2+B^s\right)c_h}$

$q_0^{hT} = \dfrac{\left(5+3B^s\right)a_{ht} + \left(1+B^s\right)a_{lt} - 2\left[\left(1+B^s\right)\overline{f} - S^* - b^s\right]}{2\left[5+3B^s+2\left(2+B^s\right)c_h\right]}$ and

$q_0^{hT} = \dfrac{\left(5+3B^s\right)a_{lt} + \left(1+B^s\right)a_{ht} - 2\left[\left(1+B^s\right)\overline{f} - S^* - b^s\right]}{2\left[5+3B^s+2\left(2+B^s\right)c_l\right]}$.

The combination of the two equations implies the thermal operator output strategy as a function of the current water stock of the hydroelectric operator. The same thing for the hydroelectric operator.

Proposition 2

The Stackelberg closed-loop game equilibrium is given by the following equations:

$$q_{ht}^T = -0.087 S_t + 0.1708 a_h + 0.077 a_l + 0.057 S^* - 0.1078 \overline{f}$$

$$q_{ht}^H = 0.3313 S_t - 0.422 a_h - 0.206 a_l - 0.2169 S^* + 0.41 \overline{f}$$

$$q_{lt}^T = -0.087 S_t + 0.1104 a_h + 0.325 a_l + 0.1 S^* - 0.1164 \overline{f}$$

$$q_{lt}^H = 0.364 S_t - 0.1246 a_h + 0.1694 a_l - 0.2383 S^* + 0.4504 \overline{f}$$

The value function associated to the hydro electrical operator is:

$$J_t^s = -0.435 S_t^2 + \left(0.2279 a_h + 0.4053 a_l + 0.2377 S^* - 0.4491 \overline{f}\right) S_t \blacksquare$$

The comparative static in the regulated and unregulated industry is given by the following table

t	*Systems*	q_{ht}^{H*}	q_{ht}^{T*}	q_{lt}^{H*}	q_{lt}^{T*}
S_t	*s*	0,331	-0,087	0,364	-0,087
	m	0,322	-0,08	0,48	-0,24
a_{ht}	*s*	-0,422	0,170	-0,124	0,110
	m	0,6	0,1	-0,165	0,3
a_{lt}	*s*	-0,206	0,077	0,169	0,325
	m	-0,11	0,22	0,4	0,3
$\overline{f}$	*s*	0,410	-0,107	0,450	-0,116
	m	0,4	-0,1	0,6	-0,3
S^*	*s*	-0,216	0,057	-0,238	0,1
	m	-0,32	0,08	-0,48	0,24

Table 2: The static comparative in a Stackelberg closed-loop game

FORCED STORAGE

We assume that the two producers observe a demand increase at the peak season equals to Δa_h. At the sequential asymmetric game, the supplementary demand energy is satisfied only by the T plant. Indeed, the H operator output at the peak season decreases by the quantity $-0,4224\Delta a_h$. The conservation of the water resources is not entirely compensated by an increase in the production of the power station. The T operator increase production by a quantity equal to $0,1708\Delta a_h$ units of energy. The position of the T operator on the market reduces the production of the H operator especially at the season h without compensation. This little increase in the production allows to the T operator profit coming from the high price of energy at the season h. The T operator adopts the same strategy at the season l but in the different sizes. Thus, the increase in the T production is evaluated at $0,1104\Delta a_h$ whereas the increase in the H production achieve only $-0,1246\Delta a_h$. This kind of game does not make it possible to the H operator to intervene by his free technology to satisfy the increase in demand at season h. On the other hand, the T operator obliges the H operator to preserve his resources in order to exert a power of market on this additional demand. In this case, the inter annual transfer, the intra-annual transfer and the technological

complementarities do not make it possible to mitigate the perverse effects of the position of the T operator.

FORCED CYCLIC HYDROELECTRIC SEEDLING

With the difference of the preceding section, an increase in the demand at the season l of a quantity equal to Δa_l requires the recourse, in different proportions, at the same time with the T station and the H power station. The production of the T operator increases by a quantity equals to $0{,}3257\Delta a_l$ whereas that of the H operator increases only by $0{,}1694\Delta a_l$ units. To satisfy this additional request, one calls more upon the T station that with the H station. The increase in H operator output comes from the intra-annual transfer from the h season to the l season. The H operator preserves $0{,}206\Delta a_l$ units of water at the season h and it uses only a quantity equal to $0{,}1694\Delta a_l$ to satisfy the additional demand of the season l. The position of the T operator on the market prevents the H operator from using all or a higher part of its transfer to the season l. It obliges the H operator to preserve $0{,}0366\Delta a_l$ units of water to satisfy any increase in demand at the season l of the following periods and not to satisfy the increase in demand in the season h. The T operator increases his production at the season h in such way that it does not compensate for the fall of production of the H operator.

CONCLUSION

In this paper, we have analyzed the effect of imperfect competition one the optimal order in operating the mixed hydrothermal system. This operating is based one the substitution of costly technology output by the costless technology output. This substitution is only possible with the potential energy transfers. In contrast, the Cournot competition in power generation favored the strategic storage of water resource. Consequently, this behavior of hydraulic operator lets the additional uses of thermal technology greater then that under regulated industry. We show that the least cost operating rule is preserved also with this hydraulic operator in the non-constraint model; but it is inverted at the binding capacity constraint of the thermal operator. Under Stackelberg competition, we have shown that the thermal operator position on the market can increase the conservation of hydraulic resources in the two seasons. In this type of competition, the inter annual transfers and the intra annual transfers do not have any role. One solution of this problem is to set, by a contract, the minimal peak

season output supplied by both operators on the market. To extend this paper, we proposes to analyze the effect of technical complementation and congestion of transmission network one the behavior of operators in the market. Future For research we proposes to study the effect of imperfect competition of each operator endowed by two heterogeneous technologies.

REFERENCES

* **Alberto T. G. J. , Johnson Sharon.A and Stedinger Jery R (1993)** : *"Comparison of two approaches for implementing multireservoir operating policies derived using stochastic dynamic programming"*, Water Resources Research, Vol 29, n° 12, December, pp3969-3980.

* **Borenstein S. and Bushnell J. (2000)**: *"Electricity restructuring: Deregulation or regulation?"*, POWER working paper PWP-074. University of California energy institute, February. www.ucei.berkeley.edu/ucei.

*** Bushnell. J. (2000)**: *"Water and power: Hydroelectric resources in the Era of competition in the western U.S"*, POWER working paper PWP-056.University of California energy institute, July. www.ucei.berkeley.edu/ucei.

* **Basar T. and Olsder G. J. (1995):** *"Dynamic Noncooperative Game Theory"*, Academic Press, Great Britain.

* **Bernard J. T. and Chatel J. (1984):** *"The role of energy limited hydro equipment in an optimal plant mix"*, Energy Economics, Vol 6, pp 139-144.

* **Blume L., Easley D. and Maureen O'Harra (1982)**: *"Characterization of optimal plans for stochastic dynamic programs*, Journal of economic theory 28, pp 221-234.

Bellman R. (1957): Dynamic programming, Princeton: Princeton University Press.

*** Crampes C. and Moreaux M. (2001)**: *"Water resources and power generation"*, International Journal of Industrial Organization, 19, pp975-997.

*** Chow G. C. (1997)**:*" Dynamic economic optimization by the Lagrange method"*, Oxford university press, New York.

*** Green R. and Newbery D. (1992)**: *"Competition in the British Electricity spot market*, Journal of Political Economy, 100 (5), pp 929

*** Kobila T. O. (1990)** : *"The choice between hydro and thermal power generation under uncertainty"*, Recent modeling approaches in applied energy, St Edmunds bury Press, Bury St. Edmunds, Suffolk, Great Britain.

*** Little J. D.C (1955)**: *"The use of storage water in hydroelectric system"*, Journal of operations research society of America, Mai, pp 187-97.

*** Mitchell D. W (2000)**: *"An analytic Riccati equation solution for tow-target discrete-time control"*, Journal of Economic Dynamics and control, 24, pp 615-622.

*** Scott T. J. and Read E. G. (1996)**:*" Modeling hydro reservoir operation in deregulated electricity market"*, International Transactions in Operational Research, vol 3 n° 3, pp 243-254.

*** Tsur Y. and Graham-Tomasi T. (1991)**: *"The buffer value of ground water with stochastic surface water supplies"*, Journal of environmental economics and management 21, pp 201-224.

*** Von der Fehr N. H. M and Sandsbraten L. (1997)**: *"Water on fire: Gains from electricity trade"*, Scand. Journal of Economics 99(2), pp281-297.

*** Wood A. J. and Wollenberg B.F. (1996)**: *"power generation, operation, and control»*. New York: John Wiley &Sons, Inc.

ESTIMACIÓN DE LOS COSTES DEL TRANSPORTE DE MERCANCÍAS EN CARGA FRACCIONADA: EL CASO DE LOS PRODUCTOS PERECEDEROS

SILVIA ANDRÉS GONZÁLEZ-MORALEJO
e-mail: silangle@upvnet.upv.es
Departamento de Economía y Ciencias Sociales
Grupo de Economía Internacional. Universidad Politécnica de Valencia

Resumen

En este trabajo se analiza el sistema de tarificación vigente en el transporte de mercancías perecederas en carga fraccionada. Utilizando datos para el año 2008, obtenidos directamente de una muestra de empresas porteadoras, se formula la estructura de precios, diferenciando entre tarifas ordinarias y precios finales y las variables que los integran, y se compara con los costes medios de producción derivados de la actividad.

Palabras Clave: Transporte por carretera, precios, bonificación, coste medio.
Área Temática: Economía de la Energía, del Ambiente, de los Transportes y de los Recursos Naturales.

Abstract

In this paper we analyze the current tariff system on transport in part load of perishable freight. Using the data of 2008, which have been directly obtained from a sample of transport companies, the prices structure is formulated by making a distinction between ordinary and final prices and also their respective variables. Later on, it is compared to the production mean costs.

Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

1. INTRODUCCIÓN

El transporte de mercancías por carretera, actividad estratégica para la industria y el comercio, ha sido desde siempre un factor de integración social, cultural y económica, y mucho más en el entorno actual de globalización. Se trata de un sector tremendamente complejo, por dos razones fundamentales: su acusado grado de atomización y los altos niveles de competencia en que se basa el desarrollo de su actividad, lo que, con el tiempo, ha configurado un mercado de gran opacidad y con prácticas de competencia en ciertos casos desleal (caracterizadas por precios inferiores a los costes marginales de los desplazamientos). La necesaria rapidez y cuidado que requieren los productos perecederos complica más, si cabe, la realidad de un sector que se ha convertido en un gran desconocido para la sociedad a la que sirve, que lo considera como un elemento agresivo y perturbador de sus movimientos.

La incertidumbre en el comportamiento de los costes energéticos, de enorme impacto para el sector, y la desaparición total de barreras para el acceso al mercado en España y el resto de países de la UE, que ha posibilitado la entrada de nuevas empresas con costes significativamente inferiores, y todo ello referido a una actividad de valor añadido limitado, han provocado que el precio final del servicio se convierta en el factor clave de elección para los operadores logísticos de frío (OLF), que son los profesionales especializados en la provisión de transporte frigorífico en carga fraccionada. En este contexto, el presente trabajo se centra en el análisis del sistema de tarificación vigente en el transporte en grupaje (carga fraccionada) de mercancías perecederas por carretera y propone la formulación de un modelo para la identificación y estimación cuantitativa de los principales componentes que integran el precio final en dicho tipo de transporte. Asimismo, y dependiendo de la disponibilidad de información, se detallará si los precios de dicho transporte reflejan exactamente el coste[1] de producción derivado de la actividad.

La preocupación por esclarecer cómo se tarifican los servicios de transporte ha suscitado, a nivel académico, el desarrollo de líneas distintas de investigación[2]. Originados en los aportes de Newbery, son numerosos los trabajos que analizan la tarificación, para cualquier tipo de usuario, por el uso de las infraestructuras. Los primeros estudios de Newbery (1988, 1989, 1991) desarrollaron el análisis de la tarificación específica de las carreteras y, posteriormente, se han publicado trabajos que determinan las ventajas e inconvenientes de cualquier cambio tarifario en el sistema de transporte por carretera (Greene *et al*., 1997; Ahlstrand, 2001). Si bien es cierto que la mayoría de estos trabajos se refieren al análisis del transporte urbano, también los hay que han analizado esta cuestión a nivel interurbano (Sansom *et al*., 2001; Vermeulen *et al*., 2004). En el caso español es destacable la reciente aportación de Álvarez *et al*. (2007), que estima la estructura de precios óptimos para el transporte interurbano por carretera, diferenciando por tipo de vehículo, vía y momento del tiempo.

Especialmente extensa es la literatura académica que estudia el transporte urbano de pasajeros, donde un organismo público controla la provisión de los servicios de transporte colectivo en ciudades y áreas metropolitanas. Esta línea de investigación ha generado resultados de notable interés económico, entre los que destacan los esfuerzos de estimación de funciones de costes, que ofrecen información valiosa sobre el tipo de economías de escala que caracterizan a la industria (Williams y Hall, 1981; Williams y Dalal, 1981; Berechman y Guliano, 1984, Berechman, 1987), la existencia de economías de alcance (Talley y Anderson, 1986) y de economías de densidad (Caves y Christensen, 1988); así como también los trabajos para la determinación de las reducciones de costes que se lograrían mediante la introducción de competencia en la industria, con aplicación empírica en países como EEUU, Reino Unido, Suecia, Australia y España (por ejemplo, véase Hensher, 1987). En España, el máximo exponente en el análisis de esta industria se encuentra en De Rus (1987, 1989, 1991), que ha abordado entre otras cosas el comportamiento de sus costes, la discriminación de precios, el sistema de subvención cruzada y la determinación de valores óptimos para las variables más significativas (precios, frecuencia, tamaño del vehículo y resultado financiero).

Por lo que conocemos es algo más escasa, sobre todo en España, la literatura económica que con carácter académico aborda los principios de la tarificación en el tráfico de mercancías[3], pues su análisis empírico conlleva graves dificultades prácticas. Los trabajos internacionales relevantes se han centrado básicamente en aspectos relacionados con la regulación del sector y la estimación de funciones de costes. Más concretamente son de destacar las aportaciones de Bayliss (1998), que ha analizado para el conjunto de la UE las causas y los efectos de la desregulación que se inició en la década de los 80, así como también la evolución que las

[1] Nos referimos a costes privados, que son aquellos soportados por la empresa porteadora que realiza la actividad de transporte.

[2] En la UE se ha asistido a un intenso debate sobre los principios de tarificación en el transporte, y muy especialmente en el de la carretera. En este sentido, los esfuerzos de la Comisión se han constatado en la publicación de importantes documentos, entre los que cabe destacar el Libro Verde *Towards fair and efficient pricing in transport* de 1995 y el Libro Blanco *European Transport Policy for 2010* de 2001.

[3] Dentro del sector de la alimentación perecedera destaca la aportación empírica de Rebollo *et al.* (2006), que estudia la evolución de los márgenes comerciales en España para los productos de alimentación en fresco, y la de Mir y Borrás (2008), que realiza el cálculo de los costes de la cadena de distribución de productos hortofrutícolas en fresco. En ambos casos se considera el precio del transporte, pero no se especifica su composición ni las variables que influyen en él, como sí hace el presente trabajo.

políticas económicas (estructura del mercado, externalidades, transparencia) y públicas (empleo, estructura industrial, diversidad regional) aplicadas en el sector, y la normativa que las desarrolla, han sufrido desde entonces. Asimismo, no podemos dejar de citar la existencia de algunos estudios empíricos que se han planteado la necesidad de regular precios (Hurley, 1995) o las implicaciones económicas que la introducción de un impuesto especial (tonelada-kilómetro) al transporte de mercancías por carretera supondría para la economía europea (Barker y Köhler, 2000). Los trabajos de Harmatuck (1991) han sentado las bases para estimar funciones de costes (basándose en un modelo de costes traslog) para operadores logísticos, a partir de las cuales se examina la naturaleza de las economías de escala y de alcance en la provisión de servicios de transporte de mercancías en EEUU. Autores como Daughety *et al*. (1985) ya obtuvieron desarrollos similares para el caso europeo.

Las limitaciones y problemas inherentes a la medición de los costes de transporte y su tarificación se han tratado en la literatura dedicada al análisis de los flujos de comercio; entre las referencias más recientes, Anderson y Van Wincoop (2004) revisan los diferentes elementos que configuran los costes del comercio, entre ellos los vinculados al transporte de mercancías. Fundamentalmente, son dos los obstáculos que constatan. En primer lugar, la evidencia empírica sobre el transporte de mercancías se ve limitada por la inexistencia de bases de datos que posibiliten su investigación, exceptuando contadas excepciones, entre las que se encuentran la *US Waterborne Database* de la Administración Marítima de EEUU, las bases de PIERS *Global Trade Intelligence* y la *Base de Datos de Transporte Internacional BTI* de la CEPAL (Naciones Unidas). Con el objetivo de paliar esta laguna, en España se está desarrollando la base de datos TRADE TRANS[4]. En el trabajo de García y Pérez (2007) se realiza una revisión de la metodología seguida para la construcción de TRADE TRANS, así como de los principales sesgos y problemas inherentes a su utilización. En segundo lugar, buena parte de la información requerida para el análisis de un sistema de tarificación, especialmente la de carácter cuantitativo (precios, costes directos, costes indirectos), pertenece al ámbito confidencial de la empresa. Se trata de información privada y protegida que las empresas se muestran muy reacias a revelar.

En este contexto, el presente trabajo propone un diseño de investigación que integra los enfoques cualitativo y cuantitativo, mediante una triangulación de técnicas, de forma sistemática y creativa, posibilitando a través de las entrevistas en profundidad el acceso a información sólida y veraz que, combinada posteriormente con la observación y el análisis estadístico descriptivo, enriquezca y haga posible la comprensión del fenómeno estudiado en su dimensión objetiva: ¿cómo se forman los precios en el transporte de perecederos a carga fraccionada?, ¿qué factores influyen en dicho proceso de tarificación y en qué medida? Después de esta introducción, la sección 2 describe los rasgos básicos de la recogida de información. En la sección 3 se formula el modelo de tarificación en la provisión de transporte frigorífico a carga fraccionada, a partir del cual se aplica un análisis estadístico descriptivo que se desarrolla en la sección 4. La sección 5 integra todos los resultados. Finalmente, en la sección 6 se discuten las principales conclusiones del trabajo.

2. GENERACIÓN DE LA INFORMACIÓN PRIMARIA

Como la información requerida para alcanzar el objetivo planteado en este trabajo no se encuentra recopilada en fuentes estadísticas ni bases de datos, se debe acudir a recogerla directamente de las empresas, que es donde se genera. Con ello, además de dotar a la investigación de una base empírica, se consigue disponer de unos datos fiables y actualizados que no podrían ser hallados de otro modo. En consecuencia, la investigación realizada se basa en un *estudio de casos*[5]*:* se ha seleccionado una muestra de 8 OLF que trabajan en carga fraccionada o grupaje y se han analizado en profundidad sus sistemas tarifarios. La información

[4] TRADE TRANS representa una novedad dada la inexistencia de este tipo de estadísticas en Europa, tanto en el ámbito académico como en la consultoría especializada.
[5] Véase en Bonache (1999) un análisis de la naturaleza de los estudios de casos y su ámbito de aplicación.

correspondiente a cada empresa[6] se ha obtenido mediante la técnica de la entrevista en profundidad mediante preguntas abiertas, dado que son mejores para que el OLF revele su auténtica experiencia.

En la elección de los OLF que conforman la muestra se ha seguido un procedimiento compuesto por varias etapas: recogida de información sobre los OLF existentes, depuración de la información y selección de las empresas a entrevistar. Se ha considerado que el estudio de casos sería ilustrativo en función de los objetivos de la investigación siempre que la elección se realice adecuadamente. Por tanto, en la designación de los OLF a analizar se ha privilegiado que todos ellos estén incluidos en el ranking de las 30 primeras empresas del sector por volumen de ventas en servicios frigoríficos[7], que se muestra en el Cuadro 1. Un segundo criterio de selección ha sido que al menos cinco de los OLF entrevistados estén dentro del top 10 en dicho ranking. De esta manera se ha conseguido que los OLF considerados en este trabajo representen el 40 por ciento de la facturación total del sector. Otros criterios de selección que se consideraron fueron la ubicación de delegaciones del OLF y el tipo de mercados que abastece. Todos los OLF entrevistados disponen de delegaciones en la Comunidad Valenciana, que es el ámbito geográfico al que queda restringido el trabajo de campo. Dicha restricción geográfica es imprescindible, pues habría dificultado enormemente la recopilación de los datos. Respecto al tipo de mercados se ha comprobado que los OLF entrevistados abastecen, como mínimo, a los mercados local y nacional.

Se ha seguido un único procedimiento para la recogida de la información. Tras hablar por teléfono con los directivos de las empresas a entrevistar, para explicarles la finalidad de la investigación y solicitar su colaboración, se les ha pedido una cita a todos los que se mostraron dispuestos a colaborar. Toda la información se ha obtenido por la autora del trabajo a partir de entrevistas personales con los directivos de las 8 empresas[8] seleccionadas.

Por lo que respecta a las variables cuantitativas, se han registrado observaciones a dos niveles distintos. El primer nivel se corresponde con la *tarifa general ordinaria por palé (TGOpalé)* de cada OLF, desglosada en lo que se denomina en el sector *tarifa interna* (desglosada a su vez tal y como se explica en el epígrafe siguiente) más un *margen bruto comercial.* Cada observación de primer nivel incluye un valor por palé en función del número de palés y del trayecto; de esta manera, se dispone de una *TGOpalé* para envíos de 1 palé, 2/3 palés, 4/6 palés y 7/10 palés[9], con origen en los principales puntos de expedición de mercancía (Madrid, Barcelona, Valencia, Sevilla y Bilbao) y destino en cualquier capital española. Se han registrado 8 observaciones de primer nivel. El segundo nivel se corresponde con el *precio final por palé (PFpalé)* ofrecido a cada cliente. Cada observación de segundo nivel incluye un valor por palé en función del número de palés y del destino; de esta manera, se dispone para cada cliente de un *PFpalé* para envíos de 1 palé, 2/3 palés, 4/6 palés y 7/10 palés, con origen del trayecto en Valencia y destino en cualquier capital española[10]. Se han registrado 300 observaciones de segundo nivel. Todas las observaciones recogidas se refieren al año 2008, pues es el año en el que se dispone de un mayor volumen de información, para servicios de transporte a temperatura controlada con mercancía paletizada sobre Palé EUR (0,80 x 1,20 m.), con una altura máxima del palé de 1,80 m. y peso máximo 800 Kg. En el caso de que las condiciones anteriores no fueran respetadas, la tarifa correspondiente sufre un incremento del 25 por ciento. Otra posible modificación tarifaria procede de la aplicación trimestral, por parte de los OLF, de una cláusula automática de revisión del precio del transporte en función de la evolución del precio del gasoil. Todas las tarifas y precios son independientes del tipo de mercancía, refrigerada o congelada.

[6] Las entrevistas se realizaron entre los meses de enero y abril de 2008.
[7] La facturación por servicios frigoríficos de las 30 empresas del ranking se distribuye entre 5 y 120 millones de euros anuales.
[8] No se citan en este trabajo por sugerencia de algunas de ellas, que prefieren permanecer en el anonimato.
[9] A partir de 10 palés resulta más rentable la contratación de un camión completo (carga completa), aunque quede espacio en la caja sin aprovechar.
[10] Hay que notar que algunas observaciones están incompletas, dependiendo de las especificaciones propias del cliente, puesto que les puede faltar el correspondiente valor por destino o por número de palés.

Cuadro 1: EMPRESAS DE GRUPAJE Y LOGÍSTICA DEL FRÍO. RANKING POR VOLUMEN DE FACTURACIÓN

Nº	EMPRESA
1	SDF IBÉRICA, S.A.
2	SALVESEN LOGISTICA, S.A.
3	GRUPO INTEGRA2 (LOGISTA)
4	EXEL IBERIA GRUPO, S.L.*
5	DHL IBERIA
6	LOGIFRIO GESTION FRIGORIFICA, S.L.
7	GRUPO LOGISTICO SANTOS, S.A.
8	CONWAY THE CONVENIENCE COMP. ESPAÑA, S.L.
9	HERMES LOGISTICA, S.A.
10	DISFRIMUR, S.L. (GRUPO)
11	GRUPO LINSER LOGISTIC & TRADE, S.L.
12	CRONOFRIO-NORESTRANS
13	ACR LOGISTICS IBERIA, S.L.
14	LOGISTICA REFRIGERADA, S.A.
15	TRANSPORTES BADOSA, S.A.
16	TRANSPORTES AGUSTIN FUENTES E HIJOS, S.L.
17	TRANSPORTES TRESSERRAS, S.A.
18	INTERLOGISTICA DEL FRIO, S.A. (INTERFRISA)
19	TRANSPORTES J. CARBO, S.A.
20	OLANO Y MUÑOZ, S.A. (TOMSA)
21	TUDEFRIGO, S.L. (GRUPO)
22	MONTFRISA, S.A.
23	GRUP-OIL, S.A.
24	LOGISTICA ALICANTINA DEL FRIO, S.L.
25	SEUR, S.A. (SEUR FRIO)
26	LOGISTICA SANCHEZ BREA, S.L.
27	TRANSFRIO, S.C.L.
28	CEFRUSA SERVICIOS FRIGORIFICOS, S.A.
29	TRANSPORTES FRIGORIFICOS BUIL, S.A.
30	TRAIGLEFER, S.L.

Fuente: Vivó, 2007.

3. FORMULACIÓN DEL MODELO DE TARIFICACIÓN

El sistema de facturación más extendido en la provisión de transporte frigorífico está compuesto por tres elementos fundamentales:

- *El coste del transporte*
- *Los gastos administrativos*
- *El seguro*

Cada uno de estos elementos, que se describen a continuación, toma un valor monetario en función de un conjunto de variables. La suma de la cuantía de las tres partidas da lugar al importe de la factura[11] que el OLF cobra a su cliente.

El *coste del transporte* deriva del servicio principal prestado por el OLF, por lo tanto es la partida más importante de la factura y la que justifica su existencia. El resto de componentes son costes auxiliares, dependen del *coste del transporte* y carecen de sentido si no se da éste último. De los tres elementos, el más complejo es el *coste del transporte*, que se construye a partir de

[11] Hemos desestimado del importe de la factura la parte correspondiente al IVA. De acuerdo con la Ley 37/1992 del Impuesto sobre el Valor Añadido, es un impuesto indirecto que recae sobre el consumo, por lo que debe recaer principalmente sobre el consumidor final. La empresa, en la mayoría de los casos, sólo hará de intermediario entre Hacienda Pública y el consumidor final, repercutiendo y soportando IVA a lo largo del desarrollo de su actividad empresarial. Sólo en el caso de que el IVA soportado por la empresa sea no deducible, pues no todas las operaciones realizadas por ésta tienen derecho a deducción, este impuesto supondrá un mayor coste para ésta, pasando a engrosar el importe del precio de adquisición o coste de producción. En nuestro caso, el IVA repercutido es el IVA soportado por la industria fabricante, pero es a su vez deducible, por lo que no supone un coste adicional para el operador o transportista.

una *TGOpalé* más una *bonificación* por cliente[12]. A su vez, la *TGOpalé* se compone a partir de una *tarifa interna* que recoge la estructura de costes directos e indirectos del OLF más un *margen bruto comercial*. La *bonificación*, que como máximo puede suponer la totalidad del *margen bruto comercial*, se aplica en función de las siguientes variables:

- Sinergias con el punto de destino (probabilidad de que el vehículo, en su viaje de regreso, lo haga cargado con nueva mercancía).
- Grado de ocupación del vehículo (alto o bajo).
- Volumen de negocio que supone el servicio contratado y frecuencia del mismo (envío único o periódico).
- Importancia del cliente que solicita el servicio (sobre los ingresos totales anuales del OLF).
- Volumen de negocio actual del OLF (necesidad de vender).

En consecuencia, el *coste del transporte* es, en la mayoría de los casos, personalizado, pues se calcula *ad hoc* para cada cliente como resultado de una negociación en la que confluyen factores tan diversos como el destino de la mercancía, el momento del año en el que se solicita el servicio[13] o la situación macroeconómica. Por tanto, para cada cliente y envío de mercancía, tomará un valor distinto. Esto da una idea de las dificultades que entraña el conocimiento profundo del *coste del transporte*, no sólo por las numerosas variables que hay que considerar en su cálculo, cuya influencia unitaria no siempre es posible de constatar (incluso para los propios OLF), sino también porque depende de información privilegiada de la empresa a la que no siempre se tendrá acceso. Un *coste del transporte* sometido a un alto grado de variabilidad es un resultado de nuestra aplicación empírica que demuestra que la política tarifaria de los OLF se basa en la diferenciación de precios, en consonancia con los costes de producción, lo que se contrapone claramente a la política tarifaria de precio único característica en el transporte de viajeros, que puede ser calificada de discriminatoria (De Rus, 1987).

Los *gastos administrativos* son los gastos asociados a la emisión y gestión de los documentos de transporte necesarios para llevar a cabo el traslado de la mercancía desde el punto de origen hasta el punto de destino. Estos gastos dependerán del ámbito en que se realice el servicio de transporte (nacional o internacional) y son gastos necesarios, pues en estos documentos queda reflejada la mercancía a transportar y sin ellos no se puede realizar el servicio. En el caso de que el transporte de la mercancía se lleve a cabo dentro del territorio nacional, el documento a tratar será el **albarán** generado por el cliente. El OLF se encarga de escanear este albarán y hacerlo llegar al destinatario a través del conductor. Una vez firmado, vuelve al OLF que se lo remite al cliente. Por esta gestión, el OLF cobra al cliente un importe fijo medio de 3 euros, en concepto de dichos *gastos administrativos*. A efectos de este trabajo, se denomina *precio final por palé* para cada servicio de transporte a la suma del *coste del transporte* más los *gastos administrativos*[14].

El *seguro* es la última rúbrica que aparece en la factura, y está constituida por el importe que paga el cliente para que el OLF asegure su mercancía durante el trayecto de transporte. De este modo, si durante el servicio prestado la mercancía sufre algún deterioro, se pierde o se demora, el OLF entregará a su cliente una indemnización previamente establecida, a nivel nacional, de acuerdo con el artículo 23 de la Ley de Ordenación del Transporte Terrestre (LOTT). Evidentemente, dicha indemnización se puede variar si existe acuerdo entre las partes y así se hace constar en el contrato.

Como se relató más arriba, la *TGOpalé* se compone a partir de una *tarifa interna* que recoge la estructura de costes directos e indirectos del OLF más un *margen bruto comercial*. La *tarifa interna*, a su vez, se puede desagregar en cuatro tipos de costes de producción distintos para el OLF: *coste de empresa, de recogida, de arrastre y de distribución*. A la hora de identificar cómo se obtiene el valor monetario que supone cada componente de coste hemos encontrado serias dificultades pues, como ya se adelantó anteriormente, constituyen una información

[12] Para algunos clientes el sistema de tarificación es algo más complejo pues la bonificación podría ser de signo negativo, es decir, se trataría de una penalización económica, normalmente en virtud de un servicio excesivamente complicado.
[13] Por ejemplo, el sector de la alimentación perecedera está afectado por un elevado grado de estacionalidad.
[14] El último componente de la factura, el seguro, dado su carácter opcional, queda excluido de nuestro análisis empírico.

privilegiada propia de la empresa que los OLF no se han brindado a revelar. A continuación se ofrece toda la información que ha sido posible recabar sobre los componentes de la *TGOpalé.*

El *coste de empresa* proviene de imputar la estructura de costes directos e indirectos. Los costes directos se refieren a la explotación del vehículo frigorífico. Dada la ausencia de estimaciones propias para las empresas de la muestra, hemos optado por ilustrar la estructura de costes medios directos que la explotación de un vehículo frigorífico articulado genera a una empresa de transportes de mercancías tipo a partir del análisis desarrollado por el "Observatorio de Costes para el Transporte de Mercancías por Carretera", del Ministerio de Fomento, y actualizado a 31 de octubre de 2007. Este coste directo corresponde a la media nacional, obtenida al ponderar los costes de cada provincia por el peso de ésta en el transporte de mercancías por carretera. Así, el coste directo se puede estimar en 1,024 €/Kilómetro recorrido. Al coste directo se le añaden los costes indirectos de gestión, personal, comercialización, etc… que en su caso soporte el OLF, que a veces resultan difícilmente cuantificables al no guardar una relación directa con el volumen de transporte realizado por la empresa. Por ello se valoran como un porcentaje, bien como un porcentaje del coste directo por kilómetro recorrido (el OLF añade un margen del 20-25 por ciento, imputando como *coste de empresa* 1,2-1,25 €/kilómetro recorrido, cantidad que se multiplica por el número de kilómetros del trayecto y se divide entre la capacidad del camión en número de palés), o bien como un porcentaje del 20-25 por ciento del sumatorio del valor de las demás partidas de coste (*recogida, arrastre* y *distribución*). El uso de porcentajes sobre el coste directo por kilómetro recorrido suele emplearse en el cálculo de tarifas para camiones completos (cargas completas), mientras que la aplicación de un porcentaje del sumatorio del valor de las demás partidas de coste es común en el cálculo de tarifas para cargas fraccionadas, lo que nos ocupa en este trabajo.

El *coste de recogida* representa el coste de recogida de mercancía en las instalaciones del cliente. Aunque es diferente para cada cliente, se valora a efectos de cálculo de la *tarifa interna* como un importe fijo.

Los OLF disponen de unas rutas preestablecidas con un punto de origen y un punto de destino. De este modo, puede suceder que la mercancía realice alguna escala, pasando de una plataforma logística de la compañía a otra, hasta llegar a destino. En cada plataforma se descarga la mercancía del camión para agruparla en otro camión junto con otra mercancía de diversa procedencia cuyo destino es también la siguiente plataforma. Esto provoca que el servicio se encarezca, pues en cada escala se incurre en un coste de manipulación, que se repercute, denominado *coste de arrastre*. Si no se realizan escalas, no hay *arrastre.* Cuando sí se realizan, a cada *arrastre* se le imputa un importe fijo.

El *coste de distribución* se refiere al reparto local hasta el punto de entrega, conocido también como reparto capilar. Junto con el *coste de empresa* constituyen las partidas más sensibles del OLF frente a su competencia, motivo por el que tampoco se ha tenido acceso a una explicación sobre la obtención de su cuantía, aunque sí se puede adelantar que para su cálculo se atribuyen por palé los costes directos del camión de reparto más la gestión correspondiente, proceso que fue descrito para el camión frigorífico articulado.

El *margen bruto comercial* que aplica el OLF se valora como un porcentaje del sumatorio del valor de las anteriores partidas de coste (*coste de empresa, recogida, arrastre* y *distribución*), normalmente supone entre un 20-25 por ciento adicional.

La expresión algebraica que recoge el cálculo de la *TGOpalé* para cada trayecto sería:

$$TGO_{palé} = T^{interna} + M^{comercial} \qquad [1]$$

$$TGO_{palé} = (C^{recogida} + C^{arrastre} + C^{distribución} + C^{empresa}) + M^{comercial} \qquad [2]$$

$$TGO_{palé} = r(C_1 + C_2 + C_3 + s\sum_{i=1}^{3} C_i) \qquad [3]$$

donde r es el porcentaje aplicado en concepto de *margen bruto comercial* (r>1), C_1 es el *coste de recogida*, C_2 el *coste de arrastre* (C_2*=0 si no hay arrastre)*, C_3 el *coste de distribución* y s el

porcentaje aplicado en concepto de *coste de empresa* (s<1). Si a [3] le incluimos la *bonificación* y los *gastos administrativos* obtenemos el *PFpalé*:

$$PF_{palé} = TGO_{palé} - B^{bonificación} + G^{administrativos} \qquad [4]$$

$$PF_{palé} = t[r(C_1 + C_2 + C_3 + s\sum_{i=1}^{3} C_i)] + 3/n \qquad [5]$$

donde t es el porcentaje aplicado en concepto de *bonificación* (t<1) y 3/n representa el *gasto administrativo* por palé (siendo n el número de palés).

4. TRATAMIENTO ESTADÍSTICO DE LOS DATOS

Siguiendo la notación empleada en De Rus (1989), la provisión de servicios de transporte de mercancías a temperatura controlada es una actividad multiproducto, pues puede hablarse de tantos productos como orígenes-destinos posibles en la red, distinguiendo por número de palés. Los datos suministrados por los OLF respecto a su sistema de tarificación nos permiten disponer, para cada trayecto, de la *TGOpalé* del OLF y, para cada cliente y trayecto, del *PFpalé*. A partir de ellos se han definido las variables a estimar en este trabajo para cada trayecto: la *TARIFA MEDIA GENERAL ORDINARIA por palé (TMGOpalé)*, el *PRECIO MEDIO FINAL por palé (PMFpalé)* y el *COSTE TOTAL MEDIO por palé (CTMepalé)*. Veamos el modelo que se ha construido para posibilitar su cálculo.

Para la obtención de la *TMGOpalé* se ha calculado el promedio de la *TGOpalé* de los 8 OLF, ponderando cada *TGOpalé* por el porcentaje que la facturación anual de cada OLF representa dentro de la facturación total. Así, la *TMGOpalé* se puede expresar del siguiente modo:

$$TMGO_{palé} = \frac{\sum_{i=1}^{8} (TGOipalé \cdot Fi)}{\sum_{i=1}^{8} Fi} \qquad [6]$$

A partir de [6], y utilizando el programa Microsoft Office Excel 2007, se ha construido un calculador automático de tarifas medias generales que permite el cálculo inmediato de las *TMGOpalé* para envíos con origen en los principales puntos de expedición de mercancía (Madrid, Barcelona, Valencia, Sevilla y Bilbao) y destino en cualquier capital española. Dado el desglose disponible de las *TGOpalé* en *tarifa interna* más *margen bruto comercial*, el calculador se ha elaborado de tal manera que ofrece al usuario la posibilidad de aplicar a cada *tarifa interna* cualquier *margen comercial*, aumentando de esta manera su potencialidad. Se dispone así de una base de datos sobre tarifas internas y tarifas generales dinámica, que permite realizar múltiples simulaciones.

Para obtener el *PMFpalé* se han promediado las 300 observaciones, una por cliente, disponibles de *PFpalé*. Y para conocer el porcentaje de descuento medio que aplican los OLF en concepto de *bonificación*, se ha pensado en aproximar dicho porcentaje de *bonificación* a partir de la desviación estándar, que considera la media aritmética de las desviaciones respecto al *PMFpalé*, calculando el porcentaje que dicha desviación estándar representa sobre el *PMFpalé*, o lo que es lo mismo, el *coeficiente de variación* (que es insensible a la magnitud de los datos). Así, se ha definido:

$$PMF_{palé} = \frac{\sum_{i=1}^{n} PFipalé}{n} \qquad [7]$$

$$Bonificación = CV = 100(\frac{s}{PMFpalé}) \qquad [8]$$

donde n es el número de observaciones de segundo nivel ($n \leq 300$) y s la desviación estándar para cada trayecto.

Según la notación empleada en Álvarez *et al.* (2007), el coste de producción soportado por el OLF en cada trayecto (que ha sido referido a lo largo de este trabajo como *tarifa interna*, compuesta por la suma de las partidas *coste de empresa, de recogida, de arrastre y de distribución)* se puede considerar una buena aproximación del *coste total medio por palé (CTMepalé).* Conocido el valor por trayecto con origen en Valencia de cada una de las anteriores partidas de coste, se ha calculado la variable *CTMepalé* como promedio de las *tarifas internas* por palé de los OLF de la muestra, ponderando por el porcentaje que la facturación anual de cada OLF representa dentro de la facturación total.

$$CTMepalé = \frac{\sum_{j=1}^{8} \langle C_1 + C_2 + C_3 + s\sum_{i=1}^{3} C_i \rangle \cdot F_j}{\sum_{j=1}^{8} F_j} \qquad [9]$$

Sabemos que la especificación de un modelo estadístico que describa correctamente la realidad que pretendemos describir es una tarea delicada y difícil, especialmente en un escenario donde ya se apuntó que la disponibilidad de información se erige como el sesgo más importante. Para conocer en qué medida el *CTMepalé* contribuye a la explicación de las diferencias en las variables *TMGOpalé* y *PMFpalé* se ha pensado en desarrollar un modelo clásico de regresión, utilizando el método de análisis por mínimos cuadrados. El objetivo de la construcción de este modelo es descubrir una formulación sencilla, pero con la suficiente precisión como para que aporte ideas útiles, que refleje correctamente cómo varían la *TMGOpalé* y el *PMFpalé* cuando varía el *CTMepalé*, así como también cómo varía este último en función de los kilómetros recorridos en cada trayecto, de la existencia de arrastres y de la complejidad del reparto capilar. La existencia de divergencias entre el modelo y la realidad nos llevará a determinar si existen otras variables suficientemente influyentes que hayan sido omitidas del modelo.

De esta forma, se ha ensayado inicialmente la estimación de un modelo lineal. Respecto a la variable *TMGOpalé*, tras un contraste de hipótesis, se ha averiguado que un modelo cuadrático ofrece una mejora en el ajuste. Por otra parte, la discreta variabilidad porcentual explicada (R^2=0,363) del *CTMepalé* en función de la distancia del desplazamiento (medida en Km) por nuestro modelo hace pensar en la necesidad de probar la inclusión de variables adicionales. En el *CTMepalé* la distancia del desplazamiento queda debidamente recogida en el *coste de empresa*; sin embargo, por lo general la partida de mayor valor es el *coste de distribución*, que representa la complicación que enfrenta el OLF en la distribución capilar, factor respecto al que no existe una información fiable para cuantificarlo, por lo que conocemos, a excepción del valor del propio *coste de distribución.* Asimismo, aunque con menor relevancia en el *CTMepalé*, el tránsito de la mercancía por otra plataforma del OLF antes de alcanzar su destino, incluido en el *coste de arrastre*, es otro factor conocido a considerar. En consecuencia, se ha ensayado la incorporación en la función de *CTMepalé* de dos variables ficticias para cuantificar también el efecto sobre los costes medios de la existencia de arrastres y la mayor o menor complejidad del reparto capilar. Los coeficientes estimados tienen significación estadística y se presentan en el Cuadro 2.

Cuadro 2. Coeficientes estimados en la modelización

	Variable Dependiente	Variable Independiente	Coeficiente b_0	Coeficiente b_1	Coeficiente b_2	R^2	Estadístico F
MODELO CUADRÁTICO	TMGOpalé	CTMepalé	8,79	1,88	-0,01	0,44	18,08
MODELO LINEAL	PMFpalé	CTMepalé	39,67	0,76		0,90	405,53
MODELO LINEAL	CTMepalé	Km	39,53	0,01		0,36	26,77
	Variable Dependiente	Coeficiente ARRASTRE	Coeficiente DISTRIBUCIÓN	Coeficiente KM	Coeficiente b_0	R^2	Estadístico F
MODELO LINEAL MÚLTIPLE	CTMepalé	18,01	45,86	0,04	40,39	0,85	84,33

Fuente: elaboración propia.

Las ecuaciones presentadas en el cuadro anterior sólo deben utilizarse para la predicción en el rango de los valores observados de las variables independientes en los que tenemos información (es decir, dentro del territorio nacional), pues fuera de ese rango no conocemos como es la relación; la interpretación de este hecho implica la invalidez de extrapolar nuestros modelos al transporte internacional, ya que este adolece de una problemática algo distinta[15].

5. RESULTADOS

A título informativo, el Cuadro 3 muestra de forma resumida una simulación de los valores de la variable *TMGOpalé*, con origen en Barcelona y Madrid y para todos los destinos, con la aplicación de un margen comercial del 15 y del 25 por ciento respectivamente. El cálculo se ha realizado con el calculador automático construido en este trabajo, que posibilita cualquier otra simulación. Nótese que la variabilidad en las *TMGOpalé* desde un mismo origen parece ser debida a diferencias en el número de kilómetros recorridos, pero también a la existencia o no de arrastres y a la mayor o menor complejidad del reparto capilar; veremos en qué medida el análisis estadístico lo confirma.

[15] Véase en Andrés (2009) un análisis aplicado al caso del transporte comunitario.

Cuadro 3. Tarifas medias generales por palé en el año 2008 (euros/palé)

Origen Barcelona, margen del 15%					Origen Madrid, margen del 25%				
Destino	1 palé	2/3 palés	4/6 palés	7/10 palés	Destino	1 palé	2/3 palés	4/6 palés	7/10 palés
ALAVA	80,53	68,07	60,41	48,33	ALAVA	82,86	70,03	62,14	49,72
ALBACETE	110,82	93,66	83,14	66,52	ALBACETE	61,74	52,19	46,32	37,07
ALGECIRAS	138,01	116,65	103,53	82,84	ALGECIRAS	99,16	83,81	74,38	59,51
ALICANTE	85,92	72,62	64,46	51,57	ALICANTE	93,87	79,34	70,42	56,34
ALMERIA	105,95	89,54	79,47	63,58	ALMERIA	92,01	77,77	69,02	55,23
ANDORRA	154,78	130,82	116,10	92,89	ANDORRA	197,35	166,80	148,04	118,45
ASTURIAS	128,70	108,78	96,54	77,24	ASTURIAS	78,04	65,97	58,54	46,84
AVILA	86,79	73,36	65,11	52,09	AVILA	42,68	36,08	32,02	25,62
BADAJOZ	123,67	104,52	92,77	74,22	BADAJOZ	72,43	61,22	54,34	43,48
BARCELONA	31,82	26,90	23,87	19,10	BARCELONA	84,53	71,44	63,41	50,73
BURGOS	88,97	75,20	66,75	53,41	BURGOS	50,82	42,96	38,13	30,51
CACERES	123,67	104,52	92,77	74,22	CACERES	72,43	61,22	54,34	43,48
CADIZ	115,32	97,48	86,52	69,22	CADIZ	86,23	72,88	64,68	51,75
CANTABRIA	83,81	70,83	62,86	50,30	CANTABRIA	93,87	79,34	70,42	56,34
CASTELLON	71,45	60,38	53,59	42,88	CASTELLON	78,04	65,97	58,54	46,84
CIUDAD REAL	98,31	83,09	73,75	59,01	CIUDAD REAL	57,86	48,91	43,40	34,73
CORDOBA	138,01	116,65	103,53	82,84	CORDOBA	86,23	72,88	64,68	51,75
CUENCA	114,18	96,50	85,65	68,53	CUENCA	66,54	56,23	49,91	39,94
GERONA	39,27	33,19	29,46	23,57	GERONA	91,55	77,37	68,68	54,94
GIBRALTAR	175,33	148,19	131,52	105,23	GIBRALTAR	116,90	98,80	87,69	70,16
GRANADA	175,33	148,19	131,52	105,23	GRANADA	86,23	72,88	64,68	51,75
GUADALAJARA	86,79	73,36	65,11	52,09	GUADALAJARA	42,68	36,08	32,02	25,62
GUIPUZCOA	69,68	58,89	52,27	41,82	GUIPUZCOA	79,22	66,96	59,42	47,55
HUELVA	115,32	97,48	86,52	69,22	HUELVA	86,23	72,88	64,68	51,75
HUESCA	84,65	71,55	63,50	50,81	HUESCA	90,64	76,61	68,00	54,40
IBIZA	164,00	157,96	156,43	156,43	IBIZA	199,44	199,44	199,44	199,44
JAEN	175,33	148,19	131,52	105,23	JAEN	116,90	98,80	87,69	70,16
LA CORUÑA	122,44	103,49	91,85	73,50	LA CORUÑA	92,47	78,15	69,37	55,50
LEON	107,55	90,90	80,68	64,55	LEON	63,93	54,03	47,95	38,37
LERIDA	46,76	39,52	35,08	28,07	LERIDA	97,19	82,15	72,91	58,33
LOGROÑO	83,81	70,83	62,86	50,30	LOGROÑO	93,87	79,34	70,42	56,34
LUGO	138,01	116,65	103,53	82,84	LUGO	98,17	82,98	73,65	58,93
MADRID	80,13	67,72	60,11	48,09	MADRID	34,59	29,24	25,94	20,76
MALAGA	177,10	149,69	132,85	106,30	MALAGA	86,23	72,88	64,68	51,75
MENORCA	164,00	157,96	156,43	156,43	MENORCA	199,44	199,44	199,44	199,44
MERIDA	177,10	149,69	132,85	106,30	MERIDA	72,43	61,22	54,34	43,48
MURCIA	85,49	72,25	64,13	51,31	MURCIA	72,78	61,51	54,60	43,69
NAVARRA	80,53	68,07	60,41	48,33	NAVARRA	82,86	70,03	62,14	49,72
ORENSE	122,44	103,49	91,85	73,50	ORENSE	92,47	78,15	69,37	55,50
PALENCIA	108,09	91,35	81,08	64,87	PALENCIA	76,12	64,35	57,11	45,69
PALMA DE MALLORCA	164,00	157,96	156,43	156,43	PALMA DE MALLORCA	199,44	199,44	199,44	199,44
PIRINEO CATALAN	72,53	61,30	54,41	43,53	PIRINEO CATALAN	114,02	96,37	85,53	68,43
PIRINEO ARAGONES	93,06	78,65	69,81	55,85	PIRINEO ARAGONES	99,66	84,23	74,75	59,81
PONTEVEDRA	110,82	93,66	83,14	66,52	PONTEVEDRA	80,02	67,64	60,03	48,03
SALAMANCA	98,31	83,09	73,75	59,01	SALAMANCA	57,86	48,91	43,40	34,73
SEGOVIA	86,79	73,36	65,11	52,09	SEGOVIA	42,68	36,08	32,02	25,62
SEVILLA	111,36	94,12	83,53	66,84	SEVILLA	80,81	68,31	60,62	48,50
SORIA	93,06	78,65	69,81	55,85	SORIA	99,66	84,23	74,75	59,81
TARRAGONA	39,27	33,19	29,46	23,57	TARRAGONA	91,55	77,37	68,68	54,94
TERUEL	93,06	78,65	69,81	55,85	TERUEL	99,66	84,23	74,75	59,81
TOLEDO	86,79	73,36	65,11	52,09	TOLEDO	42,68	36,08	32,02	25,62
VALENCIA	67,97	57,46	50,99	40,79	VALENCIA	74,24	62,75	55,70	44,56
VALLADOLID	92,14	77,88	69,12	55,31	VALLADOLID	52,11	44,04	39,10	31,28
VIZCAYA	81,34	68,75	61,02	48,82	VIZCAYA	76,51	64,67	57,39	45,92
ZAMORA	111,92	94,60	83,96	67,17	ZAMORA	63,30	53,50	47,48	37,99
ZARAGOZA	66,96	56,59	50,23	40,19	ZARAGOZA	71,70	60,60	53,78	43,03

Fuente: elaboración propia.

El Cuadro 4 ofrece los principales resultados de la variable *PMFpalé* con origen en Valencia y para todos los destinos. En él podemos observar el *PMFpalé* en función del número de palés, así como el *coeficiente de variación,* que se utiliza como una aproximación del porcentaje medio de *bonificación* en cada trayecto. La variabilidad en los *PFMpalé* parece originarse, al depender estos últimos de las tarifas generales, en diferencias en el número de kilómetros recorridos, en la existencia o no de arrastres y en la mayor o menor complejidad del reparto capilar, pero también en el porcentaje de *bonificación*. En este sentido, de la lectura del Cuadro 4 pueden extraerse algunas conclusiones interesantes:

- Como se apuntó anteriormente, los límites entre los que queda comprendido el porcentaje de *bonificación* son cero (en cuyo caso el *precio final* coincide con la *tarifa general* incrementada en los *gastos administrativos* correspondientes) y el

margen bruto comercial. Obsérvese que la máxima *bonificación* media por trayecto que hemos hallado se acerca al 22 por ciento, provocando que los *márgenes netos de rentabilidad* de los OLF, especialmente en los trayectos más solicitados, queden establecidos entre el 3 y el 5 por ciento[16]. Nuestros resultados corroboran, en este sentido, la ajustada rentabilidad que afecta al sector del transporte de mercancías.

- Los porcentajes medios de *bonificación* oscilan entre el 3 y el 22 por ciento. Los valores más bajos, hasta un 5 por ciento, se corresponden con trayectos con muy poco tráfico de mercancía, especialmente alejados como Andorra, Mérida, Algeciras y Gibraltar, o complicados como las Islas Baleares (que requieren transporte marítimo). Desde Valencia, la falta de sinergias con dichos destinos y grados muy bajos de ocupación de los camiones imposibilitan a los OLF conceder mayores *bonificaciones*. Algo muy similar es lo que ocurre con destinos como Asturias, Extremadura o las capitales andaluzas a medida que se alejan de Sevilla, que no alcanzan el 10 por ciento. En el otro extremo destacan las más altas *bonificaciones*, superando el 15 por ciento, que caracterizan a los trayectos más solicitados o más cercanos al punto de origen (por ellos se canaliza la mayor parte de la mercancía), como es la propia Comunidad Valenciana, Murcia, Cuenca, Sevilla y el País Vasco. Nuestros resultados corroboran la hipótesis de que el precio final del transporte no sólo depende del coste kilométrico, que aumenta con la distancia a recorrer, sino también del volumen de negocio que maneja el OLF con el punto de destino, que actúa determinando la existencia de sinergias y el grado de ocupación de cada vehículo. Parece razonable suponer que en los trayectos con más tráfico de mercancía se concentrarán envíos con mayor volumen y frecuencia, lo que nos lleva a pensar que la importancia del cliente que solicita el servicio o la necesidad de vender del OLF pierden relevancia como variables explicativas del porcentaje de *bonificación*.

Los resultados obtenidos a partir del modelo desarrollado, cuyos coeficientes estimados para las ecuaciones planteadas se presentaron en el Cuadro 2, permiten sostener algunas hipótesis ciertamente interesantes respecto al sistema de tarificación que nos ocupa:

- El aumento de la *TMGpalé* es cada vez menor a medida que aumenta el *CTMepalé*, aunque el modelo cuadrático que se ha ensayado no constituye un buen ajuste de los datos observados, pues sólo un 44 por ciento de la variabilidad total es explicada por el modelo (R^2=0,44). Este hecho sugiere que en el establecimiento de las *tarifas generales ordinarias* por parte de los OLF, además del *CTMepalé*, influyen también otros factores cuya inclusión en esta modelización resulta complicada por la dificultad en su medición. Concretamente, se han identificado dos factores significativos generadores de tales divergencias: en primer lugar, el porcentaje efectivo de *margen bruto comercial* aplicado por el OLF en el cálculo de sus *tarifas generales ordinarias* no es el mismo para todos los trayectos; en segundo lugar, el OLF ofrece en algunas situaciones *tarifas generales ordinarias* inferiores al *coste total medio* calculado para el desplazamiento.
- La teoría económica sugiere que existe una relación causal entre el *PMFpalé* y el *CTMepalé* y nuestro resultados apoyan esta teoría, dada la estrecha correlación hallada entre ambas variables (r=0,95). Así, a la hora de explicar la variabilidad de los *precios finales* nuestro modelo lo hace en un 90 por ciento (R^2=0,90). Este resultado sugiere que los *precios finales* que cobran los OLF sí son un fiel reflejo del *CTMepalé*, de tal manera que, la *bonificación*, como única diferencia entre la *TGOpalé* y el *PFpalé*, actúa limando las divergencias que se han constatado entre la *tarifa general* y el *coste total medio* para un palé. Expresado con otras palabras, los OLF están tarificando según el *coste total medio*. La tasa de variación de los *precios finales* con relación al *coste medio por palé* tiene un valor aproximado de 0,76.

[16] Estos valores son más propios de un margen que añade valor al producto que de un margen especulativo.

Cuadro 4. Precios medios finales por palé con origen en Valencia en el año 2008 (euros/palé)

	1 palé		2/3 palés		4/6 palés		7/10 palés	
Destino	PMF palé	CV	PMF palé	CV	PMF palé	CV	PMF palé	CV
ALAVA	94,84	18,36	88,42	19,61	83,74	20,53	77,74	15,21
ALBACETE	79,95	13,43	74,37	14,44	69,88	15,19	64,96	13,72
ALGECIRAS	120,69	3,93	112,33	3,94	106,72	3,87	99,97	4,43
ALICANTE	50,37	20,59	47,60	19,73	45,18	20,04	43,25	19,86
ALMERIA	87,84	13,32	82,26	13,71	76,92	12,75	72,58	16,16
ANDORRA	196,35	3,12	166,19	3,64	147,55	3,79	117,68	2,75
ASTURIAS	140,31	8,53	127,60	9,04	113,04	9,11	106,30	9,94
AVILA	101,06	13,63	95,10	14,81	83,22	15,37	78,14	12,28
BADAJOZ	132,72	10,86	120,19	11,21	109,83	11,19	103,21	10,92
BARCELONA	70,89	12,16	65,94	14,11	62,30	16,67	59,13	13,69
BURGOS	123,69	14,15	112,60	14,07	103,16	14,84	97,74	13,43
CACERES	132,45	9,50	120,08	9,93	109,72	9,89	103,26	10,79
CADIZ	104,12	11,95	97,08	12,31	90,41	12,57	84,30	11,38
CANTABRIA	151,34	12,60	137,85	14,45	116,80	14,94	109,60	12,10
CASTELLON	40,83	18,62	38,37	20,28	35,92	20,89	34,07	23,47
CIUDAD REAL	111,91	11,64	105,33	12,61	93,34	12,89	87,66	11,37
CORDOBA	110,99	7,86	102,29	9,49	95,33	9,50	88,51	7,89
CUENCA	126,08	19,09	115,17	19,67	103,11	19,56	97,94	19,20
GERONA	88,55	11,16	83,13	11,36	78,98	11,11	74,56	10,55
GIBRALTAR	158,41	3,57	133,89	3,57	118,83	3,56	95,10	3,43
GRANADA	114,44	7,34	105,46	6,88	97,55	6,84	89,59	7,53
GUADALAJARA	85,38	13,23	79,69	13,97	75,13	14,49	70,14	12,28
GUIPUZCOA	92,25	16,29	86,12	16,77	81,64	17,08	76,46	16,47
HUELVA	107,32	13,12	100,20	13,55	93,52	13,74	87,48	13,22
HUESCA	155,98	13,62	141,74	15,47	117,11	15,93	110,24	13,74
IBIZA	138,57	3,50	138,57	3,50	138,32	3,09	138,70	2,82
JAEN	121,86	7,45	112,58	7,12	100,58	7,04	92,89	7,19
LA CORUÑA	148,07	12,26	134,35	14,08	116,21	14,81	109,15	12,69
LEON	127,93	13,54	116,45	14,48	105,32	14,42	99,47	11,74
LERIDA	91,59	13,83	85,71	14,57	81,25	14,50	76,15	11,23
LOGROÑO	101,23	10,20	93,84	11,45	88,42	11,73	81,70	12,54
LUGO	157,29	11,70	142,99	13,55	119,57	14,10	112,10	11,48
MADRID	71,55	13,95	66,62	15,73	62,66	17,99	59,59	16,22
MALAGA	113,56	7,05	104,51	6,67	97,22	6,69	89,30	7,21
MENORCA	144,39	5,41	144,33	5,42	144,33	5,42	143,77	7,06
MERIDA	144,94	3,09	130,67	2,96	117,94	2,96	109,97	2,77
MURCIA	57,00	21,74	53,22	23,20	51,06	21,92	48,27	24,91
NAVARRA	94,94	12,57	88,42	13,46	83,65	14,11	77,88	14,62
ORENSE	148,22	12,70	134,36	14,35	116,23	15,08	109,03	12,54
PALENCIA	130,44	13,59	119,00	14,57	106,12	17,90	100,62	11,93
PALMA DE MALLORCA	138,34	3,91	138,28	3,76	138,28	3,76	138,36	3,20
PIRINEO CATALAN	123,15	12,07	101,64	10,43	85,14	6,45	72,38	10,13
PIRINEO ARAGONES	141,58	4,89	125,15	7,49	103,95	4,00	88,01	6,37
PONTEVEDRA	133,56	14,44	121,97	16,10	107,85	17,16	101,93	14,14
SALAMANCA	118,61	11,48	108,62	12,25	100,29	12,50	94,66	11,37
SEGOVIA	88,84	11,10	83,15	11,52	79,60	11,74	74,85	12,40
SEVILLA	95,99	20,19	91,07	18,45	87,26	16,16	83,21	11,73
SORIA	161,24	11,06	146,37	12,43	120,04	12,83	112,75	11,59
TARRAGONA	87,58	11,07	82,20	11,01	75,53	10,92	71,16	10,05
TERUEL	104,77	12,36	96,89	13,81	91,19	14,23	83,75	11,59
TOLEDO	98,62	12,91	89,30	13,69	79,74	14,22	74,81	12,28
VALENCIA	36,08	17,51	33,90	18,90	31,61	19,66	30,07	21,10
VALLADOLID	116,55	11,60	106,51	12,66	98,40	12,82	93,19	13,37
VIZCAYA	94,58	15,09	88,07	15,35	83,39	15,49	77,80	14,25
ZAMORA	129,14	13,93	117,53	14,56	107,25	14,39	101,27	13,19
ZARAGOZA	90,23	11,74	84,26	12,85	78,38	13,59	73,28	13,47

Fuente: elaboración propia.

- La alta variabilidad porcentual del *CTMepalé* explicada por la regresión múltiple (R^2=0,85), que ha mejorado sustancialmente respecto a la ofrecida por un modelo simple donde los costes medios son dependiente tan solo de la distancia (R^2=0,36), demuestra que la existencia de arrastres y la dificultad en el reparto capilar, junto a los kilómetros recorridos, son los costes de producción más relevantes para los OLF. Sin embargo, su variación marginal es más difícil de estimar, pues dichas variables independientes están a su vez correlacionadas entre sí (r=0,47; r=0,23; r=0,64,

respectivamente), por lo que el coeficiente de cada una de ellas no muestra el efecto que produce esa variable en todas las condiciones sino el efecto cuando se combina con las demás (coeficientes condicionados). La tasa de variación de los *costes medios por palé* en relación a la distancia del desplazamiento tiene un valor aproximado de 0,04, elevándose el *coste medio* en torno al 33 por ciento cuando hay arrastre, y hasta un 80 por ciento como efecto de un reparto capilar complicado. El error del modelo refleja la importancia de otros costes de producción cuya inclusión en esta modelización no ha sido posible porque su valor no se conoce a ciencia cierta (recuérdese que en el modelo de tarificación de los OLF se imputaron por aproximación): los costes indirectos de explotación (recogidos en el *coste de empresa*) y el *coste de recogida*.

- Los resultados obtenidos en la estimación de la función de *costes medios por palé*, aunque implícitamente, permiten sostener la hipótesis de existencia de economías de densidad en la provisión de transporte frigorífico, es decir, el coste medio de transportar un palé decrece con el número de palés transportados. En este caso es posible conocer cómo se comporta el *coste medio por palé* ante variaciones en el flujo total de palés transportados por envío: la evidencia empírica aportada por nuestro trabajo revela que el aumento de 1 palé adicional en el número de palés incluidos en el envío reduce el *coste medio por palé*, en promedio, un 3 por ciento. La presencia de economías de densidad es de suma importancia al ser este un concepto relevante cuando se negocia la política de precios, con un efecto sobre los *precios finales por palé*, en función del número de palés, que ha quedado debidamente ilustrado en el Cuadro 4 de este trabajo.

6. CONCLUSIONES

Este trabajo ha abordado la formación de precios en el transporte de mercancías en carga fraccionada, con énfasis en el caso de los productos perecederos, analizando algunas de las cuestiones más relevantes que se plantean en la determinación de los precios que van a emplearse: cuáles son y qué valor toman los componentes que integran el precio final y cómo se relacionan con el coste de producción derivado de la actividad.

El sistema generalizado de tarificación emplea precios finales calculados expresamente para cada cliente, dado que hay demandas de distinta intensidad a lo largo del año y trayectos con diferencias acusadas en los costes de producción. De esta manera, la *bonificación* que se aplica sobre la *TGOpalé* incorpora en la estructura tarifaria las diferencias más notables en los costes de producción, evitando el componente discriminatorio de la tarifa general única por trayecto. Las estimaciones obtenidas en este trabajo, reforzadas con los contrastes empleados, permiten sostener la hipótesis de que los OLF están tarificando según el *coste total medio*, y no en función de la competencia o del mercado.

Podemos concluir resaltando dos cuestiones fundamentales para la estrategia de la empresa cargadora usuaria del servicio que se derivan del contenido de este trabajo:

1. El *PFpalé* ofrecido por el OLF adolece de cierta complejidad en su cálculo y es muy variable. Se han caracterizado dos fuentes de variabilidad. Por un lado, la *TGOpalé* que se utiliza como referencia para la obtención de los *precios finales* está basada en la estructura de costes directos e indirectos de explotación del OLF; por consiguiente, tomará un valor distinto dependiendo del OLF al que se contrate. Por otro lado, a la *TGOpalé* se la aplica una *bonificación*, que depende básicamente de la red estructural del OLF (que explica la existencia de sinergias y el grado de ocupación del vehículo para cada trayecto); por consiguiente, el *PFpalé* cambiará no sólo en función del OLF que se considere, sino también para un mismo OLF en función de la época del año en que se contrate el servicio.
2. La evidencia anterior favorece a las industrias de marcado carácter competitivo, como la de la alimentación, que minimizan sus costes de producción. Si ponderan

conveniente la idoneidad de trabajar conjuntamente con dos o más OLF como proveedores, obtendrán una excelente oportunidad para lograr una reducción sustancial de sus costes de transporte.

REFERENCIAS BIBLIOGRÁFICAS

ÁLVAREZ, O., CANTOS, P. y PEREIRA, R. (2007): "Precios óptimos en el transporte interurbano por carretera". Revista de Economía Aplicada, vol. 45, págs. 155-182.

AHLSTRand, I. (2001): "The politics and economics of transport investment and pricing in Stockholm". Journal of Transport Economics and Policy, vol. 35, págs. 473-489.

andERSON, J.E. y VAN WINCOOP, E. (2004): "Gravity with gravitas: a solution to the border puzzle". American Economic Review, vol. 93, págs. 170-192.

andRÉS, S. (2009): "Intercambios comunitarios y coste del transporte en carga fraccionada: aplicación a los productos perecederos" (mimeo).

BARKER, T. y KÖHLER, J. (2000): "Changing for road freight in the EU. Economic implications of a weigh-inmotion tax". Journal of Transport Economics and Policy, vol. 34, págs. 311-332.

BAyLISS, B. (1998): "Regulation in the road freight transport sector". Journal of Transport Economics and Policy, vol. 23, págs. 113-131.

BERECHMAN, J. (1987): "Cost structure and production technology in transit". Regional Science and Urban Economics, vol. 17, págs. 519-534.

BERECHMAN, J. y GULIANO, G. (1984): "Analysis of the cost structure of an urban bus transit property". Transportation Research B, vol. 18, págs. 277-287.

BONACHE, J. (1999): "El Estudio de casos como estrategia de construcción teórica: características, críticas y defensas". Cuadernos de Economía y Dirección de la Empresa, vol. 3, págs. 123-141.

CAVES, D.W. y CHRISTENSEN, L.R. (1988): "The importante of economies of scale, capacity utilization, and density in explaining interindustry differences in productivity growth". The Logistic and Transportation Review, vol. 24, págs. 3-32.

DAUGHETy, A.F., NELSON, F.D. y VIGDOR, W.R. (1985): "An econometric analysis of the cost and production structure of the trucking industry". En Daughety, A.F. (ed): Analitical Studies in Transport Economics. Cambridge, Cambridge University Press, capítulo 2.

DE RUS, G. (1987): "Discriminación de precios y subvenciones cruzadas en transporte público". Investigaciones Económicas, vol. 11, nº 2, págs 201-218.

DE RUS, G. (1989): "Análisis del Mercado de servicios de transporte público en España: costes, demanda, precios y nivel de calidad". Investigaciones Económicas, vol. 15, nº 2, págs 229-247.

DE RUS, G. (1991): "El transporte público urbano en España: comportamiento de los costes y regulación de la industria". Investigaciones Económicas, vol. 13, nº 2, págs 207-225.

GARCÍA, L. y PÉREZ, E.M. (2007): "Fundamentos metodológicos de la base de datos TRADETRANS para el análisis de costes de transporte en España". Información Comercial Española, vol. 834, págs. 61-78.

GREEN, D.L., JONES, D.W. y DELUCCHI, M.A. (eds.) (1997): The full cost and benefits of transportation: contributions to theory, method and measurement, Berlin, Springer Verlag.

HARMATUCK, D.J. (1991): "Economies of scale and scope in the motor carrier industry". Journal of Transport Economics and Policy, vol. 25, págs. 135-151.

HENSHER, D. (1987): "Productive efficiency and ownership of urban bus services". Transportation, vol. 14, págs. 209-225.

HURLEy, W. (1995): "Is it necessary to regulate prices in freight transport markets?". Journal of Transport Economics and Policy, vol. 29, págs. 169-178.

MIR, J. y BORRÁS, F. (2008): "Los costes de la cadena de distribución de productos hortícolas en fresco". Distribución y Consumo, vol. 98, págs. 55-63.

NEWBERy, D. (1988): "Road user charges in Britain". The Economic Journal, vol. 98, págs. 161-176.

NEWBERy, D. (1989): "Cost recovery from optimally designed roads". Economica, vol. 56, págs. 165-185.

NEWBERy, D. (1991): "Pricing and congestion: economic principles relevant to the pricing of roads". Oxford Review of Economic Policy, vol. 6, págs. 22-38.

REBOLLO, A., ROMERO, J. y yAGÜE, M.J. (2006): "El coste de la comercialización de los productos de alimentación en fresco". Distribución y Consumo, vol. 85, págs. 31-51.

SANSOM, T., NASH, C., MACKIE, P., SHIRES, J. y WATKISS, P. (2001): "Surface transport costs and charges". Institute for Transport Studies, University of Leeds.

TALLEy, W.K. y andERSON, E.E. (1986): "An urban transit firm providing transit, para-transit and contracted-out services". Journal of Transport Economics and Policy, págs. 353-368.

VERMEULEN, J.P.L., BOON, B.H., VAN HESSEN, H.P., DEN BOER, L.C., DINGS, J.M.W., BRUINSMA, F.R. y KOETSE M.J. (2004): "The price of transport. Overview of the social costs of transport", Delft, CE (Centre for Energy Conservation and Environmental Technology).

VIVÓ, D. (2007): "Informe sectorial: logística del frío". Infopack, vol. 126, págs. 8-20.

WILLIAMS, M. y DALAL, A. (1981): "Estimation of the elasticities of factor substitution in urban bus transportation: a cost function approach". Journal of Regional Science, vol. 21, págs. 263-275.

WILLIAMS, M. y HALL, C. (1981): "Return to scale in the United States intercity bus industry". Regional Science and Urban Economics, vol. 21, págs. 573-584.

A STOCHASTIC WATER QUALITY INDEX, AND APPLICATION TO THE TERRITORIAL AREA OF THE CONFEDERACIÓN HIDROGRÁFICA DEL JÚCAR (SPAIN)

EDUARDO BEAMONTE CÓRDOBA
E-mail: beamonte@uv.es
ALEJANDRO CASINO MARTÍNEZ
E-mail: alejandro.casino@uv.es
ERNESTO J. VERES FERRER
E-mail: ernesto.veres@uv.es
Departamento de Economía Aplicada
Universitat de València

Abstract

Water quality can be defined using a subset of the many physical and chemical characteristics that are available. In practice, it is difficult to clearly describe water quality by separately examining the behavior of many variables. For this reason, there have been many attempts to integrate a set of physical and chemical characteristics into a unique value. In this paper, we proposed a stochastic index for measuring water quality. The great advantage of this index is that it accounts for the underlying uncertainty in quality classification that results from variations in the data for individual physical and chemical parameters. Moreover, the index can be used to analyze possible temporal changes in water quality, as well as to determine the sensitivity of the index to each of the parameters used in its calculation. We used information from the Confederación Hidrográfica del Júcar (Spain) for the period between 1990 and 2005.

Keywords: Water quality, Probabilistic index, General quality index, Statistical data analysis
Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

1. INTRODUCTION

The concept of water quality is complex because so many factors influence in it. In particular, this concept is intrinsically tied to the different intended uses of the water (Poch, 1999); different uses require different criteria. Because Spain must comply with the regulations of the European Union, the specific Spanish legislation that governs the quality of water used for human consumption (Council Directives 75/440/EEC of 16 June 1975 and 79/869/EEC of 9 October 1979) is more demanding than the legislation that governs water that is only used to support aquatic life (Council Directive 78/659/CEE of 18 July 1978) or that which governs the protection of groundwater against pollution caused by certain dangerous substances (Council Directive 80/68/CEE of 17 December 1980).

Water quality is one of the most important factors that must be considered when evaluating the sustainable development of a given region. For this reason, concerns over how to correctly measure water quality have increased in recent years. Much empirical research has demonstrated the need for a rigorous definition of water quality, including the work of Beamonte et al. (2005, 2007), Graça and Coimbra (1998), Icaga (2007), Kindler et al. (1998), Moatar et al. (2001), Prat and Munné (2000), Reisenhofer et al. (1998), Sánchez et al. (2007), and Simões et al. (2008).

Water quality must be defined based on a set of physical and chemical parameters that are closely related to the water's intended use. For each parameter, acceptable and unacceptable

values must then be defined. Water whose parameters meet the pre-established standards for a given use is considered suitable for that use. If the water fails to meet these standards, it must be treated before use.

Many physical and chemical parameters can be used to evaluate water quality or the degree of water pollution. Therefore, it is not possible in practice to clearly define water quality—either on a spatial or temporal basis—by separately examining the behavior of every individual parameter (Ministerio de Obras Públicas y Urbanismo, 1983).

The alternative, which is also difficult (Beamonte et al. 2005), consists of integrating the values of a set of physical and chemical parameters into a unique value (i.e., an overall or global index). Provencher and Lamontagne (1977) dealt with this problem in a general form. They defined a quality index for any use of water by simply determining the specifications required by that use. This indicator included various physical and chemical parameters. For each parameter, the index included a quality-value function (generally linear) that expressed the equivalence between the parameter and its quality level. These functions were defined using direct measurements of the concentration of a substance or the value of a physical parameter obtained through analyses of water samples.

The strongly subjective character of the equivalence functions is a problem with the index of Provencher and Lamontagne. In addition, these functions are not universal: they cannot be applied in all geographical areas. Because of this, it is not possible to use the index for direct comparisons of water quality between regions or countries.

Recently, a global indicator of water quality that takes advantage of the potential of statistical tools has appeared in the international literature: the probabilistic index. The objectivity of its development is its great advantage. The index takes into account the uncertainty about the quality classification that remains after the data have been observed (Beamonte et al. 2005).

In this approach, every parameter used to define water quality is considered to be a random variable. The initial and final probability distributions of each variable are then defined according to the classic Bayesian scheme. The key factor in designing this index consists of modeling the 95th percentile of the sampling distribution. This approach assumes a normal distribution of the data, either because it is inherently normal or because it can be normalized by an appropriate transformation. Alternatively, a log-normal mixed distribution can be used if many zero values exist in the measurements of some parameters.

The probabilistic index was initially defined for water intended for use in the production of drinking water. The index (I^*) can be expressed in the following form:

$$I^* = \frac{1}{6}(s_1^3 + 3s_1^2 + 2s_1) + \frac{1}{2}(s_2^2 + s_2) + v_1 + 1 \qquad (1),$$

where s_1, s_2, and v_1 are values obtained according to the probability that the different physical and chemical characteristics belong to prearranged quality intervals that have been deduced from the corresponding 95th percentile for that parameter, and by attending to the concretely probabilistic model that represents the data's behaviour. For details of this approach, see Beamonte et al. (2005).

In this paper, we adapted the index of Beamonte et al. to obtain the probabilistic index (hereafter, *PI*) from the same general perspective used by Provencher and Lamontagne (hereafter, the general quality index, *GQI*).

2. MATERIALS AND METHODS

2.1 The Hydrographic Basin of the Júcar River (Spain)

Hydrographic federations are units of the Spanish government that are charged with the control and management of the surface and underground waters of a specified region.

The territory administered by the Confederación Hidrográfica del Júcar is located in Eastern Spain. It includes all the basins that spill into the Mediterranean sea in the region between the mouth of the Segura River and the mouth of the Cenia River. This area covers 42 989 km^2 spread over the provinces of Albacete, Alicante, Castellón, Cuenca, Valencia, and Teruel, as well as a small part of Tarragona Province. The population of this area is nearly 4.5 million, with a floating population of 800 000. A complete description of the basins (in Spanish) can be found at the Web site http://www.chj.es.

Geologically, the parent materials are primarily limestone and related calcareous rocks, but with small areas of silicate rock, clayey materials, and selenium materials. The limestone materials account for 85% of the total, and include limestone and dolomites (21%), calcareous materials (38%), and marls (26%).

Basins of the the Confederación Hidrográfica del Júcar have markedly Mediterranean characteristics, with strong droughts in summer and frequently heavy rains in the autumn. Only three rivers have an average flow greater than 10 m^3/s: the Mijares River, the Turia River, and the Júcar River. The Júcar River has the highest flow, with an average annual contribution of 1825 hm^3, representing 55% of the total resources available.

The Confederación Hidrográfica del Júcar manages the following nine water-exploitation systems (data and maps available at http://www.chj.es):

- Cenia-Maestrazgo: This system includes all of the basins of the Cenia, Valviquera, Cervol, Barranco de Agua Oliva, Cervera, Alcalá, and San Miguel rivers, as well as all the coastal subbasins between the Cenia River and the municipal boundaries of Oropesa and Benicasim.
- Mijares-Plana de Castellón: This system includes all of the basins of the Mijares, Seco, Veo, and Belcaire rivers, as well as all the coastal basins between Benicasim (including its municipal area) and the provincial border between Castellón and Valencia.
- Palancia and the Valleys: This system includes the basin of the Palancia River, the coastal subbasins between the provincial borders of Valencia and Castellón, and the municipal boundaries between Sagunto and Puzol.
- Turia: This system includes all of the basin of the Turia River, as well as the basins of the Carraixet and Poyo ravines and the coastal subbasins between the northern boundary of Puzol's municipal area and the Gola del Saler.
- Júcar: This system includes all of the basin of the Júcar River, including the areas serviced by the Canal Júcar-Turia, and the coastal subbasins between the Gola del Saler and the municipal boundaries of Cullera and Tabernes de Valldigna.
- Serpis: This system includes all of the basins of the Serpis, Jaraco, and Beniopa rivers and the coastal subbasins between the southern municipal boundary of Cullera and the northern boundary of Oliva.
- Marina Alta: This system includes the basins of the Girona and Gorgos rivers and the coastal subbasins between the northern municipal boundaries of Oliva and the Algar River.
- Marina Baja: This system includes the basins of the Algar and Amadorio rivers and the coastal subbasins between the Algar River and the southern municipal boundary of Villajoyosa.
- Vinalopó-Alacantí: This system includes the basins of the Monegre, Rambla de Rambuchar, and Vinalopó rivers and the coastal subbasins between the northern municipal boundary of El Campello and the territorial area of the Confederación Hidrográfica del Segura.

2.2 Data sources for the water quality parameters

A system of measurement stations has been established along the rivers in the study area. Its purpose is to obtain samples that can be used for systematic analyses of the water quality of the rivers and, consequently, provide a warning of possible pollution problems. These stations constitute the Red Integral de Calidad del Agua (Integral Network of Water Quality). The system is designed so that a station exists at strategic points in rivers and reservoirs, and in major channels and irrigation ditches. Every station has a fixed location, with easy access, and its map coordinates are precisely known.

In the present study, we selected 22 stations capable of providing the largest quantity of information. Table 1 presents their codes, geographical data (exploitation system and province), and coordinates.

Table 1. Selected control stations. (Source: Confederación Hidrográfica del Júcar)

Station code	River (exploitation system/province)	Location (municipality and UTM coordinates - zone 30)
C302	Turia River (Turia/Teruel)	Gea de Albarracín UTMX 642506 UTMY 4474199
D201	Júcar River (Júcar/Cuenca)	Huélamo UTMX 598657 UTMY 4454219
E502	Mijares River (Mijares-Plana de Castellón/Castellón)	Montanejos UTMX 709885 UTMY 4440520
E503	Mijares River (Mijares-Plana de Castellón/Castellón)	Puebla de Arenoso UTMX 704060 UTMY 4443660
F201	Cabriel River (Júcar/Cuenca)	Pajaroncillo UTMX 610397 UTMY 4421989
F301	Turia River (Turia/Cuenca)	Santa Cruz de Moya UTMX 652250 UTMY 4424250
F401	Turia River (Turia/Valencia)	Aras de Alpuente UTMX 658762 UTMY 4411812
F702	Mijares River (Mijares-Plana de Castellón/Castellón)	Villareal UTMX 745762 UTMY 4427500
G403	Tuéjar River (Turia/Valencia)	Calles UTMX 675094 UTMY 4398527
H102	Júcar River (Júcar/Cuenca)	Alarcón UTMX 576100 UTMY 4379650
H402	Turia River (Turia/Valencia)	Chulilla UTMX 679206 UTMY 4392940
H403	Sot River (Turia/Valencia)	Chulilla UTMX 683067 UTMY 4387463
H601	Turia River (Turia/Valencia)	Manises UTMX 714072 UTMY 4378038
I402	Magro River (Júcar/Valencia)	Requena UTMX 671410 UTMY 4360201
I502	Magro River (Júcar/Valencia)	Macastre UTMX 687560 UTMY 4357320
J601	Júcar River (Júcar/Valencia)	Cullera UTMX 736075 UTMY 4339958

K202	Júcar River (Júcar/Albacete)	Albacete UTMX 608023 UTMY 4333092
K602	Verde River (Júcar/Valencia)	Alzira UTMX 720693 UTMY 4338460
L502	Cáñoles River (Júcar/Valencia)	Xátiva UTMX 710625 UTMY 4316840
M004	Mirón River (Júcar/Albacete)	Casas de Lázaro UTMX 567600 UTMY 4286800
M603	Serpis River (Serpis/Alicante)	Muro de Alcoy UTMX 724800 UTMY 4295925
O501	Vinalopó River (Vinalopó-Alacantí/Alicante)	Aspe UTMX 698951 UTMY 4245593

We considered nine water quality parameters in this study. These are the basic parameters used to calculate *GQI*, and we also used them to calculate *PI* so as to permit a more direct comparison between the two indices and because sufficient information was available to calculate both *GQI* and *PI* from data at each of the 22 selected monitoring stations, thereby permitting a statistical analysis of the two indices. The source of the information was the Confederación Hidrográfica del Júcar, and covered the period between 1990 and 2005.

3. RESULTS AND DISCUSSION

Table 2 summarizes the sample size, means, and standard deviations of the observed values from 1990 to 2005 for each parameter at each of the selected stations.

Table 2. Sample sizes, means, and standard deviations for all parameters at each control station

Station code		Total coliforms at 37ºC [a]	Conduc -tivity at 20ºC [b]	Biochemical oxygen demand [c]	Chemical oxygen demand [c]	Total phos-phates [d]	Nitrates [d]	Dissolved oxygen [c]	pH [e]	Suspended matter [f]
C302	N	116	156	190	113	118	128	161	163	193
	Mean	34250.05	594.53	1.26	5.73	0.04	4.33	9.66	8.21	15.48
	S. D.	70714.31	274.83	1.75	7.28	0.06	2.73	2.17	0.28	55.63
D201	N	97	113	158	63	52	53	144	142	160
	Mean	5305.72	632.75	1.07	4.87	0.04	1.56	10.12	8.31	14.67
	S. D.	25702.42	185.07	1.14	5.46	0.21	1.71	1.55	0.20	45.56
E502	N	111	141	179	74	62	61	157	151	179
	Mean	20265.50	739.83	1.20	6.20	0.06	5.42	10.28	8.23	7.81
	S. D.	116927.51	84.41	1.48	7.45	0.28	4.33	1.56	0.30	11.96
E503	N	134	172	220	107	99	98	178	185	220
	Mean	3671.34	769.53	1.36	3.98	0.08	5.74	10.14	8.11	28.55
	S. D.	9223.87	64.77	3.21	5.44	0.64	2.63	8.03	0.27	88.81
F201	N	132	142	191	91	83	86	177	176	191
	Mean	3768.75	652.66	1.05	4.80	0.04	4.08	9.81	8.30	18.15
	S. D.	7588.87	148.27	1.53	5.80	0.07	2.25	1.56	0.27	70.71
F301	N	43	116	143	54	34	52	134	134	142
	Mean	43447.20	1168.30	1.79	6.79	0.17	6.36	9.54	8.10	132.41
	S. D.	64161.31	298.56	3.45	9.67	0.17	3.93	2.17	0.24	342.01
F401	N	109	163	180	107	77	100	160	159	180
	Mean	28238.58	1161.50	1.34	6.95	0.11	7.51	9.61	8.29	205.83
	S. D.	71482.50	269.48	1.73	13.65	0.13	4.19	1.71	0.23	679.54
F702	N	59	72	69	49	59	58	63	65	70
	Mean	2693.36	914.63	2.12	8.59	0.31	4.51	9.78	8.13	18.76
	S. D.	13497.48	948.27	4.46	11.99	1.47	2.28	1.83	0.37	21.80
G403	N	115	130	152	91	115	105	138	144	152
	Mean	91448.63	858.68	1.66	5.24	0.14	12.74	10.11	8.43	108.96
	S. D.	146496.84	82.83	2.33	8.02	0.13	4.41	1.63	0.25	698.60
H102	N	135	165	188	92	85	86	197	193	189
	Mean	336.72	692.17	1.22	5.49	0.01	2.61	9.70	8.19	5.93
	S. D.	2001.18	138.08	1.65	5.47	0.05	3.85	1.44	0.46	7.58
H402	N	111	162	167	103	115	120	157	161	168
	Mean	956.78	1066.02	1.26	6.35	0.03	4.67	9.87	8.13	7.25
	S. D.	4037.18	143.73	1.61	6.70	0.09	2.01	1.76	0.23	4.10

Measuraments units: (a) NMP/100 mL, (b) µS/cm, (c) mg/L O_2, (d) mg/L, (e) integer values from 1 to 14 and (f) mg/L MES

Table 2. Sample sizes, means, and standard deviations for all parameters at each control station (continued)

Station code		Total coliforms at 37ºC [a]	Conduc-tivity at 20ºC [b]	Biochemical oxygen demand [c]	Chemical oxygen demand [c]	Total phos-phates [d]	Nitrates [d]	Dissolved oxygen [c]	pH [e]	Suspended matter [f]
H403	N	116	150	197	95	119	112	166	161	197
	Mean	7827.48	779.28	0.78	3.48	0.04	4.65	9.13	7.89	8.92
	S. D.	21672.27	244.73	1.00	6.31	0.11	1.75	1.13	0.32	23.26
H601	N	97	133	134	129	130	127	124	131	133
	Mean	51161.04	1208.45	1.42	5.79	0.06	22.29	10.02	8.10	14.73
	S. D.	112676.77	117.24	2.45	6.55	0.10	10.14	1.60	0.15	15.46
I402	N	110	174	163	111	84	95	170	169	166
	Mean	86374.38	1238.91	9.30	21.06	1.34	17.34	8.41	8.06	64.96
	S. D.	240760.25	367.65	16.23	14.76	0.71	13.02	2.96	0.28	442.87
I502	N	125	138	148	85	82	82	135	136	151
	Mean	3285.47	877.80	0.96	10.49	0.08	3.34	11.22	8.39	24.80
	S. D.	7876.50	109.21	1.47	8.92	0.13	2.96	4.68	0.24	116.52
J601	N	130	162	163	106	135	137	154	159	166
	Mean	22516.82	1288.06	4.24	14.95	0.41	15.77	8.93	7.94	21.62
	S. D.	67528.45	700.57	4.90	8.90	0.41	10.70	3.96	0.31	48.15
K202	N	131	151	166	105	139	131	157	156	168
	Mean	8638.96	760.22	1.19	7.38	0.11	6.70	9.06	7.96	21.75
	S. D.	63797.85	119.02	1.94	7.87	0.55	4.86	1.14	0.37	27.92
K602	N	111	149	157	101	80	93	140	145	161
	Mean	462726.23	1019.87	8.66	30.85	0.50	20.97	4.52	7.59	43.07
	S. D.	994407.51	185.08	9.32	34.09	0.57	14.00	1.72	0.27	65.10
L502	N	101	146	160	107	85	103	121	129	162
	Mean	171758.16	2901.32	39.98	49.92	1.12	30.01	4.99	7.78	64.14
	S. D.	287343.76	1592.36	121.02	44.04	1.14	32.73	2.89	0.51	324.45
M004	N	122	142	186	90	78	81	175	172	186
	Mean	859.48	581.70	0.77	10.45	0.01	7.33	9.38	8.12	14.72
	S. D.	3161.27	175.56	1.47	65.07	0.07	3.31	1.19	0.19	118.77
M603	N	75	147	137	58	37	62	143	142	140
	Mean	5390963.19	1275.94	36.68	304.03	2.31	6.65	6.88	7.95	58.71
	S. D.	16085508.95	477.58	54.07	1193.22	2.08	7.85	2.26	0.27	91.86
O501	N	128	181	166	111	91	109	177	177	167
	Mean	972464.90	14842.89	20.78	57.24	3.63	23.37	5.25	7.77	828.60
	S. D.	4677620.85	4666.41	23.85	61.37	3.36	18.29	4.06	0.33	3356.13

Measuraments units: (a) NMP/100 mL, (b) □S/cm, (c) mg/L O_2, (d) mg/L, (e) integer values from 1 to 14 and (f) mg/L MES

The highest value for mean total coliforms was recorded at station M603 (Serpis River, Alicante) and the lowest at station H102 (Júcar River, Cuenca). Mean conductivity was highest at station O501 (Vinalopó River, Alicante) and lowest at station M004 (Mirón River, Albacete). The highest mean biochemical oxygen demand was at station L502 (Cáñoles River, Valencia) and lowest at station M004 (Mirón River, Albacete). The highest mean chemical oxygen demand was observed at station M603 (Serpis River, Alicante) and the lowest was at station H403 (Sot River, Valencia). The mean total phosphates reached its maximum value at station O501 (Vinalopó River, Alicante) and its minimum value at stations H102 and M004 (Júcar River, Cuenca, and Mirón River, Albacete). Mean nitrates were highest at station L502 (Cáñoles River, Valencia) and lowest at station D201 (Júcar River, Cuenca). The highest mean dissolved oxygen content was reached at station I502 (Magro River, Valencia) and the lowest at station K602 (Verde River, Valencia). The mean pH reached its maximum value at station G403 (Tuéjar River, Valencia) and its minimum at station K602 (Verde River, Valencia). Finally, mean suspended matter was highest at station O501 (Vinalopó River, Alicante) and lowest at station H102 (Júcar River, Cuenca).

To summarize, station O501 (Vinalopó River, Alicante) has the highest means in the most parameters (three parameters), whereas stations H102 (Júcar River, Cuenca) and M004 (Mirón River, Albacete) both have the smallest means in three parameters.

Thus, water quality is generally highest in Cuenca and Albacete provinces, and lowest in Valencia and Alicante provinces. This is a logical result because the two former provinces correspond to the source regions of their respective rivers, which do not pass through large urban centers, and both provinces lack large industrial parks that would increase the risk of water pollution. The stations with the worst water quality were in Valencia and Alicante provinces, where the sizes of the urban and industrial zones are clearly greater.

We first calculated *GQI* for all measurements of each of the nine water quality parameters at each of the 22 control stations. Next, we used these values to calculate the mean *GQI* for each station. We used the same data set to calculate the corresponding *PI* value for each measurement station. Table 3 shows the resulting mean *GQI* and *PI* values for each of the 22 control stations.

Table 3. Mean *GQI* and *PI* values for each control station and the relative rankings (1 = highest value, 22 = lowest value) in brackets

Station code	*GQI*	*PI*
C302	78.51 (4)	30.37 (11)
D201	79.53 (3)	42.17 (5)
E502	68.50 (15)	36.78 (9)
E503	77.36 (7)	29.54 (12)
F201	78.06 (5)	38.53 (6)
F301	73.17 (10)	19.07 (15)
F401	69.35 (14)	19.05 (16)
F702	76.51 (8)	25.97 (13)
G403	70.35 (13)	22.12 (14)
H102	71.78 (11)	35.01 (10)
H402	64.10 (17)	39.61 (6)
H403	83.35 (2)	49.44 (2)
H601	71.40 (12)	47.84 (3)
I402	62.83 (18)	3.99 (19)
I502	74.96 (9)	44.47 (4)
J601	65.70 (16)	12.79 (17)
K202	77.77 (6)	37.13 (8)
K602	60.85 (19)	12.06 (18)
L502	57.58 (21)	1.20 (22)
M004	84.39 (1)	52.72 (1)
M603	58.47 (20)	3.98 (20)
O501	43.12 (22)	2.70 (21)

Because the measurement scales for the two indices are different, direct comparison of the values would not be meaningful. Nevertheless, comparison of the relative rankings of the stations permits a valid comparison. For both indices, the stations with the worst water quality were L502 and O501, in the Cáñoles and Vinalopó rivers (respectively), and the station with the best water quality, M004, was in the Mirón River (Albacete Province).

In addition, we considered two time subperiods: the first one includes information between 1990 and 1997, and the second includes data from 1998 to 2005.

We studied the relationship between *GQI* and *PI* to determine the extent to which the two indicators measure the same thing. For all control stations combined, Table 4 shows the correlations between the mean values of *GQI* and *PI*. In both periods (1990 to 1997 and 1998 to 2005), *GQI* was strongly and positively correlated with *PI*, both for the Pearson correlation coefficient and for the Spearman correlation coefficient (included as a non-parametric measure of correlation). The Pearson's correlation coefficients ranged from 0.740 to 0.789 ($p < 0.001$), and the Spearman correlation coefficients ranged from 0.718 to 0.789 ($p < 0.001$).

Table 4. Correlations between the mean values of *GQI* and *PI* from 1990 to 2005

	Pearson's correlation		**Spearman correlation**	
	Coefficient	***p* value**	**Coefficient**	***p* value**
Entire study period	0.789	< 0.001	0.768	< 0.001
1990-1997 period	0.773	< 0.001	0.789	< 0.001
1998-2005 period	0.740	< 0.001	0.718	< 0.001

From this point onward, we will only use *PI* for posterior analysis. To do so, we tested whether the water quality in the Confederación Hidrográfica del Júcar changed over time by analyzing the *PI* values. We compared the mean values of *PI* during the two time periods. The mean *PI* value was significantly (*t*-test for independent samples; $p = 0.001$) higher from 1998 to 2005 (40.2) than from 1990 to 1997 (21.9), indicating an improvement in water quality.

To determine the sensitivity of *PI* to the omission of each of the parameters (one at a time) from calculation of the index, we calculated the value of *PI* for all stations and all years, with one parameter at a time omitted from the calculation. For each of these calculations, we then compared the mean value of the new *PI* with the mean value of the index calculated using all the parameters (Table 5). The results reveal that the mean value of *PI* did not change significantly when any one of the following parameters was excluded from the calculation: total phosphates, nitrates, dissolved oxygen, and suspended matter. In contrast, *PI* changed significantly if any of the following parameters were excluded: total coliforms, conductivity, biochemical oxygen demand, chemical oxygen demand, and pH. Moreover, omitting the latter parameters increases the mean value of the index, indicating better water quality.

Table 5. Comparison of the mean global *PI* with the corresponding index calculated after excluding individual parameters. Differences represent the global mean *PI* minus the recalculated mean *PI*. Significance is for a *t*-test with related samples

Global mean *PI*: 27.5697			
Excluded parameter	**Recalculated mean**	**Difference**	***p* value**
Total coliforms	43.40	-15.83	**< 0.001**
Conductivity	43.39	-15.82	**< 0.001**
Biochemical oxygen demand	39.20	-11.63	**< 0.001**
Chemical oxygen demand	43.40	-15.83	**< 0.001**
Total phosphates	27.15	0.42	0.525
Nitrates	26.82	0.75	0.277
Dissolved oxygen	28.60	-1.03	0.417
pH	39.28	-11.71	**< 0.001**
Suspended matter	29.62	1.18	0.096

Table 6. Control station relative rankings, calculated after omitting the individual parameters with a significative differencial behaviour

Station code	*PI*	*Total coliforms*	*Conductivity at 20ºC*	*Biochemical oxygen demand*	*Chemical oxygen demand*	*pH*
C302	11	11	11	11	13	11
D201	5	6	8	8	9	8
E502	9	5	7	7	6	7
E503	12	12	12	12	11	12
F201	7	4	6	6	7	6
F301	15	16	15	15	15	15
F401	16	15	16	16	16	16
F702	13	14	13	13	12	13
G403	14	13	14	14	14	14
H102	10	7	10	10	5	10
H402	6	9	5	5	3	5
H403	2	2	2	2	4	2
H601	3	10	3	3	1	3
I402	19	20	20	20	20	20
I502	4	1	4	4	10	4
J601	17	17	17	17	17	17
K202	8	8	9	9	8	9
K602	18	18	18	18	18	18
L502	22	22	22	22	22	22
M004	1	3	1	1	2	1
M603	20	19	19	19	19	19
O501	21	21	21	21	21	21

Table 6 shows the classification of the control stations atending to the *PI* obtained after omitting each one of the five parameters mentioned above.

Table 7 shows the correlation between the rankings of the Table 6 and the rankings with all the parameters. As we expected, they are strongly and possitively correlated in all the cases.

Table 7. Spearman correlation coefficient between the rankings of the control stations with all the 9 parameters and the rankings omitting one parameter

Rho Spearman	*Total coliforms*	*Conductivity at 20ºC*	*Biochemical oxygen demand*	*Chemical oxygen demand*	*pH*
Rho Spearman	0.937	0.990	0.990	0.937	0.990
Sig. (bilateral)	0.000	0.000	0.000	0.000	0.000

4. CONCLUSIONS

In recent years, many papers about how to measure water quality have been published. In the present study, we compared two of the proposed indicators: a general quality index (*GQI*) currently used by the hydrographic federations in Spain and the stochastic indicator (*PI*) of Beamonte et al. (2005) that was originally designed to analyze the water intended for use in the production of drinking water. In this paper, the *PI* index was adopted independently of the possible uses of the water.

To perform this comparison, we studied nine water quality parameters for which data were recorded at 22 control stations in Spain's Confederación Hidrográfica del Júcar from 1990 to 2005. The results showed that *GQI* was strongly and positively correlated with *PI* both for the

whole study period and for two subsets of the period (Pearson's correlation coefficient ranged between 0.74 and 0.79, with $p < 0.001$ in all cases). These analyses suggest that the *PI* index has high potential for use in monitoring water quality. Using this index, the mean *PI* value was significantly higher during the second period than the first one, indicating an improvement in water quality.

In addition, we assessed the sensitivity of *PI* to the omission of each of the nine water quality parameters that we used to calculate the index. The results confirm that the mean *PI* value did not change significantly if the total phosphates, nitrates, dissolved oxygen, or suspended matter were excluded from its calculation. In contrast, the mean value of the index changed significantly if the total coliforms, conductivity, biochemical oxygen demand, chemical oxygen demand, or pH were excluded from the calculation. The exclusion of the latter parameters from the calculation increases the mean *PI* value, suggesting an increase in the estimated water quality.

REFERENCES

BEAMONTE, E., BERMÚDEZ, J., CASINO, A., VERES, E. (2005). "A global stochastic index for water quality: the case of the river Turia (Spain)". *Journal of Agricultural, Biological and Environmental Statistics* 10, 424-439.

BEAMONTE, E., BERMÚDEZ, J., CASINO, A., VERES, E. (2007). "A statistical study of the quality of surface water intended for human consumption near Valencia (Spain)". *Journal of Environmental Management* 83, 307-314.

GRAÇA, M.A.S., COIMBRA, C.N. (1998). "The elaboration of indices to assess biological water quality. A case study". *Water Research* 32, 380-392.

ICAGA, Y. (2007). "Fuzzy evaluation of water quality classification". *Ecological Indicators* 7, 710-718.

KINDLER, J., ROMAN, M., NALBERCZYNSKI, A., TYSZEWSKY, S., PUSLOWSKA, D., KLOS-TREBACZKIEWICZ, H., OSUCH-PAJDZINSKA, E., GROMIEC, M. (1998). "Balancing costs and water quality in meeting EU directives (the Upper/Middle Odra case study in Poland)". *Water Policy* 1, 283-303.

MINISTERIO DE OBRAS PÚBLICAS Y URBANISMO (1983). "La vigilancia de la contaminación fluvial". Dirección General de Obras Hidráulicas, Madrid.

MOATAR, F., MIQUEL, J., POIREL, A. (2001). "A quality-control method for physical and chemical monitoring data. Application to dissolved oxygen levels in the river Loire". *Journal of Hydrology* 252, 25-36.

POCH, M. (1999). "Las calidades del agua". Rubes Editorial S.L., Barcelona.

PRAT, N., MUNNÉ, A. (2000). "Water use and quality and stream flow in a Mediterranean stream". *Water Research* 34, 3876-3881.

PROVENCHER, M., LAMONTAGNE, M.P. (1977). "Méthode de determination d'un indice d'appréciation de la qualité des eaux selon différentes utilisations". *Ministère de Richesses Naturelles*, Québec.

REISENHOFER, E., ADAMI, G., BARBIERI, P. (1998). "Using chemical and physical parameters to define the quality of karstic freshwaters (Timavo river, North-Eastern Italy): a chemometric approach". *Water Research* 32, 1193-1203.

SÁNCHEZ, E., COLMENAREJO, M.F., VICENTE, J., RUBIO, A., GARCÍA, M.G., TRAVIESO, L., BORJA, R. (2007). "Use of the water quality index and dissolved oxygen déficit as simple indicators of watersheds pollution". *Ecological Indicators* 7, 315-328.

SIMÕES, F.S., MOREIRA, A.B., BISINOTI, M.C., GIMENEZ, S.M.N., YABE, M.J.S. (2007). "Water quality index as a simple indicator of aquaculture effects on aquatic bodies". *Ecological Indicators* 8, 476-484.

LA CALIDAD GLOBAL DE LAS AGUAS SUPERFICIALES EN LA CONFEDERACIÓN HIDROGRÁFICA DEL JÚCAR: COMPORTAMIENTO DEL ÍNDICE DE CALIDAD GENERAL Y DE LOS PARÁMETROS BÁSICOS QUE LO COMPONEN

EDUARDO BEAMONTE CÓRDOBA
E-mail: beamonte@uv.es
ALEJANDRO CASINO MARTÍNEZ
E-mail: alejandro.casino@uv.es
ERNESTO J. VERES FERRER
E-mail: ernesto.veres@uv.es
Departamento de Economía Aplicada
Universitat de València

Resumen

El Índice de Calidad General (ICG) es uno de los indicadores más utilizados en España para medir la calidad o el grado de contaminación de las aguas. En este trabajo se analiza, en el ámbito de la Confederación Hidrográfica del Júcar, el comportamiento de dicho índice en función de los parámetros básicos que lo componen. El estudio se ha desarrollado a partir de los datos observados sobre un conjunto de parámetros físico-químicos, en diferentes puntos de muestreo y a lo largo de un amplio periodo de tiempo.
Los análisis efectuados han permitido clasificar la calidad de las aguas superficiales, independientemente de sus posibles utilizaciones, y detectar los cambios producidos en la evolución temporal del ICG y de sus parámetros básicos. Además, se ha estudiado el grado de relación del ICG con sus principales determinantes, tratando de identificar los parámetros que representan mejor el comportamiento del índice general.

*Palabras Clave: A*nálisis de datos, Calidad del agua, Parámetros físico-químicos, Índice de Calidad General.
Área Temática: Economía de la Energía, del Ambiente, de los Transportes y de los Recursos Naturales.

Abstract

The General Quality Index (GQI) is a common indicator used in Spain to measure the water quality. In this work the behavior of this index is analyzed in the territorial area of the Confederación Hidrográfica del Júcar (Spain), depending on its basic component parameters. The study has carried out from the information observed on a set of physical and chemical parameters, in different sampling points and along a wide time period.
The analyses allows to classify the quality of the surface waters, independently of their posterior use. Also the changes produced in the temporary evolution of the ICG and of the basic parameters have been detected. In addition, the degree of relation between the ICG and its principal determinants has been analyzed, trying to identify the parameters that better represent the behavior of the general index.

Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

1. INTRODUCCIÓN

El agua es un recurso natural limitado, indispensable para la vida y susceptible de usos sucesivos. Se trata de un recurso que debe estar disponible no sólo en la cantidad necesaria sino también con la calidad precisa.

La calidad del agua es un término complejo en el que intervienen muchos factores. Es un concepto variable que depende del uso que se vaya a hacer de ella. En la legislación española y de la Unión Europea se recogen, para los usos más importantes del agua, una serie de requisitos que se basan en la medición de diferentes parámetros.

Con objeto de integrar los valores obtenidos sobre un conjunto de parámetros en un único valor o índice global se han propuesto distintos indicadores. El denominado Índice de Calidad General (ICG) es el más utilizado en España, por las distintas confederaciones hidrográficas, para medir la calidad o el grado de contaminación de las aguas.

Algunos de los trabajos que se han centrado en la construcción y aplicación de índices globales de la calidad del agua son: Berzas et al. (2000), Beamonte et al. (2004a y 2005), Bordalo et al. (2001), Fore et al. (1995a y 1995b), Kondratyev et al. (2002), Melloul y Collin (1998), Pesce y Wunderlin (2000), Popp et al. (2001), Schultz (2001) y Villalba et al. (1995). La creciente preocupación por la adecuada medición de la calidad del agua también se pone de manifiesto en otras investigaciones de carácter empírico: Beamonte et al. (2004b y 2007), Graça y Coimbra (1998), Moatar et al. (2001), Prat y Munné (2000), Sánchez et al. (2007) y Simões et al. (2007).

En este trabajo se ha adaptado el ICG con la finalidad de evaluar la calidad de las aguas superficiales controladas por la Confederación Hidrográfica del Júcar. A partir de los datos observados sobre un conjunto de parámetros físico-químicos, en diferentes puntos de muestreo y a lo largo de un amplio periodo de tiempo, se estudia el comportamiento del citado índice en función de los parámetros básicos que lo definen.

Los análisis efectuados en el presente estudio han permitido clasificar la calidad de las aguas de la Cuenca del Júcar independientemente de sus posibles utilizaciones. Se ha comparado el comportamiento del ICG y de sus principales determinantes en dos periodos de tiempo, detectando los cambios producidos en su evolución temporal. Además, utilizando técnicas de análisis de datos de panel, se ha estudiado el grado de relación del ICG con sus parámetros básicos más importantes, identificando aquéllos que determinan en mayor medida el comportamiento del índice general.

2. EL ÍNDICE DE CALIDAD GENERAL

La evaluación de la calidad del agua requiere la toma de muestras para obtener la información necesaria sobre los parámetros de interés. Esta evaluación permite verificar si la calidad observada en el agua es adecuada para el uso que se piensa hacer de la misma.

Existen diferentes índices de calidad que permiten sintetizar la información obtenida sobre los distintos parámetros y determinar el estado general de las aguas en función de unos rangos de calidades establecidos.

El índice más utilizado en España es el ICG, que se basa en un método desarrollado por Provencher y Lamontagne (1977) del Servicio de Calidad de las Aguas del Ministerio de Riquezas Naturales del Estado de Québec (Canadá). El método enfoca el problema de forma general y define un índice de calidad para cualquier uso posterior del agua, que permite establecer estudios comparativos de la situación de las aguas con carácter abstracto e independiente de sus posibles utilizaciones.

El resultado del ICG es un número adimensional, función de un conjunto de parámetros referentes a la calidad de las aguas medidos en una muestra. Se puede expresar matemáticamente, de forma general, así:

$$ICG = \sum_{i=1}^{n} Q_i P_i ,$$

donde *n* indica el número de parámetros (o características físico-químicas) que intervienen en el cálculo del índice; Q_i representa una función de equivalencia que transforma el valor observado en la muestra del parámetro *i* en un nivel de calidad que varía entre 0 y 100 (siendo el valor 0 el nivel pésimo y 100 el óptimo); y P_i es la ponderación correspondiente al parámetro *i*-ésimo, exigiendo que la suma de todas las ponderaciones sea la unidad con la finalidad de que el índice de calidad oscile entre 0 y 100.

En el cálculo del ICG intervienen dos tipos de parámetros: *básicos* (que siempre se consideran en el cálculo del índice); y *complementarios* (que sólo intervienen en su cálculo cuando el valor de Q_i que le corresponda sea inferior a 60).

El valor final del ICG varía entre 0 (agua muy contaminada) y 100 (agua totalmente limpia). Por ejemplo, valores inferiores a 50 indican que el agua no es adecuada para ningún uso. En la Tabla 1 se presentan las categorías de calidad del agua asociadas a los posibles valores del ICG.

Tabla 1. Categorías de calidad del agua según el valor del ICG

ICG	Calidad del agua
Entre 85 y 100	(C1) Excelente
Entre 75 y 85	(C2) Buena
Entre 65 y 75	(C3) Intermedia
Entre 50 y 65	(C4) Admisible
Inferior a 50	(C5) Inadmisible

3. PARÁMETROS, ESTACIONES DE CONTROL Y DATOS

En el estudio se han considerado los 9 parámetros *básicos* que componen el ICG: *oxígeno disuelto*, *materias en suspensión*, *pH*, *conductividad*, *demanda química de oxígeno*, *demanda bioquímica de oxígeno*, *coliformes totales*, *fosfatos totales* y *nitratos*. En la Tabla 2 se recogen estos parámetros y sus unidades de medida.

Tabla 2. Parámetros básicos del ICG

Parámetro	Denominación	Unidad de medida
Oxígeno disuelto	oxidis	mg/l O_2
Materias en suspensión	solsus	mg/l MES
pH	pH	escala
Conductividad	conduc	µS/cm a 20ºC
Demanda química de oxígeno	DQO	mg/l O_2
Demanda bioquímica de oxígeno	DBO	mg/l O_2
Coliformes totales a 37ºC	coltot	NMP/100 ml
Fosfatos totales	fosfat	mg/l
Nitratos	nitrat	mg/l

Fuente: Confederación Hidrográfica del Júcar

Los datos utilizados en el trabajo corresponden a mediciones efectuadas en 22 estaciones de control dependientes de la Confederación Hidrográfica del Júcar. En la Tabla 3 aparecen recogidos los códigos de esas 22 estaciones de control, junto con su denominación, el cauce del río, el municipio y la provincia en las que se ubican. Se ha analizado la información procedente de muestras tomadas en el periodo temporal comprendido entre los años 1990 y 2005.

Tabla 3. Estaciones de control seleccionadas y su ubicación

Código	Denominación	Cauce	Municipio (Provincia)
C302	Gea de Albarracín	Río TURIA	Gea de Albarracín (Teruel)
D201	Venta de Juan Romero	Río JÚCAR	Huélamo (Cuenca)
E502	Embalse de Arenós	Río MIJARES	Montanejos (Castellón)
E503	Puebla de Arenoso	Río MIJARES	Puebla de Arenoso (Castellón)
F201	Pajaroncillo	Río CABRIEL	Pajaroncillo (Cuenca)
F301	Las Rinconadas	Río TURIA	Santa Cruz de Moya (Cuenca)
F401	Zagra	Río TURIA	Aras de Alpuente (Valencia)
F702	Villarreal-Estación aforos	Río MIJARES	Villareal (Castellón)
G403	Calles	Río TUEJAR	Calles (Valencia)
H102	Embalse de Alarcón	Río JÚCAR	Alarcón (Cuenca)
H402	Embalse de Loriguilla	Río TURIA	Chulilla (Valencia)
H403	Sot de Chera	Río SOT	Chulilla (Valencia)
H601	La Presa (Estación de alerta)	Río TURIA	Manises (Valencia)
I402	Hortunas de Abajo	Río MAGRO	Requena (Valencia)
I502	Macastre	Río MAGRO	Macastre (Valencia)
J601	Cullera	Río JÚCAR	Cullera (Valencia)
K202	Los Frailes	Río JÚCAR	Albacete (Albacete)
K602	Alcira-río Verde	Río VERDE	Alcira (Valencia)
L502	Ayacor	Río CÁÑOLES	Játiva (Valencia)
M004	Montemayor-Aforo	Río MIRÓN	Casas de Lázaro (Albacete)
M603	Alcocer de Planes	Río SERPIS	Muro de Alcoy (Alicante)
O501	Aspe-Cola del embalse de Elche	Río VINALOPÓ	Aspe (Alicante)

Fuente: Confederación Hidrográfica del Júcar

A partir de los datos disponibles se ha calculado el valor del ICG (considerando los 9 parámetros básicos que lo componen) para cada una de las muestras tomadas en las distintas estaciones de control seleccionadas.

Además, para la realización de algunos análisis, se han distinguido dos periodos dentro del rango temporal de los datos: en el primero se han considerado los datos que abarcan desde el año 1990 hasta el año 1997, y en el segundo se han incluido los datos comprendidos entre el año1998 y el año 2005.

4. RESULTADOS

4.1 Evaluación de la calidad general del agua en la Confederación Hidrográfica del Júcar

Los valores del ICG obtenidos permiten clasificar la calidad de las aguas superficiales de la Cuenca del Júcar independientemente de sus posibles utilizaciones. En la Tabla 4 se presenta, para cada estación de control seleccionada en el estudio, el valor medio del ICG en el periodo temporal completo y en cada uno de los subperiodos en que se ha dividido éste, y la oportuna clasificación de la calidad del agua.

El valor promedio del ICG obtenido para el conjunto formado por todas las estaciones seleccionadas indica que la calidad general del agua, en la Confederación Hidrográfica del Júcar, puede calificarse como *intermedia* (categoría C3), tanto si se considera el periodo de tiempo globalmente como si se atiende a cada una de sus dos partes. No obstante, este resultado debe matizarse, pues el comportamiento no es el mismo en todas las estaciones de medición.

Considerando los resultados correspondientes al periodo temporal completo, se puede destacar que la mejor valoración corresponde a la estación M004 (situada en el río Mirón, en el municipio de Casas de Lázaro, provincia de Albacete), que prácticamente alcanza la categoría C1 (agua de calidad *excelente*). El peor resultado lo obtiene la estación O501 (situada en el río Vinalopó, en el municipio de Aspe, provincia de Alicante) donde la calidad del agua se puede calificar como *inadmisible* (categoría C5), aunque se observa una mejoría con el transcurso del tiempo.

En general, los mayores valores medios del ICG se obtienen en estaciones de control ubicadas en zonas próximas a las cabeceras de distintos ríos.

Tabla 4. Valores medios del ICG y categorías de calidad, en cada estación de control y periodo de tiempo

	Media ICG (categoría calidad)		
Código estación	**Primer periodo**	**Segundo periodo**	**Periodo completo**
C302	79.60 (C2)	77.09 (C2)	78.51 (C2)
D201	78.39 (C2)	81.51 (C2)	79.53 (C2)
E502	75.88 (C2)	57.55 (C4)	68.50 (C3)
E503	74.93 (C3)	79.93 (C2)	77.36 (C2)
F201	77.91 (C2)	78.22 (C2)	78.06 (C2)
F301	70.27 (C3)	75.09 (C2)	73.17 (C3)
F401	65.21 (C3)	72.94 (C3)	69.35 (C3)
F702	74.98 (C3)	78.98 (C2)	76.51 (C2)
G403	69.01 (C3)	71.63 (C3)	70.35 (C3)
H102	75.82 (C2)	68.42 (C3)	71.78 (C3)
H402	65.92 (C3)	62.61 (C4)	64.10 (C4)
H403	81.07 (C2)	85.74 (C2)	83.35 (C2)
H601	69.47 (C3)	72.44 (C3)	71.40 (C3)
I402	57.75 (C4)	66.45 (C3)	62.83 (C4)
I502	75.77 (C2)	74.34 (C3)	74.96 (C3)
J601	65.51 (C3)	65.92 (C3)	65.70 (C3)
K202	76.20 (C2)	79.42 (C2)	77.77 (C2)
K602	57.85 (C4)	63.50 (C4)	60.85 (C4)
L502	58.09 (C4)	57.12 (C4)	57.58 (C4)
M004	84.26 (C2)	84.51 (C2)	84.39 (C2)
M603	54.45 (C4)	63.44 (C4)	58.47 (C4)
O501	35.88 (C5)	50.77 (C4)	43.12 (C5)
Todas	69.74 (C3)	70.93 (C3)	70.34 (C3)

Para detectar la existencia de cambios temporales estadísticamente significativos en el comportamiento del ICG y los parámetros básicos que lo definen, se han comparado los correspondientes valores medios en los dos subperiodos de tiempo en los que se ha dividido el periodo muestral completo. En la Tabla 5 se recogen los niveles de significación críticos (p-valores) asociados al test *t* de comparación de medias utilizado (asumiendo o no varianzas iguales en función del resultado obtenido después de aplicar la prueba de Levene para la igualdad de varianzas). Los resultados referentes a los parámetros básicos que componen el índice general se han obtenido a partir de los valores de las respectivas funciones Q_i.

Tabla 5. Comparación de medias, del ICG y cada parámetro básico, entre periodos temporales y en cada estación de control

	p-valor del test *t* para la comparación de medias									
Código estación	**ICG**	**oxidis**	**solsus**	**pH**	**conduc**	**DQO**	**DBO**	**coltot**	**fosfat**	**nitrat**
C302	0.04	0.01	0.23	0.07	0.17	0.01	0.00	0.01	0.00	0.01
D201	0.02	0.01	0.43	0.01	0.92	0.35	0.00	0.03	0.23	0.27
E502	0.00	0.34	0.13	0.03	0.10	0.05	0.00	0.77	0.15	0.21
E503	0.00	0.19	0.04	0.25	0.25	0.77	0.00	0.03	0.24	0.60
F201	0.82	0.02	0.19	0.36	0.03	0.41	0.00	0.29	0.20	0.14
F301	0.01	0.06	0.77	0.28	0.40	0.31	0.00	0.34	0.00	0.00
F401	0.00	0.83	0.33	0.97	0.05	0.03	0.00	0.62	0.01	0.00
F702	0.05	0.08	0.81	0.42	0.34	0.01	0.00	0.48	0.07	0.58
G403	0.03	0.00	0.00	0.43	0.00	0.15	0.00	0.95	0.00	0.98
H102	0.05	0.56	0.69	0.99	0.69	0.33	0.00	0.43	0.08	0.37
H402	0.44	0.06	0.70	0.00	0.02	0.00	0.00	0.00	0.00	0.00
H403	0.00	0.02	0.08	0.00	0.00	0.01	0.00	0.00	0.00	0.01
H601	0.01	0.39	0.37	0.76	0.06	0.00	0.00	0.40	0.00	0.00
I402	0.00	0.00	0.73	0.41	0.99	0.51	0.00	0.00	0.23	0.00
I502	0.38	0.12	0.22	0.02	0.00	0.14	0.00	0.00	0.03	0.01
J601	0.79	0.10	0.20	0.49	0.12	0.23	0.33	0.00	0.00	0.97
K202	0.02	0.01	0.25	0.67	0.00	0.01	0.00	0.23	0.01	0.11
K602	0.01	0.00	0.00	0.00	0.86	0.87	0.87	0.00	0.87	0.00
L502	0.74	0.33	0.01	0.00	0.00	0.01	0.00	0.03	0.73	0.00
M004	0.84	0.46	0.14	0.12	0.04	0.06	0.00	0.63	0.23	0.39
M603	0.00	0.82	0.68	0.01	0.00	0.01	0.00	0.33	0.01	0.00
O501	0.00	0.04	0.00	0.00	0.05	0.00	0.00	0.01	0.00	0.48
Todas	0.05	0.15	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00

Los resultados reflejan que, para las estaciones de medición en conjunto, los valores medios del ICG y de la práctica totalidad de los parámetros básicos son significativamente distintos en las dos partes en las que se divide el periodo temporal completo. Además, los promedios son mayores en el segundo subperiodo de tiempo, lo que supone una mejora en la calidad del agua. Únicamente los valores medios asociados al *oxígeno disuelto* y a los *coliformes totales* parecen no experimentar cambios temporales significativos.

Respecto a los resultados obtenidos en cada una de las 22 estaciones de control analizadas, en 16 de ellas se observa que el valor medio del ICG es significativamente distinto en los dos subperiodos temporales, y en 13 de esas estaciones el cambio en la calidad media del agua ha supuesto una mejora significativa. Por otra parte, es la *demanda bioquímica de oxígeno* el parámetro cuyo valor medio presenta un comportamiento temporal significativamente distinto en el mayor número de estaciones de medición.

4.2 Relación entre el ICG y sus principales parámetros básicos

Otro de los objetivos abordados en este trabajo es el estudio de la interrelación existente entre el ICG y sus parámetros básicos más importantes, identificando aquéllos que determinan en mayor medida el comportamiento del índice general.

Los parámetros *oxígeno disuelto*, *materias en suspensión*, *pH*, *conductividad*, *demanda bioquímica de oxígeno* y *coliformes totales*, son los parámetros básicos que los técnicos valoran como "muy importantes" a la hora de determinar la calidad global del agua. Son éstos los que se van a considerar en los siguientes análisis. Los parámetros básicos restantes que integran el ICG, *demanda química de oxígeno*, *fosfatos totales* y *nitratos*, son considerados técnicamente como parámetros de "importancia débil".

Para realizar el análisis, en primer lugar y en cada una de las estaciones de control seleccionadas, se ha calculado el valor medio anual del ICG y de las Q_i asociadas a los 6 parámetros básicos considerados más importantes. Además, con objeto de obtener un conjunto de datos con información completa, se ha prescindido de los datos anteriores al año 1996. De esta forma, se ha obtenido un panel constituido por los valores medios del ICG y las Q_i de sus 6 parámetros básicos principales, en cada año del periodo temporal 1996-2005 y para cada una de las 22 estaciones de control seleccionadas.

En la Tabla 6 se presentan los valores del coeficiente de correlación lineal entre las medias anuales del ICG y de las de Q_i de los principales parámetros básicos, así como los p-valores asociados al correspondiente contraste de significatividad planteado sobre el coeficiente de correlación. En la tabla se observa que el único parámetro básico importante que no parece tener una relación significativa con el ICG es el que representa a los *coliformes totales*. Los otros 5 parámetros considerados de mayor importancia en la determinación del ICG presentan una relación significativa con el índice general. Dicha relación es positiva (o directa) en los casos del *oxígeno disuelto*, las *materias en suspensión*, la *conductividad* y la *demanda bioquímica de oxígeno*, mientras que la relación es negativa (o inversa) en el caso del *pH*.

Tabla 6. Análisis de correlaciones sobre los valores medios anuales del ICG y de sus parámetros básicos más importantes

	Correlación con el ICG	
Parámetro	**Coeficiente**	**p-valor**
Oxígeno disuelto	0.602	0.000
Materias en suspensión	0.152	0.033
pH	-0.372	0.000
Conductividad	0.646	0.000
Demanda bioquímica de oxígeno	0.615	0.000
Coliformes totales a 37ºC	0.012	0.867

Además, utilizando técnicas propias del análisis de datos de panel, se ha estudiado la influencia que tienen sobre el ICG sus 6 componentes más importantes. Se ha estimado el siguiente modelo que recoge la influencia conjunta de los 6 parámetros sobre el ICG:

$$ICG_{it} = \beta_0 + \beta_1 oxidis_{it} + \beta_2 solsus_{it} + \beta_3 ph_{it} + \\ + \beta_4 conduc_{it} + \beta_5 DBO_{it} + \beta_6 coltot_{it} + \alpha_i + u_{it} \qquad i = 1,...,22\,, \quad t = 1,...,10,$$

donde los subíndices *i* y *t* hacen referencia a la estación de control de la que proviene el dato y el momento del tiempo en que se recogió, respectivamente; $\beta_0,\ldots,\ \beta_6$ son los parámetros a estimar; α_i recoge el denominado *efecto individual*, que se considera constante a lo largo del tiempo y específico para cada estación de control; y u_{it} es el término de perturbación aleatoria.

En la Tabla 7 se presentan algunos de los resultados obtenidos en la estimación del modelo. También se presentan en esa tabla resultados de estimación de los modelos simples que recogen la influencia que tiene sobre el ICG cada uno de sus parámetros básicos determinantes.

Tabla 7. Comportamiento del ICG en función de los principales parámetros básicos

Variable dependiente: *ICG*		
	Modelo múltiple	**Modelos simples**
Variable explicativa	**coeficiente (p-valor)**	**coeficiente (p-valor)**
Oxígeno disuelto	0.1624 (0.013)	0.3285(0.000)
Materias en suspensión	0.1797 (0.000)	0.1697 (0.011)
pH	0.0629 (0.729)	-0.2505 (0.020)
Conductividad	0.1892 (0.023)	0.2104 (0.000)
Demanda bioquímica de oxígeno	0.1659 (0.000)	0.1324 (0.007)
Coliformes totales a 37ºC	0.0536 (0.351)	0.0110 (0.622)
R^2 **within**	0.4056	
Test F significatividad global	60.71 (0.000)	
Test de Hausman	181.21 (0.000)	

Nota: Los resultados de estimación del modelo múltiple se han obtenido siguiendo un enfoque de efectos fijos, mediante la aplicación de la transformación intragrupos, de forma robusta e incluyendo variables ficticias temporales. En el caso de los modelos simples, algunas de las estimaciones se han obtenido siguiendo un enfoque de efectos aleatorios, aplicando el método de los Mínimos Cuadrados Generalizados.

Las estimaciones de los coeficientes asociados a las variables explicativas del modelo múltiple se han obtenido mediante la aplicación de la *transformación intragrupos*. Se ha seguido este enfoque (denominado de *efectos fijos*) ya que el resultado del test de Hausman ha conducido al rechazo de la hipótesis de que no existe correlación entre los efectos individuales y las variables explicativas. Además, en el modelo se han incluido variables ficticias temporales. Las estimaciones y los p-valores de los coeficientes asociados al *oxígeno disuelto*, las *materias en suspensión*, la *conductividad* y la *demanda bioquímica de oxígeno*, nos permiten afirmar que cada una de estas variables tiene una influencia positiva y estadísticamente significativa sobre el ICG.

No obstante, las relaciones existentes entre las variables explicativas del modelo dificulta el análisis del efecto individual que cada una de ellas tiene sobre el ICG. Por ello, se han estimado los modelos simples que tratan de captar dicha influencia. En aquéllos casos en los que, como consecuencia del resultado del test de Hausman, no se aconseja rechazar la hipótesis de incorrelación entre los efectos individuales y las variables explicativas, se han considerado las estimaciones obtenidas mediante el método de los *Mínimos Cuadrados Generalizados* (pues en esas situaciones los estimadores obtenidos con dicho método son más eficientes que los generados con la *transformación intragrupos*). Respecto a la influencia que sobre el ICG tienen el *oxígeno disuelto*, las *materias en suspensión*, la *conductividad* y la *demanda bioquímica de oxígeno*, los resultados obtenidos en la estimación de los modelos simples confirman las conclusiones derivadas de la estimación del modelo múltiple. Del análisis de los resultados de

estimación de los modelos simples, también se desprende la existencia de una relación significativa y negativa (o inversa) entre el *pH* y el ICG.

Los resultados de estimación del modelo múltiple y de los modelos simples confirman que los *coliformes totales* no presentan una relación estadísticamente significativa con el ICG.

5. CONCLUSIONES

En este trabajo se han realizado diversos análisis a partir de los valores obtenidos después de adaptar el ICG, índice muy utilizado en las cuencas hidrográficas españolas para medir la calidad de sus aguas.

Los valores del ICG calculados con los datos muestrales observados sobre un conjunto de parámetros físico-químicos, en diferentes estaciones de control y a lo largo de un amplio periodo de tiempo, han permitido clasificar la calidad global de las aguas superficiales de la Confederación Hidrográfica del Júcar.

El valor medio del ICG obtenido para el conjunto de las estaciones de medición seleccionadas indica que la calidad general del agua, en la Confederación Hidrográfica del Júcar, puede calificarse como *intermedia*. Por otra parte, los mayores valores medios de este índice general se obtienen en estaciones de control situadas en zonas próximas a las cabeceras de distintos ríos.

El análisis comparativo de los valores medios del ICG y de los parámetros básicos que lo definen, en los dos subperiodos temporales en los que se ha dividido el periodo completo, permite concluir que, para el conjunto de todas las estaciones de control, se han producido cambios temporales estadísticamente significativos en los promedios del índice general y de la mayor parte de sus parámetros básicos. Además, los promedios son mayores en el segundo subperiodo de tiempo, lo que significa que la calidad del agua ha mejorado con el transcurso del tiempo.

Respecto a los resultados obtenidos en cada una de las estaciones de control seleccionadas, se confirma que en la mayor parte de ellas se ha producido una mejora en la calidad media de sus aguas, ya que el valor medio del ICG ha aumentado significativamente en la segunda parte del periodo muestral analizado.

En el trabajo, también se ha estudiado la interrelación existente entre el ICG y sus parámetros básicos más importantes (*oxígeno disuelto*, *materias en suspensión*, *pH*, *conductividad*, *demanda bioquímica de oxígeno* y *coliformes totales*). Sobre los valores medios anuales del índice general y de sus parámetros determinantes, se ha realizado en primer lugar un análisis de correlaciones. Únicamente los *coliformes totales* no parecen tener una relación significativa con el ICG. Los otros 5 parámetros considerados de mayor importancia en la determinación del índice general presentan una relación significativa con el mismo. Dicha relación es positiva en los casos del *oxígeno disuelto*, las *materias en suspensión*, la *conductividad* y la *demanda bioquímica de oxígeno*, mientras que la relación es negativa en el caso del *pH*.

Finalmente, utilizando técnicas de análisis de datos de panel, se ha estudiado la influencia que tienen sobre el ICG sus 6 componentes más importantes, con objeto de identificar los parámetros que determinan en mayor medida el comportamiento del índice general. Se ha estimado un modelo múltiple en el que se recoge el impacto que tienen sobre el ICG esos 6 parámetros de forma conjunta y también los modelos simples que reflejan la influencia individual de cada parámetro sobre el índice general. Los resultados de estimación del modelo múltiple revelan que el *oxígeno disuelto*, las *materias en suspensión*, la *conductividad* y la *demanda bioquímica de oxígeno* tienen una influencia positiva y estadísticamente significativa sobre el ICG. El seguimiento de estos 4 parámetros podría proporcionar una idea bastante aproximada del comportamiento del índice general y, por tanto, serían unos buenos indicadores de la calidad global de las aguas superficiales. Los resultados obtenidos al estimar los modelos simples han confirmado estas mismas conclusiones y, además, han revelado la existencia de una relación significativa y negativa entre el *pH* y el ICG.

BIBLIOGRAFÍA

BEAMONTE, E., BERMÚDEZ, J., CASINO, A. Y VERES, E. (2004a). "Un indicador global para la calidad del agua. Aplicación a las aguas superficiales de la Comunidad Valenciana". *Estadística Española*, 46, 357-384.

BEAMONTE, E., BERMÚDEZ, J., CASINO, A. Y VERES, E. (2004b). "La calidad del agua en ciertas estaciones de control del canal Júcar-Turia (periodo 1994-2001)". *Revista Española de Estudios Agrosociales y Pesqueros*, 201, 105-126.

BEAMONTE, E., BERMÚDEZ, J., CASINO, A. Y VERES, E. (2005). "A global stochastic index for water quality: the case of the river Turia (Spain)". *Journal of Agricultural, Biological and Environmental Statistics,* 10, 424-439.

BEAMONTE, E., BERMÚDEZ, J., CASINO, A. Y VERES, E. (2007). "A statistical study of the quality of surface water intended for human consumption near Valencia (Spain)". *Journal of Environmental Management*, 83, 307-314.

BERZAS, J.J., GARCÍA, L.F., RODRÍGUEZ, R.C. Y MARTÍN-ÁLVAREZ, P.J. (2000). "Evolution of the water quality of a managed natural wetland: Tablas de Daimiel National Park". *Water Research*, 34, 3161-3170.

BORDALO, A. A., NILSUMRANCHIT, W. AND CHALERMWAT, K. (2001). "Water quality and uses of the Bangpakong River (Eastern Thailand)". *Water Research*, 15, 3635-3642.

FORE, S. A., GUTTMAN, S. I., BAILER, A. J., ALTFATER, D. J. AND COUNTS, B. V., (1995a). "Exploratory Analysis of Population Genetic Assessment as a Water Quality Indicator I". *Ecotoxicology and Environmental Safety*, 30 (1), 24-35.

FORE, S. A., GUTTMAN, S. I., BAILER, A. J., ALTFATER, D. J. AND COUNTS, B. V. (1995b). "Exploratory Analysis of Population Genetic Assessment as a Water Quality Indicator II". *Ecotoxicology and Environmental Safety*, 30 (1), 36-46.

GRAÇA, M. A. S. Y COIMBRA, C. N. (1998). "The elaboration of indices to assess biological water quality. A case study". *Water Research*, 32, 380-392.

KONDRATYEV, S., GRONSKAYA, T., IGNATIEVA, N., BLINOVA, I., TELESH, I. AND YEFREMOVA, L. (2002). "Assessment of present state of water resources of Lake Ladoga and its drainage basin using Sustainable Development indicators". *Ecological Indicators*, 2, 79-92.

MELLOUL, A. J. AND COLLIN, M. (1998). "A proposed index for aquifer water-quality assessment: the case of Israel's Sharon region". *Journal of Environmental Management*, 54, 131-142.

MOATAR, F., MIQUEL, J. Y POIREL, A. (2001). "A quality- control method for physical and chemical monitoring data. Application to dissolved oxygen levels in the river Loire". *Journal of Hydrology*, 252, 25-36.

PESCE, S. F. AND WUNDERLIN, D. A. (2000). "Use of water quality indices to verify the impact of Córdoba City (Argentina) on Suquía River". *Water Research*, 34 (11), 2915-2926.

POPP, J., HOAG, D. AND HYATT, D. E. (2001). "Sustainability indices with multiple objectives". *Ecological Indicators*, 1, 37-47.

PRAT, N. Y MUNNÉ, A. (2000). "Water use and quality and stream flow in a Mediterranean stream". *Water Research*, 34, 3876-3881.

PROVENCHER, M. Y LAMONTAGNE, M.P. (1977). Méthode de determination d'un indice d'appréciation de la qualité des eaux selon différentes utilisations. *Ministère de Richesses Naturelles,* Québec.

SÁNCHEZ, E., COLMENAREJO, M.F., VICENTE, J., RUBIO, A., GARCÍA, M.G., TRAVIESO, L., BORJA, R. (2007). "Use of the water quality index and dissolved oxygen déficit as simple indicators of watersheds pollution". *Ecological Indicators* 7, 315-328.

SCHULTZ, M. T. (2001). "A critique of EPA's index of watershed indicators". *Journal of Environmental Management,* 62, 429-442.

SIMÕES, F.S., MOREIRA, A.B., BISINOTI, M.C., GIMENEZ, S.M.N., YABE, M.J.S. (2007). "Water quality index as a simple indicator of aquaculture effects on aquatic bodies". *Ecological Indicators* 8, 476-484.

VILLALBA, M., BARRAGÁN, F. J., TERNERO, M. AND JIMÉNEZ, J. C. (1995). "Hydrochemical study of an aquifer system in an agricultural area in South West Spain". *Water Research*, 29 (5), 1361-1372.

ADAPTATION AND MITIGATION POLICIES IN THE AGRICULTURAL SECTOR OF ARAGON (SPAIN)

ELENA CALVO
E-mail: ecalvo@unizar.es
Department of Economic Analysis
University of Zaragoza

JOSÉ ALBIAC
E-mail: maella@unizar.es
Department of Agricultural Economics
Agrifood Research and Technology Center
Gobierno de Aragon,

Abstract

The paper examines adaptation and mitigation measures for the primary and water resources sectors in the state of Aragon in Spain. Most of the land in Aragon, 3.7 million ha over 4.8 million, is used in agricultural and forest activities: land cultivation, livestock production and forestry. These productive activities, with significant irrigation acreage, have an important influence on ecosystem conservation and the provision of environmental services.

The loss of economic and social importance experienced by the primary sector in developed countries, and the abandonment of activities and depopulation of rural areas, has motivated a re-evaluation of the primary sector and its relationships with the ecosystems and the provision of environmental services. The environmental services provided by ecosystems in rural areas have been degrading progressively, especially after the development of intensive production processes in agricultural and livestock activities during last century. The reason for this degradation is that markets and public policies ignore the value of environmental services and ecosystems. They are externalities outside the markets, which have been unprotected by the decline of the traditional rural society of previous centuries.

The pressure on ecosystems from the growth in population and economic activities is increasing, and the environmental sustainability seems to be more and more difficult to achieve. Some authors think that consumption has to be reduced to avoid ecological disaster, while others think that global pubic policies and market institutions will solve the problems. In any case, the effort of mitigation and adaptation to climate change will be substantial (damages are evaluated at 20% GDP in the Stern report, with mitigation expenses at 1%). The problem is not only the political will to make these mitigation expenses, but the requirement of scientific knowledge and information in order to design and implement reasonable measures.

The implementation of measures such as the emissions tax proposed by the Stern report is technically, although not politically, quite simple in the energy sector, where the tax incorporates the environmental damage of emissions (once the damage is known). However in the nonpoint emissions sectors, such as the primary sector, the design and implementation of measures is much more complicated. The reason is that nonpoint emissions can not be controlled at reasonable cost, and there is not a market to introduce any tax. Control instruments for nonpoint emissions are much more sophisticated and complex, and require knowledge on biophysical processes at local scale. Furthermore, the design of incentives should take account of the strategic behavior of agents, in order to give rise to cooperation among them leading to collective action. The behavior of both developed and developing countries towards climate change, plainly shows the difficulties to achieve cooperation and collective action among countries in abating GHG emissions.

Agricultural nonpoint emissions in Aragon are important. Aragon has an extension of 4.8 million ha, of which 3.7 are used in agricultural and forest activities. Half of this acreage is cultivated land (1.9 million ha), and the other half meadows (0.4 million) and forests (1.3 million). The development of intensive agriculture and livestock activities during last century is responsible for the collapse of natural cycles of nutrients and the emission of pollution loads. Previously, these cycles were integrated with the natural environment.

The emissions of GHGs in the agricultural sector in Aragon are close to 5 million tCO_2 equivalent, which is above 20 percent of total emissions in Aragon. This percentage is well above the percentage of agricultural emissions for the whole Spain, which is 11 percent. The main GHG emissions from the agricultural sector in Aragon are those of methane and nitrous oxide from livestock (2.8 million tCO_2), and then follow the emissions from soil fertilization (1.8 million). The GHG emissions of swine manure (methane & nitrous oxide) amount to 1.9 million tCO_2 equivalent, from the large swine herd close to 5 million heads. The emissions from the enteric fermentation of cattle and sheep are 0.9 million tCO_2 equivalent. The emissions of nitrogen oxide from crop cultivation are generated by mineral and organic fertilization, used in the 400,000 ha of irrigated land and the 900,000 ha of dryland. This fertilizer contains 112,000 t of nitrogen, and producing nitrogen oxide emission loads of 1.8 million tCO_2 equivalent. Forest in Aragon are removing around 0.45 million tC per year, and the carbon stock in forest soils could be around 120 million tons.
The paper presents recommendations on adaptation and mitigation measures for the agricultural and livestock sectors. They are spatially designed by county and for the whole of Aragon, and coupled with indicators to follow their evolution. For water resources, adaptation measures are also defined, taking into account the current institutions where cooperation and collective action by stakeholders are taking place, and the domestic and European water policies.

Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

Resumen

El trabajo examina una serie de medidas de adaptación y mitigación para el sector agrario y recursos hídricos en Aragón (España). La mayoría de las tierras de Aragón, 3,7 millones de hectáreas sobre 4,8 millones, se utilizan en actividades agrícolas y forestales: tierra de cultivo, producción ganadera y forestal. Estas actividades productivas, con una importante superficie de riego, tienen una importante influencia en la conservación de los ecosistemas y en la provisión de servicios ambientales.
La pérdida de importancia económica y social experimentada por el sector primario en los países desarrollados, y el abandono de las actividades y la despoblación de las zonas rurales, ha motivado una reevaluación del sector primario y sus relaciones con los ecosistemas y la provisión de servicios ambientales. Algunos autores piensan que el consumo tiene que ser reducido para evitar el desastre ecológico, mientras que otros piensan que las políticas públicas globales y las instituciones van a resolver los problemas. En cualquier caso, el esfuerzo de mitigación y adaptación al cambio climático, será muy importante (los daños se evalúan en un 20% del PIB en el informe Stern, con unos gastos de mitigación del 1%). El problema no es sólo la voluntad política de hacer estos gastos de mitigación, sino la exigencia de información y conocimiento científico con el fin de diseñar y aplicar las medidas razonables. La aplicación de medidas tales como impuestos sobre las emisiones, propuesta por el informe Stern es técnicamente, aunque no políticamente, muy sencillo en el sector de la energía, donde se incluyen los daños al medio ambiente de las emisiones (una vez que el daño es conocido). Sin embargo, en los sectores de emisiones no puntuales, tales como el sector primario, el diseño y la aplicación de las medidas es mucho más complicado. La razón es que en las emisiones no puntuales no se puede controlar, a un coste razonable, y no hay un mercado para introducir medidas fiscales. Los instrumentos de control de las emisiones no puntuales son mucho más sofisticados y complejos, y requieren de conocimientos sobre los procesos biofísicos a escala local. Además, el diseño de incentivos debe tener en cuenta el comportamiento estratégico de los agentes, a fin de dar lugar a la cooperación entre ellos. El comportamiento de los países desarrollados en el desarrollo hacia el cambio climático, muestra claramente las dificultades para lograr la cooperación y la acción colectiva entre los países para reducir las emisiones de GEI.
Las emisiones de gases de efecto invernadero en el sector agrícola en Aragón se aproximan a los 5 millones de tCO_2 equivalente, que representan el 20 por ciento del total de las
emisiones en Aragón. Este porcentaje es superior al porcentaje de las emisiones agrícolas de toda España, aproximadamente un 11 por ciento. Las principales emisiones de GEI del sector agrícola en Aragón son las de metano y óxido nitroso de la ganadería (2,8 millones de tCO_2), y, a continuación, siguen las emisiones procedentes de la fertilización del suelo (1,8 millones). Las emisiones de GEI de estiércol porcino (metano y óxido nitroso) ascienden a 1,9 millones de tCO_2 equivalente, derivadas de la cabaña porcina, cerca de 5 millones de cabezas. Las emisiones de la fermentación entérica del ganado vacuno y ovino son 0,9 millones de tCO_2 equivalente. Las emisiones de óxido de nitrógeno de los cultivos generadas en la 400.000 hectáreas de tierras de regadío y 900.000 hectáreas de las tierras secas. Este abono contiene 112.000 t de nitrógeno, óxido de nitrógeno y la producción de las cargas de emisión de 1,8 millones de tCO_2 equivalente. Los bosques de Aragón eliminan alrededor de 0,45 millones de tC por año, y las reservas de carbono en los suelos de los bosques podrían situarse alrededor de 120 millones de toneladas.
El trabajo recoge una serie de recomendaciones sobre medidas de adaptación y mitigación para los sectores agrícola y ganadero de Aragón, junto con indicadores de seguimiento de su evolución. De los recursos hídricos, se definen las medidas de adaptación, teniendo en cuenta las actuales instituciones, la cooperación y la acción colectiva que por parte de los interesados se están llevando a cabo, y las políticas de aguas nacionales y europeas.

Área Temática: Economía de la Energía, del Ambiente, de los Transportes y de los Recursos Naturales.

INTRODUCCIÓN

Según el cuarto informe del Panel Intergubernamental sobre el Cambio Climático (IPCC 2007a), se estima un aumento de la temperatura global media de entre 1,4 a 5,8 °C en 2100 en relación a 1990, y un aumento en el nivel del mar global medio de 0,09 a 0,88 m en 2100. Los sucesos extremos y anomalías aumentarán, en forma de períodos de sequía, inundaciones, granizo, lluvias torrenciales, y cambios en la estacionalidad de la distribución temporal de las lluvias. Estos sucesos extremos y anomalías tendrán un efecto negativo sobre la disponibilidad de recursos hídricos, y causarán daños y pérdidas económicas en las infraestructuras urbanas e industriales, en las actividades agrarias, y en los biotopos y ecosistemas.

Uno de los objetivos del IPCC es evaluar el riesgo del cambio climático originado por las actividades humanas. En la línea que propone el IPCC, en este trabajo se proponen una serie de medidas para contrarrestar y difundir los efectos producidos por el cambio climático.

En el trabajo, se hace un esfuerzo inicial para exponer el estado actual del sector primario en Aragón y su influencia medioambiental, así como los cambios en la gestión de dicho sector que sirvan para contrarrestar los impactos negativos del cambio climático.

Las actividades agrícolas, ganaderas y forestales durante los últimos miles de años han conformado el capital natural del territorio. Los ecosistemas que integran este capital natural corren un peligro cierto de degradación en el caso de que no se mantenga una masa crítica de actividades tradicionales (no intensivas) beneficiosas para el medio ambiente. La desaparición de estas actividades tradicionales tendrá costes económicos, sociales y medioambientales elevados.

El enfoque multifuncional de las actividades agrícolas, ganaderas y forestales en el sector primario, en el que se combina la producción de alimentos con la protección de los ecosistemas y de los servicios medioambientales cobra ahora especial relevancia. Para ello es fundamental mantener las actividades rurales tradicionales no intensivas, y además conviene que estas actividades estén al cargo de la población agraria residual, porque su realización mediante empresas de servicios tendría un coste elevado, quizá prohibitivo, para la sociedad.

El esfuerzo de mitigación y adaptación al cambio climático de las sociedades va a ser grande. El Informe Stern (Stern 2007) estima que los efectos económicos del cambio climático supondrán una caída del 20 por cien del PIB mundial, y recomienda unos gastos de mitigación del 1 por cien del PIB. Pero el problema no es solo que exista voluntad política para realizar esos gastos de mitigación, sino que es necesario disponer de conocimientos científicos e información para diseñar las medidas apropiadas e implementarlas.

La implementación de medidas como el impuesto a las emisiones que propone el Informe Stern es técnicamente (aunque no políticamente) de una relativa sencillez en sectores como el energético, donde el impuesto incorpora el daño medioambiental de las emisiones, una vez que se conoce el daño. Sin embargo, en los sectores difusos como el sector primario el diseño e implementación de las medidas es mucho más complicado. La razón es que las emisiones contaminantes difusas no pueden controlarse a un coste razonable, ni existe un mercado en el que introducir el impuesto. Los instrumentos de control son mucho más sofisticados y complejos, y requieren generar conocimiento de los procesos biofísicos a escala local. Además los incentivos tienen que tener en cuenta el comportamiento estratégico de los agentes, para que pueda surgir la cooperación entre ellos que lleve a la acción colectiva. El comportamiento tanto de los países desarrollados como el desarrollo ante el cambio climático, muestra a las claras las dificultades para conseguir la cooperación y la acción colectiva en el control de las emisiones.

En España la reducción de precipitaciones será superior al 20%, según el escenario A1B del Summary for Policy Makers WG1 en el horizonte 2100 (IPCC 2007b), con una intensificación de la frecuencia y severidad de los fenómenos extremos de sequías e inundaciones, una fuerte caída en la disponibilidad de los recursos hídricos que puede llegar hasta el 50% en las zonas de

mayor aridez, un empeoramiento de la calidad del agua, y una disminución en las reservas tanto de los pantanos como de los acuíferos. Por otra parte los suelos perderán humedad con efectos en la vegetación, que supondrán una mayor vulnerabilidad de los bosques y cultivos ante plagas y enfermedades, y cambios en la distribución espacial. Los ecosistemas acuáticos experimentarán una progresiva degradación, en especial en las zonas de mayor aridez. La reducción de la disponibilidad de agua implicará utilizar más agua en los cultivos de regadío, y disminuirá el rendimiento de los cultivos en secano. Otros efectos serán la caída del turismo en la época estival, el aumento de los riesgos para la salud por las olas de calor y los nuevos vectores de enfermedades, y una mayor probabilidad de incendios.

La creciente preocupación sobre el cambio climático ha impulsado el interés por el potencial que tiene la vegetación y los bosques para reducir la concentración atmosférica de dióxido de carbono, gas que en términos absolutos genera la mayor contribución al calentamiento global.

El propósito de este trabajo es examinar los impactos del cambio climático sobre los recursos hídricos, la agricultura, la ganadería y los bosques en Aragón. También se analiza la relación entre el cambio climático y las políticas agrícolas, ganaderas y de bosques. Finalmente se presentan una serie de medidas para la mitigación del cambio climático y la adaptación a sus efectos.

EL SECTOR PRIMARIO DE ARAGÓN

El área de estudio está situada en el noreste de España. El territorio de Aragón abarca una extensión de 4,8 millones ha, y los sectores agrario y forestal ocupan 3,7 millones ha de este territorio. Estas actividades productivas del sector primario tienen una repercusión importante en la conservación de los ecosistemas y en la generación de servicios medioambientales.

La mitad de la superficie son tierras de cultivo (1,9 millones ha), y la otra mitad praderas (0,4 millones ha) y montes (1,3 millones ha). El sector agrario de Aragón representa el 5% del valor de la producción y el 7% de la población ocupada, y está ligado a una importante industria agroalimentaria que es la segunda rama industrial de la región con un 14% de la producción y un 12% del empleo industrial. La producción agraria alcanza los 2.600 millones de euros, de los que 1.000 los aportan las actividades de cultivo y 1.600 las actividades ganaderas. Los grupos de producciones con mayores ingresos son cereales (500 millones €), frutales (300) y forrajeras (130) en cultivos, y porcino (700), bovino carne (350) y ovino (130) en la ganadería.

La especialización productiva por provincias es cebada, arroz, alfalfa y frutales en Huesca; maíz, hortalizas y viña en Zaragoza; y olivar en Teruel. Las principales actividades ganaderas son porcino y bovino de carne, que se concentran en Huesca, ovino en Zaragoza y Teruel, y pollo y huevos en Zaragoza y Huesca.

La superficie de bosques en Aragón alcanza 1,3 millones de hectáreas. El valor económico de mercado de las actividades forestales es muy pequeño, entre 10 y 20 millones €, y no alcanza el 1 por cien de la producción del sector primario, aunque su importancia local para algunas poblaciones es significativa y sirve para sostener el uso tradicional del bosque. Las actividades que generan ingresos de mercado son madera, leña, caza y setas, mientras que otros servicios como la contribución al ciclo del agua, recreativos o paisaje están fuera del mercado y no generan ingresos. El valor económico de mercado y no mercado de los servicios medioambientales que proporcionan los bosques es la suma de los valores de uso directo, uso indirecto, opción y no uso. El valor económico de no mercado de los servicios medioambientales que proporcionan los bosques en Aragón, es muy probable que supere en al menos uno o dos ordenes de magnitud al valor económico de mercado de los bosques señalado más arriba.

Los recursos maderables se localizan en las montañas turolenses (Albarracín, Gúdar y Maestrazgo), el prepirineo y la sierra del Moncayo. Otras masas forestales son de escasa producción, difícil orografía, o repoblaciones que no han llegado a la madurez. La producción de madera alcanza los 180.000 m3 con un valor de 7 millones de euros y se concentra en Teruel y Huesca, mientras que la producción de leña alcanza 47.000 t y se concentra en Zaragoza. El

surgimiento de la agricultura y ganadería intensivas durante el último medio siglo en nuestra región es responsable del colapso de los ciclos naturales de nutrientes y la contaminación por la carga de residuos. Estos ciclos estaban más integrados con la naturaleza en la agricultura y ganadería tradicionales no intensivas de épocas anteriores.

Las principales emisiones de gases de efecto invernadero del sector agrario en Aragón provienen en primer lugar de las emisiones de metano y óxido de nitrógeno de la ganadería (2,8 millones tCO_2), y en segundo lugar de las emisiones de óxido de nitrógeno del abonado de suelos (1,8 millones). Este orden es el mismo que se observa para el conjunto de España, donde las emisiones de la ganadería (26 millones tCO_2) superan a las emisiones del abonado de suelos (21 millones tCO_2). Ahora bien, las emisiones de la ganadería en Aragón tienen una mayor importancia relativa.

Las emisiones GEI de la ganadería están formadas por metano y óxido de nitrógeno del estiércol del ganado porcino, y por metano de fermentación entérica del ganado ovino y vacuno. En Aragón se acercan a los 5 millones tCO_2 equivalente, lo que supone más del 20% de las emisiones de todos los sectores en Aragón. Este porcentaje es muy superior al 11% que representan los 47 millones tCO_2 de emisiones del sector agrario en España.

A diferencia de lo que ocurre a nivel nacional, las emisiones GEI del estiércol de la cabaña porcina superan ampliamente a las emisiones de metano de fermentación entérica del vacuno y ovino. Esto se explica porque la cabaña de porcino es enorme y alcanza las 4,5 millones de cabezas, mientras que la importancia de la cabaña de vacuno de carne y de leche es mucho menor en Aragón.

Las emisiones GEI del estiércol porcino alcanzan los 1,9 millones tCO_2 equivalente, de los que 1,4 millones corresponden a metano y 0,5 millones al óxido de nitrógeno. Las comarcas (MAPA) de Ejea, Bajo Aragón, Zaragoza y La Litera concentran un millón de toneladas de emisiones, con 700.000 toneladas de metano y 250.000 de óxido de nitrógeno (Cuadro 1). Estas enormes emisiones son consecuencia de la fuerte expansión de la cabaña de ganado porcino en las últimas décadas, que ha pasado de 1,78 millones de cabezas en 1990, a 3,5 millones en 2000, y a 4,5 millones en 2006.

Cuadro 1. Emisiones de metano del porcino en Aragón (2004)

	Cabezas porcino	Emisiones CH_4 estiércol porcino (tCO_2)
Ribargorza	151.648	47.724
Hoya Huesca	350.176	110.200
Somontano	142.965	44.991
Monegros	294.788	92.770
Litera	503.567	158.473
Bajo Cinca	383.223	120.600
Jiloca	205.526	64.679
Bajo Aragón	530.144	166.836
Hoya Teruel	37.472	11.792
Maestrazgo	39.636	12.473
Ejea	675.820	212.681
Borja	204.128	64.239
Almunia	114.653	36.081
Zaragoza	507.126	159.593
Daroca	35.267	11.099
Caspe	272.146	85.644
Total Aragón	4.528.947	1.425.260
España	25.232.000	7.940.510

Parece evidente que es necesario frenar la expansión de la cabaña de porcino, para limitar estas extraordinarias emisiones de metano y óxido de nitrógeno que suponen el 13% del total de emisiones GEI de Aragón. Pero otra razón importante que apoya la reducción de la cabaña de

porcino, es que la mayor parte del valor añadido de esta enorme cabaña se obtiene en Cataluña y no en Aragón.

Los ingresos de la carne de porcino que se obtiene de la cabaña de Aragón alcanzan los 800 millones de euros, de los que 500 millones van a Cataluña y 300 millones a Aragón. Esto se explica porque de las 580.000 t de carne que produce la cabaña de Aragón, los mataderos de Cataluña sacrifican 365.000 t mientras que los mataderos de Aragón solo sacrifican 215.000 t. Además, la mayor parte de esta producción de carne de los mataderos de Aragón se exporta a otras regiones, puesto que el consumo de carne de porcino en Aragón no supera las 26.000 t.

En cuanto al metano de fermentación entérica, las emisiones están compuestas por 0,5 millones tCO_2 equivalente del ganado ovino, y 0,4 millones tCO_2 equivalente del ganado vacuno. Las mayores emisiones de metano de fermentación entérica se generan en comarcas (MAPA) como Zaragoza, Bajo Aragón, Teruel y Ejea donde se concentra el ovino, o en comarcas como La Litera, Zaragoza, Bajo Cinca y Teruel donde se concentra el vacuno de carne (Cuadro 2).

Cuadro 2. Metano de fermentación entérica de vacuno y ovino (2004).

	Cabezas vacuno carne	Cabezas vacuno leche	Emisiones CH_4 vacuno (tCO_2)	Cabezas ovino	Emisiones CH_4 ovino (tCO_2)
Jacetania	5.662	1.166	8.696	109.104	19.748
Ribargorza	14.000	2.883	21.502	112.953	20.444
Hoya Huesca	14.479	2.981	22.236	135.736	24.568
Somontano	10.820	2.228	16.617	55.593	10.062
Monegros	9.096	1.873	13.970	72.672	13.154
Litera	72.214	14.868	110.902	59.253	10.725
Bajo Cinca	23.225	4.782	35.668	74.472	13.479
Jiloca	6.866	3.082	13.759	197.149	35.684
S. Montalbán	1.023	460	2.052	148.570	26.891
Bajo Aragón	15.423	6.922	30.905	275.989	49.954
Hoya Teruel	1.213	544	2.430	252.140	45.637
Maestrazgo	6.312	2.833	12.649	129.794	23.493
Ejea	6.060	2.092	10.934	224.311	40.600
Almunia	2.459	849	4.437	124.652	22.562
Zaragoza	28.529	9.849	51.474	285.700	51.712
Total Aragón	237.390	63.117	392.014	2.756.000	498.836
España	5.532.000	1.069.000	8.360.911	22.757.000	4.119.017

Las emisiones de óxido de nitrógeno de los cultivos son generadas por el abonado mineral y orgánico, que se utiliza en la fertilización de las 400.000 ha de regadío y las 900.000 ha de secano. Este abonado contiene 112.000 t de nitrógeno, y suponen una carga de emisiones de óxido de nitrógeno que equivalen a 1,8 millones tCO_2. La mitad de estas emisiones se generan en las comarcas de Monegros, Cinco Villas, Zaragoza, Hoya de Huesca y Bajo Cinca (Cuadro 3).

Cuadro 3. Superficies de cultivo, abonado y emisiones óxido nitrógeno del abonado (2002).

	Superficie secano (ha)	Superficie regadío (ha)	Abonado N secano (t N)	Abonado N regadío (t N)	Emisiones óxido nitrógeno (t CO_2)
Cinco Villas	56.700	58.910	3.030	11.810	234.460
Hoya de Huesca	72.350	20.510	3.870	4.050	125.120
Somontano	24.330	14.440	1.300	2.620	61.940
Cinca Medio	4.200	33.530	230	4.880	80.640
La Litera	11.600	26.040	620	3.840	70.470
Monegros	52.000	69.830	2.780	13.260	253.400
Bajo Cinca	30.270	29.060	1.620	3.960	88.140
Campo Borja	16.780	9.690	900	1.440	36.930
Ribera Alta Ebro	7.850	14.150	420	2.370	44.070
Valdejalón	26.420	14.720	1.410	1.800	50.750
Zaragoza	64.480	29.570	3.450	5.080	134.720
Ribera Baja Ebro	35.250	11.450	1.890	1.770	57.750
Bajo Aragón-Cas.	18.970	8.490	1.020	1.200	34.990
Calatayud	41.100	4.400	2.200	720	46.110
Campo Cariñena	25.230	5.500	1.350	700	32.380
Campo Belchite	44.230	2.720	2.370	410	43.850
Bajo Martín	31.680	4.300	1.700	770	38.940
Campo Daroca	37.580		2.010		31.770
Jiloca	60.560	4.630	3.240	950	66.190
Bajo Aragón	33.390	6.910	1.790	1.310	48.900
Teruel	65.560	6.810	3.510	1.250	75.160
Total Aragón	881.530	379.270	47.160	64.900	1.770.130
España	11.450.000	3.300.000			21.040.000

Las emisiones del sector agrario han seguido aumentando en los últimos años por el crecimiento de las cabañas porcina y vacuna, a pesar de la reducción de efectivos del ovino y de las menores emisiones de las actividades de cultivo. Las emisiones de óxido de nitrógeno de las actividades de cultivo están cayendo por la reducción de la superficie de cultivo en Aragón en los últimos años, y porque además se detecta una contención en la intensidad del uso de fertilizantes por unidad de superficie. Por el contrario, la expansión en un millón de cabezas de la cabaña de porcino entre 2000 y 2005, ha supuesto una carga adicional de 420.000 tCO_2 por las emisiones de metano (315.000 t) y óxido de nitrógeno (105.000 t) del estiércol.

En cuanto al papel de los bosques, realizan una fijación del carbono atmosférico que contribuye a contrarrestar las emisiones de carbono de otros sectores de la economía. El crecimiento maderero medio de los bosques de Aragón es de 1,40 m^3/ha, y el crecimiento medio es 1,87 m^3/ha en Huesca, 1,25 m^3/ha en Teruel y 0,84 m^3/ha en Zaragoza. Para un contenido de carbono de unos 275 kgC/m^3, la fijación media de carbono en cada provincia es 515 kgC/ha en Huesca, 342 kgC/ha en Teruel y 231 kgC/ha en Zaragoza. La fijación de carbono por crecimiento de la madera en Aragón alcanza las 455.000 tC, con 236,000 toneladas en Huesca, 157.000 toneladas en Teruel y 62.000 toneladas en Zaragoza. Si esta fijación de carbono que realizan los bosques se valora a 10 euros/tC, el valor de 4,6 millones de euros que se obtiene se acerca al valor de la producción maderera de la región (7 millones €).

Otro aspecto importante de la fijación de carbono por los bosques es el hecho de que la mayoría del carbono almacenado en un bosque se encuentra en el suelo y no en la vegetación. Los pocos trabajos disponibles (Sampson 1992, Matthews 1993) indican que tras la reforestación, el contenido de carbono en el suelo puede alcanzar las 70 tC/ha tras doscientos años, y hasta 100 tC/ha posteriormente. En comparación, un suelo con cultivos intensivos posee un contenido de carbono de unas 30 tC/ha. En consecuencia, la reforestación de tierras de cultivo puede servir para aumentar el stock de carbono en el suelo en 70 tC/ha. Por el contrario, el abandono de los bosques puede tener como consecuencia su degradación por incendios, y la pérdida del carbono fijado en el suelo. El stock de carbono en el suelo de los bosques de Aragón puede estimarse en unas 120 millones de toneladas.

Finalmente, el efecto en materia hídrica en la cuenca del Ebro es una disminución del caudal en los últimos decenios, que se explica por el aumento de superficie de regadío de

400.000 a 800.000 ha durante los últimos cincuenta años, y también por el aumento de la revegetación y los bosques como consecuencia del abandono de la agricultura. A partir de los informes y modelos del Tercer Informe de Evaluación del IPCC, la Confederación Hidrográfica del Ebro estima que para 2050 la reducción de caudales puede estar en un rango de entre un 5 y 10 por cien.

En cuanto a los Recursos Hídricos en Aragón, la demanda más importante es la de producción hidroeléctrica.[1] La demanda consuntiva de recursos hídricos en Aragón se acerca a los 3.000 hm^3, de los que 2.800 hm^3 se utilizan en el regadío y 180 hm^3 en la demanda de los hogares y las industrias. Las redes urbanas suministran 100 hm^3 a los hogares y 30 hm^3 a las industrias conectadas, mientras que las captaciones directas de las industrias suponen 50 hm^3. Solo la red urbana de la ciudad de Zaragoza ya suministra 70 hm^3, que representa la mitad del suministro urbano de todo Aragón.

Estas cifras muestran que la principal demanda consuntiva de recursos hídricos en Aragón es la de regadío, que absorbe más del 90% del uso consuntivo total. El cuadro 4 muestra que el mayor uso de agua está localizado en las comarcas de los grandes sistemas de riego de Bardenas (Cinco Villas, 350 hm^3), Monegros (Monegros, 540 hm^3), Cinca (Cinca Medio y Bajo Cinca, 220 y 170 hm^3), Canal de Aragón y Cataluña (La Litera, 170 hm^3), y Canal Imperial (280 hm^3).

Cuadro 4. Uso de agua por comarca y sistema de riego (hm^3).

Comarca	Uso de agua (hm^3)			
	Total	Inundación	Aspersión	Goteo
Cinco Villas	346	317	26	3
Hoya de Huesca	133	94	39	0
Somontano de Barbastro	64	25	36	3
Cinca Medio	218	142	72	4
La Litera	170	64	100	6
Monegros	544	333	210	1
Bajo Cinca	194	80	81	33
Campo de Borja	64	62	0	2
Ribera Alta del Ebro	107	101	5	1
Jalón Medio	74	40	11	23
Zaragoza	275	254	13	8
Ribera Baja del Ebro	85	53	30	2
Caspe	52	40	6	6
Calatayud	33	29	1	3
Bajo Aragón	44	42	1	1
Total Aragón	2.526	1.763	656	107

Para determinar los efectos del cambio climático sobre la disponibilidad de agua en Aragón, es necesario llevar a cabo trabajos de investigación que permitan estimar predicciones regionales fiables de variables clave del cambio climático, como temperatura, precipitación, radiación solar, humedad relativa, evapotranspiración, viento y sequías. También se requiere diseñar escenarios en los que se definan el conjunto de actividades socioeconómicas, las alternativas de tecnologías, y los usos del territorio.

En ausencia de este conocimiento, se puede hacer un ejercicio de estimación rudimentaria sobre la disponibilidad de agua en 2100. Las dos variables que parecen clave son el aumento de evapotranspiración de las plantas y la reducción de precipitaciones. Un rango razonable para el cambio de estas dos variables en el horizonte 2100 es un aumento en torno al 4% de la evapotranspiración,[2] y una reducción en torno al 20% de las precipitaciones,[3] lo que supondría una caída del caudal de agua en cuenca de entre el 20 y el 30%.

Una cuestión que tiene gran importancia es distinguir entre la margen izquierda y la margen derecha del Ebro, porque la caída de caudales en los afluentes de la margen izquierda no llega al 20% pero en la margen izquierda alcanza el 40-50%, lo que supondría una gran amenaza sobre las actividades humanas en esas cuencas.[4] Pero esa cuestión importante no se analiza en este sencillo ejercicio de estimación sobre disponibilidad de agua en Aragón en 2100.

Una caída de caudal de agua en cuenca entre el 20 y el 30%, obligaría a una reducción similar en las extracciones, lo que tendría como consecuencia una reducción de demanda consuntiva en Aragón de entre 600 y 900 hm^3, sobre el total de 3.000 hm^3 de demanda consuntiva agrícola, industrial y doméstica que se utiliza actualmente.

La reducción de extracciones se debería conseguir disminuyendo la demanda de riego de 2.800 hm^3, ya que la demanda urbana e industrial supone únicamente el 6 por cien de la demanda consuntiva y no supera los 200 hm^3. Una reducción de hasta 600 hm^3 en las extracciones para regadío, forzaría una caída en la demanda de riego de 2.800 a 2.200 hm^3, lo que podría lograrse mediante inversiones en sistemas de riego más eficientes que permitieran ahorrar agua.

El riego por inundación absorbe en la actualidad 1800 hm^3 (Cuadro 11.1), y la modernización de ese regadío permitiría ahorrar hasta 600 hm^3. Ahora bien, reducciones que alcancen o superen los 1.000 hm^3 no pueden conseguirse ya con ahorros de agua generados por inversiones en modernización de los sistemas de riego. Esas grandes reducciones solo serían factibles mediante la cesión por los agricultores de sus recursos hídricos, con el abandono de sus actividades de regadío y su sustitución por otras actividades de secano o ganaderas.

LOS RECURSOS HÍDRICOS

Las políticas de agua nacionales y europeas aún no incluyen el cambio climático, ya que los procesos de elaboración y puesta en marcha de las políticas son lentos y tienen una gran inercia. La principal política del agua en Europa es la Directiva Marco del Agua, y las principales políticas en España son el Plan Hidrológico Nacional, el segundo Plan Nacional de Calidad del Agua y el Plan Nacional de Regadíos. La Directiva del Agua promueve el principio de que los precios del agua deben aproximarse al coste completo de recuperación, incluyendo los costes de extracción, distribución, tratamiento, costes medioambientales, y valor del recurso. También se establece una combinación de límites de emisión y estándares de calidad del agua, con fechas límite para alcanzar el buen estado ecológico de las aguas. La Directiva del Agua pretende conseguir el buen estado ecológico de todas las aguas en 2015, lo que es cuestionable porque para esa fecha no van a estar disponibles los conocimientos biofísicos, cuya generación cuesta mucho tiempo y recursos sustanciales, ni tampoco podrá ponerse en marcha la necesaria acción colectiva derivada de la cooperación entre los agentes que manejan los recursos hídricos en cada subcuenca.

Recientemente, la Unión Europea ha aprobado la Directiva de Evaluación y Gestión de Riesgos de Inundación, que obliga a tomar medidas para reducir los riesgos de las inundaciones. Sin embargo, sobre la cuestión de la escasez de recursos hídricos y sequías, que es la que interesa realmente a los países del sur de Europa, no se ha conseguido una Directiva, sino únicamente una Comunicación sin valor legislativo. Ambas cuestiones, inundaciones y sequías,

[1] La demanda no consuntiva está formada por los usos de producción hidroeléctrica, refrigeración de centrales de energía y acuicultura, y para toda la cuenca el aprovisionamiento alcanza los 38.000 hm^3 en producción hidroeléctrica, 3.100 en refrigeración de centrales, y 1.000 en acuicultura.

[2] El 4% es el rango de aumento anual de evapotranspiración que para 2100 se considera en CSIRO (2007).

[3] La reducción del 20% en las precipitaciones es el rango de predicción del Cuarto Informe de Evaluación del IPCC para la península Ibérica.

[4] Ver el trabajo de Samper et al. (2007) en el que se utilizan los resultados del modelo de circulación general CGCM3 del Canadian Centre for Climate Modelling and Analysis.

son acontecimientos extremos cuya frecuencia va a aumentar como consecuencia del cambio climático.

En cuanto a las políticas de agua nacionales, el componente clave del Plan Hidrológico Nacional es el programa AGUA, que sustituyó al derogado trasvase del Ebro, y cuyo objetivo es aumentar la oferta de agua en el levante y sureste mediante desalación. En cuanto al segundo Plan Nacional de Calidad de Agua tiene como objetivo seguir reduciendo la contaminación de los centros urbanos y mejorar el abastecimiento. El Plan de Calidad plantea inversiones de 20.000 millones de euros en el horizonte 2015, para renovar las plantas de depuración y alcanzar un tratamiento terciario, proteger las zonas sensibles de las fuentes de abastecimiento, y construir depósitos de tormentas.

El Plan Nacional de Regadíos es una iniciativa importante de la administración con una financiación que supera, junto al Plan de Choque, los 6.000 millones de euros. La modernización de regadíos tiene un gran potencial para expandir cultivos rentables, conservar agua y reducir la carga de contaminación difusa, y requiere esfuerzos coordinados entre los agricultores y la administración.

Durante los últimos años, en Aragón ha habido un gran esfuerzo de modernización por parte de los agricultores con el apoyo de los gobiernos central y autónomo en el marco del Plan Nacional de Regadíos. Las inversiones totales (pública y privada) en modernización de regadío ya realizadas se acercan a los 300 millones de euros y 90.000 ha, y el nuevo Plan de Choque supondrá inversiones adicionales para alcanzar unos 400 millones de euros y 120.000 ha. De las 400.000 ha de regadío en Aragón, se ha conseguido que los sistemas de riego más modernos, por aspersión y localizado, pasen de 110.000 ha según el Censo Agrario de 1999, a 200.000 ha actuales y a 230.000 ha previstas tras el Plan de Choque.

La modernización de regadíos es una política fundamental para reducir la contaminación difusa de la agricultura y conseguir los objetivos de mejora de la calidad del agua y una herramienta fundamental para la lucha contra la escasez de recursos hídricos como consecuencia del cambio climático. Estas inversiones no garantizan la solución de todos los problemas, pero es evidente que las innovaciones tecnológicas en sistemas de riego facilitan el control privado y público de la cantidad y la calidad del agua. La realización del potencial de la modernización de regadíos requiere la coordinación y colaboración entre las autoridades del agua y las comunidades de regantes. Una caída de caudal en Aragón de 600 hm^3 en el horizonte 2100, podría ser compensada mediante la modernización de regadíos.

Otra iniciativa del Ministerio de Medio Ambiente ha sido la elaboración de planes de sequía por las confederaciones hidrográficas, que constituyen herramientas clave ante la mayor frecuencia de sequías en las cuencas peninsulares. Los planes incluyen medidas para reducir los impactos económicos, ambientales y sociales de las sequías. Estas medidas deben desarrollarse para hacer frente a los efectos del cambio climático.

El cambio climático es el debate en la sociedad aragonesa sobre la cuestión de si será necesaria una mayor regulación de agua embalsada mediante la construcción de nuevas presas, como propone el Pacto del Agua de Aragón. El problema se deriva de la desaparición progresiva de la acumulación de nieve en las montañas del Pirineo. La falta de nieve en el Pirineo, supondrá la desaparición del almacenamiento natural de recursos hídricos, y esta capacidad de almacenamiento solo podría ser sustituida por capacidad artificial mediante embalses para poder disponer de la misma cantidad de recursos hídricos. No se han construido presas en el Pirineo aragonés durante las últimas décadas, y la presión de la demanda actual sobre la capacidad de aprovisionamiento disponible es cada vez mayor. Si el cambio climático reduce significativamente el almacenamiento de nieve en las montañas del Pirineo, la construcción de los embalses del Pacto del Agua puede convertirse en una cuestión estratégica para Aragón. Esta cuestión estratégica debe ser decidida por la sociedad aragonesa, que debe elegir entre la alternativa de reducciones significativas en la utilización de agua o la construcción de infraestructuras hidráulicas que mantengan el nivel de utilización de recursos hídricos. La utilización de aguas subterráneas no puede resolver el problema de mantener el nivel de utilización de agua frente a la caída de almacenamiento natural por la nieve del Pirineo, y tampoco puede resolver la fuerte caída de caudal en los ríos de la margen derecha del Ebro. El debate admite soluciones intermedias, pero estas soluciones deben estar basadas en

conocimientos científicos e información que es necesario generar y que no está disponible. Por ejemplo no hay predicciones según escenarios sobre el caudal de las subcuencas del Ebro, ni sobre la evolución de la nieve en el Pirineo, por lo la discusión no puede estar fundamentada en hechos contrastados.

El papel de los ríos frente al cambio climático es importante, porque constituyen una defensa para la protección de los ecosistemas frente a un clima más calido y seco, que provocará la desertización de las zonas áridas y semiáridas.

Los acuíferos de Aragón y de la cuenca del Ebro no están muy degradados ya que no se están utilizando como ocurre en otras cuencas peninsulares. La mayor degradación de los recursos hídricos ligada a daños muy severos sobre los ecosistemas acuáticos se corresponde con las zonas que utilizan acuíferos en las cuencas del Júcar, Segura, Sur, Alto Guadiana y Guadalquivir. La sobreexplotación puede superar los 1000 hm^3 y está ligada a la existencia de más de un millón de pozos ilegales. El desgobierno de las aguas subterráneas es un problema serio en España, y muy difícil de arreglar porque los acuíferos son bienes comunales (bienes públicos impuros), cuya gestión sostenible requiere medidas efectivas de mando y control de las autoridades de cuenca, incentivos que consigan la cooperación de los usuarios, y el logro de la acción colectiva de los usuarios del recurso (Albiac et al. 2008).[5] Desde el punto de vista de la política medioambiental, eso es tan difícil de conseguir como la conservación de los bancos de pesca en aguas de alta mar internacionales, o que los países reduzcan significativamente sus emisiones de GEI. Ante las enormes dificultades de gestión, no se entiende el empecinamiento de las organizaciones ecologistas en desarrollar la explotación de los acuíferos en la cuenca del Ebro.

Esto no quiere decir que no se emplee el bombeo de agua, almacenamiento y generación de energía, sin necesidad de emisiones GEI. El bombeo desde un acuífero a un embalse puede ser fundamental en un momento prolongado de sequía. Pero el control por la autoridad de cuenca tiene que ser férreo.

La opción de continuar utilizando el regadío para producir alimentos es importante, en vista de que la expansión mundial de demanda de alimentos que ha provocado la subida de sus precios es una tendencia que se va a mantener a medio y largo plazo. El regadío también se puede utilizar para cultivos energéticos y fines de ocio.

Los ríos son fuentes de vida a proteger, y la cuenca del Ebro no ha experimentado la enorme presión de actividades humanas y la consiguiente degradación de la cantidad y calidad de los recursos hídricos. La conservación y protección de los ecosistemas acuáticos en el Ebro es muy superior a muchas cuencas españolas y europeas. Además de la menor presión humana, también han contribuido los esfuerzos de las instituciones de las sociedades que se asientan en la cuenca. El cambio climático va a tener unos efectos muy importantes sobre los recursos hídricos de Aragón, por lo que los ecosistemas acuáticos van a experimentar una amenaza muy seria.

OBJETIVOS PARA LA MITIGACIÓN Y ADAPTACIÓN DEL CAMBIO CLIMÁTICO EN EL SECTOR PRIMARIO Y LOS RECURSOS HÍDRICOS

Los principales objetivos para el sector primario, que abarca agricultura, ganadería y bosques; y para los recursos hídricos son los siguientes:

[5] La sobreexplotación de acuíferos y su degradación es generalizada en todos los países que utilizan acuíferos, con casos muy graves como los de China, India, Australia y Estados Unidos. Hay dos excepciones de acuíferos gestionados de forma sostenible, uno en Estados Unidos (Idaho) y otro en España (Mancha Oriental). Para tener una idea sobre las dificultades de la gestión de acuíferos, no hay más que observar el caso del Plan del Alto Guadiana (acuífero de Mancha Occidental), donde el Ministerio de Medio Ambiente pretende invertir tres mil millones de euros para acabar con la sobreexplotación de 200 hm^3 (aunque es dudoso que la enorme inyección financiera suscite la acción colectiva, como ya ha ocurrido en los últimos 15 años). Otro ejemplo es la degradación de la cuenca del Guadalquivir, que ha duplicado su superficie de regadío en los últimos quince años, con una enorme expansión del bombeo individual ilegal.

- Mantenimiento del tipo de agricultura extensiva característica del Valle del Ebro, y freno a la cada vez más importante ganadería intensiva. Es necesario evitar la transformación hacia una agricultura intensiva en capital, con grandes concentraciones productivas, que se apoya en tecnologías de cultivo agresivas en fertilizantes y pesticidas, elevado uso de materiales y energía, y la utilización a gran escala de mano de obra inmigrante ilegal. La agricultura y la ganadería extensivas son más benignas para la conservación del medio ambiente y los recursos naturales, y para proteger los ecosistemas y sus servicios medioambientales.

- Freno y reducción drástica de la ganadería intensiva de porcino, que emite el 13% de los GEI en Aragón y cuyo valor añadido no se obtiene en Aragón sino en otras regiones. Fomento de la ganadería extensiva de ovino y vacuno, que sirve para el mantenimiento de las zonas de pastos y que genera producciones autóctonas de calidad y alto valor añadido, como el jamón serrano con denominación de origen.

- Conseguir una gestión más apropiada del estiércol porcino para reducir las enormes emisiones de metano y óxido de nitrógeno, ya que la cabaña de Aragón de 4,5 millones de cabezas solo es superada por la cabaña de Cataluña. Potenciar la extensión de purín con cubrimiento, y evaluar la viabilidad de las plantas de tratamiento para producir compost, y de las plantas de biogás.

- Controlar las emisiones de metano de fermentación entérica de la cabaña de ovino y vacuno, a través de la mejora de la alimentación, la contención de la cabaña intensiva, y el aumento de eficiencia en los procesos de producción.

- Transformación de los cultivos de secano y regadío a la producción ecológica, lo que es muy sencillo en los cultivos de secano.

-Promover un uso más eficiente del abonado mineral y orgánico, que reduzca las emisiones de óxido de nitrógeno directas en parcela, como las indirectas del nitrógeno arrastrado por lixiviación hacia las capas freáticas y escorrentía. Esta reducción tiene efectos positivos adicionales sobre la mejora de la calidad del agua que beneficia a los ecosistemas acuáticos y a los usos antrópicos.

-Potenciar la gestión de bosques y la revegetación de carácter multifuncional, que promueva la fijación de carbono y favorezca la biodiversidad. Es importante mantener las actividades que generan ingresos de mercado como extracción de madera, leña, caza y setas, porque esto sirve para sostener el uso tradicional del bosque por las poblaciones locales. También hay que procurar que servicios medioambientales de los bosques como la contribución al ciclo hidrológico, soporte de biodiversidad, fijación de carbono, y recreativo y paisaje, reviertan en forma de ingresos a la población local que contribuye a la conservación de los bosques.

-Alcanzar una gestión más sostenible de los recursos hídricos en el regadío, equilibrando la introducción de tecnologías de riego avanzadas que consumen energía con el logro de una mayor eficiencia energética. También se debe tener en cuenta que las tecnologías de riego avanzadas permiten disminuir el uso del agua y fertilizantes, y reducen la contaminación por nutrientes y sales.

Los principales objetivos para el sector de recursos hídricos son los siguientes:

- Hacer frente a la importante reducción de caudales en los afluentes del Ebro, que para el horizonte 2100 puede alcanzar una caída del 20 por cien en la margen derecha, y del 40 por cien en la margen izquierda. La caída del 40 por cien de caudales en las cuencas de la margen derecha suponen una amenaza muy seria para los ecosistemas acuáticos y para las actividades humanas. En Aragón las condiciones climatológicas de la margen derecha son similares a las zonas semiáridas y áridas del levante y sureste peninsular, y los problemas van a ser compartidos.

- Compensar la disminución del almacenamiento del recurso en forma de nieve en toda la cordillera pirenaica.

- Reducir el impacto de la disminución de las lluvias y su cambio de estacionalidad, así como de la mayor frecuencia e intensidad de los fenómenos extremos de sequías y avenidas.

- Preparar los cambios necesarios en la extracción de recursos y en la calidad de los retornos, para proteger los ecosistemas acuáticos y los usos humanos esenciales.

- Es necesario preparar medidas que afronten las sequías extremas, las lluvias torrenciales, y las demás anomalías climáticas. Para ello hay que desarrollar los planes de sequía que ya se han iniciado, establecer directrices de uso del suelo, y medidas de protección de los ecosistemas.

La situación citada más arriba de degradación severa de los recursos hídricos en las cuencas del Júcar, Segura, Sur, Alto Guadiana, y Guadalquivir es consecuencia de la mala gestión de los recursos hídricos en las últimas tres décadas. Con el trasvase del Ebro se pretendía compensar los 450 hm^3 de sobreexplotación del Júcar, Segura y Sur, y los 220 hm^3 de sobreexplotación del Alto Guadalquivir (sustituyendo con agua del Ebro el trasvase Tajo-Segura, y llevando el agua de este trasvase al Alto Guadalquivir, lo que facilitó el apoyo de Castilla-La Mancha al trasvase del Ebro). El arreglo de tanto desastre de gestión era excesivo, incluso para un río como el Ebro que aún no está degradado. El caos de gestión de los recursos hídricos ha continuado en la última década alcanzando al Guadalquivir, que ya está pidiendo recursos hídricos mediante trasvases. La lógica en que se apoya esta mala gestión es que una vez que los territorios destruyen sus recursos hídricos, entonces pueden pretender que otras cuencas aún no degradadas les envíen sus recursos.[6]

Son necesarias amenazas creíbles para conseguir la cooperación que lleva a la conservación de los recursos. La respuesta de Aragón a este comportamiento estratégico gorrón, debería ser amenazar desarrollando el uso de sus recursos hídricos para impedir que otros territorios se aprovechen de los "excedentes". Esta amenaza ha de ser creíble, ante la falta de cooperación de otros territorios en la acción colectiva de protección de los recursos hídricos y los ecosistemas. El cambio climático va a desertizar la margen derecha del Ebro en Aragón, y aún más las cuencas del levante, sureste y sur peninsular. Ante el fallo de gestión de las cuencas de levante, sureste y sur, la presión para trasvasar agua del Ebro no puede más que aumentar en las próximas décadas.

DISEÑO DE POLÍTICAS: MEDIDAS PROPUESTAS DE ADAPTACIÓN Y MITIGACIÓN

Las medidas recomendadas para el cumplimiento de los objetivos deben diseñarse a escala individual, local y regional, y a su vez en cada nivel deberían proponerse medidas de adaptación y de mitigación. A nivel nacional sugerimos también una serie de medidas de investigación.

Las medidas de *adaptación* son las políticas que han de tomarse para reducir los impactos negativos del cambio climático. Las medidas de *mitigación* son políticas que reducen las emisiones GEI de las actividades del sector primario.

Medidas de *adaptación* en el sector primario:

Prever los cambios en cultivos y variedades, de forma que las actividades agrícolas extensivas del cultivo de herbáceos y leñosos, se adapten a la mayor aridez. También se puede optar por cambiar hacia una agricultura intensiva en capital, tecnología, mano de obra, agua, energía y materiales, para el caso de cultivos de alto valor añadido que se quieran mantener o potenciar. Para ello se puede seguir el ejemplo de la agricultura intensiva del sureste peninsular.

Relocalizar la ganadería extensiva de ovino y vacuno hacia las zonas que sean más apropiadas para pastos. Otra opción es seguir potenciando la ganadería porcina y vacuna intensiva en capital, tecnología y demás inputs, pero con gran impacto negativo sobre el medio ambiente y las emisiones GEI.

Estudiar los cambios en las precipitaciones y temperaturas para determinar los recursos hídricos disponibles en las cuencas, así como los cambios en las necesidades hídricas de los cultivos por la reducción de lluvias, para determinar el impacto sobre el grueso de la demanda

[6] Este comportamiento estratégico en bienes comunales se denomina en economía del medio ambiente aprovechado, gorrón o polizón (free-rider). Como se describe en La Tragedia de los Comunes (Hardin 1968), el comportamiento gorrón o polizón consigue destruir bienes comunales como los recursos hídricos, los bancos de pesca o los bosques.

de agua, que es el regadío. Utilizar la información para la planificación hidrológica y la gestión del uso de la tierra.

Ante la previsible agudización de los problemas de escasez de agua en las cuencas, será necesario disminuir las emisiones contaminantes puntuales y difusas que deterioran la calidad del agua de los ríos. Para ello se deberá planificar la reducción de la carga de contaminación urbana e industrial, y de la carga de contaminación difusa de la agricultura y la ganadería.

Las previsiones de modificación del clima se incorporarán en la gestión de los bosques y la vegetación, para facilitar la supervivencia de los ecosistemas y para favorecer en lo posible la migración de los ecosistemas hacia zonas geográficas más adecuadas.

Las principales medidas de adaptación del sector agrario y forestal están ligadas al aumento de temperaturas y la reducción de precipitaciones. El aumento de temperaturas significa una mayor evapotranspiración de las plantas, aunque el efecto principal es la reducción de precipitaciones. Ello tendrá como consecuencia un fuerte aumento de las necesidades de riego.

Las comarcas que se verán más afectadas por el aumento de necesidades de riego son las que más agua consumen: Monegros (500 hm^3), Cinco Villas (350 hm^3), Zaragoza (280 hm^3), Cinca Medio (220 hm^3) y Bajo Cinca (200 hm^3). El volumen actual de recursos hídricos utilizados en el regadío de Aragón es elevado y alcanza los 2.500 hm^3, por lo que es difícil que aumente en el futuro la disponibilidad de agua. La adaptación puede realizarse incorporando sistemas de riego más eficientes como el riego localizado y por aspersión, aunque estos sistemas a presión consumen energía. Otro mecanismo de adaptación consiste en modificar los tipos de cultivos, reduciendo las superficies de maíz y arroz que consumen grandes cantidades de agua.

Medidas de *mitigación* en el sector primario:

Se propone la reducción de la cabaña porcina pero potenciando aquella que genera producciones autóctonas de calidad y alto valor añadido, como el jamón serrano de la denominación Teruel y otros productos frescos y curados. Los responsables de la política agrícola en Aragón deben potenciar la producción y abastecimiento local de frutas y hortalizas, sustituyendo en temporada las producciones intensivas de invernadero.

A nivel comarcal y para contribuir al cumplimiento de los objetivos del Plan de Energías Renovables en utilización de energías renovables y biocarburantes para transporte, es necesario aprovechar la biomasa disponible en Aragón (residuos herbáceos y leñosos, biomasa forestal, residuos industria agroalimentaria y forestal), e incorporar 200.000 ha de cultivos energéticos para bioetanol y biodiesel. Para biodiesel se podrían dedicar 80.000 ha de oleaginosas (colza u otras), y para bioetanol se podrían destinar 40.000 ha de trigo, 30.000 ha de cebada, 20.000 ha de maíz y 20.000 ha de remolacha. Esto requiere subir los precios de los cereales para bioetanol, y aumentar los pagos de la UE para cultivos energéticos.

A nivel individual cada sujeto puede patrocinar el consumo de alimentos ecológicos de producción local, para lo que serían necesarias campañas de publicidad que convenzan a los consumidores. En el consumo de hortalizas y frutas, los individuos pueden patrocinar la compra de productos locales, en lugar de las producciones intensivas de invernadero. Los productos locales tienen una mejor calidad, sobre todo en verano cuando empeora la calidad del agua y la calidad del producto en las zonas de invernadero. Este consumo local podría estimularse para los productos en los que el nivel de consumo en Aragón está muy por encima del nivel de producción local, como es el caso de lechuga, judía verde, patata, melón, sandía, olivas y aceite de oliva. Aragón es un gran exportador a otras regiones en otros productos como tomate, cebolla, manzana, melocotón y pera.

En el consumo de carnes, el principal consumo es la carne de porcino. Ahora bien, el impacto de los cambios en el consumo de carne de porcino sobre las emisiones es bastante limitado, ya que el consumo de carne de porcino en Aragón es de solo 26.000 t,[7] mientras que la producción de carne alcanza 580.000 t de las que 215.000 t se sacrifican en Aragón y el resto en

[7] Incluye 6.000 t carne porcino transformada. La producción en España es 3,3 millones t de las que las exportaciones netas son 0,54 millones t. Como se ha señalado anteriormente, de las 580.000 t de carne de porcino que produce la enorme cabaña de Aragón (4,5 millones de cabezas), en Cataluña se sacrifican 365.000 t mientras que en Aragón solo se sacrifican 215.000 t, por lo que la mayor parte del valor añadido se va a Cataluña.

Cataluña. También puede haber una cierta influencia a través del consumo de ovino y vacuno, sobre sus producciones y por tanto sobre sus emisiones de metano. El consumo de carnes también influye sobre la producción de cereales y alfalfa, y por tanto sobre las emisiones de óxido de nitrógeno. Otros efectos del consumo alimentario sobre las emisiones de óxido de nitrógeno son los consumos de hortalizas, frutales y otros productos varios.

Las principales emisiones de gases de efecto invernadero del sector agrario en Aragón son los 1,9 millones tCO_2 equivalente que emite la enorme cabaña de ganado porcino, formadas por emisiones de metano y de óxido de nitrógeno. Las comarcas (MAPA) donde se localizan las mayores emisiones son Ejea (0,27 millones tCO_2), Bajo Aragón (0,22), Zaragoza (0,21), La Litera (0,21) y Bajo Cinca (0,16).

Las medidas de mitigación disponibles para reducir estas emisiones de metano del estiércol de porcino son: extensión del purín en los campos de cultivo, (2) utilización de balsas, (3) plantas de tratamiento y (4) plantas de producción de biogás y plantas de cogeneración. La extensión del purín en los campos de cultivo tiene la ventaja de que es la mejor manera de utilizar los nutrientes, pero hay que tener mucho cuidado con la forma de aplicación y las dosis de purín, lo que afecta a las emisiones de metano y de óxido de nitrógeno. La extensión del purín puede realizarse de dos formas, por extensión directa o por cubrimiento. La extensión directa consiste en extender el purín mediante aspersión en abanico, rampa multiboquilla, o rampa de tubos colgantes. El cubrimiento consiste en enterrar el purín, y puede hacerse mediante enterrador con reja o con disco, o por inyección. También puede utilizarse una rampa de tubos para extender y realizar después una labor de cubrimiento. El enterramiento conviene realizarlo en las dos horas posteriores a la extensión.

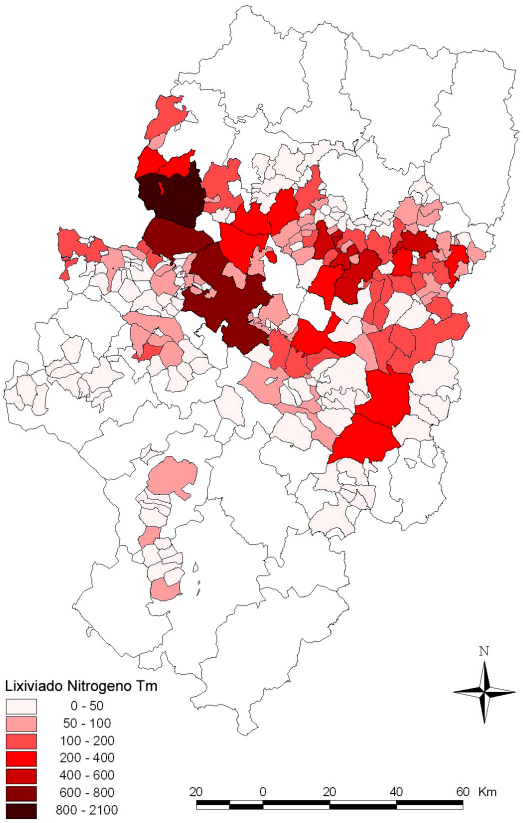
El procedimiento de balsas consiste en construir balsas de contención, donde se evapora el purín y posteriormente se extraen los residuos. Las balsas se han utilizado en los últimos años, pero no son recomendables porque generan emisiones muy elevadas de metano y de óxido de nitrógeno. En las plantas de tratamiento el purín se somete a procesos físico-químicos y biológicos aeróbicos y anaeróbicos. En estos procesos se genera agua y compost que se utiliza como abono. Las plantas de biogás producen metano a partir del purín, y en las plantas de cogeneración se calienta el purín y con el vapor de agua se produce electricidad que se coloca en la red. El problema de la cogeneración es que en la evaporación se producen emisiones contaminantes, cuyo filtrado tiene un coste demasiado elevado que hace no rentable la cogeneración.

La viabilidad de estas medidas depende sobre todo de los costes. Un coste muy importante es el coste de carga y descarga de los purines, por lo que las plantas de tratamiento que requieren la carga y descarga del purín tienen una rentabilidad dudosa. La extensión de los purines en el campo es actualmente la única opción con costes razonables en Aragón. En la Unión Europea se han dado algunos tipos de experiencias como la cooperación entre productores sin suficientes tierras (Francia) para poner en marcha plantas de tratamiento, y la limitación estricta o reducción de la cabaña porcina como ha ocurrido en Holanda, desplazando los criaderos a países como España.

En cualquier caso, las medidas de control de las emisiones del estiércol porcino deben coordinarse con la Directiva de Nitratos y la Directiva Marco del Agua. La Directiva de Nitratos ya ha introducido la obligación de realizar el balance de nitrógeno en las explotaciones ganaderas, y esta información es muy útil para el diseño de incentivos y medidas. El problema es que solo cubre a los ganaderos en zonas de acuíferos vulnerables.

A las emisiones de la cabaña de porcino, le siguen en importancia los 1,8 millones tCO_2 en emisiones de óxido de nitrógeno de las actividades de cultivo, como consecuencia del abonado con 112.000 tN de fertilizantes minerales y orgánicos. La mitad de estas emisiones de óxido de nitrógeno se concentran en las comarcas de Monegros (0,25 millones tCO_2), Cinco Villas (0,23), Zaragoza (0,14) y Hoya de Huesca (0,13).

Figura 1. Emisiones de nitrógeno del regadío en el origen (t N).



En estas cuatro comarcas es necesario mejorar las prácticas de abonado para reducir la volatilización del abonado de nitrógeno, y la carga contaminante de nitrógeno en origen (Figura 1) por escorrentía superficial y percolación. También es necesario recuperar humedales naturales o crear humedales artificiales que reduzcan la contaminación de nitrógeno y por tanto las emisiones de óxido de nitrógeno.

Las medidas disponibles para reducir la contaminación son: impuestos sobre el fertilizante nitrogenado, límites al abonado, y estándares sobre contaminación en el ambiente (en los cursos de agua) ligados a un impuesto sobre las emisiones o una penalización en forma de multa.

Estas medidas se deben diseñar e implementar conjuntamente con las medidas de la Directiva de Nitratos y de la Directiva Marco del Agua. La Directiva de Nitratos ha introducido la obligación de que los agricultores realicen el balance de nitrógeno de la explotación agrícola y tengan estas cuentas disponibles, lo que facilita el establecimiento de cualquier medida. Pero la obligación solo cubre las explotaciones agrícolas situadas en zonas de acuíferos vulnerables.

Aun con todo, el problema de estas medidas es que se refieren a contaminación difusa (es decir hay un comportamiento estratégico de los agentes), y para que sean efectivas es necesario establecer incentivos que consigan la cooperación de los agentes en el logro de la acción colectiva.

El tercer componente de de las emisiones GEI está formado por la carga de 0,9 millones tCO_2 de emisiones de metano de fermentación entérica procedente de las cabañas de ovino y vacuno. Las medidas en este caso consisten en la mejora de la alimentación del ovino y el vacuno, modificando los componentes de las raciones y desarrollando nuevas tecnologías de alimentación.

Una de las medidas consiste en establecer bases de datos accesibles a los ganaderos de porcino sobre la carga de contaminación en su explotación y en las parcelas en las que extiende el purín, y lo mismo debe hacerse para los ganaderos de ovino y vacuno. También hay que establecer bases de datos accesibles a los agricultores sobre las necesidades de riego de los cultivos, y sobre la carga de contaminación de óxido de nitrógeno de sus actividades de cultivo.

La administración debe apoyar la expansión de las producciones ecológicas, lo que es muy fácil de alcanzar en secano y también en regadío. Tambíen se debe promover entre los consumidores los alimentos ecológicos, de mayor coste y cuyo consumo es actualmente despreciable y apoyar las actividades de recreo y ocio ligadas al conocimiento de los ecosistemas, de forma que la población rural ofrezca estos servicios y los visitantes estén dispuestos a demandar estas actividades y pagarlas.

Son necesarios programas de investigaciones nacionales e internacionales sobre el desarrollo de variedades de cultivos perennes de cereales, leguminosas y oleaginosas, que sustituyan a los cultivos anuales. Las ventajas de que estos cultivos sean perennes en lugar de anuales son: creación de grandes sistemas de raíces capaces de conservar el suelo, conservación de los ecosistemas y mejora de la biodiversidad, reducción del uso de abonos y pesticidas, menor uso de agua y caída de la percolación y de la contaminación.

Medidas de *adaptación* de los recursos hídricos:

Las zonas que se verán más afectadas por el aumento de necesidades de riego son las de los grandes sistemas de regadío, que consumen más agua: Bardenas, Monegros, Cinca, Canal de Aragón y Cataluña y Canal Imperial. Será necesario elaborar planes de adaptación específicos para cada una de estas zonas.

Los recursos hídricos utilizados en el regadío de Aragón alcanzan los 2.800 hm^3, mientras que la utilización urbana e industrial alcanza los 200 hm^3. Cualquier medida de adaptación al cambio climático debe estar ligada a la reducción de la demanda de riego. Si el volumen de adaptación en el horizonte 2100 no supera los 600 hm^3, la medida más razonable es la de inversiones de modernización en los sistemas de riego, que además sirve para reducir la contaminación difusa del regadío. Si el volumen de adaptación superara los 1.000 hm^3, entonces los ahorros por modernización de regadíos no serían suficientes.

Otro mecanismo de adaptación consiste en modificar los tipos de cultivos, sustituyendo cultivos intensivos en agua como maíz y arroz, por hortalizas y frutales. De todas formas también se debe examinar el nivel de intensificación productiva que se desea, porque los efectos medioambientales de una agricultura más intensiva pueden ser negativos.

La actual organización institucional y legislativa ya garantiza la adaptación a los periodos de escasez de agua en la cuenca del Ebro, como demuestra el Plan de Sequía de la Cuenca del Ebro. Otro factor favorable es que en Aragón el regadío se realiza mediante sistemas colectivos de presas y canales controlados por la autoridad de cuenca, por lo que la respuesta de adaptación en la reasignación del agua del regadío a otros usos prioritarios es inmediata.

El fomento en la utilización por los agricultores u otros usuarios de los acuíferos mediante el bombeo individual, como propugnan algunas organizaciones ecologistas, hace mucho más difícil el control social de los recursos hídricos. La consecuencia sería que la reasignación a usos prioritarios sería mucho más complicada, y que la degradación de los ecosistemas acuáticos sería mucho más fácil.

Las estimaciones disponibles parecen indicar que la caída de caudales en la cuenca del Ebro para el horizonte 2100 estarán entre el 20 y el 30 por cien. Ahora bien, hay una diferencia crucial entre la margen izquierda y la margen derecha, que no se ha tenido en cuenta en el presente trabajo. Parece ser que el caudal en las cuencas más áridas de la margen derecha del Ebro, podría caer hasta un 40 o 50 por cien. Los riesgos que esto introduciría para la continuidad de las actividades humanas serían realmente severos, y esta cuestión merece estudiarse con detalle y profundidad.

La realización de las obras del Pacto del Agua podría convertirse es una cuestión estratégica para Aragón, si se confirma la pérdida de almacenamiento de nieve en las montañas de Pirineo aragonés y la caída de caudales en los ríos, como consecuencia del cambio climático. Esta cuestión debe ser debatida por la sociedad aragonesa, para elegir entre i) reducir la utilización de agua y no realizar obras de infraestructura, o ii) mantener la utilización de agua con nuevas infraestructuras. Este debate requiere la generación de conocimiento científico fiable sobre las consecuencias del abanico de alternativas.

El cambio climático va a aumentar fuertemente la desertización en las cuencas del levante, sureste y sur peninsular, que llevan mucho tiempo con una gestión desastrosa de los recursos

hídricos, en especial los subterráneos. Por ello, la presión para que se haga el trasvase del Ebro no puede más que aumentar en las próximas décadas. Aragón necesita llevar a cabo amenazas creíbles sobre estos territorios, para lograr la cooperación y acción colectiva en la protección de los recursos hídricos.

RESUMEN Y CONCLUSIONES

Este trabajo examina las medidas de mitigación y adaptación del sector primario y del sector de recursos hídricos en Aragón ante el cambio climático. La mayoría del territorio de Aragón se utiliza en actividades agrícolas y forestales, y estas actividades tienen un gran influencia sobre los ecosistemas y la provisión de servicios medioambientales.

La puesta en marcha de medidas como el impuesto sobre emisiones del Informe Stern (Stern 2007) es técnicamente, aunque no políticamente, sencillo en el sector de energía. Este impuesto incorporaría el daño de emisiones, una vez que el daño se haya estimado. Sin embargo en los sectores de emisiones difusas, el diseño e implementación de medidas es mucho más complicado, ya que las emisiones no pueden medirse y controlarse a un coste razonable, por lo que los instrumentos económicos no funcionan. Los instrumentos de control de las emisiones difusas son mucho mas sofisticados y requieren conocimientos sobre los procesos biofísicos a escala local. Además es necesario tener en cuenta el comportamiento de los agentes para el diseño de medidas, que permitan alcanzar la cooperación y la acción colectiva.

Se han presentado recomendaciones de adaptación y mitigación para el sector primario y para los recursos hídricos. Estas medidas deben estar ligadas a indicadores para poder seguir su evolución y cumplimiento. Las medidas de gestión del agua se definen teniendo en cuenta las actuales instituciones, en las que ya funciona la cooperación y la acción colectiva de los usuarios y grupos de interés, y de las políticas de agua nacionales y europeas.

Es necesario establecer una serie de indicadores que evalúen el cumplimiento de los objetivos y la eficacia de las medidas de mitigación y adaptación. Los indicadores más importantes para la medidas propuestas son los siguientes:

Distribución espacial a nivel de parcela del uso de la tierra por el sector primario: agricultura, ganadería y bosques, incluyendo información sobre las prácticas, tecnologías e inputs que se utilizan en cada actividad agrícola, ganadera y forestal.

Distribución espacial del consumo de energía de las actividades ganaderas y agrícolas.

Distribución espacial a nivel de parcela de la carga de emisiones de metano y óxido de nitrógeno del estiércol de la cabaña porcina, y de su proceso de transporte y destino. Distribución espacial a nivel de parcela de la carga de emisiones de metano de fermentación entérica del ovino y vacuno.

Medición de la reducción de emisiones del estiércol porcino y de la fermentación entérica lograda con cada medida y su coste. Clasificación de la eficiencia de costes de las medidas, y reevaluación del paquete de medidas.

Distribución espacial a nivel de parcela de los cultivos en secano y en regadío, la utilización de inputs, y las emisiones de óxido de nitrógeno en el origen, así como de su proceso de transporte y destino.

Distribución espacial de los cultivos energéticos y de los cultivos ecológicos.

Reducción de emisiones de óxido de nitrógeno de los cultivos lograda con cada medida y su coste. Reevaluación del paquete de medidas.

Distribución espacial a nivel de rodal de las características de los árboles, prácticas de gestión, extracciones de madera y leña, actividades de caza, obtención de setas, y servicios medioambientales provistos de no mercado. Estimación de la fijación de carbono en la vegetación y en el suelo.

Evaluación de los costes y beneficios de mercado y no mercado de las distintas alternativas de gestión del bosque, que contribuyan a la fijación de carbono.

Los indicadores para el sector hídrico que se proponen son los siguientes.

Las zonas que se verán más afectadas por el aumento de necesidades de riego son las de los grandes sistemas de regadío, que consumen más agua: Bardenas, Monegros, Cinca, Canal de

Aragón y Cataluña y Canal Imperial. Será necesario elaborar planes de adaptación específicos para cada una de estas zonas.

Distribución espacial y temporal de la cantidad y calidad de los recursos hídricos, con identificación de las masas de agua. Para ello se utilizarán los datos de caudales y de calidad de agua procedentes de la Confederación Hidrográfica del Ebro para aguas superficiales y subterráneas.

Distribución espacial y temporal de las extracciones superficiales y subterráneas, de los retornos a cuenca, y de las cargas de contaminantes puntuales y difusas. Identificación de los procesos de transporte y destino de los contaminantes.

Distribución espacial y temporal de los ecosistemas acuáticos y su estado. Impacto del volumen de caudales y de la carga de contaminantes sobre estos ecosistemas. Valoración económica del daño sobre los ecosistemas de los parámetros de cantidad y calidad del agua.

Evaluación espacial de las medidas de control de la cantidad y calidad de las masas de agua, según el coste de las medidas y el beneficio de estas medidas para los ecosistemas acuáticos.

Elaboración de indicadores de predicción de cantidad y calidad de las masas de agua para cada subcuenca de los ríos de Aragón, para los horizontes 2020, 2050 y 2100, que incorporen el conocimiento científico sobre las condiciones climáticas para esos horizontes y los escenarios de respuesta antrópica.

BIBLIOGRAFÍA

ALBIAC J., E. PLAYAN, Y. MARTINEZ. 2007. Instruments for Water Quantity and Quality Management in the Agriculture of Aragon. Water Resources Development 23(1): 147-164.

ALBIAC J., M. MEMA y E. CALVO. 2008. Sustainable Water Management and Nonpoint Source Pollution Control in Spain and the European Union. En J. Albiac y A. Dinar (eds.) The Management of Water Quality and Irrigation Technologies. Earthscan. Londres.

CSIRO (2007) "Climate Change in Australia. Technical Report 2007". CSIRO. Clayton South.

GOBIERNO DE ARAGON. 2008. Estrategia Aragonesa de Cambio Climatico. Consejeria de Medioambiente. Zaragoza.

HARDIN G. 1968. Tragedy of the Commons. Science 162: 1243-1248.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE. 2007a. Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Core Writing Team, R. K. Pachauri y A. Reisinger. Intergovernmental Panel on Climate Change IPCC. Ginebra.

INTERGOVERNMENTAL PANEL OF CLIMATE CHANGE 2007b. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change: Summary for Policymakers. Intergovernmental Panel on Climate Change IPCC. Ginebra

MATTHEWS G. 1993. The carbon content of trees. Technical paper 4. Forestry Commission. Edimburgo.

SAMPER J., D. AlvAres, B. PisaNi, M. GARCÍA VERA (2007) "Evaluación del efecto del cambio climático en los recursos hídricos en la cuenca hidrográfica del Ebro con GIS-BALAN". Ponencia presentada en las Jornadas de la Zona No Saturada del Suelo. Córdoba.

SAMPSON R. 1992. Forestry opportunities in the United Status to mitigate the effects of global warming. Water, Air and Soil Pollution 64: 157-180.

STERN N. 2007. The Economics of Climate Change. The Stern Review. Cambridge University Press. Cambridge.

UE, DESENVOLVIMENTO REGIONAL E RELAÇÃO PERIFERIA-ACESSIBILIDADE

VICTOR MANUEL PISSARRA CAVALEIRO
E-mail: cavaleiro@ubi.pt
JOSÉ RAMOS PIRES MANSO
E-mail: pmanso@ubi.pt
JORGE MIGUEL DOS REIS SILVA
E-mail: jmiguel@ubi.pt
Universidade da Beira Interior

Resumo

Um conjunto suficientemente robusto de indicadores de acessibilidade tornou-se na "pedra de toque" na busca de critérios significativos para aferição das variações na economia dos territórios. Aliás, muitos autores chamam a atenção inclusivamente para o perigo em eleger um único indicador de acessibilidade como de importância universal. Ora, ao conceito de acessibilidade está ligado o de perifericidade - realidade incontornável no contexto do desenvolvimento económico regional da UE. No entanto, avaliar o grau de perifericidade/centralidade de uma qualquer região não é tarefa fácil. Para tal, dividimos o nosso trabalho em três fases: definição de pressupostos para a escolha dos Indicadores de Acessibilidade; escolha dos impactes a observar na economia e no potencial de desenvolvimento das regiões periféricas resultantes das eventuais alterações induzidas ao nível da respectiva acessibilidade; e escolha dos Indicadores de Acessibilidade mais adequados. Terminamos com uma reflexão-síntese sobre a importância espacial dos Indicadores de Acessibilidade.

Palavras-chave: Periferia e Acessibilidade, Indicadores de Acessibilidade, Localização e Potencial de Mercado, Desempenho Económico.
Área Temática: Economia da Energia, do Ambiente, dos Transportes e dos Recursos Naturais.

Abstract

A strong enough set of accessibility indicators has become the "cornerstone of" in search of meaningful criteria for measuring the variations in the territories economy. Moreover many authors also call attention to the danger in electing a single accessibility indicator as of universal importance. However ccessibility concept is linked to that one of peripherality – a well known reality in the EU regional economic development context. Thus to assess the peripherality/centrality degree of any specific region within the EU we propose the division of this task into three main steps: definition of the requisites for an accessibility indicators choice; choice of the impacts to be observed on even the economy or the potential for development of the peripheral regions resulting from any change induced on their accessibility; and choice of the most adequate accessibility indicators to be used. Finally, we present a reflection-synthesis on the spatial importance of the accessibility indicators.

Key-words: Periphery and Accessibility, Accessibility Indicators, Location and Market Potential, Economic Performance.
Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

1 INTRODUÇÃO

O alargamento da UE e a consequente liberalização dos mercados estimularam o estudo e a previsão de fenómenos relacionados com a localização e a concentração do tecido produtivo; motivo suficiente para que se relançasse a discussão em torno da dicotomia Centro-Periferia, abrindo novos caminhos para a investigação de técnicas de avaliação do grau de perifericidade das regiões (Silva, 2005).

Do ponto de vista geográfico – ou topológico – a perifericidade pode ser encarada como a localização para além da faixa que limita determinada área, precisamente à distância máxima do respectivo centro geográfico. No entanto – e em termos de desenvolvimento regional – a perifericidade deve reflectir, igualmente, a distância à(s) actividade(s) económica(s) mais relevantes da região, do país, ou mesmo de um grupo de países. Neste sentido, a perifericidade – tal como a inversa, a centralidade - é marcada, quer pela configuração física do território, quer pela distribuição geográfica das actividades económicas no seu interior; ou seja: da relação com o desenvolvimento regional se constata que a perifericidade é sinónimo de relativa acessibilidade – ou inacessibilidade – à actividade económica, tal como esta se apresente distribuída, geograficamente, pelo território (Keeble *et al.*, 1988).

O conceito de Acessibilidade – embora ainda assim invocado segundo uma óptica estritamente geográfica – traduz as oportunidades relativas de interacção entre famílias, empresas e indústrias, distribuídas e confinadas a determinado espaço. Reflectindo os custos inerentes à distância em termos monetários, de tempo de acesso à informação, de ajustamentos organizacionais, etc., acaba por marcar as diferenças entre os graus de actividade económica das regiões, quer ao nível da natureza quer ao nível da escala.

Schürmann *et al.* (1997)[1], no trabalho intitulado "Accessibility Indicators", chamam a atenção para a dificuldade em eleger um único indicador de acessibilidade como de importância universal: "(…) on the one hand the selection of accessibility indicators will depend on their statistical interaction with the other transport and non-transport factors of the regional production function; on the other hand the accessibility indicators selected need to be sensitive to the whole range of transport investments and transport infrastructure improvements to be studied (…)". Estes autores referem ainda as dificuldades que persistem em se conseguir que os indicadores de acessibilidade reflictam o nível do serviço de transporte público. Sublinhando, no entanto, que em circunstâncias normais "(…) the main and most general measurement of accessibility change will be identical with the consumer surplus change (…)"; e acrescentam: "(…) there is value in considering the spatial impact of change which would favour use of the information in its accessibility form".

Esta argumentação permite-nos concluir que os indicadores de acessibilidade se tornaram *a pedra de toque* na busca de critérios economicamente significativos para aferição das variações na economia dos territórios. Aliás, a avaliação da perifericidade é, sem dúvida, um parâmetro incontornável no contexto do desenvolvimento económico regional da UE. Como tal – e esgrimidos vários argumentos - como é que a relativa perifericidade/centralidade de cada região podem/devem ser, realmente, avaliadas? Antes de passarmos em revista algumas soluções achamos conveniente enunciar alguns pressupostos de base, consensualmente aceites em diversos trabalhos (Keeble *et al.*, 1988):

- em primeiro lugar, o grau de perifericidade/centralidade de uma região deve ser avaliado relativamente ao volume de actividade económica mantida com outras regiões. Incorporando, sempre que possível, simultaneamente – e entre outras - indicações a respeito: dos mercados com os quais se transaccionam bens e serviços; das fontes e das oportunidades de aquisição - tanto de matéria-prima, como de componentes; e do (eventual) acesso, tanto a informação específica, como a serviços de apoio à actividade comercial;
- em segundo lugar, o índice utilizado deve avaliar a acessibilidade regional face à actividade económica em termos do vasto leque de custos de distância referidos

[1] Citados por Vickerman (2000:6).

anteriormente; evitando estreitá-lo e/ou reduzi-lo à (simples) condição de custos de transporte – tal como entendido na teoria *Weberiana* da localização industrial;

- em terceiro lugar – e de modo a reflectir, tanto o processo de integração económica da UE, como as trocas comerciais intra-comunitárias – o índice deve avaliar a acessibilidade de cada região, relativamente às actividades económicas mantidas (eventualmente) dentro de todo o espaço da Comunidade, sem limitações impostas por qualquer tipo de barreiras;
- simultaneamente, este índice deve ser expressivo da realidade, de interpretação simples, e fácil de calcular.

2 INDICADORES DE ACESSIBILIDADE

2.1 Localização e potencial de mercado

2.1.1 noTa introdutória

Tal como temos vindo a referir, a noção de perifericidade está intimamente ligada à de acessibilidade - conceito (também) não muito simples de definir, mas cujos efeitos são de fácil percepção. Neste sentido, comecemos por analisar o impacto, na economia e no potencial de desenvolvimento das regiões periféricas, de eventuais alterações induzidas ao nível da respectiva acessibilidade - muito embora cientes que melhoramentos apenas a esse nível não significam, automaticamente, um incremento na *performance* regional. Antes pelo contrário: tomados isoladamente, podem tornar essas economias mais permeáveis a mercados externos (mais) competitivos (Silva, 2005).

Em qualquer dos casos, uma questão reúne o consenso geral: para que seja possível capitalizar todos os efeitos positivos, qualquer modificação ao nível das acessibilidades necessita de ser acompanhada de políticas regionais complementares. Ora, o termo Acessibilidade – tal como é frequentemente empregue – encerra (pelo menos) dois conceitos: o de Localização e o de Potencial de Mercado (ou Económico), (Vickerman, 1995); qualquer um deles suficientemente importante para que não passe despercebido na abordagem ao tema que pretendemos efectuar.

2.1.2 Densidade populacional

Na sua forma mais simplista, a designação de Potencial de Mercado está intimamente ligada à noção de Densidade Populacional. Ora, se observarmos a distribuição dos principais centros urbanos na Europa - ou se nos debruçarmos com mais rigor dobre a densidade populacional do Velho Continente – facilmente concluímos que grande parte da população está concentrada no centro da UE, numa faixa territorial que abrange: Londres, Paris, Frankfurt e Amesterdão.

Fora deste núcleo central, emerge um número considerável de regiões metropolitanas como, por exemplo: Glasgow, Copenhaga e Hamburgo – ao Norte, e Munique, Milão e Madrid – ao Sul; as quais, isoladamente, concentram também uma população considerável. Fora deste anel, por sua vez, encontramos outros espaços territoriais (ainda) importantes, sobretudo ao redor de centros metropolitanos como Dublin, Lisboa e Atenas: constituindo áreas de povoamento mais disperso, e embora de menor peso no seio da UE, naturalmente que assumem outra relevância no contexto nacional em que se inserem.

Este quadro de distribuição populacional permite-nos estabelecer uma primeira ligação entre a acessibilidade e a *performance* económica; no entanto, "(…) such a conclusion depends on assuming that both distance and population are unambiguous determinants of access costs and spending power", (Vickerman, 1995:30).

2.1.3 Rendimento local

Se olharmos com atenção para o mapa da distribuição regional do rendimento *per capita*, concluímos que a riqueza parece concentrar-se, de igual modo, nas regiões centrais da UE. Segundo Vickerman (1995), já não é tão só uma questão de mercados alargados, mas sobretudo

de mercados ricos. No entanto, esta imagem não é tão nítida como a que se obtém da distribuição da população; ou seja: não podemos falar de um núcleo central de concentração de riqueza da UE, nem se verifica uma tendência para a diminuição gradual do rendimento *per capita* à medida que nos afastamos do Centro (Vickerman, 1995):

- em primeiro lugar, é de realçar o número significativo de *bolsas* de riqueza concentradas a alguma distância do coração da Comunidade - de entre as quais se destacam regiões como o Sul da Alemanha e o Norte de Itália, ou áreas metropolitanas como as envolventes a Lyon, a Copenhaga, a Viena e a Barcelona;
- em segundo lugar – e simultaneamente – mesmo junto ao núcleo central da Comunidade há uma faixa de regiões pobres, cujo rendimento *per capita* é (ainda) inferior à media da UE: desde o Leste de Londres ao Lorena - passando pela fronteira entre a França e a Bélgica, e ao longo da fronteira da Bélgica e da Holanda, com a Alemanha. Muito embora os problemas que estas regiões hoje enfrentam sejam, provavelmente, um reflexo das convulsões sócio-políticas que atravessaram num passado algo recente, não parece credível que este seja o único argumento para explicar um rendimento *per capita* cerca de 25.0% abaixo da média comunitária, e 50.0% do das regiões metropolitanas vizinhas.

Neste contexto, a proximidade à riqueza não significa, por si só, uma garantia de *desafogo* económico. Aliás, muitos são os casos de espaços territoriais, localizados junto a grandes áreas metropolitanas que vivem, precisamente, na *sombra* da cidade principal. No entanto, é preciso alguma cautela com as generalizações: em situações análogas, as debilidades económicas podem ser, apenas, um artefacto estatístico (Vickerman, 1995) - muito embora determinadas regiões apresentem um baixo nível de receitas, os significativos movimentos pendulares denunciam que grande parte dos seus residentes exerce actividades em regiões metropolitanas vizinhas.

No extremo oposto, as debilidades económicas sentidas por regiões mais periféricas - como as Ilhas Gregas situadas mais a Leste, por exemplo - podem ser o reflexo do afastamento em relação aos centros de riqueza; mas também não podem deixar de espelhar o *déficit* de desenvolvimento dos espaços metropolitanos adjacentes. Neste particular, estamos em crer que o principal obstáculo à melhoria da *performance* regional assenta, muito provavelmente, quer na dependência face a Atenas, quer nas dificuldades de desenvolvimento evidenciadas por toda a economia grega - e não tanto na distância que separa aquelas ilhas dos centros decisionais instalados em Bruxelas ou em Frankfurt. "Inter-regional interdependence on the macro-regional scale may thus be very significant as a determinant of a region's fortunes", (Vickerman, 1995:30).

2.1.4 Tempo de acesso

O rendimento é uma das medidas de *performance* do mercado local; uma outra é, sem dúvida, a acessibilidade física. Mesmo entrando em devida conta com as falhas e a qualidade da rede, a distância rodoviária era utilizada, até há bem pouco tempo, para avaliar o potencial de mercado em qualquer região, sobretudo com base no argumento (económico) do custo de acesso aos mercados. No entanto – e mais recentemente - a distância tem vindo a ser relegada para um plano secundário, devido a alterações entretanto verificadas, quer ao nível da qualidade das redes e da utilização, progressiva, de vários modos de transporte, quer à variedade de bens que passaram a ser transportados.

Vickerman (1995), a propósito, afirma serem estas, precisamente, as razões pelas quais o tempo de acesso é hoje de aceitação mais generalizada; acrescentando que este parâmetro:

- é um indicador neutro face aos diversos modos de transporte;
- pode ser utilizado como um único indicador mesmo em presença – simultânea - de vários modos;
- pode incorporar, directamente, restrições que traduzam falhas da rede ou alterações aos modos; e
- pode reflectir a qualidade (frequência) dos serviços.

Ao tomarmos em consideração o tempo de acesso, facilmente constatamos a alteração que este parâmetro vai operar no tradicional modelo de acessibilidade baseado, unicamente, na distância. Senão, vejamos: determinadas regiões periféricas – insulares, por exemplo – podem resultar beneficiadas se comparadas com outras, em situação análoga, mas localizadas em espaços continentais; sobretudo se as populações das ilhas utilizarem o avião como um meio de transporte regular. Em simultâneo, muitas regiões localizadas junto ao coração da UE não têm, localmente, qualquer aeródromo aberto ao tráfego comercial, estão completamente dependentes dos meios de transporte terrestres para suprirem as suas necessidades de comunicação, e é preciso despender um tempo algo considerável para se deixarem atravessar - por estrada ou ferrovia.

No entanto, esta argumentação não nos permite inferir que os espaços insulares têm (sempre) uma boa acessibilidade: regra geral, eles dependem de um único meio de transporte, com ligações pouco frequentes, e para um leque de destinos (algo) limitado. "More importantly they will be perceived as having poor accessibility. Perception may be the most important factor here, governing the views of both inward investors and potential purchasers worried about reliability of supplies", (Vickerman, 1995:31).

2.1.5 Potencial económico

Tradicionalmente, parâmetros como o Rendimento Local e o Tempo de Acesso estão na base da avaliação do Potencial Económico das regiões. Clark (1969), Keeble *et al.* (1982, 1988), e Vickerman (1996), a propósito, afirmam que o potencial económico é uma medida da área de mercado, resultante do somatório das receitas de outras regiões, e com ajustamento às respectivas distâncias; ou seja:

$$P_i = \sum_j \frac{y_j}{d_{ij}^2} \tag{1}$$

… onde: P_i é o Potencial Económico da região i, y_j é a receita de cada região j, e d_{ij} é a distância entre as regiões i e j.

Embora haja indicações precisas sobre a forma como os resultados devem ser interpretados - e sendo certo que o potencial económico nos permite retirar, *à priori*, informação sobre a dimensão do mercado afecto a cada região - regra geral eles têm vindo a ser utilizados na produção de mapas que, simplesmente, mostram linhas de igual potencial. Também não será de estranhar que seja, precisamente, junto ao coração da UE que este parâmetro apresente os valores mais consideráveis: "(…) since this is close to the highest levels of concentration of income and wealth achieved", (Vickerman, 1995:32).

Mas a questão principal em torno do potencial económico prende-se com o facto de este fazer concentrar os níveis de rendimento em *centróides*, utilizando a distância como factor repulsivo – em consonância com os custos de transporte – dando uma noção (errada) de redução uniforme à medida que nos afastamos do Centro. Além do mais – e muito embora o potencial económico possa reflectir os níveis de rendimento *per capita* das regiões periféricas - não cremos que este parâmetro possa encerrar, em si mesmo, toda a informação relacionada com a riqueza de experiências que tornam único cada território. Neste sentido, em 1988 Keeble *et al.* - no trabalho realizada para a (então) CEC, e invocando outras aplicações nos Estados Unidas da América, na Comunidade Europeia e, inclusivamente, no interior do Reino Unido – propõem uma ligeira mudança na formulação da expressão (1):

$$P_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{m_j}{d_{ij}} + \frac{m_i}{d_{ii}} \tag{2}$$

… onde: P_i é o Índice de Acessibilidade/Perifericidade da região i, m_j é o volume da actividade económica de cada região j, d_{ij} é a distância entre as regiões i e j, e o quociente entre m_i e d_{ii} traduz a contribuição geral do volume da actividade económica da própria região i. A partir desta expressão (2), Keeble *et al.* (1988) tecem – de entre outros – os seguintes comentários:

- m_i e m_j - reflectem o PIB_{pc} da região i e das regiões j, respectivamente;
- d_{ij} - é a distância mais curta, por estrada, entre as principais cidades - ou aglomerados populacionais – das regiões i e j;
- m_i / d_{ii} – ao traduzir a contribuição geral do volume da actividade económica da própria região, este quociente pressupõe uma identificação/avaliação do custo interno de distância (d_{ii}), dando-nos uma estimativa da vitalidade da actividade económica inter-regional; ou seja:

$$d_{ii} = \frac{1}{3}\sqrt{\frac{Área_i}{\pi}} \tag{3}$$

Invocando Rich (1980), Keeble *et al.* (1988) afirmam que a distância assim encontrada é, aproximadamente, um terço do raio de um círculo com a mesma área da região i; traduzindo, na pratica, a realidade de muitas das regiões da Comunidade, onde se observa uma concentração da actividade económica junto aos principais núcleos metropolitanos. "The very marked concentration of high levels of GDP in certain small, highly urbanised regions of the Community, such as Greater London, the West Midlands and Greater Manchester conurbations, Bremen, Hamburg and West Berlin, results in a particularly high intra-regional contribution to the overall peripherality index for these regions, reflecting of course the very considerable possibilities for intra-regional trade which they enjoy", (Keeble *et al.*, 1988:16).

2.2 Tipologias de base

2.2.1 Actividades e impedâncias

"Accessibility is the main "product" of a transport system. It determinates the locational advantage of a region relative to all regions (including itself)", (ME&P, 2001:112). Como tal, os indicadores de acessibilidade devem procurar avaliar os benefícios que as empresas e a população em geral podem colher em determinada região, e resultantes da existência e utilização das infraestruturas de transporte. No entanto, os indicadores de acessibilidade mais simples apenas tomam em consideração as infraestruturas de transporte intra-regionais, e exprimem-se em termos do comprimento total de estradas ou do número de estações de caminho de ferro (Biehl: 1986, 1991); ou, ainda, no tempo de viagem aos nós mais próximos das redes inter-regionais mais significativas (Lutter *et al.*, 1993)[2].

Muito embora este tipo de indicadores contenha informações de muita utilidade sobre a região em si mesma, eles não conseguem incorporar, em rede, os diversos componentes do sistema de transporte: nem os que fazem parte da região nem, tão pouco, os de ligação a outros espaços no exterior (ME&P, 2001). Neste sentido Bökemann[3], em 1982, apresenta uma série de indicadores de acessibilidade que, embora, de formulação mais complexa, tomam já em devida atenção a conectividade das redes de transporte; distinguindo, precisamente, a rede em si mesma – os nós e as respectivas ligações – e as eventuais actividades/oportunidades por ela proporcionadas.

Em termos gerais, a acessibilidade é apresentada dividida em duas funções - cada uma delas representando entidades distintas: por um lado, as actividades/oportunidades a ser alcançadas; e, por outro, o esforço – tempo, distância ou custo - para as alcançar:

[2] Citados por ME&P (2001).

[3] Citado por ME&P (2001).

$$A_i = \sum_j g(W_j) f(C_{ij}) \quad (4)$$

… onde: A_i é representa a acessibilidade da região i, W_j é a actividade que se pretende alcançar na região j, e C_{ij} é o custo total para alcançar a região j a partir da região i; $g(W_j)$ é a função Actividade, e $f(C_{ij})$ é a função Impedância.

Na expressão (4), as funções $g(W_j)$ e $f(C_{ij})$ aparecem-nos sob a forma de um produto - isto é, são pesos uma da outra; demonstrando, claramente, a sua importância – e interdependência – no cálculo de A_i. Neste sentido, podemos afirmar que a Acessibilidade (A_i) de uma determinada região (i) é o resultado do somatório de todas as actividades (j) que a partir dela é possível alcançar, ponderadas pelo facilidade/dificuldade (C_{ij}) nessa ligação. Esta abordagem é mais uma generalização do conceito de Potencial de Stewart (1947). Neste particular, as actividades a alcançar fora da região i funcionam como *atractores*, enquanto as distâncias para as alcançar desempenham o papel de *repulsores*. Ou seja: quanto maior o número de destinos atractivos em regiões j, e quanto mais facilitado lhes for o acesso a partir da região i, tanto maior será o grau de acessibilidade da própria região (i).

Esta definição de acessibilidade é tomada em consideração tendo por base a ideia de Destino. De igual modo, também é possível enunciar uma outra, agora baseada no conceito de Origem: quanto maior for a população que habite nas regiões j, e quanto mais facilmente puder viajar até à região i, tanto melhor sera considerada a acessibilidade da própria região i. No entanto: "Because the symmetry of most transport connections, destination-oriented and origin-oriented accessibility tend to be highly correlated", (ME&P, 2001:113).

Em qualquer dos casos, as funções $g(W_j)$ e $f(C_{ij})$ podem assumir formas diferentes consoante o contexto em que forem utilizadas; e resultado de várias combinações que, desta forma, é possível obter, também a expressão (4) assume aspectos igualmente distintos. O Quadro 1 mostra, precisamente, as combinações mais usuais daquelas funções, donde resultam outros tantos tipos (distintos) de indicadores de acessibilidade.

Quadro 1 – Tipologias dos Indicadores de Acessibilidade

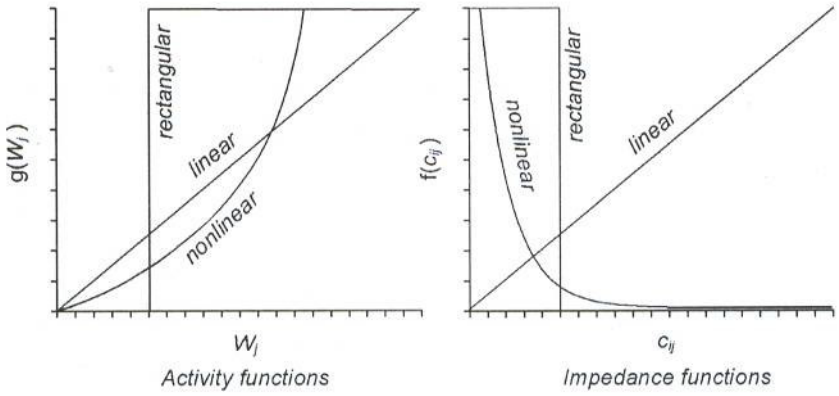
Tipos de Acessibilidade	Função Actividade g (W_j)	Função Impedância f (C_{ij})
Custo da Deslocação	W_j | 1 se $W_j \geq W_{min}$ 0 se $W_j < W_{min}$	C_{ij}
Acessibilidade Diária	W_j	1 se $C_{ij} \leq C_{máx}$ 0 se $C_{ij} > C_{máx}$
Acessibilidade Potencial	W_j^{α}	$\exp(-\beta C_{ij})$

Fonte: Adaptado de ME&P (2001)

… onde: W_{min} e $C_{máx}$ são constantes, e α e β são parâmetros.

A Figura 1 mostra, graficamente, os formatos que as várias funções $g(W_j)$ e $f(C_{ij})$ podem assumir. De notar, precisamente, que os tipos de acessibilidade atrás enunciados resultam de diferentes combinações de funções rectangulares, lineares e não lineares.

Figura 1 – Funções de Actividade e de Impedância Utilizadas em Indicadores de Acessibilidade



Fonte: ME&P (2001:114)

2.2.2 Custo da deslocação

Partindo do pressuposto de que a partir de determinada região nem todos os destinos possíveis de alcançar são importantes para a aferição da respectiva acessibilidade, este indicador permite seleccionar, de entre eles, o conjunto de maior relevância, do qual fazem parte, por exemplo, todas as cidades e/ou regiões acima de um determinado tamanho ou nível de atracção – representado por W_{min}. Por conseguinte, este indicador mede o total dos custos de deslocação, acumulados para um determinado grupo de destinos.

Na sua formulação mais simples, o indicador aceita/pondera de igual modo todos os destinos pertencentes ao grupo de *tamanho* pré-seleccionado, rejeitando todos os outros - situação correspondente à função de actividade rectangular descrita na Figura 1. No entanto, esta não é a regra geral: em muitas situações – a que corresponde uma função de actividade linear, os destinos podem vir a ser ponderados pelo respectivo tamanho. Quanto à função de impedância, ela mantêm-se sempre de características lineares, pois "(…) does not take into account that more distant destinations are visited less frequently", (ME&P, 2001:114):

$$A_i = \sum_j g(W_j)C_{ij}\text{, onde: } g(W_j) = W_j \begin{cases} 1, se W_j \geq W_{\min} \\ 0, se W_j < W_{\min} \end{cases} \quad (5)$$

No entanto – e de forma a torná-lo de mais fácil comparação – este indicador pode transformar-se no custo médio de deslocação para o conjunto de destinos tomados em consideração, bastando, para tal, efectuar o quociente entre a expressão (5) pelo total de destinos, ou deslocações, ou actividades *atractoras*:

$$A_i = \frac{\sum_j g(W_j).C_{ij}}{\sum_j g(W_j)}\text{, onde: } g(W_j) = W_j \begin{cases} 1, se W_j \geq W_{\min} \\ 0, se W_j < W_{\min} \end{cases} \quad (6)$$

Em qualquer dos casos - (5) e (6) – o indicador "(…) expresses a disutility, i.e. the lower its value the higher the accessibility", (ME&P, 2001:114).

Este tipo de indicadores apresenta uma grande vantagem: é de fácil interpretação, sobretudo se traduzido em unidades de uso corrente como, por exemplo, o valor médio do custo e/ou do tempo de deslocação. No entanto – e segundo ME&P (2001) – também enferma de uma desvantagem algo considerável, que acaba por ser basilar: não consegue traduzir o facto de que os destinos mais longínquos são, regra geral, visitados com menor frequência; por conseguinte,

o resultado final está dependente, em grande medida, da escolha inicial do conjunto de destinos - isto é, do livre arbítrio com que se escolhe W_{min}, o limiar do parâmetro W_j.

2.2.3 Acessibilidade diária

Este indicador tem por base dois pressupostos, ambos subjacentes ao simples acto de viajar entre uma origem e um destino: existe uma razão plausível para tal deslocação; e ela é justificável desde que possa ser efectuada dentro de um intervalo de tempo, considerado aceitável.

Na génese deste indicador está o conceito apresentado por Törnqvist, em 1970[4], segundo o qual um empresário só se disporá a efectuar uma viagem de negócios no mesmo dia, se a distância entre a origem e o destino puder ser percorrida, em cada sentido, no espaço de três a cinco horas. Segundo ME&P (2001), é precisamente por esta comparação com a viagem de negócios no próprio dia que este indicador toma a designação de de "Acessibilidade Diária":

$$A_i = \sum_j W_j f(C_{ij}) \text{, onde: } f(C_{ij}) = \begin{cases} 1, se C_{ij} \leq C_{máx} \\ 0, se C_{ij} > C_{máx} \end{cases} \quad (7)$$

… onde: $C_{máx}$ é o limiar do tempo de deslocação.

O indicador de acessibilidade diária é um caso particular do de acessibilidade potencial - que analisaremos em seguida; ou seja: de acordo com o tempo de deslocação pré-definido, os destinos são ponderados pelo respectivo tamanho, não sendo considerados quaisquer outros para além desse limite (temporal). Em termos gráficos – e tal como resulta da Figura 1 – estamos em presença de uma função de actividade linear, conjugada com uma função de impedância rectangular. Este tipo de indicadores – e tal como sucede com o anterior de custo/tempo de deslocação – tem a vantagem de poder expressar-se em termos muito acessíveis; como por exemplo – e no caso vertente – na população que é possível alcançar em determinado espaço de tempo. No entanto, partilham algumas desvantagens: à semelhança do que sucede com aquele em relação ao parâmetro W_{min}, também este indicador depende, em grande medida, do livre arbítrio com que se escolhe o limiar de tempo máximo de deslocação – $C_{máx}$; valor a partir do qual todos os destinos são rejeitados.

2.2.4 Acessibilidade potencial

Segundo ME&P (2001), este indicador baseia-se no princípio de que a atracção produzida por determinado destino é directamente proporcional ao respectivo *tamanho*, e inversamente proporcional à distância - ou ao custo/tempo de deslocação. Neste sentido, especial atenção passa a ser dada, quer ao tamanho do destino, quer à distância a que ele se encontra. O tamanho do destino é, regra geral, representado pela população local, ou por um qualquer indicador económico regional como, por exemplo, o PIB ou o Rendimento.

Neste indicador, a função actividade pode ser tanto linear como não linear. No entanto, também é possível encontrar o elemento *atractor* W_j ponderado por um expoente α superior à unidade, "(…) to take account of agglomeration effects, i.e. the fact that larger facilities may be disproportionally more attractive than smaller ones", (ME&P, 2001:115).

Quanto à função impedância, ela assume características não lineares. Aliás – e regra geral – ela aparece sob a forma de uma exponencial negativa, onde um parâmetro β se encarrega de atribuir um peso mais significativo aos destinos mais próximos, em detrimento dos mais afastados:

$$A_i = \sum_j W_j^{\alpha} \exp(-\beta . C_{ij}) \quad (8)$$

[4] Citado por ME&P (2001).

De acordo com ME&P (2001:115), as primeiras versões deste indicador "(…) used an inverse power function reminiscent of Newton's gravity model. (…) This form was proposed by Hansen as early as 1959 and is therefore called "Hansen" accessibility":

$$A_i = \sum_j \frac{W_j}{C_{ij}^{\alpha}} \tag{9}$$

Este tipo de indicadores - porque alicerçado em princípios estocásticos de maximização da utilidade – é de maior robustez se comparado com os anteriores. No entanto, também possui desvantagens algo consideráveis: por um lado, na existência de parâmetros que é necessário calibrar; e, por outro, os resultados obtidos nem sempre são de fácil interpretação como acontece, por exemplo, com o tempo de deslocação ou com a população a alcançar. "Therefore, potential indicators are frequently expressed in percent of average accessibility of all regions or, if changes of accessibility are studied, in percent of average accessibility of all regions in the base year of the comparison", (ME&P, 2001:116).

2.3 Outros contributos

ME&P (2001) – citando Ruppert (1975) – assegura que a partir das três configurações básicas acabadas de referir derivam uma quase infinidade de versões de indicadores de acessibilidade; sendo os mais importantes susceptíveis de agrupar da seguinte forma:

- Acessibilidade Multimodal – os indicadores de acessibilidade modal podem ser apresentados separadamente, de forma a evidenciarem as diferenças entre modos de transporte. Ou, alternativamente, distintas acessibilidades modais são integradas num mesmo indicador para que este possa expressar, em determinado local, o efeito combinado de modos alternativos;
- Acessibilidade Intermodal – focando especial atenção em dois ou mais modos, tomam em consideração as viagens intermodais que, eventualmente, se estabelecem entre eles. Estes indicadores são os mais indicados para aplicar em cadeias logísticas de transporte de mercadorias, com diferentes combinações possíveis de modos de carga e de terminais; como, por exemplo: comboios de mercadorias em conjugação com camiões, ou – no caso de transporte de passageiros – "(…) mode combinations such as rail-and-fly or car access to railways", (Spiekermann e Neubauer, 2002:10);
- Acessibilidade Intra-regional – debruçados, em particular, sobre a distribuição irregular da acessibilidade em determinada cidade e/ou região;
- Acessibilidade com Base na Actividade e no Actor – tomam em consideração, para determinado local, os diferentes requisitos de acessibilidade, tanto para as actividades como para os actores directamente envolvidos.

Rietveld e Bruinsma (1998:33), a propósito desta diversidade, afirmam que a acessibilidade é específica "(…) per type of user and per trip purpose", isto é: "(…) accessibility becomes trip purpose specific"; e enumeram até onze possibilidades distintas para a operacionalização/definição deste indicador - cada uma delas recorrendo, geralmente, a informações muito precisas sobre a localização dos *nós* na *rede*, as distâncias entre eles, e os respectivos custos de transporte.

Na sequência dos trabalhos conduzidos precisamente daqueles autores também ME&P (2001) se refere a alguns casos particulares de indicadores de acessibilidade aos quais não podemos deixar de prestar atenção. Sobretudo por uma razão fundamental: temos vindo a enunciar algumas das abordagens que, no contexto da UE – ou mesmo da Europa – têm servido de base para a construção de indicadores de acessibilidade. Os quais – e de uma forma geral - se podem dividir em dois grupos: os de âmbito local e os de âmbito interurbano/inter-regional. Enquanto uns são puramente de carácter topológico, outros tomam em consideração a dispersão das próprias actividades. No entanto, ME&P (2001:61) chama a atenção, muito em particular, para a falta de um método de cálculo dos indicadores de acessibilidade que reflictam "(…) the latest GIS developments and could also include social weights". Neste sentido – e após cuidada revisão de todas as alternativas – ele acaba por apontar três soluções:

- O indicador Direccionado ("Pointer") – de aplicação no contexto do transporte de passageiros, e que tem a seguinte formulação:

$$POINTER_i(TL,TT) = \sum_j \left[POP_j (TL - 2T_{ij}) \right] \quad \text{, com: } T_{ij} \leq TT \tag{10}$$

… onde:
TL – total/limite de tempo, que inclui: o das viagens entre a origem (i) e o destino (j) - em ambos os sentidos - e o necessário para que, no destino, seja possível desenvolver a actividade pretendida;
TT – limite para o tempo de deslocação num único sentido; caso este valor seja excedido, o respectivo destino é excluído;
POP_j – população que é possível alcançar no destino j;
T_{ij} – tempo de ligação entre i e j.
Este indicador pode ser utilizado, por exemplo, para avaliar a possibilidade de um empresário, que residindo numa região (i), pretenda alcançar um vasto leque de clientes distribuídos por outras regiões (j), mas tendo como restrição determinado limite de tempo, num dia (normal) de trabalho.

- O indicador de Equidade ("Equity") – goza de algumas particularidades em relação ao anterior; nomeadamente: não entra em linha de conta com o tempo *gasto* no destino; e atribui um peso mais elevado à população que se encontra na proximidade, reduzindo-o consideravelmente à medida que o tempo de viagem (à população de destino) aumenta:

$$EQUITY_i(TT) = \frac{\sum_j \frac{POP_j}{e^{\sqrt{T_{ij}}}}}{\sum_j POP_j} \quad \text{, com: } T_{ij} \leq TT \tag{11}$$

- O indicador de Coesão ("Cohesion") – de utilização mais generalizada na avaliação da acessibilidade em situações de longa distância, este indicador mede o potencial de interacção entre pessoas; ou seja: avalia o número total de pessoas que podem ser alcançadas em determinado intervalo de tempo (limite) de viagem:

$$COHESION_i(TT) = POP_i \sum_j POP_j \quad \text{, com: } T_{ij} \leq TT \tag{12}$$

Face a determinado sistema de transporte, este indicador avalia o conjunto de oportunidades que é possível alcançar num certo espaço territorial; identificando, consequentemente, as regiões (periféricas) menos acessíveis a tais oportunidades.

2.4 Proposta de síntese

Tal como temos vindo a constatar, as transformações económicas e sociais das últimas décadas têm servido de cenário – e de justificação – para o "(…) desenvolvimento de uma multiplicidade de indicadores que, sob a denominação geral de [de] acessibilidade, têm utilizações diversas, (…) não possuem os mesmos objectivos, e não reflectem os mesmos fenómenos", (Figueira, 2001:10). Bökemann *et al.* (1997)[5] – a este propósito – sistematizam tais preocupações do seguinte modo:

- numa primeira fase, os trabalhos sobre a acessibilidade reportavam-se à distância entre dois quaisquer pontos de uma zona de referência; ou seja: a acessibilidade de um ponto dependia da sua posição face aos restantes pontos do sistema. Nesta conformidade, a

[5] Citados por Figueira (2001).

acessibilidade era sinónimo de proximidade, quer em termos de distância – em sentido estrito, quer em termos de custo e/ou tempo – em sentido lato;

- em trabalhos posteriores, a acessibilidade passou a ser encarada como o número de oportunidades para participar em determinadas actividades; conceito que incorporava, (já) não só a rede de transportes, mas também a necessidade de analisar o perfil dos utilizadores, e de quantificar a atractividade dos vários destinos em potência;
- recentemente, a acessibilidade surgiu como um benefício acrescido para um grupo de pessoas que se encontram em determinado contexto, e que podem utilizar um sistema de transportes específico.

Figueira (2001) socorrendo-se, precisamente, deste (amplo) conceito de *distância social* - e assumindo que se trata de uma medida de proximidade entre actividades humanas que permite aquilatar da facilidade em as alcançar a partir de um certo local e utilizando determinado meio de transporte - deduz que as principais componentes de um qualquer indicador de acessibilidade são as seguintes:

- a Base a que se aplica - acessibilidade de um ponto, ou entre dois pontos;
- a medição da Distância entre pontos;
- a Conexão dos pontos com o sistema de comunicações – permitindo *cobrir* a distância entre eles;
- o Esforço – distância, tempo, custo, etc. – necessário para cobrir tal distância;
- a Relação com o tipo de actividade na qual os utilizadores querem participar.

Neste contexto, também podemos afirmar que os indicadores de acessibilidade diferem no respectivo grau de complexidade; ou seja:

- os mais simples tomam em atenção, unicamente, as infraestruturas de transporte da própria região, avaliando: o comprimento das estradas e/ou das linhas de caminho de ferro; o número de auto-estradas e/ou de estações de caminho de ferro; ou o tempo de ligação aos nós mais próximos das redes de transporte de nível superior. Este tipo de indicadores – embora portador de uma informação muito preciosa sobre a própria região – não reflecte, contudo, a realidade: muitas das actividades mais importantes são efectuadas, precisamente, fora da região em apreço;
- os mais complexos avaliam a conectividade das redes de transporte, fazendo uma distinção entre a rede em si mesma e as actividades/oportunidades que através dela é possível alcançar; incluindo na sua formulação um termo de impedância que descreve a facilidade/dificuldade em tal ligação, expressa em termos de distância, de tempo, ou de outra qualquer medida de custo.

Figueira (2001), apoiando-se, sucessivamente, nos trabalhos de Izquierdo e Monzon (1992), Chatelus (1997) e Viegas e Figueira (1999), propõe uma divisão geral dos indicadores de acessibilidade da seguinte forma: Topológicos, de Ponderação Agregada, e Bilaterais; reservando algum destaque, em último lugar – mas ainda a este propósito – ao conceito de Interconectividade.

3 CONSIDERAÇÕES FINAIS

O principal objectivo de uma qualquer infraestrutura de transportes é o de permitir a interacção entre espaços territoriais; ou seja, o de facilitar a mobilidade de pessoas e de mercadorias, na prossecução de várias actividades: sociais, culturais e económicas (Silva, 2005).

No plano do desenvolvimento territorial, a qualidade de tal infraestrutura – em termos de capacidade, conectividade, velocidade de deslocação, etc. – é fundamental na determinação, também da qualidade de localizações alternativas; isto é, da vantagem competitiva de uma localização sobre as demais, e em cuja avaliação entram os diversos conceitos/indicadores de acessibilidade que temos vindo a analisar. Neste contexto, os indicadores de acessibilidade assumem uma importância acrescida face a um diversificado leque de características espaciais,

de entre as quais Wegener *et al* (2000, 2002)[6] destacam: Origens, Destinos, Impedâncias, Restrições, Barreiras, Tipos de Transporte, Modos, Escalas, Equidade e Dinâmicas (Quadro 2).

Quadro 2 – Importância Espacial dos Indicadores de Acessibilidade

Característica Espacial	Importância do Indicador de Acessibilidade
Origem	Os indicadores de acessibilidade podem ser avaliados tomando em consideração diferentes grupos populacionais (sociais ou etários), diferentes ocupações (negócios, viajantes ou turistas), e diferentes actores económicos (industrias ou empresas).
Destino	Os indicadores de acessibilidade podem avaliar determinado espaço territorial face a oportunidades, actividades e bens, tais como: população, actividades económicas, universidades ou atracções turísticas. A função de actividade pode ser rectangular (todas as actividades até determinado volume), linear (todas as actividades de determinado volume), ou não linear (de forma a que possa expressar efeitos de aglomeração).
Impedância	O termo que representa a função de impedância espacial pode ser uma função de um ou vários atributos inerentes às ligações entre as áreas em apreço, tais como: distância (Euclidiana ou distância em rede), tempo de deslocação, custo da viagem, conveniência, confiança ou segurança. A respectiva função de impedância pode ser rectangular (todos os destinos dentro de valores de impedância pré-definidos), linear (com base no valor médio), ou não linear (p.ex., sob a forma de uma exponencial negativa).
Restrição	As ligações entre espaços territoriais podem estar sujeitas a limitações impostas, ou por regulamentos específicos (limites de velocidade, restrições de acesso a determinado tipo de veículos, ou a um limite máximo de horas de condução), ou por restrições de capacidade (gradientes rodoviários, ou determinados níveis de congestionamento).
Barreira	Em complemento às impedâncias que se estabelecem nas ligações entre territórios, há que tomar em consideração, de igual modo, quer as barreiras imateriais (p.ex., políticas, económicas, legais, culturais ou linguísticas), quer as ligações imateriais (p.ex., regras – implícitas e explícitas – na articulação do tecido empresarial).
Transporte	São tomados em consideração as viagens de carácter pessoal, ou o transporte de mercadorias, ou ambos.
Modo	Os indicadores de acessibilidade podem ser calculados tendo por base os modos: rodoviário, ferroviário, marítimo ou aéreo. Os indicadores de acessibilidade multimodais (abrangendo as deslocações efectuadas recorrendo a mais do que um modo) congregam vários indicadores de acessibilidade modais.
Escala	Os indicadores de acessibilidade às escalas continental, transnacional, ou regional, podem necessitar de informações com diferentes resoluções espaciais; nomeadamente ao nível: das dimensões das áreas e das representações das redes, dos acessos em cada área e dos terminais em cada nó, e dos tempos de transbordo.
Equidade	Os indicadores de acessibilidade podem ser aplicados a determinados espaços territoriais de forma a poderem ser identificadas as desigualdades de níveis de acessibilidade entre áreas: ricas e pobres, centrais e periféricas, urbanas e rurais, nodais e intersticiais.
Dinâmica	Os indicadores de acessibilidade podem ser calculados em momentos distintos, para que seja possível avaliar eventuais alterações e impactes de acessibilidade induzidas, quer por investimentos em infraestruturas, quer pela implementação de políticas de transporte.

Fonte: Adaptado de Spiekermann e Neubauer (2002)

No entanto, este precioso meio de diagnóstico complementar, do ordenamento e do desenvolvimento territorial, constituído pelo vasto leque de indicadores de acessibilidade colocados à nossa disposição, não nos pode fazer esquecer o caminho ainda a percorrer no sentido de os tornar (cada vez) mais consentâneos com a realidade. Neste sentido, Vickerman (1995) alerta para a necessidade de que os trabalhos de investigação subsequentes continuem a não descurar, especificamente, três áreas:

[6] Citados por Spiekermann e Neubauer (2002).

- Acessibilidade Intra e Inter-regional – é necessário combinar a qualidade dos sistemas de transporte com os padrões de qualidade das próprias regiões, para que seja possível avaliar o verdadeiro alcance das (eventuais) alterações de acessibilidade em todo o território em apreço: "There is a danger that we concentrate on the accessibility of major gateways when the real constraint on development is internal access to those gateways", (Vickerman, 1995:37);
- Papel do Transporte de Mercadorias – a acessibilidade pode (agora) ser avaliada, muito facilmente, em termos de tempo, sobretudo se aplicada ao transporte de passageiros; como tal – e em algumas situações – a acessibilidade relativamente ao transporte de mercadorias parece ter sido relegada para um plano secundário. Curiosamente, a partir do momento em que as soluções multimodais se tornaram mais comuns, também o transporte de mercadorias passou a enfrentar algumas dificuldades "(…) particularly interesting […], and although freight shippers are often considered to be excessively committed to a single mode, they are nevertheless more likely to be well-informed and influenced by changes in both accessibility and cost than many passengers", (Vickerman, 1995:38);
- Estabilização/Confiança – estes conceitos, abordados, simultaneamente, na perspectiva dos custos inerentes à própria rede e na da elaboração dos *stocks* de mercadorias, devem acompanhar as preocupações subsequentes, quer em matéria de acessibilidade intra e inter-regional, quer na discussão sobre o papel do transporte de mercadorias. Tal como oportunamente constatamos, tais custos tornaram-se, regra geral, de mais fácil avaliação qualitativa do que quantitativa; aliás, em situações em que o transbordo de mercadorias se efectua segundo um plano pré-determinado, eles podem nem sequer conduzir a valores excessivos. Em casos onde, por exemplo, os camiões têm de efectuar travessias por meio de *ferry-boats*, tais momentos podem coincidir com os períodos de descanso impostos pelos/aos condutores - donde os respectivos custos de tempo de transporte se tornam mais pequenos do que inicialmente previsto. Neste contexto – e tal como acontece em muitas situações - a graduação dos custos não depende da rede em si mesma, mas do trajecto escolhido. Vickerman (1995), a propósito, afirma que os custos de tempo inerentes à travessia do Túnel da Mancha são muito maiores nos trajectos (mais curtos) entre Londres e Paris (ou Bruxelas), do que nas travessias (mais longas) entre Manchester e Milão. "Where this impinges on firms is not typically through the level of transport costs, which tend to be relatively low in relation to total costs, but through the effect which variations in these costs (especially unpredictable variations) have on the need for higher levels of inventory holdings than would otherwise be necessary (especially where just-in-time production procedures are in use)", (Vickerman, 1995:38).

REFERÊNCIAS BIBLIOGRÁFICAS

FIGUEIRA, F. (2001) *Medição da Acessibilidade no Contexto do Desenvolvimento Regional. Alterações de Acessibilidade Regional na Península Ibérica*, Tese de Mestrado, Lisboa, Instituto Superior Técnico.

KEEBLE, D., J. OFFORD e S. WALKER (1988) *Peripheral Regions in a Community of Twelve Member States*, Luxembourg, Office for Official Publications of the European Communities.

KEEBLE, D., P. OWENS e C. THOMPSON (1982) Regional Accessibility and Economic Potential in the European Community, *Regional Studies*, 16, 419-432.

ME&P (Marcial Echenique & Partners, Ltd) (ed) (2001) *EUNET/SASI Report*, Luxembourg, Office for Official Publications of the European Communities.

RIETVELD, P. e F. BRUINSMA (1998) *Is Transport Infrastructure Effective? Transport Infrastructure and Accessibility: Impacts on the Space Economy*, Berlin, Springer-Verlag.

RODRIGUES, C., E. CASTRO, R. CARVALHO, D. POLÓNIA, M. DUARTE e D. TRIGO (1996) Serviços Avançados de Telecomunicações nas Regiões Menos Favorecidas. O Projecto MOSAICC, in

Associação Portuguesa para o Desenvolvimento Regional (ed), *IV Encontro Nacional da APDR*, Covilhã, Universidade da Beira Interior.

SILVA, J. (2005) *As Acessibilidades como Factor do Desenvolvimento de Regiões Periféricas: o Caso da Beira Interior*, Tese de Doutoramento em Transportes, (ISBN: 972-8790-54-6, UBI, 2006), Lisboa, Instituto Superior Técnico.

SPIEKERMANN, K. e J. NEUBAUER (2002) *European Accessibility and Peripherality: Concepts, Models and Indicators*, Stockholm, Nordregio.

STEWART, J. (1947) Empirical Mathematical Rules Concerning the Distribution and Equilibrium of Population, *Geographical Reviews*, 37, 461-485.

VICKERMAN, R. (1995) Accessibility and Peripheral Regions, in H. Coccossis e P. Nijkamp (eds), *Overcoming Isolation. Information and Transportation Networks in Development Strategies for Peripheral Areas*, Heidelberg, Springer-Verlag, 29-40.

VICKERMAN, R. (1996) Location, Accessibility and Regional Development: the Appraisal of Trans-European Networks, *Transport Policy*, 2, 4, 225-234.

VICKERMAN, R. (2000) Core-Periphery and the Cohesion-Competitiveness Debate, in Regional Science Association International (ed), *6th World Congress of the RSAI*, Lugano, RSAI.

A ACESSIBILIDADE COMO UM FACTOR CHAVE DO DESENVOLVIMENTO DO TURISMO EM REGIÕES PERIFÉRICAS. O CASO DA BEIRA INTERIOR.

JOANA FIDALGO VAZ REIS SILVA
JORGE MIGUEL DOS REIS SILVA
E-mail: jmiguel@ubi.pt
MARGARIDA MARIA FIDALGO DA COSTA VAZ
E-mail: mvaz@ubi.pt
Universidade da Beira Interior

Resumo

Turismo e Acessibilidades estão interligados. Por um lado o turismo facilita e (não raramente) estimula a mudança das acessibilidades e, por outro lado, as acessibilidades abrem novas perspectivas de destinos turísticos contribuindo para o desenvolvimento de novos (e correlacionados) modelos de negócio. A Beira Interior está localizada perto da fronteira com Espanha e é encarada normalmente como uma região periférica. Nos últimos anos a debilidade da indústria do turismo tem vindo a ser justificada com base na falta e/ou na qualidade das acessibilidades sobretudo as relativas à infra-estruturas de transporte. Afim de aferir se efectivamente a debilidade do sector do turismo na Beira Interior pode (ou não) ser imputada à falta e/ou à qualidade das infra-estruturas de transporte desenvolvemos uma metodologia que permita identificar os ganhos de acessibilidade no passado recente, os previsíveis num horizonte temporal aceitável em termos de planeamento, e os ainda possíveis num cenário assimptótico de forte acessibilidade generalizada, permitindo, desta forma, explicitar os ganhos de acessibilidade já alcançados e os ainda possíveis de alcançar. Esta metodologia permite não só evidenciar quer as direcções geográficas quer as regiões parceiras com as quais a Beira Interior poderia vir a obter contribuições significativas no desenvolvimento do seu potencial turístico, mas também perspectivar a futura instalação na região de novos modelos de negócio associados a esta actividade económica, em condições de alta competitividade.

Palavras-chave: Acessibilidades, Turismo, Regiões Periféricas, Desenvolvimento Económico.
Área Temática: Economia da Energia, do Ambiente, dos Transportes e dos Recursos Naturais.

Abstract

Tourism and accessibilities are interlinked. On one hand tourism is a driving factor for and (often) a stimulator of change in accessibilities; on the other hand accessibilities themselves opened new touristic destination perspectives even helping to develop locally correlated new business models. Beira Interior is located close to the Spanish border and traditionally seen as a peripheral region. In the last years the weaknesses of the tourism industry have been justified on the basis of shortness and/or lack of quality in accessibilities, mainly those concerning transport infrastructures. In order to evaluate whether there is in fact a case in favour of accessibility shortage we developed a methodology that would allow us to identify the accessibility gains in the recent past, the ones foreseeable in the usually adopted planning periods and the ones possible in a asymptotic scenario of strong generalised accessibility, enabling this way to make explicit identification of the gains already achieved and the ones still possible. Finally, this methodology allows evidencing either geographic directions or partner regions for which Beira Interior could get significant contributions in the development of its tourist potential, also targeting the future installation under high competitiveness conditions of new business models associated with this economic activity.

Key-words: Accessibilities, Tourism, Peripheral Regions, Economic Development.
Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

1. INTRODUÇÃO

A Beira Interior[1] é uma região localizada no centro do país junto à fronteira com Espanha e considerada por muitos como periférica. Partindo da sua marginalidade geográfica relativamente aos centros nacionais e internacionais mais desenvolvidos, as acessibilidades[2], nomeadamente as infra-estruturas de transporte, assumem um papel estratégico na demolição daquilo que poderíamos chamar "barreiras geográficas à entrada".

De acordo com Silva (2005) a disposição do relevo e os principais cursos de água condicionaram a estruturação das principais vias de comunicação (internas e externas) da Beira Interior. Também o atraso que sistematicamente se tem vindo a verificar na beneficiação de alguns dos itinerários principais e complementares que atravessam a região, inviabiliza a imprescindível permeabilização do território quer entre o Norte e o Sul, por um lado e, por outro, entre o Litoral português e o Interior espanhol. Quanto à rede interna, *capilar*, não tem merecido a devida atenção por parte das entidades competentes pese embora ela ser duplamente importante para o desenvolvimento integrado da região (Silva, 2005): internamente – nas ligações mais *finas* entre localidades, mas também nas ligações estratégicas ao território espanhol – numa lógica de cooperação transfronteiriça que urge preservar e incrementar.

Pese embora estas debilidades, nos últimos anos a situação evoluiu favoravelmente[3], sendo de esperar que a região fique dotada das ligações verdadeiramente importantes após concluídos todos os projectos rodoviários, ferroviários e aéreos (em obra e planeados). Se durante muitos anos a escassez e/ou a falta de qualidade das infra-estruturas de transporte foi um dos argumentos para justificar os baixos índices demográficos e a fraca dinâmica económica desta região, a concretização daqueles projectos fará com que tal justificação deixe de ter sentido.

Turismo e acessibilidades estão interligados. Por um lado, o turismo é um factor pull de acessibilidades e (muitas vezes) um factor push estimulador de mudança em acessibilidades; por outro lado as acessibilidades abrem novos destinos e novas perspectivas para destinos turísticos já existentes, podendo ajudar a desenvolver localmente novos modelos de negócios correlacionados.

Como actividade económica e social, o turismo está naturalmente integrado num quadro de desenvolvimento socioeconómico global, sendo reflexo, e simultaneamente reflectindo, as particularidades de um determinado estádio de desenvolvimento (Vaz, 2005). Sendo uma actividade fortemente ancorada no território, a interligação do turismo com o sistema económico local dependerá necessariamente da estratégia de desenvolvimento da região onde se ancora. A estrutura socioeconómica e as ligações inter-sectoriais existentes irão determinar o potencial do turismo na geração de efeitos multiplicadores sobre o rendimento, o emprego e o consumo dessa região.

À semelhança de outras actividades económicas, também os pontos fracos do turismo nesta região têm sido justificados com base na localização periférica da Beira Interior, muitas vezes associada à escassez de acessibilidades e/ou à falta de qualidade delas, sobretudo as relativas à infra-estruturas de transporte.

A fim de avaliar se há, de facto, um caso de escassez de acessibilidades foi utilizada uma metodologia desenvolvida por Silva (2005) que permite identificar os ganhos de acessibilidade ocorridos no passado recente, os ganhos previsíveis num horizonte temporal considerado aceitável (em termos de planeamento) e os ganhos possíveis num cenário assimptótico de forte acessibilidade generalizada, permitindo desta forma tornar explícita a identificação dos ganhos de acessibilidade já alcançados e os ainda possíveis de alcançar.

Tirando partido da localização geográfica da Beira Interior relativamente ao litoral e, principalmente, a Espanha, este método permite identificar quer direcções geográficas quer

[1] Região definida pelo conjunto das quatro NUT III da região Centro de Portugal - Serra da Estrela, Beira Interior Norte, Beira Interior Sul e Cova da Beira, que corresponde genericamente aos Distritos da Guarda e de Castelo Branco.

[2] Compreendendo o termo "acessibilidades" quer as infra-estruturas de transporte quer as de telecomunicações, neste trabalho será dada ênfase às primeiras. Tem-se no entanto presente a importância que actualmente as telecomunicações assumem na vida socioeconómica de qualquer região e também a sua crescente importância para o turismo.

[3] Construção da A.23– Auto-Estrada entre Torres Novas e Guarda, e da A.25 - eixo transversal entre Aveiro e Vilar Formoso, passando pela Guarda, e que assegura a ligação com a fronteira espanhola.

regiões parceiras com as quais a Beira Interior pode obter contribuições significativas no desenvolvimento do seu potencial turístico, visando também a futura localização de novos modelos de negócio associados a esta actividade económica, sujeita cada vez mais a grandes pressões competitivas.

2. A CRESCENTE PROCURA TURÍSTICA DE ZONAS DO INTERIOR

De acordo com um estudo feito por Vaz e Dinis (2007), as zonas do litoral português[4], continuam a ser as mais procuradas pela maioria dos portugueses e estrangeiros, nelas se registando os valores mais elevados para todos os indicadores analisados, quer do lado da oferta quer da procura. Contudo, há indícios de alteração da dinâmica turística nacional com as zonas do litoral a crescerem a um ritmo mais lento (3,6%) do que as zonas do interior (4,3%) que pode ser consequência quer da saturação de alguns destinos tradicionais quer de eventuais alterações nas motivações turísticas, considerando que os aspectos naturais e culturais destas regiões do interior constituem a base da sua oferta turística.

A crescente procura dos destinos do interior e os níveis de saturação turística baixos (em particular a região Centro) dão a estas regiões maiores potenciais de crescimento. Contudo, elas são maioritariamente procuradas por portugueses, verificando-se um período de permanência dos turistas bastante reduzido, o que evidencia um tipo de turismo mais vocacionado para curtas estadias ou associado ao conceito de *touring*.

No que respeita à região da Beira Interior, o quadro seguinte põe em evidência duas grandes debilidades: as baixas taxas de ocupação-cama dos estabelecimentos hoteleiros[5] e a grande concentração no mercado turístico português.

Quadro 1 – Nº Hóspedes, Estada Média e Taxa Ocupação Beira Interior (dados de 2007)

	Hóspedes				Estada média	Taxa ocupação-cama (líquida)
	Portugueses	Estrangeiros	Total	% Estrangeiros		
Serra da Estrela	25134	1833	26967	7%	1,8	27,0
Beira Interior Norte	83044	18232	101276	18%	1,3	21,6
Beira Interior Sul	68744	10730	79474	14%	1,8	29,7
Cova da Beira	122676	9221	131897	7%	1,7	32,3
Total	**299598**	**40016**	**339614**	**12%**		

Fonte: Elaboração própria com base no Anuário Estatístico da Região Centro, INE (2007)

No mesmo período de tempo (ano de 2007), os estabelecimentos hoteleiros[6] das vizinhas regiões espanholas de Castilla y Léon e Extremadura receberam, no seu conjunto, cerca de 5,7

[4] Principalmente o Algarve, o Norte Litoral e a Madeira.

[5] A falta de informação estatística em Portugal sobre o número de hóspedes dos alojamentos em turismo rural, inibe a comparação com as regiões espanholas. Segundo a Encuesta de Ocupación en Alojamientos Turísticos Extrahoteleros publicada pelo Instituto Nacional de Estadística (INE), no mês de Janeiro de 2009, a comunidade autónoma de Castilla Y León voltou a ser o destino preferido em alojamento de turismo rural em Espanha. Esta comunidade regista 18,29 por cento do total de dormidas do país neste tipo de estabelecimentos. Como refere Vaz (2005) as unidades de Turismo em Espaço Rural têm vindo a afirmar-se no contexto regional, não só como alternativa de qualidade aos alojamentos tradicionais, como também em resposta às novas motivações da procura. Esta oferta de alojamento tem toda a oportunidade numa paisagem marcadamente rural e onde ainda se podem encontrar casas e solares de grande valor histórico e/ou arquitectónico, para além de todo um conjunto de edificações tradicionais ao abandono, à espera de serem objecto de reconstrução e decoração cuidadas e utilizadas para fins turísticos, à semelhança, aliás, do que vem acontecendo na vizinha Espanha, onde também as unidades de Turismo em Espaço Rural ganham cada vez maior expressão.

milhões de hóspedes, dos quais cerca de 21% estrangeiros. Similarmente à Beira Interior, também os hóspedes nestas regiões espanholas permanecem durante pouco tempo; contudo o elevado nº de hóspedes contribuirá de alguma forma para que as taxas de ocupação-cama sejam mais elevadas do que as verificadas na região portuguesa.

Quadro 2 – Nº Hóspedes em Castilla Y León e Extremadura (dados referentes a 2007)

	Hóspedes				Estada média	Taxa ocupação-cama
	Espanhóis	Estrangeiros	Total	% Estrangeiros		
Castilla y León	3691966	1019021	4710987	21,63%	1,76	37,15
Extremadura	830308	151022	981330	15,39%	1,77	34,63
Total	**4522274**	**1170043**	**5692317**	**20,55%**		

Fonte: Elaboração própria a partir do Instituto Nacional de Estadística (2009)[7]

Vaz (2005: 111) citando Gómez (1998) refere que a crescente mobilidade associada ao lazer e ao turismo traduz-se nas crescentes deslocações aos espaços urbanos pelo período de um dia ou por um período normalmente não superior a cinco dias, por motivos culturais, recreativos ou profissionais, sendo de notar que o turismo de curta duração cresceu 3,5 vezes mais que as férias de longa duração.

Segundo Hall (2004) em consequência das inovações tecnológicas no sector dos transportes, as viagens de um só dia estão a tornar-se cada vez mais importante no alargamento das escalas espaciais, tornando ainda mais ténue a fronteira entre turismo, lazer ou migrações. De acordo com Hall *et al*. (2004) o turismo é apenas uma forma de lazer orientado para a mobilidade temporária e constitui parte dessa mobilidade, ambas sendo moldadas por e simultaneamente moldando as práticas contemporâneas de consumo, produção e estilos de vida. O turismo contextualiza-se na mobilidade que caracteriza as sociedades actuais (Lebre, 1998), exigindo mudança de lugar no espaço (Pearce & Butler, 1993)[8].

Este facto reporta para Ritchie *et al.* (2003) quando os autores sublinham uma particularidade que decorre da própria natureza do turismo e que pode determinar uma diferente relação concorrencial entre os destinos - a sua relativa proximidade geográfica, bem como dos seus mercados, pode permitir-lhes integrar a sua oferta num mesmo pacote turístico ou, então, os seus esforços de marketing conjunto podem expandir a dimensão do mercado na medida em que, isolados, poderiam não conseguir uma tão grande procura.

Neste contexto tem toda a oportunidade a posição defendida por Vaz e Dinis (2007) quando referem a necessidade de algumas localidades e sub-regiões tirarem partido da vizinhança e actuarem mais como aliadas do que como concorrentes, havendo por isso a necessidade de um destino turístico identificar as relações competitivas e complementares mais significativas para cada segmento de mercado, de modo a tornar claro quem são os seus concorrentes directos e quem é que complementa a sua oferta.

3. AFERIÇÃO DAS ACESSIBILIDADES

Segundo Hall (2004) a mobilidade tem surgido nos últimos anos como um conceito-chave na área das ciências sociais no entanto a sua aplicação nos estudos do turismo tem sido relativamente limitada. Na perspectiva deste autor (*op.cit*., 2004) o turismo continua a ser classificado, e os seus fluxos medidos, por dados estatísticos e concepções reducionistas de viagem, que servem para distinguir o turismo de outras formas de mobilidade. Como

[6] Que não incluem as unidades de turismo rural, à semelhança de Portugal.

[7] http://www.ine.es/jaxi/tabla.do?type=pcaxis&path=/t38/bme2/t11/e162eoh/l0/&file=16000c1.px

[8] *The treatment of tourism exclusively as a form of leisure is too limiting. The nature of tourism should also be sought, for instance, in our attitude to space (...). In this sense, tourism could simply be regarded as the process of the change of place in space by an individual or a group of people, a temporary change of the place of sojourn.*(Pearce & Butler, 1993, p.11).

consequência, há muitas vezes pouco cruzamento estatístico ou intelectual com outras áreas envolvidas com a mobilidade, como é o caso dos transportes, migração ou outros domínios da diáspora e das networks globais. Em certo sentido essa limitação já começou a ser reconhecida pela Organização Mundial do Turismo, ao admitir a categoria de viagens de um dia como uma forma de comportamento turístico (ONU, 1994, citada em Hall, 2004: 2).

Como refere Silva (1993) uma área de destino turístico deverá possuir boa acessibilidade a partir das principais vias de comunicação, devendo o sistema de circulação ser eficiente. Segundo Fadda e Fancello (2004) o desenvolvimento gerado pelo turismo é o resultado não apenas da possibilidade de criar empregos relacionados com o turismo, mas também da facilidade de acesso aos sítios seleccionados. Tal significa que para aumentar a atractividade de um produto turístico, a integração do turismo com as infra-estruturas de transportes assume um papel fundamental.

3.1 Aspectos Metodológicos[9]

A análise da acessibilidade, real e potencial, da Beira Interior tem por base o quadro das redes rodoviárias ibéricas em 1992[10], 2006[11] e 2020[12], entrando em conta neste último quadro com a projectada rede ibérica ferroviária de VA/AV. A intenção é comparar a acessibilidade em cada um destes cenários com um outro, ideal (assimptótico), no âmbito do qual fosse possível realizar deslocações para qualquer destino, em linha recta, a uma velocidade média de 100 km/h.

De acordo com as características inerentes às redes rodoviária e ferroviária de cada um dos cenários, verificou-se até onde era possível viajar em 5 horas, com intervalos de 1 hora[13], a partir da Guarda, Covilhã e Castelo Branco. A acessibilidade de cada cidade foi avaliada, em cada limiar de tempo de viagem, pelo somatório do total de população atingida ao longo de cada um dos percursos. Ou seja: muito embora as ligações tivessem por origem *aquelas* cidades em particular e por destino uma qualquer outra localidade, a população total considerada foi a correspondente ao somatório da de todos os concelhos alcançados, incluindo a do de origem e a do de destino. Nos casos em que a cidade de destino se situava em território espanhol, o critério para a população deixou de ser o do concelho para passar a ser o da província[14]. No cenário de 2020, nos casos em que alguns troços da ligação se efectuaram também pela rede de comboio de VA/AV, apenas foi contabilizado o total da população correspondente aos concelhos (em Portugal) e/ou às províncias (em Espanha) das respectivas estações de paragem.

Constatou-se, na altura, que dos planos da REFER[15] fazia parte a construção de uma linha de caminho de ferro, simultaneamente de VA e de AV, entre Aveiro e Salamanca, que permitiria, não só unir o Porto e Madrid em menos de 3 horas mas também configurar um nova rota de passageiros e de mercadorias entre o Norte de Portugal e o Norte/Centro da Europa – via Irún, e a Europa Mediterrânica – via Port-Bou (Silva, 2004). Embora as autoridades competentes ainda o não tivessem definido com rigor admitia-se, no entanto, que o respectivo traçado viesse a beneficiar a cidade de Viseu. Porém, há estudos recentes que defendem, simultaneamente, a construção de mais estações do que as inicialmente previstas e a circulação do maior número de comboios possível em toda a rede de VA e de AV, de forma a poder captar

[9] Para o desenvolvimento deste ponto bem como do ponto 3.2, baseámo-nos essencialmente em Silva (2005).

[10] 1992 - logo após ter sido dado por concluído o eixo rodoviário Porto – Lisboa.

[11] 2006 - no contexto de todos os investimentos previstos no respectivo QCA.

[12] 2020 - data em que se prevê estarem realizados, não só o actual PRN mas também grande parte das RTE de Transportes (Dalton, 2003).

[13] Subdivididos em escalões de 15 minutos.

[14] Os dados demográficos relativamente: aos concelhos portugueses, são os publicados pelo INE e dizem respeito ao último Censo (INE, 2003); às províncias espanholas, são os constantes do Anuario Estadístico 2003 disponibilizado pelo Ministerio de Fomento espanhol (Ministerio de Fomento, 2003).

[15] REFER – Rede Ferroviária Nacional, EP.

também um maior número de tráfegos locais[16]. Assim sendo, admitimos que em cada linha as estações e o respectivo número de comboios que as serviriam eram as constantes do Quadro 3.

No entanto, sempre que em cada percurso se justificava uma mudança do modo rodoviário para o ferroviário, e vice-versa, entrámos em linha de conta com algum tempo de espera. Em primeiro lugar, e sempre, 20 minutos - o tempo que considerámos necessário para entrar/sair da cidade e para aceder/deixar a estação de comboios. Em segundo lugar, um tempo (TUe_i) de acesso a qualquer uma das estações (i) servida por **uma** única linha[17], e em função do respectivo número de serviços diários (Sd_i), (J. M. Viegas Consultores, 2000). No caso das estações servidas por **várias** linhas[18], este tempo (TVe_i) foi calculado através do quociente entre a soma do máximo da população dos pares de concelhos/províncias correspondentes às estações nos extremos de cada linha (j, k) ponderada pelo TUe_i da respectiva linha, e o somatório do máximo da população dos pares de concelhos/províncias correspondentes a essas mesmas estações[19].

Quadro 3 – Estações e Total de Serviços Diários nas Linhas de VA e de AV

Linhas de VA e de AV	Estações	Total de Serviços (por dia)
Lisboa – Porto	Lisboa	20
	Santarém	4
	Leiria	4
	Coimbra	6
	Aveiro	4
	V. Nova de Gaia	8
	Porto	20
Lisboa – Madrid	Lisboa	20
	Pragal	4
	Évora	4
	Elvas/Badajoz	8
	Cáceres	8
	Madrid	20
Lisboa - Faro	Lisboa	6
	Pragal	2
	Évora	2
	Beja	2
	Faro	6
Porto – Vigo	Porto	10
	Braga	4
	Vigo	10
Porto - Madrid	Porto	6
	V. Nova de Gaia	6
	Aveiro	2
	Viseu	2
	Guarda	2
	Salamanca	2
	Valladolid	2
	Madrid	6

Fonte: TIS,pt

[16] Ainda que, mesmo assim, cada estação não seja servida por todos os comboios que por lá passem mas apenas por alguns, numa frequência a ser ponderada por vários factores ao longo do tempo. Os referidos estudos admitem ainda a possibilidade das linhas de Velocidade Alta e de Alta Velocidade poderem vir a ter paragens a cada 50 km e 100 km, respectivamente.

[17] Casos de Santarém, Leiria, Coimbra, Elvas/Badajoz, Cáceres, Beja, Faro, Braga, Vigo, Viseu, Guarda, Salamanca e Valladolid.

[18] Casos de Lisboa, Aveiro, V. Nova de Gaia, Porto, Lisboa, Pragal, Évora e Madrid.

[19] No caso da estação de origem ser servida por mais de uma linha, o índice j coincide com o índice i.

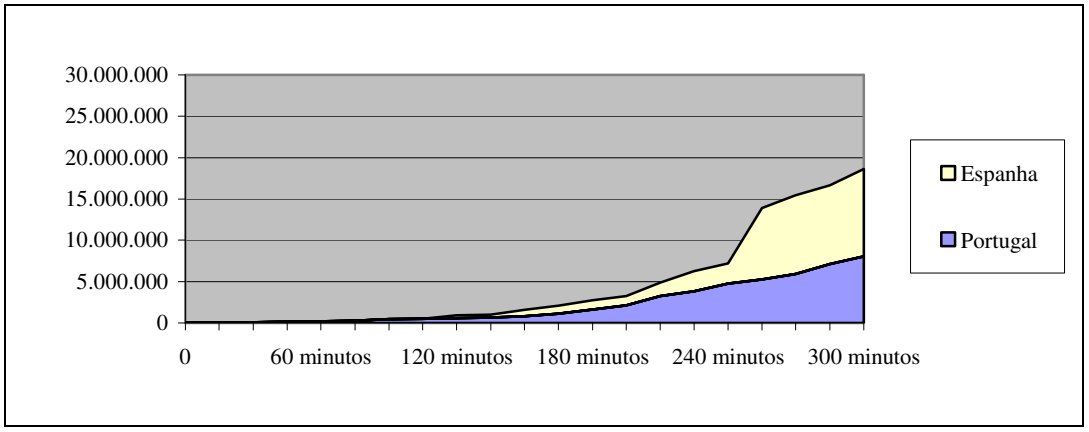
Para o cálculo das acessibilidades rodoviária e ferroviária em cada cenário, para além das bases de dados em Arc View, Arc Info e Access que foram cedidas pela TIS,pt, pela Prof.ª Elisabete Silva e pela Eng.ª Fátima Figueira, utilizou-se o aplicativo GeoMedia Transportation Manager – versão 5.2, do programa GeoMedia Professional – versão 5.2, da Intergraph Corporation.

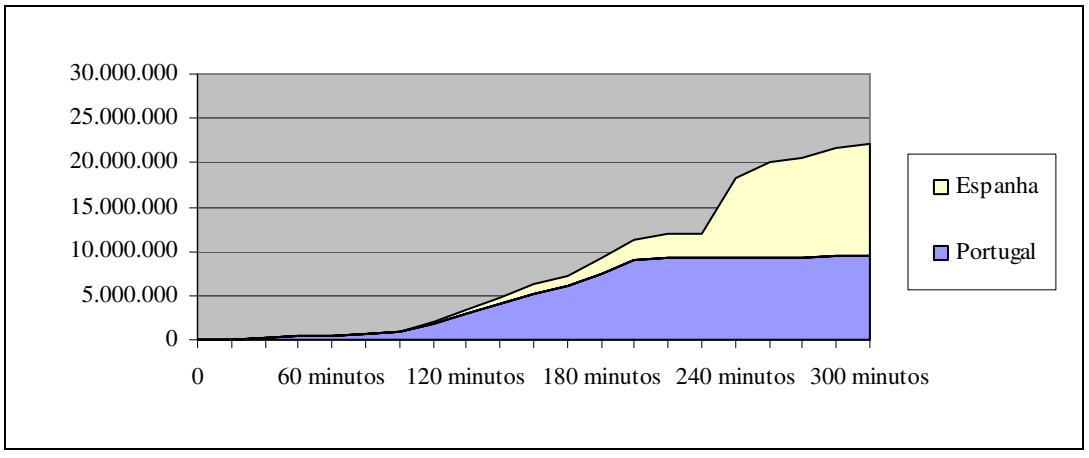
3.2 Guarda, Covilhã e Castelo Branco

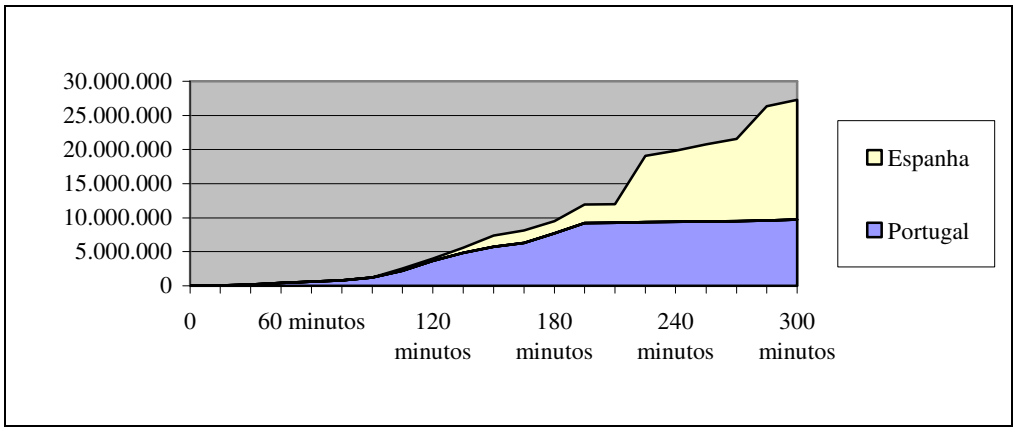
Começamos por analisar a acessibilidade da Guarda, Covilhã e Castelo Branco, na Beira Interior, numa perspectiva longitudinal, isto é, tentando perceber o total de população que era possível atingir a partir de cada uma destas cidades no espaço de 5 horas, em cada um dos cenários: 1992, 2006, 2020 e Assimptótico. Além do mais, dividimos o contributo populacional assim obtido em duas partes – a portuguesa e a espanhola, de forma a melhor visualizar o contributo de cada país para a acessibilidade de cada cidade.

A primeira cidade que analisámos foi a Guarda (Figura 1). No cenário de 1992 verificámos que apenas entre os 75 minutos (0,273 milhões de habitantes) e os 90 minutos de viagem (0,481 milhões) a população possível de atingir superava a confinada, unicamente, à Beira Interior (0,337 milhões) - um isolamento que podia ser explicado, em boa medida, pelo défice de infraestruturas de acessibilidade de qualidade existente na altura. Só a partir dos 90 minutos de viagem se começava a vislumbrar um aumento significativo da população alcançada, que às 2 horas atingia 0,944 milhões de habitantes, às 3 horas 2,745 milhões, às 4 horas 7,188 milhões e às 5 horas 18,648 milhões.

Figura 1 – Contributos de Portugal e de Espanha para a Acessibilidade da Guarda em 1992, em 2006 e em 2020, em habitantes por tempo de viagem



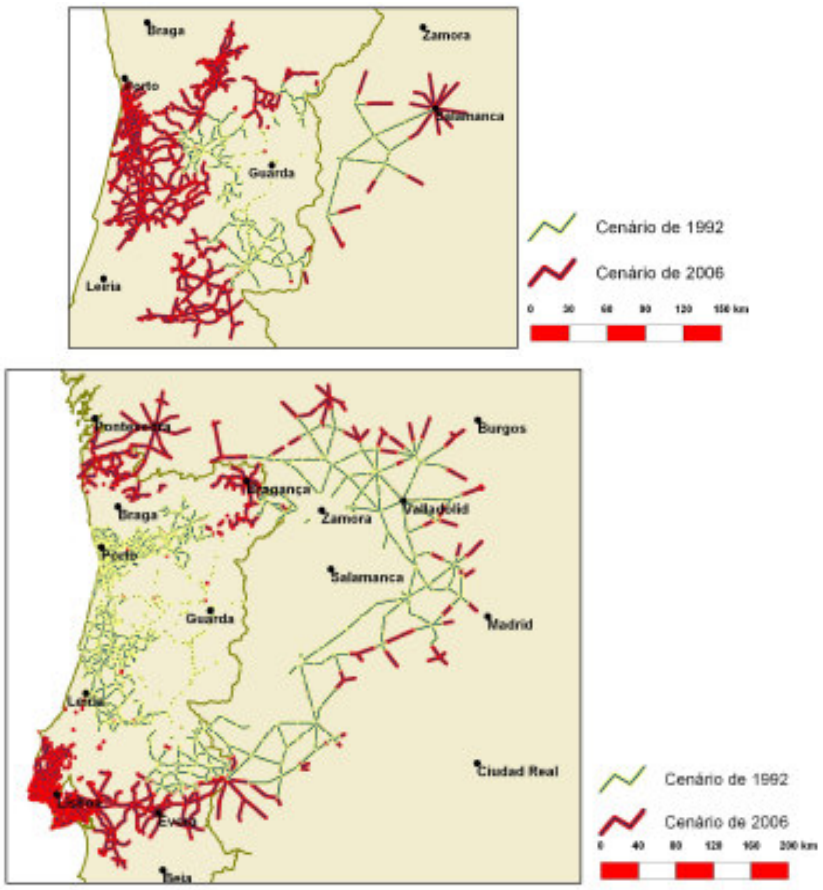




Fonte: Cálculo das Acessibilidades

Esta situação altera-se significativamente em 2006, sobretudo por entretanto ter aberto ao tráfego o IP.5 - permitindo ligações mais fáceis ao Litoral e ao Norte do país mas também a Espanha, e a A.23 – facilitando as ligações ao Sul. Repare-se, em primeiro lugar, que o tempo necessário para alcançar uma população equivalente à da Beira Interior, que era sensivelmente igual a 80 minutos em 1992, é agora de apenas 45 minutos (0,347 milhões de habitantes); isto é, foi substancialmente reduzida a barreira do isolamento. Aliás, é logo a partir dos 45 minutos de viagem que se começa a verificar o incremento da acessibilidade, especialmente significativa em 2 horas (onde se atingem 3,297 milhões de habitantes) e em 4 horas (com cerca de 18,283 milhões), (Figura 2).

Figura 2 – Acessibilidade da Guarda nos Cenários de 1992 e 2006, para Limiares de Tempo de Viagem de 2 horas e de 4 horas



Fonte: Cálculo das Acessibilidades

Se em 1992 as 2 horas de viagem permitiam atingir a quase totalidade da Raia Central Ibérica (0,944 milhões de habitantes), em 2006 as mesmas 2 horas permitem alcançar toda a Região Centro do País e viajar até Salamanca, em Espanha (3,297 milhões de habitantes na totalidade). Do mesmo modo, se em 1992 com 4 horas de viagem era possível alcançar uma população de 7,188 milhões de habitantes, em 2006 com as mesmas 4 horas chega-se a 18,283 milhões. Aliás, para além de ser possível alcançar agora o Norte, o Centro e parte significativa do Sul de Portugal, também se pode viajar mais facilmente até ao Norte (Pontevedra, Burgos) e Centro (Zamora, Valladolid, Madrid) de Espanha.

Ou seja: em 1992 eram necessárias 2 horas para se atingir a população que em 2006 se atinge em pouco mais de 90 minutos, 3 horas em 1992 para se atingir a mesma que em menos de 2 horas em 2006, e 4 horas em 1992 para alcançar a mesma que em menos de 3 horas em 2006. Em 2006, com 5 horas de viagem a Guarda está ao alcance de 22,086 milhões de habitantes - mais 3,438 milhões do que em 1992.

Na transição de 2006 para 2020, isto é, tomando agora em consideração também a rede de caminho de ferro de VA e de AV em toda a Península Ibérica, podemos verificar que para o mesmo tempo de viagem o acréscimo de população não é muito significativo: em 2020, com 2 horas de viagem será possível alcançar 3,890 milhões (contra 3,297 milhões em 2006), em 3 horas 9,477 milhões (contra 9,264 milhões em 2006), e em 4 horas 19,841 milhões (contra 18,283 milhões em 2006). Apenas para as 5 horas de viagem tal incremento é superior, na ordem de 5 milhões de habitantes (22,086 em 2006 contra 27,295 em 2020).

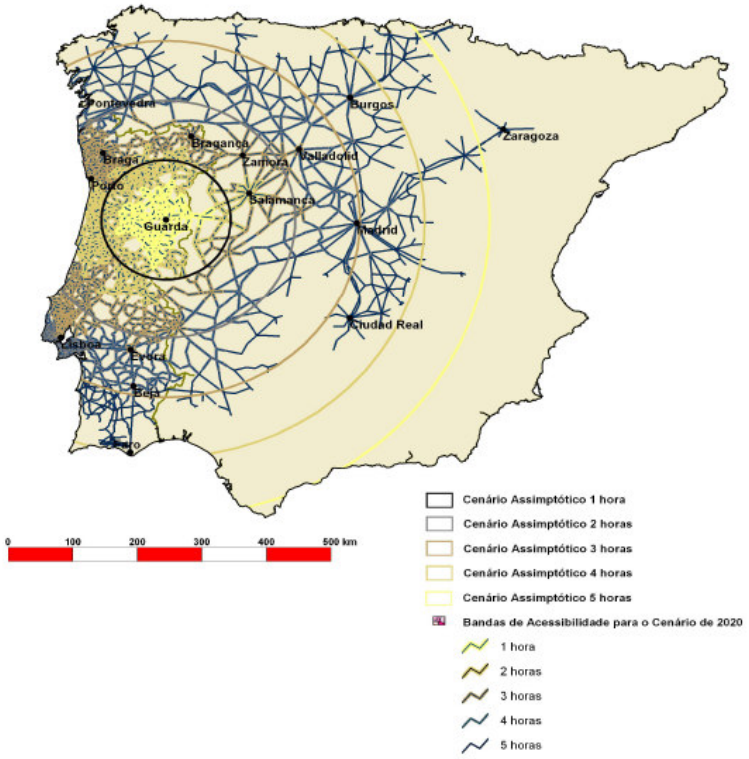
Há ainda outros aspectos que não podemos deixar de referir. Repare-se que o tempo necessário para superar uma população equivalente à da Beira Interior era, para a cidade da Guarda, sensivelmente igual a 80 minutos em 1992, e que se reduz para 45 minutos em 2006 - onde se mantém, praticamente, em 2020 (0,402 milhões de habitantes). Se na transição de 1992 para 2006 tal facto pode ser explicado pela entrada em funcionamento do IP.5/A.25 e da A.23, então como explicar a realidade patenteada após 2006? Certamente que já não com o argumento da falta de infraestruturas de acessibilidade de qualidade. De facto, o que parece estar agora em causa é a falta de *peso* próprio deste espaço territorial em matéria de população.

Ainda assim, repare-se que no cenário de 1992, a partir das 4 horas de viagem, o *peso* da população espanhola aumentava consideravelmente, sobretudo pela facilidade das ligações a Espanha - mais do que ao Litoral português. Nos cenários subsequentes atinge-se a quase totalidade do território português logo após as 3 horas de viagem, tempo a partir do qual se começa a verificar o (forte) incremento da população alcançada em Espanha – tal como seria de esperar.

A Figura 3 compara o cenário de 2020 com um quadro (assimptótico) de acessibilidade ideal, no âmbito do qual fosse permitido viajar da Guarda para qualquer destino, em linha recta, a uma velocidade média de 100 km/h. Numa situação destas, em 3 horas de viagem seria possível alcançar, para além da quase totalidade do território nacional[20] também boa parte do espanhol, sobretudo quase toda a Galiza e parte significativa do Centro, incluindo Madrid. Em 4 horas seria possível atingir todo o território nacional, ultrapassar Madrid e alcançar Burgos e Ciudad Real. Em 5 horas chegar-se-ia quase a Irún e a Zaragoza. No entanto, repare-se que esta é também a leitura possível do cenário para 2020, altura em que a cidade da Guarda se encontrará servida por uma linha de caminho de ferro de Velocidade Alta. Neste sentido podemos afirmar, com algum rigor, que para tempos de viagem na ordem das 3 e 4 horas, em 2020 a cidade da Guarda estará com uma acessibilidade igual a 100% e 80%, respectivamente, da considerada ideal.

[20] Neste particular, sempre que nos referirmos ao território nacional estamos a excluir, naturalmente, os arquipélagos dos Açores e da Madeira.

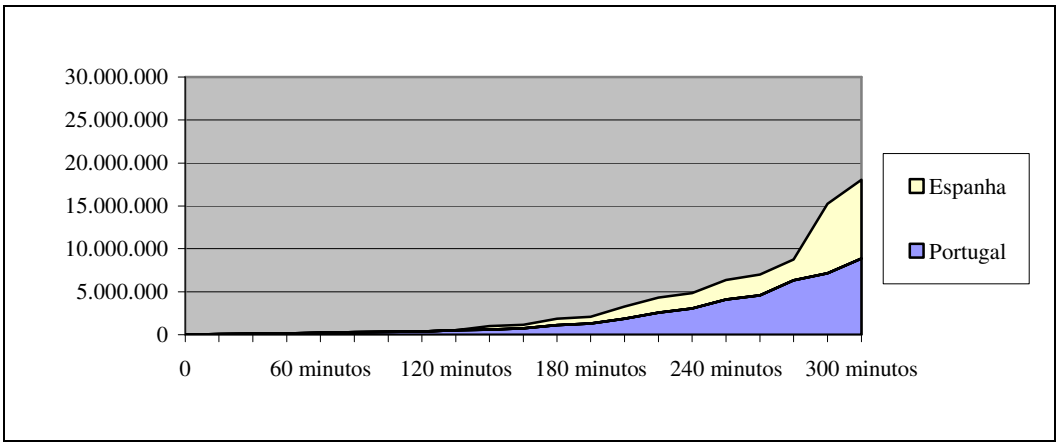
Figura 3 – Acessibilidade da Guarda nos Cenários de 2020 e Assimptótico até 5 horas de Viagem

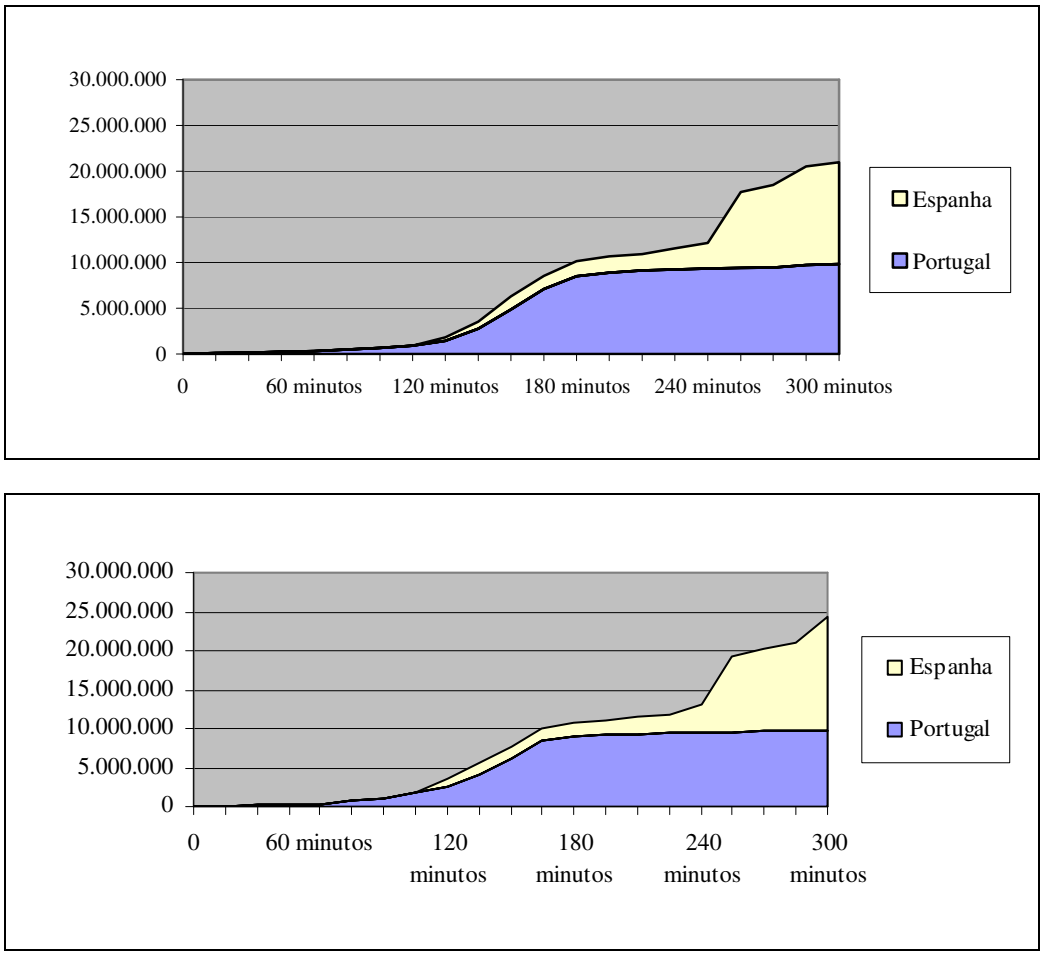


Fonte: Cálculo das Acessibilidades

Depois da Guarda a cidade que analisámos foi a Covilhã (Figura 4). No cenário de 1992 verificámos que até cerca de 90 minutos de viagem a população possível de atingir (0,346 milhões de habitantes) a partir desta cidade estava confinada, praticamente, à da Beira Interior, um *isolamento* que tal como para a cidade da Guarda podia ser explicado pelo défice de infraestruturas de acessibilidade de qualidade que à data se fazia sentir. De igual modo, só a partir dos 90 minutos de viagem se começava a sentir um ligeiro aumento da população alcançada que às 2 horas atingia 0,524 milhões, às 3 horas 2,067 milhões, às 4 horas 6,366 milhões e às 5 horas 18,042 milhões.

Figura 4 - Contributos de Portugal e de Espanha para a Acessibilidade da Covilhã em 1992, em 2006 e em 2020, em habitantes por tempo de viagem

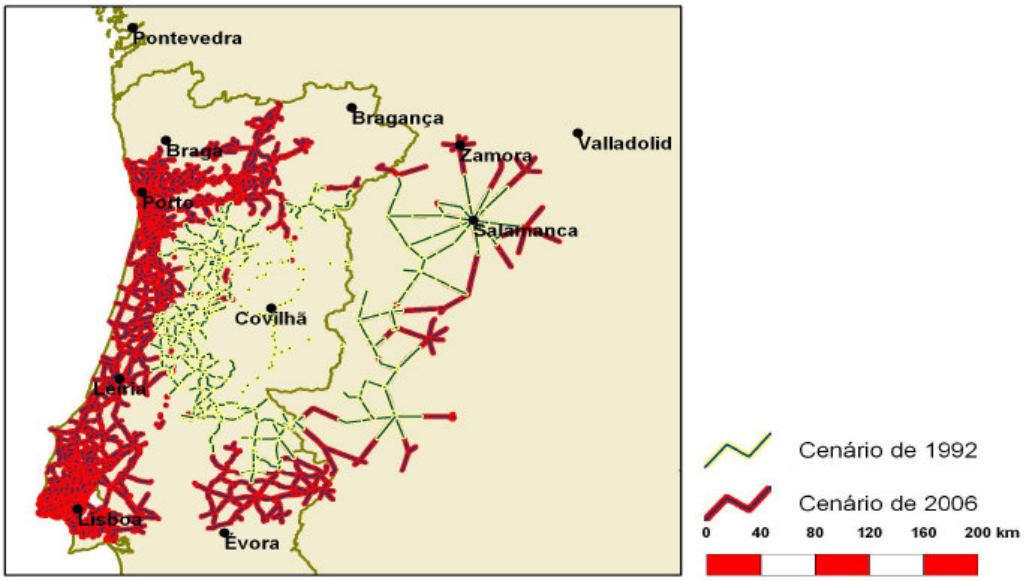




Fonte: Cálculos das Acessibilidades

Esta situação altera-se em 2006, sobretudo por entretanto ter aberto ao tráfego o IP.5 (a Norte) a A.23 e o IC.6 (este de ligação directa a Coimbra e, por conseguinte, a todo o Litoral Centro). No entanto, porque a Covilhã está mais longe do IP.5/A.25 do que a Guarda, o tempo que necessita para alcançar uma população equivalente à da Beira Interior, que era de 90 minutos em 1992, recua menos de 30 minutos em 2006 (contra os 35 minutos da Guarda), fixando-se, sensivelmente, em 1 hora. Aliás, é precisamente a partir da 1 hora de viagem que se começa a verificar o incremento da acessibilidade, especialmente significativo em 3 horas (Figura 5): se em 1992 as 3 horas de viagem permitiam alcançar 2,067 milhões de habitantes – correspondentes a grande parte do interior da Região Centro do país e da Raia Central espanhola até Salamanca, em 2006 as mesmas 3 horas permitem alcançar 10,130 milhões, de Braga a Évora e até Salamanca. Curiosamente, o afastamento da Covilhã relativamente à fronteira de Vilar Formoso (IP.5/A.25) levam-na a privilegiar, neste cenário, os contactos em território nacional.

Figura 5 – Acessibilidade da Covilhã nos Cenários de 1992 e 2006, para um Limiar de Tempo de Viagem de 3 horas



Fonte: Cálculo das Acessibilidades

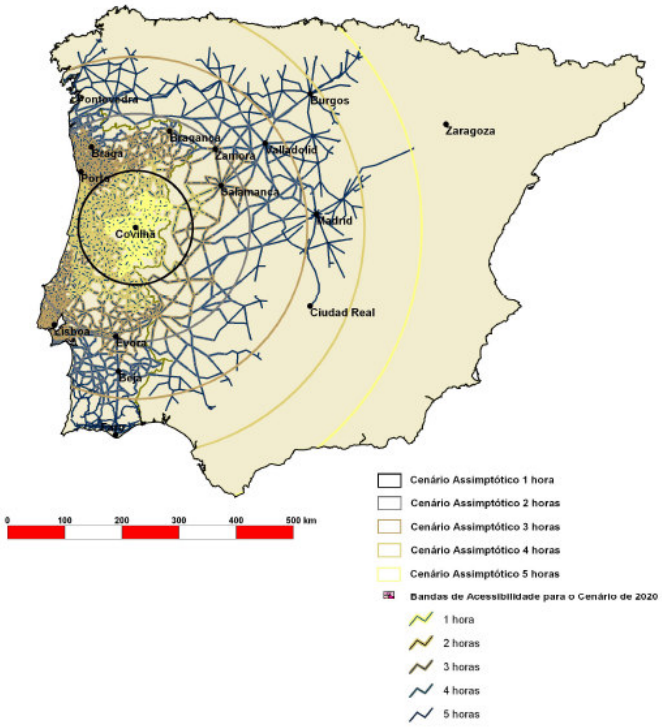
Em 1992 eram necessárias 3 horas para atingir a população (2,067 milhões de habitantes) que em 2006 se atinge em pouco mais de 2 horas (1,835 milhões); mas com 3 horas de viagem em 2006 é possível alcançar (10,130 milhões) bem mais do que em 4 horas no cenário de 1992 (6,366 milhões). È curioso verificar, de igual modo, que no cenário de 2006 a Covilhã não regista um grande incremento da acessibilidade das 3 para as 4 horas, sobretudo porque já consegue alcançar boa parte do território nacional sem, no entanto, penetrar muito em Espanha. Porém, a situação inverte-se das 4 para as 5 horas: a Covilhã está ao alcance de 20,983 milhões de habitantes, mais 2,941 milhões do que em igual tempo de viagem no cenário de 1992.

Na transição de 2006 para 2020 assinale-se o facto do tempo que a Covilhã necessitará para alcançar uma população equivalente à da Beira Interior apenas recuar ligeiramente em relação à 1 hora (permitindo que se alcancem 0,373 milhões de habitantes). Neste contexto, em 2020 será possível alcançar em 2 horas 3,698 milhões de habitantes, o dobro da população que em igual tempo de viagem em 2006 (1,835 milhões). Mas em tudo o resto ambos os cenários se assemelham: exceptuando a diferença de 3,344 milhões que se verifica para 5 horas de viagem – que em muito se deve ao comboio de VA com uma estação prevista para a vizinhança da Guarda, tudo leva a crer que a partir de 2006 a acessibilidade da Covilhã se mantenha, praticamente, inalterada.

Tal como sucedia para a Guarda em 45 minutos, o tempo necessário à Covilhã para alcançar uma população equivalente à da Beira Interior não recua significativamente para além de 60 minutos em qualquer um destes últimos cenários. De facto, após entrarem em pleno funcionamento a A.25, a A.23 e o IC.6, o que parece estar agora em causa é a falta de *peso* deste espaço territorial em matéria de população. Ainda assim – e tal como sucede para a Guarda, nos cenários de 2006 e 2020 atinge-se a quase totalidade do território português logo após as 3 horas de viagem, altura a partir da qual se começa a fazer sentir o *peso* da população espanhola – mais significativo nas 4 e nas 5 horas.

A Figura 6 compara os cenários de 2020 e o Assimptótico. Neste último, em 3 horas seria possível alcançar, praticamente, todo o país e também grande parte da Galiza e do Centro de Espanha. Em 4 horas seria possível atingir todo o país, Madrid, Burgos e Ciudad Real. No entanto, mesmo em 5 horas de viagem não seria possível atingir as fronteiras de Espanha com a França. Mas também esta é uma leitura possível do cenário para 2020. Neste contexto podemos afirmar, com algum rigor, que para tempos de viagem na ordem das 3 e 4 horas, em 2020 a cidade da Covilhã – tal como a Guarda, estará com uma acessibilidade igual a 100% e 80%, respectivamente, da considerada ideal.

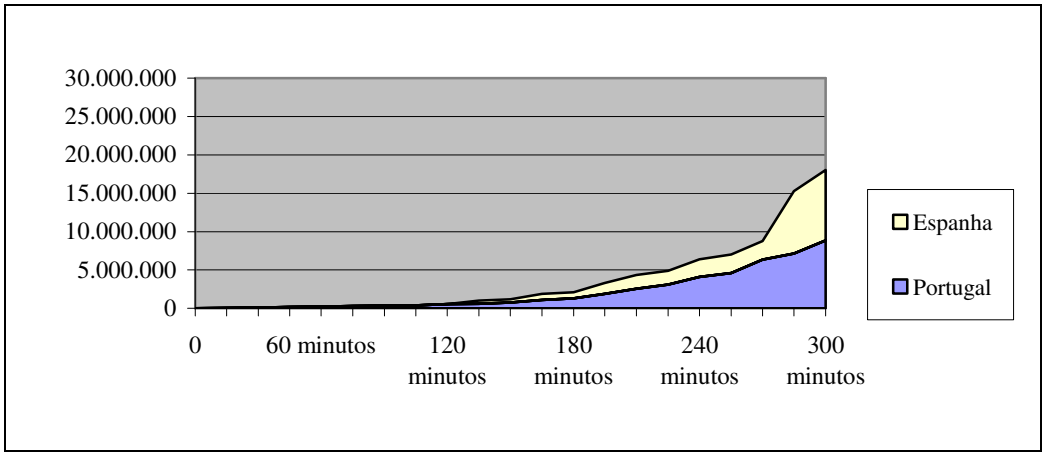
Figura 6 - Acessibilidade da Covilhã nos Cenários de 2020 e Assimptótico até 5 horas de Viagem

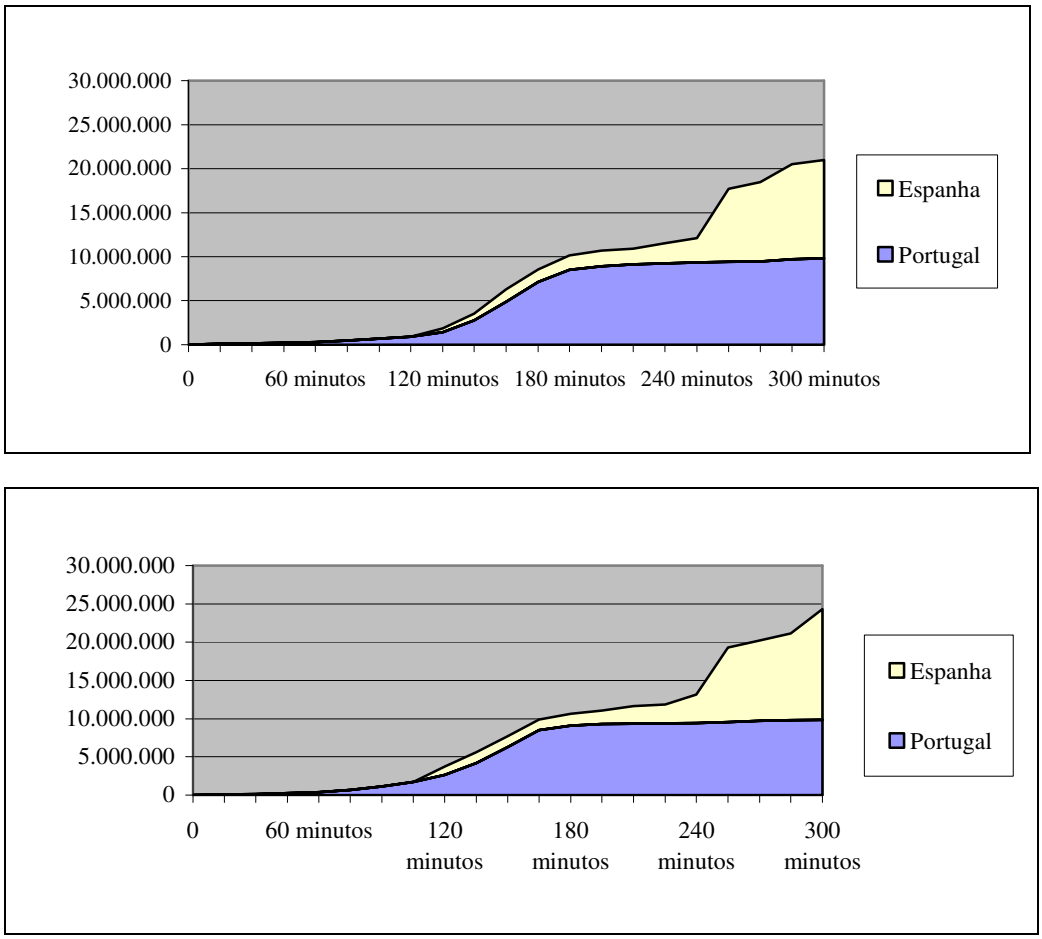


Fonte: Cálculo das Acessibilidades

Em terceiro lugar analisámos a cidade de Castelo Branco (Figura 7). À semelhança da Guarda (80 minutos), mas ao invés da Covilhã (90 minutos), o tempo necessário em Castelo Branco para alcançar uma população equivalente à da Beira Interior, em 1992, estava perto dos 80 minutos de viagem (0,250 milhões de habitantes em 75 minutos, e 0,766 milhões de habitantes em 90 minutos). Não que a cidade estivesse dotada de boas acessibilidades, mas sobretudo porque sendo de entre elas a localizada mais a Sul no Eixo Urbano da Beira Interior, tal posição sempre lhe conferiu melhores ligações a Lisboa e ao Sul do país. Repare-se como logo a partir de 1 hora de viagem se começava a fazer sentir o aumento da população possível de alcançar, que às 2 horas rondava cerca de 1,644 milhões de habitantes, às 3 horas 3,274 milhões, às 4 horas 9,365 milhões e às 5 horas 18,379 milhões.

Figura 7 – Contributos de Portugal e de Espanha para a Acessibilidade de Castelo Branco em 1992, em 2006 e em 2020, em habitantes por tempo de viagem



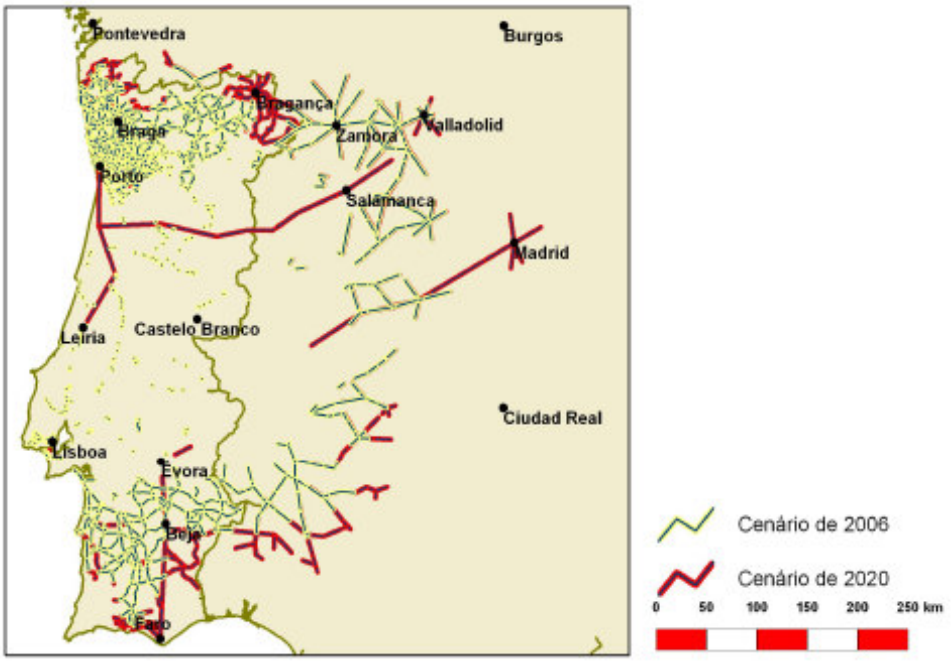


Fonte: Cálculos das Acessibilidades

Esta situação sofre algumas alterações no cenário de 2006, sobretudo por entretanto ter sido aberto ao tráfego a A.23, o IP.6 e o IC.8. Neste cenário é evidente o incremento da população alcançada em território nacional logo a partir da 1.ª hora de viagem; o contributo da população espanhola é sempre muito inferior, sobretudo até às 4 horas de viagem. A partir daí ele torna-se mais significativo, fruto das ligações transfronteiriças proporcionadas pelo IC.31. Deste modo, em 2 horas é possível alcançar 4,236 milhões de habitantes, 8,631 milhões em 3 horas, 11,678 milhões em 4 horas e 22,802 milhões em 5 horas. Ou seja: em 2006, com 2 horas de viagem é possível alcançar 2.6 vezes mais população do que em 1992 se atingia em 2 horas, e quase mais 1 milhão do que em 3 horas; e em 4 e 5 horas de viagem é possível alcançar mais 2,313 milhões e 4,423 milhões, respectivamente, do que em igual tempo em 1992.

Na transição de 2006 para 2020 não há alterações significativas para a acessibilidade até às 2 horas de viagem. De facto, é a partir das 3 horas que surge o incremento mais significativo (de 8,631 milhões de habitantes passa para 9,677 milhões), ainda mais assinalável com 4 horas de viagem (de 11,678 milhões de habitantes passa para 17,536 milhões). A Figura 8 mostra bem como este incremento se deve, em boa medida, ao contributo da rede ferroviária de VA e de AV, onde alguns dos troços surgem perfeitamente assinalados. De igual modo, em 2020 – com 24,902 milhões, será possível alcançar, em 5 horas, mais 2,1 milhões de habitantes do que em igual período em 2006 (22,802 milhões).

Figura 8 – Acessibilidade de Castelo Branco nos Cenários de 2006 e 2020, para um Limiar de Tempo de Viagem de 4 horas

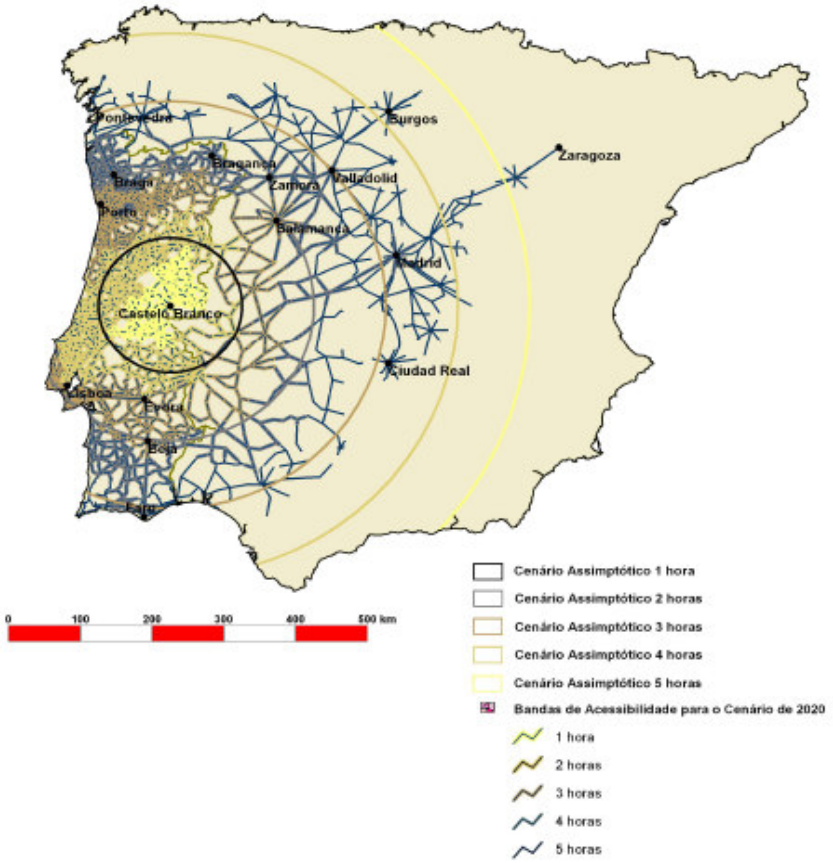


Fonte: Cálculo das Acessibilidades

Tal como sucedia para a Covilhã (60 minutos) e para a Guarda (45 minutos), também aqui o tempo necessário para alcançar uma população equivalente à da Beira interior não recua para além de 60 minutos em qualquer um dos últimos cenários. O que demonstra, efectivamente, que após estarem em pleno funcionamento a A.25, a A.23, o IC.6, o IP.6, o IC.8 e o IC.31, o que parece estar em causa em Castelo Branco – neste particular, e na Beira Interior – em geral, é a falta de *peso* próprio desta região em matéria de população. Tal como para a Covilhã e para a Guarda, nos cenários de 2006 e 2020 atinge-se a quase totalidade do território português logo após as 3 horas de viagem, limiar de tempo a partir do qual se começa a fazer sentir o *peso* da população espanhola – mais significativo ainda nas 4 e nas 5 horas.

A Figura 9 compara os cenários de 2020 e o Assimptótico. Neste último, em 3 horas seria possível alcançar todo o país, e grande parte da Galiza e do Centro de Espanha. Em 4 horas seria possível atingir Madrid e Ciudad Real. Muito embora com 5 horas de viagem fosse possível atingir Burgos, ainda assim Castelo Branco ficaria aquém de alcançar a fronteira de Espanha com a França. No entanto, também esta é uma leitura possível do cenário para 2020. Neste sentido podemos afirmar, com algum rigor, que para tempos de viagem na ordem das 3 e 4 horas, em 2020 a cidade de Castelo Branco – tal como a Covilhã e a Guarda, estará com uma acessibilidade igual a 100% e 80%, respectivamente, da considerada ideal.

Figura 9 – Acessibilidade de Castelo Branco nos Cenários de 2020 e Assimptótico até 5 horas de Viagem



Fonte: Cálculo das Acessibilidades

4. CONCLUSÃO

Toda a literatura salienta a importância das boas acessibilidades para possibilitar o desenvolvimento das regiões, ainda que se torne cada vez mais evidente que as acessibilidades, por si mesmas, não constituem o único - nem talvez o principal, factor crítico. Na Beira Interior, tal como em qualquer outra região, o desenvolvimento está presente sob a forma de um conjunto de iniciativas potenciais. No entanto, para que elas se concretizem é necessário, por um lado que sejam mitigadas as barreiras que lhe servem de obstáculos e, por outro, que os quadros regulamentares que as envolvem sejam suficientemente estáveis para que os impulsionadores de tais iniciativas avaliem e sintam a existência de condições para que elas venham a ter sucesso.

A análise da evolução temporal das acessibilidades das principais cidades da Beira Interior mostra ganhos consideráveis no passado recente e uma pequena margem de ganho remanescente no futuro. Daqui resulta que importa melhorar, sobretudo, as ligações rodoviárias de âmbito sub-regional, e concentrar os esforços tendentes à obtenção do desenvolvimento por via de outros factores que não apenas as infraestruturas de acessibilidade.

Esta metodologia permite-nos demonstrar: em que domínios de acessibilidade é ainda desejável à Beira Interior obter melhorias e em que domínios os pequenos ganhos, ainda possíveis, são conquistáveis apenas a custos muito elevados; quais as *direcções* geográficas e as regiões-parceiras relativamente às quais a Beira Interior pode vir a obter contributos significativos para ganhos de acessibilidade em tais escalões mais deficitários; e em que domínios de acessibilidade a região dispõe de vantagens comparativas com o Litoral, permitindo abrir perspectivas para a busca de sectores e de actividades económicas que se poderiam fixar na Beira Interior em condições de elevada competitividade. Em suma, ela

permite-nos obter uma panorâmica muito completa do contributo já conquistado, e ainda possível de alcançar, da melhoria das acessibilidades às cidades da Beira Interior como instrumento ao serviço do desenvolvimento desta região.

Em face destes resultados somos de opinião que a Beira Interior deve ponderar, comparativamente a outras regiões mais ricas e prósperas e que habitualmente lhe servem de referência, as vantagens competitivas que lhe vão ser proporcionadas pelo incremento da acessibilidade. Afinal, uma das vantagens da Beira Interior resulta do próprio fenómeno da Globalização: a região não deve olhar só para a sua trajectória e para os seus objectivos específicos, mas tem de passar a atender aos tipos/estilos de relacionamento resultantes de *janelas* de oportunidade que se lhe podem vir a abrir, precisamente para além das 2 horas de viagem, e que em boa medida podem passar a influir e a condicionar o seu próprio desenvolvimento. Ou seja, é possível encontrar situações de vantagem comparativa devendo, para tal, buscar-se os tipos de investimentos capazes de potenciar actividades que possam suscitar relações de vantagem económica recíproca com regiões situadas para além de 3 horas de percurso a partir da Beira Interior, como é o caso de alguns destinos turísticos espanhóis.

Viajar é uma parte intrínseca do turismo, sendo as infra-estruturas de transporte o elo essencial entre a origem e o destino, facilitando a circulação quer dos visitantes quer dos próprios fornecedores do turismo. Vimos que o turismo de curta duração cresceu 3,5 vezes mais que as férias de longa duração e que as viagens de um só dia estão a tornar-se cada vez mais importantes, sendo consideradas por Genna (2006) como um meio de aumentar o crescimento do bem-estar social e o consequente aumento dos rendimentos destinados a consumo e viagens. Também o mesmo autor refere que a procura de mobilidade a médio e longo curso, que tem o maior impacto sobre a maior parte das infra-estruturas, aumenta com o processo de integração económica europeia, a crescente globalização dos mercados internacionais, a redução dos custos de transporte, o crescimento do PIB per capita, e a crescente procura de serviços turísticos.

Os factos apresentados vêm sublinhar a importância de a região da Beira Interior tirar partido da sua vizinhança, nomeadamente com regiões-parceiras de Espanha no sentido de integrarem no mesmo "pacote" turístico uma oferta de qualidade e diversificada que, embora tendo presente o pouco tempo de permanência dos turistas, os faça vir à região, incentivando a sua descoberta numa próxima viagem mais demorada. O conceito de *touring* pode ser um conceito integrador dos vários interesses pois exige que os turistas, na sua mobilidade, percorram vários sítios durante a mesma viagem. Ora as regiões onde esses sítios se localizam podem obter sinergias se, por exemplo, numa mesma campanha de marketing, divulgarem em cada local outros sítios, uma vez que, mesmo que nessa viagem não sejam visitados, podem despertar o interesse para uma próxima visita (Vaz e Dinis, 2007). Para isso, torna-se necessária (Vaz, 2005) uma estruturação da oferta com multi-destinos que compreenda uma maior cooperação entre os vários agentes envolvidos, de modo a obter "massa crítica" que possa dar resposta a uma procura com maior mobilidade e frequência das visitas; destinos esses que, paradoxalmente, têm, ou devem ter, como elemento comum, a diferença. Se para Ritchie *et al.* (2003) nenhum destino compartilha com outro o mesmo perfil competitivo porque cada destino tem o seu próprio mix de tradições, valores, objectivos e estilos que lhe dão características próprias, segundo Maillat e Kebir (1998) a lógica territorial só pode ser mantida se os actores tiverem compreendido a necessidade de estimular a mobilidade e a reprodução de recursos específicos, coordenando as actividades produtivas com base num sistema de relações comerciais e não comerciais baseado na confiança... e no longo prazo, pois como defende a teoria dos meios inovadores, a lógica territorial implica uma acção simultânea das redes de interacção e das dinâmicas de aprendizagem.

REFERÊNCIAS BIBLIOGRÁFICAS

DALTON, G. (2003) *Van Miert Report and EU Investment. Presentation to the IC Plenary Session. 13 November 2003*, disponível através do site: http://www.uic.asso.fr (último acesso: 16 de Novembro de 2004).

FADDA AND FANCELLO (2004) An experimental evaluation of the variables in planning all inclusive holiday packages, in *European Transport Conference,* Strasbourg, France, disponível em: http://etcproceedings.org/paper/an-experimental-evaluation-of-the-variables-accessibility-and-quality-of-service.

GENNA, G. (2006) The Transportation System in the Sicilian Touristic Development, in *European Regional Science Association*, series: ERSA Conference Papers, number ersa06p652, disponível em: http://www-sre.wu-wien.ac.at/ersa/ersaconfs/ersa06/papers/652.pdf.

HALL, C. (2004) Tourism and Mobility, in *CAUTHE Conference*, Brisbane, Queensland.

HALL, C. AND D. MÜLLER (eds.) (2004) *Tourism, Mobility and Second Homes: Between Elite Landscape and Common Ground*, Clevedon, Channelview Publications.

INE (2007) *Anuário Estatístico da Região Centro*, Lisboa, Instituto Nacional de Estatística.

INE (2003) *Censos 2001. XIV Recenseamento Geral da População. IV Recenseamento Geral da Habitação*, Lisboa, Instituto Nacional de Estatística.

INSTITUTO NACIONAL DE ESTADÍSTICA (2009), disponível em: http://www.ine.es/jaxi/tabla.do?type=pcaxis&path=/t38/bme2/t11/e162eoh/l0/&file=16000c1.px.

JMVIEGAS CONSULTORES (2000) *Estudos Complementares para a Rede Ferroviária de Alta Velocidade*, Lisboa, Instituto Nacional do Transporte Ferroviário.

LEBRE, A. (1998) **A** Urbanística do Lazer e do Turismo, em *Sociedade e Território. Revista de Estudos Urbanos e Regionais. Dossier Territórios do Lazer & do Turismo, Nº 28, Agosto, 23-30*, Porto, Ed. Afrontamento.

MAILLAT, D. AND L. KEBIR (1998) The learning region and the territorial production systems, *Working paper*, disponível em: http://www.unine.ch/irer/wp9802b.

MINISTERIO DE FOMENTO (2003) *Anuario Estadístico 2003*, disponível em: http://www.mfom.es (último acesso: 28 de Novembro de 2004).

PEARCE, D. AND R. BUTLER (1993) (ed.) *Tourism Research. Critiques and Challenges*, London and New York, Routladge in Association with The International Academy for the Study of Tourism.

RITCHIE, J. AND I. GEOFFREY (2003) *The Competitive Destination. A sustainable tourism perspective*, U.K., CabiPublishing.

SILVA, J. (2005) *As Acessibilidades como Factor do Desenvolvimento de Regiões Periféricas: o Caso da Beira Interior*, Tese de Doutoramento em Transportes, (ISBN: 972-8790-54-6, UBI, 2006), Lisboa, Instituto Superior Técnico.

SILVA, L. (2004) *Rede Ferroviária para o Século XXI*, Palestra proferida no âmbito da disciplina de Gestão de Transporte Aéreo, da Licenciatura em Engenharia Aeronáutica, Covilhã, Universidade da Beira Interior.

SILVA, JOÃO ALBINO (1993) Áreas de Destino Turístico – Introdução metodológica à sua avaliação e desenvolvimento, em *Serviços e Desenvolvimento numa região em mudança*, Coimbra, Comissão de Coordenação da Região Centro.

VAZ, M. (2005) Expressão Regional do Desenvolvimento Turístico, Tese de Doutoramento, Covilhã, Universidade da Beira Interior.

VAZ, M. E A. DINIS (2007) Turismo no Litoral Versus Turismo no Interior Português. O destino turístico Serra da Estrela, em *Estudos regionais - Revista Portuguesa de Estudos Regionais*, nº 14, 1º quadrimestre, 71-98, Coimbra, APDR-INE.

A UE E O DESENVOLVIMENTO DAS SUAS REGIÕES: ACESSIBILIDADE E RELAÇÃO CENTRO-PERIFERIA

JOSÉ RAMOS PIRES MANSO
E-mail: pmanso@ubi.pt
JOANA FIDALGO VAZ REIS SILVA
E-mail: fidalgo.jo@gmail.com
JORGE MIGUEL DOS REIS SILVA
E-mail: jmiguel@ubi.pt
Universidade da Beira Interior

Resumo

A distribuição espacial das actividades não é fruto do acaso e surge como resultado de oportunidades e de estratégias de localização pensadas de forma a alcançar objectivos específicos: de maximização do lucro dos produtores ou da utilidade dos consumidores, de optimização da utilização dos recursos naturais e humanos, de estratégias militares ou das instituições ou, inclusivamente, de optimização na obtenção de prazeres individuais. Mas porque o tema do desenvolvimento local e regional não se esgota na análise das teorias de localização e dos modelos que o explicam propomo-nos abordá-lo no contexto da UE e das suas regiões, sobretudo das periféricas, não só analisando sob diversas ângulos a relação Centro-Periferia, mas também definindo alguns pressupostos de base para os Indicadores de Acessibilidade a utilizar em tal análise. Terminamos apresentando alguns dos modelos mais recentes que, no contexto da UE, melhor ilustram o envolvimento da Acessibilidade com a relação Centro-Periferia.

Palavras-chave: Centro e Periferia, Periferia e Acessibilidade, Periferia e Desenvolvimento, Desempenho Económico.
Área Temática: Economia da Energia, do Ambiente, dos Transportes e dos Recursos Naturais.

Abstract

The spatial distribution of activities is not fruit of chance and emerges as a result of opportunities and strategies for the location thought to achieve specific objectives: to maximize producers profit or consumers utility, to optimize natural and human resources use, to accomplish with military or institutions strategies, or even to obtain individual pleasures. But because local and regional development theme is not limited to location theories and correlated models we will address it in EU and its regions context, particularly the peripheral ones, not only analyzing the center-periphery relationship by several angles, but also setting the basic assumptions for those accessibility indicators used in such analysis. Finally, we present some of the most recent models which better describe accessibility and center-periphery relationship in the context of EU territory.

Key-words: Center and Periphery, Periphery and Accessibility, Periphery and Development, Economic Performance.
Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

1 CENTRO E PERIFERIA

1.1 Assimetrias Regionais

Apesar de todas as iniciativas Comunitárias, os resultados obtidos (ainda) não são suficientes para contrariar esta realidade: no limiar do séc. XXI as disparidades regionais no seio da UE persistem, lembrando o muito que ainda há a fazer neste domínio. Inclusivamente – e à medida que a Europa caminha no sentido da plena integração - são vários os factores que

podem vir a concorrer para o agravamento de tais assimetrias, a exemplo do que acontece com a livre circulação das mercadorias e dos factores de produção (Hitiris, 1994). Tal mecanismo beneficiando, regra geral, as regiões tecnologicamente mais desenvolvidas, mais eficientes e mais competitivas - e dispondo de técnicas de produção mais sofisticadas - conduz, necessariamente, a uma tendência generalizada, em todo o território da UE, para a progressiva substituição das actividades económicas menos eficientes por outras mais eficientes (Silva, 2005).

Paralelamente, as empresas melhor apetrechadas procuram as zonas mais desenvolvidas da Comunidade para aí se instalarem - facto que também não concorre para que os diferenciais de preço/produtividade, entre os Estados membros e/ou Regiões, se anulem. Num passado recente – e num esforço de reposição da competitividade - tais diferenciais sempre se poderiam atenuar recorrendo a ajustes pontuais nas taxas de câmbio; presentemente, a integração económica veda às regiões de mais débil produtividade a desvalorização da respectiva moeda: a integração económica, a liberdade de comércio, e o incremento da concorrência e da mobilidade dos factores de produção, são os únicos mecanismos que permitem o equilíbrio de preços, quer dos factores de produção, quer das mercadorias.

Neste contexto – e em face, tanto da incapacidade na exploração de economias de escala, como da perda de competitividade - as regiões menos desenvolvidas apresentam-se em manifesta desvantagem face às mais produtivas.

Assumindo a diversidade como uma realidade incontornável, e esquecendo um modelo (utópico) de aplicação universal, a política regional da UE tem vindo a caminhar, progressivamente, no sentido de contextualizar a competitividade inerente a cada território, na globalidade económico-social - necessariamente complexa - em que estes se inserem. Neste sentido, a discussão em torno da política regional da UE deslocou-se do tema da igualdade inter-regional para outro entendimento da realidade, agora prestando mais atenção à promoção e à eficiência das várias regiões, valorizando especificidades (Costa, 1996).

A política regional da UE é colocada, assim - e face à realidade - perante novos desafios: a par da correcção das desigualdades verificadas à partida, torna-se necessária uma gestão dos territórios, em função das opções estratégicas acordadas mas, simultaneamente, subjacentes à dinâmica inerente a cada espaço; sobretudo pelas consequências - quer políticas quer sociais – que a negação de tal evidência pode representar, inclusivamente na *performance* das economias dos respectivos Estados (Armstrong e Taylor, 2000).

1.2 Antecedentes Históricos

As assimetrias são, no fundo, inerentes ao funcionamento dos próprios sistemas económicos: embora intoleráveis do ponto de vista da equidade são, muitas vezes, o resultado da eficiência sistémica. E "(…) se existe, muitas vezes, uma dimensão cumulativa nestes desequilíbrios, à escala de um país mas, sobretudo, à escala do mapa-mundo, em tempo histórico, os centros e as periferias mudam", (Guerra, 1998:13). De facto, nos séculos XV e XVI Portugal foi o Centro da Europa e do Mundo - hoje é entendido como um espaço periférico da UE. O Reino Unido também já foi o Centro da Europa e do Mundo – mas, no pós-guerra, perdeu tal privilégio em favor dos Estados Unidos da América.

Na Idade Média, os locais centrais estavam associados aos nós de passagem e/ou confluência das principais rotas comerciais - terrestres e marítimas, dando origem às cidades mais importantes – ao Norte e ao Centro da Europa, e nas Costas do Atlântico e do Mediterrâneo, respectivamente. Estas redes urbanas são definidas por Nunes da Silva (1994:52), como "(…) the polynuclear centrality of the upper Medieval age". Elas representavam, no fundo, o centro do poder – político, religioso, económico e cultural - onde o destino do Velho Continente era discutido e traçado. Os restantes centros urbanos – com uma esfera de influência mais restrita - constituíam a periferia, tomando a cargo as produções locais e assegurando as funções administrativas de base.

Durante o século XIX, as cidades reforçaram o seu papel de centros de desenvolvimento, e a Inovação tornou-se, cada vez mais, uma realidade urbana. "The contradiction between centrality and peripherality physically corresponded to urban and rural territories", (Nunes da Silva, 1994:52). No plano internacional, os locais centrais tornaram-se, inquestionavelmente, sinónimos de poder económico e de actividades científicas e culturais, reflexo da importância atribuída – e conquistada - pelas cidades.

No século XX – e sobretudo com o incremento dos transportes terrestres na década de 50 - assistimos ao alastrar da mancha urbana em torno das grandes aglomerações europeias. A periferia assume-se, neste contexto, também, como o espaço na vizinhança dos locais centrais. A Perifericidade deixa de ser considerada como apanágio exclusivo dos espaços rurais já que os espaços urbanos contêm, igualmente, a sua própria periferia.

No contexto das regiões, dos países e do espaço Europeu, o conceito de periferia está associado, não só à distância – física, cultural e psicológica – em relação a determinada centralidade, mas também à fraca densidade das economias de aglomeração, e ao baixo grau de autonomia e de capacidade de decisão - ao ível supra-nacional, no plano económico. Nunes da Silva (1994:53), defende, a propósito, que em cada país a centralidade associada a uma cidade ou a uma aglomeração urbana é, sobretudo, função do papel que desempenha nas redes internacionais de decisão e negócio, se comparada com a importância específica aos níveis regional e nacional; e sublinha: "(…) the international role of a city clearly influences it's national and regional role".

No fundo, a perifericidade é "(…) um fenómeno multidimensional; daí que, não sendo uma fatalidade, é uma realidade dificilmente contrariada pela vontade programada [dos Homens], que tendem a monitorar, sobretudo políticas públicas unidimensionais", (Guerra, 1998:13).

1.3 Dinâmicas Regionais

O conceito de "perifericidade" está directamente relacionado com algo (ou alguém) localizado, precisamente, na "periferia" - isto é: nos limites ou na circunferência – em relação a um "centro".

Historicamente, as regiões periféricas desenvolveram-se como se de sistemas fechados se tratassem, "(…) in which the society grew in harmony with the economy", (Giannopoulos, 1994:40). Isto significa que as expectativas da população em torno de questões como a mobilidade, a informação, os bens materiais – a qualidade de vida, em geral – eram, sensivelmente, da mesma ordem de grandeza da capacidade económica em produzir tais bens e serviços, considerados necessários. À medida que se desenrolaram (e incrementaram) as inovações tecnológicos e os fluxos de informação, aquelas expectativas cresceram mais rapidamente do que o suportado, quer pelo tecido económico, quer pelo sistema produtivo.

Simultaneamente, a UE iniciou a caminhada para a plena integração – política e económica – das Regiões e dos Estados. Desde logo, tudo quanto as economias *locais* não pudessem satisfazer passou a ser importado de outras (já) em diferentes estados de desenvolvimento. Desta forma, estavam criadas as condições para o aparecimento do círculo vicioso: desequilíbrio no comércio externo, inflação, desemprego, etc. (Giannopoulos, 1994).

Nesse sentido, o problema das regiões periféricas parece confinar-se à dicotomia entre as expectativas sociais e a capacidade de resposta das economias locais. Encontramos resposta para tais expectativas fomentando o desenvolvimento sustentável; ou seja, quebrando o círculo vicioso: do baixo nível de desenvolvimento, das poucas oportunidades de emprego, da drenagem de recursos humanos e materiais, da escassez das oportunidades de desenvolvimento, etc.. Se, pelo contrário, este círculo não for interrompido, as repercussões são várias e podem atingir níveis (quase) irreversíveis. Regra geral – e por razões essencialmente de carácter político - as autoridades responsáveis reagem a este ciclo/círculo vicioso com uma série (avulsa) de iniciativas, de repercussões a curto ou a médio prazos, acabando por agravar os desequilíbrios. Neste sentido, Giannopoulos e Boulougaris (1995:257), sublinham que "(…) a long term, stable political strategy that recognises the nature of the problem and follows a consistent approach to its alleviation, is the first prerequisite for reversing the negative spiral of

economic decline in peripheral regions"; acrescentando: "(…) in this sense one could also state that the problems (…) are also political".

1.4 Conceitos de Periferia

1.4.1 Nota Introdutória

O alargamento da UE e a consequente liberalização dos mercados estimularam o estudo e a previsão de fenómenos relacionados com a localização e a concentração do tecido produtivo; motivo suficiente para que se relançasse a discussão em torno da dicotomia Centro-Periferia, abrindo novos caminhos para a investigação de técnicas de avaliação do grau de perifericidade das regiões.

Do ponto de vista geográfico – ou topológico – a perifericidade pode ser encarada como a localização para além da faixa que limita determinada área, precisamente à distância máxima do respectivo centro geográfico. No entanto – e em termos de desenvolvimento regional – a perifericidade deve reflectir, igualmente, a distância à(s) actividade(s) económica(s) mais relevantes da região, do país, ou mesmo de um grupo de países. Neste sentido, a perifericidade – tal como a inversa, a centralidade - é marcada, quer pela configuração física do território, quer pela distribuição geográfica das actividades económicas no seu interior; ou seja: da relação com o desenvolvimento regional se constata que a perifericidade é sinónimo de relativa acessibilidade – ou inacessibilidade – à actividade económica, tal como esta se apresente distribuída, geograficamente, pelo território (Keeble *et al.*, 1988).

O conceito de Acessibilidade – embora ainda assim invocado segundo uma óptica estritamente geográfica – traduz as oportunidades relativas de interacção entre famílias, empresas e indústrias, distribuídas e confinadas a determinado espaço. Reflectindo os custos inerentes à distância em termos monetários, de tempo de acesso à informação, de ajustamentos organizacionais, etc., acaba por marcar as diferenças entre os graus de actividade económica das regiões, quer ao nível da natureza quer ao nível da escala.

1.4.2 Abordagem Económica

O debate em torno do impacte da perifericidade – ou inacessibilidade – das regiões na respectiva actividade económica vem de longa data. No início da segunda metade do século XX Schultz[1] argumentava – a este propósito – que "(…) the existing economic organisation works best at or near the centre of a particular matrix of economic development". No entanto, a importância relativa atribuída pela Teoria Económica Neoclássica à dimensão espacial parece demonstrar que os principais contributos nesta matéria vieram, não tanto da área da teoria económica mas, sobretudo, da Economia Aplicada e da Geografia Económica.

Myrdal (1957) e Hirschmann (1958) procuraram estabelecer uma relação entre a economia das regiões centrais – mais dinâmicas e apresentando um crescimento auto-sustentado, com a economia das regiões periféricas – menos dinâmicas e com um desenvolvimento dependente, em grande medida, do crescimento do centro. Neste sentido, as economias periféricas funcionariam como um receptáculo das ondas de desenvolvimento que emanavam das economias centrais, numa articulação de efeitos a que Myrdal (1957), apelidou de *spread effect*. Simultaneamente, o crescimento das economias centrais seria estimulado, precisamente devido à procura adicional de bens e de serviços que era gerada a partir do desenvolvimento das economias periféricas – o *backwash effect*, na terminologia de Myrdal (1957).

Este duplo efeito na distribuição dos benefícios resultantes do processo de desenvolvimento regional veio a ser invocado, mais tarde, por Kaldor (1966, 1970) - para explicar as taxas relativas de crescimento económico, e por Dixon e Thirlwall (1975) – para demonstrarem a forma como as vantagens competitivas iniciais podiam sair reforçadas num processo de optimização do investimento.

[1] Citado por Keeble *et al.* (1988:2).

Esta perspectiva da relação Centro-Periferia é, segundo Vickerman (2000:3), algo limitada, porque não confere o devido destaque aos custos do espaço, "(…) except implicitly in the rate of diffusion of new technologies or the speed with which demand preferences adjust. Space is thus captured, but hidden, in the relevant elasticities of the model".

O incremento das economias de escala - associado aos progressos no sentido da plena integração europeia - acabam por também influir, directamente, na relação Centro-Periferia, em matérias como o crescimento da indústria e o desenvolvimento económico regional. A localização central, porque facilita o contacto com áreas de mercado mais amplas conduz - em regra - a economias de escala na produção, a custos mais baixos dos produtos, e a aumentos significativos das vendas. Evidentemente que as empresas localizadas em regiões centrais podem também incorrer em custos mais elevados relacionados, por exemplo, com a mão-de-obra ou o valor do espaço; inversamente, as da periferia "(…) may to some extent be able to trade off low labour charges against high transport/distance costs", (Keeble *et al.*, 1988.5).

Toda esta discussão releva a diversidade de custos de distância que – do ponto de vista teórico - podem vir a ocorrer em regiões periféricas, simultaneamente afectando a eficiência competitiva das empresas ali sediadas, e reduzindo o investimento das provenientes do exterior. Mas para além do mais, o impacte da perifericidade pode ainda vir a ser intensificado "(…) by powerful negative multiplier mechanisms with respect to new increments of investment, which arise indirectly from the basic competitive disadvantage of relative location, through the latter's historic conditioning of regional social, economic and organisational structures", (Keeble *et al.*, 1988:6).

A abordagem económica ao conceito de perifericidade – ou inacessibilidade – precisa de ser devidamente enquadrada com o que Vickerman (2000:9) classifica de "(…) specific spatial objectives" e que oscilam, na prática, entre "(…) making a region more accessible to the outside and making it work more efficiently within".

1.4.3 Abordagem Geográfica

Este modelo aborda a relação Centro-Periferia de uma forma distinta da do anterior: na sua forma mais simplista, ele admite que as áreas de mercado se alteram como resposta às oscilações dos custos atribuídos, quer aos espaços, quer aos transportes (Vickerman, 2000), introduzindo o conceito de Potencial Económico como forma de expressar, num único indicador, a distribuição espacial da actividade económica.

Na perspectiva de Vickerman (2000), as principais críticas a esta abordagem (geográfica) são as seguintes:

- a descrição da economia é muito agregada: variáveis como o Rendimento ou o Produto Interno Bruto não traduzem fielmente a importância nas variações estruturais, quer dos sectores, quer dos mercados;
- assumindo como idênticas as necessidades de transporte dos vários sectores, o modelo apenas releva como necessária uma única função de impedância. Ora, os custos de transporte variam, consideravelmente, de sector para sector. Aliás, mesmo de região para região há variações na estrutura de mercado que originam diferenças assinaláveis no custo efectivo dos transportes, inclusivamente para além do reflectido nas funções de custo com base, unicamente, na distância.

Apesar das críticas, esta abordagem assume uma dupla importância: através de um só indicador é possível inferir, quer da distribuição da actividade económica, quer das perturbações na relação Centro-Periferia motivadas por alterações ao nível dos transportes.

1.4.4 *Nova* Geografia Económica

Segundo Vickerman (2000), a abordagem na óptica da *nova* geografia económica assenta em três princípios basilares:

- reconhece a importância dos mercados não competitivos – característica de qualquer economia espacialmente diferenciada;
- incorpora distintas estruturas sectoriais em diferentes regiões;
- reconhece a importância de funções de custo de transporte, específicas, por sector.

A principal crítica a esta abordagem é a de que, na realidade, não é *nova*: as primeiras referências que nos chegam, a este propósito, remontam aos trabalhos pioneiros de Predöhl e de Hotelling, em 1928 e 1929, respectivamente - "(…) just that recent developments have made the problems much more tractable, "tricks" as Krugman (1998) calls them", (Vickerman, 2000:5).

2. PERIFERIA E ACESSIBILIDADE

2.1 Desempenho Económico

Schürmann *et al.* (1997)[2], no trabalho intitulado "Accessibility Indicators", chamam a atenção para a dificuldade em eleger um único indicador de acessibilidade como de importância universal: "(…) on the one hand the selection of accessibility indicators will depend on their statistical interaction with the other transport and non-transport factors of the regional production function; on the other hand the accessibility indicators selected need to be sensitive to the whole range of transport investments and transport infrastructure improvements to be studied (…)".

Estes autores referem ainda as dificuldades que persistem em se conseguir que os indicadores de acessibilidade reflictam o nível do serviço de transporte público. Sublinhando, no entanto, que em circunstâncias normais "(…) the main and most general measurement of accessibility change will be identical with the consumer surplus change (…)", Schürmann *et al.* (1997)[3]; e acrescentam: "(…) there is value in considering the spatial impact of change which would favour use of the information in its accessibility form".

Esta argumentação permite-nos concluir que os indicadores de acessibilidade se tornaram *a pedra de toque* na busca de critérios economicamente significativos para aferição das variações na economia dos territórios (Silva, 2005).

2.2 Pressupostos De Base

A perifericidade é, sem dúvida, um parâmetro incontornável no contexto do desenvolvimento económico regional da UE. Como tal – e esgrimidos vários argumentos - como é que a relativa perifericidade/centralidade de cada região podem/devem ser, realmente, avaliadas? Antes de passarmos em revista algumas soluções achamos conveniente enunciar alguns pressupostos de base, consensualmente aceites em diversos trabalhos (Keeble *et al.*, 1988):

- em primeiro lugar, o grau de perifericidade/centralidade de uma região deve ser avaliado relativamente ao volume de actividade económica mantida com outras regiões. Incorporando, sempre que possível, simultaneamente – e entre outras - indicações a respeito: dos mercados com os quais se transaccionam bens e serviços; das fontes e das oportunidades de aquisição - tanto de matéria-prima, como de componentes; e do (eventual) acesso, tanto a informação específica, como a serviços de apoio à actividade comercial;
- em segundo lugar, o índice utilizado deve avaliar a acessibilidade regional face à actividade económica em termos do vasto leque de custos de distância referidos anteriormente; evitando estreitá-lo e/ou reduzi-lo à (simples) condição de custos de transporte – tal como entendido na teoria *Weberiana* da localização industrial;

[2] Citados por Vickerman (2000:6).
[3] Citados por Vickerman (2000:6).

- em terceiro lugar – e de modo a reflectir, tanto o processo de integração económica da UE, como as trocas comerciais intra-comunitárias – o índice deve avaliar a acessibilidade de cada região, relativamente às actividades económicas mantidas (eventualmente) dentro de todo o espaço da Comunidade, sem limitações impostas por qualquer tipo de barreiras;
- simultaneamente, este índice deve ser expressivo da realidade, de interpretação simples, e fácil de calcular.

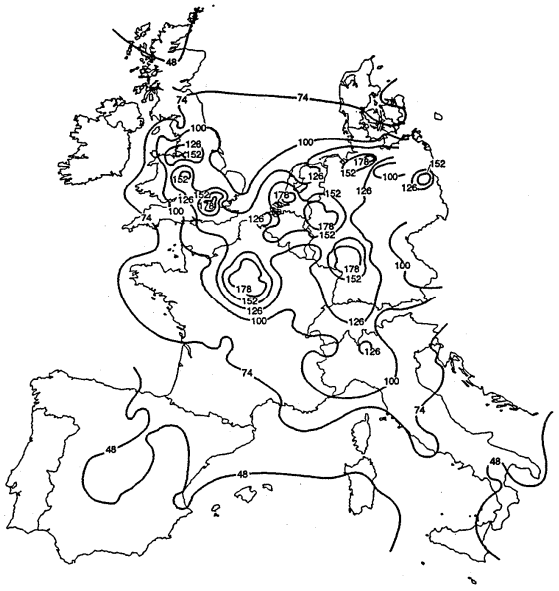
2.3 A Realidade Europeia

Tem vindo a ser publicado, ao longo dos anos, um vasto leque de trabalhos envolvendo a acessibilidade e a relação Centro-Periferia - muito embora em grande parte deles se prefira uma abordagem regional e/ou nacional em detrimento da (verdadeira) dimensão europeia. Diga-se, no entanto – e em abono da verdade – que as últimas duas décadas trouxeram novo fôlego à investigação, com o aparecimento de autores que apostaram, precisamente, em modelos mais abrangentes.

Com base nos trabalhos desenvolvidos, quer por Rietveld e Bruinsma (1998), quer por Wegener *et al.* (2000, 2002)[4], Spiekermann e Neubauer (2002) e Spiekermann (2002) defendem que os contributos mais significativos nesta matéria se resumem a nove modelos, descritos, em seguida, por ordem cronológica.

Keeble *et al.* (1982, 1988) são mandatados pela DG XVI da CE para analisar as diferenças económicas entre regiões - centrais e periféricas - no seio da Comunidade, e investigar, simultaneamente, se algumas (diferenças) de entre elas podem ser explicadas relativamente à posição (geográfica e económica) que ocupam. Neste sentido, é desenvolvido um modelo gravitacional (de potencial), onde as actividades de destino têm por base o PIB, e a impedância é descrita tomando em consideração os custos de distância, por estrada. Os resultados são expressos em termos de Índice de Potencial Económico, e representados no mapa da Europa através de curvas de nível (Figura 1).

Figura 1 – Acessibilidade na Europa: o Modelo de Keeble *et al.* (1982, 1988)

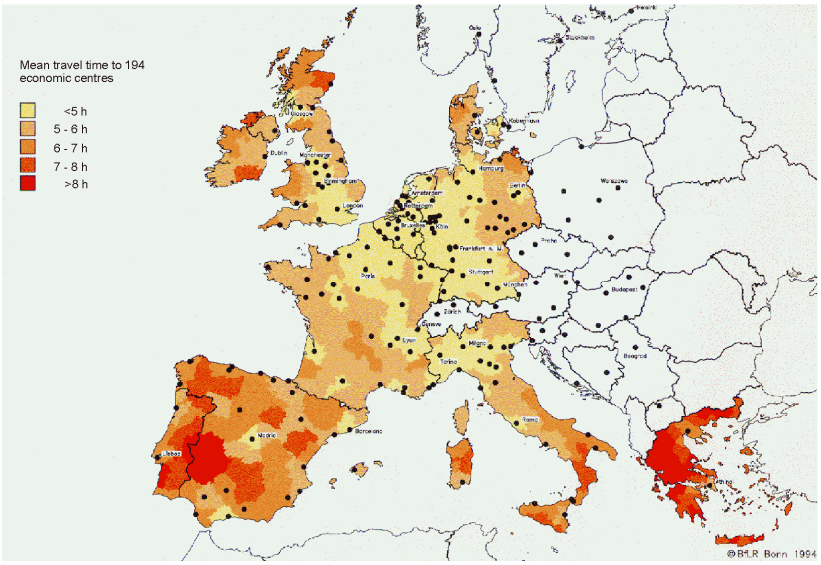


Fonte: Spiekermann (2002:7)

[4] Citados por Spiekermann e Neubauer (2002).

O Bundesforschungsanstalt für Landeskunde und Raumordnung (Lutter *et al.*, 1992, 1993)[5] - num trabalho desenvolvido para a DG REGIO da CE - avalia a acessibilidade das NUT III das regiões dos (então) doze Estados membros, em face do tempo médio de percurso (multimodal) que os separam de 194 centros económicos espalhados pelo território europeu. Neste estudo, são ainda tomados em consideração outros destinos: as três aglomerações populacionais, a paragem de comboio rápido e o aeroporto mais próximos. Simultaneamente, é calculado um indicador de acessibilidade diária, com base no número de pessoas que podem ser alcançadas em três horas, utilizando a ligação mais rápida. Este trabalho toma em atenção os modos de transporte rodoviário, ferroviário e aéreo, e em duas situações distintas: na presença e na ausência de investimentos em infraestruturas (então) planeados, tais como: auto-estradas, linhas de caminho de ferro de alta velocidade, e determinadas ligações aéreas (Figura 2).

Figura 2 – Acessibilidade na Europa: o Modelo de Lutter *et al.* (1992, 1993)

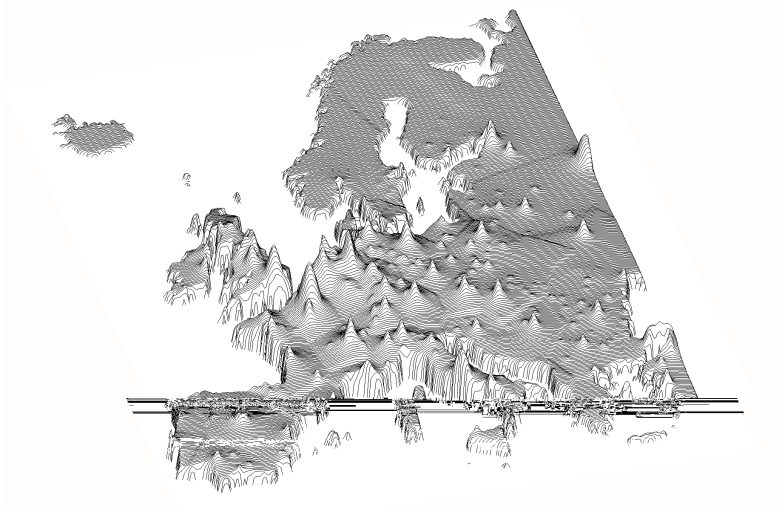


Fonte: Spiekermann (2002:7)

Spiekermann e Wegener (1994[6], 1996) desenvolvem plataformas tridimensionais, com base em tecnologia GIS, evidenciando a acessibilidade - diária e potencial - na Europa, numa primeira fase só para a rede ferroviária e, posteriormente, integrando também os modos rodoviário e aéreo. No modelo, as plataformas quase homogéneas são obtidas a partir da subdivisão da Europa em cerca de 70.000 quadrículas - de 10 km de largura - e calculando a acessibilidade de cada uma em relação a todas as outras. A afectação da população, por cada quadrícula, é feita a partir da informação sobre a existente em cada um das NUT III das várias regiões, e aplicando – por hipótese - determinado gradiente exponencial negativo à densidade populacional em redor dos centros populacionais. O tempo de viagem - no acesso de cada quadrícula ao nó mais próximo da rede – é estimado com base numa velocidade média de deslocação de 30 km/h (Figura 3).

[5] Citados por Spiekermann e Neubauer (2002).
[6] Citados por Spiekermann e Neubauer (2002).

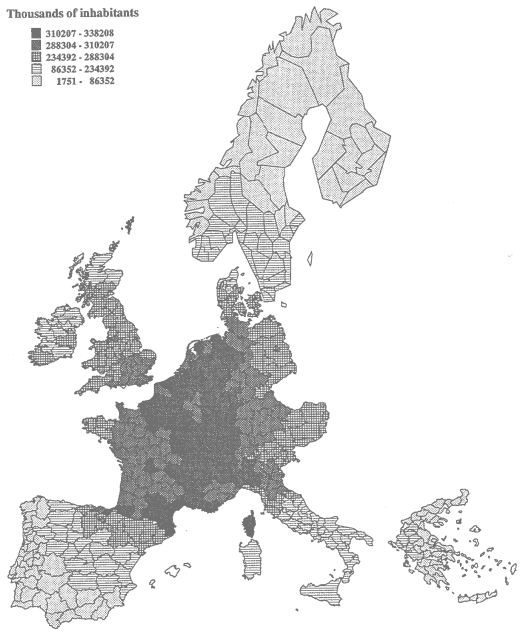
Figura 3 – Acessibilidade na Europa: o Modelo de Spiekermann e Wegener (1994, 1996)



Fonte: Spiekermann e Neubauer (2002:34)
e Spiekermann (2002:7)

No trabalho sob responsabilidade da UTS, Chatelus e Ulied (1995)[7] apresentam vários indicadores de acessibilidade para avaliar os impactes provocados pelas redes Transeuropeias nas NUT II das regiões da UE. Neste documento insere-se o indicador FreR(T) que, dirigido ao transporte de mercadorias determina, em simultâneo, quer a dimensão do mercado que pode ser alcançado num certo intervalo de tempo, quer a população que pode ser abrangida em um, dois ou três dias, através da mais rápida ligação rodoviária e/ou ferroviária; respeitando, inclusivamente, as restrições (legais) de tempo de condução impostas aos condutores dos camiões (Figura 4).

Figura 4 – Acessibilidade na Europa: o Modelo de Chatelus e Ulied (1995)

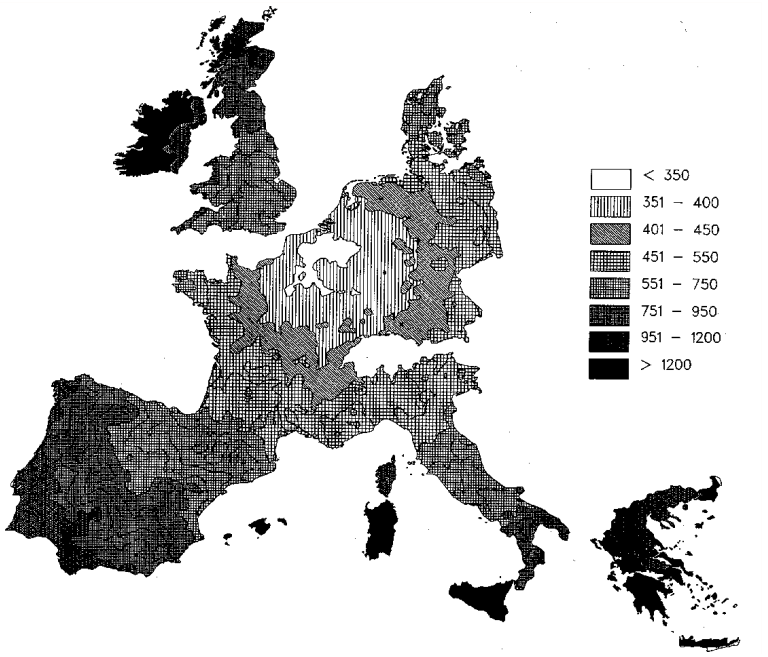


Fonte: Spiekermann (2002:7)

[7] Citados por Spiekermann e Neubauer (2002).

Em 1996, investigações desenvolvidas por Gutiérrez *et al.* e por Gutiérrez e Urbano debruçam-se sobre o tempo médio de deslocação – por estrada e por caminho de ferro – entre cerca de 4.000 nós de uma rede multimodal europeia de transporte, e 94 aglomerações populacionais com mais de 300.000 habitantes; e contemplando duas situações distintas: na presença e na ausência de investimentos em infraestruturas (então) planeados. Embora sujeitos à devida ponderação, os tempos de deslocação por estrada incluem, quer os momentos em que os camiões efectuam travessias em *ferry--boat*, quer as situações em que se torna incontornável a travessia de centros populacionais congestionados. Os tempos de deslocação por caminho de ferro incluem – para além dos constantes nos horários previstos - os dispendidos nos acessos rodoviários, e as penalizações inerentes: ao transbordo entre modos rodoviário-ferroviário (60 min) e ferroviário-marítimo/*ferry-boat* (180 min), e aos atrasos nas mudanças de bitola ferroviárias entre Espanha-França (30 min), (Figura 5).

Figura 5 – Acessibilidade na Europa: o Modelo de Gutiérrez e Urbano (1995, 1996)

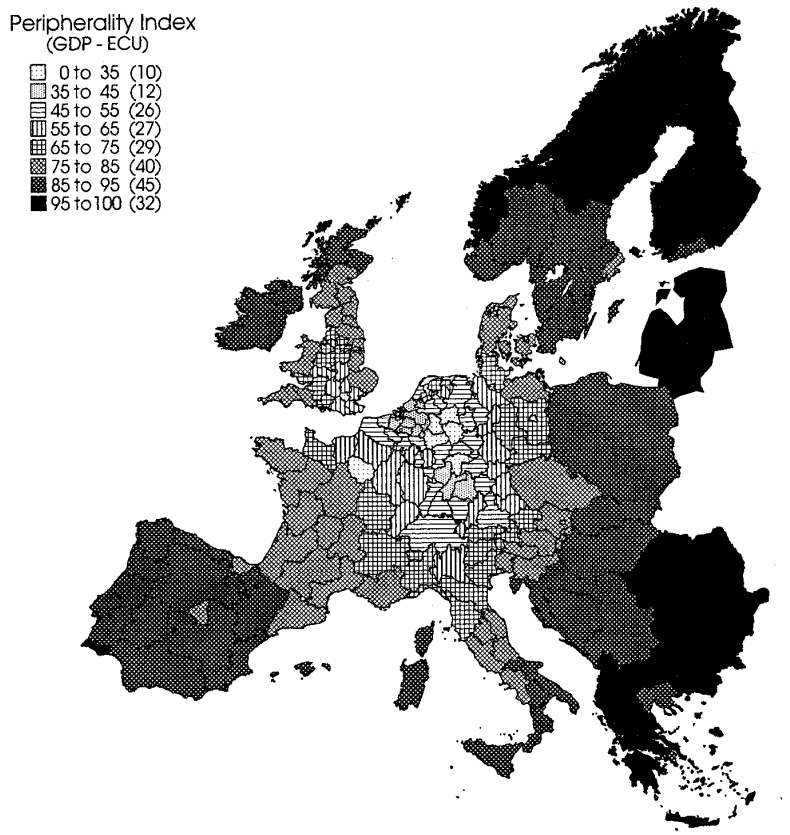


Fonte: Spiekermann (2002:8)

Em trabalhos efectuados para a DG REGIO da CE, Copus (1997, 1999)[8] apresenta indicadores de acessibilidade para as NUT II e NUT III das várias regiões, com base em "(…) road-based potential measures of the Keeble type", (Spiekermann e Neubauer, 2002:14). O modelo toma em consideração distintos valores de velocidade média consoante: o nível/qualidade de/da ligação rodoviária; o tempo (real), quer de travessia por *ferry-boat*, quer de acesso aos respectivos terminais; o (eventual) atraso provocado pelo atravessamento das fronteiras internas da UE; e o tempo de descanso consignado na Lei para os condutores dos camiões. Neste contexto, a acessibilidade é apresentada como um índice de perifericidade "(…) derived as the inverse standardised to the interval between zero (most central) and one hundred (most peripheral)", (Spiekermann e Neubauer, 2002:14), (Figura 6).

[8] Citados por Spiekermann e Neubauer (2002).

Figura 6 – Acessibilidade na Europa: o Modelo de Copus (1997, 1999)

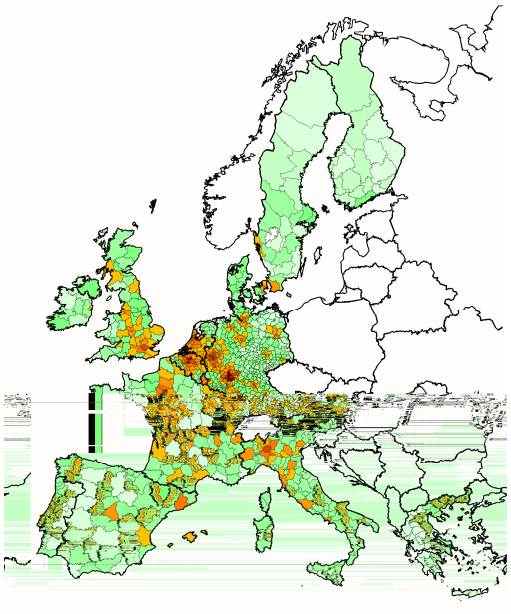


Fonte: Spierkermann (2002:8)

Também no âmbito da DG REGIO da CE, Wegener *et al.* (2000, 2002)[9] propõem a elaboração de indicadores de referência que descrevam a posição geográfica das NUT III das várias regiões europeias. Neste sentido – e para além das referências geográficas, físicas e culturais – surgem três indicadores: dois que avaliam a acessibilidade rodoviária e ferroviária face à população; e um terceiro, de aplicação ao modo aéreo, cuja preocupação central é a actividade económica, expressa em termos de PIB. Desta forma, a acessibilidade torna-se num indicador da verdadeira dimensão das áreas de mercado para vários agentes envolvidos: tanto para os fornecedores de bens e serviços - se correlacionada com os núcleos populacionais; como para os fornecedores "(…) of high-level business services", (Spiekermann e Neubauer, 2002:14) – se correlacionada com a actividade económica. Em qualquer das situações, a acessibilidade é apresentada sob a forma de um índice, cuja referência é a acessibilidade média europeia (Figura 7).

[9] Citados por Spiekermann e Neubauer (2002).

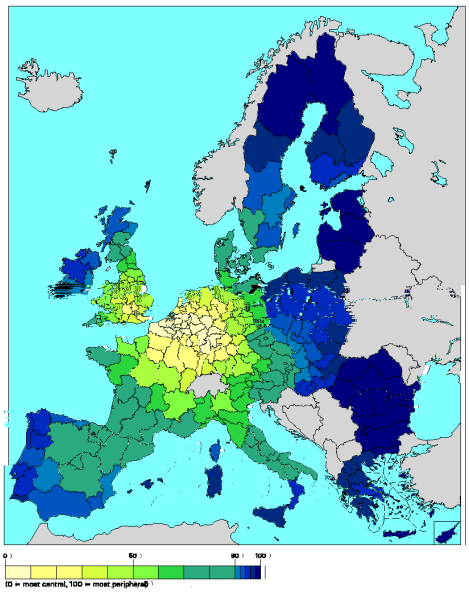
Figura 7 – Acessibilidade na Europa: o Modelo de Wegener *et al.* (2000, 2002)



Fonte: Spiekermann (2002:8)

A propósito do Relatório de Coesão elaborado pela CE em 2001, Schürmann e Talaat (2000)[10] apresentam, um ano antes, um índice (de potencial) de perifericidade incorporado num sistema de informação geográfica - com indicadores para o transporte rodoviário de pessoas e de mercadorias - e cujas actividades de destino são o PIB, a População ou a Mão-de-obra. Este trabalho abrange, não só as NUT III das várias regiões da UE, mas também os espaços territoriais equivalentes nos diversos países candidatos. Neste contexto, o índice de perifericidade é apresentado sob duas formas: com base na média europeia – tal como no modelo de Wegener *et al.* (2000, 2002); ou sob a forma de um intervalo entre 0 e 100 – como defende Copus (1997, 1999), (Figura 8).

Figura 8 – Acessibilidade na Europa: o Modelo de Schürmann e Talaat (2000)

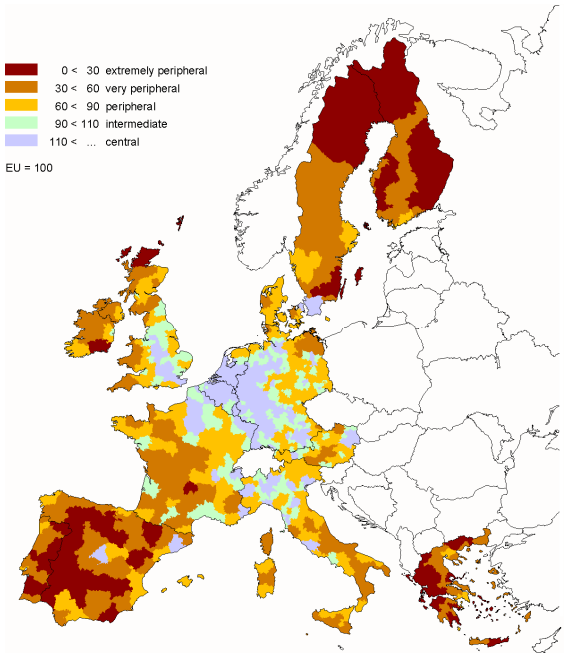


Fonte: Spiekermann e Neubauer (2002:8) e Spiekermann (2002:8)

[10] Citados por Spiekermann e Neubauer (2002).

Mais recentemente, Spiekermann *et al.* (2002)[11] - com base em cálculos logarítmicos da acessibilidade potencial - desenvolvem um indicador (multimodal) capaz de integrar as contribuições individuais de cada modo de transporte na definição de centralidade e de perifericidade. Fürst *et al.* (2000)[12], a propósito, defendem que: "(…) multi-modal indicators are considered to have much more explanatory power with respect to regional economic performance than any accessibility indicator based on a single mode only". Neste sentido, o modelo de Spiekermann *et al.* (2002) permite colocar em evidência a situação das NUT III em que se dividem as regiões da UE, graduando-as em: centrais, intermédias, periféricas, muito periféricas e extremamente periféricas (Figura 9).

Figura 9 – Acessibilidade na Europa: o Modelo de Spiekermann *et al.* (2002)



Fonte: Spiekermann (2002:14)

Passados em revista os (nove) modelos europeus que - na opinião de Spiekermann e Neubauer (2002) e de Spiekermann (2002) – traduzem os contributos mais significativos na matéria ao longo das duas últimas décadas, facilmente constatamos, não só o vasto leque de abordagens possíveis, como também o grau de importância atribuído ao conceito de acessibilidade.

Na opinião de Spiekermann e Neubauer (2002) – e apesar das diferenças que acabam por os caracterizar - os (nove) modelos de acessibilidade europeia gozam também de algumas características comuns; nomeadamente:

- mais de metade utiliza a Acessibilidade Potencial como Indicador Principal; os restantes socorrem-se, ou da Diária, ou da dos Custos de Deslocação. No entanto, alguns modelos estão aptos a apresentar cálculos com base em qualquer uma das metodologias;
- as Origens são, regra geral, Centróides, correspondentes às NUT III e NUT II das várias regiões que constituem os Estados membros da UE. Poucos são os que apresentam uma representação espacial mais detalhada;

[11] Citados por Spiekermann e Neubauer (2002).
[12] Citados por Spiekermann e Neubauer (2002:14).

- as actividades de Destino dividem-se, normalmente, por dois tipos de modelos: a População ou o PIB - para os de tipo Potencial, e um grupo pré-definido de aglomerações - no caso dos de Custo de Deslocação;
- na grande maioria dos casos – e em detrimento do Custo da Deslocação - o termo de Impedância escolhido é o Tempo de Viagem;
- os modelos que se debruçam sobre o Transporte de Mercadorias invocam, como restrição, as paragens (legalmente estatuídas) para descanso dos condutores;
- as Barreiras à deslocação surgem, regra geral, sob a forma de atrasos no atravessamento das fronteiras; apenas Keeble *et al.* (1882, 1988) fazem uma referência explícita às "(…) trade barriers", (Spiekermann e Neubauer, 2002:14);
- na sua grande maioria, os modelos são construídos com base no Transporte Individual; apenas dois – de entre eles – consideram, também, o Transporte de Mercadorias;
- cerca de metade dos modelos tomam em atenção, apenas um Modo de transporte - e, regra geral, o Rodoviário; os restantes apresentam soluções para diferentes modos, e alguns – de entre estes – utilizam mesmo redes Multimodais.

O Quadro 1 sintetiza as principais conclusões a retirar destes modelos relativamente, quer às disparidades espaciais, quer às alterações comportamentais ao longo do tempo.

Quadro 1 – Modelos de Acessibilidade para a Europa: Equidade e Dinâmicas

Autor(es)	**Disparidades Espaciais**	**Alterações Comportamentais ao Longo do Tempo**
Keeble *et al.* (1882, 1988)	Nítido padrão Centro-Periferia	As disparidades aumentaram, relativamente a períodos anteriores
Lutter *et al.* (1992, 1993)	Embora existindo, a respectiva dimensão/extensão depende das actividades de destino consideradas	As regiões periféricas experimentaram melhorias ao nível dos tempos de viagem, enquanto a acessibilidade diária aumentou nas regiões centrais
Spiekermann e Wegener (1994, 1996)	Nítido padrão Centro-Periferia, ao qual acrescem as disparidades Centro-Interior em todos os países Europeus	Aumentaram as disparidades, sobretudo devido à implementação das Redes Transeuropeias
Chatelus e Ulied (1995)	Nítido padrão Centro-Periferia	Diminuição das disparidades
Gutiérrez e Urbano (1995, 1996)	Nítido padrão Centro-Periferia	Diminuição das disparidades, sobretudo devido à implementação das Redes Transeuropeias
Copus (1997, 1999)	Nítido padrão Centro-Periferia	-
Wegener *et al.* (2000, 2002)	Diferentes padrões na dicotomia Centro-Periferia, dependendo dos modos de transporte em apreço	Aumento ou diminuição das disparidades, consoante o indicador em apreço
Schürmann e Talaat (2000)	Nítido padrão Centro-Periferia no caso do modo de transporte rodoviário	As melhorias são mais evidentes nos países candidatos à integração na UE
Spiekermann *et al.* (2002)	Nítido padrão Centro-Periferia, mas apresentando graus de perifericidade muito distintos	-

Fonte: Adaptado de Spiekermann e Neubauer (2002)

No que respeita à Disparidades Espaciais, repare-se como todos os modelos assinalam um nítido padrão Centro-Periferia; alguns sublinhando, inclusivamente, diferenças consideráveis de acessibilidade no interior das próprias regiões, centrais ou periféricas.

Quanto às Alterações Comportamentais ao longo do tempo, os modelos não são unânimes nas respectivas apreciações: os que se baseiam em indicadores como os Custos de Deslocação justificam o incremento da coesão pelos investimentos públicos em matéria de infraestruturas; outros, mais cautelosos – tais como os de Acessibilidade Diária ou Potencial – perspectivam um

aumento das disparidades regionais, precisamente em consequência dos investimentos em infraestruturas de transporte.

Em suma, uma avaliação global aos nove modelos de acessibilidade não é tarefa fácil, sobretudo porque nenhum deles é capaz de abranger, simultaneamente, todos os aspectos considerados mais relevantes. Repare-se, por exemplo, que grande parte deles toma por base o transporte individual de passageiros ignorando o de mercadorias - muito embora pareça ser este último, precisamente, o que mais importa em contexto das regiões periféricas. "However, empirical work has shown that road accessibility by using car or trucks is very highly correlated and that car accessibility can be used as a proxy for truck accessibility", (Spiekermann e Neubauer, 2002:15). Por outro lado – e regra geral - muitos dos modelos apresentam-se directamente relacionados, apenas com determinados sectores da economia. Deste modo – e por exemplo - ao darem especial destaque ao transporte individual, os modelos acabam por ser conotados sobretudo com o sector dos serviços, esquecendo as inúmeras ligações que o transporte estabelece, naturalmente, com outros sectores (Spiekermann e Neubauer, 2002).

Apesar do vasto leque de modelos ao nosso dispor, não há nenhum que satisfaça, simultaneamente, todos os requisitos nas mais variadas circunstâncias. Spiekermann e Neubauer (2002:15) afirmam, a propósito, que: "(…) models that are superior in a certain dimension are often behind in others"; e acrescentam: "(…) there is no model available that would be able to calculate accessibility for a spatially detailed representation of Pan-Europe for personal and freight transport for all transport modes including multi and inter-modal trips for different indicator types and destinations activities and that has a database that allows assessments for different points in time, i.e. past, current and future accessibility patterns".

3 PERIFERIA E DESENVOLVIMENTO

Na prática, a dicotomia Periferia-Desenvolvimento pode ser abordada de duas formas distintas (Silva, 2005):

- do ponto de vista meramente *estatístico* - e da experiência que colhemos no dia-a-dia - podemos afirmar que o conceito de periferia é, regra geral, sinónimo de baixa densidade populacional, de baixo índice de desenvolvimento, de dificuldades acrescidas no incremento das actividades económicas, e de baixo nível de produtividade;
- outra forma de olhar para a realidade é segundo uma abordagem *mais* analítica: embora não fugindo à realidade negativista da condição periférica, é necessário identificar, de igual modo, os sinais positivos da relação Periferia-Desenvolvimento, maximizando-os. Neste contexto – e contrariamente ao que é usual afirmar-se – "(…) "peripherality" should not necessarily mean a low level of development", (Giannopoulos e Bolougaris, 1995:260).

Ora, tanto na perspectiva sócio-económico como na territorial, o desenvolvimento das regiões depende, regra geral, dos tradicionais factores de produção: Trabalho, Terra e Capital. Nos espaços europeus mais periféricos escasseiam, precisamente, o Trabalho e o Capital - como factores de produção, enquanto a Terra parece querer assumir, progressivamente – sobretudo nas zonas rurais – um papel de *equilíbrio* numa sociedade cada vez mais urbanizada.

Se forem despoletados os necessários mecanismos de compensação pela falta, sobretudo de boas redes de transportes e de telecomunicações, tais espaços acabam por atrair mais habitação, mais emprego, mais actividades turísticas, e mais condições para o lazer, transformando-se – de forma sustentável - em reservas, paisagísticas e culturais, reforçando e valorizando – consequentemente – o papel basilar do espaço no seio das regiões periféricas.

Espaços deste tipo em regiões mais acessíveis - embora escassos - são muito disputados; porém, é nas regiões de menor acessibilidade que os podemos (ainda) encontrar, com relativa facilidade. Daí que Giannopoulos e Bolougaris (1995:260) defendam, precisamente, que: "(…) this is the first and perhaps most important impact of the provision of transport on the "development" of peripheral areas".

Quanto aos restantes factores de produção – Trabalho e Capital – também a este respeito a postura das regiões periféricas, no futuro, não terá de ser, necessariamente, de fragilidade. Aliás, as pressões económicas para a descentralização de actividades – afastando-as, progressivamente, das áreas centrais mais congestionadas com o argumento da protecção ambiental, bem como as pressões políticas para a descentralização dos processos de tomada de decisão, constituem-se como os sinais mais promissores de um *regresso* anunciado de ambos os factores, aos espaços periféricos.

O alcance das oportunidades que assim emergem em muito dependerá das condições locais que forem sendo criadas, "(…) since peripheral areas in Europe today are internally differentiated to the extent that it is increasingly difficult to analyze them in terms of "averages", (Giannopoulos e Bouloungaris, 1995:260). Além do mais, verificam-se amplos limites de variação em termos das características estruturais mais comuns, e antevêem-se desenvolvimentos algo divergentes em matéria de implementação de políticas e de processos. Um cenário assim traçado exige a concertação de estratégias locais, tanto ao nível da dotação em serviços, como na reestruturação das actividades económicas.

Giannopoulos e Bouloungaris (1995:261) defendem ainda que não deve ser perdido de vista "(…) the potential conflict with the ongoing infrastructure development policies for the peripheral areas, which are to an increasing degree, based on the priorities of the urbanized society and, more importantly, as seen from an international "central" perspective" – numa clara alusão ao ênfase que hoje se coloca em matérias como a Procura e a Competição, que chegam a constituir-se - para os espaços rurais - como mais uma dificuldade acrescida.

Relativamente a esta dicotomia Periferia-Desenvolvimento, Giannopoulos e Bouloungaris (1995) defendem precisamente a adopção de quatro linhas (prioritárias) de actuação; nomeadamente:

- para que as regiões se possam desenvolver, é necessário que saibam/possam utilizar o respectivo Potencial. Para tal, a modernização das redes de transporte e de telecomunicações é um requisito necessário, mas não suficiente: para além da criação de programas específicos de apoio financeiro, é preciso que os operadores das redes de serviços se consciencializem de que os investimentos em acessibilidades não são um fim, em si mesmo, mas um meio de facilitar a adopção de outras – e melhores – práticas, num ambiente cada vez mais competitivo;
- o aumento da procura é um factor essencial para estratégias de desenvolvimento dos espaços regionais. No âmbito de um programa integrador, de ampla abrangência regional, é necessário prestar às infraestruturas de carácter social – educação, cultura, etc. – uma atenção tão cuidada como a dada às de carácter eminentemente económico; o que pressupõe uma cooperação entre distintas entidades, a diferentes níveis;
- as obrigações de serviço público não devem ser tomadas como referência indiscutível na aplicação das políticas de desenvolvimento regional; muito embora, na prática, se devam analisar cuidadosamente determinadas medidas – tais como, por exemplo, a subsidiação de tarifas ou os serviços de custo diferenciado – à luz de circunstâncias específicas: recursos existentes, meios tecnológicos disponíveis, etc.;
- por forma a contrariar a inércia *centralizadora* do desenvolvimento tecnológico das redes de infraestruturas e de serviços, as regiões periféricas necessitam de um forte apoio, não só em termos gerais de intervenção pública, mas também em matéria de I & D – Investigação e Desenvolvimento, sobretudo ao nível dos Transportes e das Telecomunicações.

REFERÊNCIAS BIBLIOGRÁFICAS

ARMSTRONG, H. e J. TAYLOR (2000) *Regional Economics and Policy*, Third Edition, Oxford, Blacwell Publishers Ltd.

COSTA, P. (1996) Novas Formas de Organização do Espaço em Portugal: Uma Visão Estratégico-Prospectiva, in Associação Portuguesa para o Desenvolvimento Regional (ed), *IV Encontro Nacional da APDR*, Covilhã, Universidade da Beira Interior.

DIXON, R. e A. THIRLWALL (1975) A Model of Regional Growth Rate Differences on Kaldorian Lines, *Oxford Economic Papers*, 27, 201-214.

GIANNOPOULOS, G. (1994) Peripherality: Perceptions, Reality and the Role of Transport in Addressing the Issues – The Greek Perspective, in D. Evans e A. Smyth (eds), *European Peripherality: States of Mind?*, Ulster, University of Ulster, 39-51.

GIANNOPOULOS, G. e G. BOULOUGARIS (1995) Overcoming Isolation and the Role of Transport: the Case of the Aegean Islands, in H. Coccossis e P. Nijkamp (eds), *Overcoming Isolation. Information and Transportation Networks in Development Strategies for Peripheral Areas*, Berlin, Springer-Verlag, 254-268.

GUERRA, A. (1998) *Globalização e Competitividade. O Posicionamento das Regiões Periféricas*, Lisboa, Gabinete de Estudos e Prospectiva Económica do Ministério da Economia.

GUTIÉRREZ, J. e J. URBANO (1996) Accessibility in the European Union: the Impact of the Trans-European Road Network, *Journal of Transport Geography*, 4, 15-25.

GUTIÉRREZ, J., R. GONZÁLEZ e G. GÓMEZ (1996) The European High-Speed Train Network: Predicted Effects on Accessibility Patterns, *Journal of Transport Geography*, 4, 227-238.

HIRSCHMANN, A. (1958) *The Strategy of Economic Development*, New Haven, Yale UP.

HITIRIS, T. (1994) *European Community Economics,* Third Edition, Harvest-Weatsheaf.

HOTELLING, H. (1929) Stability in Competition, *Economic Journal*, 39, 41-57.

KALDOR, N. (1966) *Causes of the Slow Rate of Growth of the United Kingdom Economy*, Cambridge, Cambridge UP.

KALDOR, N. (1970) The Case for Regional Policies, *Scottish Journal of Political Economy*, 17, 337-347.

KEEBLE, D., J. OFFORD e S. WALKER (1988) *Peripheral Regions in a Community of Twelve Member States*, Luxembourg, Office for Official Publications of the European Communities.

KEEBLE, D., P. OWENS e C. THOMPSON (1982) Regional Accessibility and Economic Potential in the European Community, *Regional Studies*, 16, 419-432.

MYRDAL, G. (1957) *Economic Theory and Underdeveloped Regions*, London, Duckworth.

NUNES DA SILVA (1994) Dealing with Peripherality: a non Conformist Perspective, in D. Evans e A. Smyth (eds), *European Peripherality: States of Mind?*, Ulster, University of Ulster, 52-66.

PREDÖHL, A. (1928) The Theory of Location in its Relation to General Economics, *Journal of Political Economy*, 36, 371-390.

RIETVELD, P. e F. BRUINSMA (1998) *Is Transport Infrastructure Effective? Transport Infrastructure and Accessibility: Impacts on the Space Economy*, Berlin, Springer-Verlag.

RODRIGUES, C., E. CASTRO, R. CARVALHO, D. POLÓNIA, M. DUARTE e D. TRIGO (1996) Serviços Avançados de Telecomunicações nas Regiões Menos Favorecidas. O Projecto MOSAICC, in Associação Portuguesa para o Desenvolvimento Regional (ed), *IV Encontro Nacional da APDR*, Covilhã, Universidade da Beira Interior.

SILVA, J. (2005) *As Acessibilidades como Factor do Desenvolvimento de Regiões Periféricas: o Caso da Beira Interior*, Tese de Doutoramento em Transportes, (ISBN: 972-8790-54-6, UBI, 2006), Lisboa, Instituto Superior Técnico.

SPIEKERMANN, K. (2002) *Peripherality and Accessibility to the European Agglomeration Centres and International Markets*, Stockholm, Nordregio.

SPIEKERMANN, K. e J. Neubauer (2002) *European Accessibility and Peripherality: Concepts, Models and Indicators*, Stockholm, Nordregio.

SPIEKERMANN, K. e M. WEGENER (1996) Trans-European and Unequal Accessibility in Europe, *European Journal of Regional Development*, 4, 35-42.

VICKERMAN, R. (2000) Core-Periphery and the Cohesion-Competitiveness Debate, in Regional Science Association International (ed), *6th World Congress of the RSAI*, Lugano, RSAI.

EL PAPEL DE TRANSPORTISTAS Y SUS CLIENTES EN LA CONSECUCIÓN DE UN SISTEMA DE TRANSPORTE SOSTENIBLE

MARÍA DOLORES CARO VELA
E-mail: mdcarvel@upo.es
Universidad Pablo de Olavide

Resumen

Un sistema de transporte de mercancías eficaz y eficiente a la vez que compatible con el medioambiente es objetivo fundamental de la política de transporte europea desde hace aproximadamente dieciséis años. Desde entonces la política de transporte nacional y regional ha incorporado este objetivo a su planificación.
Una de las medidas principales que han desarrollado los planes para conseguir este objetivo, es el desarrollo y potenciación de modos de transporte menos contaminantes que la carretera, como el ferrocarril y el short sea shipping.
En este sentido, las políticas regionales se han dirigido fundamentalmente al desarrollo de las infraestructuras, sobre todo de los modos menos contaminantes. Sin embargo la existencia de infraestructura sólo es un requisito para que se opte por la intermodalidad, porque los agentes últimos que deciden sobre la utilización de los modos de transporte son cargadores y usuarios finales (demandantes) y empresas transportistas(oferentes), que eligen según los criterios de mercado. Este trabajo analiza la importancia de desarrollar por parte de la Administración un conjunto de medidas dirigidas estos agentes de mercado para lograr el éxito en políticas de transporte sostenibles con el medioambiente.

Palabras clave: política de transporte-medioambiente-planificación de transporte.
Área Temática: Economía de la Energía, del Ambiente, de los Transportes y de los Recursos Naturales.

Abstract

A freight transport system efficiently and effectively at the same time compatible with the environment is fundamental objective of European transport policy for about sixteen years. Since then the policy of national and regional transport has incorporated this goal planning.
One of the main measures that have developed plans for achieving this aim is the development and promotion of less polluting transport modes that road, rail and short sea shipping.
In this sense, regional policies have been directed primarily to infrastructure development, especially the less polluting modes.
However, the existence of infrastructure is only one requirement for the choice of intermodality, because the agents who decide on the recent use of transport modes, haulage contractors and customers to choose as market criteria.
This paper analyzes the importance of achieving a broad consensus and cooperation between different authorities and actors in these markets for success in sustainable transport policies to the environment.

Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

1. INTRODUCCIÓN

La finalidad de la política de transportes sostenible es crear sistemas de transportes que satisfagan las necesidades de nuestra sociedad desde el punto de vista económico, social y medioambiental. (European Commisionn, 2006).

Un sistema de transporte de mercancías eficaz y eficiente a la vez que compatible con el medioambiente es objetivo fundamental de la política de transporte europea desde hace

aproximadamente dieciséis años. Desde entonces la política de transporte nacional y regional ha incorporado este objetivo a su planificación.

El Libro Blanco de 2001 (European Comisión, 2001) fundamentó su contenido en el objetivo de desvincular el transporte del crecimiento económico y lograr un progresivo cambio modal (reequilibrio de la demanda del transporte a favor del transporte marítimo y ferroviario), fijando también especial atención en acometer la materia de la congestión y en fomentar la seguridad vial. Estos objetivos no se han conseguido, por lo que la propia Comisión ha elaborado una Revisión del Libro Blanco, (2006) que manteniendo los mismos objetivos, pretende paliar los errores de dicho Libro Blanco, principalmente ambicionando revitalizar el ferrocarril y marítimo con medidas más directas hacia estos modos.[1] La revisión del Libro Blanco apuesta por un conjunto de medidas de transporte más amplio y más flexible, buscando soluciones geográficamente diferenciadas en el que se apliquen métodos de legislación a medida o de cooperación mejorada.

La posición geográfica de las regiones, las características del comercio, la distancia de los envíos y los tipos de mercancías transportadas influyen decisivamente en las cuotas de mercado que ostenta cada modo en las distintas regiones. Vassallo y Fagan (2007) encontraron que casi el 83% de la diferencias en las cuotas de mercando del ferrocarril entre Europa y EE.UU. se debe probablemente a las diferencias inherentes o naturales, principalmente la geografía, los recorridos y los tipos de mercancía transportadA.

Basándonos en la orientación hacia la flexibilidad y enfoque geográficamente diferenciado, el artículo pretende analizar el caso particular del transporte de mercancías en la Comunidad Andaluza, estudiando las características del transporte y las medidas de política sostenible planificadas. ¿Son suficientes para alcanzar los objetivos propuestos?

Para ello en el apartado 2 se realiza un diagnóstico del sistema de transporte de mercancías en Andalucía con la finalidad de detectar los puntos débiles en materia medioambiental. El apartado 3 analiza las medidas de política de transporte sostenible planteadas en la planificación nacional y regional y en el apartado 4 se estudia el efecto que estas acciones pueden tener en el sistema de transporte en la Comunidad Andaluza y propone algunas medidas de política alternativas para mejorar la sostenibilidad del sistema, destacando el fortalecimiento de la competitividad de las empresas de logística andaluzas y la colaboración y consenso entre las diferentes Administraciones y los agentes de mercado para lograr el éxito en políticas de transporte sostenibles con el medioambiente.

2. DIAGNÓSTICO DEL SISTEMA DE TRANSPORTE DE MERCANCÍAS EN ANDALUCÍA. PERSPECTIVAS DE SOSTENIBILIDAD.

El sistema de transporte, tal y como funciona en la actualidad, es una amenaza para la sostenibilidad en el sentido de que contribuye a la emisión de gases de efecto invernadero, ruido y fragmentación del espacio urbano.

En Andalucía, al igual que en el conjunto de U.E. se está produciendo un intenso crecimiento del transporte en paralelo con el crecimiento económico y al incremento del grado de interrelación con las economías de nuestro entorno. El volumen global de las mercancías consumidas o producidas por la población andaluza en el año 2007 es aproximadamente de 600 millones de toneladas, incluyendo tanto las operaciones regionales como las exteriores y en éstas tanto las nacionales como las internacionales. Esta cifra supone casi triplicar la correspondiente a 1998, resultado de una tasa de crecimiento de casi el 13% anual acumulativo, valor que se ha acelerado hasta 15,6% en los últimos años. (Junta de Andalucía, 2008). La distribución modal de los tráficos de mercancías, en los años 2005 y 1998 se representa a continuación:

[1] Véase "Por una Europa en Movimiento. Revisión Intermedia del Libro Blanco del Transporte de 2001 de la Comisión Europea."

Gráfico 2.1. Distribución modal de toneladas de mercancías transportadas en Andalucía. Año 2005.

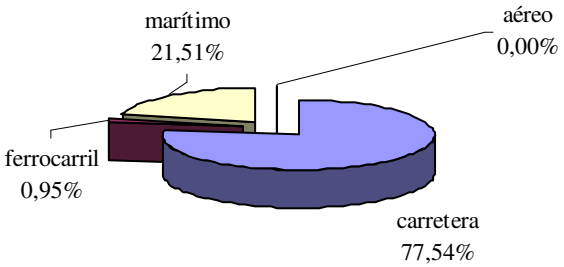
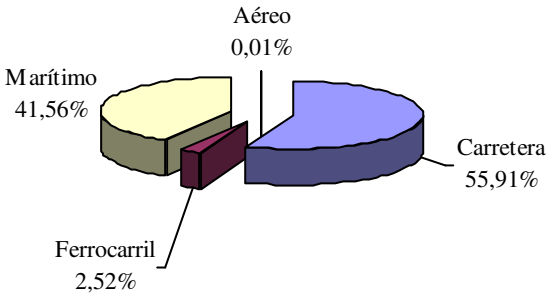


Gráfico 2.2. Distribución tráfico de mercancías en Andalucía. Año 1998



Fuente: Elaboración propia a partir de: Encuesta Permanente de transportes de mercancías por carretera. Años 1998 y 2005. Anuario Estadístico de Puertos del Estado. Año 2002. Anuario Estadístico de Andalucía. Año 2002 y 2006. **y** *Significación del transporte en la economía andaluza.* Consejería de Obras Públicas y Transportes. Año 2001.

El reparto modal revela la preponderancia de la carretera, que desplaza más de las tres cuartas partes de las mercancías transportadas. El transporte marítimo en segundo lugar, mueve algo más de una quinta parte del total, siendo poco relevantes los volúmenes movidos por el ferrocarril y, menos aún, por vía aérea si bien este último cobra significado en el transporte de determinas mercancías de elevada relación precio/peso.

Puesto que el transporte de mercancías por carretera es ampliamente mayoritario, más del 77% y es el modo menos sostenible, es interesante analizarlo detenidamente para estudiar si determinadas mercancías pueden dirigirse hacia otros modos de manera eficiente.

En los últimos años se ha desarrollado bastante la investigación sobre la intermodalidad. La U.E. ha subvencionado proyectos europeos que abordan esta temática incluidos en el IV, V y VI Programa Marco Comunitario de Investigación y Desarrollo Tecnológico; como ejemplos SPIN Project (Análisis del potencial del transporte intermodal), LOGIQ (Proceso de toma de decisiones en transporte intermodal), IQ (Mejorando la calidad del transporte intermodal),TRILOG, etc…

Los estudios concluyen que en la elección sobre el modo de transporte influyen entre otros los siguientes factores Distancia del transporte, frecuencia y regularidad de los envíos, tipología de la mercancía, cantidad que se transporta, compensación y estructura de la carga, tamaño de la empresa cargadora, organización logística, coste y calidad de los servicios y factores de entorno.

Uno de los factores principales es la distancia del transporte. Así, por ejemplo, Kreutzberger ,E (2008) concluye que el ferrocarril puede competir con la carretera a partir de distancias mayores a 600 kms en el transporte puerta a puerta en Europa. *

En la siguiente tabla podemos observar como en EU-15 el ferrocarril comienza a tener cuotas de mercado mayores al 9% a partir de 150-200 kms. , mientras que en España la cuota del 9% se alcanza a partir de los 500 kms[2].

Tabla 2.1. Distribución modal toneladas transportadas en EU-15. por distancias. Año 2001

Distancias (kms) U.E.	Carretera	Ferrocarril	Fluvial
0-49	96,99%	2,03%	0,99%
50-149	92,70%	4,30%	3,01%
150-499	88,38%	**9,02%**	2,60%
500-	88,78%	**10,47%**	0,75%

Distancias (kms) España	Carretera	Ferrocarril
0-49	99,75%	0,25%
50-149	99,57%	0,43%
150-499	94,91%	**5,09%**
500-	90,86%	**9,14%**

Fuente: Energy and Transport in figures. 2003. Elaboración propia

Aunque no conocemos la distancia de los recorridos en la Comunidad Andaluza, disponemos de los datos atendiendo al ámbito de actuación. Se distingue entre transporte intrarregional (dentro de la región de Andalucía), interregional (transporte con otras Comunidades Autónomas) e internacional. En 2007 la distribución es la que sigue:

Tabla 2.2. Distribución del transporte de mercancías por carretera atendiendo al ámbito de actuación

Ámbito	Operaciones(%)	Carga(t)	Movilidad(t-km)
Internacional	0,84%	1,60%	18,64%
Interregional	9,22%	14,00%	55,33%
Intrarregional	89,93%	84,41%	26,03%

Fuente: Encuesta Permanente de Transporte de Mercancías por Carretera 2007. Elaboración propia.

2.1. Transporte Internacional

El transporte internacional es el que presenta menor volumen de operaciones de transporte, carga y movilidad.

Los flujos de transporte de mercancías por carretera entre Andalucía y Europa a través de Irán y la Junquera se presentan a continuación:

[2] Puede ocurrir que las diferencias en las cuotas de mercado del ferrocarril entre U.E. y España se deban a la situación geográfica, las distancia de los recorridos y los tipos de mercancía transportada, así como la diferencia entre los anchos de vía entre ambos territorios.

Tabla 2.3.

Tráfico Internacional por carretera entre Andalucía y Europa			
	Exportaciones (Tn)	**Importaciones (Tn)**	**Total (Tn)**
Andalucía-Europa	**3.521.029**	**1.741.600**	**5.262.629**
	Más de **4.300** Camiones/ Semana	Más de **2.100** Camiones/Semana	Más de **350.000** Camiones/Año

Fuente: Observatorio de los Pirineos 2004.

Los principales destinos de las mercancías son Alemania, Francia y Reino Unido.

La Junta de Andalucía estima que del total de 5.262.629 Tns de transporte internacional por carretera en 2004, aproximadamente 35% se podría captar por el transporte intermodal (carretera+ferrocarril, carretera-marítimo) lo que supondría el 0,55% del total de carga en Andalucía que se transporta actualmente por carretera.

Teniendo en cuenta que el 70% de los productos que se transportan por carretera en el ámbito internacional son de origen animal o vegetal y corresponden a productos perecederos, frescos o refrigerados como hortalizas, tomates, frutas, etc., la dificultad está en encontrar un sistema intermodal fiable para las mercancías refrigeradas, puesto que este tipo de mercancías exige también un alto grado de fiabilidad.

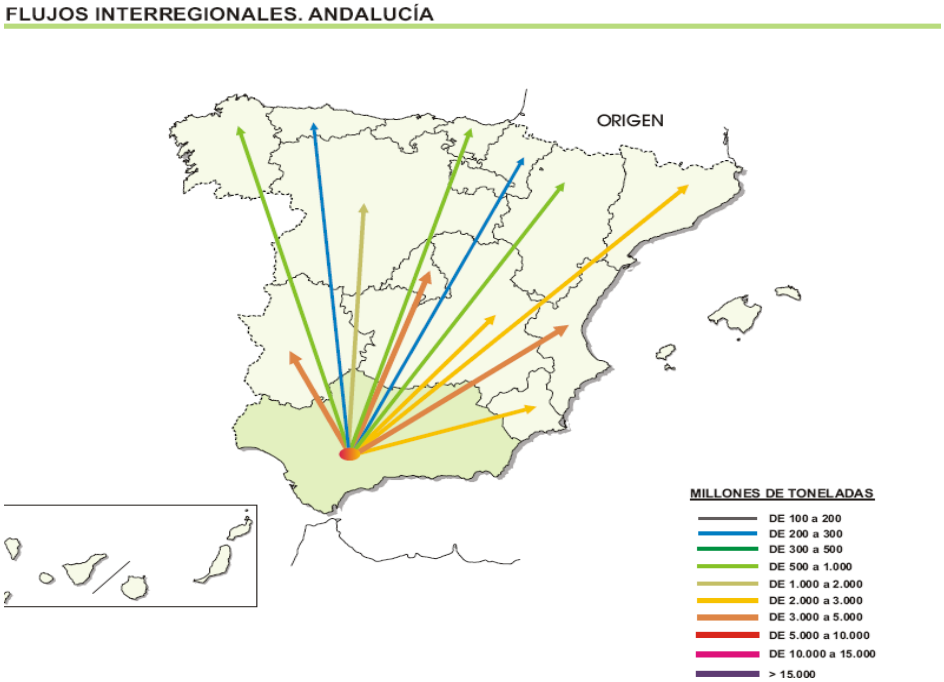
Una práctica de utilización del ferrocarril en este tipo de mercancías es el mercado de Saint Charles en Perpignan, que reexpide más de un millón y medio de toneladas de fruta y hortalizas frescas, envía trenes de carga hasta París (Rungis) que ahorran tiempo y descongestionan las autopistas francesas. En Suiza también se ha desarrollado este tipo de transporte para reducir el paso de camiones por los Alpes.

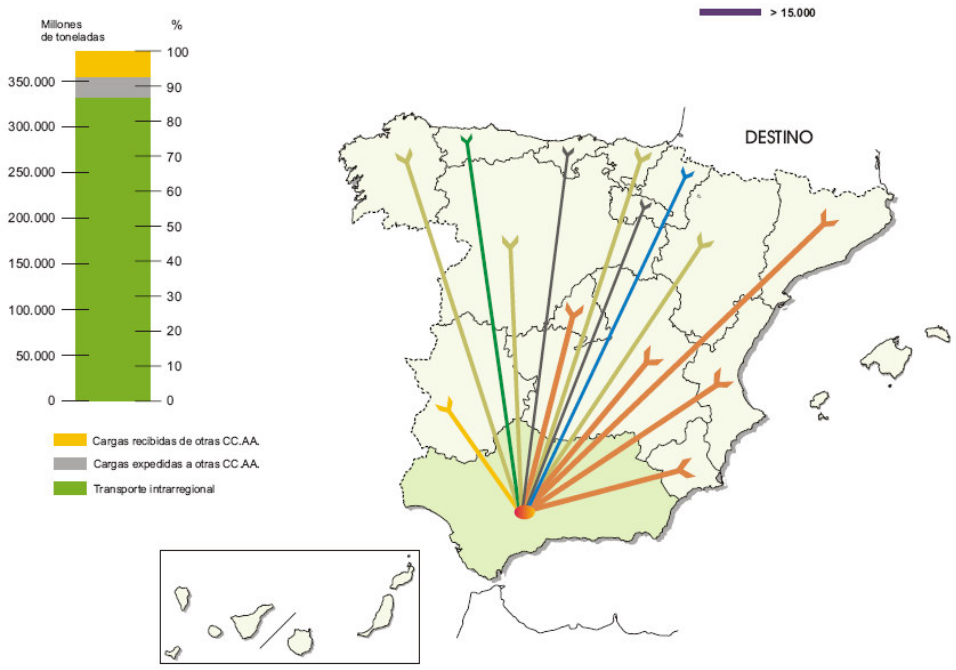
2.2. Transporte Interregional

Aproximadamente el 99% de las operaciones de transporte y más del 98% de las toneladas transportadas por carretera corresponden al tráfico nacional.

El transporte interregional representa 9,22% de las operaciones y 14% de la carga transportada. Los principales flujos de Andalucía con las demás regiones se representan en el siguiente mapa:

Gráfico 2.3



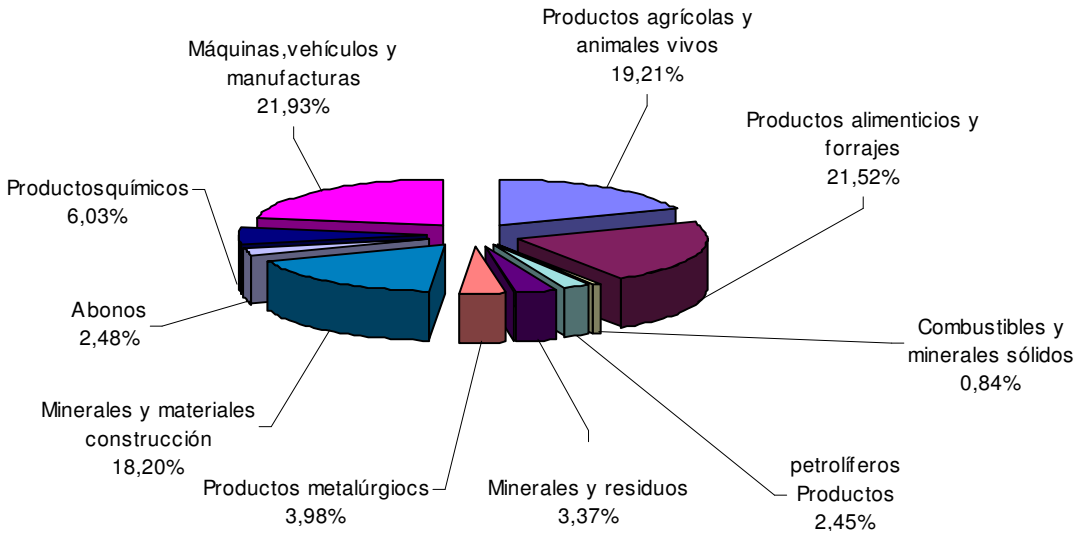


Fuente: Encuesta Permanente de Transportes de Mercancías por Carretera 2007.

Los principales puntos de origen y destino del transporte interregional en Andalucía por carretera ordenados de mayores a menores flujos son: Extremadura, Comunidad Valenciana, Comunidad de Madrid, Castilla-La Mancha, Murcia y Cataluña,

En lo que se refiere a los tipos de productos, los principales expedidos son: Productos alimenticios y agrícolas (40,73%), materiales de construcción (18,2%) y máquinas, vehículos y objetos manufacturados (21,93%).

Gráfico 2.4. Transporte interregional. Mercancías expedidas. Andalucía. Año 2007.



Fuente: Elaboración propia a partir de: Encuesta Permanente de transportes de mercancías por carretera. Año 2007

El tráfico interregional en Andalucía se caracteriza por flujos de relativamente corto recorrido, de productos agrícolas y alimenticios que necesitan un alto grado de fiabilidad y rapidez en el transporte y materiales de construcción con bastante dispersión en los destinos y por tanto con escasas posibilidades de utilizar el transporte combinado.

2.3. Transporte Intrarregional

Andalucía se caracteriza por ser la Comunidad Autónoma española con mayores porcentajes de flujos dentro de la propia región. **El tráfico intrarregional** representa en torno al 89% de las operaciones de transporte por carretera en Andalucía y 85% de las toneladas transportadas. Esta característica se acentuó en el periodo 2004-2006, coincidiendo con el alza del sector de la construcción, gran consumidor de transporte por carretera de corto recorrido. En los últimos dos años la ratio intrarregional ha comenzado a descender a consecuencia de la crisis en la construcción.

El recorrido medio de las mercancías movidas en el ámbito intrarregional no llega a los 40 kilómetros lo cual indica que la gran mayoría de estos movimientos no rebasa el ámbito provincial; de hecho, el 86,1% de las toneladas movidas en el interior de la región lo son internamente a cada provincia. En este tipo de transportes predominan los recorridos en vacío, que representan ente el 47% y 49% del total de las operaciones de transporte.

El tipo de mercancía más frecuente en transporte intrarregional e intramunicipal son los materiales de construcción seguido muy de lejos por máquinas, vehículos y objetos manufacturados.[3]

Podemos concluir por tanto que el transporte de mercancías en Andalucía está dominando por muy cortos recorridos, lejos de los 500-600 kms. A partir de los cuales el ferrocarril comienza a ser competitivo en el transporte puerta a puerta. Si los centros de producción están cercanos a los recursos o mercados, el uso del transporte intermodal es muy difícil, debido al aumento del coste que produce el acarreo terrestre entre terminales.

3. LA PLANIFICACIÓN SOSTENIBLE DEL TRANSPORTE DE MERCANCÍAS. UNA APLICACIÓN A ANDALUCÍA.

En este apartado se analizan los objetivos y medidas de política de transporte sostenible planteadas en la planificación nacional y regional, estudiando el caso particular de la Comunidad Autónoma andaluza.

El Plan de Infraestructura para la Sostenibilidad del Transporte en Andalucía, PISTA 2007-2013, es un instrumento de planificación ideado para hacer compatible un cierto desarrollo infraestructural en el sistema de transporte con los requerimientos de la sostenibilidad.

Se enmarca dentro de la planificación del transporte en El Libro Blanco de la UE (revisión 2006) y el Plan Estratégico de Infraestructura y Transporte (PEIT) 2005-2020 del Gobierno de España. Entendiéndose las actuaciones del PISTA correspondientes a la Administración Central dentro del marco temporal de la Planificación del Estado.

Los objetivos del PISTA referentes al transporte de mercancías son los siguientes:

- Consolidar un sistema intermodal de transporte de mercancías que actúe como factor al servicio del desarrollo territorial y productivo de Andalucía, garantizando el uso combinado de los modos más eficientes y dando prioridad al transporte ferroviario y marítimo frente a la carretera.
- Desarrollar el sistema de áreas logísticas de Andalucía como red interconectada de nodos multimodales, dotándola de las infraestructuras que permitan una correcta articulación con el conjunto de los sistemas de transportes y desarrollar sus funciones de intermodalidad.
- Garantizar la accesibilidad del territorio y los sistemas productivos a la red de centros logísticos para el transporte de mercancías.

[3] Un análisis más detallado del transporte por carretera en Andalucía se puede ver en Caro, M.D. "Oportunidades de la intermodalidad carretera-ferrocarril en el transporte de mercancías en Andalucía.". Consejería de Obras Públicas y Transportes. Junta de Andalucía.

- Impulsar estrategias de cooperación, complementariedad y coordinación en las políticas relacionadas con el transporte de mercancías entre los puertos comerciales y los aeropuertos andaluces.
- Fortalecer el sector del transporte de mercancías y desarrollar el mercado, adaptando las entidades operadoras públicas y privadas a las nuevas realidades y objetivos e impulsando la creación de empresas operadoras logísticas de ámbito europeo mediante las oportunas alianzas o acuerdos.
- Fortalecer el papel de Andalucía en el transporte de mercancías y el desarrollo de la logística con el Norte de África.

Fundamentalmente son objetivos dirigidos hacia la creación de infraestructuras que potencien otros modos de transporte como alternativa o complementarios a la carretera.

Las medidas específicas en que se concretan estos objetivos son:

Desarrollo de la Red de Áreas Logísticas de Andalucía, conexión exterior del sistema logístico, acciones para el desarrollo de intermodalidad potenciando el transporte ferroviario, marítimo, aeropuertos y conexión de redes viarias y ferroviarias. Algunas medidas transversales son: Programas para mejorar la formación y el uso de nuevas tecnologías, más control de los vehículos para mejorar la seguridad, potenciar la implantación de Sistemas de Gestión Ambiental y potenciar el conocimiento del sector.

La gran apuesta del PISTA en transporte de mercancías es la ordenación de una red de áreas y zonas logísticas que, con características de intermodalidad, sea capaz de procurar un cambio modal y una mejora de la eficiencia en el transporte de mercancías.

Gráfico 3.1. Red de Áreas logísticas Andalucía



Fuente:Red Logística de Andalucía. Consejería de Obras Públicas y Transporte. Junta de Andalucía.2008.

Los nodos se convierten en centros de producción y de distribución regionales, con equipamientos intermodales y bien conectadas con las redes de transporte nacionales e internacionales.

Se ubican en los principales puntos de producción-distribución de las mercancías y con conexiones con las grandes redes ferroviarias y viarias.

4. EFECTOS ESPERADOS DE LAS MEDIDAS DE SOSTENIBILIDAD PROPUESTAS EN EL PISTA

La inversión en la red de Áreas Logísticas de Andalucía puede tener efectos positivos en cuanto a la competitividad de la economía y el territorio andaluz. Estas Plataformas Logísticas favorecen por una parte la productividad en la manipulación de las mercancías y garantizan una buena conectividad entre la carretera y el resto de los modos de transporte (ferrocarril, marítimo, etc.) a la vez que permiten la concentración de actividades que generan valor añadido, lo que las convierte en motor de desarrollo económico.

De esta forma se pretende captar parte de los tráficos de trasbordo del Estrecho de Gibraltar para fijar actividades generadoras de valor añadido en la Comunidad, (actualmente solo permanece en el Puerto de Algeciras el 10% de las 1.362.889 TEUS que se trasbordan) así como derivar al Transporte Marítimo de Corta Distancia parte de los tráficos de paso Norte de África-Europa, generando mayor actividad en Andalucía. Esto supondría un impulso al tráfico internacional. En este contexto el ferrocarril tiene la oportunidad de combinarse con la carretera y marítimo para realizar unos servicios de transporte altamente competitivos.

Sin embargo, la construcción de la red de áreas logísticas intermodales no garantiza por si misma la mejora en la sostenibilidad ambiental[4].

Por una parte, porque como hemos expuesto en el apartado 2, el transporte de mercancías que actualmente se realiza por las carreteras andaluzas tiene pocas oportunidades para pasar a transporte intermodal. Las. características de los flujos de transporte hacen poco atractivo el cambio hacia otros modos.

Por otra parte porque los agentes últimos que deciden sobre la utilización de los modos de transporte son cargadores y usuarios finales (demandantes) y empresas transportistas(oferentes), que eligen según los criterios de mercado. Aunque el sector público puede regular para crear las condiciones marco adecuadas con la finalidad de que las empresas consigan un funcionamiento eficiente y de alta calidad de las cadenas logísticas. (European Comisión, 2007).

En las últimas décadas en EEUU y la UE la imposibilidad social de mantener las políticas clásicas de infraestructuras, sobre todo en el transporte urbano de viajeros, llevó a la formulación de políticas de otros tipos:Políticas de oferta, demanda y ordenación del Territorio. Los mejores resultados se han conseguido en aquellos territorios que han aplicación simultánea un conjunto de políticas. Se trata de aplicar ciertas medidas que favorecen los modos y las prácticas ambientalmente positivas y al mismo tiempo otras que dificultan y penalizan las conductas o los modos perjudiciales. (Torres, V. 2007).

Sería interesante formular un conjunto de medidas dirigidas hacia oferentes y demandantes de transporte de mercancías

4.1. Las Políticas de Oferta

Buscan maximizar el rendimiento de las infraestructuras existentes. En este sentido se podrían diseñar políticas dirigidas a:

Mejorar la organización logística:

1. Apoyar la adquisición de mayor tamaño por parte de los operadores y empresas logísticas en Andalucía para que puedan adquirir la dimensión mínima que les permita ofrecer unos servicios de transporte intermodales. [5]

2. Mejorar la comunicación e intercambio de información entre todos los agentes de la cadena logística.

[4] Diferentes autores europeos (como los Ingenieros de Caminos Offner, francés, o Izquierdo, español), contrastando los efectos esperados de diversos proyectos con los reales, han llegado a calificar de "mito de las infraestructuras", esa creencia ampliamente arraigada que tiende a supervalorar los efectos positivos de las mismas en el desarrollo económico, y a ignorar sus impactos negativos. Los efectos globales de su construcción y funcionamiento pueden ser muy diferentes, si no contradictorios, con los objetivos buscados con su construcción, por la gran absorción de recursos de inversión y por una serie de efectos cruzados.

[5] La media de vehículos por empresa en el transporte de mercancías en Andalucía está en torno al 3.

3. Mejorar la formación en el sector para que, por ejemplo, los agentes adquieran la capacidad necesaria para utilizar las TICS.

4. Estudiar fórmulas organizativas y apoyar la cooperación entre modos de transporte

5. Facilitar el acceso a los programas de ayudas de la UE para la puesta en marcha de servicios de transporte intermodales.

Mejoras tecnológicas:

Apostar por mejoras tecnológicas que permitan mejor uso de vehículos y sistemas de transferencia entre modos de transporte.

4.2. Políticas de demanda:

Pretenden reducir la movilidad o modificar las tendencias de la demanda para reducir el uso del transporte por carretera, desviando a los usuarios hacia otros modos de transporte. menos contaminantes. La Gestión de la demanda de Transportes incentiva ciertas formas de consumir o actuar penalizando otras.

En este sentido la **educación** es una importante herramienta de política para fomentar el el cambio de conducta del usuario hacia un transporte sostenible. Se podría introducir en los planes de formación. o Otras opciones pueden ser campañas de información y sensibilización

Actualmente se está estudiando por parte de la U.E. las medidas para internalizar en los precios de los servicios los costes externos medioambientales. Se analiza esta medida para concienciar a los usuarios sobre los costes medioambientales y transformar gradualmente el sistema de transporte. El flujo de ingresos obtenidos por la utilización de los modos de transporte más contaminantes se propone invertirlo en la financiación de infraestructuras más sostenibles medioambientalmente.

5. CONCLUSIONES

Un sistema de transporte de mercancías eficaz y eficiente a la vez que compatible con el medioambiente es objetivo fundamental de la política de transporte europea desde hace aproximadamente dieciséis años. Desde entonces la política de transporte nacional y regional ha incorporado este objetivo a su planificación.

La revisión del Libro Blanco (2006) apuesta por conseguir este objetivo con un conjunto de medidas de transporte más amplio y más flexible, buscando soluciones geográficamente diferenciadas en el que se apliquen métodos de legislación a medida o de cooperación mejorada.

Basándonos en la orientación hacia la flexibilidad y enfoque geográficamente diferenciado, el artículo pretende analizar el caso particular del transporte de mercancías en la Comunidad Andaluza, estudiando las características del transporte y las medidas de política sostenible planificadas, preguntándose si son suficientes para alcanzar los objetivos propuestos.

La planificación estratégica de infraestructura y transporte en Andalucía tiene como objetivo principal la creación de infraestructuras que potencien otros modos de transporte como alternativa o complementarios a la carretera, concretándose fundamentalmente en la creación de una Red de Areas Logísticas intermodales en Andalucía.

Sin embargo, el transporte de mercancías que actualmente se realiza por las carreteras andaluzas tiene pocas oportunidades para ser captado por el transporte intermodal. Las. características de los flujos de transporte hacen poco atractivo el cambio hacia otros modos. Por otra parte, los agentes últimos que deciden sobre la utilización de los modos de transporte son cargadores y usuarios finales (demandantes) y empresas transportistas(oferentes), que eligen según los criterios de mercado.

Concluimos que la construcción de infraestructuras logísticas no garantizan por si misma la mejora en la sostenibilidad ambiental por lo que se propone además un conjunto de medidas

dirigidas a que los agentes de mercado realicen un uso más eficiente de los modos de transporte sostenible.

6. BIBLIOGRAFÍA

BADÍA, M. (2007). "La contribución de las infraestructuras de transporte al desarrollo de la co-modalidad: la Red Logística de Andalucía" Jornada técnica sobre la red logística de Andalucía y los sistemas multimodales de transportes. Sevilla, Septiembre 2007

COMISIÓN DE LAS COMUNIDADES EUROPEAS.(2007) "Plan de acción para la logística del transporte de mercancías" COM/2007/0607 final .

COMISIÓN DE LAS COMUNIDADES EUROPEAS. (2006) "Por una Europa en movimiento. Movilidad sostenible para nuestro continente. Revisión intermedia del Libro Blanco del transporte de la Comisión Europea de 2001" CONM 314.

COMISIÓN EUROPEA. (2003). "Europa en la encrucijada. La necesidad de un transporte sostenible". Commission of the European Communities "Study on Freight Integrators". (2003). Berlin.

COMISIÓN EUROPEA. DIRECCIÓN GENERAL DE ENERGÍA Y TRANSPORTE. (2003) "Energy Transport in Figures", 2003.

CONSEJERÍA DE OBRAS PÚBLICAS Y TRANSPORTES. JUNTA DE ANDALUCÍA. (2008). "Plan de Infraestructuras para la Sostenibilidad del Transporte en Andalucía (PISTA 2007-2013)" Sevilla 2008.

DIRECCIÓN GENERAL DE PROGRAMACIÓN ECONÓMICA Y PRESUPUESTARIA. MINISTERIO DE FOMENTO.(1995-2007) "Encuesta Permanente de Transporte de Mercancías por Carretera. Años 1995-2007." Madrid.

DUQUE, J.L. (2000) "Significación del transporte en la economía andaluza", Consejería de Obras Públicas y Transportes, Junta de Andalucía.

EUROPEAN COMISIÓN, (2009). "The Future of Transport". Focus Group`Report. Brusselas.

EUROPEAN COMISIÓN, (2001). "White Paper, European Transport Policy for 2010: Time to Decid". COM370.

EUROPEAN ENVIRONTMENT AGENCY (2009). "Transport at a crossroads". N3/2009. EEA Report.

GARCÍA-MENÉNDEZ, L. MARTÍNEZ-ZARZOSO, I. and PIÑERO DE MIGUEL, D. (2004) "Determinants of mode choice between road and shipping for freight transport." Journal of transport economics and Policy. Volume 38, Part 3. September 2004. pp.447-466.

INSTITUTO ESTADÍSTICO DE ANDALUCÍA. CONSEJERÍA DE ECONOMÍA Y HACIENDA. (2003-2006) Anuario Estadístico de Andalucía. Años 2003-2006. Sevilla.

KREUTZBERGER,E. (2008)"Distance and time in intermodal goods transport networks in Europe: A generic approach". Transportation Research Part A 42 pp. 973–993

MINISTERIO DE FOMENTO (2005) "La aportación del transporte por carretera a la intermodalidad" Madrid.

MINISTERIO DE FOMENTO. (2004) "Plan Estratégico de Infraestructuras y Transportes. PEIT" Madrid.

MINISTERIO DE FOMENTO.(2001) "Plan Estratégico para el sector del transporte de mercancías por Carretera". Dirección General de Transportes por Carretera. Madrid.

PUERTOS DEL ESTADO.(2002) "Tráfico Portuario 2001". Anuario Estadístico de Puertos.

TORRES, V. (2007)"De la política de infraestructuras a la movilidad sostenible"

VICENTE CASTEJÓN. JORNADAS SOBRE TRANSPORTE, Las Palmas, 10 de Mayo, 2007

TSAMBOULAS,D. et AL (2007) "Assessment of a transport policy potential for intermodal mode shift on a European scale"Transportation Research Part A. 41 pp.715-733.

VASSALLO, J.M. and FAGAN, E. (2007). "Nature or nurture: why do railroads carry greater freight share in the United States than in Europe?" Transportation, 34. pp.177-193.

Impacto de Variáveis Socioeconómicas sobre as Perdas Comerciais de Energia Elétrica

KARLOS EDUARDO ARCANJO DA CRUZ
E-mail: karlosarcanjo@hotmail.com
FRANCISCO S. RAMOS
E-mail: fsr@ufpe.br
Universidade Federal de Pernambuco

Resumo

A tarifa de energia elétrica brasileira é das mais caras do mundo, decorrente principalmente dos impostos e os altos índices de perdas de energia. Estas perdas decorrem principalmente do furto de energia. A maior parte da literatura sobre perdas comerciais faz uma abordagem qualitativa sobre os fatores socioeconômicos relacionados ao furto de energia, tais como violência e renda, não se detendo o suficiente em mensurar quantitativamente o impacto dessas variáveis. Este trabalho desenvolve um modelo para analisar a influência das perdas comerciais nas empresas do Brasil e outro modelo para analisá-la entre países. No primeiro modelo encontrou-se uma correlação positiva do índice de perdas totais com a violência e uma correlação negativa com o nível de organização da empresa, a renda per capita e o percentual de clientes rurais. O segundo modelo encontrou uma correlação negativa com o percentual de habitantes rurais e com a eficiência do país.

Palavras-chave: Perda Comercial, Furto de Energia, Violência, Eficiência.
Área Temática: Economia da Energia, do Ambiente, dos Transportes e dos Recursos Naturais.

Abstract

The price of Brazilian electric energy is one of the highest of the world, due mainly to the taxes and the high indexes of losses. These losses are caused mainly by the theft of energy. Most of the literature on commercial losses makes a qualitative approach about the socioeconomic factors related to the theft of energy (violence, income), and not enough about measuring the quantitative impact. This work shows a model to analyze the influence of commercial losses in the Brazilian electrical companies and another model to analyze commercial losses among countries. The first model shows a positive correlation of the index of total losses with the violence and a negative correlation with the level of organization of the company, the per capita income and the percentile of rural customers. The second model found a negative correlation with the percentile of rural inhabitants and with the efficiency of the country.

Key words: Commercial loss of energy, theft of energu, violence, efficiency
Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

1. INTRODUÇÃO

O Brasil tem a tarifa de energia elétrica (EE) entre as mais caras do mundo, e um dos principais motivos para isso, evidentemente, são os encargos e os tributos (CINTRA, 2007). Entretanto, um outro fator que também merece ser abordado se refere às perdas. Segundo dados do *World Development Indicators* – WDI, o Brasil perdeu 16,85% da energia adquirida no ano de 2004, em que a média mundial foi de aproximadamente 8,83% (WORLD BANK, 2004). Esta perda é calculada pela diferença entre a energia injetada no sistema e a energia faturada. Uma parcela das perdas de EE ocorre devido a interações físicas da corrente elétrica com os materiais – são as perdas técnicas. Outra parcela é a energia consumida por alguns domicílios, mas não faturada – são as perdas comerciais de energia ou perdas não-técnicas.

As perdas comerciais podem ocorrer devido a erros de medição, fraudes nos medidores, ligações clandestinas, problemas administrativos das concessionárias, entre outros. No entanto, a ANEEL (2006a) e Lima (2005) afirmam que a maior parcela das perdas comerciais ocorre devido ao furto de EE. Segundo Lima (2005), em apenas um ano as perdas comerciais de energia chegam ao patamar de 4,5% do mercado de EE (282,26 TWh/ano em 2004). A ANEEL (2006a) estima que as perdas comerciais de EE foram de aproximadamente R$ 5 bilhões para o ano de 2004, equivalendo a aproximadamente US$ 1,7 bilhões (câmbio R$/US$ 2,92).

O furto de energia não é um problema encontrado apenas no Brasil. Nesbit (2000) estima que as empresas americanas perderam aproximadamente entre US$ 1 bilhão e US$ 10 bilhões com o furto de energia em 1998. Já Power (1999) estima que o Paquistão perdeu aproximadamente US$ 500 milhões com o mesmo problema para o ano de 1999.

Um raciocínio linear mostra que se a quantidade de energia perdida for reduzida, será necessário comprar menos energia, o que conseqüentemente reduzirá o valor da tarifa de energia elétrica (ANEEL, 2006b). A redução das perdas técnicas é uma questão de engenharia e de relação custo-benefício. No entanto, a redução das perdas comerciais envolve fatores socioeconômicos e organizacionais.

Os trabalhos que têm como objeto de estudo as perdas comerciais, normalmente, propõem modelos para aumentar a fiscalização, ou melhorá-la, com o uso de novas tecnologias para direcionar os agentes fiscalizadores como Calili (2005) que propõe a adoção de técnicas de redes neurais para direcionar os agentes a quais domicílio fiscalizar. As técnicas coercitivas são importantes, porém, trazem dois problemas: o problema da reincidência do delito (REIS, 2005), e o problema da inadimplência, pois muitos clientes detectados não têm condições financeiras de pagar a energia que consomem.

Quase todos os trabalhos nessa área costumam fazer uma análise qualitativa sobre fatores socioeconômicos que estão associados ao furto de energia, como Vieiralves (2005) ao analisar as causas do furto de energia elétrica na cidade de Manaus, ou Reis (2005) ao analisar as mesmas causas no Estado do Rio de Janeiro. No entanto, poucos trabalhos fazem uma análise quantitativa sobre o impacto de tais fatores sobre as perdas. Uma exceção é o trabalho de Gümüsdere (2004) que apresenta uma análise do impacto de variáveis socioeconômicas sobre o percentual de energia perdida na Turquia.

Uma análise deste tipo de perdas para o Brasil não é simples. A estrutura de distribuição é complexa, contando com 64 concessionárias distribuidoras, sendo 24 privadas, 21 privatizadas, 4 municipais, 8 estaduais e 7 federais (ANEEL, 2005). A ANEEL repassa a responsabilidade no combate às perdas comerciais para as concessionárias, com uma estrutura de incentivos que estabelece uma meta de redução do percentual de perdas que a concessionária deve atingir, sob pena de arcar com o excedente de perdas. Compete então às concessionárias investir em tecnologia e campanhas publicitárias, como a conscientização dos clientes quanto ao furto de energia. Entretanto, existem problemas socioeconômicos associados que não são gerenciáveis pela concessionária, como a violência. Portanto, o ideal é definir quais fatores socioeconômicos estão associados ao problema das perdas comerciais, pois dessa forma pode-se definir qual a melhor política para redução de perdas e quais agentes serão responsáveis por aplicá-las. À medida que o nível de perdas for reduzido, os ganhos podem ser repassados para as tarifas, trazendo-as a valores mais condizentes com a realidade financeira da população brasileira, criando, assim, um círculo virtuoso.

Este trabalho efetua uma análise tanto entre países como para empresas do mesmo país. Ele está dividido da seguinte maneira: a seção 2 discute as causas das perdas de energia elétrica, a 3 apresenta as formas de combates ao furto de EE no contexto mundial e nacional, a 4 discute a correlação existente entre as perdas comerciais de EE e as variáveis socioeconômicas, a 5 apresenta a metodologia que foi aplicada, na seção 6 encontra-se o desenvolvimento, a aplicação e os resultados do modelo para a análise entre países, enquanto que na seção 7 é feita a mesma discussão para o modelo de análise entre as concessionárias e a seção 8 apresenta os principais resultados e as conclusões.

2. CAUSAS DAS PERDAS DE ENERGIA ELÉTRICA

As perdas de EE, conforme visto anteriormente, são classificadas em técnicas e comerciais. As perdas técnicas ocorrem devido aos efeitos físicos do sistema: estão associadas ao próprio sistema, à sua qualidade e à forma como ele está sendo usado. Dentre os efeitos físicos pode-se destacar os quatro seguintes: efeito Joule, perda por histerese, correntes de Foucault e efeito Corona (para uma explicação minuciosa ver Luna (2006)).

As perdas comerciais podem ser definidas como a parcela da energia que foi consumida por determinadas unidades consumidoras, mas não foi faturada (VIEIRALVES, 2005). Ou seja, quando é subtraído do total de energia injetada no sistema o total de energia faturada têm-se as perdas totais. Assim, calculando-se as perdas técnicas e subtraindo-as das perdas totais, resta uma parcela de perdas que não são explicadas por características técnicas do sistema.

O motivo do surgimento dessas perdas é diverso, contudo só há dois grupos de agentes capazes de ocasioná-las: as concessionárias e os consumidores. Utilizando como referência a classificação feita por Smith (2004) e por Ibrahim (1999), pode-se extrair uma classificação para as formas em que os agentes ocasionam as perdas comerciais.

2.1 Perdas por ação do consumidor:

a. **Fraude**: É quando o consumidor tenta enganar a empresa. Uma maneira de fazer isso é adulterar o medidor com o intuito do mesmo apresentar uma medição abaixo do patamar consumido.

b. **Desvio (Furto)**: Esse tipo de ação ocorre quando o consumidor desvia energia propriamente dita. Uma das maneiras é utilizar uma ligação que leve a energia para casa sem passar pelo medidor. Há diversas técnicas de desvio de energia elétrica, como por exemplo, algumas pessoas possuem equipamentos que têm um consumo alto e desviam a energia no ramal de ligação, antes do medidor, para alimentar esses equipamentos, enquanto que os demais permanecem alimentados por meio do medidor. Há outros consumidores que em determinados dias alimentam a casa por meio de uma ligação que não faz uso do medidor e nos demais dias, alimentam a residência por intermédio do medidor, reduzindo assim o consumo mensal.

c. **Ligação de clandestinos**: trata-se de ligação que não é reconhecida pela concessionária. Um indivíduo pode não ter medição em sua casa por diversos motivos, entre eles, pode ocorrer o caso do domicílio não ter atendido às normas técnicas ou até mesmo o caso em que na região onde mora não há rede elétrica próxima. Então, para não ficar sem energia, este faz uma ligação ilegal até a sua casa, o que sai mais barato do que a concessionária fazendo, pois ele não se preocupa com os padrões e com a segurança das instalações. Em outros casos, o consumidor pode fazer uma ligação clandestina para atender uma pequena demanda, como uma máquina de caldo de cana, por exemplo.

d. **Faturamento incorreto**: caracteriza-se pelo fato do consumidor induzir o leiturista a se corromper, fazendo uma leitura equivocada da medição com o intuito de reduzir o consumo faturado. É importante lembrar que o leiturista pode cometer um erro de leitura sem a intenção. Nesse caso, é um faturamento incorreto também, porém, não ocorre por ação do consumidor.

2.2 Perdas por deficiência da concessionária

a. **Problemas de Medição**: O medidor pode não funcionar corretamente, faturando abaixo do esperado. Há também os casos dos erros de leitura, já que muitas vezes as leituras são visuais, podendo o agente responsável cometer erros na aquisição dos dados.

b. **Erros de Cadastro do Cliente**: caso a empresa não tenha os dados do cliente corretos, o leiturista pode não encontrar o domicílio e assim não faturar a conta de energia.

c. **Faturamento Incorreto**: por diversos motivos, a fatura pode sair incorreta, desde um erro de leitura até um erro no sistema. Esses casos podem levar a um subfaturamento, reduzindo assim a arrecadação da concessionária.

d. **Perdas Administrativas**: ocorre quando a empresa deixa de faturar algum cliente. Um exemplo disso é quando o responsável pela digitação dos dados de faturamento comete um erro ao passar a informação ou perde dados de alguns clientes, não podendo assim faturá-lo.

3. FORMAS DE COMBATE

O problema das perdas comerciais é combatido em todo o mundo, uma "guerra" que envolve países independentemente de sua riqueza ou cultura. As empresas de eletricidade da Inglaterra têm utilizado várias tecnologias para reduzir o furto de energia, como, por exemplo, a substituição dos tradicionais medidores eletromecânicos por eletrônicos. Também têm sido aplicados medidores pré-pagos, cuja utilização se dá por meio de um cartão de crédito ou de uma chave eletrônica de transferência de crédito. Porém, essa técnica apresentou várias formas de tentativa de fraude, como o uso de cartões falsificados, ou até mesmo a colocação de materiais elétricos no lugar da chave eletrônica (DICK & MACEY, 1999).

Tewari e Shah (2003) analisaram a política de implantação de medidores pré-pagos na África do Sul. Segundo os autores, o sistema traz benefícios para a concessionária, pois reduz o custo de leitura, de postagem das faturas e de corte de energia[1] dos clientes inadimplentes. Para o cliente também é interessante, pois ele tem um maior controle sobre o seu consumo de energia, como também o valor da tarifa é reduzido e não há o transtorno do corte e religação, caso a conta de energia não seja paga. Porém, segundo administradores da empresa, o custo dos medidores pré-pagos ainda é alto e não reduz de forma significativa o furto de energia.

Ghajar e Khalife (2003) propõem um processo de leitura automática de medição para os grandes consumidores e os considerados problemáticos (aqueles que já foram detectados em casos de furto). Com a análise de projetos de investimento, percebeu-se que para o caso do Líbano, que tem um índice de perdas de 50%, é viável a utilização de processo de leitura automático.

Apesar do surgimento de muitas tecnologias antifurto, muitas empresas usam a fiscalização para reprimir o furto de energia e, associado a ela, alguma técnica para direcionar os agentes de fiscalização. Na Tailândia, a *Provincial Energy Authority of Thailand* – PEA é a empresa responsável pela transmissão e distribuição de energia para todo o país, com exceção de Bangladesh. Ela estabeleceu um protocolo de fiscalização de acordo com a categoria do cliente: para clientes acima de 69 kV, a propriedade é fiscalizada ao menos uma vez por ano. Para outros clientes, é feita a fiscalização por varredura[2] ou baseando-se em dados históricos do cliente. Com esses tipos de fiscalização, a empresa conseguiu detectar entre outubro de 2000 e junho de 2001, para clientes abaixo de 12kV e acima ou igual à 12kV, 1.167 e 127 casos de infração, respectivamente. Dado que a empresa tinha 11.399.150 clientes, foram detectados cerca de 0,011% dos clientes como infratores, em aproximadamente oito meses (SURYANMONGKOL, 2002).

No Paquistão, em 1998, o governo empregou 35.000 homens do exército para recuperar a energia e a água perdidas e, assim, frear o furto. Eles foram de casa em casa, sob a orientação da empresa responsável, checando adulteração no medidor, e só em 2003, o exército tinha descoberto 100.993 ocorrências de furto de energia e prendido ou penalizado 1.188 pessoas (SMITH, 2004)

Na China, oficiais detectaram aproximadamente 10.000 casos de furto de energia durante uma fiscalização feita no ano de 1998 (WORLD, 2002). A polícia britânica também está investigando os casos de furto de energia não só devido ao furto em si, mas também por acreditar haver uma conexão entre produtores de droga e furto de energia, pois para produzir a droga necessita-se de uma grande quantidade de energia (ENERGY, 2005).

O governo brasileiro também entrou na briga para combater o furto de energia elétrica por meio de fiscalização. Segundo Brasil, o Rio de Janeiro iria treinar 200 policiais especialmente para fiscalizar o furto de energia, pois, segundo o secretário de Segurança Pública da época,

[1] Entende-se por corte de energia o procedimento que a concessionária faz para suspender o fornecimento de energia elétrica.
[2] É quando são investigados todos os estabelecimentos que atendem a um determinado critério, como segmento econômico. É importante observar que nesse tipo de fiscalização não há inicialmente nenhuma indicação de furto para o estabelecimento.

"quem furta energia não está roubando apenas da concessionária, e sim, da sociedade" (BRASIL, 2005).

As empresas brasileiras também têm investido no combate ao furto, como é o caso da Escelsa, que investiu fortemente nos últimos anos para reduzir as perdas comerciais, dividindo o ataque em duas partes: (i) melhoria na tecnologia de medição e (ii) fiscalização. Para se ter uma idéia, no ano de 2005 foram feitas 258.573 inspeções do tipo varredura, ou se baseando em dados do cliente ou mediante denúncia. Com isso, conseguiu-se retirar 142.801 ligações clandestinas, reduzindo o índice de perdas comerciais de 5,6% para 5,1% nos meses subseqüentes às ações (ESCELSA, 2005).

Como o custo da fiscalização é alto, uma vez que pessoas utilizadas no processo devem ser capacitadas tecnicamente e ainda devem receber um salário relativamente alto, para evitar a corrupção, é provável que a empresa utilize um processo de inteligência para tornar a fiscalização mais eficiente.

As técnicas para definir o direcionamento das fiscalizações são diversas. Entre elas, pode-se citar: a análise na variação do consumo, clientes com consumo zero, denúncia, varredura, etc. Porém, com a evolução da informática, surgiram métodos computacionais que podem melhorar a eficiência da fiscalização. Calili (2005) partiu da idéia de que os modelos tradicionais de inteligência de fiscalização não são eficientes, pois não incorporam o fator sazonalidade na variação de consumo (o que eles fazem é simplesmente fiscalizar clientes que reduziram o consumo acima de um percentual). Com base nessa hipótese, ele aplicou técnicas de redes neurais a um banco de dados dos clientes da ELEKTRO, na cidade do Guarujá. Com o uso de técnicas de redes neurais, separou os clientes em grupos e por meio de Lógica Fuzzy, tentou detectar quais grupos poderiam ser classificados em adimplentes, inadimplentes ou fraudadores. Com essa técnica, ele conseguiu detectar 68,2% dos fraudulentos nos grupos de maior incidência de fraude.

Empresas brasileiras têm utilizado também tecnologias antifurto para combater as perdas comerciais. Nesse sentido, Reis (2005) faz um estudo de caso sobre a aplicação de um processo de leitura automática, chamado concentrador, em uma determinada região da empresa Ampla. Esse equipamento tem a capacidade de fazer leituras do cliente em tempo real e fazer o corte e religação de energia elétrica através de simples comando do operador. Além da instalação do equipamento, a empresa tomou outras ações, como legalizar clientes que estavam clandestinos e mudar a estrutura da rede (elevando as redes de tensões mais baixas para dificultar o acesso dos clientes). Ele observou uma redução inicial nas perdas comerciais, porém com o passar do tempo as perdas voltaram a crescer. O autor acredita que tal fato ocorreu porque os clientes se acostumaram com a nova tecnologia e descobriram novas formas de furtar energia. Contudo, o autor acredita que a empresa ganhou em dois aspectos: o primeiro é que a empresa conseguiu identificar um conjunto de clientes fraudadores, e o outro é que reduziu o custo do uso de leituristas.

No que se refere às perdas técnicas, Méffe (2001) desenvolveu uma metodologia para o cálculo por segmento. Tal metodologia visa ter uma melhor informação a respeito das perdas técnicas e, assim, poder definir as melhores políticas para a redução das perdas. Ele define uma faixa percentual de perdas por segmento (Tabela 1) onde se observa que o transformador de distribuição é um dos segmentos que acarreta mais perdas técnicas ao sistema.

Tabela 1 – Nível médio de perdas de energia em relação ao total de perdas técnicas

Segmento	Faixa Esperda (%)
Sistema de alta tensão	16-25
Subestação de distribuição	9-17
Rede primária	15-30
Transformadores de distribuição	15-55
Rede secundária	3-30
Ramais de ligação	0,5-6,0
Medidores de energia	3-7
Outros	5-11

Fonte: Meffé (2001)

De fato, três fatores determinam a quantidade de perdas no sistema elétrico: a topologia, a operação do sistema e sua qualidade. Quando a empresa realiza o projeto de engenharia para a construção ou ampliação do sistema elétrico, entende-se que ela deve considerar a projeção de crescimento do mercado faturado de energia, tendo como meta principal o atendimento ao consumo presente e futuro. Entende-se que as perdas comerciais não fazem parte do projeto original. Logo, essas perdas podem levar os equipamentos a trabalhar sobrecarregados, reduzindo a eficiência deles e aumentando a quantidade de perdas técnicas (ANEEL, 2006b).

Como uma parcela das perdas técnicas ocorre por razão das perdas comerciais, é possível dividir as perdas técnicas em três categorias:

- Perdas Técnicas por Consumo Regular: são as ocorridas devido ao consumo faturado e, conseqüentemente, previsto;
- Perdas Técnicas por Consumo Irregular: são as ocorridas devido ao sobrecarregamento causado pelas perdas comerciais (não fazem parte do projeto);
- Perdas Técnicas Fio (totais): é o resultado da soma das perdas técnicas por consumo regular mais as perdas técnicas por consumo irregular.

Neste trabalho, ao se falar em perdas técnicas estar-se-á fazendo referência às perdas técnicas fio.

4. CORRELAÇÃO ENTRE VARIÁVEIS SOCIOECONÔMICAS E PERDAS DE ENERGIA ELÉTRICA

Diversas variáveis socioeconômicas estão associadas às perdas de energia elétrica, entre as quais pode-se considerar como as mais importantes a violência, a renda, aspectos culturais e a organização da empresa.

4.1 Violência e as perdas comerciais

A violência é um dos fatores sempre comentados na literatura como associado ao problema das perdas comerciais: constata-se que cidades com maiores índices de perdas também apresentam altos índices de violência. Em cidades como o Rio de Janeiro, os agentes muitas vezes são impedidos pelo crime organizado de fiscalizar determinadas regiões (CALILI, 2005). Com o intuito de corrigir essa deficiência, muitas empresas brasileiras de energia têm feito convênios com a respectiva Secretaria de Segurança Pública para poder fiscalizar determinadas regiões por meio da autoridade policial. Entretanto, o alto índice de violência leva as autoridades policiais a relevarem o furto de energia diante de outros crimes "mais importantes", como homicídios, assaltos à mão armada, seqüestros relâmpagos, etc (COELHO, 2004).

A ANEEL reconhece que existem áreas onde há impossibilidade de repressão ao furto de energia elétrica e recomenda nessas áreas a adoção de políticas socioeconômicas para reprimir o furto de energia (ANEEL, 2006b).

4.2 Cultura e as perdas comerciais

Algumas concessionárias têm feito campanhas educativas tentando conscientizar os clientes a respeito dos malefícios do furto e criar um canal de comunicação maior com os clientes. Bassaikhan (2003) afirma que é necessária uma constante comunicação com o público, com o intuito de influenciar o comportamento do consumidor. Uma das formas de se comunicar com o cliente é por meio de propagandas usando meios de informações. Vieiralves (2005) entende que o baixo nível educacional da população influencia na cultura do desperdício e do furto de energia. Isso porque o nível de consumo médio mensal por habitante da região que ele analisou (cidade de Manaus) é muito alto em relação à média nacional, o que indica um desperdício muito grande de energia por parte da população. Entretanto, como uma parte da população não tem como pagar pelo que consome, alguns indivíduos preferem furtar energia a reduzir o

consumo. Esse ciclo vicioso, desperdício + consumo alto+ furto, o autor credita principalmente ao baixo nível educacional.

4.3 Renda e as perdas comerciais

Segundo Cintra (2007), a tarifa de energia elétrica no Brasil supera a de muitos paises, inclusive paises como Reino Unido, Espanha, França, Suécia e Estados Unidos, cuja maior parte da geração é térmica, diferentemente do Brasil, que é hidráulica. Ou seja, apesar de ter uma forma de energia mais barata, cobra mais caro. Uma justificativa para isso é a alta carga tributária brasileira, que chega a corresponder a 34% da tarifa de energia elétrica.

Os governos de alguns Estados brasileiros tentam atenuar o impacto da tarifa no orçamento do consumidor mediante uma política de redução da alíquota de ICMS sobre a tarifa de energia, para algumas classes de consumidores. Em alguns casos, aplicam a isenção total sobre uma determinada classe, como no caso de Pernambuco onde os consumidores de baixa renda (consumo médio mensal de até 220 kWh) são isentos de ICMS (MOREIRA, 2007).

Costa (2005) afirma que os aumentos exagerados das tarifas comerciais e residenciais contribuíram para a elevação do nível de furto de energia e inadimplência, aumentando assim o nível das perdas comerciais, principalmente nas empresas menos eficientes.

Com certeza existe uma correlação entre perdas e tarifa, mas essa correlação pode não ser devido à influência da tarifa nas perdas, e sim o contrário, pois na tarifa de energia elétrica normalmente estão incluídos os custos total ou parcial com as perdas, inclusive no Brasil.

Fankhauser e Tepic (2007) analisaram vários paises em transição e perceberam que, na média, o custo da água, energia e aquecimento não ocupam mais do que 4% do orçamento familiar. Contudo, para as famílias consideradas mais pobres, apenas o peso da eletricidade no orçamento chega a ser de mais de 10% da renda familiar. Os autores acreditam que uma política melhor de tarifa poderia, inclusive, reduzir a inadimplência. A questão não é o valor da tarifa, e sim o seu impacto na renda.

4.4 Organização da empresa e as perdas comerciais

Ibrahim (1999) considera que é possível reduzir as perdas de energia no sistema elétrico por meio de um plano de redução de perdas. Ele estabelece uma classificação para as empresas de acordo com o nível de perdas, onde as empresas com os menores índices de perdas são consideradas corporações com excelente organização, e as empresas com os piores índices são chamadas de empresas com organização deficiente.

Costa (2005) analisou a concessionária Ampla que se encontra no Estado do Rio de Janeiro e percebeu que a falha nos sistemas de informação da empresa e a falta de alinhamento entre os diversos setores prejudicaram a adoção de estratégias que pudessem combater o furto de energia. Além disso, o autor detectou uma série de dados inconsistentes que impossibilitaram a tomada de decisão. Com isso, ele concluiu que a maturidade do modelo de gestão contribui para a redução das perdas, desde que sejam considerados os fatores socioeconômicos em que a empresa está inserida.

A ANEEL (2006b) acredita que algumas empresas distribuidoras, caso tivessem uma melhor gestão, poderiam reduzir as perdas comerciais. Espera-se, portanto, uma correlação negativa entre o nível de organização da empresa e as perdas comerciais de energia.

5. METODOLOGIA

Com o objetivo de analisar o impacto das variáveis socioeconômicas sobre as perdas comerciais de energia elétrica, foram desenvolvidos dois modelos: o primeiro procura fazer uma análise macro, ou seja, uma análise entre países com o intuito de detectar a influência da eficiência governamental nas perdas comerciais de energia elétrica; o segundo faz a análise entre as empresas de um mesmo país, no caso, o Brasil.

O modelo utilizado com o intuito de analisar a influência das variáveis socioeconômicas nas perdas comerciais pode ser visto na equação 1, Observa-se que as perdas comerciais dependem da renda, do nível de eficiência governamental do país, da violência e do nível cultural.

$$\text{Perdas Comerciais} = f(\text{Eficiência, Violência, Renda, Cultura}) \quad (1)$$

Smith (2004) afirma que um aumento no nível de eficiência governamental reduz as perdas comerciais, Coelho (2004) considera que um aumento na violência leva a um aumento nas perdas comerciais e Costa (2005) evidencia que os maiores índices de perdas se encontram no segmento de clientes de baixa renda e favelados.

As perdas técnicas podem ser mensuradas pela equação 2. Caso deseje-se medir as perdas totais, basta somar as perdas comerciais com as perdas técnicas, conforme pode ser observado na equação 3.

$$\text{Perdas Técnicas} = f(\text{Topologia, Qualidade, Operação do sistema}) \quad (2)$$

$$\text{Perdas Totais} = \text{Perdas Técnicas} + \text{Perdas Comerciais} \quad (3)$$

Algumas variáveis citadas nos modelos não são fáceis de mensurar. Por isso, foi necessário utilizar *proxies* para essas variáveis.

6. O MODELO ENTRE OS PAÍSES

Como variável dependente, optou-se por utilizar o índice de perdas totais, que é o percentual de energia gerada que é perdida desde a fonte geradora até chegar ao consumidor, incluindo o furto de energia. Sabe-se que o objetivo é mensurar as perdas comerciais. No entanto, é difícil a separação entre perdas técnicas e comerciais (STRAUCH, 2002). Devido às dificuldades de cálculos e dificuldade de acesso ao valor das perdas comerciais por países, optou-se pelo Índice de Perdas total – IP -, como variável dependente.

As perdas totais são explicadas tanto por variáveis que influem nas perdas técnicas como variáveis que influem nas perdas não-técnicas. Logo, para expressar as perdas totais optou-se por utilizar variáveis associadas aos dois tipos de perdas. Nesse item, serão descritas as variáveis independentes utilizadas.

A variável Densidade de Consumo – DC é uma *proxy* para a topologia da rede. Ela é calculada pela divisão entre o consumo total no ano pela área de concessão no mesmo ano, na unidade de kWh por km^2. Gümüsdere (2004) considerou que a quantidade de linha de baixa tensão em quilômetros dividida pela quantidade de consumidores é uma boa forma de medir a topologia da rede. No entanto, tal variável não é fácil de obter devido à dificuldade de acesso aos dados das empresas. Por causa disso, optou-se por utilizar densidade de consumo, pois quanto mais denso o consumo menores serão as perdas percentuais para levar a energia até o consumidor. Logo, espera-se uma correlação negativa entre o IP e a DC.

A Variável Grau de Ruralidade – GR foi utilizada como *proxy* para mensurar as perdas técnicas. Ela é obtida da relação entre o número de habitantes em áreas rurais e o número de habitantes totais do país. O uso do GR decorre do fato de que quanto maior a proporção de consumidores rurais maiores serão as perdas técnicas com a transmissão e com o uso de sistemas rurais de distribuição elétrica, como o sistema Monofilar de Retorno por Terra - MRT[3], que acarretam perdas maiores ao sistema que os sistemas convencionais. Espera-se, portanto, uma correlação positiva entre o IP e o GR.

A variável Eficiência Governamental – EF foi utilizada com o intuito de medir a qualidade do serviço público e a sua capacidade em formular políticas e aplicá-las. Essa variável é um indicador calculado por Kauffman, Kraay e Mastruzzi (2006) de um conjunto de indicadores denominado *Governance Matters*. De acordo com o autor, ela mede a qualidade dos serviços

[3] Sistema Monofilar de Retorno por Terra: sistema de transmissão com um fio, cujo retorno é a própria terra. Esse tipo de transmissão reduz o custo de construção da rede elétrica devido ao uso de menos condutor e estruturas, porém aumenta as perdas de energia por transmissão.

públicos, das instituições e das autoridades, o seu grau de independência de pressões políticas e a capacidade de formulação e aplicação de políticas.

Gümüsdere (2004) acredita que a capacidade do governo de arrecadação e o seu poder de aplicar políticas influenciam as perdas comerciais. Visto que o governo perde com arrecadação toda vez que a energia é furtada, é de esperar que ele queira adotar estratégias para não perder essa arrecadação. Espera-se, portanto, uma correlação negativa entre perdas comerciais e Eficiência Governamental.

Optou-se pelo o uso da Taxa de Homicídio por grupo de 100.000 habitantes – TH como uma *proxy* para a violência, devido à dificuldade de se medir por outros métodos. Também em função das diferenças entre subnotificação[4] e confiabilidade nos dados, e a falta de homogeneidade entre unidades da Federação (CERQUEIRA & LOBÃO, 2004). Espera-se uma correlação positiva entre TH e IP, pois, segundo Coelho (2004), a violência é um dos fatores que influenciam as perdas comerciais de energia, devido à dificuldade de fiscalização em áreas violentas, que acarreta um sentimento de impunidade.

Alguns autores discordam do uso da taxa de homicídio como uma *proxy* para violência, como Shrader (2001), que afirma que nem sempre há uma correlação entre o número de homicídios e a violência que se deseja mensurar, Contudo, Wilsem (2004) encontrou uma correlação evidente entre altas taxas de homicídio e furto para paises do leste europeu e da Europa ocidental.

A variável Renda per Capita – RPC será utilizada para mensurar se a opção de furtar energia decorre de problemas financeiros. Como Fankhauser e Tepic (2007) detectaram que muitas vezes o custo da energia chega a valer 10% do orçamento familiar, é possível que o comprometimento com outros bens necessários possa levar o individuo a furtar energia. Espera-se uma correlação negativa entre a RPC e o IP, pois quanto maior a RPC menor será o interesse do individuo em furtar energia para reduzir o impacto da tarifa no orçamento.

6.1 Os dados

A proveniência dos dados é a seguinte:

- O IP, a DC, o GR e a RPC foram obtidos no *World Development Indicators* para o ano de 2002 (WORLD BANK, 2004);
- A TH foi obtida na Organização Mundial de Saúde para o ano de 2002 por meio do documento (WHO, 2004);
- A variável EF foi obtida no *World Bank* por meio do documento (KAUFFMAN, KRAAY & MASTRUZZI, 2006).

A amostra dos dados é composta por 90 países, abrangendo todos os continentes (Tabela 2). O continente com mais países na amostra é o europeu, com 34 países, seguido pela Ásia, com 21; e os continentes com menos países são a Oceania e a América do Norte[5], com apenas 2 países para o ano[6] de 2002.

A América Central é o continente que apresenta o maior IP, com 22,3. Em seguida está a América do Sul, com 15,09. Por sua vez, a América do Norte apresenta o menor IP, de 7,69. É importante observar que apenas dois continentes possuem IP abaixo de 10. Um fato preponderante é a superioridade da densidade de consumo da América do Norte em relação aos demais continentes.

[4] Notificação abaixo do valor real dos dados sobre violência

[5] Apesar de América do Norte, América do Sul e América Central serem subcontinentes do continente americano, o trabalho os tratará como continente.

[6] O ano de 2002 foi um ano de racionamento de energia elétrica no Brasil, fato este ocorrido devido à matriz energética diferenciada do Brasil e às políticas adotadas para o setor elétrico brasileiro. Portanto, as causas do racionamento foram intrínsecas ao País, e como na amostra há 90 países, não haverá problemas em utilizar tal ano.

Tabela 2 – Valores médios das variáveis por continente

Continente	QUANT	IP (%)	DC (kWh/km²)	EF	GR (%)	RPC (US$)	TH*
América Central	8	22,30	1260,75	-0,32	49,95	3101,56	13,39
América do Norte	2	7,69	15040,61	1,86	20,27	29282,13	3,43
América do Sul	10	15,09	1434,53	-0,33	24,63	3243,99	17,36
África	11	15,01	1080,75	-0,17	49,74	1731,25	10,62
Ásia	21	12,52	3701,02	0,15	41,21	7279,13	4,18
Europa	34	11,26	6862,44	0,95	30,24	14811,71	3,69
Oceania	2	9,95	9822,53	1,94	13,25	18174,62	1,34

Fonte: Elaborado pelo autor por meio do Eviews 5.1
* a variável TH possui a unidade de número de homicídios por grupo de 100.000 habitantes
**A coluna "quant" refere-se ao número de países analisadas por continentes;
***Os dados foram calculados mediante os dados da amostra

A América do Norte e a Europa são os continentes com maior eficiência governamental, com 1,86 e 0,95, respectivamente, enquanto o com menor eficiência é a América do Sul, com -0,33. A diferença entre América do Norte e Europa se deve ao fato de que na América do Norte foram considerados na amostra apenas dois países: Estados Unidos e Canadá. O que não ocorre na Europa, que tem vários paises e muito deles com baixa eficiência, como a Albânia. Isso reduz o nível de eficiência do continente como um todo.

A América do Sul é o continente com o maior TH, de 17,36, superando a África. Por sua vez, a Oceania é o com menor, com 1,34. É importante informar que a variável TH não considera homicídios devido à guerra, e é provavelmente por isso que a África tem um TH bem abaixo da América do Sul.

A análise descritiva dos dados pode ser observada na Tabela 3, onde se percebe um alto desvio padrão para a variável IP, o que evidencia a grande heterogeneidade das perdas entre os países. No entanto, a média está próxima à mediana, isto denota que os valores estão bem distribuídos. O país que teve o menor IP foi o Paraguai, com 3,24, e o que obteve o maior IP foi o Haiti, com 51,01. A França ficou com um IP de 5,58 e a Alemanha de 4,51, valores bem abaixo do valor médio do IP na Europa, 11,26.

A variável DC tem um desvio padrão muito alto (5,13 MWh por km^2) e uma distância entre média e mediana também alta, isso deixa claro que alguns países têm uma densidade de consumo muito alto em relação aos demais. O país com a maior DC foi a Islândia, com 27,8 Mwh per capita, e o país com a menor DC foi a Etiópia, com 0,027 Mwh per capita.

Tabela 3 – Análise Descritiva das Variáveis

Variável	Média	Mediana	Máximo	Mínimo	Desvio Padrão
IP	13,38	10,83	51,01	3,24	8,77
DC	4488,23	2836,18	27850,70	27,37	5137,33
EF	0,40	0,41	2,31	-1,64	1,05
GR	36,10	34,64	88,64	0,00	19,25
RPC	9472,23	3918,76	48558,00	125,95	11280,30
TH	6,92	3,30	43,18	0,23	9,01

Fonte: Elaborado pelos autores.

A variável GR tem uma mediana próxima da média, o que indica que não há valores tão grandes ou pequenos que enviese a amostra. No entanto, o país com o maior GR é Trinidad e Tobago, com 88,64, e o com menor GR é Singapura, com praticamente 0, seguido pela Bélgica, com 2,63.

A variável RPC tem um desvio padrão alto e a média está distante da mediana, o que demonstra que, além de haver uma grande diferença entre os países, há países muito ricos que enviesam a amostra à direita. O país com a menor RPC é a Etiópia.

A variável TH tem um desvio padrão de 9,01 homicídios por grupo de 100.000 habitantes. O país que teve o maior TH em 2002 foi a África do Sul, com 43,18, seguido da de El Salvador, com 38,38. Por sua vez, os países desenvolvidos tiveram um baixo TH, como França, Japão e Alemanha, que apresentaram um TH de aproximadamente 0,6.

O nível de eficiência governamental foi desenvolvido para ter uma média próxima de zero. Os países que têm os piores níveis de eficiência têm um EF negativo, como Haiti, com -1,52. O maior EF foi obtido por Singapura, com 2,3, seguido da Suíça, com 2,25.

Todos os valores altos de desvio padrão entre as variáveis independentes servem para reforçar a idéia de que a grande heterogeneidade nas perdas ocorre, entre outros fatores, devido ao fato de existirem diferenças marcantes entre os países.

6.2 Estimação

A função utilizada segue a forma expressa na equação 4. Percebe-se que há cinco variáveis independentes, sendo que três influenciam o modelo negativamente e as demais, positivamente. Além disso, pode-se dizer que DC e GR estão associados às características técnicas do sistema elétrico do país, enquanto EF está associado às características organizacionais do país, e as demais, RPC e TH, estão relacionadas com as características socioeconômicas do país. Como não foi encontrado uma *proxy* satisfatória para o nível cultural, tal variável foi omitida do modelo.

$$IP = f(\underset{(-)}{DC}, \underset{(+)}{GR}, \underset{(-)}{EF}, \underset{(-)}{RPC}, \underset{(+)}{TH}) \quad (4)$$

Os dados correspondem ao ano de 2002. Percebe-se na equação 5 que todas as variáveis independentes são lineares, com exceção do DC, que foi utilizado o logaritmo de DC. O modelo descrito na equação 5 foi o que obteve o melhor R^2 ajustado, o melhor critério de Akaike e de Schwarz, dentre os modelos testados, que foram os modelos log-linear, o modelo linear e o modelo semi-logarítmo.

$$. \quad IP_i = \beta_1 + \beta_2 \log(DC_i) + \beta_3 GR_i + \beta_4 RPC_i + \beta_5 TH_i + \beta_6 EF_i + u_i \quad (5)$$

6.3 ANÁLISE DOS RESULTADOS

Ao utilizar a técnica dos Mínimos Quadrados Ordinários – MQO, o valor da estatística de Durbin-Watson foi de 1,76. Isso indica que não se pode rejeitar as hipóteses nulas de ausência de autocorrelação positiva e negativa. No entanto, ao aplicar o teste de White para testar a hipótese nula de homocedastidade, foi rejeitada tal hipótese, o que tornou necessário a aplicação de uma medida corretiva. Foi aplicada a correção de erros padrão e variância em heterocedasticidade segundo White. O resultado obtido da regressão é observado na Tabela 4. É importante destacar que os resultados obtidos após a correção não alteraram de forma significativa o resultado original, de tal modo que a heterocedasticidade não se mostrou um problema sério no conjunto de dados.

A variável EF foi significativa para um nível de significância de 1%, e o coeficiente, de -4,26, está conforme o esperado, ou seja, quanto mais eficientes os países, menores serão as perdas comerciais. Isso mostra que a capacidade de um país em definir estratégias e fiscalizar influencia no IP, o que está de acordo com (SMITH, 2004).

Tabela 4 – Resultado da Estimação por Países

Variável	Coeficiente	Erro Padrão	Valor - t	Valor – p
EF	-4,2609	1,4337	-2,9718	0,0039
GR	-0,1109	0,0422	-2,6277	0,0102
LOG(DC)	-3,1564	1,2543	-2,5164	0,0138
RPC	0,0001	0,0001	0,9407	0,3496
TH	-0,0107	0,0705	-0,1510	0,8803
C	42,5652	9,6351	4,4177	0,0000

R^2 ajustado = 0,47

Fonte: Elaborado pelo autor por meio do Eviews 5.1

O GR foi significativo para um nível de significância de 5%. O coeficiente, de -0,11, foi diferente do esperado, de acordo com a literatura, porém está de acordo com o resultado obtido pela Fundação Getúlio Vargas em um relatório apresentado pela mesma que analisou duas regiões da empresa AMPLA, uma rural e outra urbana e constatou que as pessoas interioranas tinham menos propensão a furtar energia, ao qual foi justificado como um indicativo de que os valores culturais dos habitantes de cidades do interior prevalecem sobre o desejo de furtar a energia inibindo a prática do delito (FGV/UFF, 2003). Ou seja, uma boa parcela de pessoas do interior deixa de furtar energia por questões morais. Todavia, o fato de o coeficiente do GR ter sido negativo não quer dizer que ele não seja uma *proxy* para a topologia da rede, e sim que essa variável está associada a outros valores que influenciam mais no IP do que a topologia.

A variável DC foi significativa para um nível de significância de 1%. e o coeficiente angular de -3,15 apresentaÿÿinal esperadÿÿ mÿÿtrÿÿdo que quanto mais denso o consumo menores serão as perdas para se transmitir energia ao consumidor. Esse resultado mostra que quando se comparam duas concessionárias, o simples fato de uma ter uma densidade de consumo superior à outra, pode significar que a empresa com a menor DC vai ter uma dificuldade maior para manter o nível de IP num mesmo patamar que a outra. Isso evidencia que as concessionárias devem ser tratadas de formas diferentes, de acordo com suas características.

A RPC mostrou-se não-significativo, o que pode dar a impressão de que a renda não influencia no IP, no entanto, foi utilizado o PIB per capita como *proxy* para a renda do indivíduo e a comparação do PIB entre os países não é algo fácil, visto que a estrutura de preços entre os produtos de uma mesma cesta difere de país para país.

A taxa de homicídio mostrou-se não-significativa na regressão. Isso pode ter ocorrido por três motivos: (i) a taxa de homicídio não é uma boa *proxy* para violência quando se faz uma análise por país, o que contradiz Wilsen (2004); (ii) a violência não é uma das causas das perdas comerciais, e sim outros fatores mais importantes; (iii) o efeito da violência só pode ser detectado numa análise micro, como por exemplo, numa análise por Estados.

O modelo conseguiu explicar 47% das variações do IP, os outros 53% são explicados por variáveis que não foram inclusas, mas que estão associadas ao IP, como, por exemplo, a penalidade imposta ao consumidor que varia de país para país.

Com os resultados obtidos na análise por países foi possível explicar algumas informações que já são comentadas na literatura, como a correlação existente entre eficiência governamental e índice de perdas, porém esses resultados não são o suficiente para análise que o trabalho se propõe a executar. Visto que, o efeito da renda seria analisado melhor num estudo entre as empresas de um mesmo país, além de que, fatores que se sabe variar muito de uma região para outra, como o nível de violência, foram agregados numa análise macro. Portanto, é interessante analisar o furto de energia numa análise menos agregada.

7. O MODELO PARA AS EMPRESAS

Como variável dependente, optou-se por utilizar o índice de perdas totais que pode ser calculado de duas formas: dividindo-se as perdas totais pela energia total injetada no sistema ou

dividindo-se as perdas totais pela energia faturada. A justificativa para o uso do IP é a mesma considerada no modelo para países.

Neste modelo foi utilizada como *proxy* para a eficiência organizacional da empresa as variáveis DEC e IASC, a serem explicadas adiante. A metodologia de cálculo de algumas variáveis mudou do modelo entre países para o entre empresas, o que torna interessante explicá-las novamente.

A variável Densidade de Consumo – DC – é uma *proxy* para a topologia da rede. Ela é calculada pela divisão entre o consumo total no ano pela área de concessão, no mesmo ano, na unidade de GWh por km^2.

A Variável Grau de Ruralidade – GR – foi utilizada como uma *proxy* para mensurar as perdas técnicas. Ela é obtida da relação entre o número de consumidores rurais e o número de consumidores totais da concessionária. Espera-se uma correlação positiva entre o IP e o GR pelos motivos já comentados anteriormente.

A ANEEL criou no ano de 2000 o Índice ANEEL de Satisfação do Consumidor – IASC, que tem como finalidade principal a análise da qualidade da concessionária utilizando a opinião do consumidor como parâmetro básico. O critério utilizado para a pesquisa foi a seleção aleatória de 300 consumidores por concessionária. Os clientes selecionados são visitados e é feita uma série de perguntas visando avaliar: a qualidade percebida, o valor percebido, a satisfação global, a confiança no fornecedor e a fidelidade. Feita a pesquisa, o IASC é calculado através de uma média ponderada entre os índices que se encontram no item satisfação global (ANEEL, 2006c). Esse índice pode direcionar a ANEEL em suas políticas e ações estratégicas, além de ser utilizado na composição da tarifa.

Assim sendo, foi utilizada a variável IASC como uma *proxy* para a capacidade organizacional da empresa. Isso porque tal índice influencia no valor da tarifa da empresa. Portanto, a empresa só não melhora o índice por dois motivos: (1) falta de interesse em aumentar o faturamento; (2) problemas organizacionais que a impossibilitam de aumentar o IASC. Claramente, a segunda opção mais plausível. Sob outro aspecto, uma empresa que se preocupa em ter um bom índice de satisfação também se preocupa em reduzir as perdas administrativas de energia, pois estas também afetam o faturamento. Logo, espera-se uma correlação negativa entre o IASC e o IP.

Segundo o Departamento Nacional de Água e Energia Elétrica – DNAEE a Duração Equivalente de Interrupção por Consumidor – DEC explicita o espaço de tempo que, em média, cada consumidor de um dado conjunto considerado ficou privado de fornecimento de energia elétrica em um período determinado de tempo (DNAE, 1978). O DEC pode ser calculado através da equação 6.

$$DEC = \frac{\sum_{i=1}^{n} Ca(i) \cdot t(i)}{Cs} \qquad (6)$$

onde:
DEC = duração (em horas)
i = número de interrupções variando de 1 a n
Ca(i) = número de consumidores, do conjunto considerado, atingidos nas interrupções;
t(i) = tempo de duração das interrupções (i), em horas;
Cs = número total de consumidores do conjunto considerado.

A DEC foi utilizada como uma *proxy* para a organização da empresa pois, ocorrendo a falta de energia, o tempo para restabelecimento dependerá da organização da empresa e de sua eficiência em detectar a falha. Logo, quanto maior o DEC, menos organizada a empresa. Espera-se, portanto, uma correlação positiva entre o DEC e o IP. A idéia de que o DEC está associado à organização da empresa também é compartilhada pela CEEE, pois, segundo Moura (2002), a empresa incorporou o índice DEC no seu programa de participação de resultados para os funcionários. E os gerentes da empresa acreditam que o DEC e o FEC são os dois fatores que mais influenciam no gerenciamento estratégico da empresa.

Quanto à correlação entre o IASC e o DEC, espera-se que seja pequena, pois a metodologia de cálculo é diferente: ambos são uma *proxy* para a organização da empresa, todavia, o primeiro é baseado no sentimento do cliente em relação à empresa – portanto algo difícil de mensurar. O

outro é baseado na qualidade do sistema e na capacidade da empresa em responder às falhas, que é mais fácil de mensurar.

Da mesma forma que apresentado no modelo para países, optou-se pelo uso da Taxa de Homicídio por grupo de 100.000 habitantes – TH como uma *proxy* para a violência. Espera-se uma correlação positiva entre TH e IP, conforme citado anteriormente (COÊLHO, 2004).

A variável média de anos de escolaridade para pessoas acima de 25 anos – ESC será utilizada como uma *proxy* do nível cultural da população. Espera-se que quanto mais instruída a população maior a sua consciência quanto ao furto de energia. A idéia de utilizar a escolaridade também foi adotada por FGV/UFF (2003), que encontraram uma correlação negativa entre escolaridade do chefe de família e furto de energia.

A variável Renda Per Capita – RPC será utilizada para mensurar se a opção de furtar energia decorre de problemas financeiros. Como Foi detectado que muitas vezes o custo da energia chega a valer 10% do orçamento familiar, é possível que o comprometimento com outros bens necessários possa levar o individuo a furtar energia (FRANKHAUSER & TEPIC, 2007).

É possível utilizar a tarifa média de energia para analisar o seu peso no orçamento, no entanto existem três fatores que levaram a não utilizá-la: (1) no reajuste tarifário, é considerada a energia perdida no ano anterior, ou seja, um aumento nas perdas de energia *ceteris paribus* leva a um aumento na tarifa do ano posterior, podendo assim haver uma autocorrelação entre tarifa e perdas de energia; (2) não se encontrou dados para a tarifa média por concessionária para os anos em análise (2000, 2001, 2002); e (3) conforme já analisado na revisão da literatura, a diferença de renda per capita dos Estados consegue detectar o peso da tarifa média no orçamento, visto que as tarifas são muito próximas umas das outras entre regiões. Espera-se uma correlação negativa entre a RPC e o IP, pois quanto maior a RPC menor será o interesse do individuo em furtar energia para reduzir o impacto da tarifa no orçamento.

7.1 Os dados

O IP foi obtido em Coelho (2004) para os anos 2000, 2001 e 2002. As variáveis DC e GR foram calculadas mediante dados obtidos no site da ABRADEE para os anos de 2000, 2001 e 2002. A variável DEC foi obtida no site da ABRADEE: www.abradee.org.br, para os anos de 2000, 2001 e 2002. O IASC foi obtido em ANEEL (2006c) para os anos de 2000, 2001 e 2002. Já ESC, TH e RPC foram obtidos, para os anos de 2000 à 2002, no IPEA, diretamente no site www.ipea.gov.br.

A amostra dos dados é composta por 33 concessionárias, que compreendem a maioria dos Estados brasileiros, com exceção do Acre, Amapá e Piaui[7], para o período de 2000 a 2002. A amostra total é composta de 99 unidades individuais no tempo.

A análise descritiva dos dados pode ser observada na Tabela 5, onde se percebe um alto desvio padrão para a variável IP, o evidencia a grande heterogeneidade das perdas entre as concessionárias. Entretanto, a média está próxima à mediana, isso mostra que os valores estão bem distribuídos. A concessionária que obteve o menor IP foi a AESSUL 5,33, e a que obteve o maior foi a CERON 35,80.

Tabela 5 – Análise Descritiva das Variáveis

Variáveis	Média	Mediana	Máximo	Mínimo	Desvio Padrão
IP	14,925	13,300	35,800	5,330	7,115
DC	0,408	0,066	8,269	0,002	1,327
DEC	21,189	16,670	77,980	5,820	13,842
ESC	5,752	5,764	8,543	3,904	1,044
GR	5,183	5,204	14,009	0,064	3,534
IASC	63,514	63,550	75,050	52,620	5,156
RPC	5,881	5,724	14,442	1,616	2,940
TH	29,650	25,474	59,600	10,241	12,743

Fonte: Elaborado pelos autores

[7] Para esses Estados, não foi possível obter a variável IP no período desejado

A variável DC tem um desvio padrão muito alto (1,3 GWh por km^2) e uma distância entre media e mediana também alto, o que evidencia que algumas concessionárias têm uma densidade de consumo muito alta em relação às demais, como no caso da Eletropaulo, que tem uma DC de 8,269 GWh por km^2, enquanto a média é 0,408 GWh por km^2.

A variável DEC tem um desvio padrão alto. A concessionária que tem o maior DEC é a CERON (77,98), e a que tem o menor DEC é a CPFL (5,82). Essa grande diferença entre o maior e o menor valor é possível, pois o DEC anual fornece a quantidade de horas sem energia por consumidor no ano. Logo os consumidores da CERON, em média, ficaram 77,98 horas sem energia no ano. As concessionárias da Região Norte normalmente têm valores altos, como, por exemplo, a ELETROACRE, que não faz parte do conjunto de concessionárias analisado. Ela teve o DEC de 64,08, em 2000, e conseguiu reduzir para 17,63, em 2002. A CEMAR, no Nordeste, teve 67,64, em 2001, e conseguiu reduzir também.

A variável GR tem uma mediana próxima da média. O maior GR da amostra é de apenas 14,01% pertencente à RGR e o menor é de 0,0006 da Eletropaulo.

A variável ESC não demonstra um grande desvio padrão nem uma distância muito grande entre média e mediana.

A variável RPC tem um desvio padrão baixo, no entanto a diferença entre o maior RPC e o menor é muito grande, mostrando que a diferença de renda per capita entre alguns Estados ainda é muito grande.

A variável TH tem um desvio padrão alto, o que mostra a heterogeneidade dos dados entre os Estados. Além disso, a mediana está abaixo da média, indicando que alguns Estados estão com o índice de violência muito alto, o que enviesa a amostra à direita, como são os casos dos Estados de Pernambuco e do Rio de Janeiro.

Todos os valores altos de desvio padrão entre as variáveis independentes servem para reforçar a idéia de que a grande heterogeneidade nas perdas ocorre, entre outros fatores, devido ao fato de existir diferenças socioeconômicas marcantes entre os Estados.

7.2 A Regressão

A função utilizada segue a forma expressa na equação 7. Percebe-se que há sete variáveis independentes, sendo que três influenciam o índice de perdas negativamente e as demais, positivamente. Além disso, pode-se dizer que DC e GR estão associados às características técnicas da empresa, enquanto DEC e IASC estão associados às características organizacionais da empresa, e as demais, ESC, RPC e TH, estão relacionadas com as características socioeconômicas do meio em que as concessionárias estão inseridas.

$$IP = f(\underset{(-)}{DC}, \underset{(+)}{DEC}, \underset{(-)}{ESC}, \underset{(+)}{GR}, \underset{(-)}{IASC}, \underset{(-)}{RPC}, \underset{(+)}{TH}) \qquad (7)$$

O modelo utilizado para a análise empírica foi baseado na técnica de dados de painel balanceado com o objetivo de conseguir menor colinearidade entre as variáveis, mais grau de liberdade e maior eficiência (BALTAGI, 2001). Uma forma de utilizar Dados em Painel é colocando as informações empilhadas, desprezando o ano e fazendo simplesmente um OLS. Normalmente, o resultado fornece um baixo valor da estatística d de Durbin-Watson, sugerindo uma autocorreleção ou erro de especificação. Portanto, seria ingênuo utilizar tal regressão (GUJARATI, 2004).

Uma forma de resolver tal problema é considerar que podem ocorrer fatores ao longo do tempo que possam afetar o comportamento dos indivíduos de forma igual entre eles, tais como racionamento de energia ou mudanças na política nacional. Para captar o efeito no tempo, o modelo considera os coeficientes angulares constantes e os interceptos variando no tempo e constante entre os indivíduos. O modelo é conhecido como modelo de efeitos fixos no tempo. A equação é representada pela equação 8.

O modelo de efeitos fixos no tempo será o utilizado para a pesquisa em questão, porque será considerado que os efeitos que ocorreram nos sistemas elétricos afetaram as concessionárias de forma igual, além de este ser o mesmo modelo adotado por Gümüsdere (2004) para analisar o IP entre as províncias para a Turquia.

$$IP_{it} = \beta_{1t} + \beta_1 DC_{it} + \beta_2 DEC_{it} + \beta_3 IASC_{it} + \beta_4 ESC_{it} + \beta_5 GR_{it} + \beta_6 RPC_{it} + \beta_7 TH_{it} + u_{it} \quad (8)$$

7.3 Análise dos resultados

A técnica utilizada para a regressão foi o FGLS (*Feasibly Generalalized Least Square*) para fazer a estimação. Como essa técnica permite a utilização de estruturas diferentes para a matriz de variância, optou-se por utilizar a estrutura período SUR (*Period Heterokedasticity and Serial Correlation*), com o intuito de corrigir a correlação serial e heterocedasticidade detectada. Utilizando-se o software *Eviews* 5.1 para fazer a estimação, foi obtido o resultado exposto na Tabela 6.

A variável DC mostrou-se não-significativa para um nível de 10%. Tal fato talvez justifique que as perdas técnicas não explicam a grande heterogeneidade das perdas, ou simplesmente que a DC de consumo não é uma boa *proxy* para a topologia da rede. No entanto, a variável DC pode ter se mostrado não-significativa devido à qualidade da amostra.

A variável DEC foi significativa, isso evidencia que o nível organizacional da empresa, sua capacidade de administração e de coordenação de manutenção e inspeção influencia no IP. Nesse caso, a sua influencia ocorre, principalmente, nas perdas comerciais. O coeficiente de 0,154 está de acordo com a literatura, ou seja, quanto maior o DEC menor o nível organizacional da empresa e maior será o IP.

Tabela 6 – Resultado da Estimação por Concessionária

Variável	Coeficiente	Desvio Padrão	Valor t	Valor p
DC	-0,2815	0,5235	-0,5377	0,5921
DEC	0,1540	0,0398	3,8664	0,0002
ESC	0,3394	1,3661	0,2484	0,8044
GR	-0,5141	0,2030	-2,5330	0,0131
IASC	0,0040	0,0733	0,0548	0,9564
RPC	-0,9219	0,5165	-1,7848	0,0777
TH	0,1187	0,0526	2,2560	0,0265
C	14,1367	7,6302	1,8527	0,0672

R^2 ajustado = 0,71

Fonte: Elaborado pelos autores com o uso do software Eviews 5.1

A variável ESC foi não-significativa, indicando que a diferença de escolaridade entre regiões pode não influenciar na variação de perdas entre as concessionárias, o que contradiz o estudo da FGV/UFF (2003) e o comentário de Vieiralves (2005), em termos de nível cultural. Se for considerada a escolaridade como uma *proxy* para o nível cultural, esse resultado indica que possivelmente a diferença de nível cultural entre regiões pode não influenciar de forma considerável nas perdas comerciais. Esse resultado não era o esperado. Esperava-se que as pessoas mais instruídas tivessem menos interesse em furtar, no entanto isso não é verificado pelo modelo.

O GR foi significativo para um nível de significância de 5%. No entanto, o coeficiente de -0,5 foi contrário ao resultado esperado, porém esse mesmo resultado foi encontrado por FGV/UFF (2003) ao qual eles justificaram como um indicativo de que os valores culturais do povo do interior prevalecem impedindo o furto de energia, ou seja, uma boa parcela de pessoas do interior deixa de furtar energia por questões morais. Esse resultado também foi encontrado na análise por países[8].

O IASC foi não-significativo, o que mostra que apesar de o DEC ter sido significativo o mesmo não ocorreu com o IASC, pois talvez haja outros fatores que influenciam no IASC, além

[8] Outro fato que pode justificar a correlação entre GR e IP negativa é que a tarifa praticada para a energia destinada à irrigação é baixa em muitos Estados Brasileiros, o que pode desestimular o individuo a furtar energia, no entanto, é aplicado tarifas baixas também para os consumidores da classe Baixa Renda na cidade e, nem por isso há uma correlação negativa entre renda per capita e furto de energia.

do nível organizacional da empresa possivelmente torne não significativa a correlação entre furto de energia e o índice. Como o IASC mede a satisfação do cliente em relação da empresa, pode-se afirmar que não há uma correlação significativa entre insatisfação e furto de energia, ou seja, o individuo não furta energia por estar insatisfeito com a empresa. Este resultado é oposto ao obtido por Gümüsdere (2004).

O RPC foi significativo a um nível de significância de 10%, e o coeficiente de 0,11 está de acordo com a literatura, isso pode significar que o impacto da tarifa de energia na renda pode influenciar no furto de energia.

A variável TH foi significativa a um nível de significância de 5%. O coeficiente de 0,11 está de acordo com a literatura, logo, possivelmente, a dificuldade de fiscalização em regiões violentas pode influenciar no furto de energia e, conseqüentemente, no IP.

8. RESULTADOS E CONCLUSÃO

O presente trabalho procurou identificar quais fatores socioeconômicos estão associados às perdas comerciais de energia elétrica, para então definir a melhor política de combate às perdas comerciais de energia elétrica e quais agentes são os mais adequados a aplicá-la. Isso porque a tarifa de energia elétrica brasileira está entre as mais caras do mundo e um dos componentes responsáveis pela a alta tarifa são as perdas de energia elétrica.

A ANEEL tem se preocupado em garantir a modicidade tarifária, prova disso, é que a agência estabeleceu metas de redução de perdas de energia elétrica para as concessionárias. Muitos países, como a Turquia, Paquistão, Líbano e Inglaterra, também têm se preocupado com o furto de energia – parcela das perdas de energia –. E para reduzir o furto, eles têm utilizado diversas tecnologias como o uso de medidores pré-pagos e processos de leituras automáticas. No entanto, uma das técnicas mais utilizada pelas empresas no combate ao furto de energia, tanto no Brasil quanto no Mundo, é a fiscalização.

Deve-se saber que técnicas coercitivas, como a fiscalização, têm problemas associados a sua aplicação, como a reincidência do delito (REIS, 2005), ou a forte correlação negativa existente entre furto de energia e inadimplência, quando um aumenta o outro diminui (FGV/UFF 2003).

O uso de novas tecnologias, como estruturas de redes diferenciadas ou processos de leituras automáticas, reduzem o furto de energia em curto prazo. Porém, no longo prazo, o agente infrator se acostuma com a nova tecnologia e descobre novas formas de praticar o delito.

Além das possibilidades técnicas para se inibir o furto de energia, existe ainda a possibilidade de adoção de políticas socioeconômicas com esse intuito. Nesse sentido, foram desenvolvidos dois modelos para analisar as variáveis socioeconômicas associadas às perdas de energia elétrica.

A análise dos resultados mostrou que existe uma correlação entre as perdas de energia e a violência, caracterizada pela taxa de homicídios por grupo de 100.000 habitantes, o que condiz com Coelho (2004). Portanto, as concessionárias estão corretas em afirmar que a violência influencia no furto de energia elétrica, e esse fato pode estar associado ao sentimento de impunidade que surge na população residente em áreas violentas, devido às dificuldades de fiscalização encontradas em tais regiões.

Apesar de existirem apenas dois tipos de agentes capazes de ocasionar as perdas comerciais: clientes e concessionárias, o governo surge como um dos responsáveis pelas perdas, isso porque, na análise entre países constatou-se que a capacidade de um país em definir e aplicar políticas, bem como a qualidade de seus representantes, influencia no índice de perdas totais.

Isso é evidente quando se percebe que uma boa parte da tarifa é destinada a pagar encargos e tributos (aproximadamente 36% no caso do Brasil), logo em toda a energia que é furtada o governo perde em arrecadação, portanto ele tem o interesse em reduzir o furto de energia e só não atinge tal objetivo devido à sua ineficácia ou ao seu descaso.

As empresas também são responsáveis pelas perdas comerciais, conforme foi constatado na análise do modelo entre as empresas, portanto à medida que determinadas empresas aumentem a

sua eficiência organizacional, provavelmente, conseguirão reduzir o seu índice perdas de energia elétrica, conforme afirma (ANEEL, 2006b).

O uso de propagandas para aumentar o custo moral das pessoas com relação ao furto de energia, também pode ser uma política eficiente, porque a análise indicou que a cultura do povo rural inibe o furto, isso possivelmente pode estar associado ao custo moral de ser detectado e sofrer um constrangimento público diante dos vizinhos e amigos.

O fato de repassar o custo das perdas ao consumidor pode ser um agravante para o aumento das perdas de energia elétrica, tendo em vista que, a análise de resultados constatou que a renda per capita influência no índice de perdas. Isso pode significar que algumas pessoas praticam o ato de furtar a energia para atenuar o peso da fatura de energia no orçamento. Portanto, quando se aumenta a tarifa de energia mais pessoas podem praticar o delito devido a este motivo.

Por outro lado, o fato de o cliente honesto não ter culpa quanto ao furto de energia não o isenta de pagar pelo furto. Isso porque o cliente ao pagar pelo furto, sente-se mais estimulado a denunciar agentes fraudadores com o intuito de reduzir a sua tarifa, conforme afirma Lima (2005).

É possível, então, a redução das perdas comerciais, mediante uma política de estimulo às empresas a se tornarem mais eficientes, o que já é feito no Brasil pela ANEEL, estímulo ao uso de campanhas educativas nas escolas, na mídia e nas comunidades, bem como, subsídio a programas sociais que visem reduzir o nível de violência e melhorar a qualidade de vida da população.

Apesar dos resultados encontrados se mostrarem estatisticamente suficientes para balizar o desenvolvimento de uma política de redução de perdas, algumas considerações devem ser feitas. Primeiro, devido à escassez de dados, faltou analisar melhor a questão de como a penalidade sobre o cliente infrator afeta o furto de energia. Além disso, o fato de não ter conseguido analisar o índice de perdas comerciais isoladamente, levou a dificuldades de análise nas variáveis independentes, como no caso do Grau de Ruralidade, que inicialmente era uma *proxy* para as perdas técnicas, mas após os resultados percebeu-se estar associado também as perdas comerciais, outra variável de difícil análise foi a DEC que está associada ao nível organizacional da empresa, mas também, está associada a qualidade técnica dos equipamentos, isto porque, a DEC depende do número de falhas e do tempo que durou cada falha.

Para uma melhor análise dos impactos das variáveis socioeconômicas sobre as perdas comerciais de energia é importante fazer um estudo menos agregado, talvez uma análise por bairros, para identificar se o nível de renda per capita, o nível de violência da região, a escolaridade e outras variáveis socioeconômicas influenciam o furto de energia, mas para isso ocorrer é necessário uma parceria com as concessionárias distribuidoras de energia elétrica.

Dado que existe um custo para se reduzir o furto de energia elétrica, de modo que é possível definir um nível "aceitável" de furto de energia elétrica – situação em que o custo marginal do investimento no combate ao furto é igual à receita marginal recuperada -, sugere-se para trabalhos posteriores o estudo da aplicação de benefícios e penalidades aos indivíduos, bem como, do investimento em fiscalização, com o intuito de definir qual o nível aceitável de furto de energia elétrica de cada concessionária e assim estabelecer as metas de redução de perdas de energia elétrica, garantindo a modicidade tarifária à população.

9. REFERÊNCIAS BIBLIOGRÁFICAS

ANEEL (2006a). Por dentro da conta de luz: informações de utilidade pública. Recife.

ANEEL (2006b). Nota técnica 26/2006 de 23 de maio de 2006. *Tratamento regulatório das perdas de energia nas tarifas dos sistemas de distribuição de energia elétrica***.** Nota técnica 26/2006 de 23 de maio de 2006.

ANEEL (2006c). *Relatório geral IASC – 2005***.** Brasília.

ANEEL (2005). *Atlas Nacional de Energia Elétrica***.** Brasília.

BALTAGI, H. Badi (2001). *Econometric Analysis of Panel Data***.** Ontário: John Wiley & Sons.

BASSAIKHAN, D., et. al. (2003). Commercial losses of UB Electric distribution network shareholdin company and their reducing**.** *Proceedings of the 7th Korea-Russia International Symposium.* KORUS.

BRASIL, C. I. (2005). Rio treinará policiais para combater furto de energia elétrica. *Agência Brasil,* 17 de julho de 2005. Disponível em: <http://www.radiobras.gov.br/materia.phtml?materia=232614&editoria=NA>. Acesso em: 30/07/2007.

CALILI, Rodrigo Flora (2005). *Desenvolvimento de sistema para detecção de perdas comerciais em redes de distribuição de energia elétrica***.** Dissertação de Mestrado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro, 157p.

CERQUEIRA, Daniel e LOBÃO, Waldir (2004). Determinantes da criminalidade: arcabouços teóricos e resultados empíricos. *IPEA*.

CINTRA, Marcos (2007). Energia e telefonia: tributos e chiados. Conjuntura Econômica. V. 61, n.4. Abril.

COÊLHO, Jorge Luiz Moreira (2004). *Até que ponto as variáveis socioeconômicas, culturais e de impunidade contribuem, direta ou indiretamente, para a prática de processos irregulares (furto de energia), gerando perdas comerciais de energia elétrica*. Monografia (Especialização em Gestão de Negócios do Programa MBA do Departamento de Administração), Universidade Federal de Pernambuco. Recife.

COSTA, Artur Fernando de Souza (2005). *A maturidade do modelo de gestão para redução dos níveis de perdas comerciais: o caso de uma distribuidora de energia elétrica***.** Dissertação de Mestrado (Mestrado em sistemas de gestão), Universidade Federal Fluminense. Niterói.

DICK, A. J. e MACEY, R (1999). Revenue protection in a competitive supply environment. *Matering and Tariffs for Energy Supply*. IEE, n. 462 Maio.

DNAEE, Portaria n. 46, 1978.

ENERGY (2005) probe uncovers drug farms. *BBC News*. Londres, 10 de junho de 2005. Disponível em: < http://news.bbc.co.uk/2/hi/uk_news/4080602.stm >. Acesso em: 25/06/2007

ESCELSA. *Relatório de Administração e de Responsabilidade social 2005*. Vitória, 2005.

FANKHAUSER, Samuel and TEPIC, Sladjana (2007). Can poor consumers pay for energy and water? An affordability analysis for transition countries. *Energy Policy*, V. 35, n2. Fev.

FGV/UFF (2003). *O Componente Social das Perdas e a Inadimplência na Área de Concessão da CERJ*. Niterói.

GHAJAR, Raymond F. e KHALIFE, Joseph (2003). Cost/benefit of an AMR system to reduce electricity theft and maximize revenues for Électricité du Liban. *Applied Energy***.** v.76, n. 2003, p. 25-37. Jan.

GUJARATI, Damodar N (2004). *Econometria Básica*. São Paulo: Pearson Makron Books.

GÜMÜSDERE, Eray (2004). *Theft and losses in Turkish Electricity sector – Empirical analysis and implications for tariff design***.** Dissertação (Dissertação de mestrado em Economia) – Graduate School of Arts and Social Sciences. Sabanci University, Sabanci. 142p.

IBRAHIM, Ernad S. (1999). Management of loss reduction projects for power distribution systems. *Electric Power systems Research*. Volume 55, pages 49-56

KAUFFMAN, Daniel, KRAAY, Aart e MASTRUZZI, Massino (2006). *Governance Matters V: Agregate and Individual Governance Indicators for 1996-2005*. Word Bank. Setembro.

LIMA, Davi Antunes (2005). Textos para disrcursão, 2. *Convergência Tarifária: remédio regulatório para o livre acesso*. Brasília: ANEEL.

LUNA, Aelfo Marques (2006). *Materiais de engenharia elétrica***.** Recife: do autor.

MÉFFE, André (2001). *Metodologia para o cálculo de perdas técnicas por segmento de distribuição***.** Dissertação (Mestrado em Engenharia) – Escola Politécnica da Universidade de São Paulo. São Paulo.

MOREIRA, Aluísio (2007). *Eduardo baixa a conta de energia de 700 mil famílias.* GOVERNO DE PERNAMBUCO. 08 de maio de 2007. Disponível em: <http://www.pe.gov.br/gabinete_civil/news1_.asp?idnoticia=818&argumento=icms> acessado em: 10/07/2007.

NESBIT, Bil (2000). Thieves lurk – the sizable problem of stolen electricity**.** *Electrical World T&D*. September/October.

POWER (1999) Theft: an insidious menace. *Power Economics*. Julho.

REIS, Cláudia Zuccolotto (2005). *Eficácia de soluções tecnológicas para reduzir o furto de energia elétrica em empresas distribuidoras: estudo de caso***.** Dissertação (Mestrado em Engenharia de Produção) – Pontifícia Universidade Católica do Rio de Janeiro. Rio de Janeiro.

SHRADER, Elisabeth (2001). Methodologies to Measure the Gender Dimensions of Crime and Violence. *Poverty Reduction and Economic Management Latin America and Caribbean Region the World Bank*. World Bank. Julho.

SMITH, B. (2004). Thomas. Electricity theft: a comparative analysis. *Energy Policy*, v. 32, n.2004, p. 15-21.

STRAUCH, Mariana Torres. *Desenvolvimento de metodologia para o cálculo de perdas elétricas em redes de baixa tensão*. Dissertação (Mestrado profissional em engenharia) – Universidade de Salvador. Salvador, 2002.

SURYANMONGKOL, Dan (2002). *Non-Technical losses in Electrical power systems***.** Dissertação (Mestrado em Ciências) – Ohio University. Ohio.

TEWARI, D. D. e SHARH, Tushaar (2003). An assessment of South African prepaid electricity experiment, lessons learned, and their policy implications for developing countries. *Energy Policy***.** V. 31 n. 2003 p. 911-927.

VIEIRALVES, Eduardo de Xerez (2005). *Proposta de uma metodologia para avaliação das perdas comerciais dos sistemas elétricos. O caso Manaus***.** Dissertação (Mestrado em Planejamento de Sistemas Energéticos) - Universidade Estadual de Campinas. Campinas.

WILSEN, Johan van (2004). Criminal Victimization in Cross-National Perspective. *European Society of Criminology*. London: SAGE Publications.

WORLD (2002): Asia-Pacific the great volt robbery. *BBC News*. [S. l.], 2 de maio de 2002. Disponível em: <http://news.bbc.co.uk/2/hi/asia-pacific/206395.stm>. Acesso em: 25/06/2007.

WORLD BANK (2004). *2004 World Development Indicators*, Washington, DC.

WHO - World Health Organization (2004). World Heath Report 2004. *WHO*. Dezembro.

ASSESSMENT OF PRODUCTIVE EFFICIENCY IN IRRIGATION AREAS IN ANDALUSIA

NOELINA RODRÍGUEZ FERRERO
E-mail: nrferrer@ugr.es
MANUEL SALAS VELASCO
E-mail: msalas@ugr.es
Mª TERESA SÁNCHEZ MARTÍNEZ
E-mail: tsanchez@ugr.es
Department of Applied Economics
University of Granada

Abstract

Expansionist policies in irrigation have led to an exhaustive use of water which has depleted rivers and aquifers, with high added costs in terms of hidden subsidies for irrigation farming. The scarcity of water has become an increasing social and economic worry for those responsible for regulating water policies and has created problems among users, especially when used as a production factor. The main objective of this study is to conduct an empirical analysis of the differences that exist in terms of economic efficiency in the 156 irrigation areas in Andalusia, since in order to achieve better management of irrigation water it is indispensable to know this parameter and to be able to carry out comparative studies aimed at identifying the causes, which cause the differences in terms of efficiency. In conclusion, this article offers a study of economic rationality in the use of these scarce resources, particularly water.

Key words: Irrigation, economic efficiency, DEA, tobit model.
Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

Resumen

Las políticas expansionistas del regadío han dado lugar a un uso exhaustivo del agua, esquilmando los ríos y acuíferos, con altos costes añadidos en términos de subvenciones encubiertas a la agricultura de regadío. La escasez de agua se ha convertido en una preocupación social y económica cada vez mayor para los responsables políticos y ha fomentado la conflictividad entre sus usuarios, especialmente cuando ésta se usa como factor de la producción. La mejora de la eficiencia en la asignación del agua de riego es una cuestión que preocupa a planificadores hidrológicos y economistas pues no es posible seguir gestionando el agua de riego con las pautas actuales. El principal objetivo de este trabajo es analizar empíricamente las diferencias que existen en la eficiencia económica de las 156 áreas de riego andaluzas y poder hacer estudios comparativos encaminados a identificar las causas, que motivan las diferencias en la eficiencia. Este trabajo aporta, en definitiva, un estudio sobre la racionalidad económica en el empleo de recursos escasos, entre ellos el agua.

Palabras clave: Regadío, eficiencia económica, DEA, modelo tobit.
Área Temática: Economía de la Energía, del Ambiente, de los Transportes y de los Recursos Naturales.

1. INTRODUCTION

Andalusia currently has an irrigated surface area of around 900,000 hectares, some 22.86% of its farmed land and 10.2% of its territory (see Figure 1). This area generates 57% of final agricultural production and 60% of total employment in this sector, and therefore water is a key factor in the development of Andalusian agriculture. Andalusian irrigation systems have increased constantly as a consequence of a long-term policy of regulation and distribution of

hydric resources aimed at agriculture. The common pattern of water policies since the beginning of the twentieth century, following the regenerationist model, was the development of hydraulic regulation infrastructures in order to have sufficient water capacity in periods of drought and to convert drylands into irrigation areas.

FIGURE 1. Location of Andalusia in Europe



The idea of technical progress as the motor for the regeneration of Spain was a defining point of the collective consciousness at the end of the 19th century and in the first few decades of the 20th century, and the concept of independent subsistence came to the main and irrefutable argument for the extension of irrigation since the beginning of the past century. Nevertheless, the lack of economic resources in our country impeded the implantation of this model, which in fact began to be used with some frequency at the beginning of the 1950s. Later in the 1960s and 1970s, the economic development associated with the construction of infrastructures came to be a indisputable social myth.

These expansionist policies in irrigation have led to an exhaustive use of water which has depleted rivers and aquifers, with high added costs in terms of hidden subsidies for irrigation farming. The scarcity of water has become an increasing social and economic worry for those responsible for regulating water policies and has increased tensions among users, especially when used as a production factor. Improving efficiency in the assignation of irrigation water, due to its scarcity, is a question that preoccupies hydrological planners and economists since over the last century has become a common and free heritage and the protection of the environment has come to be an important question and a more politically correct concept (Aguilera-Klink, 2003). In this sense, The European Water Framework Directive considers the conditions necessary to conserve and recuperate the proper ecological state of water. As a consequence, the Andalusian irrigation system will suffer even more hydric tension than it has already undergone due to quantitative reasons[1] and also due to ecological reasons (MAPA, 2001).

[1] The growth in irrigation in Andalusia has taken place at a rate of at least 30% over the forecasts for 2012 made in the current Hydrological Plans without the construction of the regulation works included in these plans.

In the Andalusian irrigation system there is a wide range of agricultural production; moreover, there are major differences regarding the factors that shape this activity: water supply and origin, irrigation systems and climatic and edaphological conditions; the land ownership structure has also influenced these differences. It is possible to highlight two main types of irrigation agriculture:

- Irrigation in the interior, where there is a predominance of extensive crop farming, olive and other fruit trees, usually with a low degree of technology although there is a clear tendency towards modernization. The production of these crops is protected through the subsidies provided by an entity called *FEOGA- Guarantee*, and therefore their economic viability depends to a certain extent on the Common Agricultural Policy (CAP). A representative case of this type of agriculture is found in the olive groves of Jaén, an area in which there is currently a predominance of localized irrigation.

- Irrigation on the coast, specializing in crops with a greater economic value, using a higher degree of technology and with a predominance of localized irrigation in these areas. As the main representatives of this type of agriculture one can highlight strawberries in Huelva and greenhouse agriculture on the coasts of the provinces of Almería and Granada (Rodríguez-Díaz *et al.* 2004). Unlike what happens in the case of extensive crops, the economic viability of irrigation for intensive crops does not depend on the CAP, since the prices of fruits and vegetables are determined by market factors and by competitiveness. Given their dedication to exportation, the main threat for this type of producers stems from the increasing competition from North African Mediterranean countries and the Southern hemisphere (Rico-Amorós, 2006).

Consequently, the socioeconomic efficiency and the future perspectives for these two types of agriculture are different. Another factor which creates instability in the Andalusian irrigation system is the insufficient availability of water, in the Guadalquivir River Basin due to an excess in the irrigated surface area and in the Mediterranean and Atlantic Basins due to the risk of overexploitation of the aquifers. Therefore, in general agricultural producers face some uncertainty regarding the assignation of water (López-Martos and Rodríguez-Ferrero, 1997; Calatrava and Garrido, 2005).

The scarcity of water along with the principle of cost recuperation, introduced by the *European Water Framework Directive*, make it absolutely necessary for the management of this resource to be based on engineering, legal and economic criteria when dealing with an activity of this nature. Therefore, the main objective of this study is conduct and empirical analysis of the differences that exist in terms of economic efficiency in 156 irrigation areas in Andalusia[2]; consequently, in order to improvement irrigation water management it is indispensable to know this parameter and to be able to conduct comparative studies aimed at identifying the causes, whether they are modifiable or not, that lead to these differences in terms of efficiency.

In conclusion, this article offers a study of economic rationality in the use of scarce resources, specifically water. The methodological questions associated with the calculation of efficiency, its determinants and its temporary changes are dealt with in the next section. In the third section, there is a description of the variables used and the results obtained from the analysis carried out along with the main factors which explain the efficiency of all of the irrigation areas in Andalusia. Finally, in the fourth section there is a summary of the main conclusions and some recommendations which can be deduced from this study.

[2] This is the basic unit of planning in the irrigation system in Andalucía and is established by the agricultural authorities in such a way that Andalusia is divided into 156 irrigation areas.

2. METHODOLOGY

2.1. Efficiency analysis in irrigation areas in Andalusia

The concept of efficiency refers to the relationship that exists between the resources or *inputs* that are involved in the production process (industrial, agricultural, educational, etc.) and the *outputs* generated by this process (cars, agricultural products, educational results, etc.). This means that efficiency indicates the productivity of the resources. It can be said that a production process is efficient if it obtains the maximum from a product or *output* from some given *inputs* or resources (orientation *output*); or if a given *output* is obtained the lowest possible quantity of *inputs* (orientation *input*).

The concept of efficiency, applied to irrigation areas in Andalusia, would mean that these areas obtain a determined agricultural production minimizing the consumption of resources such as water, labour, etc. An irrigation area will be considered relatively efficient if there is no other one which is technologically similar in the sample which obtains the same value for gross production using less than one productive factor and no more than the remaining ones (efficiency in terms of *input*). Therefore, better management of irrigation in general would require the implementation of policies aimed at increasing the efficiency of the *inputs*, particularly water, since "more can be achieved with less water" through better management (Allan, 1999).

An appropriate method to assess the productive efficiency of irrigation areas is Data Envelopment Analysis (DEA). DEA methodology, which was established by Charnes *et al*. (1978), is a non-parametric approach to assess the performance of producers included in a sample, called DMUs (decision making units), which convert multiple *inputs* into multiple *outputs*[3]. In our case, the DEA seeks to assess the degree of skill with which an irrigation unit carries out the technical process of converting the *inputs* into *outputs*, comparing this activity with that of the other sample units which, when technologically homogeneous, are relatively efficient. In summary, the DEA method would consider as efficient those irrigation areas which, in relative terms, consume fewer *inputs* for a given *output*; which can help those responsible for water policies to know where the application of the water will create the greatest benefit.

The DEA is formulated as a problem of mathematical optimization (Expression 1). In general, if there are *n* producers to be assessed, who use *m* factors (*inputs*) and obtain *s* products (*outputs*), DEA calculates the relative technical efficiency of each producer (1, 2, …, n). Specifically, DMU_j consumes x_{ij} of *input i* and produces y_{rj} of *output r*. We assume that $x_{ij} \geq 0$ y $y_{rj} \geq 0$ and, moreover, it is assumed that each DMU has at an *input* and an *output* with a positive value. The set of production possibilities, which implicitly consider the original CCR model (Charnes, Cooper and Rhodes, 1978), verifies the free availability of *inputs* and *outputs*, constant performances of scale and convexity.

$$\text{Minimize } \theta - \varepsilon \left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ \right)$$

Subjet to

$$\sum_{j=1}^{n} \lambda_j \, x_{ij} + s_i^- = \theta x_{io} \qquad i = 1,2,\ldots,m$$

$$\sum_{j=1}^{n} \lambda_j \, y_{rj} - s_r^+ = y_{ro} \qquad r = 1,2,\ldots,s$$

$$\lambda_j \geq 0 \qquad j = 1,2,\ldots,n$$

$$s_i^- \geq 0; \; s_r^+ \geq 0$$

Expression 1

[3] See Johnes (2004) for further details on efficiency measurement.

In Expression 1, λ_j and θ are the variables of the model; and s_i^- and s_r^+ represent slacks of the *input* and *output*, respectively. In this case, we are assessing the efficiency of the DMU "o" (DMUj = DMUo)[4]. Solving this model of programming J times, once for each DMU, give us a vector estimation of technical efficiency. A unit is 100% efficient if and only if $\theta^* = 1$ $\{\theta^* (= \min \theta)\}$ and $s_i^- * = 0$ and $s_r^+ * = 0$ for all of *i* and *r*. A score equal to 1, along with null values for the slacks, indicates that the unit in question obtains a given production based on the minimum consumption of the available resources and the conditions under which it operates. A score which is lower than 1 ($\theta^* < 1$) indicates that the productive unit assessed is technically inefficient, and thus can reduce the quantities used for all its *inputs* in a proportion equal to (1 - θ^*).

2.2. Determinants of efficiency

Normally, as in our case, there will be panel data i.e. a set of irrigation areas which are observed at different moments in time. For each year, we will carry out an efficiency study with the DEA methodology proposed previously, obtaining an efficiency index for each irrigation area.

Furthermore, the efficiency study is validated if it can provide an analysis of the causes of the results. This is frequently carried out in the second stage, regressing the efficiency values obtained in the first stage on a series of explanatory variables which include specific determinants for each irrigation area. As there is a small number, *T*, of observations for each sample, it might seem sensible to consider estimating an econometric model with each one of the T cross sections and then compare the evolution of the coefficients of the model over time. There are two main reasons to decide against this strategy (Arellano and Bover, 1990; Novales, 1993). One of them is that the search for efficiency would suggest the need for simultaneous estimation with the T cross sections, i.e. with all the panel data, to optimally exploit the fact that the irrigation areas in each cross section are the same; furthermore, the panel data allow the estimation of models that take into account permanent differences between the irrigation areas although these are not observed[5].

The general specification of an econometric model of panel data is:

$$y_{it} = x_{it}'\beta + \eta_i + \upsilon_{it} \qquad (i = 1,\ldots,N; t = 1,\ldots T)$$

Expression 2

Where y_{it} is the dependent variable, in our case, the efficiency of irrigation area *i* in year *t*; x_{it} is a vector of observable explanatory variables (crop alternatives, land ownership structure, water distribution system for each plot of land and the level of dedication of the farmer to agricultural activity); β is the vector of parameters to be estimated; η_i is an individual effect[6]; and υ_{it} a term of perturbation.

In the case of irrigation areas it is known that there are correlations between the unobservable random variable η_i and the observable characteristics x_{it} (for example, the type of crop depends, crucially, on the type of soil and the climatic conditions). In this case, it would be a *fixed effects model*. Moreover, given the fact that the efficiency indexes have values between 0 and 1 (dependent variable), from the econometric perspective this poses a problem of censorship in the data which would bias the OLS (ordinary least square) estimations made in the

[4] The variable θ represents overall technical efficiency; λ_j is the vector of n elements that represent the influence that each DMU has on the determination of the assessed DMU.

[5] If this heterogeneity remains relatively constant over time, the panel can solve the problem. For example, the climatic and edaphological characteristics can be considered as relatively constant and are usually unobservable.

[6] This includes the effect of unobservable variables, characteristic of each irrigation area, which are stable over time (such as climatic and edaphological conditions).

sample of producers. In order to solve this, the procedure chosen in this study consists of estimating through maximum likelihood a *fixed effects tobit model*[7].

2.3. Changes in productivity

The information provided by the DEA method is static in nature and, obviously, in historical studies it is essential to offer a time analysis. The Malmquist productivity index which, on the one hand, is also based on the previous method, provides this dynamic perspective[8]; this index was originally introduced in the field of consumer theory (Malmquist, 1953) and was later applied to the measurement of productivity by Caves *et al.* (1982), in a context of production functions, and by Färe *et al*. (1994) in a non-parametric (DEA) context.

The Malmquist index between the periods t and t+1, taking as a reference the technology of period t ($M^{t}_{t,t+1}$), is the quotient between the efficiency index of the observation corresponding to the last period, assessed regarding the efficiency frontier of the first period (E^{t}_{t+1}), and the efficiency index of the observation of the first period, compared to the frontier of that same moment (E^{t}_{t}). An index higher than the unit means that the productivity has increased from period t to t+1 whilst values lower than 1 means a step back in this respect.

It is possible to decompose the variation of the productivity in two components, in such a way that the previous index is equal to the product of both. The first one includes the variation due to the movement of the efficient frontier, so that it expresses the degree to which the sector has undergone a technical change. In a similar way to the cases described previously, values which are higher than the unit would indicate technological progress and lower values would mean a technological recess[9]. In the second case, change in efficiency or *catching-up*, expresses the variation attributable to the improvement in relative performance of the unit with respect to the best ones in each period, i.e. its approach to the efficient frontier, if the value is higher than 1, or its distancing from this frontier, if it is lower than the unit[10].

3. EMPIRICAL FINDINGS

3.1. Data and variables

The values for water efficiency in the different areas have been obtained from the information provided by a study entitled *Inventarios y Caracterización de los Regadíos de Andalucía* (Inventories and Characteristics of Irrigation in Andalusia) published by the Andalusian Department for Agriculture and Fisheries (*Consejería de Agricultura y Pesca*, 1996, 2002)[11]: the irrigated surface for each irrigation area, its crop alternatives, the gross productivity of these crops, its water consumption, along with the average size of the exploitation, the land ownership structure, the manner of application of water to the land, the characteristics of the network of water distribution and the degree of dedication of the farmer. Nevertheless, given that this data base did not contain any information related to the exploitation costs, we have used the calculations made for each crop in the in the different irrigation areas by Rodríguez-Ferrero and Sánchez-Martínez (2007) within the framework of The Water Accounts for Andalusia (*Las Cuentas del Agua en Andalucía,*) funded by the Andalusian Water Institute (*Instituto del Agua*)[12].

[7] The *tobit model* was proposed by James Tobin in 1958 (Tobin, 1958).
[8] Changes in the total productivity of the factors through the Malmquist index can be calculated only when there is some panel data.
[9] The technological development shows the product increases that could be achieved, from one period to another, without altering the *input* quantities used. This last case could happen through the introduction of new techniques or production methods.
[10] For further details, see Färe *et al*. (1985) and Bjurek (1996), among others.
[11] These are the only years for which this information has been published.
[12] A differential characteristic with respect to other studies, including Rodríguez-Díaz *et al.* (2004) and Amores and Contreras (2009), which do not consider exploitation costs.

TABLE 1. Definition of the variables used in the analysis of productivity in irrigation areas (1996 and 2002)

Variables incorporated in the first stage (DEA method)	
Input Variables:	
SURFACE	Agricultural Surface (measured in hectares) which in a climatologically normal year is irrigated regardless of the duration or the number of irrigation sessions, including temporary irrigation or extra irrigation.
EX_ COST	Exploitation Cost. This includes: labour, seeds, fertilizers, machinery, energy and water costs.
WATER_CON	Water Consumption: Volume of water per surface unit (m^3/ha) which is supplied by the system for a traditional irrigation season.
Output Variable:	
GROSS_PRO	Gross Production (euros/ha): average performance (kg/ha) multiplied by the average price obtained by the farmer (euros/kg).
Variables incorporated in the second stage (*tobit regression analysis*)	
Explanatory or Independent variables	
EXPSIZE	Average size of the exploitation (number of hectares).
OWNER	Type of land ownership (percentage of explotations under ownership).
I_ASPERS	Percentage of surface area irrigated by aspersion.
I_GRAVITY	Percentage of surface area irrigated by gravity irrigation system.
I_LOCALI	Percentage of surface area irrigated by localized irrigation.
DITCH	Percentage of surface area irrigated by ditches.
REINDITCH	Percentage of surface area irrigated by reinforced concrete ditches.
PIPES	Percentage of surface area irrigated by pipes.
ONLYACTIV	Percentage of irrigated surface area in which the farmer is devoted to agriculture as his/her only activity.
FIRSTACTIV	Percentage of irrigated surface area whose business owner is devoted to agriculture as his/her only activity.
SECACTIV	Percentage of irrigated surface area whose business owner is devoted to agriculture as a secondary activity.
CEREAL	Percentage of irrigated surface area devoted to cereals.
MAIZESUN	Percentage of irrigated surface area devoted to maize and sunflower.
COTBEEFOD	Percentage of irrigated surface area devoted to cotton, beetroot and fodder.
STRAWBERRIES	Percentage of irrigated surface area devoted to strawberries.
FRUIT	Percentage of irrigated surface area devoted to fruits with a stone and citrus fruits (oranges and lemons).
SUBTROPI	Percentage of irrigated surface area devoted to subtropical fruits (avocado, custard apple, mango).
OPEN AIR_GREEN	Percentage of irrigated surface area devoted to open air and plastic greenhouse vegetables.
OLIVES	Percentage of irrigated surface area devoted to olives.
RICE	Percentage of irrigated surface area devoted to rice.
Dependent variable:	
INDEX	Efficiency indexes (from 0 to 1) obtained from DEA.

Source: authors' calculations.

3.2. DEA results

In Table I there is a summary of the information regarding the variables used in the first stage (efficiency study) and in the second stage (explanatory factors of efficiency). The results of the measurement of the efficiency of the irrigation areas through the DEA technique (first stage) are shown in Maps 1, 2 and 3 in Tables 2 and 3. Table 4 and Maps 1 and 2 include the results of the factors that have most impact on the highest or lowest efficiency (second stage). The analysis carried out demonstrates the great heterogeneity that exists in the Andalusian irrigation system regarding efficiency in terms of water use and other production factors. As will be explained, this phenomenon is due to the crop alternatives, which depend on the climatic and edaphological characteristics of the area, on the predominant irrigation system, on the land ownership structure and even on tradition.

In the efficiency study it has been considered that the objective of the irrigation areas is to achieve a determined production (gross production per hectare) consuming the lowest possible quantity of water and minimizing exploitation costs[13].

TABLE 2. Degree of relative efficiency in Andalusian irrigation areas

	Irrigated surface (ha)		% Irrigated surface	
Efficiency (%)	**1996**	**2002**	**1996**	**2002**
Efficiency = 100	28,587	23,929	3.51	2.68
100 > Efficiency ≥ 95	11,670	32,211	1.43	3.61
95 > Efficiency ≥ 80	481,743	234,105	59.04	26.22
80 > Efficiency ≥ 60	256,372	559,597	31.42	62.66
Efficiency < 60	37,548	43,167	4.60	4.83
Total	815,920	893,009	100.00	100.00

Source: authors' calculations.

From the results obtained, it is observed that the efficient areas have little presence in the irrigated surface area of Andalusia (Table 2). In 1996 these areas were in the South East, on the coasts of Almería and Granada[14], and in the South West, on the coast of Huelva[15] (Map 1). In 2002 there were also efficient areas in the North East of Andalusia, in the province of Jaén and some parts of the province of Granada (Map 2). In general, the relative efficiency of irrigation in Andalusia is, on average, around 80%, as can be observed in Table 3, which indicates that Andalusian agricultural sector could achieve the same production value consuming 20% fewer resources, particularly water.

In 2002, average efficiency decreased in relation to 1996 as can be observed in Table 3. This fact was due to three factors: the two first ones were related to the increase in exploitation cost, especially labour, and the decrease in the prices which were charged by farmers due the competition from other countries[16]. According to Reig and Picazo (2002, p. 15): "...the rapid technical change that has taken place in this sector has led to such significant increases in the productivity of agricultural resources that changes in agricultural supply have taken place at a much higher rhythm than changes in demand, and consequently there has been a drop in agricultural prices in relation to other sectors of the economy, or in relation to those good which

[13] The irrigated surface area (hectares) is also introduced in the DEA as an *input*. It has been considered to be a DEA model oriented towards *input* that assumes constant returns to scale.

[14] Specifically, Los Guiraos, Campo de Nijar and Poniente, in Almería and Motril/ Salobreña, and therefore not all of the irrigation areas situated on the coasts of Granada and Almería have the same degree of efficiency, as this depends on the percentage of irrigated surface area devoted to open air and greenhouse vegetables, along with the water supply.

[15] Irrigation areas in Condado Litoral and Palos de Moguer.

[16] Agricultural production should increase added value, *in situ*, through techniques of selection and presentation of their products and by improving the channels of commercialization since, if this does not happen, the products will not be competitive with other countries near Europe such as the North African Mediterranean countries (Rodríguez-Ferrero and Sánchez-Martínez, 2008).

are used to calculate the Consumer Price Index". Finally, the third factor is the fact that the consumption of water per hectare has increased all over the Andalusian irrigation area by 3%.

Map 3 shows the change in productivity in the irrigation areas of Andalusia. The total change in productivity is given by the Malmquist productivity change index. The irrigation areas which improved their productivity between 1996 and 2002 have an index higher than 1, and they represent 17.41% of the irrigated surface area (155,548 hectares); on the contrary, those whose productivity decreased in 2002 in comparison with 1996 have an index lower than 1, and account for 80.53% (719,141 hectares); the rest, 2.06% of irrigation (18,410 hectares), have maintained the same level of productivity. As has been stated previously, irrigation in Andalusia reduced its productivity by 7% in the period between the two years analysed (Malmquist index equal to 0.93).

3.3. What factors explain efficiency in irrigation areas?

In order to answer this question, it is necessary to estimate a *tobit fixed effects model* from the efficiency indexes in the first stage (Table 4). The important problems of multicollineality (a linear relation between independent variables) observed in Model 1 make it necessary to estimate model 2. From this it can be inferred that the different crop alternatives in irrigation areas represent the main factor that explains economic efficiency/inefficiency (see Table 4 and Maps 1 and 2). The agricultural products which contribute most to efficiency are, in this order, strawberries, open air and greenhouse vegetables, olive trees, subtropical fruits, winter extensive crops, and general fruit trees; the opposite happens in the case of rice. The following comments can be made in regarding these crops:

Strawberry growing has had most impact on efficiency, although this crop has exhausted its potential for direct consumption for commercial reasons. Furthermore, derived products along with the research which would allow for improvement in the organoleptical characteristics of strawberries without a reduction in quality could permit the use of a greater surface area for this crop. Moreover, irrigation of rice with underground water next to a strawberry growing area is illogical, although it does currently occur.

Vegetables, both grown in the open air and in plastic greenhouses, improve efficiency but an increase in production of these crops will only be possible if there is an improvement in their preparation and presentation in the market place.

Olive trees in Andalusia are almost always irrigated, although it would be necessary to stop this activity in the least efficient cases, in mountainous areas where in addition to their low productivity there has been an alarming increasing increase in erosion in Andalusian river basins, and introduce the activity in the flat plains near to the River Guadalquivir, in the East of the province of Córdoba and in the province of Jaén.

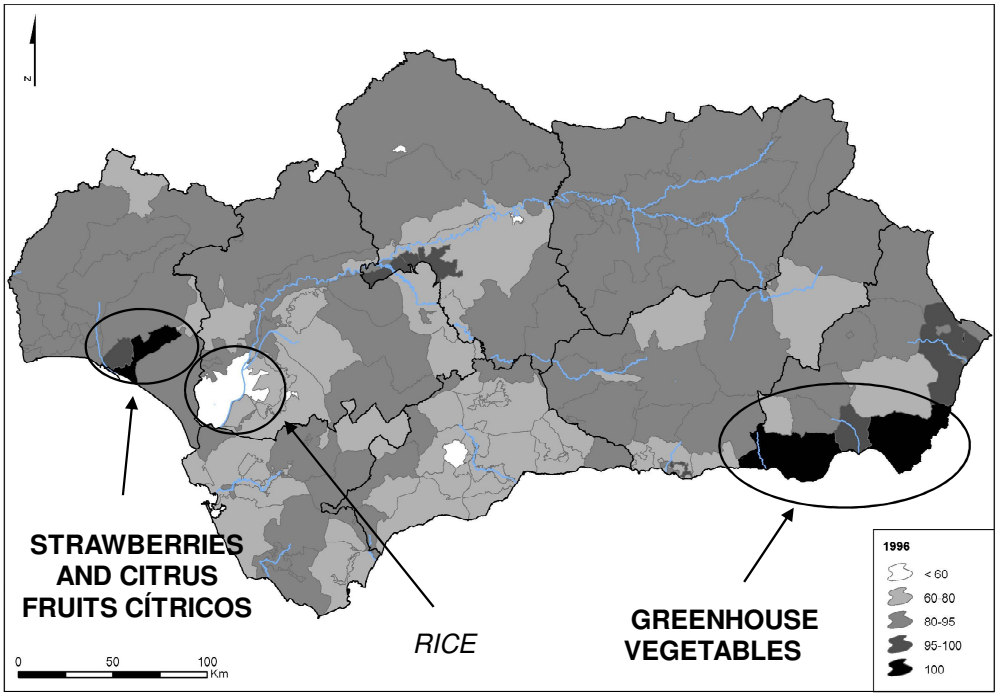
The irrigation of subtropical crops would allow for their expansion in the provinces of Granada and Málaga, improving their commercialization and this irrigation should come from the least efficient systems. It would be a step in the right direction if farmers stopped growing sugar cane in these areas. A change towards fruit trees in the Western plains of Córdoba and in all of these plains in the province of Seville would also be advisable.

Rice is the crop which has the most negative effect on efficiency and, taking into account the fact that part of the surface area is irrigated with ecological water sources in estuary of the River Guadalquivir, this crop needs a change in its water source at the beginning of the estuary, where there is a confluence of continental and transitional waters, along with research into the viability of a change in crop since recently there have been several experiences with open air vegetables in spring, when there has been little or no irrigation required due to the rainfall in the zone.

Another significant variable is the farmer's level of dedication to agricultural activity, and thus if the farmer is not fully dedicated to this activity this fact has a negative influence on efficiency in the use of productive factors since they do not obtain the maximum benefit from

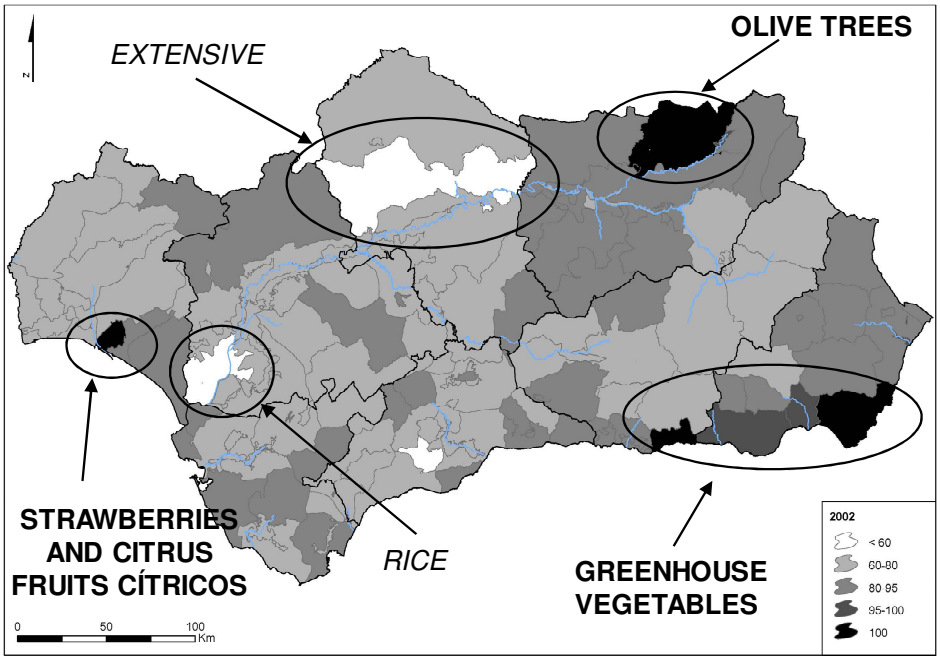
these productive factors by concentrating on crops that are easy to manage from a distance but which create a lower added value.

Map 1. Efficiency in Andalusian irrigation areas (1996)



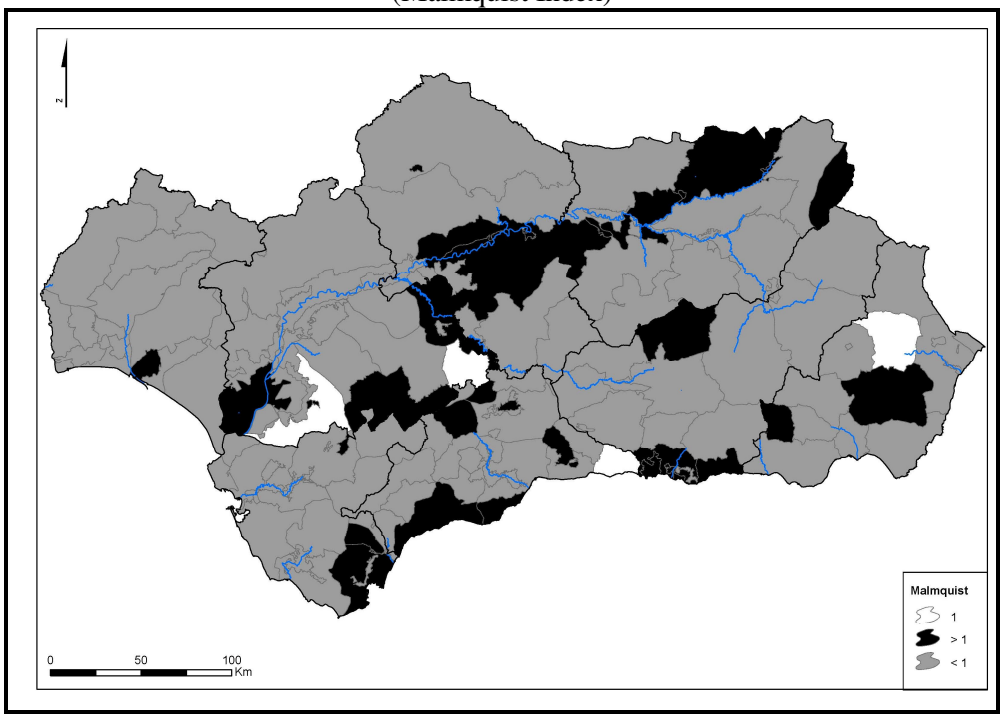
Source: authors' calculations.

Map 2. Efficiency in Andalusian irrigation areas (2002)



Source: authors' calculations.

Map 3. Changes in efficiency in Andalusian irrigation areas between 1996 and 2002 (Malmquist Index)



Source: authors' calculations.

TABLE 3. Efficiency in Andalusian irrigation areas (summary)

	Average value	Maximum value	Minimum value
Efficiency in 1996	81.28	100.00	43.92
Efficiency in 2002	79.36	100.00	47.18
Total change in productivity (Malmquist)	0.93	1.44	0.66
• Change in efficiency	0.97	1.48	0.69
• Technological change	0.95	1.03	0.85

Source: authors' calculations.

TABLE 4. Determinants of efficiency in Andalusian irrigation areas. Estimations of tobit fixed effect models

	MODEL 1			MODEL 2			Average value of the variable
	Coefficient		t-*student*	Coefficient		t-*student*	
Constant	0.810		1.397	0.675	**	18.500	
EXPSIZE	9.62E-04		1.399	6.47E-04		0.936	7.130
OWNER	2.23E-04		0.511	3.78E-04		0.851	80.324
I_ASPERS	-2.04E-04		-1.500				17.998
I_GRAVITY	7.73E-05		0.576				37.896
I_LOCALI	5.38E-05		0.379	-6.89E-05		-0.968	36.324
DITCH	8.09E-04		0.469	-1.94E-04		-0.938	12.948
REINDITCH	1.02E-03		0.598				27.163
PIPES	1.05E-03		0.613				60.097
ONLYACTIV	-1.41E-03		-0.255				57.520
FIRSTACTIV	-1.55E-03		-0.281	-5.21E-04		-1.374	34.765
SECACTIV	-4.08E-03		-0.732	-3.06E-03	**	-3.595	7.666
CEREAL	5.03E-04		1.565	1.24E-03	**	4.286	16.893
MAIZESUN	-2.06E-03	**	-4.601				10.899
COTBEEFOD	-2.64E-05		-0.176	6.64E-05		0.446	13.250
STRAWBERRIES	3.50E-03	**	5.054	4.75E-03	**	7.592	1.266
FRUIT	-4.91E-04		-1.440	6.64E-04	**	2.718	16.870
SUBTROPI	1.40E-04		0.289	1.26E-03	**	2.956	2.604
OPEN AIR_GREEN	1.81E-03	**	5.162	2.93E-03	**	11.526	18.394
OLIVES	9.62E-04	**	2.849	2.08E-03	**	9.185	23.262
RICE	-1.97E-03	**	-3.755	-9.42E-04	*	-1.911	1.149
Sigma	7.57E-02	**	27.745	7.81E-02	**	27.751	
Log-likelihood Function	438.785			426.955			

Dependent Variable: INDEX
** Coefficient statistically significant at 5% * Coefficient statistically significant at 10%

Source: authors' calculations.

4. CONCLUSIONS AND FINAL RECOMMENDATIONS

The forecasts for growth in hydric resources in Andalusia, along with the foreseeable increase in irrigation in this autonomous community in line with social demand and regional policies, will make it impossible to continue managing irrigation in the same pattern as at present. It is indispensable to introduce economic parameters into water management in order to achieve an efficient use which will permit farmers to reach the same level of production with a lower consumption of water. Therefore, this study seeks to determine the causes of efficiency in the use of productive resources, fundamentally water.

DEA applied to Andalusian irrigation areas allows us to know in the first stage the relative efficiency of these areas. The panel data which have been used also allow us to know whether or not these areas improve their level of efficiency over a period of time (specifically from 1996 to 2002). In the second stage, regression analysis makes it possible to delimit the main factors which explain efficiency or inefficiency – such as crop distribution, the level of dedication of the farmer to agricultural activity, the water distribution system or the land ownership structure – which can be used in the creation of public water policies to improve productivity in the less efficient areas.

One application of this analysis would be to concentrate on those factors which improve efficiency and can reduce water consumption, a factor which is of paramount importance in the Andalusian irrigation system and which is expected to undergo further restrictions with the application of the European Water Framework Directive.

Crop alternative is the factor that most influences efficiency and, therefore, it is necessary to concentrate more on this variable without forgetting the climatic and edaphological restrictions. The agricultural products which contribute most to efficiency are, in this order: strawberries, open air and greenhouse vegetables, olive trees, subtropical fruits, extensive winter crops and fruits; the opposite occurs in the case of rice.

The irrigation areas which have the greatest extension of these crops are the most efficient. In 1996 these areas were on the coasts of Huelva and Almería, whose main crops were strawberries in the case of Huelva (77.2% of the irrigated surface area) and open air and greenhouse vegetables in the case of Almería (90.8%). In 2002 the parts of Almería situated on the flatlands in the West of the province underwent a slight decrease in efficiency; nevertheless, the area known as Campo de Nijar was still efficient. Another area which joined this group was a part of the province of Huelva known as Palos de Moguer where strawberries occupied 80.5% of the irrigation area, an irrigation area on the coast of Granada where 99% of the irrigated surface area was devoted to open air and greenhouse vegetables and another area of the province of Jaén where olives trees grown with localized irrigation accounted for 80% of the irrigated surface area.

Another factor which also influences efficiency, although negatively in this case, is the farmer's dedication to agricultural activity as a secondary activity.

The average efficiency of Andalusian irrigation decreased slightly in 2002 due to the fact that the efficient areas occupied 16% less of the surface area and those areas which had a relative efficiency of 80%-95% were reduced to 52%. There was a loss of efficiency in the coastal areas of Almería because the exploitation costs increased in a greater percentage than prices charged by farmers, and this also occurred in practically all inland areas, with the exception of those with a predominance of olive trees and those in which cereals are substituted by this crop. It is necessary to point out that by mechanizing the collection of olives has led to a reduction in exploitation costs.

Finally, any process of expansion of the irrigation system should always take into account the dual objective of increasing efficiency and reducing water consumption and, therefore, it would only be recommendable if it aims to change certain crops or substitute drylands with irrigation systems. Furthermore, consideration should be given to abandoning those irrigation systems whose efficiency is very low, whilst taking into account the importance of population stability and urban planning.

REFERENCES

AGUILERA-KLINK, F. (2003): "Gestión autoritaria *versus* gestión democrática del agua", *Revista Archipiélago,* num. 57, pp. 34-42.

ALLAN, T. (1999): "Productive Efficiency and Allocative Efficiency: Why Better Water Management May Not Solve the Problem", *Agricultural Water Management*, vol. 40, pp.71-75.

AMORES, A.F., and CONTRERAS, I. (2009): "New Approach for the Assignment of new European Agricultural Subsidies Using Scores from Data Envelopment Analysis: Application to Olive-growing Farms in Andalusia (Spain)", *European Journal of Operational Research*, vol. 193, pp. 718-729.

ARELLANO, M., and BOVER, O. (1990): "La econometría de datos de panel", *Investigaciones Económicas,* vol. XIV (1), pp. 3-45.

BJUREK, H. (1996): "The Malmquist Total Factor Productivity Index", *Scandinavian Journal of Economics*, vol. 98, num. 2, pp. 303-313.

CALATRAVA, J., and GARRIDO, A. (2005): "Modelling Water Markets under Uncertain Water Supply", *European Review of Agricultural Economics,* vol. 32, pp. 119-142.

CAVES, D., CHRISTENSEN, L., and DIEWERT, E. (1982): "The Economic Theory of Index Numbers and the Measurement of Input. Output and Productivity", *Econometrica*, November, pp. 1393-1414.

CHARNES, A., COOPER, W.W., and RHODES, E. (1978): "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research*, vol. 2, pp. 429-444.

CONSEJERÍA DE AGRICULTURA Y PESCA DE LA JUNTA DE ANDALUCÍA. (1996, 2002)*: Inventario y Caracterización de los Regadíos de Andalucía*.

FÄRE, R., GROSSKOPF, S., and LOVELL, C.A.K. (1985): *The Measurement of Efficiency of Production*. Kluwer-Nijhoff Publishing, Boston.

FÄRE, R., GROSSKOPF, S., LINDGREN, B., and ROOS, P. (1994): "Productivity Developments in Swedish Hospitals: A Malmquist Output Approach". In Charnes, A., Cooper, W., Lewin, A., and Seiford, L. (eds.): *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic Publishers, Dordrecht.

JOHNES, J. (2004): "Efficiency measurement". In Johnes, G., and Johnes, J., (eds): *International Handbook on the Economics of Education*, Edward Elgar, Cheltenham, pp. 613-742.

LÓPEZ-MARTOS, J., and RODRÍGUEZ-FERRERO, N. (1997): "La gestión empresarial del agua". In Naredo, J.M.: *La economía del agua en España*, Fundación Argentaria-VISOR, Madrid, pp. 227-233.

MALMQUIST, S. (1953): "Index Numbers and Indifference Surfaces", *Trabajos de Estadística,* num. 4, pp. 209-242.

MAPA (2001): *Plan Nacional de Regadío. Horizonte 2008*, Ministerio de Agricultura, Pesca y Alimentación, Madrid.

NOVALES, A. (1993): *Econometría*, McGraw-Hill, Madrid.

REIG, E., y PICAZO, A. (2002): *La agricultura española: Crecimiento y productividad*, Ed. Caja de Ahorros del Mediterráneo, Alicante.

RICO-AMORÓS, A.M. (2006): "Políticas agrarias. Eficiencia socioeconómica y retos de futuro en los regadíos intensivos", *Boletín de la A.G.E.*, num. 41, pp. 113-149.

RODRÍGUEZ-DÍAZ, J.A., CAMACHO-POYATO, E., and LÓPEZ-LUQUE, R. (2004): "Application of Data Envelopment Analysis to Studies of Irrigation Efficiency in Andalucía", *Journal of Irrigation and Drainage Engineering*, May/June, pp. 175-183.

RODRÍGUEZ-FERRERO, N., and SÁNCHEZ-MARTÍNEZ, M.T. (2007): *La productividad del agua en el regadío andaluz*, Editorial Universidad de Granada. Granada.

RODRÍGUEZ-FERRERO, N., and SÁNCHEZ-MARTÍNEZ, M.T. (2008): "Un análisis de la eficiencia socioeconómica del agua en el regadío andaluz", *Revista Española de Estudios Agrosociales y Pesqueros*, num. 217, pp. 183-208.

TOBIN, J. (1958): "Estimation for Relationships with Limited Dependent Variables", *Econometrica,* vol. 26 (1), pp. 24-36.

EL MECANISMO DE DESARROLLO LIMPIO COMO ALTERNATIVA PARA LA REDUCCIÓN DE EMISIONES DE CARBONO

JOHANNA MAGALY ALVARADO ESPEJO
E-mail: jmalvaradox@utpl.edu.ec

Centro de Asesoría y Desarrollo Empresarial y Social (CADES)
Universidad Técnica Particular de Loja

Resumen

En el Ecuador el sector energético ha presentado una mayor actividad en cuanto a la inclusión del componente del MDL, especialmente en lo que se refiere a generación eléctrica basada en fuentes renovables. Las iniciativas de proyectos hidroeléctricos bajo el MDL presentadas en el país, mantienen la misma tendencia mundial con el 71% del total de proyectos.

Se analiza el caso del proyecto Hidroeléctrico Chorrillos, ubicado en la provincia de Zamora Chinchipe, el cual muestra la relación entre generación de energía hídrica y la reducción de emisiones de carbono. Entre las múltiples ventajas de plantear este tipo de proyectos, está la de obtener beneficios ambientales y económicos, en la parte ambiental reduce las emisiones de CO_2 y en lo que se refiere a beneficios económicos permite vender por anticipado reducciones de emisiones, generando ingresos para financiar la ejecución y operación del proyecto, lo cual no se obtiene con proyectos convencionales.

Palabras Clave: Mecanismo de desarrollo limpio, reducción de emisiones y desarrollo sostenible.
Área Temática: Economía de la Energía, del Ambiente, de los Transportes y de los Recursos Naturales.

Abstract

Ecuador in the energy sector has provided more activity as regards the inclusion of the CDM component, especially in regard to electricity generation based on renewable sources. The initiatives of hydroelectric projects under the CDM in the country, maintain the same global trend with 71% of all projects.
Case is analyzed Chorrillos Hydroelectric Project, located in the province of Zamora Chinchipe, which shows the relationship between power generation and water reducing carbon emissions. Among the many advantages of this kind of project is to obtain environmental and economic benefits in reducing the environmental emissions of CO_2 and in terms of economic benefits in advance to sell emission reductions, generating revenue to fund implementation and operation of the project, which is not obtained with conventional projects.

Keywords: Clean Development Mechanism, emission reduction and sustainable development.
Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

1. INTRODUCCIÓN

Los altos grados de contaminación del ambiente por los que atraviesa el planeta, debido a las distintas actividades humanas y en especial por las actividades productivas, tienen como resultado la creciente acumulación de Gases de Efecto Invernadero (GEI) en la atmósfera, provocando un calentamiento global, el cual tiene impactos negativos en la vida de los seres vivos, cuyos efectos serán irreversibles a menos que se ponga en práctica medidas para limitar las emisiones de estos gases.

El Tercer Informe de Evaluación del Grupo Intergubernamental de Expertos sobre el Cambio Climático (IPCC) 2001, señala que las tres cuartas partes de las emisiones de CO_2

antropogénicas que se han producido en los últimos 20 años son consecuencia de la quema de combustibles fósiles para procesos industriales, transporte y generación de energía.

El Comité Nacional del Clima en su estudio y evaluación "Bases para la Comunicación Nacional" (2001), afirma que el Ecuador al igual que en el resto del mundo, las principales fuentes de emisiones de CO_2 son causadas por la utilización de combustibles fósiles para la generación de energía, la deforestación y la actividad petrolera que tiene impactos en los servicios climáticos por la destrucción de bosques amazónicos para su explotación. Así encontramos que el cambio en el uso del suelo y silvicultura aporta con el 69,5% de CO_2, el sector energético con el 28,8% y la producción de cemento el 1,7%.

El informe Stern "La verdad del cambio climático" (2007) señala: "Si no actuamos, los costes y riesgos totales del cambio climático equivaldrán a una pérdida anual permanente de al menos un 5% del PIB mundial. Si se toma en consideración un conjunto más amplio de riesgos y efectos, los daños estimados podrían elevarse hasta el 20% o más de ese PIB. Por el contrario, los costos de actuar podrían limitarse hasta no superar, por año el 1% del PIB global", lo que demuestra la necesidad de tomar medidas inmediatas para minimizar pérdidas económicas, ambientales y sociales.

La información obtenida por los modelos de predicción del IPCC, obliga a los organismos nacionales e internacionales a tomar acciones emergentes para mitigar los efectos del cambio climático a través de políticas e instrumentos ambientales, que en la mayoría de las veces se establecen en acuerdos y convenciones internacionales, los cuales contienen objetivos obligatorios y cuantificados de limitación y reducción de GEI (SEOÁNEZ, 2002).

Como respuesta al cambio climático, la Convención Marco de las Naciones Unidas sobre Cambio Climático (CMNUCC) y el Protocolo de Kyoto establecieron las bases de un mercado para reducciones de emisiones de GEI, creándose entre otros el Mecanismo de Desarrollo Limpio (MDL), el cuál tiene el potencial de ayudar a los países en desarrollo a alcanzar un desarrollo sostenible mediante la promoción de inversiones ambientalmente amigables por parte de gobiernos o empresas de los países industrializados, además está diseñado para ayudar a los países industrializados a cumplir con la meta para 2012 de reducir sus emisiones a un nivel inferior en no menos de 5%, respecto de los niveles de 1990; dentro de éste mecanismo se negocian Certificados de Emisiones Reducidas (CER's).

Dentro de este contexto los proyectos MDL permiten a los emisores de los países desarrollados invertir en proyectos que reduzcan emisiones de tal manera que estas puedan ser utilizadas como parte de sus obligaciones. De esta manera, la implementación de proyectos de eficiencia energética adquiere mayor relevancia a la luz de las nuevas evidencias científicas y efectos del calentamiento global, así como de la reducción de las reservas de petróleo a nivel mundial. La coincidencia de ambos fenómenos hace suponer que la transición a un nuevo modelo de desarrollo, con nuevos patrones de producción y consumo sobre la base de fuentes de energía limpia, se convierte en una necesidad económica y social.

2. MERCADO DE CERTIFICADOS DE EMISIONES REDUCIDAS

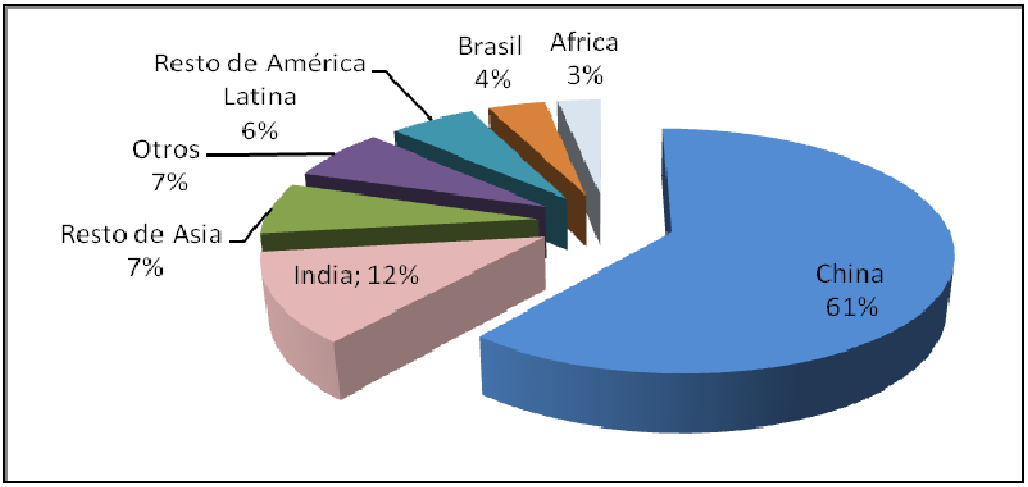
2.1 Oferta

La oferta de CER's, proviene de países en vías de desarrollo que buscan cumplir con su meta de reducción de emisiones. Como se observa en el gráfico 1, básicamente Asia domina el mercado con el 80% de volúmenes transados para el año 2006, con la presencia de China[1] con un 61%, aunque en el año 2005 poseía el 73%; sus proyectos se enfocan a sectores prioritarios como energías renovables (eólica, hídrica y biomasa), y la eficiencia energética que ha mejorado en el sector industrial la recuperación y utilización del metano.

[1] Líder por segundo año consecutivo.

India posee el 12% de mercado teniendo un ascenso del 3% desde el 2005, mientras que el resto de Asia ostenta un 7%, donde los registros de proyectos son financiados por entidades en la espera de vender CER´s (gráfico 1) (CAPPOR, 2007).

Gráfico 1. Oferta proyectos MDL 2006



Fuente: Banco Mundial. State and Trends of the Carbon Market 2007. Mayo 2007

Por su parte América Latina considerada como pionera en este mercado, en el año 2006 contó con el 10% de transacciones relacionadas al MDL, en esta región se encuentra liderando Brasil en primer lugar con el 4%.

La Comisión Económica para América Latina (CEPAL) (2006), señala que en la región, Brasil es el mayor potencial exportador de créditos de carbono, seguido por Colombia, Panamá, Costa Rica y Perú. América Latina se ha convertido en la región de países en desarrollo más activa en este mercado emergente con alrededor de USD 210,6 millones en créditos de carbono. El potencial de la región está en proyectos de biomasa, cogeneración con bagazo2 o energía basada en otra biomasa; en la gestión de desechos sólidos animales, granjas en el sector agrario; los proyectos hidroeléctricos y los Desechos Sólidos Municipales (DSM).

El continente africano para el mismo año conserva el 3% de mercado, es importante señalar que otros países de Africa Sub - Sahariana abarcan la tercera parte de este mercado.

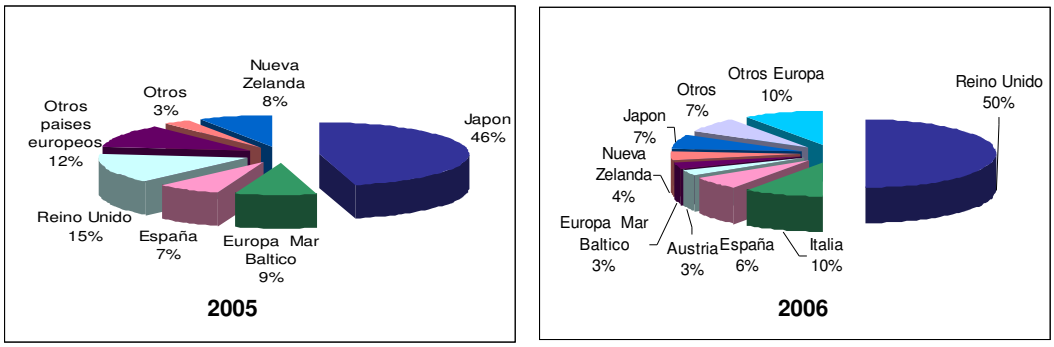
2.2 Demanda

La demanda internacional de CER's proviene básicamente de aquellos países que han ratificado el Protocolo de Kyoto y que se encuentran enlistados en el Anexo B del mismo.

Como se observa en el gráfico 2, la demanda de proyectos cambia radicalmente, en el año 2006 se encuentra dominada por compradores europeos con la participación del 86% del mercado. Japón en el 2005 representa el 46%, y para el 2006 el principal demandante es Reino Unido con el 50%. Italia representa el 10% del mercado debido a las adquisiciones de compañías privadas. España se mantiene en el mercado con un 6%, de los cuales el 25% de adquisiciones son gubernamentales. Austria con el 3% de volúmenes de mercado transados.

2 Residuo de la caña de azúcar.

Gráfico 2. Demanda proyectos MDL



Fuente: Banco Mundial. State and Trends of the Carbon Market 2007. Mayo 2007

2.3 Precios

A partir de la entrada en vigencia del Protocolo de Kyoto, los volúmenes transados y los precios han mostrado no sólo dinamismo si no también una volatilidad importante. El acceso a información de calidad es un requisito importante para participar en este mercado; sin embargo existe información limitada especialmente en lo que se refiere a precios, consecuencia de la falta de un sistema central para las transacciones de carbono, por lo tanto, es difícil comparar precios / cantidades en todo el mercado.

El promedio de precios por unidad de CER´s en el mercado primario fue de USD 5,15 en 2004, USD 7,04 en 2005 y USD 11,56 en el primer trimestre de 2006.

> *"Desde el punto de vista microeconómico, los CER´s son bienes "homogéneos" ya que su calidad puede ser definida con parámetros objetivos, y son además "sustitutos perfectos" de las EUA's*[3]*. Además, a diferencia de las EUA's, son "ahorrables" para siguientes períodos de compromiso, lo cual los hace muy atractivos para tomar posición y para acceder al mercado secundario. En consecuencia, podría pensarse que el precio de los CER´s no debiera ser inferior a los de las EUA´s. Otra ventaja de los CER´s es que son transables sin restricciones en todos los mercados de los países industrializados que ratificaron el Protocolo de Kyoto: UE, Japón y Canadá. A diferencia, las EUA's no son válidas para cumplimiento fuera del Mercado Europeo de Derechos de Emisión (EU ETS)"* (OYHANTCABAL, 2007).

3. EL SECTOR ELÉCTRICO Y EL MECANISMO DE DESARROLLO LIMPIO

El sector energético, dado su carácter de soporte del sistema productivo de la sociedad esta más relacionado con el medio ambiente. Este sector, comprendiendo la extracción, producción, transporte y uso de la energía, es la fuente más importante de gases de efecto invernadero[4]. Los principales gases de efecto invernadero producidos por el sector energético son el CO_2 y el CH_4 procedentes de la quema de combustibles fósiles, así como el de las minas de carbón (en disminución), y de las instalaciones de hidrocarburos y gas (Ministerio De Industria, Turismo y Comercio, 2003).

El informe *Little Green Data Book (2007),* sostiene que los combustibles fósiles, se utilizan para generar el 66% de la electricidad del mundo. En el Medio Oriente, la proporción de combustibles fósiles en la generación de electricidad llega a 93%, mientras que la cifra respectiva es de 82% en Asia Oriental y el Pacífico y en Asia Meridional. Al otro extremo del espectro está América Latina y el Caribe, donde el 38% de la electricidad se produce con combustibles fósiles.

[3] Siglas en inglés de los permisos de emisión del Esquema de Comercio de Emisiones de la Unión Europea.
[4] Representa el 80% de las emisiones de CO_2 a escala mundial.

Ríos[5] (2007) en su artículo "Eficiencia Energética, Cambio Climático y Desarrollo Social", señala que en las próximas tres a cuatro décadas, el planeta requerirá cada vez más energía (2 a 3% por año) debido al elevado crecimiento vegetativo y crecimiento económico, principalmente de los países en vías de desarrollo, que con justa razón requieren más energía para desarrollarse y disminuir en algo la brecha que los separa de los países desarrollados.

El impacto adverso causado por el consumo y producción de energía puede mitigarse reduciendo el consumo o sustituyendo la generación con combustibles fósiles por el uso de energía renovable. Aún no se obtiene pleno provecho del enorme potencial de las energías eólica, solar, hidráulica, geotérmica, de biomasa y oceánica. Estas fuentes de energía son viables desde el punto de vista técnico y económico. Además, suministran energía en forma perenne y limpia[6]. De esta manera las energías renovables podrían solucionar muchos de los problemas ambientales, como el cambio climático, los residuos radiactivos, las lluvias ácidas y la contaminación atmosférica.

De esta manera la sustitución de fuentes de energía contaminantes por proyectos MDL basados en fuentes de energía renovables para la producción de electricidad es una de las principales medidas que el sector energético está impulsando, dentro del contexto del Protocolo de Kyoto, con objeto de luchar contra el cambio climático.

Bajo este contexto la necesidad de gestionar el cambio climático ha llevado a introducir al Mecanismo de Desarrollo Limpio en proyectos de energía renovable como son los Hidroeléctricos para contribuir a la reducción de emisiones globales.

PROYECTOS HIDROELÉCTRICOS A NIVEL MUNDIAL BAJO EL MDL

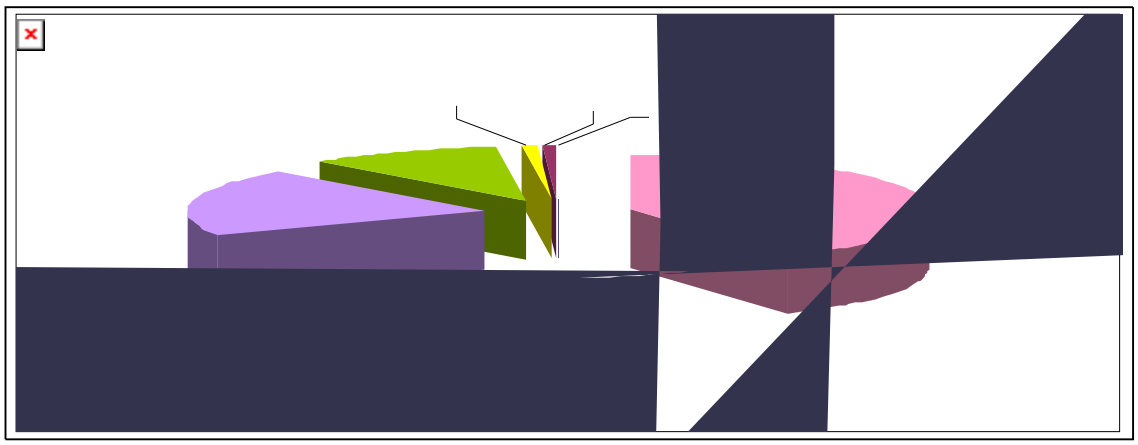
El potencial eléctrico de origen hidráulico aún sin aprovechar es enorme, apenas se utiliza el 17% a escala mundial, cifra que se reduce al 8% en el Tercer Mundo. Por ello, en las últimas décadas la promoción de estas instalaciones se ha incrementado. En la actualidad hay 36.327 grandes embalses, que almacenan 5.500 km^3 de agua. Según datos del World Watch (1997) citado en Ministerio De Industria, Turismo y Comercio de España (2003), la producción mundial de energía hidroeléctrica supera anualmente los 2.000 TWh de producción, lo que representa el 20% de la producción mundial de electricidad.

El poder considerar proyectos basados en fuentes de energías renovables como proyectos MDL puede impulsar el desarrollo de los mismos al convertir dichos proyectos en rentables, a través de los ingresos adicionales obtenidos por los certificados de CO_2 asociados. Esto significa que cada proyecto de energía renovable tiene un efecto positivo en el ambiente, tanto a nivel local como global.

En el gráfico 3 se puede observar que dentro de los proyectos de energía renovable, los hidroeléctricos mantienen mayor participación con el 41,1% del total, seguido por proyectos de biomasa con el 26,8%, viento con el 19,9%, biocombustibles con el 10,7% y el restante 1,5% se distribuye en geotérmico, solar y mareomotriz. Concluyendo la participación de proyectos hidroeléctricos tienen mayor participación en el mercado mundial del MDL.

[5] Secretario Ejecutivo de la Organización Latinoamericana de Energía (OLADE), por el periodo (Enero 2006 – Octubre 2007).
[6] El complemento indispensable es un uso adecuado de la energía, es decir, la eficiencia y el ahorro.

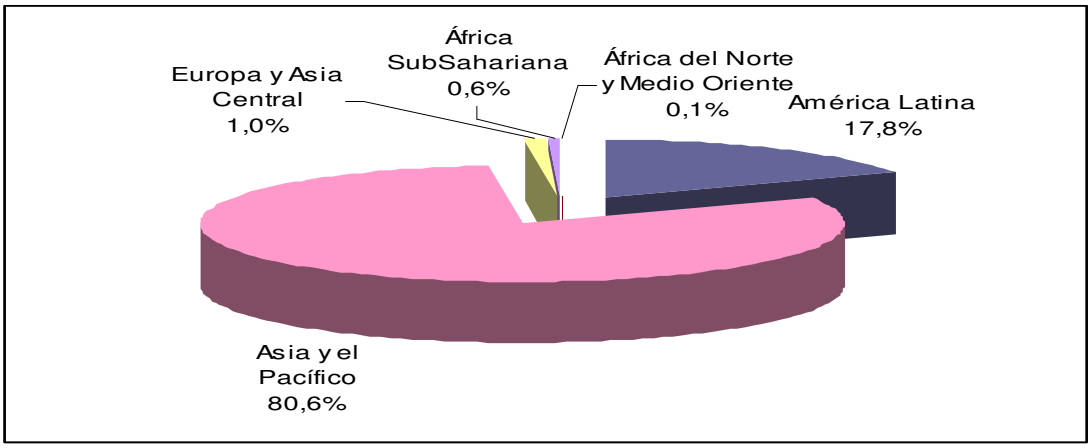
Gráfico 3. Tipos de proyectos de energía renovable MDL



Fuente: UNEP Risoe MDL / JI Pipeline y Análisis de la base de datos, enero 8 de 2008

En el gráfico a **c**ontinuación se puede observar que Asía y el Pacífico y América Latina tienen mayor participación en el mercado hidroeléctrico MDL con el 98,4% del total. China e India acogen el mayor número de proyectos en Asia, y en América Latina el país con mayor participación en este mercado es Brasil con 52 proyectos; sin embargo China con 463 proyectos tiene la mayor participación mundial.

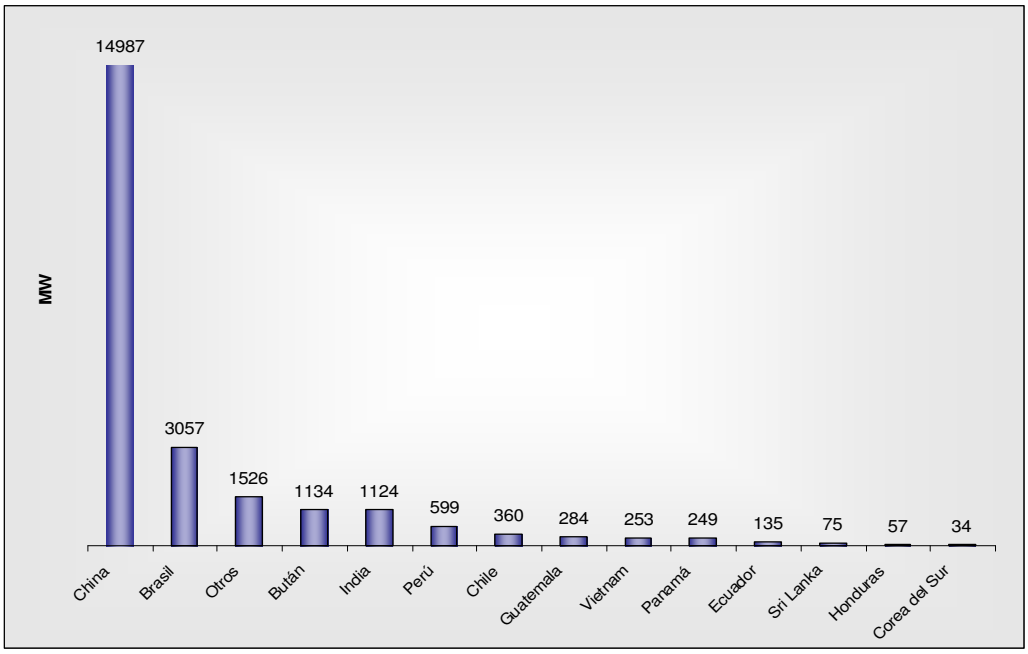
Gráfico 4. Proyectos hidroeléctricos MDL distribuidos por región



Fuente: UNEP Risoe MDL / JI Pipeline y Análisis de la base de datos, enero 8 de 2008

El gráfico muestra los MW que generan los proyectos registrados en el MDL, el total de generación mundial es de 23.874 MW, contribuyendo en gran medida China, país líder en este tipo de proyectos con una generación de 14.987 MW, sin embargo Ecuador tiene una participación menor con 135 MW, por la razón que es un mercado nuevo y se encuentra en desarrollo. Pese a ello nuestro país representa el 7% en la oferta de América Latina.

Gráfico 5. Energía generada (MW) por proyectos hidroeléctricos MDL por países



Fuente: UNEP Risoe MDL / JI Pipeline y Análisis de la base de datos, enero 8 de 2008

3.1 Proyectos MDL implementados en Ecuador

El Ecuador interviene en el Mercado de Carbono a través de proyectos MDL, los cuales se dirigen a los sectores energético, forestal, industrial, y gestión urbana; que se describen a continuación.

Desarrollo energético: Se desarrollan proyectos de generación eléctrica basada en fuentes renovables. En este punto también se pueden añadir proyectos de reducción de pérdidas en la transmisión y distribución de electricidad e iniciativas de interconexión; así como, iniciativas de eficiencia energética tanto en la oferta como en la demanda de electricidad, en general representan el 53% de proyectos MDL.

Fomento forestal: Comprende proyectos de reforestación y manejo de bosques con una participación del 24%.

Procesos industriales: Proyectos por el aumento de la eficiencia energética y el mejoramiento en la gestión de procesos y el manejo de residuos líquidos en distintas ramas industriales del país, que interviene con 14%.

Gestión urbana: Abarca proyectos concebidos en el mejoramiento de la gestión de los residuos sólidos y líquidos urbanos; así como iniciativas de mejoramiento en la gestión de transporte urbano colectivo, con el 9%.

En relación con el número de proyectos que se desarrollan dentro del MDL en Ecuador existen 18 proyectos; incluidos los validados, en solicitud de registro y los registrados[7] (cuadro 1).

Cuadro 1. Proyectos MDL en Ecuador

Estatus de los Proyectos	Total
En validación	8
Solicitud de Registro	1
Registrados	9
TOTAL	**18**

Fuente: UNED Risoe MDL / JI Pipeline y Análisis de la base de datos, enero 8 de 2008

[7] La información está basada en los datos oficiales que aparecen en el sitio Web del MDL y no incluye aquellos proyectos que están siendo formulados y que no han pasado a la etapa de validación.

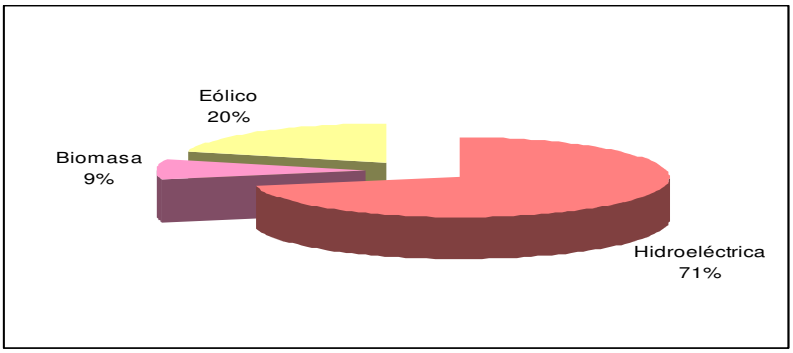
3.2 Proyectos Hidroeléctricos del ECUADOR basados en el MDL

En Ecuador la mayoría de problemas ambientales se originan en la explotación y consumo de forma irracional de los recursos naturales energéticos, especialmente del petróleo. La responsabilidad del sistema energético del país en el efecto invernadero va a crecer sustancialmente en el futuro cercano debido al crecimiento poblacional y a las metas de desarrollo socioeconómico.

Según el censo de Población y Vivienda de 2001[8], la cobertura energética del país es de 89,67%, lo cual refleja que existe ineficiencia en el sector, puesto que el Ecuador cuenta con un alto porcentaje de recursos energéticos, los cuáles aún no han sido explotados.

El sector energético/eléctrico nacional ha sido uno de los de mayor actividad en cuanto a la inclusión del componente MDL en proyectos especialmente en lo que se refiere a generación eléctrica basada en fuentes renovables El portafolio[9] presenta 35 iniciativas distribuidas en proyectos de generación hidroeléctrica (71%), eólica (20%) y biomasa (9%) como se presenta en el gráfico a continuación (CORDELIM, 2006).

GRÁFICO 6. Clasificación de proyectos de energía limpia

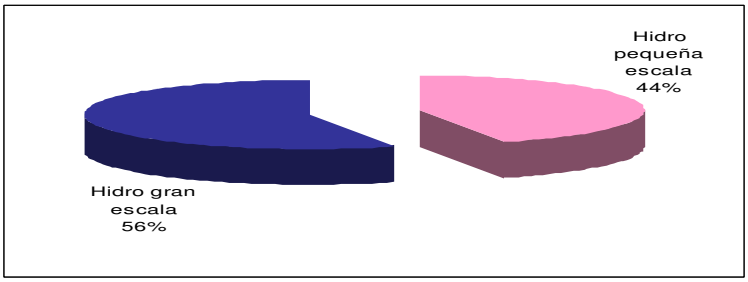


Fuente: CORDELIM. Proyectos MDL en Ecuador, Portafolio Indicativo. Septiembre 2007.

Ecuador posee 25 iniciativas de proyectos Hidroeléctricos dentro del MDL, de las cuales 9 proyectos se encuentran oficialmente registrados por la Junta Ejecutiva, listos para vender los Certificados de Reducción de Emisiones.

En lo relacionado al alcance de los proyectos hidroeléctricos, de las 25 iniciativas en el proceso de calificación como MDL, los que mantienen una mayor representación son de gran escala (56%) aportando con 384,5 MW al sistema eléctrico nacional si todos estos proyectos entran en operación. Sin embargo el porcentaje de los proyectos de pequeña escala es casi similar (44%) con 74,28 MW.

Gráfico 7. Proyectos hidroeléctricos según su tamaño



Fuente: CORDELIM. Proyectos MDL en Ecuador, Portafolio Indicativo. Diciembre 2007.

[8] Instituto Nacional de Estadísticas y Censos (INEC).
[9] Para mayor detalle del portafolio revisar: www.cordelim.net

La central "Chorrillos" es una pequeña hidroeléctrica de 3.96 MW , ubicada en el cantón Zamora, provincia de Zamora Chinchipe, en la zona sureste del Ecuador, ésta central será conectada al Sistema Nacional Interconectado (SNI), contribuyendo con una producción media de 23.178 MW/h por año. Para la construcción y operación del proyecto se creó la Compañía de Economía Mixta "HIDROZAMORA C.E.M. El proyecto hidroeléctrico es considerado como un proyecto de energía renovable que contribuye al desarrollo sostenible del país, debido a que la producción de energía estará dirigida a áreas estratégicas de la economía. La hidroeléctrica aprovechará los recursos hídricos de la región oriental del país, en este caso de la cuenca de la Quebrada "Chorrillos".



Actualmente la central hidroeléctrica "Chorrillos" se encuentra en fase de construcción y prevé iniciar sus operaciones durante el año 2009. En lo referente a las etapas del proceso MDL el proyecto se encuentra presentando la documentación para acceder a la Aprobación Nacional del país anfitrión que en este caso es Ecuador.

La metodología es elegida en base a la tipología del proyecto, siendo el proyecto Chorrillos" de pequeña escala, se utiliza la metodología para la Generación de Electricidad para Suministro de una Red de Distribución (AMS – I.D.), establecida por la Junta Ejecutiva, la cual tiene

menores requerimientos tanto a nivel metodológico y de desarrollo del Diseño Final del Proyecto (PDD) como también en el proceso de Validación, que en la metodología dirigida a proyectos de mediana y gran escala.

El Proyecto Chorrillos cumple con las condiciones de aplicabilidad de la metodología seleccionada: La actividad de proyecto es una planta hidroeléctrica a filo de agua. El proyecto entregará electricidad al sistema de distribución que es o hubiera sido suministrado por al menos una unidad generadora operando con combustible fósil como se establece en AMS I.D. La capacidad instalada del proyecto es 3.96 MW, inferior a 15 MW (el límite máximo permitido para ser considerada como una actividad de proyecto de pequeña escala).

3.2.1 Línea base

De acuerdo con la categoría de proyecto y metodología correspondientes, la línea de base corresponde a la energía producida por la unidad de generación renovable (MW/h) multiplicada por un coeficiente de emisión (tCO_2e/MW/h) calculado de manera transparente y conservadora utilizando el margen combinado[10].

Cálculo del Factor de Emisión según el margen combinado: De esta forma, según la metodología, la información clave para determinar el escenario de línea de base está dada en el siguiente cuadro.

Cuadro 2. Información para el cálculo de emisiones reducidas

Fecha	Fuente
Generación eléctrica del proyecto (GE)	23.178 MW/h año. HIDROZAMORA C.E.M.
Factor de emisión de margen combinado (FE)	0.67 t CO_2e/MW/h. CORDELIM[11],

Fuente: HIDROZAMORA C.E.M.2007.

Con esta información se determinó la reducción de emisiones (RE), de la siguiente manera:

$$RE = FE \cdot GE$$
$$RE = 0.67\, tCO_2e / MWh * 23178\, MWh / año$$
$$RE = 15.529\, tCO_2e / año$$

Por lo tanto, al aplicar el factor de emisión oficial de la red de Ecuador, las reducciones de emisiones resultantes son 15.529 tCO_2e/año, aplicando la formula establecida.

3.2.2 Demostración de adicionalidad

La demostración de la adicionalidad se la realizó a través del análisis de barreras:

Barreras de Inversión

Es importante considerar que el costo de un mini proyecto hidroeléctrico es realmente alto por la escasas oportunidades que presenta el país, para obtener recursos financieros, así como el obtener préstamos en el sector público y privado, para poder superar los altos costos en la etapa de construcción de éste tipo de proyectos. Debido a esto, el análisis financiero debe usar tasas de descuento bastante altas, dando más importancia a las fuentes adicionales de ingresos, como la negociación de CER's.

[10] Es una combinación de los factores del Margen de Operación (MO) y del Margen de Construcción (MB).
[11] http://www.cordelim.net/

Barreras Tecnológicas

A pesar de que la tecnología utilizada en estas plantas hidroeléctricas es bastante conocida en Ecuador y América Latina, existen barreras de naturaleza tecnológica y logística asociadas con su aplicación y mantenimiento. El equipamiento para la operación de esta pequeña planta hidroeléctrica no es producido en Ecuador; debe ser importado desde el extranjero. Esto representa un problema para el desarrollador del proyecto, porque se depende de las importaciones para establecer y mantener las nuevas instalaciones y también de la disponibilidad en el extranjero para la entrega de dichos equipos.

Para permitir el correcto funcionamiento del proyecto, HIDROZAMORA C.E.M. tuvo que capacitar a su personal para que adquiriera conocimientos tanto en la transmisión de electricidad como en estrategias operacionales y de precios para el mercado eléctrico. Dicha capacitación y construcción de capacidad no podría lograrse sin una importante inversión. Los incentivos del MDL ayudarán a disminuir dichos gastos.

Barreras Institucionales

En Ecuador todo el proceso para obtener los permisos y licencias legales y administrativas para operar plantas hidroeléctricas toma más tiempo que los permisos necesarios para operar proyectos en plantas de energía térmica.

La etapa de construcción del proyecto ha presentado por un lado demoras debido a la falta de recursos económicos, sin embargo con la venta anticipada de CER´s se ha logrado en parte continuar; y por otro lado por los excesivos tramites burocráticos

3.3 Comercialización de CER'S

Negociación del Contrato de Emisiones Reducidas (ERPA)

HIDROZAMORA C.E.M. mediante este acuerdo vendió por adelantado (antes que empiece a operar la hidroeléctrica) a través de MGM[12], certificados de carbono por los tres primeros años, a un precio de USD. 7, lo cual permitió obtener recursos económicos para continuar con la construcción de la planta; sin embargo ésta venta se la puede realizar en el momento que éstos certificados son emitidos.

Los desarrolladores del proyecto junto con el broker MGM[13], han necesitado aproximadamente tres meses para coordinar las actividades, en este caso el Broker asume todos los costos que implica el proceso del MDL.

3.4 Análisis de beneficios financieros, socio – económicos e impactos ambientales

3.4.1 Beneficios financieros

El análisis financiero en proyectos MDL, al igual que en cualquier otro proyecto, permite determinar los beneficios o pérdidas en los que se puede incurrir al realizar la inversión, analizar la viabilidad del proyecto y sirve de ayuda en la toma decisiones. Este análisis se enfoca a estudiar las distintas posibilidades de generación de energía limpia, estimando la rentabilidad para los distintos escenarios del proyecto dependiendo del precio de los CER's.

El análisis financiero se ha realizado tomando como referencia el Flujo Neto de Caja de la Central Hidroeléctrica Chorrillos, el proyecto tiene una vida útil de 50 años, tomando en cuenta los siguientes costos e ingresos:

[12] Empresa internacional intermediaria especializada en la adquisición de CER´s en América Latina. Tiene el respaldo de la empresa japonesa J Power de Japón, no sólo para adquirir CER´s si no también, si fuera necesario, pueden contribuir con capital para el desarrollo de los proyectos. www.mgminter.com

13 En el caso de la hidroeléctrica el proceso se lo realizo junto con un intermediario financiero debido a la experiencia para buscar las mejores condiciones financieras para la venta de Cer´s, sin embargo es un proceso que lo puede asumir la empresa sin necesidad de acompañamiento de un bróker.

Inversión Referencial: A noviembre del 2007 el presupuesto referencial de inversión del proyecto es de USD 7.508.572,59.

Capital de Trabajo: Capital de trabajo es de USD 420.000,00.

Ingresos: Los ingresos provienen de dos fuentes: la venta de energía y de la venta de CER´s (cuadro 3). Con el propósito de cuantificar los beneficios que el proyecto tendría como producto de la venta de su producción energética, se ha tomado como referencia para el análisis los precios de venta de potencia y energía aprobados por el Consejo Nacional de Electricidad (CONELEC), que es de USD 0.055 por kilovatio. El precio de los certificados son los considerados en la negociación con MGM que es USD 7, sin embargo para realizar un análisis de sensibilidad se ha tomado en cuenta los precios de USD 5 y USD 10.

Además se consideran los ingresos por inversiones a plazo fijo en pólizas de USD 1.372.183,80.

Cuadro 3. Ingresos del proyecto

Ingreso por Potencia Remunerable
Ingreso por Energía Primaria
Ingreso por Energía Secundaria
Ingreso por venta de CER's
Ingreso por venta anticipada(3 años) de CER's
Ingreso por inversiones L.P (Bco Loja)

Fuente: HIDROZAMORA C.E.M. Flujo Neto de Caja. 2007.

Costos: Se ha realizado un cálculo detallado de los costos de operación y mantenimiento para el Proyecto, mediante una esquematización del personal requerido y una estimación de los costos de los materiales e insumos a utilizarse; así mismo los costos financieros, relacionados a seguros, por reposición y por préstamo que se presentan en el cuadro 4.

Cuadro 4. Costos del proyecto

Costos Operativos
Costos de Operación y Mantenimiento
Beneficios a la comunidad (obras de compensación)
Imprevistos 1%
Costos Financieros
Seguros
Costos Financieros por mora en pago de energía
Costo de Reposición de Turbinas y Equipo de Control (intereses)
Costo de Reposición Generadores y Equipo de Electromec. (intereses)
Costo de Reposición de Línea de Transmisión y Accesorios (intereses)
Costo financiero Crédito Inicial (intereses)

Fuente: HIDROZAMORA C.E.M. Flujo Neto de Caja. 2007.

Reposición de equipos: Los costos por concepto de reposiciones intermediarias de equipos durante la vida útil del proyecto de generación se consideran a partir del año 32 hasta el 45 de operación del proyecto.

Tasa de Descuento: La tasa de descuento utilizada en los procesos de actualización de los flujos y cálculo del Valor Actual Neto (VAN) es el 9.2% por la ponderación realizada del aporte del capital de los accionistas y el préstamo. La tasa de descuento para el cálculo del TIR es de 10%.

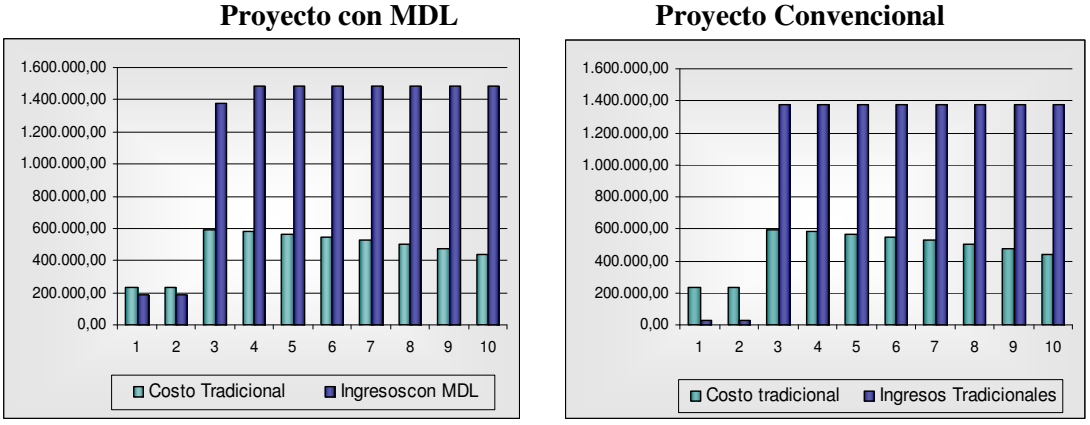
3.4.1.1 Análisis de la adicionalidad financiera por la venta de CER´s

La Central Hidroeléctrica "Chorrillos" con la implementación del MDL obtendrá recursos económicos adicionales por la venta de CER´s; lo que contribuirá al cierre financiero del proyecto, el cual no se hubiera logrado sin el MDL debido a existencia de barreras de financiamiento, propias de este tipo de proyectos.

Al comparar los ingresos del proyecto sin MDL, con los ingresos del proyecto financiado con MDL, como se observa en el gráfico 9[14], podemos concluir que el proyecto con MDL tiene ingresos adicionales, contrario a lo que ocurre en un proyecto convencional. Los costos en ambos casos son los mismos debido a que los costos de transacción del MDL son asumidos por MGM[15].

En el proyecto con MDL desde el primer año de operación ya se obtienen ingresos por la venta anticipada de CER´s, anualmente se venderán 15.529,26 ton de CO_2 anuales durante los 50 años de vida útil del proyecto, obteniéndose ingresos adicionales de USD 108.704,82 al año; sin embargo HIDROZAMORA C.E.M. ha conseguido vender CER´s anticipados durante los tres primeros años de operación mientras que siendo un proyecto convencional no contaría con este ingreso adicional, correspondiéndole solo los ingresos por la venta de energía a partir del tercer año de operación, lo que tal vez podría ser un obstáculo para que el proyecto se lleve a cabo.

Gráfico 9. Adicionalidad financiera



Fuente: HIDROZAMORA C.E.M Flujo Neto de Caja. 2007.

Análisis de Sensibilidad

En esta parte se analizó el efecto que produce el cambio del precio de los certificados en la rentabilidad del proyecto en un escenario pesimista, estable y optimista; a un precio de los CER´s de USD 5, USD7 y USD 10 respectivamente y sin MDL.

- ✓ **Valor Actual Neto (VAN)**

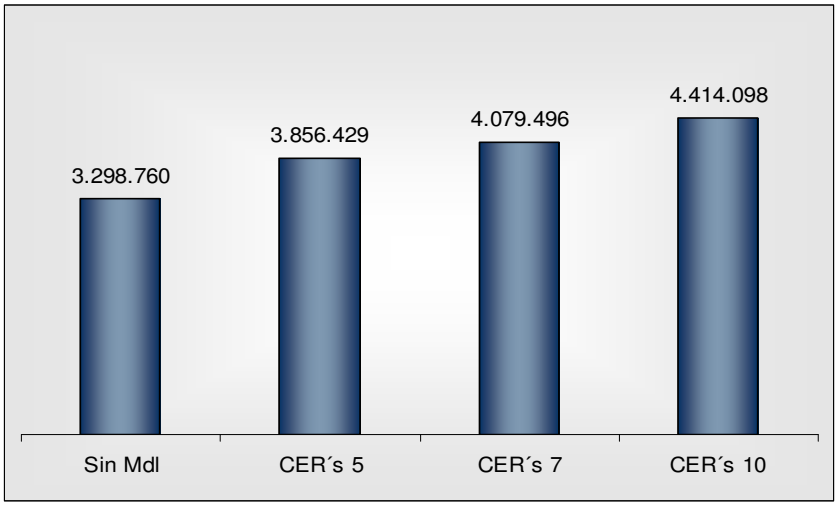
La valoración de los valores futuros en el presente obtenidos a través del flujo de caja de HIDROZAMORA C.E.M. sin la venta de CER's muestra un valor de USD 3.298.760, en un escenario pesimista con un precio de USD 5 por CER's el proyecto tiene una valoración de USD 3.856.429 y a un precio de USD 10, escenario optimista USD 4.414.098. En todos los

[14] Por fines didácticos el grafico se lo ha realizado para 10 años, sin embargo el proyecto tiene un horizonte de vida de 50 años.
[15] Broker de Hidrozamora C.E.M..

casos el valor presente es positivo por lo que se espera rentabilidad del proyecto, aunque no se venda certificados.

El comportamiento del valor presente neto a los cambios del precio de CER´s se muestra en forma gráfica a continuación:

Gráfico 10. Análisis de sensibilidad: Valor Presente Neto



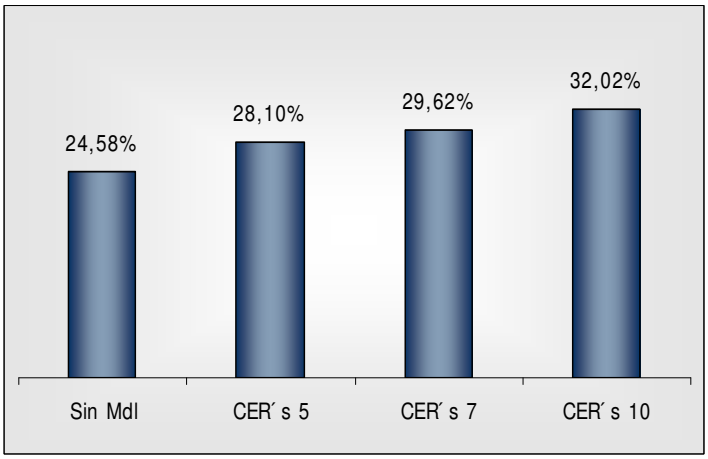
Fuente: HIDROZAMORA C.E.M. Flujo Neto de Caja. 2007.

Del análisis de éste indicador se concluye que el proyecto es rentable independientemente de la venta de CER's, sin embargo se puede obtener un mayor valor del proyecto con la venta y conforme mejoren los precios en el mercado.

✓ **Tasa Interna de Retorno (TIR)**

Como se observa en el gráfico siguiente, la TIR aún si los ingresos del MDL es alta (24,58%); y si se considera el impacto del MDL en los diferentes escenarios pesimista, moderado y optimista la TIR tiene un mayor rendimiento siendo de 28,10%, 29,62% y 32,02% respectivamente. El valor de este indicador aumenta a un mayor precio de los CER´s, por lo que en principio, la rentabilidad de la declaración de un proyecto como MDL es muy atractiva. Sin embargo hay que considerar los altos riesgos que conlleva actualmente la aprobación como MDL de estos proyectos, debido a que es un proceso largo y complicado (Coto y Morera, 2006).

Gráfico 11. Análisis de sensibilidad: Tasa Interna de Retorno



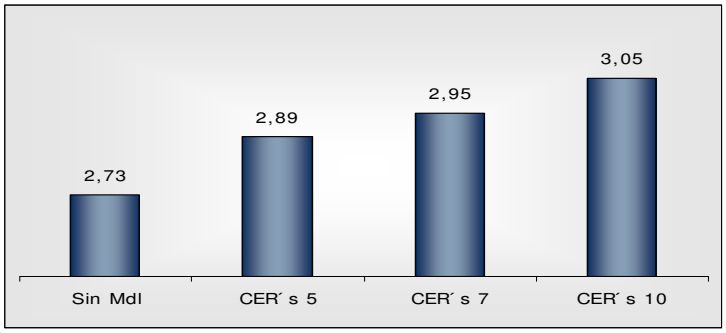
Fuente: HIDROZAMORA C.E.M. Flujo Neto de Caja. 2007.

De los resultados obtenidos se puede observar que en todos los casos el proyecto es rentable desde el punto de vista financiero. Además al comparar las TIR del proyecto sin la venta de CER's con aquellas obtenidas considerando la venta, se observa que en el primer caso la TIR es menor, demostrando una mayor rentabilidad al ser un proyecto MDL.

✓ **Relación Costo – Beneficio**

La relación Costo – Beneficio en todos los casos es atractiva (gráfica 14) para los inversionistas de HIDROZAMORA S.A, por cada dólar invertido se obtendrá una beneficio de USD 1,73 sin la venta de CER's y las expectativas siguen mejorando al vender los certificados a precios mayores hasta obtener USD 2,05 de beneficio.

Gráfico 12. Análisis de sensibilidad: Relación Costo Beneficio



Fuente: HIDROZAMORA C.E.M. Flujo Neto de Caja. 2007.

3.4.2 Beneficios Socio – Económicos

La falta de proyectos de generación eléctrica al Sistema Nacional Interconectado (SNI), provoca un déficit en el mercado de energía. En el 2001, se encontró que en la provincia de Zamora Chinchipe de 16.191 viviendas sólo cuentan con electrificación 11.877 viviendas, existiendo una cobertura energética de 73,67%, siendo una de las provincias con mayor escasez energética, después de Orellana, Morona Santiago, Napo y Sucumbios; cabe recalcar que está tendencia entre otras causas se debe a que en la Región Amazónica se dificulta el acceso de servicios básicos, por cuestiones de cobertura por infraestructura y escasa inversión. El déficit se refleja en caídas de tensión y cortes de suministro frecuentes en la ciudad de Zamora; por lo cual el Ilustre Municipio de está provincia realizó las gestiones necesarias para iniciar con la fase de preinversión del proyecto Hidroeléctrico "Chorrillos".

La construcción del proyecto persigue resolver los problemas de demanda de energía eléctrica en la zona y de tener como objetivo la generación de energía limpia para que la venta de energía futura del proyecto permita la ejecución de obras de infraestructura y proyectos productivos que colaboren con el desarrollo sostenible de la región.

Se debe resaltar que el proyecto conlleva únicamente impactos socio-económicos positivos debido a que sus objetivos deben estar acordes a los de desarrollo sustentable de la región para que pueda ser considerado como MDL.

- Estimulará la economía local, a través de la creación de fuentes de trabajo; durante el tiempo de construcción del proyecto, aproximadamente 150 personas en mano de obra no calificada y 7 técnicos, mientras que en la etapa operativa contara con 16 personas permanentes; contribuyendo a disminuir el nivel de desocupación.
- Aporte para resolver el problema eléctrico del país, ya que generará 3.96 MW, lo que permitirá reducir el déficit ante la demanda de energía eléctrica nacional. La provincia de Zamora Chinchipe, dejaría de ser un punto receptor de energía para convertirse en un punto generador de energía eléctrica.
- Capacitación de los trabajadores involucrados en actividades de instalación, operación y mantenimiento eléctrico, lo cual contribuye a mejorar sus conocimientos, habilidades y destrezas técnicas para el mejor desempeño de sus funciones y mantener condiciones más seguras de trabajo, minimizándose así el riesgo de accidentes laborales.

- La generación de recursos económicos provenientes de la venta de excedentes de energía permitirá mejorar el ingreso económico de los trabajadores al verse beneficiados con la participación en las utilidades empresariales, tal como lo contempla la legislación laboral vigente en Ecuador.
- Posibilidad de que las empresas asentadas en la zona desarrollen con normalidad sus actividades productivas en caso de presentarse daños o problemas operacionales que reduzcan la capacidad de generación de la empresa que suministra energía al sector. Esto permitirá mantener la productividad y la competitividad de las industrias locales.
- Los ingresos derivados del proyecto se los destinará para Programas de Educación y Salud y reinversión en este tipo de Proyectos en la provincia de Zamora.
- La reactivación económica del comercio en la zona a través de proveedores de materiales de construcción, maquinaria, herramientas, alquiler de equipos y vehículos, provisión de alimentos, etc.
- Evitará racionamientos en los períodos de estiaje, que frecuentemente ocurren entre octubre – marzo de cada año, que se calcula impactan en la economía.

3.4.3 Impactos Ambientales

La Hidroeléctrica Chorrillos durante la etapa de construcción y operación del proyecto causará alteraciones en el ecosistema, para lo cual HIDROZAMORA C.E.M. ha realizado estudios de impacto ambiental para determinar la magnitud de las mismas.

3.4.3.1 Impactos positivos al ambiente

Los impactos positivos del proyecto están concentrados principalmente en la fase operativa del proyecto, y se han catalogado como de alta significancia, ya que están relacionados con los siguientes beneficios:

- El principal impacto es la Producción de energía *limpia*, que generará adicionalidad ambiental de 15.529,26 ton de CO_2 anuales durante los 50 años de vida útil del proyecto, las cuales se emitirían si se creará un Central Termoeléctrica, la cual utilizaría combustible fósil para su funcionamiento. Al no realizarse el proyecto no hubiera existido reducción de emisiones de Gases de Efecto Invernadero.
- Al considerar el tamaño de la central hidroeléctrica, la incidencia en el medio ambiente por la generación de metano es nula, lo que no ocurre en las grandes hidroeléctricas.

3.4.3.2 Impactos negativos al ambiente

Los impactos negativos[16] más importantes se prevé ocurrirán en la fase de constructiva y operación del proyecto, y están relacionados con aquellos que se generan en las actividades de excavaciones y movimientos de tierras, que conllevan afectaciones ambientales, tales como:

- Generación de emisiones de particulado sólido (polvo), gases y ruido por la operación de equipos y maquinaria pesada y tránsito vehicular.
- Formación de áreas inestables, en el área de los túneles de excavación subterránea.
- Afectación a las geoformas naturales, pérdida de suelos, vegetación y hábitats naturales.
- Posible contaminación de suelos y aguas debido a la inadecuada disposición de los excedentes de excavaciones.
- Afectación a la belleza escénica.
- Pérdida de vegetación natural y hábitats naturales y su fauna silvestre asociada.

[16] Sin embargo para corregir los impactos ambientales negativos con la implementación y operación del proyecto HIDROZAMORA C.E.M. ha desarrollado un Plan de Manejo Ambiental.

4. CONCLUSIONES

Para la puesta en marcha de proyectos con el componente MDL, es preciso señalar algunos factores críticos como los beneficios económicos y sociales, la contribución a la reducción de emisiones de CO_2, y las barreras que existen dentro de éste mecanismo

El poder considerar proyectos basados en fuentes de energías renovables como proyectos MDL puede impulsar el desarrollo de los mismos al convertir dichos proyectos en rentables, a través de los ingresos adicionales obtenidos por los certificados de CO_2 asociados. Esto significa que cada proyecto de energía renovable tiene un efecto positivo en el ambiente, tanto a nivel local como global.

Los costos de transacción para ser considerado un proyecto como MDL son elevados, por lo que es importante contratar los servicios de un broker para que éste se encargue de todo el proceso y a su vez asuma éstos costos para obtener mejores beneficios.

El precio del CER´s impacta directamente sobre los resultados de la financiación con MDL. Es crítico planear el momento y las condiciones de mercadeo y venta del proyecto MDL, para así maximizar el precio. Cuando el proyecto se encuentra en etapa de operación los CER´s tienen un mayor valor, sin embargo la compra venta se la puede realizar en etapas previas para financiar costos de construcción, a lo cual un proyecto convencional no puede acceder, lo que podría ser un obstáculo para que el proyecto se lleve a cabo.

En el proyecto "Chorrillos" desde el primer año de operación ya se obtienen ingresos por la venta anticipada de CER´s, anualmente se venderán 15.529,26 ton de CO_2 anuales durante los 50 años de vida útil del proyecto, obteniéndose ingresos adicionales de USD 108.704,82 al año; parte de los ingresos obtenidos por la venta de CER´s serán reinvertidos en el componente socio-económico de la comunidad.

La venta de CER´s constituye un flujo adicional de recursos económicos al Proyecto, sin embargo la rentabilidad se encuentra condicionada por el precio de los CER´s. Como se muestra en la tabla a continuación se puede observar que los indicadores financieros al incorporar la venta de certificados de carbono a un precio de 7 USD mejoran.

INDICADOR	Valor sin MDL	Valor con CER´s (7 USD)
VAN	3.298.760	4.079.496
TIR	24.58%	29.62%
C/B	2.73%	2.95%

Como principal impacto ambiental positivo la Hidroeléctrica Chorrillos contribuirá a la reducción de los efectos del cambio climático; dentro del componente socio económico: la generación de empleo y el suministro de energía serán los principales impactos positivos. La ejecución del proyecto no comprometerá zonas de valor arqueológico ni áreas naturales protegidas.

En Ecuador el proceso para obtener los permisos y licencias legales y administrativas para operar plantas hidroeléctricas toma más tiempo que los permisos necesarios para operar proyectos en plantas de energía térmica, al igual que el proceso del Mecanismo de Desarrollo Limpio el cual es largo y complicado resultando el gran obstáculo para que se den éste tipo de proyectos.

De acuerdo con las características particulares del sector eléctrico nacional y las barreras descritas, en ausencia de la actividad de proyecto, la generación eléctrica del Proyecto Chorrillos se habría producido usando generadores conectados al SNI, principalmente a través de unidades térmicas que generan emisiones por la quema de combustibles fósiles.

5. REFERENCIAS BIBLIOGRÁFICAS

BANCO MUNDIAL. 2007. "Little Green Data Book"

BAUER, Mariano, entre otros. 2005. "Metodologías para la Implementación de los Mecanismos Flexibles de Kyoto – Mecanismo de Desarrollo Limpio en Latinoamérica". OLADE.

BROWN, Lester. 2004. Salvar el Planeta. Plan B; Ecología para un mundo en Peligro. Barcelona, España. Paidos.

BUTZE A, Walter. 2006. "Permisos de contaminación negociables: un instrumento de mercado para la regulación ambiental". Universidad Autónoma Metropolitana – Azcapotzalco. México.

CAMINOSCA. MEMORIA EJECUTIVA. 2003. "Estudios y Diseño Definitivo de la Central Hidroeléctrica Chorrillos".

CAPPOR, Karen. 2007. State and Trends of the Carbon Market 2007". Banco Mundial.

CASTRO, Carlos. 2004."Revisión de arreglos institucionales y portafolio indicativo". CORDELIM. Ecuador.

CEPAL. 2006. "Mercado de energías renovables y mercado del carbono en América Latina: Estado y perspectivas".

CÓMITE NACIONAL SOBRE EL CLIMA. 2001. "Bases para la Comunicación Nacional"

CORDELIM. 2005. "Procedimiento de la Autoridad Nacional para la Emisión de Carta de Aprobación a Proyectos de Pequeña Escala del Mecanismo de Desarrollo Limpio (MDL)". Ecuador.

CORDELIM.2007. Proyectos MDL en Ecuador, Portafolio Indicativo. Ecuador

COTO Oscar y Morera Liliana. 2006. "El MDL en América Latina y el Caribe. Lecciones Aprendidas – El Salvador". OLADE

EGUREN, Lorenzo. 2007. "Mercado de energías renovables y mercado del carbono en América Latina: Estado de situación y perspectivas". CEPAL. Chile.

ESTUPIÑAN Felix, entre otros. (___). "Mercado de Emisiones de CO2 en el sector eléctrico". España.

HIDROZAMORA C.E.M. 2007. "Informe 2006 Gerencia General". Ecuador.

Instituto de Estudios Económicos de la Bolsa de Cereales. 2005. "Situación y Tendencias del Mercado del Carbono en 2005".

JACOBO, Alejandro.2001. "El Medio Ambiente como Propiedad. Los permisos transables". Universidad Nacional de Córdova.

MARTINEZ Julia, entre otros. 2004. "Cambio Climático: una visión desde México". México.

MGM. 2007. "Oportunidades de negocios en el marco del Protocolo de Kyoto para las empresas chilenas".

Ministerio de Industria, Turismo y Comercio. 2003. "Energía y Cambio Climático". España

NEEFF, Till, entre otros. 2007. "Guía sobre los Mercados y la Comercialización de Proyectos MDL Forestales". Centro Agronómico Tropical de Investigación y Enseñanza, CATIE. Costa Rica.

OYHANTCABAL, Walter. 2006. "Servicios ambientales: situación y perspectivas del mercado de carbono". Uruguay.

PROCLIM. 2004. El Mecanismo de Desarrollo Limpio: "Guía Práctica para Desarrolladores de Proyectos".

RIOS, Alvaro. 2007. "Eficiencia Energética, Cambio Climático y Desarrollo Social". Organización Latinoamericana de Energía (OLADE).

SANCHEZ, Ana, entre otros. 2005 "Cambio climático. Efectos sobre los sectores productivos (energéticos, industrial y otros"). Instituto Nacional de Trabajo, medio ambiente y salud (ISTAS). España.

SEOÁNEZ, Mariano. 2002 "Tratado de la Contaminación Atmosférica; Problemas, tratamiento y gestión". España.

STERN, Nicholas. 2007. "La verdad del cambio climático". Barcelona, España.

UNED Risoe MDL / JI. 2008. "Pipeline y Análisis de la base de datos". actualiazación enero 8 de 2008

Universidad Nacional de Educación a Distancia (UNED). 2006. "Energía y Desarrollo Sostenible". Publicación. España.

IMPACTOS DA INICIATIVA RENOVÁVEIS NA HORA NA MICROGERAÇÃO SOLAR FOTOVOLTAICA

KATHRIN CALHAU
E-mail: kathrin.c@inteli.pt
INTELI – Inteligência em Inovação
LUIS RORIZ
E-mail: luis.roriz@ist.utl.pt
Instituto Superior Técnico

Resumo

O presente artigo faz uma análise da legislação aplicável às centrais de microgeração - Decreto-Lei n.º 363/2007 - com vista à determinação das perspectivas de evolução da microgeração fotovoltaica em Portugal, e identificação de factores críticos para o seu sucesso.
Conclui-se que dificilmente será possível atingir as metas propostas pelo governo, designadamente um crescimento da potência instalada a nível nacional de 20% e a instalação de165 MW até 2015.
Os resultados obtidos revelam ainda que o sucesso do incentivo proposto depende do equilíbrio entre a evolução da potência instalada a nível nacional e da evolução dos custos dos equipamentos no mercado internacional (factor exógeno) e da evolução do custo da instalação e elaboração de projecto no mercado nacional (factor endógeno).
Assim, sugere-se um acompanhamento próximo e a implementação de medidas que possam limitar o número de licenças atribuídas em determinados anos, ou promover uma taxa de aprendizagem mais elevada dos agentes nacionais.

Palavras-chave: energias renováveis; energia fotovoltaica; microgeração; taxa de aprendizagem.
Área Temática: Economia da Energia, do Ambiente, dos Transportes e dos Recursos Naturais.

Abstract

The present paper carries out an analysis of the legal framework applicable to microgeneration photovoltaic systems – Decree-Law n.o 363/2007 – in order to determine the future perspectives concerning photovoltaic solar cells technology and determine key factor for its success.
The central conclusion is that this framework isn't likely to achieve the goals set out by the government, such as a 20% annual growth of installed capacity, achieving 165 MWp in 2015.
Additionally, the results reveal that the success of the recently implemented framework depends on the equilibrium between the growth of the installed capacity, the systems price evolution of equipments in the international markets (external factor), and the price evolution of the installation of systems, such as the labour work and the project elaboration that depend on the national market (internal factor).
Therefore, a close surveillance is suggested as well as the implementation of instruments that might limit the number of systems installed in some years, or that promote a high learning rate regarding national players.

Key-words: renewable energy; photovoltaic energy; microgeneration; learning rate.
Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

1. INTRODUÇÃO

A definição de produção de electricidade em regime especial (PRE) consiste na produção de electricidade com incentivos à utilização de recursos endógenos e renováveis. Actualmente o enquadramento de PRE contempla três vertentes específicas: cogeração; produção de energia eléctrica a partir de recursos renováveis, combustíveis nacionais, e resíduos industriais, agrícolas ou urbanos e; produção de energia eléctrica de muito pequena dimensão – com potência inferior ou igual a 5,75 kW – vulgarmente designada de microgeração.

A vertente da microgeração é regulada pelo Decreto-Lei n.º 363/2007, lançado em Novembro de 2007 com a designação de *Renováveis na Hora*, e que tem por objectivo promover a instalação de sistemas de produção de energia de pequena escala através de fontes renováveis, incluindo a solar, biomassa, bombas geotérmicas e mini-eólicas, assim como a micro-cogeração. Neste sentido, para além de um processo de licenciamento simplificado, o referido decreto-lei estabeleceu um regime de remuneração bonificado aplicável às instalações com potência de ligação até 3,68 kW, que utilizem energias renováveis como fonte de energia primária, e que contemplem uma área mínima de 2 m^2 de painéis solares térmicos para aquecimento de água.

No lançamento desta iniciativa, a meta proposta pelo Ministério da Economia e Inovação (MEI) era atingir cerca de 165 MW (MEI, 2007) em sistemas de microgeração até 2015, o que pressupunha uma instalação no primeiro ano de 10 MW e um crescimento de 20% da potência nova instalada em anos subsequentes. Subjacente a esta meta, o objectivo segundo o MEI seria induzir um aumento da procura, que por sua vez estimulasse a criação e crescimento de novas empresas, o aumento de concorrência e uma consequente redução de custos. Como tal, o enquadramento prevê uma redução da tarifa do regime bonificado em função da potência instalada, de forma a forçar a descida de preços de mercado e evitar uma distorção do mercado.

Desde que a iniciativa entrou em vigor[1], tem-se verificado uma procura intensa, o que levou à definição de datas para a realização de registos e limites de potência a atribuir em cada data. Os últimos dados publicados no portal *Renováveis na Hora*[2] mostram que desde a entrada em vigor da iniciativa, aproximadamente 7 338 registos foram realizados, traduzindo-se numa potência superior a 25 000 kW. Contudo, verifica-se que nem todos os registos realizados se concretizaram. Até à data, foram instalados 2 132 sistemas, perfazendo 7 454 kW de potência instalada, o que constitui um valor expressivo, sobretudo se comparado com valores de potência instalada em microgeração no âmbito enquadramentos jurídicos anteriores. Verifica-se que mais de 90% dos sistemas instalados no âmbito do Decreto-Lei n.º 363/2007 são sistemas solares fotovoltaicos, pelo que as análises aqui efectuadas se concentram apenas nesta tecnologia.

Neste contexto, o objectivo do estudo realizado consiste na determinação dos efeitos a médio prazo que a aplicação do Decreto-Lei n.º 363/2007 poderá ter, designadamente no que respeita à potência instalada a nível nacional. Para tal, recorre-se a uma metodologia de análise de investimento de sistemas de microgeração fotovoltaicos, considerando a evolução de preços dos sistemas fotovoltaicos de acordo com uma curva de aprendizagem.

2. METODOLOGIA

A metodologia usada propõe uma análise previsional do investimento associado a um sistema de microgeração fotovoltaica, ao abrigo do Decreto-Lei n.º 363/2007, e à luz de diferentes cenários de desenvolvimento da microgeração em Portugal, no que se refere à tarifa aplicável e ao preço dos sistemas solares fotovoltaicos.

[1] 2 de Abril de 2008

[2] 17 de Abril de 2009

Optou-se por tomar como critério de avaliação do investimento o período de recuperação, procurando medir o tempo necessário para a recuperação do investimento por via do *cash-flow* de exploração acumulado. Para a escolha deste critério, pesou o facto de ser um critério de fácil leitura e por ser, de uma forma geral, o critério mais relevante (para além do valor do investimento) na tomada de decisão de um potencial microprodutor[1].

O *cash-flow* de exploração resulta da conta previsional de exploração e representa a diferença entre as receitas e as despesas da exploração. A consideração do factor tempo impõe a actualização dos *cash-flows* anuais, a uma determinada taxa (correspondente à taxa de remuneração de uma alternativa equivalente de igual risco) e a sua comparação com o montante investido, para determinar se a recuperação do investimento é garantida. A actualização dos *cash-flows* anuais para o início da vida útil do projecto de investimento (definido como ano 0) faz-se através do factor de desconto $(1+r)^{-k}$, em que r é a taxa adoptada para a actualização e k corresponde ao ano (ordenado por ordem crescente desde o ano 0) a que respeita o *cash-flow*.

Assim, o valor actualizado do *cash-flow* de um certo ano k é dado por:

$$CF_0 = CF_k \times (1+r)^{-k} \qquad \textit{Eq. 1}$$

A consideração da inflação requer cálculos previsionais de proveitos e custos a preços correntes, admitindo para o efeito uma determinada taxa de inflação para as várias componentes dos custos e proveitos.

De seguida apresentam-se os vários pressupostos adoptados para o cálculo do período de retorno.

2.1. Caracterização do sistema

O sistema considerado, com vista ao acesso ao regime bonificado previsto no Decreto-Lei n.º 363/2007, cumpre um conjunto de requisitos:

- Dispor de colectores solares térmicos para aquecimento de água com um mínimo de 2 m² de área de colector, excepto para condomínios;
- Os condomínios deverão realizar auditoria energética ao edifício;
- A potência de ligação deverá ser igual ou inferior a 3,68 kW e a metade da potência contratada para consumo.

Assim, tomando por hipótese um produtor individual e doméstico, cuja potência contratada seja 6,9 kW[2], a potência de ligação máxima para o sistema solar fotovoltaico será de 3,45 kW. Sendo obrigatório que o valor da potência nominal do inversor seja inferior à potência de ligação registada, assume-se um inversor com potência nominal de 3,4 kW. Em termos de potência instalada de módulos fotovoltaicos, assume-se um total de 4,0 kWp, o que equivale aproximadamente a uma potência de 3,6 kW em condições de operação, face às perdas inerentes ao sistema. Quanto à tecnologia, considera-se a utilização de módulos de silício cristalino.

Relativamente ao sistema solar térmico, assume-se um sistema de circulação forçada, e parte-se do pressuposto que está conforme o Regulamento dos Sistemas Energéticos de Climatização em Edifícios (RSECE), no que refere ao isolamento do depósito.

Assim, a simulação efectuada, baseia-se nos sistemas com as características apresentadas na Tabela 1. Assumiu-se ainda uma vida útil para o conjunto dos sistemas de 15 anos.

[1] O critério do período de recuperação apresenta, contudo, algumas limitações. Uma das mais relevantes é o facto de desprezar o aspecto rendibilidade, pois assim é tão justificável um investimento que se recupere num certo prazo e só possa produzir efectivamente rendimentos por mais um dia, como outro investimento que se recupere no mesmo prazo mas possa produzir rendimento por muitos anos mais.

[2] Em rigor, a potência contratada é expressa em kVA. Porém, para os efeitos considerados neste capítulo considera-se 1 kVA = 1 kW.

Tabela 1 – Caracterização do sistema de microgeração

Sistema Solar Fotovoltaico		
Colocação módulos	Inclinação: 30º Horizonte: Livre	Azimute: 0º (Sul) Sombreamento: Nenhum
Módulos PV	Silício policristalino N.º módulos: 32 (16 em série / 2 *strings*) Potência unitária: 125 Wp Área total: 36 m^2	Potência total: 4,0 kWp
Inversor	Potência Nominal: 3,4 kW	
Sistema Solar Térmico		
Colocação colector	Inclinação: 45º Horizonte: Livre	Azimute: 0º (Sul) Sombreamento: Nenhum
Colector	Plano N.º módulos: 1 Área total: 2,3 m^2	
Sistema auxiliar	Tipo: Esquentador a gás propano	Eficiência: 55%
Bomba de circulação	Potência: 150 W (10 horas/dia)	

Para o cálculo da produção de energia dos sistemas solar fotovoltaico e solar térmico respectivamente, utilizaram-se as ferramentas PVSYST 4.33[1] (desenvolvida pela Universidade de Genebra, para o dimensionamento, simulação e análise de sistemas fotovoltaicos) e F-Chart[2] (criada por S.A. Klein e W.A. Beckman, para a simulação e análise de sistemas solares activos e passivos) respectivamente. No caso do sistema solar fotovoltaico obteve-se o valor de 5 668 kWh/ano. Para o sistema solar térmico, estabeleceu-se o valor de 3 621 kWh/ano em termos de necessidades totais, 63% das quais colmatadas por via dos colectores solares, equivalendo a um total de 2 288 kWh/ano, e que será valorizado ao preço do gás propano.

Refira-se ainda que o conjunto dos sistemas cumpre o disposto no Regulamento das Características de Comportamento Térmico dos Edifícios (RCCTE), relativamente ao dimensionamento de sistemas solares térmicos. De acordo com este regulamento, os sistemas solares térmicos devem ser dimensionados na base de 1 m^2 de colector por ocupante. Porém, em alternativa à utilização de colectores solares térmicos, podem ser utilizadas outras tecnologias de produção de energia com base em fontes renováveis, que captem anualmente uma quantidade de energia equivalente à dos colectores solares. Ora, atendendo ao facto que são já considerados 2 m^2 de colectores solares, o conjunto dos sistemas solar térmico e solar fotovoltaico, conforme se calculou, equivalem à produção aproximada de um sistema solar térmico de área superior 7 m^2, ou seja, um sistema adequado a 7 ocupantes (de acordo com o RCCTE).

2.2. Tarifa de venda de electricidade aplicável à microgeração

Conforme mencionado será considerado o regime de remuneração bonificado, segundo o qual a tarifa aplicável toma o valor de 0,65 €/kWh para os primeiros 10 MW instalados a nível nacional. Por cada 10 MW instalados adicionalmente, o valor da tarifa diminui 5%.

Porém, para uma determinada instalação, esta tarifa vigora apenas nos primeiros 5 anos de funcionamento da mesma, após os quais, aplica-se a tarifa em vigor para os sistemas de microgeração que entrem em funcionamento no dia 1 de Janeiro do ano em causa.

[1] http://www.pvsyst.com/
[2] http://www.fchart.com/

Este facto introduz, naturalmente, uma incerteza na análise económica de um sistema solar fotovoltaico ao abrigo do Decreto-Lei n.º 363/2007, já que não é possível saber com exactidão qual o valor da tarifa após os primeiros cinco anos.

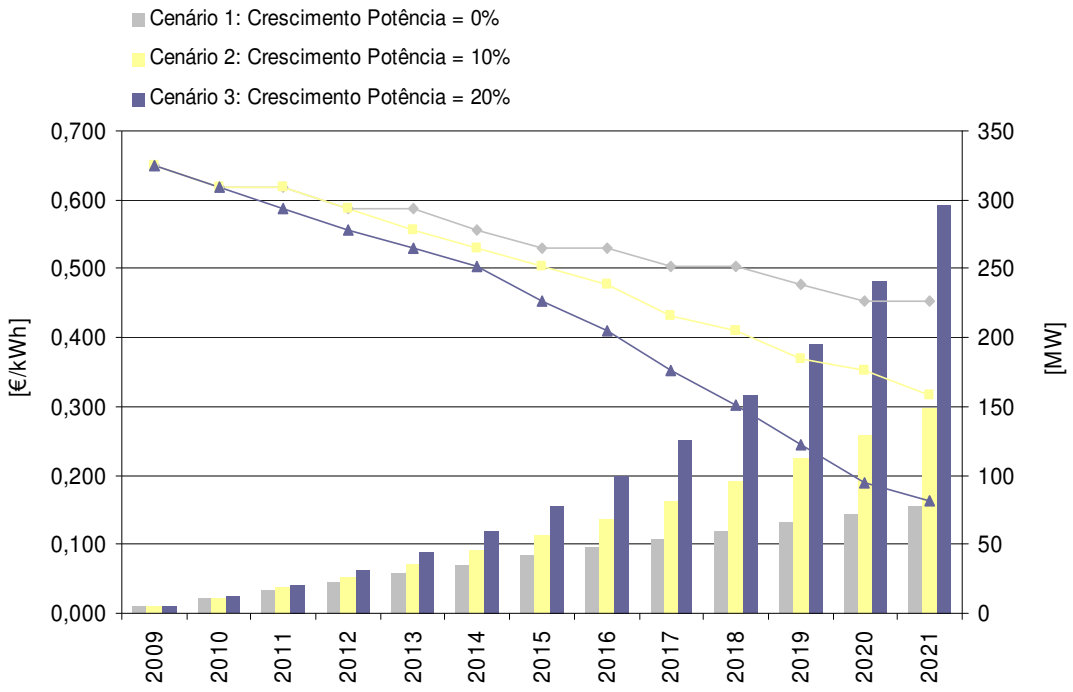
Deste modo, são calculados resultados para diferentes cenários, relativos a diferentes taxas de instalação de potência nacional de sistemas de microgeração, tendo como referência a potência instalada entre início de Abril e final de Dezembro de 2008 (4 600 kW):

- Cenário 1 – taxa de crescimento anual nula (isto é, todos os anos é instalada a mesma potência em sistemas de microgeração);
- Cenário 2 – taxa de crescimento anual de 10%; e
- Cenário 3 – taxa de crescimento anual de 20%.

Em qualquer dos casos, considera-se como limite inferior da tarifa de venda, o regime de remuneração geral, igual ao tarifário aplicável pelo comercializador de último recurso (CUR) do fornecimento à instalação de consumo. Aqui, assumiu-se a tarifa simples de venda a clientes finais em baixa tensão (BT) para 2009 – 0,1211 €/kWh (ERSE, 2008).

O Gráfico 1 mostra a evolução das tarifas (representadas pelas linhas, com leitura no eixo das ordenadas principal – à esquerda) segundo o Decreto-Lei n.º 363/2007 a aplicar em instalações novas, para os três cenários diferentes considerados, bem como o crescimento da potência acumulada até 2021 (representada pelas colunas e com leituras no eixo das ordenadas secundário – á direita).

Gráfico 1 – Evolução da potência instalada e tarifas aplicáveis



Constata-se que para o Cenário 1, obtém-se em 2021 uma potência global instalada de 78 MWp, o que corresponde a uma tarifa a aplicar em instalações novas de 0,454 €/kWh. O Cenário 2 leva a um nível de potência de 149 MWp, e a uma tarifa de 0,317 €/kWh. Por fim, à luz do Cenário 3, a potência acumulada atinge o valor global de 296 MWp e uma tarifa de 0,163 €/kWh.

2.3. Preço dos sistemas fotovoltaicos e respectiva evolução futura

De acordo com o "National Survey Report of PV Power Applications in Portugal", (IEA PVPS, 2005), o preço chave-na-mão para um sistema interligado à rede eléctrica com uma potência até 10 kW em Portugal varia entre 6 e 8,5 €/Wp. Já em 2008, no "Trends in Photovoltaic Applications – Survey report of selected IEA countries between 1992 and 2007" (IEA PVPS, 2008), os valores reportados para Portugal, para 2007, foram de 5 a 6 €/Wp para sistemas ligados à rede com potência inferior a 10 kW.

Com vista à confirmação dos dados reportados, fez-se uma consulta de mercado para um sistema fotovoltaico com potência 4 kWp – o sistema em causa na presente análise.

Uma vez que a análise de investimento é feita na óptica do utilizador final, inclui-se: o IVA à taxa reduzida (12%); a taxa de registo da instalação de microprodução estabelecido pela Portaria n.º 201/2008, que é de 256,3 € acrescida do IVA à taxa reduzida (12%). Por outro lado, considerou-se o valor máximo de dedução à colecta resultante dos encargos com a aquisição de equipamentos para aproveitamento de energias renováveis, estabelecido pela Proposta de Lei n.º 226/X – 796 €.

Considerando todos os valores mencionados, do ponto de vista do utilizador, o preço do investimento chave-na-mão do sistema fotovoltaico ronda os 6,15 €/Wp.

A evolução do preço de venda dos sistemas fotovoltaicos depende, em grande medida, da evolução dos custos, sendo de uma forma geral 20 a 40% mais elevado do que o custo dos sistemas fotovoltaicos, uma vez que os preços incluem as margens de lucro dos fornecedores, projectistas e instaladores. Aqui considera-se que a evolução dos preços é paralela à evolução dos custos, pelo que nesta secção apenas se faz referência aos custos.

O custo de um sistema fotovoltaico é composto pelo custo dos módulos fotovoltaicos[1] e o custo de outros elementos, como sejam o inversor, a instalação eléctrica, a estrutura de suporte e a integração do sistema, e a instalação propriamente dita, que inclui a mão-de-obra e a execução do projecto. Ao conjunto destes componentes, dos quais se excluem os módulos, é dada a designação de *Balance of System* (BOS).

A redução dos custos dos sistemas fotovoltaicos depende, portanto, da evolução da redução do custo dos módulos e do custo do BOS. Uma das formas mais comuns de avaliar a redução dos custos é por via da curva de aprendizagem, que tem origem na constatação de que os operários em instalações fabris se tornam mais eficientes à medida que produzem mais unidades (Nemet, 2006). A experiência é expressa utilizando produção acumulada, e a alteração do custo representa a medida de aprendizagem. De acordo com esta metodologia, a experiência adquirida com uma determinada tecnologia, é expressa através de uma taxa de aprendizagem, isto é, a percentagem à qual o custo por unidade diminui de cada vez que se duplica a produção acumulada (van der Zwaan, 2003):

$$Custo_t = Custo_0 \left(\frac{q_t}{q_0} \right)^{-b} \qquad Eq.\ 2$$

em que:

- $Custo_t$ representa o custo unitário no momento t [€/W];
- $Custo_0$ é o custo unitário inicial, ou seja, no momento 0. Para efeitos deste estudo considera-se que o momento inicial é o primeiro dia de 2009;
- q_t é a capacidade acumulada instalada no momento t [kW];
- q_0 é a capacidade acumulada instalada no momento 0;
- b representa a elasticidade da aprendizagem, calculado da seguinte forma:

$$b = \frac{\ln t_p}{\ln 2} \qquad Eq.\ 3$$

sendo

[1] 50 a 60% para sistemas ligados à rede segundo Watanabe (2007) e 55% segundo o IEA PVPS (2007)

– t_p a taxa de progressão, definida por:

$$t_p = 1 - t_a \qquad \textit{Eq. 4}$$

– t_a a taxa de aprendizagem.

No que respeita aos módulos, importa referir que o seu custo depende do fornecedor, do tipo de tecnologia e da dimensão, mas sendo o mercado de módulos fotovoltaicos internacional, a diferença de preços entre países deverá ser nivelado gradualmente.

A evolução dos custos dos módulos tem sido amplamente estudada por via da curva de aprendizagem, obtendo-se, tipicamente, taxas de aprendizagem entre os 10 a 30%. Uma taxa de aprendizagem normalmente utilizada em várias tecnologias do sector da energia é de 20% (van der Zwaan, 2003). As taxas de aprendizagem são baseadas em dados históricos de tecnologias que emergiram, atingiram a fase de comercialização e que têm vindo a amadurecer. São vários os exemplos de tecnologias que não seguiram esta evolução e que se extinguiram precocemente (em prol de outras tecnologias concorrentes, por exemplo). Por conseguinte, não é garantido que as taxas de aprendizagem observadas até à data, em matéria de tecnologias fotovoltaicas, se mantenham na evolução futura das mesmas.

Porém, de acordo com a literatura, existe margem para uma redução de custos, mesmo para a tecnologia mais madura – silício cristalino – sendo vários os avanços tecnológicos que sugerem a manutenção da redução de custos de acordo com uma taxa de aprendizagem durante as próximas décadas. De acordo com Hoffman (2006), a curva de aprendizagem histórica relativa ao preço dos módulos no mercado mundial em função da produção acumulada de módulos demonstra uma diminuição de 20% de preço com a duplicação acumulada de produção. A mesma fonte refere que a taxa de aprendizagem no futuro poderá ser ligeiramente inferior:

Com o aumento progressivo das economias de escala, o custo da matéria-prima tem vindo a contribuir cada vez mais para o custo dos módulos. Com vista a contabilizar este factor, a taxa de aprendizagem deverá ser reduzida para 18%.

Optou-se, para efeitos dos cálculos efectuados, considerar este valor para a taxa de aprendizagem relativa aos módulos fotovoltaicos – 18%.

A curva de aprendizagem não contém directamente uma escala temporal, pelo que é necessário estabelecer uma taxa de crescimento de potência instalada, que permita estimar a redução dos custos temporalmente.

Para a presente simulação, considerando que o mercado dos módulos fotovoltaicos é internacional, presume-se uma taxa de crescimento (a nível mundial) de 25% para os próximos anos, conforme sugerido por Hoffman (2006). Para o custo inicial (2009) adopta-se o valor de 3,38 €/Wp (55% x 6,15 €/Wp), e para potência instalada acumulada inicial (também a nível mundial) o valor de 10 GW.

A taxa de aprendizagem relativa ao BOS, não terá sido ainda estudada de forma tão exaustiva como a taxa de aprendizagem dos módulos. Shum (2008) discute que o sistema BOS é constituído essencialmente por componentes eléctricos provenientes de mercados maduros, não dependentes do mercado solar fotovoltaico, e que de uma forma geral, a aprendizagem destes componentes está esgotada (com excepção dos casos em que ocorrem inovações tecnológicas). Shum afirma ainda que a aprendizagem associada ao BOS deve-se, sobretudo, à experiência acumulada da execução de projecto, integração e instalação do sistema fotovoltaico.

No entanto, a componente BOS está fortemente sujeita ao enquadramento local[1], existindo um espaço para a redução de custos resultante da evolução da curva de aprendizagem associada à instalação, designadamente a mão-de-obra e a elaboração do projecto.

Um estudo (Schaeffer, 2003), dedicado ao mercado alemão, considerado um mercado dinâmico, determinou, para o período de 1992 até 2000, que a curva de aprendizagem do BOS apresenta uma taxa próxima dos 22%.

[1] Existem diferenças substanciais nos standards impostos no que respeita a ligação à rede, bem como aos regulamentos aplicáveis à integração de sistemas fotovoltaicos em edifícios, em diferentes países.

Para a presente análise assumem-se (para cada um dos cenários relativos às taxas de crescimento da potência a nível nacional) três sub-cenários possíveis relativamente à taxa de aprendizagem da componente BOS:

- Cenário optimista, com uma taxa de aprendizagem de 20%;
- Cenário intermédio com uma taxa de aprendizagem de 10%; e
- Cenário pessimista com uma taxa de aprendizagem nula.

2.4. Preço do sistema solar térmico

Para o preço do sistema solar térmico, adopta-se o valor de 875 €/m^2, sugerido como valor mínimo para sistemas domésticos pela iniciativa pública AQSpP promovida pela DGEG.

2.5. Tarifa de gás propano

Com vista ao cálculo da redução da factura de energia térmica, resultante da energia fornecida pelos colectores solares térmicos, em detrimento do esquentador a gás propano cuja eficiência adoptada é de 55%, toma-se para o custo do gás propano o valor de 0,138 €/kWh[1].

2.6. Custos de exploração

Como custos de exploração, tem-se o custo de manutenção dos sistemas, para o qual se assume o valor de 200 €/ano[2].

Além disso, há a considerar o custo da energia eléctrica consumida pela bomba de circulação do sistema solar térmico. Assumindo que a bomba funciona em média 10 horas por dia, durante todo o ano, o consumo da mesma ronda os 550 kWh/ano. Toma-se, para o preço da electricidade, tarifário aplicável pelo CUR do fornecimento à instalação de consumo – 0,1211 €/kWh.

2.7. Taxa de inflação

A definição do valor para a taxa de inflação teve por base o objectivo primário do Banco Central Europeu (BCE) em matéria de política monetária – manter a estabilidade dos preços. No médio prazo a meta estabelecida para o conjunto da Zona Euro consiste em manter a inflação abaixo, mas próxima dos 2%.

No entanto, e atendendo a que o Nível de Preços é inferior, em Portugal, à média europeia, é expectável que a sua progressão se faça a um ritmo ligeiramente superior, como aliás, se tem verificado em anos anteriores[3]. Desta forma, o valor assumido para a presente análise será de 2,5%.

2.8. Taxa de actualização

Para a taxa de actualização, adopta-se o valor de 5%, tendo pesado para esta escolha a taxa de rendibilidade das Obrigações de Tesouro a 10 anos, publicados pelo Banco de Portugal[4].

---

[1] Consulta de mercado.

[2] Consulta de mercado.

[3] Índice Harmonizado de Preços no Consumidor (Portugal): 1996 – 2,3%; 1997 – 2,4%; 1998 – 2,8%; 1999 – 2,3%; 2000 – 2,9%; 2001 – 4,4%; 2002 – 3,6%; 2003 – 3,3%; 2004 – 2,4%; 2005 – 2,3%; 2006 – 3,1%; 2007 – 2,5% (Fonte: INE).

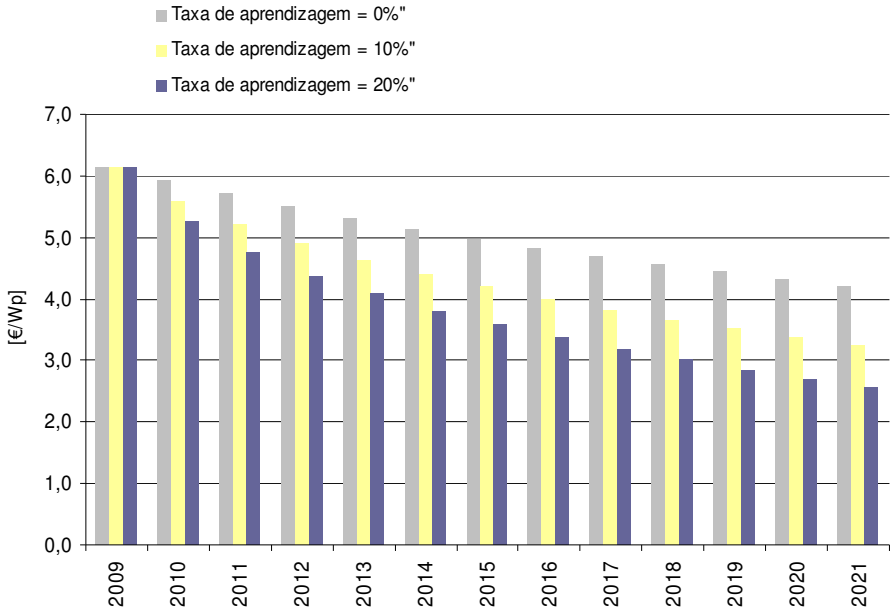
[4] Taxa de rendibilidade das OT a 10 anos (médias mensais de valores diários): Nov.2007 – 4,36; Dez.2007 – 4,47; Jan.2008 – 4,31; Fev.2008 – 4,27; Mar.2008 – 4,36; Abr.2008 – 4,52; Mai.2008 – 4,6; Jun.2008 – 4,96; Jul.2008 – 4,95; Ago.2008 – 4,69; Set.2008 – 4,66; Out.2008 – 4,56 (Fonte: Banco de Portugal).

3. RESULTADOS

3.1. Cenário 1 – Taxa de crescimento da microgeração fotovoltaica em Portugal nula

O Gráfico 2 apresenta a evolução do preço por cada kWp instalado considerando nulo o crescimento da potência instalada (que se traduz na instalação anual de 6 133 kWp) e diferentes taxas de aprendizagem da componente BOS.

Gráfico 2 – Cenário 1: Evolução do preço da instalação fotovoltaica



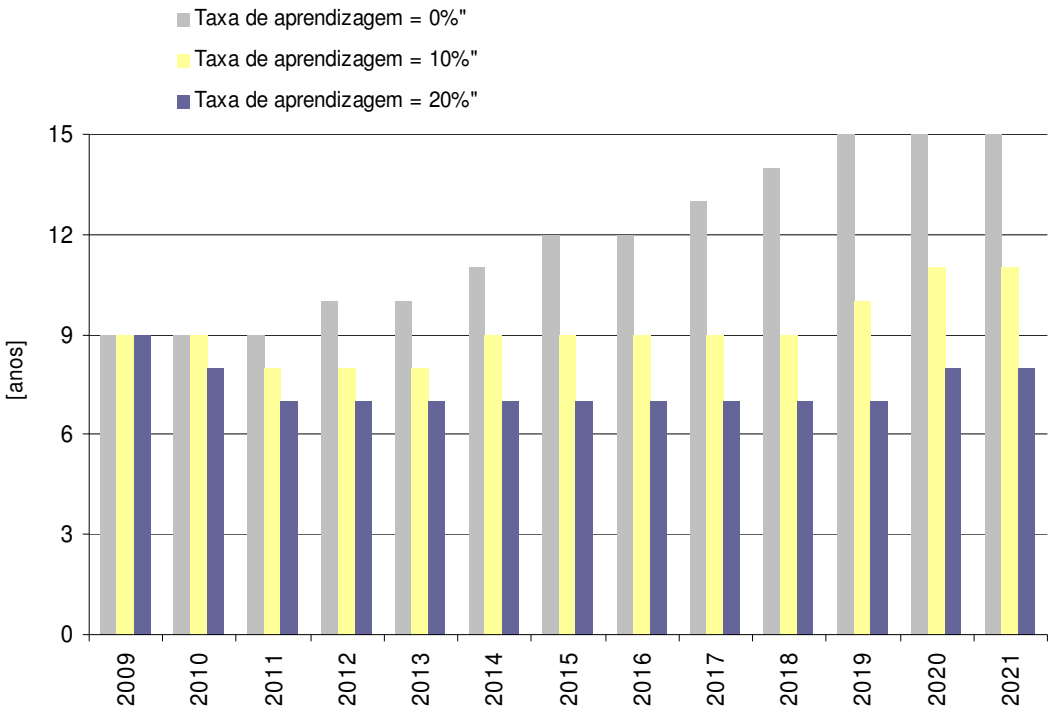
Assim, fixando a taxa de aprendizagem para os módulos solares fotovoltaicos de 18% e o preço inicial (1 de Janeiro de 2009) de 6,15 €/kWp, e adoptando uma taxa de aprendizagem da componente BOS de 0%, o preço da instalação será de 4,21 €/kWp em 2021.

Caso se verifique uma taxa de aprendizagem para a componente BOS de 10%, o preço em 2021, rondará os 3,24 €/kWp.

Por fim, uma taxa de aprendizagem de 20% leva a um preço próximo de 2,56 €/kWp em 2021.

O Gráfico 3 apresenta os períodos de retorno do investimento da instalação, considerando as diferentes taxas de aprendizagem.

Gráfico 3 – Cenário 1: Período de retorno do investimento[1]



Convencionando que para um período de retorno igual ou superior a 15 anos o investimento não é atractivo, conclui-se que no caso de se verificar uma taxa de aprendizagem de 10% ou 20% para a componente BOS dos sistemas fotovoltaicos, o investimento manter-se-á atractivo no longo prazo, embora, como observado no Gráfico 1, a potência instalada a nível nacional neste cenário fique muito aquém das expectativas do governo (inferior a 80 MWp). Uma taxa de aprendizagem nula poderá comprometer os efeitos da legislação actualmente em vigor, já que a partir de 2019 o investimento num sistema de microgeração fotovoltaica poderá deixar de ser atractivo.

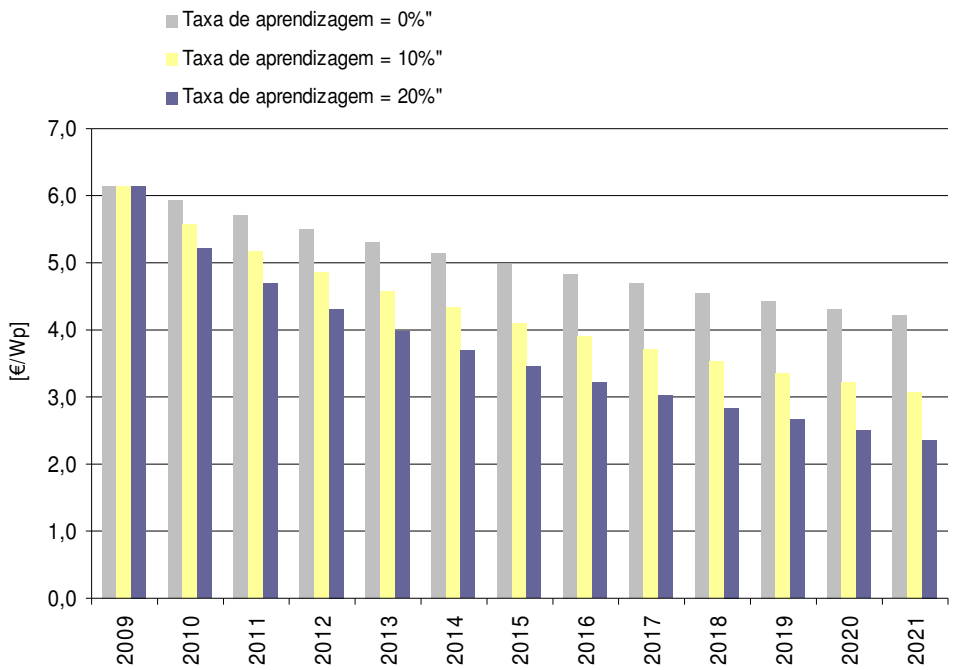
3.2. Cenário 2 – Taxa de crescimento da microgeração fotovoltaica em Portugal de 10%

Com uma perspectiva mais optimista, o segundo cenário considerado assume que a potência instalada em sistemas de microgeração fotovoltaica cresce 10% por ano, sendo que no primeiro ano se considera a instalação de 6 133 kWp, no segundo ano de 7 421 kWp, no terceiro de 8 163 kWp, e assim sucessivamente.

O Gráfico 4 apresenta a evolução do preço por cada kWp instalado, considerando as diferentes taxas de aprendizagem: 0%, 10% e 20%.

[1] Optou-se, na representação gráfica, limitar o eixo das ordenadas, pelo que um período de retorno de 15 anos traduz-se num período de retorno igual ou superior a 15 anos.

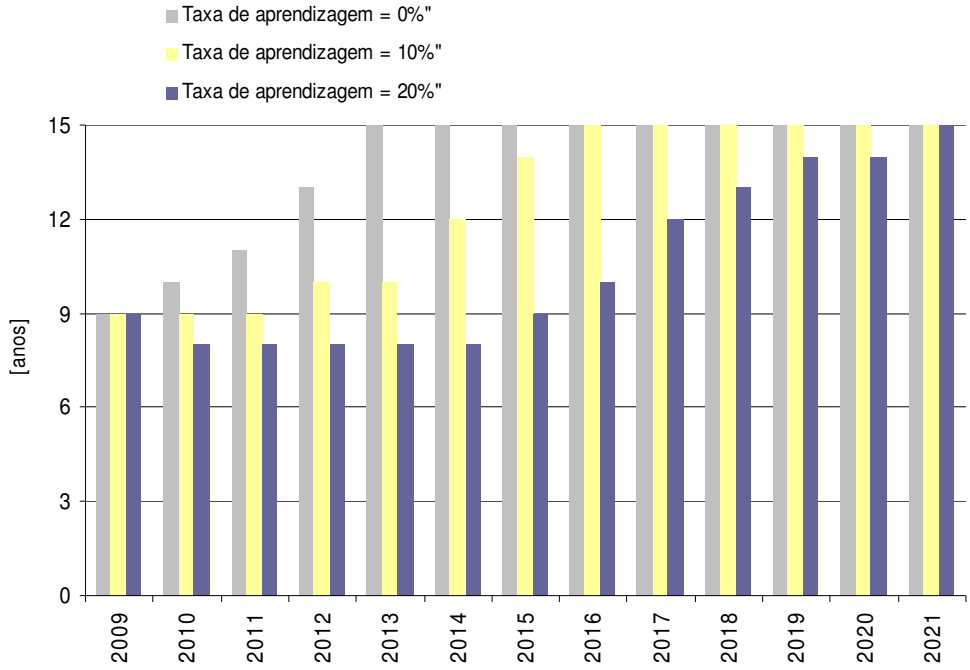
Gráfico 4 – Cenário 2: Evolução do preço da instalação fotovoltaica



Constata-se que, para a taxa de aprendizagem da componente BOS de 10%, o preço da instalação deverá rondar 4,21 €/kWp em 2021. Para uma taxa de aprendizagem para a componente BOS de 10%, o preço em 2021, deverá atingir 3,08 €/kWp. Considerando uma taxa de aprendizagem de 20%, o preço será próximo de 2,35 €/kWp em 2021.

O Gráfico 5 apresenta os períodos de retorno para uma instalação implementada nos diferentes anos.

Gráfico 5 – Cenário 2: Período de retorno do investimento



Embora se verifique uma redução de preço mais acelerada do que para o cenário anterior (Cenário 1) verifica-se que a redução da tarifa é igualmente mais acelerada, devido ao

crescimento mais rápido da potência instalada, levando a que mais rapidamente o período de retorno atinja valores iguais ou superiores a 15 anos. Caso se verifique uma taxa de aprendizagem nula, no que respeita à componente BOS, é possível que em 2013 o investimento deste tipo de instalações deixe de ser atractivo. Este cenário poderá levar à estagnação da implementação destes sistemas em 2013, o que, de acordo com o Gráfico 4, se traduziria numa capacidade instalada acumulada inferior a 40 MWp.

Por outro lado, no caso de uma taxa de aprendizagem elevada, o investimento num sistema fotovoltaico só atinge o período de retorno igual ou superior a 15 anos por volta de 2020, traduzindo-se aproximadamente a 130 MWp, um valor mais interessante, embora inferior à meta estabelecida pelo governo.

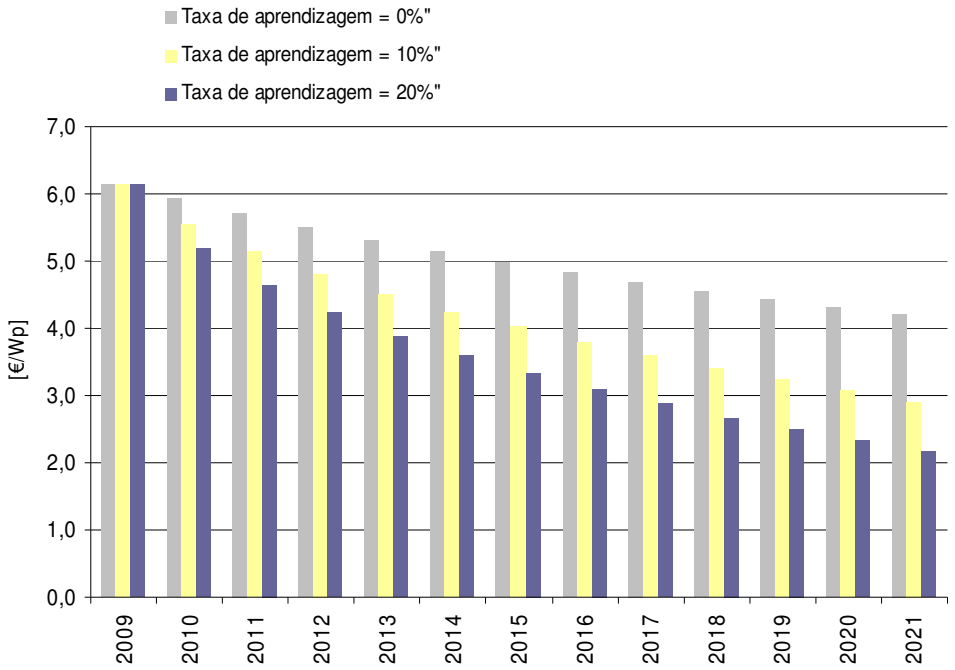
3.3. Cenário 3 – Taxa de crescimento da microgeração fotovoltaica em Portugal de 20%

Por fim, considera-se o terceiro cenário, mais ambicioso do ponto de vista do crescimento anual da potência instalada – 20%, que terá sido assumido pelo governo quando o Decreto-Lei n.º 363/2007 foi apresentado.

Partindo do valor de 6 133 kWp instalados no primeiro ano, no segundo ano considera-se a instalação de 7 360 kWp, no terceiro de 8 832 kWp, e assim sucessivamente.

O Gráfico 6 apresentado abaixo mostra a evolução do preço por cada kWp instalado três taxas de aprendizagem consideradas – 0%, 10% e 20%.

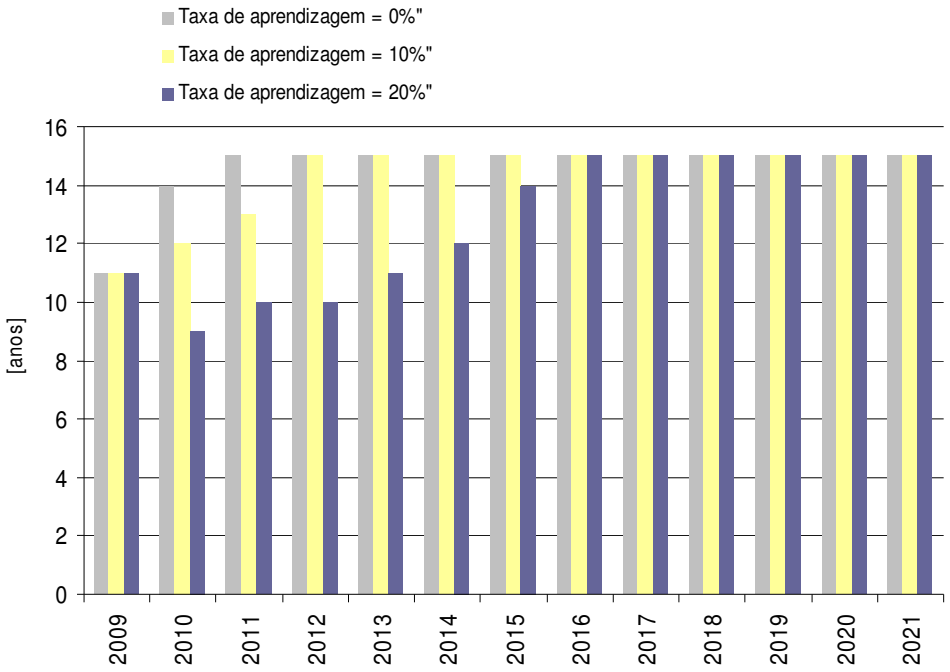
Gráfico 6 – Cenário 3: Evolução do preço da instalação



Verifica-se, para uma taxa de aprendizagem da componente BOS de 0%, um preço, em 2021, de 4,21 €/kWp. Uma taxa de aprendizagem de 10% leva a um preço de 2,92 €/kWp, e, para uma taxa de aprendizagem de 20%, um preço próximo de 2,17 €/kWp em 2021.

O Gráfico 7 apresenta o período de retorno do investimento de uma instalação nos diferentes anos considerados e para as diferentes taxas de aprendizagem.

Gráfico 7 – Cenário 3: Período de retorno do investimento



Do gráfico constata-se que um crescimento anual da potência instalada de 20%, é um cenário pouco provável, no médio e longo prazo, uma vez que leva à redução da tarifa a um ritmo demasiado acelerado, face à redução de preços de instalação. Consequentemente, mesmo que a taxa de aprendizagem seja elevada, ou seja, 20%, um investimento num sistema de microgeração fotovoltaica deixará de ser interessante a partir de 2015, equivalendo a uma potência instalada total de 78 MWp, de acordo com o Gráfico 1.

4. DISCUSSÃO

A avaliação de um programa de promoção para a implementação de uma tecnologia emergente deve ser baseada num conjunto de critérios, designadamente: a eficácia de disseminação; o custo para a sociedade; as melhorias de desempenho da tecnologia; a conformidade de mercado e; a redução de custos (Lopez-Polo, 2007).

A análise aqui efectuada foca-se no impacto do desempenho da tecnologia e da redução de custos na eficácia da disseminação. A ideia subjacente à definição do tarifário, e a sua progressão, em função do crescimento da potência acumulada, tem por objectivo incentivar a redução dos custos associados aos sistemas. Segundo Haas:

O desenvolvimento de curto prazo de um mercado imperfeito – o que normalmente se verifica para uma tecnologia emergente – é caracterizado pelo aumento de custos se a procura aumenta. Porém, este facto induz o surgimento de novas empresas, aumentando o nível de concorrência e uma consequente descida de preços. No longo prazo, os preços descem até ser atingido um mercado competitivo perfeito.

O desafio do programa lançado pelo governo consiste em implementar um incentivo que permita induzir um mercado e paliar as suas imperfeições. Para atingir esse objectivo é lançada uma tarifa atractiva que evolui em função da potência instalada, forçando a uma maior eficiência do mercado. Este é um requisito fundamental para que o sector de serviços de energia fotovoltaica possa desenvolver-se de forma sustentada.

O cenário mais ambicioso, isto é, com um crescimento anual de 20%, leva a uma capacidade instalada expressivamente mais interessante do que os outros dois cenários, mas dos resultados

obtidos verifica-se que é um cenário pouco provável, uma vez que a descida das tarifas se efectuará a um ritmo demasiado acelerado.

Um crescimento anual nulo, torna a legislação viável (para as diferentes taxas de aprendizagem), mas não permite atingir valores de potência instalada próximos do pretendido.

No cenário intermédio, assumindo um crescimento anual de 10%, os valores de potência instalada são mais interessantes. Porém, a sua viabilidade a médio prazo, requer uma taxa de aprendizagem elevada.

Os resultados obtidos para os três cenários e diferentes taxas de aprendizagem da componente BOS estão resumidos nas Tabela 2 e Tabela 3.

Tabela 2 – Evolução de potência instalada para os três cenários

	Potência Instalada [kWp]		
Ano	Crescimento Anual: 0%	Crescimento Anual: 10%	Crescimento Anual: 20%
01-01-2009	4.602	4.602	4.602
01-01-2010	11.502	12.192	12.882
01-01-2011	18.402	20.541	22.818
01-01-2012	25.302	29.725	34.742
01-01-2013	32.202	39.827	49.049
01-01-2014	39.102	50.940	66.219
01-01-2015	46.002	63.164	86.822
01-01-2016	52.902	76.610	111.546
01-01-2017	59.802	91.401	141.215
01-01-2018	66.702	107.671	176.817
01-01-2019	73.602	125.567	219.540
01-01-2020	80.502	145.254	270.808
01-01-2021	87.402	166.909	332.329

Tabela 3 – Período de retorno do investimento para três cenários

	Período de Recuperação [anos]								
	Crescimento Anual: 0%			Crescimento Anual: 10%			Crescimento Anual: 20%		
Ano	TA=0%	TA=10%	TA=20%	TA=0%	TA=10%	TA=20%	TA=0%	TA=10%	TA=20%
01-01-2009	9	9	9	9	9	9	11	11	11
01-01-2010	9	9	8	10	9	8	14	12	9
01-01-2011	9	8	7	11	9	8	≥15	13	10
01-01-2012	10	8	7	13	10	8	≥15	≥15	10
01-01-2013	10	8	7	≥15	10	8	≥15	≥15	11
01-01-2014	11	9	7	≥15	12	8	≥15	≥15	12
01-01-2015	12	9	7	≥15	14	9	≥15	≥15	14
01-01-2016	12	9	7	≥15	≥15	10	≥15	≥15	≥15
01-01-2017	13	9	7	≥15	≥15	12	≥15	≥15	≥15
01-01-2018	14	9	7	≥15	≥15	13	≥15	≥15	≥15
01-01-2019	≥15	10	7	≥15	≥15	14	≥15	≥15	≥15
01-01-2020	≥15	11	8	≥15	≥15	14	≥15	≥15	≥15
01-01-2021	≥15	11	8	≥15	≥15	≥15	≥15	≥15	≥15

TA – Taxa de Aprendizagem

É importante salientar que a metodologia da curva de aprendizagem proporciona um modelo atractivo, uma vez que pode ser construído com apenas duas séries temporais empíricas – dados de custo e produção. Com efeito, existem diversas indústrias, nas quais se verifica que os custos decrescem de acordo com a curva de aprendizagem (incluindo o sector de produção de energia). Contudo o modelo da curva de aprendizagem apresenta algumas fragilidades, que devem ser

consideradas na análise dos resultados, designadamente: (1) a sensibilidade da localização temporal das reduções de custo a pequenas alterações da taxa de aprendizagem; (2) a impossibilidade de introduzir descontinuidades na taxa de aprendizagem, particularmente relevantes nas fases precoces da evolução de uma tecnologia; (3) o facto de ignorar factores técnicos específicos (como os limites teóricos de eficiência); (4) o facto de se aplicar a metodologia ao sector de indústria agregado em detrimento do nível da fábrica, o que implica à partida uma disseminação de conhecimento entre empresas perfeita; (5) o único determinante para a redução de custos consiste na produção acumulada, a qual representa a experiência adquirida, ignorando o conhecimento obtido de outras fontes, por via de actividades de I&D e de outras indústrias, por exemplo, e por fim; (6) o facto de ignorar as alterações de qualidade, tais como o aumento de eficiência, ou aumento de fiabilidade (Nemet, 2006). Apesar destas críticas, o modelo da curva de aprendizagem tem vindo a sobreviver sem grande alterações, como uma base para a previsão de evoluções técnicas, contribuindo para definição de políticas públicas e estratégias empresariais.

Apesar das fragilidades mencionadas, os resultados obtidos permitem concluir que os efeitos do incentivo proposto pelo MEI dependem em grande medida da evolução dos custos dos equipamentos no mercado internacional (factor exógeno) e da evolução do custo da instalação / taxa de aprendizagem no mercado nacional (factor endógeno), já que um aumento forte do ritmo das instalações de equipamentos pode fazer subir o período de recuperação do investimento para 15 anos ou mais, inviabilizando o próprio incentivo.

Assim, o incentivo proposto apenas será eficaz se se cumprirem um conjunto de condições. Algumas não são susceptíveis de controlo por parte do governo português, mas atendendo à análise efectuada, é razoável pensar em medidas que possam limitar o número de licenças atribuídas em determinados anos, ou promover uma taxa de aprendizagem mais elevada dos agentes nacionais.

No que respeita a aprendizagem associada a sistemas de microgeração fotovoltaica, uma das maiores dificuldades 0 prende-se com a diversidade de questões que podem surgir no que respeita a adaptação de um sistema a cada cliente. Este aspecto reflecte a elevada flexibilidades dos sistemas, mas acarreta algumas dificuldades para a redução de custos da componente BOS, uma vez que dificulta a aprendizagem entre diferentes projectos.

Por um lado, sugere-se a aposta na formação standard de profissionais dos solar fotovoltaico, com o objectivo de desenvolver um conhecimento comum entre diferentes projectistas e instaladores que facilite a comunicação e consequentemente a transferência de conhecimento.

Por outro lado, o desenvolvimento de uma plataforma para apoio na integração de sistemas, dirigida não só a profissionais, mas também a utilizadores. Esta plataforma deverá ter como objectivos fazer a ponte entre utilizadores e profissionais, de forma a disseminar as necessidades dos utilizadores e garantir a qualidade dos profissionais, mas também servir de infra-estrutura para a troca de experiências entre profissionais com vista à transferência de conhecimento e definição de standards.

BIBLIOGRAFIA

ALBERTH S. (2008) Forecasting technology costs via the experience curve — Myth or magic? *Technological Forecasting & Social Change*, Vol. 75, pp. 952–983.

DIRECÇÃO-GERAL DE ENERGIA E GEOLOGIA (2004) Utilização de Colectores Solares para Aquecimento de Água no Sector Doméstico, www.aguaquentesolar.com.

DIRECÇÃO-GERAL DE ENERGIA E GEOLOGIA (2008) *Renováveis – Estatísticas rápidas Outubro 2008*, disponível em www.dgeg.com

ENTIDADE REGULADORA DOS SERVIÇOS DE ENERGIA (2009) Tarifas e preços para a energia eléctrica e ouros serviços em 2009, Lisboa, ERSE.

EUROPEAN PHOTOVOLTAIC TECHNOLOGY PLATFORM (2007) The Status of the PV Industry http://www.eupvplatform.org/

HOFFMAN, W. (2006) PV solar electricity industry: Market growth and perspective. *Solar Energy Materials & Solar Cells*, Vol. 90, pp 3285–3311.

INTERNATIONAL ENERGY AGENCY – Photovoltaic Power Systems. (2007) *Cost and Performance Trends in Grid-Connected Photovoltaic Systems and Case Studies* http://www.iea-pvps.org/.

INTERNATIONAL ENERGY AGENCY – Photovoltaic Power Systems (2006) *National Survey Report of PV Power Applications in Portugal* – 2004, www.iea-pvps.org/.

INTERNATIONAL ENERGY AGENCY – Photovoltaic Power Systems (2008) *Trends in Photovoltaic Applications – Survey report of selected IEA countries between 1992 and 2007*, www.iea-pvps.org/.

KAZMERSKI, L. L. (2006) Solar photovoltaics R&D at the tipping point: A 2005 technology overview. *Journal of Electron Spectroscopy and Related Phenomena* Vol.150, pp. 105–135.

LOPEZ-POLO, A. & HAAS, R. & SUNA, D. (2007) *Economical Drivers and Market Impacts of Urban PV*, Intelligent Energy Europe, PV Upscale, WP5- Deliverable 5.1.

MINISTÉRIO DA ECONOMIA E DA INOVAÇÃO, Programa de Incentivos à Modernização da Economia, disponível em www.prime.min-economia.pt

NEMET, G. F. (2006) Beyond the learning curve: factors influencing cost reductions in photovoltaics. *Energy Policy*, Vol. 34, pp. 3218–3232.

PORTUGAL, MINISTÉRIO DA ECONOMIA E INOVAÇÃO (2007) *Política de Energia – Março de 2005 a Março de 2007*, Vol.I, www.min-economia.pt

PORTUGAL, MINISTÉRIO DA ECONOMIA E INOVAÇÃO (2007) *Micro-produção renovável na hora*, apresentação do MEI em Setembro de 2007.

PORTUGAL, MINISTÉRIO DA ECONOMIA E INOVAÇÃO (2007) Decreto-Lei n.º 363/2007, Diário da República, 1.ª série, n.º 211 de 2 de Novembro de 2007, pp. 7978-7984.

PORTUGAL, PRESIDÊNCIA DO CONSELHO DE MINISTROS (2005) *Estratégia nacional para a energia*, Resolução do Conselho de Ministros n.º 169/2005, Diário da República, I Série-B, n.º 204 de 24 de Outubro de 2005, pp. 6168-6176.

PORTUGAL, MINISTÉRIO DA ECONOMIA E DA INOVAÇÃO (2006) Decreto-Lei n.º 29/2006, Diário da República, I Série-A, n.º 33 de 15 de Fevereiro 2006, pp. 1189-1203.

PORTUGAL, MINISTÉRIO DAS OBRAS PÚBLICAS, Transportes e Comunicações (2006) *Regulamento dos Sistemas Energéticos de Climatização em Edifícios*, Decreto-Lei n.º 79/2006, Diário da Republica, I Série-A, n.º 67 de 4 de Abril de 2006, pp. 2416-2468.

PORTUGAL, MINISTÉRIO DAS OBRAS PÚBLICAS, Transportes e Comunicações (2006) *Regulamento das Características de Comportamento Térmico dos Edifícios* Decreto-Lei n.º 80/2006, Diário da Republica, I Série-A, n.º 67 de 4 de Abril de 2006, pp. 2468-2513.

Portugal, Assembleia Da República (2008) Proposta de Lei n.º 226/X, www.parlamento.pt.

PORTUGAL, MINISTÉRIO DA ECONOMIA E DA INOVAÇÃO (2008) Portaria n.º 201/2008, Diário da Republica, 1.ª Série, n.º 38, de 22 de Fevereiro pp. 1172-1173.

PORTUGAL, MINISTÉRIO DA ECONOMIA E DA INOVAÇÃO (2008) Portaria n.º 201/2008, Diário da República, I Série, n.º 38 de 22 de Fevereiro 2008, pp. 1172-1173.

PV POLICY GROUP, (2006) European Best Practice Report - Assessment of 12 national policy frameworks for photovoltaics, www.pvpolicy.org/.

SCHAEFFER, G. (2003) Experience curve as policy tool—the case of PV. Workshop on a Tool for Energy Policy Analysis and Design, Paris, 23 January 2003.

SHUM, K.L. & WATANABE, C. (2008) Towards a local learning (innovation) model of solar photovoltaic deployment. *Energy Policy*, Vol. 36, pp 508-521.

Van der Zwaan, B., Rabl, A. (2003) Prospects for PV: A Learning Curve Analysis. *Solar Energy*, Vol. 74, pp. 19-31.

GESTÃO AMBIENTAL E A CONTABILIDADE: SUBSÍDIO PARA O CRESCIMENTO DAS EMPRESAS E O DESENVOLVIMENTO NACIONAL

GILVAN DIAS DE LIMA
E-mail: gilvandl@gmail.com
THIAGO HENRIQUES FERNANDES
E-mail: thiagohenriques@gmail.com
Universidade Federal de Campina Grande

Resumen

Nota-se atualmente uma preocupação crescente, não somente com a perpetuação das empresas, mas também com a continuidade do ambiente onde elas estão inseridas. Percebe-se então que a questão ambiental pode ser uma grande fonte de investimento, e a contabilidade ambiental, procura demonstrar isso, aproximando cada vez mais a linguagem dos grupos ambientalistas com o meio empresarial. Assim, a contabilidade ambiental propõe uma abordagem que discute dois aspectos relevantemente diferentes, porém complementares: de um lado a busca constante das empresas em aumentar seu patrimônio e, de outro a utilização de recursos naturais para transformar os ativos em lucro e promover o desenvolvimento sustentável. Nesta circunstância, o presente trabalho procura apresentar respostas, mostrando como a contabilidade pode ser utilizada para a valorização mercadológica das empresas, atuando como fator de diferenciação ante a concorrência e discutindo a necessidade da mensuração e apresentação das contas ambientais nos demonstrativos contábeis.

Palabras Clave: Gestão ambiental e contabilidade, Crescimento empresarial, Desenvolvimento nacional.
Área Temática: Economía de la Energía, del Ambiente, de los Transportes y de los Recursos Naturales.

Abstract

it is seen recently, one growing concern not only for the perpetuation of the business, but also the continuity of the environment where they are inserted. It is seen that the environment issue can be a great source of investments, and the environmental accounting has trying to demonstrate that, approaching more by more the interests from the environmentalists and business man. So the environmental accounting, aims to show one approach that discusses two different aspects, but complementary: on one hand the constant search of companies to increase their wealth, and the use of other natural resources to turn assets into profit and promote national development sustainable. Within such a circumstance is that this paper seeks answers present, showing how the accounts can be used for the recovery of marketing companies, acting as a differentiation factor at the competition and discuss the need for measurement and presentation of environmental accounts in the financial statements.

Keywords: Environmental management and accounting. Business growth. National development.
Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

1.1 – ENTENDENDO A AMBIÊNCIA

O meio ambiente revela ser o conjunto de todos os organismos bióticos – vivos – e abióticos, como a água, solo, a (atmosfera) e energia solar, que formam a biosfera.

A atmosfera permite a existência de vida na terra, mantém a temperatura terrestre, impede ações de meteoritos, e fornece gases indispensáveis aos seres vivos. O solo permite a sobrevivência do homem disponibilizando os recursos vegetais, animais e minerais. Para a sobrevivência animal e vegetal percebemos a necessidade imprescindível da água, presente

através dos oceanos, rios, gelo e fontes subterrâneas; como também, da energia solar, que provê o calor que beneficia a animais e vegetais. (KRAEMER, 2004, p. 34)

A cadeia alimentar é formada por uma série de elementos que contribuem para a eficaz continuidade da natureza. Os animais dependem das plantas numa seqüência de ligação denominada de cadeia alimentar. As plantas precisam da água, do dióxido de carbono e luz para realizar a fotossíntese. É de esta cadeia alimentar que depende a sobrevivência dos seres vivos. O homem, como o principal agente destruidor da natureza, tem um importante dever de preservá-la para sua continuidade.

Desde o século XIX existia a preocupação com o meio ambiente; porém no século XX mais precisamente na década de 70 do século passado, repercutiu de modo mais intenso na sociedade, onde o problema passou de uma responsabilidade localizada, para uma responsabilidade globalizada. Com a deteriorização do meio ambiente, o seu futuro tornar-se bastante obscuro. Logo, a sociedade deve estar consciente da limitação dos recursos ambientais. A humanidade deve reconhecer que agredir a ambiência põe em perigo a sobrevivência de sua própria espécie, e pensar que o que está em jogo não é uma causa nacional ou regional, mas a existência da humanidade como um todo. É a vida que está em jogo. Podemos conceber um ecossistema sem o homem, porém não encontramos o homem sem estar inserido em algum ecossistema.

Acerca da utilização dos recursos naturais a Constituição Federal de 1988 nos é oportuna:

> *"Todos têm o direito ao meio ambiente ecologicamente equilibrado, bem de uso comum do povo e essencial à sadia qualidade de vida, impondo-se ao poder público e á coletividade o dever de defendê-lo e preserva-lo para as presentes e futuras gerações".*

Porém, é válido salientar que não se devem exaurir tais recursos e sim buscar o equilíbrio entre a produção e a preservação num processo de equilíbrio ecológico e que as pessoas a cada dia o reconhecem como sendo atividades economicamente e ambientalmente corretas e sustentáveis.

1.2 – IMPACTOS AMBIENTAIS

Poluição é o nome dado a qualquer degradação das condições ambientais, do habitat de uma coletividade humana. É uma perda mesmo que relativa, da qualidade de vida em decorrência de mudanças ambientais. Existem diversos agentes que provocam a poluição com o petróleo, gás nocivos, agrotóxicos, etc. Fazendo um levantamento histórico observa-se que a questão da poluição já acontecia em tempos passados, no império romano, mas tratava-se de um problema localizado. Com a industrialização e o surgimento dos grandes centros urbanos a poluição cresceu bastante e deixou de ser um problema local para se tornar um problema global. (MAZON, 1979, p.84).

1.3 – IMPACTOS NACIONAIS

Em 1984, tragédia sacudiu Cubatão (SP). Duas fortes explosões e o incêndio que se seguiu, provocados por um Vazamento de gasolina nos dutos da Petrobrás, incendiaram mais de 1000 casas n bairro de Vila Socó e mataram mais de 150 pessoas.

Em 2000, um duto avariado da Petrobrás permitiu que 1,3 milhão de litros de óleo contaminasse o maior cartão postal do país, a Baía de Guanabara. Esse desastre ambiental levou o governo estadual a criar o grupo Gestor da Baía da Guanabara, formado por todos os órgãos da administração estadual ligados ao meio ambiente.

Em 2001, duas explosões seguidas destruíram a Plataforma de Processamento de Petróleo P-36, em Macaé, no Campo do roncador, Bacia de Campos, no Estado do Rio de Janeiro. A

plataforma afundou a 120 quilômetros da costa, colocando em risco duas espécies de aves migratórias marinhas, que estão ameaçadas de extinção. (KRAEMER, 2004, p. 45)

1.4 – IMPACTOS INTERNACIONAIS

Em outubro de 1966, o mal funcionou do sistema de refrigeração de uma usina de Detroit causa o derretimento parcial do núcleo do reator. Na central nuclear de Chernobyl, na cidade de Pripyat (próximo a Kiev, na Ucrânia), um acidente provocou em 1986 o aumento da temperatura normal da câmara do reator atômico, em 2204 graus centígrados para 3000 graus centígrados. Isso danificou a cobertura de grafite que envolve o urânio atômico, que começou a queimar em contato com o oxigênio e derreteu o próprio núcleo de urânio do reator. Com isso foi liberada alta dose de radioatividade para a atmosfera, que em menos de uma semana se dispersou por vários países da Europa. (KRAEMER, 2004, p. 38)

1.5 – HISTÓRICO DA CONTABILIDADE

A contabilidade surgiu desde os tempos em que o homem primitivo necessitou de informações e registros para conhecer seu patrimônio – objeto central da contabilidade. O surgimento do comércio, e posteriormente da indústria, contribuiu para a formação da base científica da contabilidade, agregando conhecimentos baseados nas necessidades de registros e mensuração que surgiram nas idades média e moderna, possibilitando a formação de diversos segmentos como a contabilidade comercial, de custos, auditoria, dentre outros.

A partir do século XIX, com novas formas de divisão do trabalho e da produção, ampliaram consideravelmente a produção de riquezas através de conquistas de novos mercados de consumo, exigindo grandes investimentos e financiamentos que permitiram a formação dos mercados de capitais, especialmente com o surgimento das bolsas de valores como a de Londres e Nova York – EUA. Tais atitudes necessitaram do uso mais aprofundado da contabilidade. (KRAEMER, 2004, p.19)

1.5.1 - A nível internacional

O Século XX foi cenário de diversas agressões ao meio ambiente o que levou a uma conscientização da sociedade de que a preservação dele tem de ser discutido a nível internacional, e de que deve haver uma mudança cultural e educacional em relação à visão ambiental e ecológica. Assim surgiu na década de 70 em várias partes do mundo, movimentos de setores organizados da sociedade que passaram a discutir formas para minimizar os impactos de agressão a ambiência, destacando-se os decorrentes de atividades econômicas.

A conferência Mundial sobre o Meio Ambiente das Nações Unidas realizada em Estocolmo em 1972 com grande repercussão internacional discutiu a relação do homem com o meio ambiente e de mecanismos para tratar de questões ambientais. A Organização das Nações Unidas - ONU instituiu o Programa das Nações Unidas para o Meio Ambiente - PNUMA.

Em 1975 foi realizado em Belgrado um Seminário Internacional de Educação, com a participação de vários países, tendo como resultado a Carta de Belgrado estabelecendo que toda ação de preservação ambiental deveria passar por um programa de educação ambiental.

Para Ferreira (2003:12):

> *"O desenvolvimento de programas de educação ambiental pelos países devem conscientizar as crianças e dar conhecimento sobre a importância de cuidar do meio ambiente, preparando profissionais com habilidades especificas e conhecimento da ciência ambiental com condições de tomar decisões no futuro em relação a questão ambiental".*

Em 1992 houve a Conferência das Nações Unidas sobre o Meio Ambiente e Desenvolvimento a ECO-92 no Rio de Janeiro tendo como objetivo elaborar estratégias para diminuir os efeitos de deterioração ambiental e promover o desenvolvimento sustentável.

O Programa das Nações Unidas para o Desenvolvimento – PNUD, destaca a nível internacional o Protocolo de Quioto que após varias discussões desde 1997 foi ratificado por 141 países (Estado Unidos não assinou) e entrou em vigor em fevereiro de 2005 onde os países participantes se comprometem a reduzir a emissão de gases poluentes em 5,2% entre 2008 e 2012 em comparação com o nível de 1990. Foi uma tarefa difícil dos organismos internacionais e da ONU para que o protocolo pudesse entrar em operação.

1.5.2 – No Brasil

Em 1981 com a Lei 6.938 foram estabelecidos objetivos e instrumentos da Política Nacional de Meio Ambiente, uma Lei para assegurar o desenvolvimento econômico com a preservação ambiental e qualidade de vida. Esta Lei criou o Sistema Nacional do Meio Ambiente que através do Conselho Nacional do Meio Ambiente - CONAMA, que é composto por representantes dos poderes políticos em diferentes níveis, instituições civis, órgãos ambientais estaduais e municipais e de setores organizados da sociedade.

A Resolução CONAMA n° 001/86 constitui um marco da política ambiental brasileira, pela instituição da obrigatoriedade de que as empresas elaborem um Estudo de Impacto Ambiental – EIA para identificar os tipos de impactos ambientais provocados pelas atividades, as conseqüências e os meios preventivos, divulgando os resultados através do Relatório de Impacto Ambiental – RIMA.

O Art. 225, da Constituição de 1988 inclui na longa declaração de direitos e deveres individuais e coletivos, o direito de todos a um meio ambiente ecologicamente equilibrado.

No Brasil surgiram iniciativas tanto por parte do governo quanto da iniciativa privada visando implementar ações de proteção ao meio ambiente, como o Protocolo Verde em 1995, um acordo entre o Instituto Brasileiro do Meio Ambiente e Recursos Naturais Renováveis - IBAMA e o Banco Nacional de Desenvolvimento Econômico e Social - BNDES e alguns outros bancos, cujo objetivo é subsidiar a aquisição de tecnologias antipoluentes.

De maneira geral cabe destacar a importância da legitimação das Leis, ações e políticas ambientais, junto aos outros países, ao setor produtivo e a sociedade em geral, que deve entender que são instrumentos institucionais a serviço do bem coletivo, da preservação do meio ambiente e melhoria da qualidade de vida.

2 – A EMPRESA E A QUESTÃO AMBIENTAL

A questão ambiental passou a ser vista como uma fonte de investimentos para o meio empresarial, pois a sociedade como um todo, vem cobrando cada vez mais a preservação dos recursos ambientais, uma vez que estes afetam diretamente o meio em que está inserida.

Para garantir a continuidade, as empresas devem ser capazes de atender as exigências e necessidades da sociedade onde estão inseridas, apresentando um alinhamento entre os interesses econômicos com os interesses da sociedade, adotando medidas que muitas vezes produzem resultados econômicos menores no curto prazo, mas que a longo prazo será vantajoso Segundo Barbieri (2004:99):

> *"As preocupações ambientais dos empresários são influenciadas por três grandes conjuntos de forças que se interagem reciprocamente: o governo, a sociedade e o mercado. Se não houvesse pressões da sociedade e medidas governamentais não se observaria o crescente envolvimento das empresas em matéria ambiental".*

Algumas empresas vêm inserindo a questão ambiental nas premissas básicas de seus negócios, adotando o gerenciamento ambiental. Além das grandes empresas existem pequenas e médias empresas, envolvidas e preocupadas com a preservação ambiental.

Muitas empresas vêm adotando o processo de reciclagem como forma de reduzir custos e também diminuir o impacto negativo de suas atividades sobre o meio ambiente. As empresas de papel e celulose, um dos setores mais críticos em termos de poluição já vêm adotando metodologias avançadas nos seus processos produtivos, objetivando eliminar os resíduos, próprios da atividade, pois como diz Barbieri (2004:111):

> *"A abordagem ambiental estratégica significa tratar sistematicamente as questões ambientais para proporcionar valores aos componentes do ambiente de negócio da empresa que os diferenciem dos seus concorrentes [...]. A empresa que se antecipa no atendimento dessas novas demandas por meio de ações legítimas e verdadeiras acaba criando um importante diferencial estratégico".*

Com relação ao mercado internacional as empresas interessadas estão se adaptando aos padrões de qualidade ambiental para obtenção do certificado ISO 14000, atendendo as normas definidas pela *Internacional Organization for Standardization* (ISO), fundada em 1947, com sede em Genebra na Suíça. O objetivo principal da ISO é criar normas internacionais de padronização que representem e traduzam o consenso dos diferentes países.

Esse certificado é importante para continuar comercializando em países que não admitem a entrada de produtos que possam agredir o meio ambiente local ou que tenham sido produzidos por processos ecologicamente não recomendáveis.

Em alguns setores econômicos para que a empresa se constitua e funcione torna-se necessário a apresentação de um Estudo de Impacto Ambiental - EIA e o Relatório de Impacto ao Meio Ambiente - RIMA aos órgãos técnico-governamentais responsáveis, que devem estar de acordo com os parâmetros definidos para o desenvolvimento sustentável, pois conforme Carvalho e Ribeiro (2000), esses documentos tem sido acrescentados nas exigências de algumas instituições financeiras para análise e concessão de crédito de maneira a aumentar as garantias de continuidade e melhorar as condições de seus clientes honrarem seus compromissos bancários.

Na visão geral a empresa precisa demonstrar que, além dos aspectos econômicos e financeiros, também atua com preocupação social e ambiental. Este comportamento fortalece a imagem da empresa junto à sociedade atraindo melhores empregados e fornecedores, maiores investimentos, tendo um diferencial para enfrentar a concorrência, pois os consumidores além dos preços adquirem marcas ligadas às empresas cidadãs que em suma são socialmente responsáveis.

2.1 - A empresa e a ambiência

Com a evolução dos tempos, nota-se uma grande preocupação com as questões ambientais, pois a perpetuação das empresas, depende da preservação da ambiência em que estão inseridas.

Franco (1999; p.38), afirma que:

> *"As empresas que protegem o meio ambiente são bem vistas pelo consumidor e por investidores, já havendo, em alguns países, fundos especializados em investir nestas empresas. Esses fundos têm crescido e apresentado rentabilidade, porque as empresas são também mais lucrativas".*

O interesse com a preservação do meio ambiente, além de favorecer a empresa e garantir seu perpetuar na sociedade, traz consideráveis melhorias na vida de todos os cidadãos. Se esta

preocupação inicia nas organizações empresariais, logo os indivíduos também começam a perceber o quanto é importante a preservação inclusive para sua própria sobrevivência.

Tal preocupação faz com que a própria administração comece a modificar sua maneira de realizar seus negócios, revendo seu plano de vendas, de controle, de produção, de marketing e outros. Enfim, leva a administração empresarial a tomar novas decisões que trarão benefícios importantes para a diferenciação da empresa na sociedade, e desta maneira aumentar sua capacidade de atrair mais consumidores, possibilitando incremento em seus negócios e consequentemente, mais lucro.

Empresas que ainda não se preocupam com padrões de qualidade, que não procuram utilizar-se de equipamentos para minimizar o impacto ambiental à sua volta, são severamente punidas, pois com a nova mentalidade dos consumidores, de que é necessária a preservação, empresas que não se adequarem a este padrão, simplesmente estarão condenadas ao abandono de seus consumidores. Desta forma, ou se enquadram no novo sistema, ou estarão prestes a ver seu fim muito mais próximo do que imaginam.

De certa maneira a questão ambiental passou a ser vista como uma fonte de investimento, pois a sociedade em si vem cobrando cada vez mais a preservação dos recursos ambientais, uma vez que estes estão afetando diretamente o meio em que elas se inserem.

A necessidade da preservação do meio ambiente, deixou de ser preocupação de grupos ambientalistas e de organizações não-governamentais, e passou a ser uma responsabilidade da humanidade, pois se percebe a importância do mesmo, não só para o ser humano, mas também para as empresas que usufruem dele, e até então não tinham uma preocupação neste sentido.

Na visão de Barbieri (1997, p.199):

> *"O crescimento da consciência ambiental, ao modificar os padrões de consumo, constitui uma das mais importantes armas em defesa do meio ambiente. Quando a empresa busca capturar oportunidades através do crescente contingente de consumidores responsáveis através de ações legítimas e verdadeiras, essas ações tendem a reforçar ainda mais a consciência ambiental, criando um círculo virtuoso, na qual a atuação mercadológica, marketing verde, como querem alguns, torna-se um instrumento de educação ambiental".*

Para que as empresas percebam realmente a importância da preservação da ambiência, é necessário que incorporem em seu planejamento estratégico e operacional um adequado programa de gestão ambiental, compatibilizando seus objetivos com sua inserção no meio em que vive.

Com toda esta evolução, cabe ao contador transformar a contabilidade em uma ferramenta essencial na perpetuação das empresas, capaz de aproximar a linguagem contábil da linguagem dos ambientalistas, procurando demonstrar as adequações necessárias à empresa no sentido de obter os resultados esperados diante de seu investimento na preservação e subsistência do meio ambiente.

Ao analisar historicamente a evolução da preservação ambiental e sua aplicação, percebe-se que é uma questão muito recente, pois foi, a partir de 1972 com a Primeira Conferência Mundial sobre o Meio Ambiente realizada em Estocolmo, que deu-se início a grande preocupação com o meio ambiente para a perpetuação da humanidade.

Considerando este aspecto, Ferreira (2003, p.15.) apresenta quatro princípios da ecologia que interferem no dia a dia das empresas e da sociedade:

- Toda entidade em particular está ligada a todo resto;
- Tudo vai para algum lugar;
- Você não pode conseguir as coisas de lugar nenhum;
- A natureza é sábia.

Estes princípios demonstram que as empresas e a sociedade devem rever seus conceitos com relação ao meio ambiente, unindo empresários e ambientalistas na busca de soluções para a melhoria do planeta.

3 – DANO AMBIENTAL

Como bem jurídico, o meio ambiente pode ser objeto de danos em decorrência da atividade humana. Silva (1998, p.71), com base no artigo 225, § 3º, da Constituição Federal do Brasil de 1988, define dano ecológico como:

> "Qualquer lesão ao meio ambiente causado por condutas ou atividades de pessoa física ou jurídica de direito público ou direito privado".

Segundo Nery (1985, p.130):

> *"Podemos conceituar dano ecológico como sendo prejuízo causado ao meio ambiente"*.

Em caráter legislativo, encontra-se uma definição do que seja poluição na Lei 6.938, de 31 de agosto de 1981, que estabelece a Política Nacional do Meio Ambiente:

> *"Poluição é a degradação da qualidade ambiental resultante de atividades que direta ou indiretamente prejudiquem a saúde, a segurança e o bem-estar da população, criem condições adversas às atividades sociais e econômicas, ou afetem as condições vitais, estéticas ou sanitárias do meio ambiente, ou ainda lancem matérias ou energia em desacordo com os padrões ambientais estabelecidos"*.

Tem-se estabelecido, através das exposições doutrinárias, a necessidade de três requisitos para a configuração do dano ambiental, do qual decorre o dever de indenização: a anormalidade; a periodicidade e a gravidade.

A gravidade se verifica pela transposição dos limites de absorção das agressões que possuem os seres humanos, e principalmente os elementos naturais.

O dano deve também ser periódico, não bastando para a sua configuração uma eventual emissão de poluição. Mas é preciso que se evidencie o sentido que se dá à periodicidade. Aqui, ela não assume a noção cotidiana de transcurso de certo lapso temporal; consiste na necessidade de que a poluição se verifique por tempo suficiente para a produção de um dano substancial e grave, não sendo relevantes, por exemplo, as emissões momentâneas de odores.

A anormalidade é aferida em comparação com o que se aceita como decorrência normal de determinada atividade; ocorre quando se constata uma modificação substancial nos elementos naturais circunvizinhos à atividade, em grau suficiente para fazer com que os mesmos percam total ou parcialmente suas propriedades de uso.

Pela junção desses requisitos, não será cabível discutir o dano ambiental e, conseqüentemente, sua reparação, quando as alterações causadas ao meio ambiente por atividades comerciais ou industriais forem 'normais e toleráveis', só merecendo repressão quando atingirem níveis intoleráveis e prejudiciais à comunidade. O julgado apresentado a seguir demonstra a transposição dos limites de tolerância como configuradora do dano ambiental:

> "A prova pericial demonstrou, de maneira inconteste, a poluição ambiental provocada pelo

> derramamento de óleo da barcaça 'Gisela', e pouco importa que o estuário de Santos já estivesse poluído, eis que o fato transbordou dos limites do tolerável, danificando áreas marginais, e seria, mesmo, contra-senso inadmitir poluição apenas porque já preexistia elemento poluidor no local [omissis]. (TJSP - 3ª C. – Ap. – Rel. Toledo Cézar – j. 07.04.87 – RT 620/69)."

O artigo 225 da Constituição Federal do Brasil prescreve ser o meio ambiente um bem de uso comum do povo, ou seja, é uma propriedade de toda a coletividade. Não é um conjunto de propriedades individuais, é um direito indivisível, logo cada indivíduo em si, não é considerado titular do direito ao meio ambiente.

O dano ecológico sempre existiu, e o Direito criou algumas soluções que rapidamente tornaram-se envelhecidas e superadas, mas nunca houve um hiato tão grande entre o Direito e a realidade fática, vez que os danos ecológicos sofrem verdadeira agravação geométrica, por influência do considerável perigo criado pelas conquistas científicas que não foram seguidas de iguais provisões de cautela por parte dos seus responsáveis.

O que existe hoje para a reparação dos danos ambientais é insatisfatório, tendo em vista que os instrumentos para tal ação são precipuamente de direito privado, e como já ficou estabelecido, o direito ao meio ambiente é meta-individual, devendo ser regulado por disposições de direito público mescladas às já existentes normas privatísticas.

Pelas próprias características do dano ambiental, os agredidos que se propõem a pleitear sua reparação deparam-se com inúmeros problemas, como a comprovação da ocorrência do prejuízo, do seu montante, e de suas causas, sem contar as situações em que se prevê um dano futuro, mas não se consegue mensurá-lo.

Essa dificuldade não passou desapercebida por Lucarelli (1994, p.15), que afirma:

> *"Ora, por sua natureza, o dano ambiental é de dificultosa comprovação, eis que seus efeitos são, em maior parte, invisíveis ao olho do indivíduo desatento, não se verificando, em regra, de imediato, além da expressiva quantidade de casos em que é impossível especificar um responsável pelo dano, já que nas proximidades há um numeroso número de empresas ou poluidores em potencial".*

Pelas poucas dificuldades relacionadas já é possível antever o que sucederia numa ação de reparação de danos ambientais caso aplicada a regra da teoria geral das provas, que impõe a quem alega um fato o ônus de prová-lo. Em matéria ambiental é de grande valia a aplicação da construção basicamente jurisprudencial da culpa presumida, que transfere para o acusado de lesionar o meio ambiente o ônus de demonstrar a regularidade de sua conduta.

Um último ponto a ser levantado diz respeito aos danos futuros causados por atividade poluidora. Além de ser de difícil mensuração, existe todo um problema interno da ordem jurídica para trabalhar com eventos futuros. Somente estudos científicos conseguem demonstrar com relativa certeza as dimensões que um dano ambiental poderá assumir com o transcurso do tempo, e no direito brasileiro não se admite uma decisão judicial baseada em probabilidades, o que pode deixar sem reparação danos vultosos ao ambiente.

Assim, a doutrina tem se posicionado no sentido de ser necessária a efetiva ocorrência do dano para se pleitear uma reparação, independentemente da data em que ocorreu o fato causador deste dano.

3.1 – Reparação de danos ambientias

O dano ambiental possui duas vias de reparação: uma administrativa e outra judicial. Pela via administrativa, o Poder Público impõe aos agressores do meio ambiente ações tendentes a recompor as áreas lesadas e recolocá-las na sua situação antes da lesão (*status quo ante*), ou o mais próximo possível disso. Exemplo desta situação é a exigência em que as madeireiras devem possuir áreas de reflorestamento idênticas (ou quase) às por elas desmatadas. Outra forma é aplicação de tributos majorados sobre as atividades poluentes, com o destino dos fundos arrecadados para atividades de preservação ambiental.

Mas a configuração da responsabilidade civil é uma forma de reparação do dano ambiental através do Poder Judiciário, independente das ações administrativas de preservação e reconstituição ambiental. A partir do momento em que se fixa o dever de indenizar, surge o problema de saber quais são as formas de reparar os danos.

A primeira preocupação do autor da ação de indenização deve ser a de procurar, em juízo, obter a recomposição, ou seja, a volta à situação anterior, do ambiente deteriorado, desde que isso seja possível. Pretende-se, inicialmente, que o agente poluidor reconstitua a paisagem natural deteriorada por sua atividade, ou pelo menos a deixe num estado semelhante ao que existia antes da poluição.

Subsidiariamente, quando for impossível a recomposição, deve ser solicitada indenização pecuniária pelo dano ambiental, que necessariamente precisa ser revertida a um fundo de preservação ambiental.

Como última forma de reparação, deve ser solicitada a condenação em obrigação de fazer ou não fazer, como a colocação de filtros em chaminés ou a abstenção de incineração de determinados produtos.

A responsabilização do degradador ambiental deixou de exigir a comprovação de culpa, bastando à constatação do nexo de causalidade entre o ato e a lesão sofrida para a imputação de pena. Criou-se um regime de reparação autônomo na área ambiental, com responsabilização objetiva.

Desta maneira, a Lei da Política Nacional do Meio Ambiente trata da responsabilidade civil de maneira mais ampla, porém, restrita ao dano ambiental. A Constituição Federal do Brasil de 1988 (CF/88) vale lembrar, em seu art. 225, § 3º, recepcionou a Lei 6.938/81, Lei da Política Nacional do meio ambiente, e deixou intacta a responsabilização objetiva do causador do dano ambiental.

3.2 - Mensuração dos danos ambientais

A questão da mensuração dos danos ambientais pela empresa é problemática, pois em muitas situações envolvem variáveis complexas, devido a dificuldade em medir o valor real de uma perda relativa a alguma agressão provocada, como por exemplo, contaminação de solo por resíduos de uma empresa, ou o impacto ambiental ocasionado pelo derramamento de óleo no oceano o que leva também há dificuldades de avaliar e determinar o custo necessário para aquisição de produtos e tecnologia para recuperação das perdas ocasionadas.

Dessa forma a mensuração relativa a muitos passivos ambientais, torna-se de difícil execução devido a falta de meios e técnicas adequadas para o seu reconhecimento ou pela difícil identificação imediata.

Assim, para Ribeiro (2005:104), as dificuldades para a mensuração dos gastos na área ambiental na ocorrência do fato gerador, podem ser amenizadas pelas técnicas estatísticas e matemáticas ao dispor da contabilidade, pelos estudos de riscos e experiências anteriores. È melhor lançar valores aproximados que omitir totalmente despesas que poderão afetar o resultado do período.

Para um melhor reconhecimento do quanto será necessário para restauração ou recuperação pela lesão provocada, assim como medir o valor de danos causados, é necessário um estudo técnico e especifico da área atingida e deteriorada.

É papel da contabilidade analisar os fatos e dentro do possível aplicar o princípio da confrontação entre receitas e despesas reconhecendo e contabilizando as transações que envolvem desembolso por parte da empresa no momento em que ocorre o fator gerador.

Para Ferreira (2003:101):

> *"As contas de despesas quando tem relação com o processo produtivo devem acompanhar a contabilização do custo de produção, sendo que as contas de receitas e despesas relativas ao meio ambiente podem ser organizadas da seguinte forma:* ***Despesas Ambientais:*** *Recuperação de áreas degradadas, Depreciação de equipamentos e Outras despesas ambientais.* ***Receitas ambientais:*** *Receita de serviços e Receita de venda de material reciclado".*

É importante para a empresa analisar os montantes envolvidos, visto que, em alguns casos os danos ao meio ambiente envolvem valores monetários consideráveis o que podem afetar o valor econômico e o patrimônio da empresa.

4 - A CONTABILIDADE AMBIENTAL

Visto que a questão ambiental tem sido muito discutida, conseqüentemente a questão da mensuração desse patrimônio torna-se motivo de estudo. A contabilidade pode evidenciar em suas demonstrações contábeis os eventos e transações econômico-financeiro que tem impacto no meio ambiente. Desta forma a contabilidade passa não só a evidenciar a evolução econômica das empresas, mas também as contribuições e responsabilidade destas com o meio ambiente.

Assim a contabilidade Ambiental não se refere a uma nova Contabilidade, ela é sim uma especialização da Contabilidade que visa identificar os ativos e passivos ambientais, bem como as questões econômicas que se relacionam com a gestão ambiental de uma organização. O desenvolvimento da Contabilidade Ambiental é resultado da necessidade de oferecer informações adequadas às características de uma gestão ambiental.

Percebe-se que a contabilidade ambiental é mais bem caracterizada através da divisão em contabilidade financeira ambiental e contabilidade gerencial ambiental. A contabilidade financeira ambiental visa atender aos clientes externos, com informações padronizadas que atendam aos princípios contábeis, através do fornecimento de registros contábeis ambientais, planos de contas e contabilização dos principais eventos relativos ao meio ambiente.

Em 1996, o Instituto Brasileiro de Contadores – IBRACON, atual Instituto dos Auditores Independentes do Brasil, publicou a Norma de Preservação Ambiental - NPA 11, que instituiu "*os liames entre a Contabilidade e o Meio Ambiente*", incumbindo a Contabilidade de "*participar dos esforços em favor da defesa e proteção contra a poluição e as agressões à Vida Humana e à Natureza*". Essa norma também comenta que as diretrizes ecológicas devem ser refletidas nas demonstrações contábeis e relatórios da administração, as quais "*devem se ajustar a um novo mundo, no qual metas ambientais, além das econômicas, terão de ser alcançadas rigorosamente*".

Devido ao aumento de usuários interessados nessas informações os princípios, critérios e técnicas utilizados pela contabilidade vêm evoluindo e se aperfeiçoando, para que a situação da empresa seja mais bem analisada. Assim, a contabilidade pode contribuir para solucionar problemas ambientais por meio da identificação e informações específicas sobre despesas e custos ambientais.

A contabilidade ambiental surge da necessidade de demonstrar adequadamente as informações decorrentes da adoção de uma política e normas de gestão ambiental pela empresa.

Ribeiro (2005:45) é oportuno quando afirma:

> *"Podemos definir como objetivo da contabilidade ambiental: identificar, mensurar e esclarecer os eventos e transações econômico-financeiros que estejam relacionados com a proteção, preservação e recuperação ambiental, ocorridos em um determinado período, visando a evidenciação da situação patrimonial de uma entidade."*

Para contabilização dessas transações é necessária a definição de normas e regras para o registro dos gastos, custos, despesas, provisões, assim como de divulgação de informações qualitativas.

Desta forma a contabilidade ambiental tem por objetivo registrar e controlar as atividades desenvolvidas ou em andamento, que possam vir causar qualquer dano ao meio ambiente, bem como as ações destinadas a amenizar ou extinguir as degradações, devendo ser registrada em contas contábeis específicas, na data de sua ocorrência, de acordo com os Princípios Fundamentais de Contabilidade estabelecidos pela Resolução 750/93 do Conselho Federal de Contabilidade – CFC.

4.1 - Princípios contábeis e contabilidade ambiental

De forma geral, os princípios direcionam o conhecimento científico da contabilidade sem perder o seu rumo, ou seja, sem percorrer caminhos ilegais que tomem contraditórios o pensamento contábil. Como representa a base científica toma todas as normas e convenções subordinadas aos princípios. Segundo o Conselho Federal de Contabilidade, e os princípios e convenções fundamentais são:

Entidade Contábil: controle sobre o uso dos recursos naturais, diferenciados dos recursos dos sócios. Logo, referente a impactos ambientais causados por determinada entidade, não podem ser reconhecidos por outra entidade.

Continuidade: prevê a continuidade dos recursos naturais à disposição da empresa, permitindo o crescimento da empresa. Do ponto de vista ambiental, no entanto, deve ser vista com restrições, pois os recursos naturais são finitos.

> [...] essa continuidade indefinida pode colaborar fortemente para a descontinuidade do planeta. Para evitar isso, esse postulado deveria limitar-se ao crescimento, evitando o esgotamento dos recursos naturais indispensáveis à vida humana, tais como o oxigênio, ozônio, alimentos, etc. (FERREIRA, 2003, P.94)

Oportunidade: as informações ambientais devem ser registradas e disponibilizadas no tempo oportuno para permitir a ação ambiental de preservação do meio ambiente, refletindo a integridade e mutações patrimoniais.

Registro pelo Valor Original: os impactos ambientais devem ser registrados pelos valores originais das transações, expressos a valor presente na moeda do país.

Atualização Monetária: caso o impacto ambiental somente seja percebido pela empresa daqui a 10 ou 20 anos depois, é muito importante que haja o reconhecimento da alteração do poder aquisitivo da moeda nacional.

Competência: os fatos geradores relacionados ao meio ambiente devem ser incluídos na apuração do resultado, das receitas e despesas, no período em que ocorrerem.

Prudência: é condição importante para o reconhecimento dos riscos relativos ao meio ambiente que colocam em risco o patrimônio da entidade.

Confrontação: devemos reconhecer as receitas juntamente com as despesas necessárias para sua obtenção. Por exemplo: a venda de um produto deveria ser reconhecida no momento com o custo de degradação causado para produzir esse mesmo produto.

Objetividade: é preciso estimar fisicamente, se possível, os impactos ambientais e depois seus efeitos.

Materialidade: são os impactos cujos efeitos puderem ser percebidos de forma aparente e forem economicamente mensuráveis, a partir da poluição medindo fisicamente o que aconteceu.

Evidenciação: é a confirmação da informação necessária que definirá o registro contábil. É uma prestação de contas à sociedade de como o meio ambiente está sendo afetado pela gestão ambiental de determinada empresa.

Todo evento econômico que altere o patrimônio da entidade e que esteja ligado direta ou indiretamente ao meio ambiente é um fator gerador e deve, de alguma forma mesmo por estimativas ser contabilmente reconhecido, mesmo que apenas nas notas explicativas. A determinação do fato gerador dos impactos ambientais é determinante para o reconhecimento dos eventos que modificam o patrimônio no momento em que ele ocorre.

No tocante a evidenciação a contabilidade ambiental também vem sendo utilizada para valorização da imagem da empresa junto ao consumidor, sendo um fator de diferenciação com a concorrência, tornando-se evidente a importância e necessidade da inclusão de contas de natureza ambiental nas demonstrações contábeis elevando-a ao imaginário social sob o título de empresa cidadã.

4.2 - Finalidades e Objetivos da Contabilidade Ambiental

A finalidade da contabilidade ambiental é, dentre outras metas, realizar a gestão ambiental e seu controle, visando reduzir custos e despesas operacionais e melhorar a qualidade dos produtos; também, busca atender as crescentes exigências legais e normativas como obrigar os diretores a controlar mais seus riscos ambientais, sob pena de multas e de indenizações.

Como a empresa está submetida cada vez mais a pressões internas e externas, têm aumentado a demanda da atuação dos parceiros sociais, como os clientes, empregados, organizações ecológicas, seguradoras, comunidade local, acionistas, administração pública, bancos e investidores. Para atender a estas finalidades é importante conhecer quais são os objetivos dessa contabilidade. Basicamente é necessário:

- Identificar se a empresa cumpre ou não com a legislação ambiental existente;
- Identificar se existe gestão ambiental ou orientação quanto ao processo decisório;
- Comprovar a evolução das tendências e atuação ambiental da empresa através do tempo;
- Observar se cumpre com os objetivos ambientais fixados pela companhia;
- Identificar oportunidades para melhor gestão dos aspectos ambientais, e oportunidades estratégicas, como vantagens competitivas como o uso científico da contabilidade ambiental;
- Obter informação específica para fazer frente à solicitação dos usuários da contabilidade ambiental.

4.3 – Classificação da contabilidade ambiental

Segundo a Agência de Proteção Ambiental - EPA - a contabilidade ambiental pode ser estruturada sob três aspectos:

- Contabilidade Nacional;
- Contabilidade Gerencial ambiental;
- Contabilidade Financeira Ambiental.

Contabilidade Nacional Ambiental: reflete sua preocupação quanto às unidades físicas ou monetárias de acordo com o consumo de recursos naturais no país de origem em que esta ocorrendo o registro contábil, quer sejam recursos renováveis ou não renováveis. A

contabilidade nacional tem um enfoque macroeconômico, particularmente no estoque e nos fluxos de recursos naturais, em custos ambientais, em custos de externalidade, dentre outros. (KRAEMER, 2004, p.150)

Contabilidade Gerencial Ambiental: é a identificação, a estimação, a análise e o uso do custo de informações para tomada de decisão dentro de uma organização. Tem por objeto o processo de identificação, compilação e análise de informação, principalmente para fins internos, gerenciais. Está dirigida a gestão de resultados, que compreende a produção, os custos e as receitas, em especial para tomada de decisões. A contabilidade gerencial ambiental é, portanto, a contabilidade de gestão com foco centrado em balanços de massa, em fluxos de materiais, fluxo de energia e na informação do custo ambiental, ou seja, centrados na produção empresarial. É dirigido com enfoque ao usuário interno, como os administradores da empresa que irão utilizar tais informações (registros contábeis) para tomada de decisões. A utilização da contabilidade e gestão ambiental – CGA – pode proporcionar grandes reduções de custos na gestão de resíduos, dado que os custos de manuseio e deposição de resíduos são relativamente fáceis de definir e de atribuir a produtos específicos. Existem outros mais difíceis de avaliar como os custos e despesas ambientais, incluindo os custos da deterioração da imagem da empresa; riscos e responsabilidade ambiental de acordo com a legislação específica. Os dados da CGA são aplicáveis em vários campos de atuação como:

- Avaliação anual dos custos e despesas ambientais;
- Preços de produtos;
- Elaboração de orçamentos empresariais;
- Avaliação e análise de alternativas de investimentos;
- Cálculo de custos, economias e benefícios de projetos;
- Design e implementação do sistema de gestão ambiental (SGA);
- Avaliação do desempenho ambiental.

A contabilidade de gestão ambiental – CGA: representa assim uma abordagem combinada que facilita a transição de informação da contabilidade financeira e da contabilidade de custos para aumentar a eficiência de materiais, e reduzir o impacto e risco ambiental. Enquanto a contabilidade gerencial ambiental - CGA – visa atender aos usuários internos, como administradores, diretores e acionistas, a contabilidade financeira ambiental - CFA – visa atender à necessidade dos stakeholders, como acionistas, fornecedores, governo, bancos, financeiras, comunidade, dentre outros. (KRAEMER, 2004, p.150-157)

Contabilidade Financeira Ambiental: está relacionada com a preparação dos demonstrativos contábeis, baseadas nos princípios fundamentais da contabilidade. Nesse contexto, objetiva a mensuração, a avaliação e a informação das responsabilidades ambientais do ponto de vista financeiro e sua divulgação aos usuários. Esta contabilidade tem o foco particular em relatar custos, despesas e passivo ambiental da responsabilidade empresarial e outras questões ambientais significativas. (KRAEMER, 2004, p.150)

4.4 - A contabilidade e a relevância dos itens econômico-financeiro ambiental

A segregação das informações econômico-financeira de natureza ambiental passou a ser de fundamental relevância nas demonstrações contábeis das empresas, tendo em vista que envolvem recursos monetários significativos e influenciam no processo decisório. Essas informações sobre os investimentos, obrigações decorrentes de falhas preventivas e o consumo de recursos utilizados para corrigir os impactos orientam os usuários das demonstrações contábeis sobre a postura da empresa em relação à questão ambiental.

A contabilidade tem o desafio de tratar das questões econômico-financeiras e demonstrar a necessidade da existência do equilíbrio entre a atividade econômica e a exploração dos recursos naturais, contribuindo para o desenvolvimento sustentável**,** evidenciado a importância do respeito a capacidade de renovação dos recursos da natureza.

Desenvolvimento sustentável pode ser conforme Ribeiro (2003:6):

> *"Um processo de mudança no qual a exploração de recursos, o direcionamento de investimentos, a orientação do desenvolvimento tecnológico e as mudanças institucionais estão em harmonia e valorizam o potencial, presente e futuro, para a humanidade satisfazer suas necessidades."*

Ou ainda como afirma GUILHERME (2001, p.24):

> *"Este termo, foi criado em 1897, pela ex-ministra da Noruega, Sra. Gro Harlem Brudthand. O termo sustentável foi empregado com o objetivo de demonstrar que este novo processo de desenvolvimento despertaria mundialmente a necessidade de descobrir novas formas de alternativas, em todos os processos de produção, de prover os escassos recursos naturais atuais, e no futuro, utilizáveis. É, portanto, um instrumento que beneficia a sociedade quanto ao uso dos recursos naturais para gerações futuras".*

A segregação das informações ambientais poderá permitir a análise dos investimentos realizado na área ambiental por meio da comparação do montante de recursos aplicados e dos benefícios obtidos em termos de redução ou eliminação de poluentes, ou ainda pela confrontação dos gastos para eliminação de falhas operacionais e o total das penalidades que seriam devidas caso o meio ambiente fosse agredido.

A Contabilidade tem instrumentos para demonstrar a relação da empresa com o meio ambiente, como os investimentos em bens e direitos de natureza ambiental utilizados em medidas preventivas e de recuperação, os passivos ambientais em decorrência dos investimentos, do cumprimento de penalidades provenientes de agressões a ambiência, a segregação dos recursos consumidos em cada período visando a redução de poluentes, ou em decorrência do cumprimento de procedimentos burocráticos para obtenção dos certificados e atendimento a legislação ambiental.

O acompanhamento e análise de todas as atividades relacionadas ao gerenciamento ambiental, a visualização dos benefícios decorrentes dos gastos incorridos, permite à contabilidade elaborar demonstrações sobre o desempenho econômico-financeiro da empresa e evidenciar os riscos de descontinuidade caso os aspectos ambientais sejam desconsiderados.

Desta forma a contabilidade passa a ter papel fundamental na redução e eliminação dos custos da empresa e na otimização de seus resultados. Os resultados de cada período possibilitam analisar o total de recursos consumidos comparativamente aos resultados obtidos em termos econômicos, financeiros, ambientais e sociais. Permitindo também a análise do comportamento dos custos durante o período de duração, o que permite visualizar a evolução dos esforços empreendidos e sua eficácia.

As informações de natureza econômico-financeira são relevantes para avaliar a situação da empresa e a perspectiva de continuidade, assim como, interagir com o processo do desenvolvimento econômico sustentável.

4.5 – Demonstrações contábeis aplicadas a ambiência

Utilizando-se algumas demonstrações contábeis exigidas pela Lei 6404/76, as entidades podem demonstrar, através da sua implementação, seu comprometimento com o meio ambiente, as medidas tomadas pela sua gestão, os investidores e o seu resultado ambiental. Quanto às formas de divulgação no âmbito da contabilidade há duas linhas de pensamentos: uma propõe a implementação de um novo relatório apenso às demonstrações contábeis, tratando somente das

questões ambientais; e a outra sugere a inclusão destes dados nas atuais demonstrações mantendo o padrão já utilizado, mas apresentando contas e notas explicativas específicas.

Entendemos que a segunda alternativa seria a mais adequada (pelo menos a curto prazo) já que se trata de aspectos inseridos no contexto operacional das empresas e, inclusive, para atender a necessidade imediata de a contabilidade estar melhor informando seus usuários sobre o real valor patrimonial das empresas, sem maior perda de tempo em um processo de elaboração e implantação de um novo relatório. No anexo E consta às demonstrações contábeis do balanço patrimonial ambiental e demonstração do resultado do exercício. (RIBEIRO, 1992, p. 129)

4.5.1 – Lançamentos contábeis e balancete de verificação

As contas patrimoniais e de resultado na área ambiental surgiram devido à ausência de registros contábeis dos fatos geradores ocorridos continuamente nas empresas. A partir desta necessidade surgiram planos de contas ambientais, que organizaram de modo seqüenciais os grupos de contas a serem utilizadas para a escrituração contábil ambiental.

Graças ao empreendedorismo de empresas preocupadas quanto aos impactos ambientais decorrentes de suas atividades produtivas, e a ação marcante de vários agentes da sociedade como grupos ambientais – Greenpeace – e organizações não-governamentais permitiram o surgimento de ações mais fortes do governo com multas procurando solucionar os problemas ambientais.

Os lançamentos contábeis no livro diários desta empresa apresentaram contas como: despesas de treinamento, livros técnicos ambientais, equipamentos e materiais ambientais, empréstimos e financiamentos ambientais, fornecedores de insumos ambientais, dentre outros. O balancete de verificação é um demonstrativo contábil que relaciona as contas patrimoniais e de resultados ambientais. É comum haver a ocorrência de ajustes a esse balancete, inclusive para atender aos princípios fundamentais de contabilidade, o autor apresenta alguns ajustes que podem ocorrer num balancete de verificação. (KRAEMER, 2004, p.65-73)

Ajustes sobre os salários ambientais em que incidem os encargos sociais (INSS, FGTS, Férias, 13º salário, etc); depreciação de equipamentos ambientais; amortização de despesas de treinamento; redução de custos com matérias-prima.

4.5.2 - Planos de contas ambientais

É o conjunto de contas que deve ser utilizado por uma entidade (empresa) na contabilização de suas operações, ou seja, de todos os fatos contábeis mensuráveis monetariamente, a fim de auxiliar o contador no uso em potencial deste instrumento, atendendo as exigências dos usuários de informação contábil. Sendo um conjunto estruturado de contas numeradas ou codificado racionalmente, apresentado nos anexos C e D, o que permite seu uso em sistemas de contabilidade manuais, mecânicos, ou sistemas contábeis eletrônicos.

Apresentam contas sintéticas dos grupos de contas do ativo, passivo e patrimônio líquido, receitas, despesas; como também, contas analíticas, como por exemplo: clientes Ambientais; Participações em Fundos de Investimentos Ambientais; Equipamentos Ambientais, Provisões para depreciação e Exaustão Ambientais Acumuladas, Empréstimos e Financiamentos, Obrigações, Multas por danos ambientais, Indenizações por danos ambientais, reservas de Lucros, receitas ambientais, dentre diversas contas existentes.

4.5.3 – Contas patrimoniais e de resultado

A contabilidade ambiental propõe evidenciar a relação das empresas com o meio ambiente. Isto é proposto através de inúmeras demonstrações contábeis, que tem em sua origem nos registros contábeis, em contas patrimoniais e de resultado. As contas patrimoniais mais utilizadas são agrupadas em dois grupos: ativos ambientais e passivos ambientais. O ativo ambiental é formado pelo ativo circulante e permanente. O ativo circulante é formado pela conta estoque; o ativo permanente é formado pelas contas diferido, imobilizado, provisão para

desvalorização, depreciação e goodwill. Quanto a este último, o seu valor, numa conceituação moderna, corresponde à diferença entre o valor atual de toda a empresa, ou seja, sua capacidade de geração de lucros futuros, e o valor econômico de seus ativos apresentando, portanto, uma característica residual. (Monobe, 1986, p.65)

Já o passivo ambiental é formado pelas despesas do exercício atual, resultado de exercícios anteriores, risco ambiental potencial e provisões para contingências. As contas de resultado mais utilizadas são: Receitas Ambientais, Custos Ambientais, Despesas ambientais e Perdas Ambientais.

4.5.3.1 – Ativos ambientais

Com a necessidade de mensurar os bens tangíveis e intangíveis relacionados a ambiência, podemos definir ativos ambientais como sendo recursos econômicos controlados com a finalidade de gerar benefícios para a empresa e o meio ambiente onde está inserida.

Segundo Santos, Silva e Souza (2001:91):

> *"São considerados ativos ambientais todos os bens e direitos destinados ou provenientes das atividades de gerenciamento ambiental, podendo estar na forma de capital circulante ou capital fixo. O capital circulante (capital de giro) é o montante aplicado para a realização da atividade econômica da empresa, sendo composto pelas disponibilidades e pelo ativos realizáveis a curto e longo prazo."*

Pode-se ainda dizer que eles são recursos econômicos controlados por uma entidade, como resultado de transações ou eventos passados, dos quais se espera obter benefícios econômicos futuros, e que tenham, por finalidade, o controle, preservação e recuperação do meio ambiente.

Os ativos ambientais representam os estoques dos insumos, peças, acessórios etc., utilizados no processo de eliminação ou redução dos níveis de poluição e de geração de resíduos, os investimentos em máquinas, equipamentos, instalações etc., adquiridos ou produzidos com intenção de amenizar os impactos causados ao meio ambiente, os gastos com pesquisas, visando ao desenvolvimento de tecnologias modernas, de médio e longo prazo, desde que constituam benefícios ou ações que irão refletir nos exercícios seguintes. (KRAEMER, 2004, p.176)

As características dos ativos ambientais são diferentes de uma organização para outra em virtude de apresentarem processos operacionais distintos, por exemplo: uma refinaria de petróleo tem um processo complexo enquanto uma fábrica de refrigerantes seu processo é simples. O ativo ambiental é classificado em ativo circulante e permanente.

Os ativos ambientais são formados pelos estoques de insumos, equipamentos e acessórios, investimentos, ou seja, todos os bens e direitos destinados a gerar benefícios à empresa e que tenham relação com a preservação e recuperação da ambiência. Assim como o imobilizado utilizado diretamente no processo de produção ou aqueles não ligados diretamente a produção, como por exemplo aqueles com o objeto de proteger e recuperar a ambiência. Serão demonstrados os gastos com pesquisa e desenvolvimento de projetos que poderão gerar resultados futuros, ou seja, em mais de um exercício, como por exemplo, investimentos no desenvolvimento de projeto para produtos ambientalmente corretos.

Portanto, podemos destacar como ativos ambientais as aplicações de recursos em atividades relacionadas com o controle da poluição e tecnologias para preservação e diminuição dos impactos ambientais, devendo-se levar em consideração que as características desses ativos são diferentes de uma empresa para outra, devido os processos operacionais distintos da atividade econômica.

Os ativos ambientais são bens e direitos, os quais possuem diferentes prazos de realização, e que para cada tipo de empresa haverá um plano de conta especifico, com estrutura para registrar e permitir comparações com as transações realizadas.

Desta forma se destaca o papel do profissional da contabilidade, que além dos métodos já utilizados, deverá utilizar-se de métodos gerenciais e critérios diferenciados para classificação dos bens tangíveis e intangíveis da empresa de forma a permitir aos usuários das informações melhor avaliação das ações ambientais da empresa.

4.5.3.2 – Passivo Ambiental

O passivo ambiental fica amplamente conhecido pelas empresas que os possuem agridem significadamente o meio ambiente e, dessa forma, têm que pagar vultuosas quantias a título de indenização a terceiros, de multas e para a recuperação de áreas danificadas (RIBEIRO, Egratão, 2000). Devemos ressaltar que os passivos ambientais também apresentam conotações positivas, quando são usadas em atitudes ambientalmente responsáveis. Exemplo, na manutenção de sistema de gerenciamento ambiental, os quais requerem pessoas para sua operacionalização. Tais sistemas exigem ainda a aquisição de insumos, máquinas, equipamentos, instalações para o seu funcionamento. Tais investimentos podem ser financiados por fornecedores ou por meio de instituições de créditos, como Banco do Nordeste e Banco do Brasil.

Os passivos ambientais podem ter como origem qualquer evento ou transação que reflita a interação da empresa com o meio ecológico, cujo sacrifício de recursos econômicos se dará no futuro como aquisição de insumos que serão inseridos no processo operacional para que este não produza resíduos tóxicos, despesas de manutenção e operação de "departamento" de gerenciamento ambiental inclusive mão-de-obra, gastos para recuperação e tratamento de áreas contaminadas (máquinas, equipamentos, insumos), pagamento de multas por infrações ambientais, gastos para compensar danos irreversíveis, inclusive os relacionados à tentativa de reduzir o desgaste da imagem da empresa perante a opinião pública (KRAEMER, 2004, p.180). Um ponto em que está pouco avançado no que se refere à Contabilidade ambiental é a mensuração ou quantificação dos resultados alcançados pelo Sistema de Gestão Ambiental de determinados elementos patrimoniais os quais são considerados intangíveis. Por exemplo, valor de um patrimônio florestal, ou de um certificado de qualidade ambiental (ISSO 14000). Até que ponto pode atribuir a esta certificação o aumento do faturamento, lucratividade e competitividade da empresa que a conquista?

O objeto da mensuração é o evento econômico e este deve representar uma decisão tomada. As decisões mais importantes relacionadas com a atividade do meio ambiente estão vinculadas a aspectos operacionais, econômicos e financeiros dessa gestão. A mensuração desses eventos deve ter capacidade de estabelecer como unidade de mensuração a unidade monetária; permitir a avaliação dos ativos pelos benefícios futuros que ele poderá propiciar à entidade; permitir a previsão de passivos efetivos pelo valor presente da dívida; permitir a previsão de passivos contingentes, pelo valor presente da expectativa de restrições futuras sobre os ativos; que a moeda esteja isenta das variações em seu poder, ou seja, use taxas de inflação que possam corrigir isso; evidenciar os efeitos de tempo sobre os capitais, aplicando as taxas de juros pertinentes a cada evento; estabelecer um preço de transferência para os serviços prestados ou produtos gerados que venham ser entregue a terceiros; nesse caso, deve-se utilizar o preço de mercado a vista por um serviço de igual qualidade e especificação; que, para os consumos efetivos, seja utilizado o menor preço de mercado por seu valor a vista, para especificação e qualidade iguais. (FERREIRA, 2003, p.53)

4.5.4 – Receitas Ambientais

Segundo Santos, Silva e Souza (2001:92*):*

> *"O objetivo principal da implantação da gestão ambiental não é gerar receita para a empresa, e sim, desenvolver uma política responsável acerca dos problemas ambientais".*

Na contabilidade o que se entende por receitas ambientais, seria o lucro obtido na venda de resíduos e sucatas que normalmente seriam descartadas ao lixo, se a empresa não negociasse. A empresa também pode gerar receita através do processo de reciclagem, elaborando produtos ou gerando matéria-prima para venda ou uso próprio.

É importante lembrar que os consumidores cada vez mais exigentes estão procurando produtos de empresas com responsabilidade social, ecológica e ambiental o que leva a empresa que tem uma política de gestão ambiental a gerar mais receita. Esta seria uma receita ambiental intangível e de difícil mensuração, visto que o valor da empresa é aumentado devido ao seu compromisso com as causas ambientais.

4.6 – Gastos (custos e despesas) ambientais

Conforme Santos, Silva e Souza (2001:93):

> *"Custos e despesas ambientais são gastos (consumo de ativos) aplicados direta ou indiretamente no sistema de gerenciamento ambiental do processo produtivo e em atividades ecológicas da empresa. Quando aplicados diretamente na produção, estes gastos são classificados como custo, e se forem aplicados de forma indireta são chamados de despesas".*

Desta forma os custos ambientais estão associados ao gerenciamento de uma gestão responsável dos impactos da atividade da empresa no meio ambiente e relacionado com a proteção e preservação desse meio ambiente, assim como custos incorridos para atender os objetivos e exigências da legislação.

As despesas ambientais podem ser consideradas como os gastos relacionados ao gerenciamento ambiental, incorridos na área administrativa, como recursos humanos, área de compras, financeiro, ou seja, áreas que desenvolvem atividades ligadas á proteção do mesmo. As despesas devem ser apropriadas no período de ocorrência, obedecendo-se o principio de competência.

Na visão de Ribeiro (2005:55), os custos e despesas não poderão ser associados a um único processo produtivo, e havendo dificuldades em determinar com precisão o exato período de competência, a contabilidade deve utilizar-se dos instrumentos tradicionais para alocar os custos e despesas nos períodos de forma segregada.

É necessário estudo e análise para uma adequada identificação quanto à segregação dos custos em operacionais e não operacionais, visto que em alguns casos a identificação dos custos ambientais torna-se difícil, uma vez que, os gastos ambientais além dos benefícios à natureza, também trazem melhoria e eficiência nas operações da empresa.

4.7 - A evidenciação do resultado das informações contábeis ambientais

À contabilidade cabe o papel de divulgar os eventos econômicos ou fatos contábeis relativos às atividades da empresa por meio dos demonstrativos publicados, e relatando os principais fatos não evidenciados nos demonstrativos contábeis tradicionais.

Nos lembra Ribeiro (2005:141):

> "*Em relação a divulgação das informações ambientais há duas linhas de pensamento, uma que propõe a implementação de um novo relatório apenso às demonstrações contábeis, o qual trataria somente das questões ambientais; e outra que sugere a inclusão desses dados na atuais demonstrações, mantendo o*

modelo já utilizado, mas apresentado contas e notas explicativas específicas".

A implementação de novos relatórios ainda vai depender de pesquisas e estudos por profissionais da auditoria e especialistas em contabilidade ambiental no desenvolvimento de novas técnicas e métodos de trabalho, além de regulamentação e normas.

No momento a contabilidade dispõe das demonstrações contábeis tradicionais, além das notas explicativas, relatórios da administração ou quando necessário a empresa poderá utilizar outros meios de divulgação.

Notas Explicativas - As notas explicativas devem exercer um papel de complementação, detalhamento de fatos existentes nas demonstrações financeiras. Quando houver casos de difícil mensuração de valores de custo, estes deverão ser detalhados na nota explicativa. Além de critérios quanto a avaliação de estoques ambientais, depreciação, ativação de gastos, apropriação de despesas e valor do lucro destinado a utilização no meio ambiente.

Relatório da Administração - Através do relatório da administração a empresa pode expressar a sua posição em relação às políticas ambientais adotadas, assim como objetivos alcançados. Informações importantes, principalmente aos interessados na situação financeira, visto que, as ações da empresa, em relação ao meio ambiente podem causar grande influência nos lucros futuros.

A empresa também poderá utilizar-se de outros meios para divulgar suas ações em relação ao meio ambiente, como boletins informativos e divulgação através da rede mundial de computadores, direcionado a todos os interessados, e tendo como objeto a imagem da empresa junto ao acionista e a sociedade em geral.

5 – CONSIDERAÇÕES FINAIS

Com a crescente conscientização ambiental dos consumidores e às exigências dos investidores e acionistas da empresa cada vez mais interessados na forma pelas quais as operações da empresa estão sendo conduzidas em relação à questão ambiental, está influenciando em uma mudança de comportamento das empresas e fazendo com que o conceito de administração se amplie, incluindo a variável ambiental.

O mercado não mais aceita os descasos no tratamento dos recursos naturais. A legislação torna-se mais rígida, obrigando as empresas a encarar com seriedade e responsabilidade a variável ambiental em sua estratégia operacional e na tomada de decisão.

A contabilidade encontra-se envolvida nesse contexto como principal instrumento de comunicação da empresa com a sociedade. Desde seu surgimento, a contabilidade teve como função primordial o acompanhamento das atividades econômicas, evidenciando a situação patrimonial das empresas aos principais usuários das informações contábeis, administrativas, econômicas, sociais e ambientais, acompanhando sua evolução no decorrer do tempo.

A evolução e diferenciação das necessidades dos usuários fizeram surge a contabilidade ambiental com o objetivo principal de mensurar e evidenciar as transações que afetam ambiência, através da criação de parâmetros de mensuração e registro que permitam o acompanhamento da convivência da empresa com o meio ambiente e a evolução econômica e patrimonial de tal relação.

Desta forma a contabilidade cumpre seu papel no sentido de auxiliar as empresas nas suas tomadas de decisão, gerando informações de integração entre o desenvolvimento econômico e a manutenção da qualidade de vida, cumprindo assim a sua função social refletindo as ações da empresas com relação ao meio ambiente e contribuindo para uma gestão ambiental responsável.

Conhecer a ambiência e a amplitude das conseqüências de sua exploração torna-se de fundamental importância para que os gestores avaliem quais danos ambientais sua empresa está ocasionando e como estes danos podem ser prevenidos e seu impacto amenizado.

Por esse e outros aspectos a contabilidade é uma ferramenta essencial para a identificação e mensuração dos ativos, passivos, receitas e despesas ambientais, com o objetivo de alertar para

a necessidade de programas de controle, recuperação e preservação do meio natural onde elas, as empresas, estão inseridas.

No tocante a confirmação ou não das hipóteses, pode-se agora após ter verificado a literatura existente e a nossa disposição afirmar que, ambas se confirmam tanto a de que a contabilidade pode e deve ser usada na mensuração das informações econômicas financeiras das entidades, no intuito de aprimorar seus resultados, assim como a segunda servindo de parâmetro no sentido de que através dela pode-se verificar os registros e os reflexos do cumprimento das legislações impostas a cada situação na qual as entidades estejam inseridas, assim como servindo de aparato para aqueles que se interessam por tais informações pois através delas também se poderá inferir perspectivas de lucrabilidades futuras a depender do grau de conscientização dos consumidores diretos e indiretos, ou mesmo dos mercados onde estas empresas atuam.

Portanto, o meio ambiente é uma riqueza que precisa ser preservada e as empresas que tiverem o devido cuidado que o assunto requer, tendem a melhorar seu desempenho econômico e financeiro e aumentar seu valor de mercado simplesmente pelo fato de bem cuidar do planeta, desta forma evitando punibilidades por parte de órgãos competentes e acima de tudo por quem conscientizado a cada dia poderá ser fator predominante na continuidade das operações das entidades, um senhor todo poderoso e que pode se quiser levar a falência qualquer empresa, e este é o cliente de cada corporação. Pois mais importante que legislar sobre a matéria ambiental seria a educação promulgada vastamente acerca do assunto, e isto aclara uma certeza mais do que leis precisamos de cidadãos politicamente corretos e cumpridores do seu dever social.

6 - REFERÊNCIA BIBLIOGRÁFICA

BARBIERI, José Carlos. **Desenvolvimento e Meio Ambiente.** São Paulo: Vozes, 1997.

CASTRO, C. de M. **A Prática da Pesquisa. São Paulo**: Me Graw – Hill,1977.

CHIZZOTTI, A. **Pesquisa em Ciências Humanas e Sociais**. 5ª ed. São Paulo: Cortez, 2001.

Constituição Federal: promulgada em 05 de outubro de 1988, atualizada em 1994.

FERREIRA, Araceli de S. Contabilidade Ambiental. **Uma Informação para o Desenvolvimento Sustentável**. São Paulo: Atlas, 2003.

FERREIRA, Aracéli Cristina de Sousa. **Contabilidade Ambiental – Uma Informação para o Desenvolvimento Sustentável**. São Paulo: Editora Atlas, 2003.

FRANCO, Hilário. **A Contabilidade na Era da Globalização**. São Paulo: Editora Atlas, 1999.

IUDÍCIBUS, Sérgio de. **Manual de Contabilidade das Sociedades por Ações**. São Paulo: Atlas, 1987.

IUDÍCIBUS, Sérgio de. **Teoria da Contabilidade /** Sérgio de Iudícibus- 7ª ed.- São Paulo: Atlas, 2004.

KRAEMER, Maria Elisabeth Pereira. TINOCO, João Eduardo Prudêncio. **Contabilidade e Gestão Ambiental**. São Paulo: Atlas, 2004.

LAKATOS, E. M. e MARCONI, M de A. **Fundamentos de Metodologia Científica**. São Paulo: Atlas, 1998.

LUCARELLI, Fábio Dutra. **Responsabilidade civil por dano ecológico. Revista dos Tribunais**. São Paulo, ano 83, v. 700, p. 7-26, fev. 1994.

MACHADO, Paulo Affonso Leme. **Direito ambiental brasileiro**. São Paulo: Malheiros Editores, 1992.

MAZON, Rubens. **Em direção ao novo paradigma de gestão ambiental: tecnologias limpas ou prevenção de poluição**. Revista de administração de empresas, são Paulo: 1979.

MONOBE, M. **Contribuição à mensuração e contabilização do *goodwill* não adquirido**. 1986. 183 f. Tese (Doutorado em Ciências Contábeis). Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, São Paulo.

PAIVA, Paulo Roberto de. **Contabilidade ambiental evidenciação dos gastos ambientais com transparência e focada na prevenção**. São Paulo: Atlas, 2003.

Revista do Conselho Regional de Contabilidade do Rio Grande do Sul. A Contabilidade e os Contadores a Serviço do Judiciário. **Proposta para Avaliação de Ativos Ambientais**. Maio – 2003.

RIBEIRO, Osni Moura. **Contabilidade Básica**. São Paulo: Saraiva, 1998.

RIBEIRO, Maisa de Souza. **Contabilidade e Meio Ambiente**. Dissertação de Mestrado, FEA/USP, 1992.

SILVA, Moacyr de Lima. **Contabilidade Geral**. São Paulo: Érica, 1998. CRCSC & VOCÊ/ Conselho Regional de Contabilidade de Santa Catarina. – v.1, n.1, Florianópolis: CRCSC, 2001.

TINOCO, J. E. P., KRAEMER, M. E. P. **Contabilidade e Gestão Ambiental**. São Paulo: Atlas, 2004.

ÁREA V

CRECIMIENTO ECONÓMICO

CRESCIMENTO ECONÓMICO

LA EVOLUCIÓN DE LA PARIDAD DE PODER DE COMPRA EN LA UE-27

ALEJANDRO RODRÍGUEZ CARO
E-mail: arcaro@dmc.ulpgc.es
SANTIAGO RODRÍGUEZ FEIJOÓ
CARLOS GONZÁLEZ CORREA
Departamento de Métodos Cuantitativos en Economía y Gestión
Universidad de Las Palmas de G.C.

Resumen

En el presente trabajo se estudia la evolución hacia un mercado europeo común desde la perspectiva de la capacidad de compra, debido a la importancia que los posibles desajustes en la misma pueden provocar en un mercado único, como pone de manifiesta el hecho de la convergencia en precios como uno de los cinco objetivos establecidos en el tratado de Maastrich. Con este objetivo se han analizado las paridades de poder de compra ofrecidos por Eurostat para el conjunto de la Unión Europea de los 15 y 27 Estados, utilizando modelización dinámica a través de kernels estocásticos.

Palabras clave: Paridad de poder de Compra, Convergencia, Unión Europea.
Área Temática: Economía del Crecimiento económico.

1. INTRODUCCIÓN

La Teoría de la Paridad de Poder de Compra (PPC) fue enunciada formalmente por Gustav Cassel en 1918. Siguiendo a este autor, La Teoría de la Paridad de Poder de Compra se cumple cuando "la tasa de cambio de equilibrio iguala el poder adquisitivo de una moneda en el interior (de un país) con lo que dicha moneda podría comprar en el exterior si fuera cambiada por una moneda extranjera" [citado en Spiegel (1996), p. 690]. Del enunciado de esta teoría, se deduce el carácter de comparación territorial de la misma en términos de precios, incluyendo tanto el precio interior en cada uno de los territorios que se compara como el precio mediante el cual se interrelacionan, esto es, el tipo de cambio.

En las últimas décadas el estudio de la PPC se ha hecho, si cabe, más relevante debido a varios factores. Uno de ellos se debe a que en el mundo desarrollado cada vez se acepta con mayor nivel de generalidad que la política monetaria debe tener como objetivo la estabilidad de los precios, debido a que sus movimientos pueden producir desequilibrios importantes al originar cambios en las decisiones de producción y consumo. Un reflejo de esta factor lo tenemos al observar la importancia que la Comunidad Europea le otorga al objetivo de estabilidad de precios. Otro factor está relacionado con el proceso de globalización en el cual se encuentra inmersa la Economía. En este sentido, el proceso de globalización ha disparado el volumen de intercambios entre países, con lo cual también se ha multiplicado el riesgo de los flujos monetarios debido a los movimientos de los precios, tanto internos como externos. El incremento de estos riesgos puede provocar nuevamente que el mercado no funcione adecuadamente y realice una asignación de recursos ineficientes. Un tercer factor se encuentra en la declaración de intenciones que explícitamente enuncia la Comunidad Europea en su Tratado Constitutivo, cuando establece que para alcanzar una calidad de vida elevada y una gran cohesión social es fundamental el control de la Paridad de Poder de Compra entre sus países miembros. En este caso, no solo es importante la estabilidad de los precios, sino que se hace necesario definir una política que garantice la convergencia de la Paridad de Poder de Compra de todos los países miembros. Esto supone que, si en el mercado que configura el conjunto de la Comunidad Europea siguen conviviendo distintas monedas, la PPC tenderá a medio plazo a complementar y, al final, a sustituir al uso de los Índices de Precios de Consumo.

En este trabajo se usan los datos de las paridades de poder de compra y el estudio de su convergencia se ha realizado a través del análisis de la dinámica de la distribución. Para ello, se han estimado los correspondientes kernels estocásticos, así como, la aplicación de la técnica de cadenas de Markov con el fin de caracterizar la evolución a largo plazo de esta variable, mediante la correspondiente distribución ergódica. Esta metodología, propuesta por Quah (1993a), (1993b), (1994), (1996a), y (1996b) para la distribución de ingresos entre distintos países o regiones, se aplica al estudio temporal de la paridad de poder de compra.

2. RESULTADOS INFLACIÓN

Comenzamos este trabajo con los índices de precios de consumo armonizados mensuales con base en el 2005, para el período Enero-2006 hasta Febrero-2008. A partir de los mismos se han obtenido las inflaciones interanuales para cada país y finalmente se ha restado estos valores al valor de la inflación para la UE25, para cada uno de los meses disponibles.

En la tabla 1 se presentan el porcentaje de meses en los que la inflación interanual de cada país se encuentra por debajo del valor de la UE25 para dicho período.

Tabla 1: Porcentaje de meses con inflación por debajo de media UE25

Pais	Porcentaje	Pais	Porcentaje
Belgium	85,07%	Luxembourg	52,24%
Bulgaria	21,64%	Hungary	1,49%
Czech_Republic	44,78%	Malta	68,66%
Denmark	91,79%	Netherlands	76,12%
Germany	97,01%	Austria	92,54%
Estonia	20,90%	Poland	33,58%
Ireland	35,82%	Portugal	44,03%
Greece	23,88%	Romania	0,00%
Spain	38,81%	Slovenia	6,72%
France	89,55%	Slovakia	21,64%
Italy	78,36%	Finland	99,25%
Cyprus	59,70%	Sweden	91,04%
Latvia	41,04%	United_Kingdom	90,30%
Lithuania	53,73%		

En la misma se puede observar claras diferencias entre los distintos países, algunos como Hungría, Rumanía y Eslovenia apenas tienen meses por debajo de la inflación media de la UE25, mientras que, en el otro extremo se encuentran países como Dinamarca, Alemania, Finlandia, Suecia, Reino Unido que no bajan del 90% de meses por debajo de la media de la UE25.

Ilustración 1: Estimación Kernel UE-25

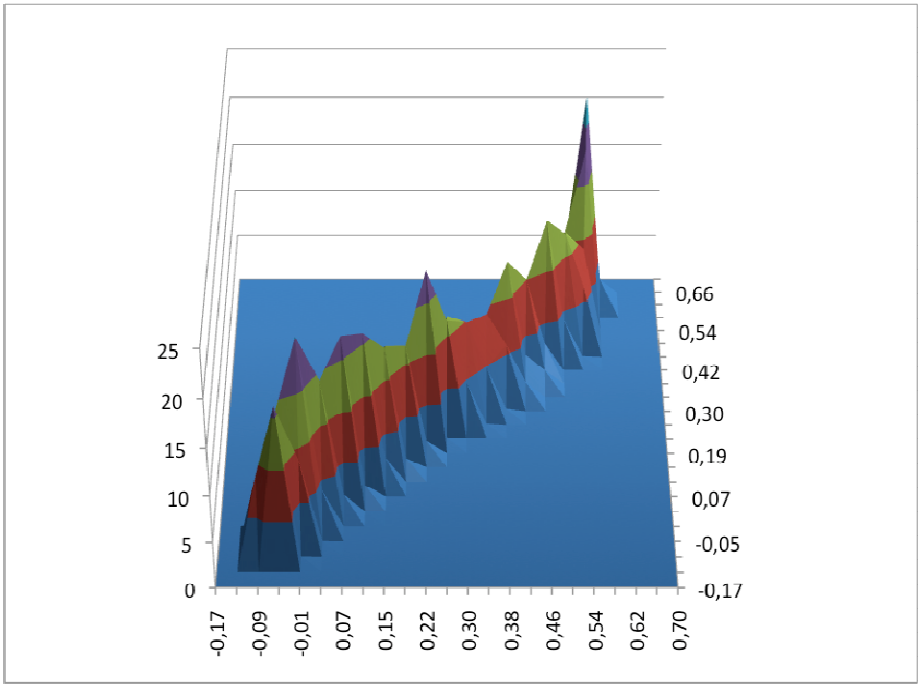
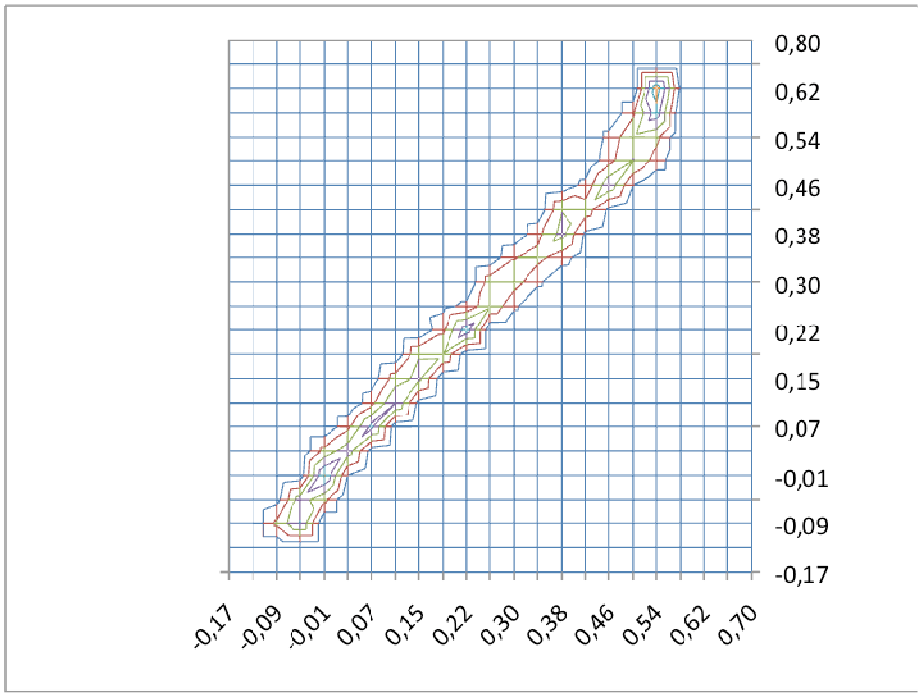


Ilustración 2: Estimación Kernel UE25



En la gráfica de estimación del correspondiente kernel para el conjunto de la UE se observa un claro comportamiento convergente. A pesar de encontrarse los picos de densidades en los extremos de la distribución, el "movimiento" esperado de la misma es hacia el interior de la misma. Esto se puede observar en una mayor densidad de probabilidad hacia valores inferiores en períodos futuros. Esta tendencia convergente se puede observar más claramente en la ilustración de las distribuciones ergódicas en la Ilustración 7, donde se muestra la evolución a largo plazo de la distribución de densidad bivariante de la Ilustración 1. Como se puede apreciar, la distribución presenta un único pico central, hacía donde tenderían los valores de la inflación interanual de cada uno de los países.

Además de presentar la estimación para el conjunto de los 25 países de la UE, se presentan las estimaciones para los países dentro de la UE antes de la última ampliación (UE15) en la Ilustración 3 E ilustración 4,así como para los diez países de reciente incorporación (UE16-27) en la Ilustración 5 y la Ilustración 6. Las estimaciones de las correspondientes distribuciones a largo plazo también se presentan en la

La diferencia fundamental entre los dos conjuntos de países residen en los rangos de valores de inflación dentro del período estudiado, siendo los países UE15 los mas estables, con valores muy cercanos a los valores medios de la UE. Los países de reciente incorporación presenta valores mas alejados, lo cual redunda también en la estimación de la distribución de largo plazo, donde también se presenta una mayor dispersión, aunque asimétrica hacia la derecha, es decir, hacia inflaciones mayores de la media de la UE.

Ilustración 3: Estimación Kernel UE15

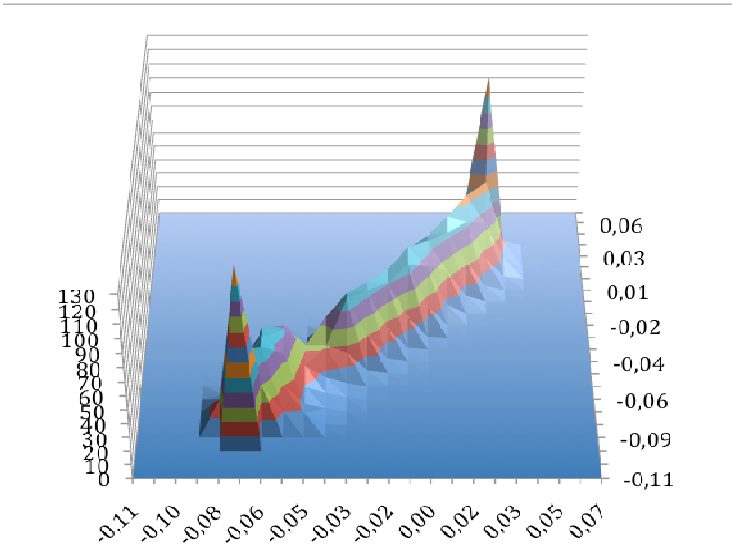


Ilustración 4: Estimación Kernel UE15

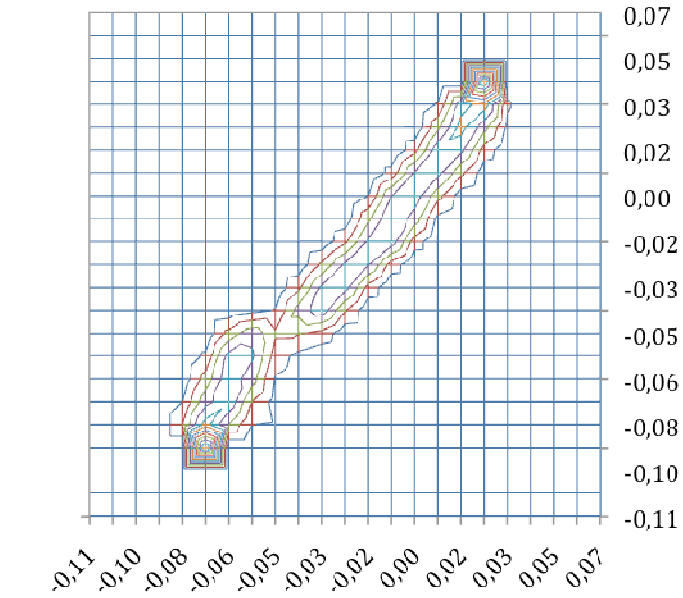


Ilustración 5: Estimación Kernel UE16-27

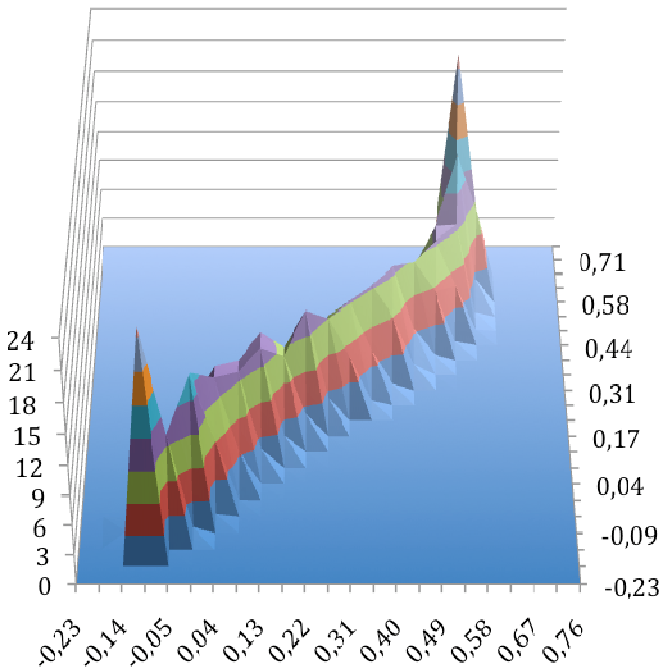


Ilustración 6: Estimación Kernel UE16-27

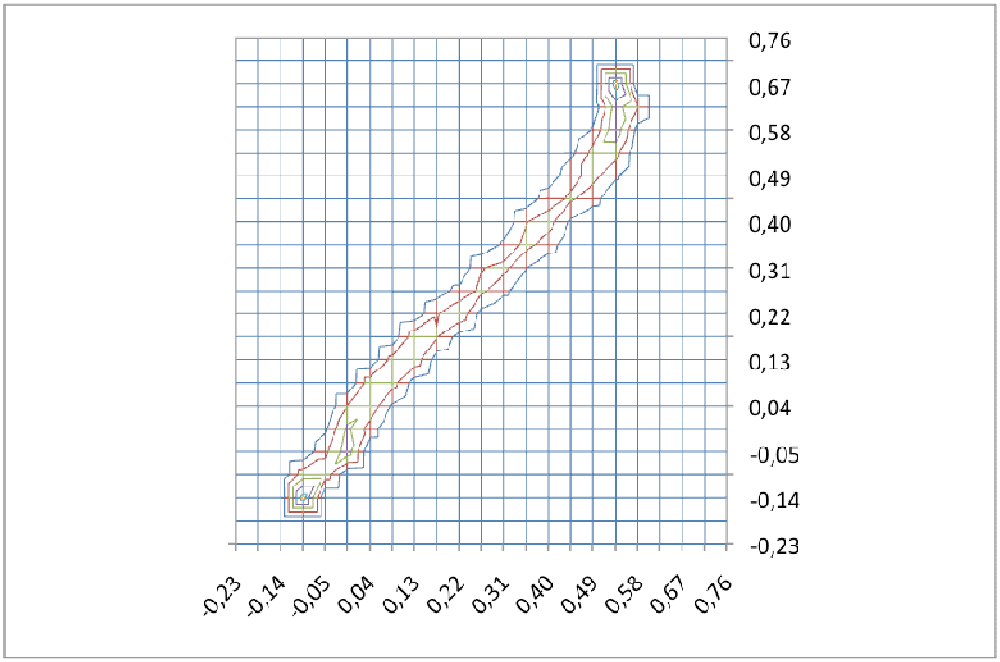
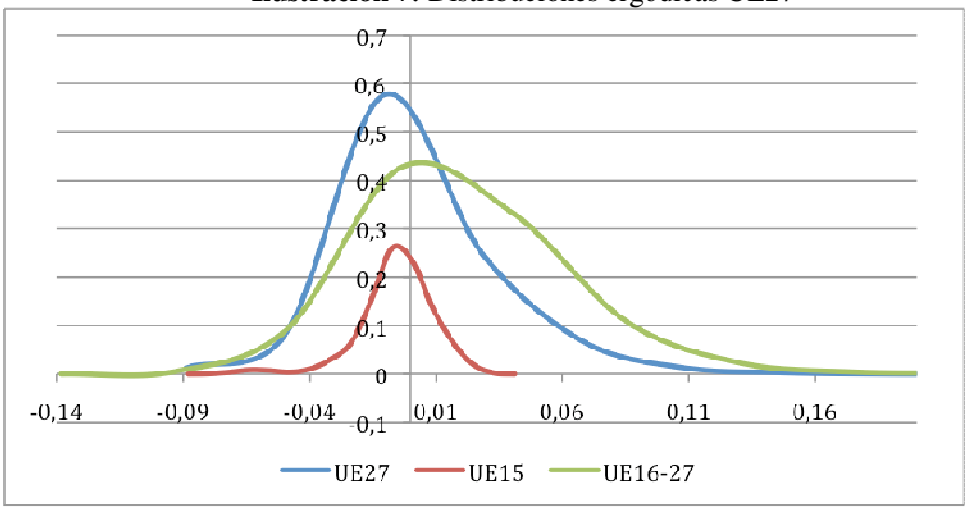


Ilustración 7: Distribuciones ergódicas UE27



Sin embargo, la convergencia en inflación (paridad relativa de poder de compra) no implica la convergencia en la paridad de poder de compra absoluta. No solo es necesario que las variaciones de precios sean convergentes, ya que la convergencia en precios es posible con variaciones divergentes. La convergencia depende de la situación relativa de precios entre los distintos países.

3. PARIDAD ABSOLUTA DE PODER DE COMPRA

En este apartado se han utilizado la serie de índices comparativos del nivel de precios. Serie de carácter anual disponible desde 1995 a 2006, tomando la media de la UE el valor 100 para cada uno de los años estudiados. Esta serie está calculada como el cociente entre la Paridad de Poder de Compra de un país, y el tipo de cambio nominal de la moneda contra el Euro. Son índices que están puestos en relación a otro país o grupo de éstos, en el caso que tratamos, el conjunto de la Unión Europea que toma valor 100 en el índice para cada instante del tiempo. La ventaja de esta serie está en comparar directamente los niveles de precios de los distintos países miembros, sin embargo, todavía no hay un número de años suficientes que permitan realizar estimaciones por grupos de países o períodos de tiempo.

La Tabla 2 nos muestra en porcentaje de años sobre el total disponible, en los que el índice comparativo de nivel de vida se encuentra por debajo de la media de la UE. En la misma se puede observar una alta persistencia de los países miembros, esperable dado el pequeño tamaño temporal de la que se dispone.

Tabla 2: Porcentaje de Años por debajo de la media UE

Pais	Porcentaje	Pais	Porcentaje
Belgium	0,00%	Luxembourg	0,00%
Bulgaria	100,00%	Hungary	100,00%
Czech Republic	100,00%	Malta	100,00%
Denmark	0,00%	Netherlands	0,00%
Germany	0,00%	Austria	0,00%
Estonia	100,00%	Poland	100,00%
Ireland	16,67%	Portugal	100,00%
Greece	100,00%	Romania	100,00%
Spain	100,00%	Slovenia	100,00%
France	0,00%	Slovakia	100,00%
Italy	75,00%	Finland	0,00%
Cyprus	100,00%	Sweden	0,00%
Latvia	100,00%	United Kingdom	16,67%
Lithuania	100,00%		

Los resultados de la estimación del kernel estocástico correspondiente a la transición de un año a otro, se puede ver en la Ilustración 8 y la Ilustración 9, mientras que la distribución de largo plazo asociada se muestra en la Ilustración 10. El comportamiento es muy similar al comentado para la inflación interanual, es decir, la forma de la distribución presenta dos picos de probabilidad en los extremos, mientras que el movimiento esperado de la distribución es hacia el medio de la misma. Esto provoca que la distribución esperada a largo plazo tienda a la convergencia.

Ilustración 8: Estimación Kernel 27

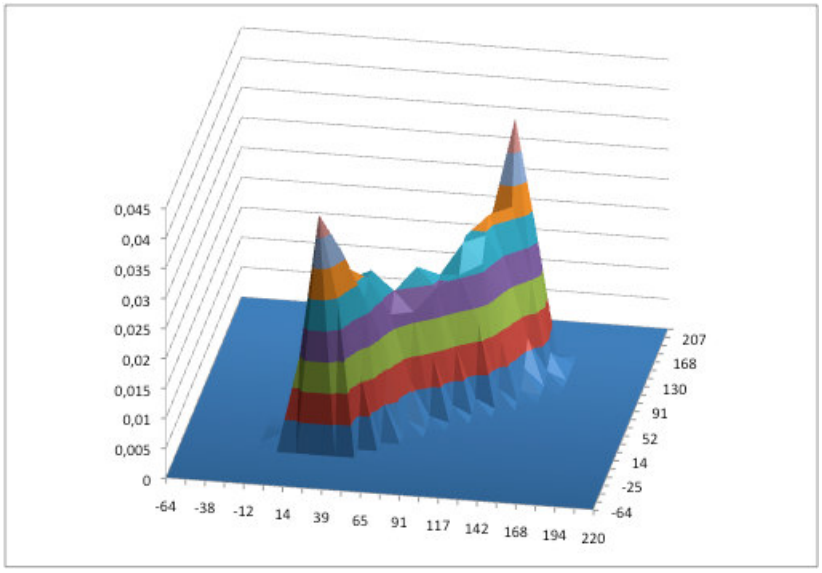
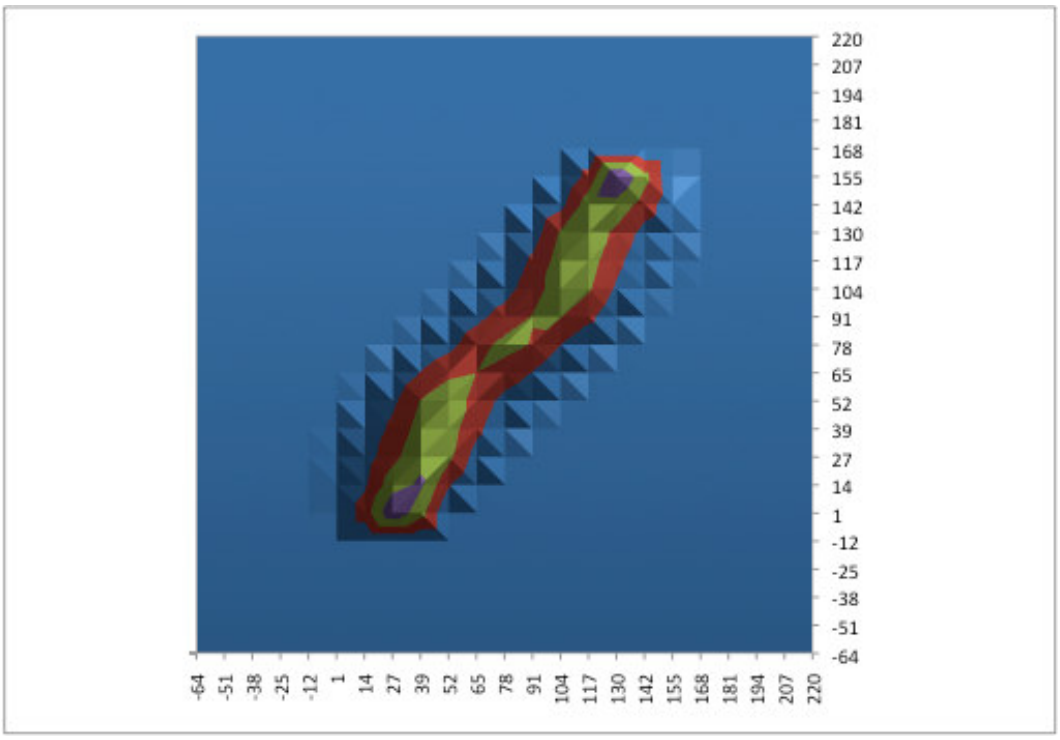
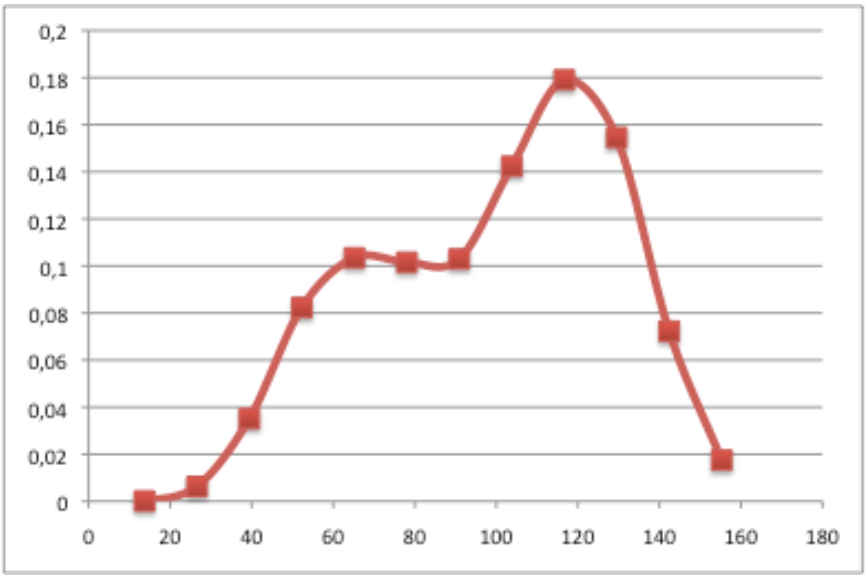


Ilustración 9: Estimación Kernel 27



Sin embargo, la distribución ergódica no presenta un solo pico de forma tan clara como la de la Ilustración 7, sino que se aprecia una pequeña sobre los valores 60-80.

Ilustración 10: Estimación distribución ergódica UE27



CONCLUSIONES

Durante los últimos diez años parece observarse un proceso de convergencia en los precios. En general, países con precios mas bajos han visto como su inflación superaba la media de la UE de tal forma que sus precios se tienden hacia los valores medios de la UE. A la inversa para aquellos países donde los precios superan la media de la UE, que mantienen inflaciones interanuales por debajo de la media de la UE.

BIBLIOGRAFÍA

ANDERSON T.W., GOODMAN, L.A. (1957). "Statistical Inference about Markov Chains". Annals of Mathematical Statistics 28 (1), 89-110

CANCELO, J., FERNÁNDEZ, A., RODRÍGUEZ, F., URRESTARAZU, I. y GOYENECHE, J. (2000). "Paridad de Poder de Compra en el Mercosur: Un Análisis a Partir de la Evolución a Largo y Mediano Plazo del Tipo de Cambio Real", Quantum, nº 11.

JOHNSON, P.A. (2000). "A nonparametric analysis of income convergence across the US states". Economics Letters 69, 219-223.

KULLBACK, S., KUPPERMAN, M., KU, H.H. (1962). "Test for Contingency Tables and Markov Chains". Technometrics 4 (4), 573-608.

LÓPEZ-BAZO, E., VAYÁ, E., MORA, A.J., SURIÑACH, J. (1999). "Regional Economic Dynamics and Convergente in the European Union". Annals of regional Science 33, 343-370.

Le GALLO, J. (2001). "Space-time analysis of GDP disparities among European regiones : A Markov chains approach". LATEC - Document de travail - Economie 2001-06.

MAGRINI, S. (1999). "The Evolution of Income Disparities Among the Regions of the European Union". Regional Science and Urban Economics 29, pp 257-281.

NEVEN, D., GOUYETTE, C. (1995). "Regional convergence in the European Community". Journal of Common Market Studies 33, 47-65.

PEKKALA, S. (1999). "Regional convergence across the finnish provincies and subregions, 1960-94". Finnish Economic Papers 12 (1), 28-40

QUAH, D.T. (1993). "Empirical cross-section Dynamics in Economic Growth". European Economic Review 40, 1353-1375.

QUAH, D.T. (1993). "Galton's fallacy and Tests of the Convergence Hypothesis". Scandinavian Journal of Economics 95, 427-443.

QUAH, D.T. (1996). "Regional Convergence clusters across Europe". European Economic Review 40, 951-958.

QUAH, D.T. (1996). "Empirics for Economic growth and convergence". European Economic Review 40, 1353-1375.

QUAH, D.T. (1996). "Convergence Empirics across economies with (some) capital mobility". Journal of Economic Growth 1, 95-124.

QUAH, D.T. (1997). "Empirics for growth and distribution: stratification, polarization, and convergence clubs". Journal of Economic Growth 2, pp. 27-59.

QUAH, D.T. (2000). "The TSRF Reference Manual". http://econ.lse.ac.uk/~dquah

RODRÍGUEZ, S. GONZÁLEZ, C. y RODRÍGUEZ, A. (2002), "La Teoría de la Paridad Relativa del Poder de Compra entre Territorios Económicos que Tienen una Misma Moneda: Una Aplicación a las Comunidades Autónomas Españolas", Estadística Española, 44, nº 150, pp. 229-256.

RODRÍGUEZ, S. RODRÍGUEZ, A. DÁVILA, D. (2003). "Methods for Quarterly Disaggregation Without Indicators; a Comparative Study Using Simulation". Computacional Statistics and Data Analysis 43, pp 63-78

RODRÍGUEZ, S., GONZÁLEZ, C. y RODRÍGUEZ, A. (2004), "¿Cómo Medir los Cambios en La Paridad de Poder de Compra a Partir de los Índices de Precios de Consumo y los Tipos de cambio", Estadística Española, nº 157.

TSIONAS, E.G. (2002). "Another Look at Regional Convergente in Greece". Regional Studies 36 (6), 603-609.

ANÁLISIS EMPÍRICO DE LA DISPERSIÓN DE PRECIOS EN BIENES HOMOGÉNEOS EN UN MERCADO FRAGMENTADO

SANTIAGO RODRÍGUEZ FEIJOÓ
E-mail: srfeijoo@dmc.ulpgc.es
CARLOS GONZÁLEZ CORREA
ALEJANDRO RODRÍGUEZ CARO
Dpto. de Métodos Cuantitativos en Economía y Gestión
Universidad de Las Palmas de Gran Canaria

Resumen

La dispersión de los precios es una variable que la literatura ha relacionado tanto con la cantidad de información que tienen los consumidores como con la capacidad que tienen las empresas para tener mayor poder del mercado. En este trabajo se analiza empíricamente la dinámica de precios que se encuentra detrás de la dispersión. La hipótesis que mantenemos es que la dispersión de los precios es el elemento que permite obtener el equilibrio entre la cantidad de información que está dispuesto a adquirir el consumidor con la capacidad que tienen las empresas para obtener mayor poder de mercado, mediante la diferenciación de los servicios adjuntos a la venta del producto y a las dificultades que puede generar para incrementar los costes de búsqueda de información del consumidor.

Palabras clave: Economía de la búsqueda, dispersión, precios, poder de mercado e IPC.
Área Temática: Economía del Crecimiento económico.

Abstract

The dispersion of prices is one variable that the literature has linked both with the amount of information consumers have, and with the ability of companies to have greater market power. This paper empirically examines the dynamics of prices that is behind that dispersion. Our hypothesis is that the dispersion of prices is the element that allows a balance between the amount of information consumer are willing to purchase with the ability of firms to obtain greater market power through differentiation of services attached to the product and the difficulties that can lead to consumers to a higher costs of finding information.

Thematic Area: Growth Economics.

1. INTRODUCCIÓN

La dispersión de precios y, en consecuencia, el incumplimiento de la ley de precio único, es objeto de estudio para la Economía de la Búsqueda (search Theory). Para Stigler (1961) la dispersión de los precios es la medida de la cantidad de información que tiene el consumidor. Sin embargo, Tommasi (1995) pone en duda dicha relación y afirma que la falta de información de los consumidores sobre la distribución de precios de un determinado bien genera poder de mercado para los vendedores, lo que aprovechan éstos para subir los precios, pero no encuentra justificación para que ello afecte a la dispersión de los precios.

Desde que Stigler (1961) afirmó que la dispersión de los precios era la medida de la ignorancia de los consumidores, muchos autores la han utilizado como una medida de la información del consumidor y la han explicado en función de la inflación, esperándose una relación positiva entre ambas variables. Esta conclusión tiene su fundamento teórico en la combinación de los modelos de fijación de precios de las empresas y la teoría según la cual los consumidores no están dispuestos a realizar una investigación exhaustiva del mercado. En este

marco, como referencia Van Hoomissen (1988), diferentes autores han demostrado que las empresas pueden maximizar beneficios generando una distribución de precios con más de un único precio para cada producto. En este caso, si la dispersión de precios está relacionada con la ignorancia en el mercado, la inflación debiera ir acompañada de incremento en la dispersión de precios. Nuevamente Van Hoomissen (1988) cita varios trabajos en los cuales se encuentra evidencia a favor de una relación positiva entre la inflación y la dispersión relativa de los precios.

Sin embargo, Reinsdorf (1994), utilizando datos de precios para nueve ciudades de Estados Unidos, obtiene una relación negativa entre ambas variables, apoyando indirectamente el planteamiento teórico de Tommasi (1995).

Rodríguez, González y Rodríguez (2009) analizan los cambios de precios diferenciando crecimientos y decrecimientos por quintiles y encuentran que las reducciones de precios son más fuertes al incrementar el quintil del precio, mientras que los crecimientos de precios son muy similares entre quintiles, excepto cuando nos encontramos en una situación de picos de inflación, en cuyo caso, son los precios más bajos los que mayores crecimientos medios presentan. Estos resultados apoyarían las conclusiones de Reinsdorf (1994) en el sentido de que existe relación entre dispersión e inflación y que ésta puede ser de signo negativo, puesto que, en situaciones de aceleración de la inflación, si los precios más bajos son los que crecen relativamente más, la dispersión se reduciría a la vez que la inflación se dispara.

En lo que sigue el trabajo se estructura en cuatro partes. En el epígrafe segundo planteamos las hipótesis que se van a contrastar. En el punto tercero describimos los datos con los cuales se contrastarán las hipótesis enunciadas anteriormente. En el punto cuarto contrastaremos dichas hipótesis y en el punto quinto se enuncian las principales conclusiones del estudio.

2. PLANTEAMIENTO DE LAS HIPÓTESIS A CONTRASTAR

En la mayoría de los estudios empíricos revisados se dispone de información individual de precios, entre la que se encuentra el establecimiento, pero normalmente se renuncia a esta información y lo máximo con lo que se trabaja es con el tipo de establecimiento.

En este trabajo nosotros tomamos como elemento fundamental el establecimiento. Nuestra hipótesis es que los establecimientos son heterogéneos como lo son los consumidores. En este sentido, aunque se analicen precios de un bien único, por tanto homogéneo, los servicios que lleva aparejados y que son propios del establecimiento pueden ser distintos y ello tiene como consecuencia que el consumidor se sienta predispuesto a pagar un precio superior si considera que el establecimiento le añade un valor diferencial1. En este caso, el establecimiento tiene mayor poder de mercado y su consecuencia será que sistemáticamente nos encontraremos con un grupo reducido de establecimientos cuyos precios son superiores al resto.

Por otra parte, los consumidores son reacios a cambiar sus pautas de consumo, pero lo harían siempre y cuando el nivel de ahorro que obtuvieran compensase el coste en el que se ve obligado a incurrir para realizar la selección de un establecimiento más barato. De hecho lo hacen, tal y como señalan Berné et al.(2008) al encontrar que el 81% de los hogares con los que trabajan utilizan tres o más establecimientos para realizar sus compras familiares. Sin embargo, cuando abordan los factores que pueden estar detrás de estos cambios de establecimiento únicamente la frecuencia de la compra, el volumen total de la cesta de la compra, el nivel de ingreso del hogar, su tamaño y el área geográfica se muestran estadísticamente significativas a la hora de explicar el cambio de establecimiento. Son todas variables estructurales, que cambian lentamente y no explican las elevadas cifras de la frecuencia del cambio de los precios y de su reducida duración [Álvarez y Hernando (2006), Rodríguez, González y Rodríguez (2009)]. Otras variables, tales como la cantidad de tiempo disponible por el consumidor, su edad, su situación laboral, la etapa del ciclo de vida de la familia o el tamaño de la población en la que reside el consumidor son estadísticamente significativas en el trabajo de Berné et al. (2008)

[1] Este valor diferencial puede tener múltiples orígenes: diferentes niveles de confianza del consumidor en los establecimientos, distintos niveles de atención al cliente, facilidad para realizar la compra de una cesta de productos, localización, etc.

En este escenario es de esperar, por un lado, que los establecimientos pueden incrementar su poder de mercado encareciendo la obtención de información para el consumidor y, por otro lado, los consumidores podrían quedar más satisfechos con su consumo incrementando el esfuerzo para disponer de más información. La actuación de uno y otro en el mercado nos lleva a concluir que la dispersión es innata al mercado, pero si el mercado es competitivo debiera poder diferenciarse claramente al grupo de establecimientos con mayor poder de mercado de aquellos que compiten vía precios. De la misma manera, la existencia de establecimientos con un menú de costes más bajos implica que sus precios tendrían que corresponder mayoritariamente con los precios mínimos ofertados por el mercado.

En esta línea, nuestra primera hipótesis es:

H_1: Hay unos pocos establecimientos que aportan un plus adicional al producto, lo cual les da mayor poder de mercado.

La evidencia a favor de esta hipótesis la encontraremos si observamos que para todos los instantes de tiempo es el mismo conjunto de establecimientos el que ofrece el precio más alto. Por el contrario, si el precio más alto lo presentasen distintos establecimientos a lo largo del tiempo, ello apoyaría la idea de que no existe una aportación diferencial por parte del establecimiento y lo que pretenden éstos es ganar poder del mercado aprovechándose de la desinformación que generan en el consumidor y los costes en los que le obligaría a incurrir para poder mantenerse informados. Esto último también vendría apoyado si se observa que para unos productos un mismo establecimiento es el más caro, mientras que para otros no lo es.

La segunda hipótesis que pretendemos contrastar se refiere a los establecimientos con precios más bajos.

H_2: Hay unos pocos establecimientos que tienen un menú de costes significativamente más bajo que el resto.

La evidencia a favor de esta hipótesis la obtendríamos si comprobásemos que un reducido conjunto de establecimientos siempre tienen el precio más bajo. En caso contrario, de lo que estaríamos hablando es de competencia y/o de que las empresas juegan con los costes de obtener información por parte de los consumidores para crear confusión y poder fijar en diferentes periodos un precio superior al mínimo. Si se observase que para un mismo establecimiento, para unos productos presenta el precio mínimo mientras que para otros no, se presentaría menos evidencia estadística a favor de la hipótesis planteada.

El análisis de los establecimientos con precios mínimos nos permite obtener una medición del coste en el cual tiene que incurrir el consumidor para obtener información, así como calcular el IPC mínimo que debe asumir el consumidor para mantener su nivel de vida. Este IPC sería el IPC que se obtiene por parte de un consumidor que siempre comprase en el establecimiento más barato.

La tercera y última hipótesis que pretendemos contrastar se refiere a la influencia del tamaño del mercado en la distribución de precios. Ello es posible dado que disponemos de información para un territorio geográficamente fragmentado compuesto por siete islas, formando cada una de ellas un mercado que imposibilita la existencia de una competencia real entre establecimientos de distintas islas. Por tanto, la tercera hipótesis que planteamos es:

H_3: El tamaño de los mercados tiene un claro efecto sobre la distribución de precios que se traduce en menor competencia cuanto menor tamaño tenga el mercado, básicamente por el mayor poder de mercado de las empresas.

3. LOS DATOS

Los datos que se usan son observaciones mensuales de precios que recoge el Instituto Canario de Estadística (ISTAC) con el fin de elaborar la operación estadística Cesta de la Compra. Encuesta de Precios-Canarias. La metodología puede consultarse en ISTAC(2008).

La muestra está formada por registros de precios de dos productos para un período máximo comprendido entre septiembre del año 2003 y febrero de 2008. Para cada precio individual se dispone de los siguientes datos: la identificación del producto, etiquetados como producto 1 y 2, el año y el mes en que fue medido el precio, el código identificativo del establecimiento en el que se observó el precio (un total de 99 establecimientos) y la isla. El producto 1 pertenece a la rúbrica "Productos lácteos", el dos a la rúbrica "Leche" y ambos se integran en el grupo COICOP[2] 01.1. Los productos están perfectamente definidos incluyendo la marca y el empaquetado, siendo ambos productos de consumo habitual y repetitivo, pero el segundo de naturaleza más básica que el primero. En total se dispone de 198 series temporales con una duración máxima que abarca desde septiembre de 2003 hasta febrero de 2008. Cada una de estas series mide, a lo largo del tiempo, el precio de uno de los dos productos en un establecimiento concreto, que a su vez pertenece a una de las siete islas que forman la Comunidad Autónoma de Canarias. El tamaño máximo de estas series es de 54 datos, pero existe un importante número de datos faltantes. De hecho, para el producto 1 únicamente se dispone de 35 series completas, lo cual supone un 35,71% del total de rachas completas. Para el producto 2 no hay ninguna serie completa debido a que no se dispone del dato del mes de mayo del año 2005, pero se dispone de 55 series (56,12%) a las cuales únicamente les falta dicho dato.

Con esta estructura, finalmente se dispone de un total de 8.860 precios, de los cuales 4.384 pertenecen al producto 1 y 4.476 al producto 2. Además, los datos se reparten entre las siete islas que componen la Comunidad Autónoma de Canarias, numeradas de la 1 a la 7, de las cuales, en términos de población y tamaño de mercado, la 3 y 4 son las de mayor tamaño, las 1, 2 y 6 de tamaño intermedio y la 5 y la 7 las más pequeñas.

El primer análisis que se realiza con estos datos se refiere a su nivel de representatividad. Para ello contamos como punto de referencia con el Índice de Precios de Consumo que elabora el Instituto Nacional de Estadística (INE) para las dos rúbricas en estudio para todo el período muestral, junto con sus ponderaciones en el conjunto del IPC para los años 2007 y 2008. Centrándonos en estas ponderaciones (ver tabla 1), se observa que son dos rúbricas que tienen mayor peso en Canarias que en el conjunto de España y para Canarias la rúbrica Productos Lácteos es la de mayor peso dentro del grupo de alimentación, mientras que la rúbrica Leche es la séptima o sexta más importante de dicho grupo dependiendo del año, aunque ya a mucha distancia de la primera.

Tabla 1. Ponderaciones IPC (Tantos por mil)

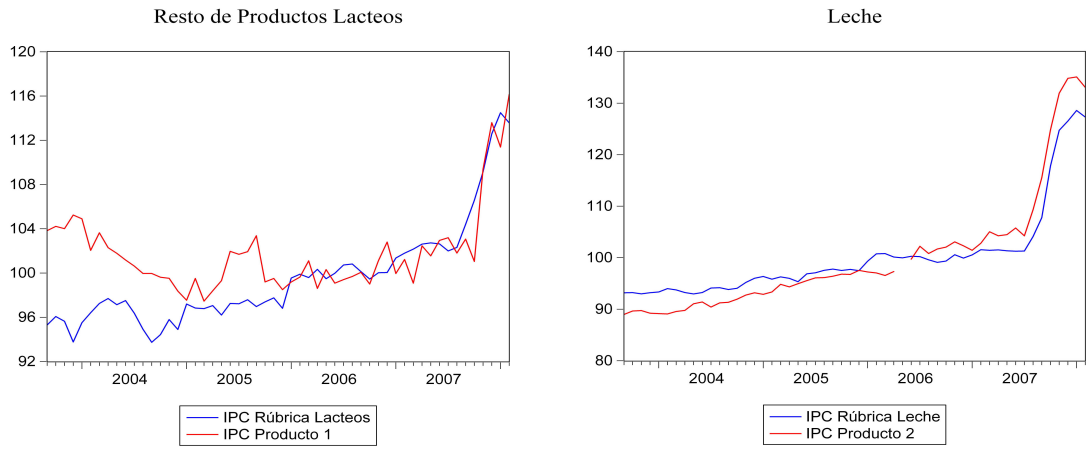
Rúbrica/Grupo	2007		2008	
	España	Canarias	España	Canarias
Productos Lácteos	16,624	22,897	15,560	21,430
Leche	11,475	13,504	11,748	13,805
Grupo Alimentación y Bebidas no Alcohólicas	220,555	228,192	202,793	209,781

Fuente: Instituto Nacional de Estadística
Elaboración Propia

Con el objeto de analizar el nivel de representación de los datos, en la ilustración 1 se muestran los índices de precios elaborados por el INE para Canarias y para las rúbricas de Productos Lácteos y Leche y los que se obtienen con los precios de los que disponemos de los productos 1 y 2 respectivamente[3].

[2] Siglas de la Classification Of Individual COnsumption by Purpose, United Nations Statistics Division
[3] Para la elaboración de los índices de precios con los datos disponibles se han utilizado las fórmulas no ponderadas de Dutot y Jevons, no observándose diferencias estadísticamente significativas entre ambos índices.

Ilustración 1. Relación entre IPC rúbrica INE e IPC elaboración propia



Para la rúbrica Productos Lácteos el índice de precios elaborado con los datos disponibles del producto 1 presenta una correlación de 0,68 observándose velocidades de precios claramente distintas en la primera mitad de la serie. En consecuencia, los resultados que obtengamos con el producto tendrán una validez limitada para el conjunto de la rúbrica. Este correspondencia limitada era de esperar, puesto que esta rúbrica representa en torno al 10% del gasto del grupo de alimentación, por lo cual sintetiza la evolución de precios de múltiples productos con distintos niveles de complementariedad entre ellos. Sin embargo, la relación que mantiene el IPC de la rúbrica Leche con el Índice de Precios elaborado con los datos del producto 2 es superior a 0,98, indicando el alto nivel de representatividad de los datos muestrales.

El análisis por islas muestra que en ambos productos los precios son más grandes en las islas más pequeñas, lo cual, siguiendo a Tommasi (1995), implicaría que los establecimientos tienen mayor poder de mercado en las islas más pequeñas con respecto a las islas de mayor tamaño. Sin embargo, las inflaciones medias son estadísticamente iguales en todas las islas, no así su varianza. Se diferencian ambos productos en que la correlación entre la inflación para el conjunto de las siete islas y la inflación de cada una de ellas es estadísticamente más baja para el producto 1 que para el producto 2. Estas correlaciones se muestran en la tabla 2. El triángulo superior derecho contiene las correlaciones entre las inflaciones de las islas y del conjunto correspondientes al producto 1, mientras que el triángulo inferior izquierdo se refiere a las correlaciones de las inflaciones para el producto 2.

Tabla 2. Correlaciones entre inflaciones por islas y productos

Isla		Producto 1							
		Isla 1	Isla 2	Isla 3	Isla 4	Isla 5	Isla 6	Isla 7	Canarias
Producto 2	Isla 1		0,039	0,170	0,238	-0,191	0,414	0,197	0,496
	Isla 2	0,412		0,174	0,157	0,065	0,305	-0,115	0,479
	Isla 3	0,089	0,370		0,357	0,073	0,216	-0,076	0,620
	Isla 4	0,195	0,422	0,611		-0,025	0,333	0,083	0,715
	Isla 5	0,311	0,323	0,211	0,416		-0,003	-0,058	0,138
	Isla 6	0,156	0,346	0,466	0,601	0,446		0,060	0,656
	Isla 7	0,427	0,031	0,270	0,050	0,216	0,242		0,157
	Canarias	0,547	0,674	0,708	0,786	0,619	0,721	0,407	

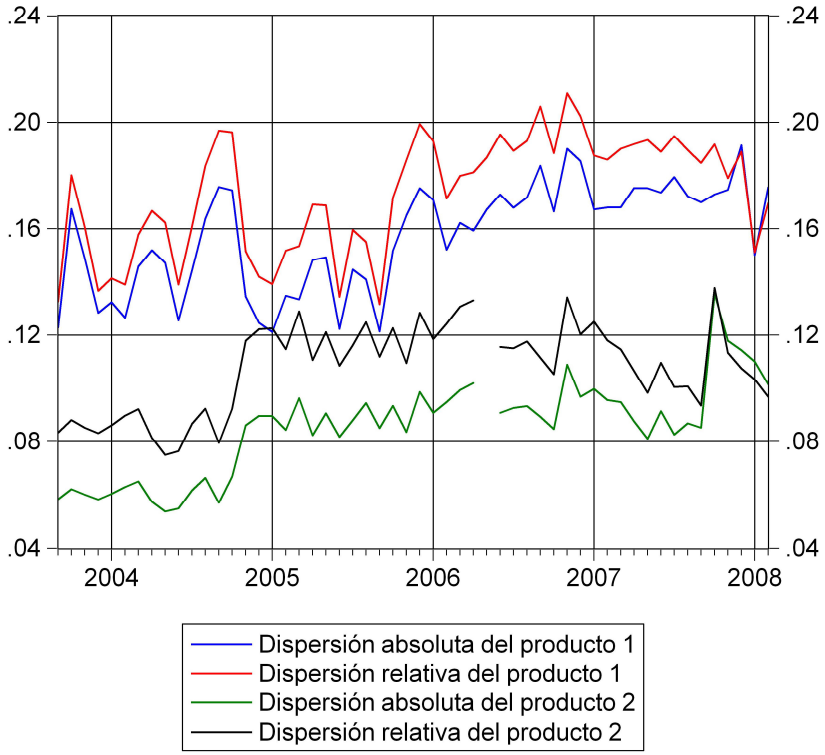
Elaboración propia

En un segundo análisis de los datos nos centramos en la dispersión, objeto central del trabajo. La ilustración 2 muestra la evolución temporal de las dispersiones para el conjunto total de precios, tanto en términos absolutos, medida como desviación típica de los precios, como relativa, dividiendo dicho valor por el precio medio para cada instante de tiempo. Ambas dispersiones se han calculado para cada uno de los productos.

Como se puede observar, las dispersiones, tanto las absolutas como las relativas, son distintas para ambos productos y presentan unas evoluciones también distintas entre productos, pero no así entre absolutas y relativas. En los cuatro casos coinciden en que presentan un valor mayor al final del período analizado que al principio.

También se observa una considerable caída de la dispersión en los últimos meses que se justifica por unos meses de fuerte inflación (ver ilustración 1) con mayor crecimiento de los precios más bajos. Siguiendo a Stigler (1961) tendríamos que decir que los consumidores son en el año 2008 más ignorantes del mercado que en el año 2003. El análisis por islas muestra un comportamiento muy disperso y con fuertes oscilaciones, lo que no permite identificar patrones generales de comportamiento.

Ilustración 2. Dispersiones absolutas y relativas



En los anexos 1 y 2 se muestran los resultados de las estimaciones del modelo genérico (1), en donde y es la dispersión total y por islas, en unos casos absoluta y en otros relativa[4], x es el IPC o la inflación, ε es la perturbación aleatoria y α, β son los parámetros del modelo. El anexo 1 se corresponde con las regresiones para el producto 1 y el anexo 2 para el producto 2.

$$y_i = \alpha + \beta x_i + \varepsilon_i \qquad (1)$$

[4] La dispersión absoluta se calcula como la desviación típica de los precios y la dispersión relativa como el cociente de la dispersión absoluta y la media aritmética de los precios.

En cada uno de estos anexos, la columna 1 contiene la variable y, en la columna 2 se indica la variable x, la tercera columna el valor estimado de β, en la cuarta su nivel de significación estadística y el la última columna el valor del coeficiente de determinación. Los resultados confirman la relación positiva, estadísticamente significativa, entre el IPC (o lo que es lo mismo, con los precios) y la dispersión absoluta y en menor medida la relativa y no tanto entre la inflación y la dispersión.

4. EXPLICACIÓN DE LA DISPERSIÓN DE LOS PRECIOS SOBRE LA BASE DE LOS ESTABLECIMIENTOS

El objetivo de este apartado es buscar evidencia a favor o en contra de las tres hipótesis planteadas. Los resultados a lo que se llega en los apartados anteriores son a nivel agregado, bien por isla o bien para el conjunto de la comunidad autónoma, y lo que ponen de manifiesto es la gran heterogeneidad que presenta la dispersión a medida que reducimos los tamaños de los mercados y es, por tanto, utilizando la información de los mercados más pequeños en dónde podemos encontrar la información necesaria para abordar las hipótesis planteadas.

El instrumento de análisis que vamos a utilizar es, para cada isla, la matriz de establecimientos por tiempo, identificando los establecimientos más caros y los más baratos (anexos 3 al 16). En estos gráficos se representa en el eje de abscisas el tiempo y en el de ordenadas los establecimientos, indicando mediante un círculo de color azul un precio máximo y un cuadrado de color rojo un precio mínimo. Cada anexo se corresponde con una isla y un producto. El análisis visual de estos anexos revela los siguientes patrones de comportamiento: i) hay más establecimientos que presentan precios mínimos que precios máximos, ii) el resultado anterior es mas evidente para el producto 2 (el de consumo más básico y repetitivo) que para el producto 1, iii) en las islas menores los precios extremos se concentran en un menor número de establecimientos y iv) si el consumidor quisiera consumir siempre en un establecimiento de precio mínimo se vería obligado a cambiar de establecimiento con una alta frecuencia.

Con el objeto de cuantificar estos resultados, a partir de los datos que sustentan este conjunto de ilustraciones, calcularemos distintos estadísticos. Para su definición, en primer lugar, denotamos por $X\max_{j,t}^{i,a}$ y $X\min_{j,t}^{i,a}$ a las variables definidas como:

$$X\max_{j,t}^{i,a} = \begin{cases} 1 & \text{si en la isla } i \text{ el producto } a \text{ presenta un precio máximo en el establecimiento } j \text{ en el instante } t \\ 0 & \text{En caso contrario} \end{cases}$$

$$X\min_{j,t}^{i,a} = \begin{cases} 1 & \text{si en la isla } i \text{ el producto } a \text{ presenta un precio mínimo en el establecimiento } j \text{ en el instante } t \\ 0 & \text{En caso contrario} \end{cases}$$

En donde[5]:

$$i = \{1,2,3,4,5,6,7\}; a = \{1,2\}; j = \{1,2,...,j_i\}; t = \{sep.2003, oct.2003, ..., feb.2008\}$$

$$j_1 = 12; j_2 = 12; j_3 = 22; j_4 = 22; j_5 = 10; j_6 = 16; j_7 = 5; C_t = \text{Cardinal de } t$$

Además definimos como rutas de máximos al conjunto de rutas formadas por el conjunto de establecimientos que permiten consumir para todos y cada uno de los instantes de tiempo en establecimientos de precios máximos ($R\max_h^{i,a}$). De igual manera definimos las rutas de mínimos para el caso de precios mínimos y la denotaremos por $R\min_h^{i,a}$. Llamaremos ruta óptima de máximos a aquella ruta de máximos que obligue al consumidor a realizar el menor

[5] Para el producto 2 el índice *t* excluye el mes de mayo del año 2005 al no disponer de datos para ese instante de tiempo.

número de cambios de establecimientos. La denotaremos por $R\max_o^{i,a}$. De forma similar definimos la ruta óptima de mínimos, $R\min_o^{i,a}$.

A partir de estas variables se pueden definir las siguientes medidas:

$M_1^{i,a}$: Probabilidad de que un consumidor del producto *a* al seleccionar un establecimiento al azar en la isla *i* elija uno de precio máximo.

$$M_1^{i,a} = \frac{\sum_{j=1}^{j_i}\sum_{t} X\max_{j,t}^{i,a}}{j_i * C_t} \tag{2}$$

$M_2^{i,a}$: Probabilidad de que un consumidor del producto *a* al seleccionar un establecimiento al azar en la isla *i* elija uno de precio mínimo.

$$M_2^{i,a} = \frac{\sum_{j=1}^{j_i}\sum_{t} X\min_{j,t}^{i,a}}{j_i * C_t} \tag{3}$$

$M_3^{i,a}$: Peso del número de cambios de establecimiento en $R\max_o^{i,a}$.

$$M_3^{i,a} = \frac{\text{Número de cambios de establecimientos presentes en } R\max_o^{1,a}}{C_t} \tag{4}$$

$M_4^{i,a}$: Peso del número de establecimientos en $R\max_o^{i,a}$.

$$M_4^{i,a} = \frac{\text{Número de establecimientos presentes en } R\max_o^{1,a}}{j_i} \tag{5}$$

$M_5^{i,a}$: Peso del número de cambios de establecimiento en $R\min_o^{i,a}$.

$$M_5^{i,a} = \frac{\text{Número de cambios de establecimientos presentes en } R\min_o^{1,a}}{C_t} \tag{6}$$

$M_6^{i,a}$: Peso del número de establecimientos en $R\min_o^{i,a}$.

$$M_6^{i,a} = \frac{\text{Número de establecimientos presentes en } R\min_o^{1,a}}{j_i} \tag{7}$$

También definimos un conjunto de medidas para cada isla, artículo y establecimiento, que denotaremos por $M_{7,j}^{i,a}$ y $M_{8,j}^{i,a}$

$M_{7,j}^{i,a}$: Peso del establecimiento j como establecimiento del precio máximo dentro del conjunto de establecimientos de precios máximos.

$$M_{7,j}^{i,a} = \frac{\sum_{t} X \max_{j,t}^{i,a}}{\sum_{j=1}^{j_i} \sum_{t} X \max_{j,t}^{i,a}} \tag{8}$$

$M_{8,j}^{i,a}$: Peso del establecimiento j como establecimiento del precio mínimo dentro del conjunto de establecimientos de precios mínimos.

$$M_{8,j}^{i,a} = \frac{\sum_{t} X \min_{j,t}^{i,a}}{\sum_{j=1}^{j_i} \sum_{t} X \min_{j,t}^{i,a}} \tag{9}$$

Por último se definen seis nuevas medidas, cuyo objetivo es cuantificar los niveles de concordancia entre los precios de ambos productos en los establecimientos con precios máximos y mínimos. Estas medidas se denotan por M_9^i a M_{14}^i y se definen en las ecuaciones (10) a (15).

$$M_9^i = \frac{\text{Número de veces en las cuales un establecimientos presenta precios máximos simultáneamente para ambos productos}}{\text{Número de veces en las cuales un establecimientos presenta precio máximo para el producto 1}} \tag{10}$$

$$M_{10}^i = \frac{\text{Número de veces en las cuales un establecimientos presenta precios máximos simultáneamente para ambos productos}}{\text{Número de veces en las cuales un establecimientos presenta precio máximo para el producto 2}} \tag{11}$$

$$M_{11}^i = \frac{\text{Número de establecimientos con precios máximos simultáneamente para ambos productos}}{\text{Número total de establecimientos}} \tag{12}$$

$$M_{12}^i = \frac{\text{Número de veces en las cuales un establecimientos presenta precios mínimos simultáneamente para ambos productos}}{\text{Número de veces en las cuales un establecimientos presenta precio mínimo para el producto 1}} \tag{13}$$

$$M_{13}^i = \frac{\text{Número de veces en las cuales un establecimientos presenta precios mínimos simultáneamente para ambos productos}}{\text{Número de veces en las cuales un establecimientos presenta precio mínimo para el producto 2}} \tag{14}$$

$$M_{14}^i = \frac{\text{Número de establecimientos con precios mínimos simultáneamente para ambos productos}}{\text{Número total de establecimientos}} \tag{15}$$

En las tablas 3, 4 y 5 se muestran los valores que se obtienen para todas y cada una de las medias definidas.

El análisis de las medidas $M_1^{i,a}$ y $M_2^{i,a}$ nos indica que hay diferencias tanto por producto como por tamaño de la isla. En concreto, $M_1^{i,a}$ muestra que la probabilidad de consumir en establecimientos de precio máximo es la misma para ambos productos, pero a mayor tamaño del mercado menor es la probabilidad de seleccionar un establecimiento de precio máximo, pasando de una probabilidad del 5% en las islas más grandes a valores entorno al 10% en las medianas y

del 20% en las más pequeña. Sin embargo, el análisis de $M_2^{i,a}$ indica que la probabilidad de consumir en un establecimiento de precio mínimo es significativamente mayor para el producto 2 que para el producto 1 y, además, para el producto 1 cuanto mayor sea el tamaño del mercado menor es la probabilidad de elegir el establecimiento más barato, mientras que para el producto 2 es en las islas de mayor tamaño la que presentan una mayor probabilidad de consumir en un establecimiento de precio mínimo.

La comparación de ambas medidas entre islas muestra que en la islas medianas y grandes la probabilidad de comprar en un establecimiento de precio mínimo es mayor que la de comprar a uno de precio máximo, mientras que en las islas más pequeñas tienden a igualarse ambas probabilidades.

Las medidas $M_3^{i,a}$ a $M_6^{i,a}$ se calculan sobre las series óptimas de precios máximos, medidas $M_3^{i,a}$ y $M_4^{i,a}$, y sobre las series óptimas de precios mínimos, medias $M_5^{i,a}$ y $M_6^{i,a}$. Antes de analizar los resultados de estas medidas es de destacar que la inflación media calculada con estas series óptimas son estadísticamente iguales, lo cual en absoluto implica que el consumidor incurra en el mismo gasto. Simplemente significa que las velocidades medias relativas de precios son las mismas.

Tabla 3. Resultados de las medidas generales de análisis

Medida	Isla						
	1	2	3	4	5	6	7
$M_1^{i,1}$	0,12	0,09	0,05	0,05	0,10	0,09	0,20
$M_1^{i,2}$	0,14	0,11	0,05	0,05	0,10	0,10	0,21
$M_2^{i,1}$	0,15	0,12	0,11	0,08	0,10	0,10	0,21
$M_2^{i,2}$	0,20	0,20	0,25	0,35	0,12	0,22	0,22
$M_3^{i,1}$	0,54	0,26	0,30	0,50	0,37	0,50	0,26
$M_3^{i,2}$	0,21	0,25	0,21	0,23	0,23	0,55	0,17
$M_4^{i,1}$	0,67	0,58	0,38	0,36	0,80	0,56	0,80
$M_4^{i,2}$	0,58	0,36	0,32	0,32	0,30	0,67	0,80
$M_5^{i,1}$	0,57	0,54	0,56	0,50	0,11	0,48	0,44
$M_5^{i,2}$	0,38	0,43	0,28	0,28	0,15	0,32	0,19
$M_6^{i,1}$	0,50	0,75	0,52	0,45	0,40	0,56	0,80
$M_6^{i,2}$	0,58	0,64	0,32	0,45	0,30	0,53	0,60

Los resultados de las medidas $M_3^{i,a}$ y $M_6^{i,a}$ cuantifican el esfuerzo que tendría que realizar el consumidor para poder consumir siempre en una ruta óptima, tanto de precios máximos como de precios mínimos, medido este esfuerzo en términos de cambio de establecimiento. Por ejemplo, en la isla 1 y para el producto 1, un individuo que quisiera consumir siempre en un establecimiento de precio máximo, de cada 100 compras en 54 de ellas las realizaría en un establecimiento distinto al que consumió en el instante anterior y ello supondría tener que visitar el 67% de los establecimientos disponibles. De igual manera, un consumidor de la isla 1 del producto 1 para consumir siempre en un establecimiento de precio mínimo, al menos tendría que cambiar de establecimiento con respecto a la compra anterior en el 57% de las compras, lo cual le supondría tener que recorrer el 50% de los establecimientos.

El análisis de estas medidas nos muestra que estos esfuerzos son ligeramente menores para el producto 2 que para el producto 1, pero siempre por encima del 20% en términos de cambios de establecimientos con respecto a la compra anterior y del 30% en términos de

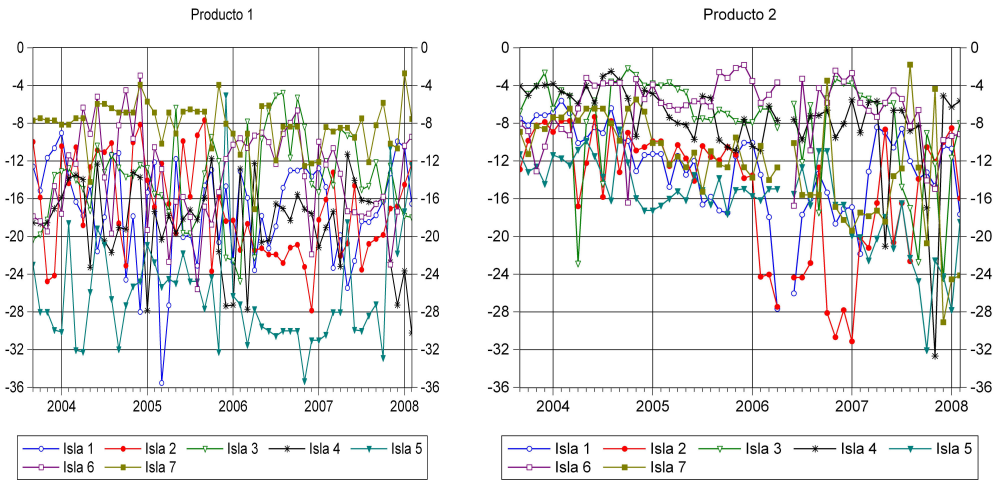
establecimientos utilizados. Es decir, para que un consumidor consuma siempre al precio máximo o mínimo, tendría que cambiar de establecimiento, como mínimo cada cuatro compras y ello le obliga a visitar, como mínimo, el 30% de los establecimientos disponibles. Para el producto 1 estas cifras se elevan hasta tener que cambiar de establecimiento cada 3 compras y visitar a más del 50% de los establecimientos. El efecto tamaño del mercado es menos evidente en este caso no pudiéndose identificar una pauta general.

Este nivel de esfuerzo que le permitiría al consumidor comprar siempre en un establecimiento de precio mínimo debe completarse con una medida del ahorro que ello le supondría. Para ello hemos calculado la diferencia relativa, tomando como referencia el precio medio, entre el precio medio y el precio mínimo por producto y por isla, indicando la tasa porcentual así obtenida la reducción de gasto que obtendría el consumidor al seleccionar siempre un establecimiento de precio mínimo frente a uno de precio medio. Los resultados de esas tasas de ahorro se muestran en la ilustración 3. Gráficamente se observan diferencias por producto y por isla. Para el producto 1 son las dos islas de menor población las que delimitan las posiciones extremas del ahorro. La isla 5 con un valor medio del 26,1% de ahorro y la isla 7 con el 8,3%, reflejando todas las islas un comportamiento del ahorro bastante estable, cada una de ellas a su nivel. El ahorro en el producto 2 es ligeramente más reducido, pero con una clara tendencia a incrementarse en el tiempo, volviendo a ser la isla 5 la que tiende a delimitar el ahorro máximo, con un 16.5% de ahorro, pero ahora es la isla 6 junto con la 3 y 4 las que menores niveles de ahorro presentan. La justificación a este resultado podría ser, para las islas 3 y 4 la mayor competencia al ser las de mayor población y en cuanto a la isla 6 la existencia de una producción propia para dicho producto.

Las medidas $M_{7,j}^{i,a}$ a $M_{8,j}^{i,a}$ miden el nivel de concentración de cada establecimiento como establecimiento que presenta precio máximo y mínimo respectivamente.

En términos generales, existe una mayor concentración en los establecimientos que presentan precios máximos con respecto a los que presentan precios mínimos, en el sentido de que entre dos establecimientos por cada isla concentran un porcentaje muy elevado de precios máximos. La concentración es menor para los establecimientos de precios mínimos y, además, a mayor tamaño de marcado la concentración es menor, hasta el punto de que en las islas 3 y 4, las de mayor tamaño, ningún establecimiento, ni para el producto 1 ni para el 2, llega a concentrar el 20% de los precios mínimos.

Ilustración 3. Niveles de ahorro por producto e isla



Valores medios							
	Isla 1	Isla 2	Isla 3	isla 4	Isla 5	Isla 6	Isla 7
Producto 1	-17,1%	-17,1%	-13,9%	-18,6%	-26,1%	-13,6%	-8,3%
Producto 2	-12,7%	-15,0%	-7,7%	-7.6%	-16,3	-6,5%	-12,1%

El conjunto de resultados anteriores apoyarían la hipótesis H_1, existencia establecimientos con mayor poder de mercado; rechazarían la hipótesis H_2, no hay un grupo reducido de establecimientos con costes menores que el resto y respaldarían la hipótesis H_3, el tamaño de los mercados influye en la distribución de los precios.

Tabla 4. Resultados de las medidas de análisis por establecimiento

Medida		Isla													
		1		2		3		4		5		6		7	
$M_{7,j}^{i,1} \mid M_{8,j}^{i,1}$	j=1	0,01	0,09	-	-	0,38	0,04	-	0,15	-	0,94	0,04	0,05	0,11	0,30
	j=2	0,29	-	0,07	-	-	-	0,02	0,02	0,54	-	0,03	0,01	0,05	0,11
	j=3	0,05	-	0,27	0,03	-	0,05	-	0,19	0,02	-	0,08	0,05	0,82	0,07
	j=4	0,07	-	0,10	0,12	-	0,01	0,12	-	0,16	0,02	0,36	-	-	0,02
	j=5	0,09	-	-	0,29	-	0,08	-	0,03	0,02	0,02	0,23	0,04	0,02	0,50
	j=6	-	0,23	0,03	0,01	0,02	0,03	-	0,09	-	-	0,09	0,01		
	j=7	0,03	0,18	0,44	-	0,09	-	-	0,15	0,04	-	0,03	0,05		
	j=8	0,36	-	0,03	0,03	-	0,02	0,09	-	0,02	0,02	-	0,02		
	j=9	0,11	-	0,02	0,08	0,05	0,02	-	0,01	0,07	-	0,01	-		
	j=10	-	0,11	0,02	0,07	0,16	0,01	-	-	0,14	-	0,04	0,01		
	j=11	-	0,27	0,02	0,28	-	0,11	0,05	-			0,03	0,02		
	j=12	-	0,11	-	0,11	-	0,06	0,34	-			-	0,28		
	j=13					-	0,19	-	-			-	0,01		
	j=14					0,02	0,08	0,02	0,10			0,04	0,16		
	j=15							0,17	-			-	0,28		
	j=16					0,25	-	0,17	0,01			0,03	-		
	j=17					-	0,07	-	0,19						
	j=18					0,02	0,07	-	-						
	j=19					-	0,06	-	0,01						
	j=20					-	-	0,02	0,03						
	j=21					-	0,04	-	0,01						
	j=22					-	0,05	-	0,01						
$M_{7,j}^{i,2} \mid M_{8,j}^{i,2}$	j=1	-	0,18	0,02	-	-	0,09	-	0,07	-	0,10	0,03	0,11	-	0,12
	j=2	0,40	0,05	0,30	0,03	0,56	-	-	0,08	-	0,21			0,52	0,04
	j=3	-	-	0,14	0,03	-	0,08	-	0,06	-	0,34	0,04	0,04	0,04	0,82
	j=4	0,01	0,04	0,00	0,03	-	-	0,03	-	0,54	-	0,13	0,02	0,29	-
	j=5	0,13	0,02	0,00	0,11	-	0,03	-	0,09	-	0,02	0,32	0,01	0,16	0,02
	j=6	0,01	0,13			-	0,07	-	0,08	-	-	0,03	0,03		
	j=7	0,02	0,15	0,49	0,00	0,02	0,03	-	0,09	-	-	0,01	0,11		
	j=8	0,40	0,05	0,03	0,17	-	0,01	0,21	-	-	0,34	0,08	0,11		
	j=9	-	0,01	0,02	0,22	0,02	0,02	-	0,07	0,31	-	0,03	0,15		
	j=10	0,01	0,10	0,00	0,19	-	0,06	0,02	0,03	0,15	-	0,17	0,02		
	j=11	-	0,21	0,00	0,08	-	0,03	0,02	0,01			0,07	0,08		
	j=12	0,01	0,06	0,00	0,14	0,07	0,05	0,48	0,02			0,05	0,12		
	j=13					-	0,09	-	0,01			0,01	-		
	j=14					-	0,08	-	0,08			-	0,08		
	j=15					-	-	0,06	-			0,03	0,12		
	j=16					0,28	-	0,08	0,05			0,01	-		
	j=17					0,02	0,06	-	0,06						
	j=18					-	0,05	0,05	0,03						
	j=19					-	0,07	0,03	0,04						
	j=20					-	0,04	0,02	0,06						
	j=21					0,02	0,05	-	-						
	j=22					0,02	0,08	-	0,06						

(-) Valor más pequeño que 0,01
Elaboración propia

Los datos que reflejan las medidas $M_{7,j}^{i,a}$ y $M_{8,j}^{i,a}$ también se pueden usar para analizar la concordancia de máximos y mínimos entre productos. En líneas generales, en los datos de la tabla 4 se observa poca concordancia en los establecimientos en términos de precios máximos y mínimos para los dos productos. La excepción es la isla 1, en la cual los establecimientos 2 y 8 presentan una alta frecuencia como establecimiento de precio máximo para ambos productos y,

además, el establecimiento 11 se destaca como un establecimiento que presenta precio mínimo para ambos productos con una frecuencia por encima del 20%. Aparte de la isla 1, los únicos establecimientos con una frecuencia elevada conjuntamente para ambos productos son el establecimiento 12 para la isla 4 y el establecimiento 4 para la isla 6, en ambos casos, como establecimientos de precios máximos.

Sin embargo, la medida de la concordancia entre productos que se observa con las medidas $M_{7,j}^{i,a}$ y $M_{8,j}^{i,a}$ es en frecuencia, lo que no garantiza que de forma simultánea el mismo establecimiento presente el precio máximo o mínimo en el mismo instante de tiempo, este último nivel de concordancia lo miden las medidas M_9^i a M_{14}^i, las tres primeras para los precios máximos y las tres últimas para los precios mínimos. . Por lo tanto, los datos que reflejan las medidas $M_{7,j}^{i,a}$ y $M_{8,j}^{i,a}$ se deben de complementar, para el análisis de la concordancia de máximos y mínimos, con las medidas M_9^i a M_{14}^i recogidos en la tabla 5.

Del análisis de estas medidas lo primero que destaca es que la probabilidad de que consumiendo en un establecimiento con precio máximo o mínimo en alguno de los dos producto también estemos consumiendo a precio máximo o mínimo el otro producto nunca supera el 53% en el caso de los máximos y 50 % en los mínimos. No se detectan patrones de comportamiento por el tamaño de las islas con la salvedad de la isla 7, cuyo nivel de concordancia es realmente mínimo, tanto para precios máximos como para precios mínimos, siendo en este caso menor la concordancia para los primeros que para los segundos. Para esta isla, la comparación del gráfico del anexo 15 con el del anexo 16 muestra un comportamiento por establecimientos totalmente contrario para ambos productos.

Tabla 5. Nivel de concordancia entre establecimientos de precios extremos

	Isla						
	1	2	3	4	5	6	7
M_9^i	$\frac{33}{76}=0,43$	$\frac{31}{59}=0,53$	$\frac{14}{55}=0,25$	$\frac{15}{58}=0,26$	$\frac{13}{56}=0,23$	$\frac{17}{77}=0,22$	$\frac{2}{55}=0,04$
M_{10}^i	$\frac{33}{90}=0,37$	$\frac{31}{63}=0,49$	$\frac{14}{54}=0,26$	$\frac{15}{62}=0,24$	$\frac{13}{54}=0,24$	$\frac{17}{76}=0,22$	$\frac{2}{56}=0,04$
M_{11}^i	$\frac{2}{12}=0,17$	$\frac{3}{12}=0,25$	$\frac{1}{22}=0,05$	$\frac{3}{22}=0,14$	$\frac{2}{10}=0,20$	$\frac{5}{16}=0,31$	$\frac{1}{5}=0,20$
M_{12}^i	$\frac{34}{98}=0,35$	$\frac{22}{76}=0,29$	$\frac{43}{125}=0,34$	$\frac{47}{94}=0,50$	$\frac{6}{54}=0,11$	$\frac{29}{85}=0,34$	$\frac{7}{56}=0,13$
M_{13}^i	$\frac{34}{130}=0,26$	$\frac{22}{119}=0,18$	$\frac{43}{286}=0,15$	$\frac{47}{412}=0,11$	$\frac{6}{62}=0,10$	$\frac{29}{178}=0,16$	$\frac{7}{57}=0,12$
M_{14}^i	$\frac{6}{12}=0,50$	$\frac{7}{12}=0,58$	$\frac{14}{22}=0,64$	$\frac{12}{22}=0,55$	$\frac{1}{10}=0,10$	$\frac{6}{16}=0,38$	$\frac{2}{5}=0,40$

Elaboración propia

Las medidas M_{11}^i y M_{14}^i muestran comportamientos distintos. De esta manera, M_{11}^i destaca por el reducido número de establecimientos que en algún momento del tiempo presentan precios máximos simultáneamente en ambos productos, mostrando indicios además de que a mayor tamaño del mercado menor número de establecimientos que simultanean en el tiempo precios máximos para los dos productos. Por el contrario, M_{14}^i obtiene valores genéricamente más altos que M_{11}^i, estando en las islas más pobladas por encima del 0,50. Es decir, en estas islas más de la mitad de los establecimientos, en algún momento presentaron precios mínimos simultáneamente para ambos productos

En consecuencia, el análisis realizado sobre la concordancia de los establecimientos en cuanto a precios máximos y mínimos rechaza las hipótesis H_1 y H_2 y aportarían evidencia a favor de la hipótesis H_3. Por tanto, la combinación de estos resultados con los obtenidos con las

anteriores medidas no es concluyente sobre la primera hipótesis, rechazaría la segunda y aportaría evidencia a favor de la tercera.

5. CONCLUSIONES

El análisis realizado permite concluir que en el mercado conviven:

a) Establecimientos con precios sistemáticamente máximos al concentrar entre pocos establecimientos un porcentaje elevado de precios máximos. Sin embargo, estos establecimientos son distintos para cada producto. Este comportamiento vendría explicado por un establecimiento que aporta un plus adicional al producto, pero cuando el bien es básico para el consumidor está dispuesto a bajar temporalmente el precio de dicho producto básico con la finalidad de atraer al consumidor al establecimiento. Este plus adicional en las islas menores puede estar justificado por la localización y la falta de competencia. Por tanto, los datos aportan evidencia a favor y en contra de la primera hipótesis planteada.

b) Todos los resultados analizados tienden a rechazar la hipótesis de que hay un conjunto reducido de empresas que tienen en menú de costes significativamente más bajo que el resto. En torno a un 15% de puntos de venta para el producto 1 y del 25% para el producto 2 son puntos de venta a precios mínimos, siendo el nivel de concentración por establecimiento menor cuanto mayor es el tamaño del mercado y no observándose diferencias por el hecho de que el producto sea de consumo más o menos básico. Forman un grupo voluminoso de establecimientos que compiten vía precios y que adquieren poder de mercado cambiando su posición relativa en la distribución de precios

c) Este amplio conjunto de empresas, que en algún momento ofertan un precio mínimo, incrementan su precio sabiendo que el consumidor no está dispuesto a adquirir la información que le permitiría comprar siempre a un precio mínimo, dado su elevado coste en términos de los cambios de establecimiento y de número de puntos de compra a los que se vería obligado el consumidor para mantenerse en una ruta de precio mínimo, a pesar de que ello le permitiría obtener un significativo ahorro. Esto queda de manifiesto por el hecho de que para que un consumidor se mantenga siempre en una ruta de precioÿÿmíniÿÿs se vería obligado a visitar ÿÿÿÿorno al 50% de los establecimientos disponibles y a cambiar de establecimiento, con respecto a la compra anterior, en una de cada dos compras.

d) El tamaño del mercado mantiene una relación inversa con el precio y no tanto con la inflación y con la dispersión. Sin embargo, las posiciones de los establecimientos con mayor poder de mercado y con menor menú de costes son muchas más claras que en las islas menores. A pesar de ello, el reducido número de establecimientos limita la generalización de los resultados.

Por tanto, la dispersión de los precios es el resultado de una combinación de factores relacionados con la coexistencia de un reducido número de empresas que tienen mayor poder de mercado, lo que le permite poner unos precios más altos, con otras, muchos más numerosas, que compiten para alcanzar mayor cuota de mercado vía precios. Las primeras utilizan reducciones puntuales de precios para atraer a consumidores y las segundas fijan precios temporalmente por encima del mínimo sabiendo que el consumidor, que está habituado a variar de establecimiento, lo tenga que hacer de forma aleatoria, ya que los cambios aleatorios de precios incrementan el coste de la información que necesita el consumidor para actuar de forma eficiente en el mercado.

BIBLIOGRAFÍA

ÁLVAREZ, L. y HERNANDO, I. (2006). "Price Settting Behavior in Spain: Evidencia from CONSUMER PRICE MICRODATA". *Economic Modelling*, vol. 23, pp. 699-716.

BERNÉ, C., MARTÍNEZ, N., GARGALLO, P. y SALVADOR, M. (2008). "Análisis del Comportamiento Variado de Compra de los Hogares Españoles: Una Aproximación Bayesiana", *Estadística española*, vol. 50, nº 169, pp. 479-521.

ISTAC (2008). Cesta de la Compra. Encuesta de Precios-Canarias. http:// www.gobiernodecanarias.org/istac.

LACH, S. y TSIDDON, D. (1992). "The Behavior of Prices and Inflation : An Empirical Analysis of Dissagregated Price Data", *Journal of Political Economy*, Vol. 100, pp. 349-389

REINSDORF, M. (1994). "New Evidence on the Relation Between Inflation and Price DISPERSIÓN", *American Economic Review*, Vol. 84, pp. 720-731

RODRÍGUEZ, S., GONZÁLEZ, C, y RODRÍGUEZ A. (2009): "Aproximación al Comportamiento de los Precios en una Economía Insular. El Caso de Canarias", 23 Congreso Internacional de Economía Aplicada, ASEPELT 2009.

STIGLER, GEORGE (1961). "The Economics of Information", *Journal of Political Economy*, Vol. 69, pp. 213-225

TOMMASI, M. (1995). "No sea Ignorante: La Dispersión de Precios no es una medida de la Ignorancia en el mercado", *Estudios de Economía*, Departamento de Economía de la Facultad de Ciencias Económicas y Administrativas de la Universidad de Chile, Vol. 22, nº 1, pp. 1-15

VAN HOOMISSEN, T. (1988). "Price Dispersion and Inflation: Evidence from Israel", *Journal of Political Economy*, Vol. 96, pp. 1303-1314

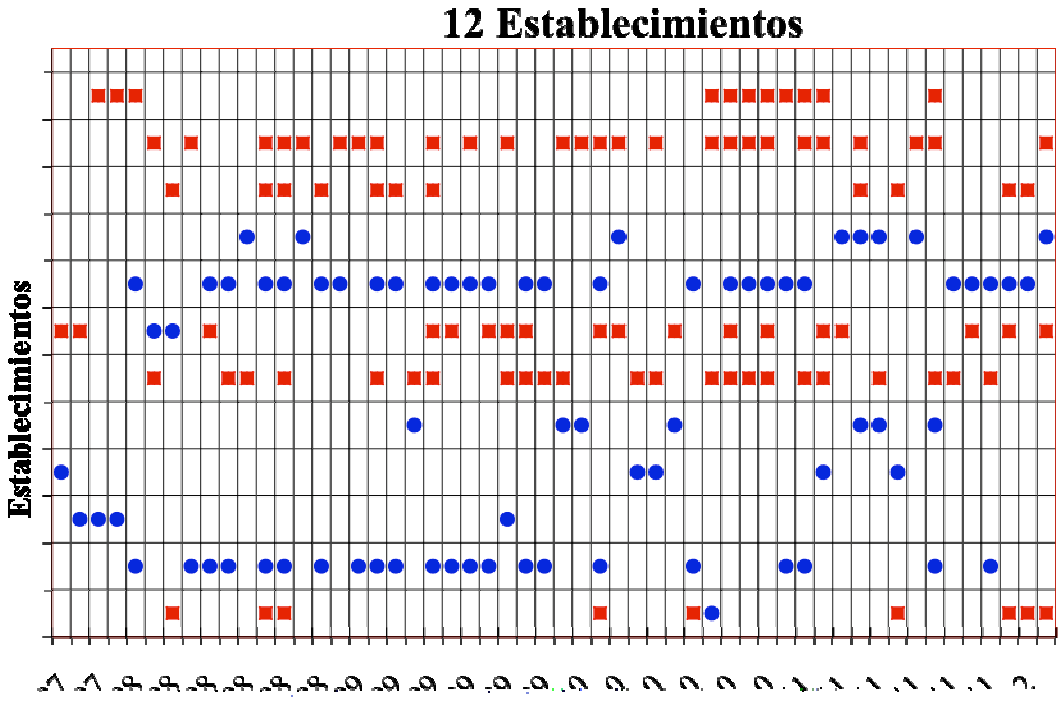
Anexo 1. Resultados de los modelos para explicar la dispersión del producto 1

Endógena	Exógena	Coeficiente estimado	Significación (prob.)	R^2
Dispersión total	IPC Grupo Lácteos	0,0022	0,0001	0,25
Dispersión isla 1	IPC isla 1	0,0029	0,0012	0,18
Dispersión isla 2	IPC isla 2	0,0015	0,0245	0,09
Dispersión isla 3	IPC isla 3	0,0033	0,0000	0,31
Dispersión isla 4	IPC isla 4	0,0037	0,0005	0,21
Dispersión isla 5	IPC isla 5	0,0035	0,0088	0,13
Dispersión isla 6	IPC isla 6	0,0021	0,0026	0,16
Dispersión isla 7	IPC isla 7	0,0025	0,017	0,11
Dispersión total	Inflación Grupo Lácteos	0,0013	0,5737	0,01
Dispersión isla 1	Inflación isla 1	0,0016	0,2683	0,02
Dispersión isla 2	Inflación isla 2	0,0008	0,7667	0,00
Dispersión isla 3	Inflación isla 3	0,003	0,0174	0,11
Dispersión isla 4	Inflación isla 4	0,003	0,0196	0,10
Dispersión isla 5	Inflación isla 5	-0,0001	0,9149	0,00
Dispersión isla 6	Inflación isla 6	0,003	0,0091	0,13
Dispersión isla 7	Inflación isla 7	0,0014	0,0921	0,05
Dispersión relativa total	IPC Grupo Lácteos	0,0015	0,02	0,10
Dispersión relativa isla 1	IPC isla 1	0,0017	0,06	0,69
Dispersión relativa isla 2	IPC isla 2	0,002	0,12	0,04
Dispersión relativa isla 3	IPC isla 3	0,003	0,001	0,19
Dispersión relativa isla 4	IPC isla 4	0,002	0,045	0,08
Dispersión relativa isla 5	IPC isla 5	0,0016	0,17	0,04
Dispersión relativa isla 6	IPC isla 6	0,0013	0,08	0,06
Dispersión relativa isla 7	IPC isla 7	0,0015	0,13	0,04
Dispersión relativa total	Inflación Grupo Lácteos	0,0001	0,96	0,00
Dispersión relativa isla 1	Inflación isla 1	0,0006	0,66	0,00
Dispersión relativa isla 2	Inflación isla 2	-0,0001	0,96	0,00
Dispersión relativa isla 3	Inflación isla 3	0,003	0,04	0,08
Dispersión relativa isla 4	Inflación isla 4	0,002	0,14	0,04
Dispersión relativa isla 5	Inflación isla 5	-0,0008	0,44	0,01
Dispersión relativa isla 6	Inflación isla 6	0,002	0,04	0,08
Dispersión relativa isla 7	Inflación isla 7	0,0008	0,27	0,02

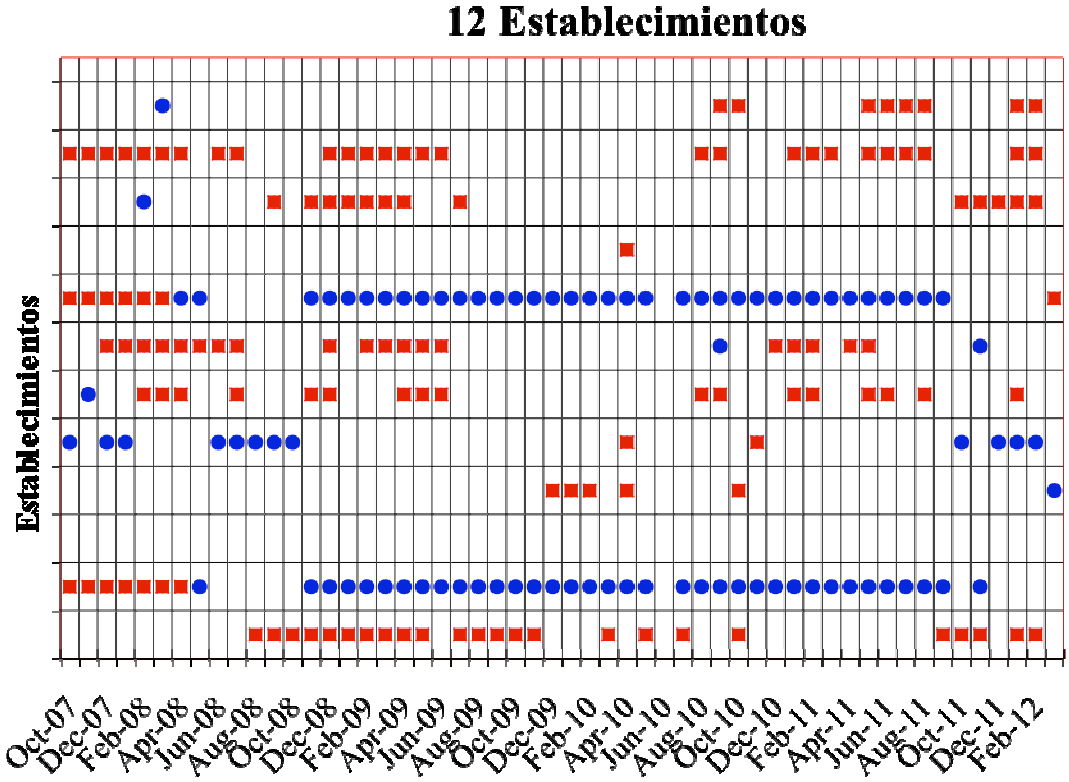
Anexo 2. Resultados de los modelos para explicar la dispersión del producto 2

Endógena	Exógena	Coeficiente estimado	Significación (prob.)	R^2
Dispersión total	IPC Grupo Leche	0,0014	0,000	0,50
Dispersión isla 1	IPC isla 1	0,00085	0,005	0,15
Dispersión isla 2	IPC isla 2	0,00024	0,48	0,01
Dispersión isla 3	IPC isla 3	0,001	0,000	0,65
Dispersión isla 4	IPC isla 4	0,00097	0,000	0,23
Dispersión isla 5	IPC isla 5	0,0028	0,000	0,65
Dispersión isla 6	IPC isla 6	0,0015	0,000	0,59
Dispersión isla 7	IPC isla 7	0,0015	0,000	0,41
Dispersión total	Inflación Grupo Leche	0,005	0,000	0,24
Dispersión isla 1	Inflación isla 1	0,0005	0,68	0,00
Dispersión isla 2	Inflación isla 2	-0,0007	0,50	0,01
Dispersión isla 3	Inflación isla 3	0,0043	0,000	0,36
Dispersión isla 4	Inflación isla 4	0,0025	0,11	0,05
Dispersión isla 5	Inflación isla 5	0,0056	0,000	0,27
Dispersión isla 6	Inflación isla 6	0,0057	0,000	0,29
Dispersión isla 7	Inflación isla 7	-0,0009	0,33	0,02
Dispersión relativa total	IPC Grupo Leche	0,0004	0,10	0,05
Dispersión relativa isla 1	IPC isla 1	0,00	0,83	0,00
Dispersión relativa isla 2	IPC isla 2	-0,0006	0,14	0,04
Dispersión relativa isla 3	IPC isla 3	0,0005	0,000	0,27
Dispersión relativa isla 4	IPC isla 4	0,0004	0,19	0,03
Dispersión relativa isla 5	IPC isla 5	0,0017	0,000	0,33
Dispersión relativa isla 6	IPC isla 6	0,0011	0,000	0,33
Dispersión relativa isla 7	IPC isla 7	0,0009	0,002	0,18
Dispersión relativa total	Inflación Grupo Leche	0,002	0,11	0,05
Dispersión relativa isla 1	Inflación isla 1	-0,0008	0,62	0,01
Dispersión relativa isla 2	Inflación isla 2	-0,0022	0,13	0,05
Dispersión relativa isla 3	Inflación isla 3	0,0033	0,000	0,29
Dispersión relativa isla 4	Inflación isla 4	0,0009	0,58	0,01
Dispersión relativa isla 5	Inflación isla 5	0,0051	0,000	0,32
Dispersión relativa isla 6	Inflación isla 6	0,0053	0,000	0,27
Dispersión relativa isla 7	Inflación isla 7	-0,0014	0,10	0,08

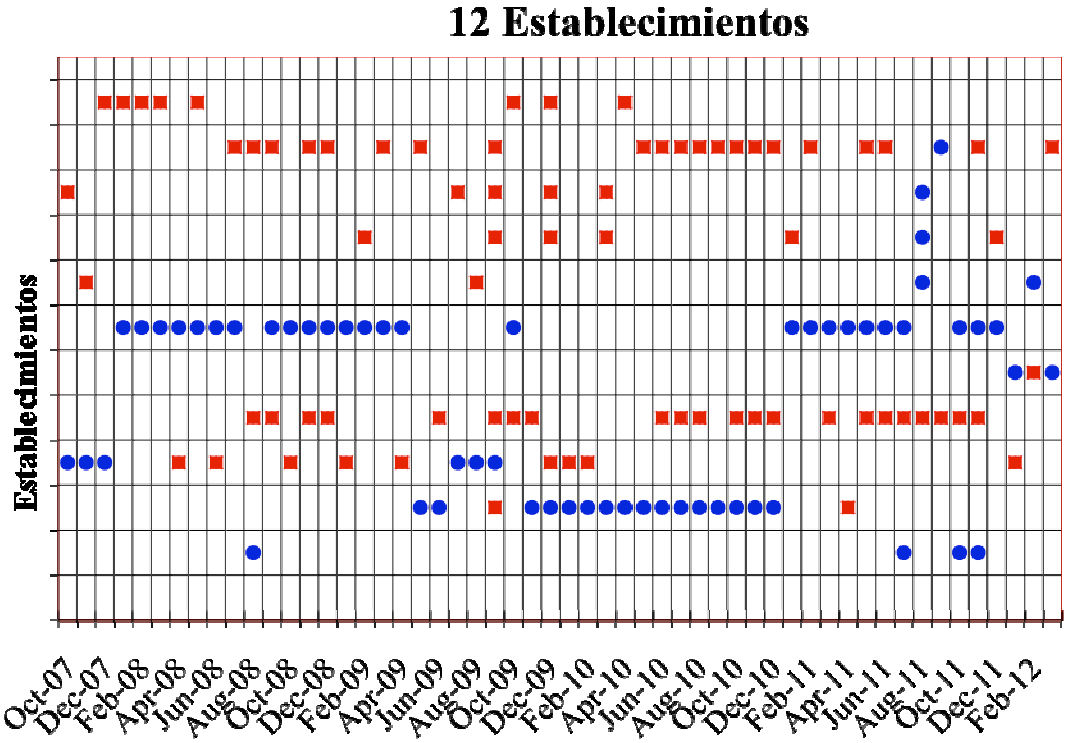
Anexo 3. Gráfico de precios máximos (●) y mínimos (■). Isla 1. Producto 1



Anexo 4. Gráfico de precios máximos (●) y mínimos (■). Isla 1. Producto 2



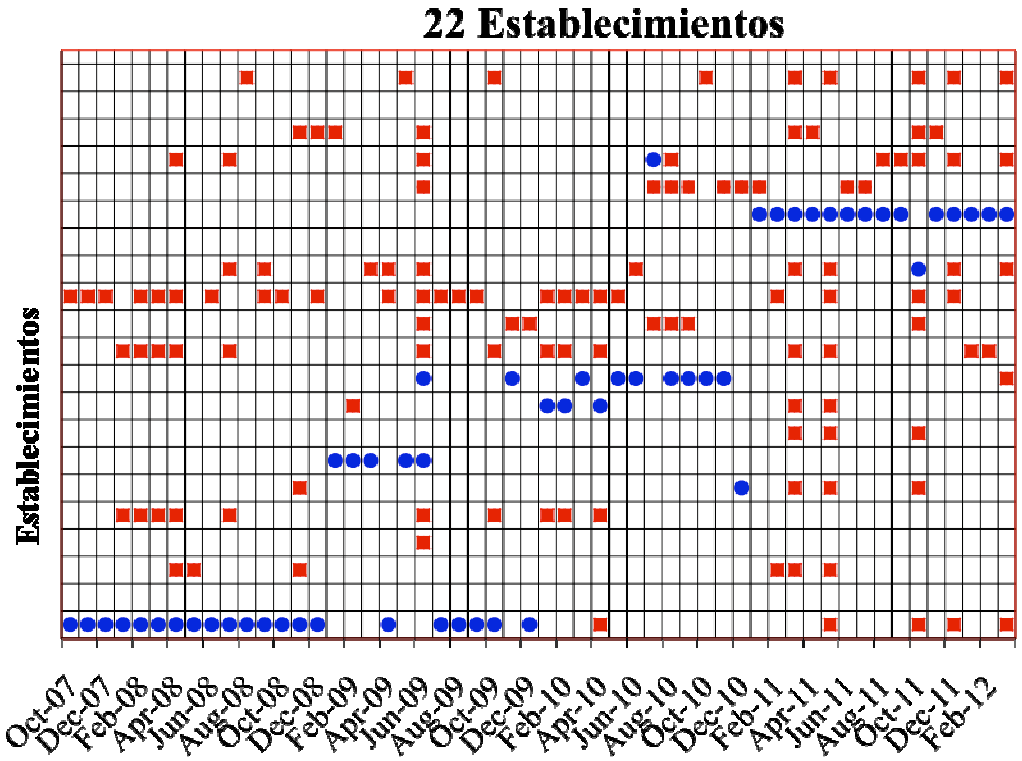
Anexo 5. Gráfico de precios máximos (●) y mínimos (■). Isla 2. Producto 1



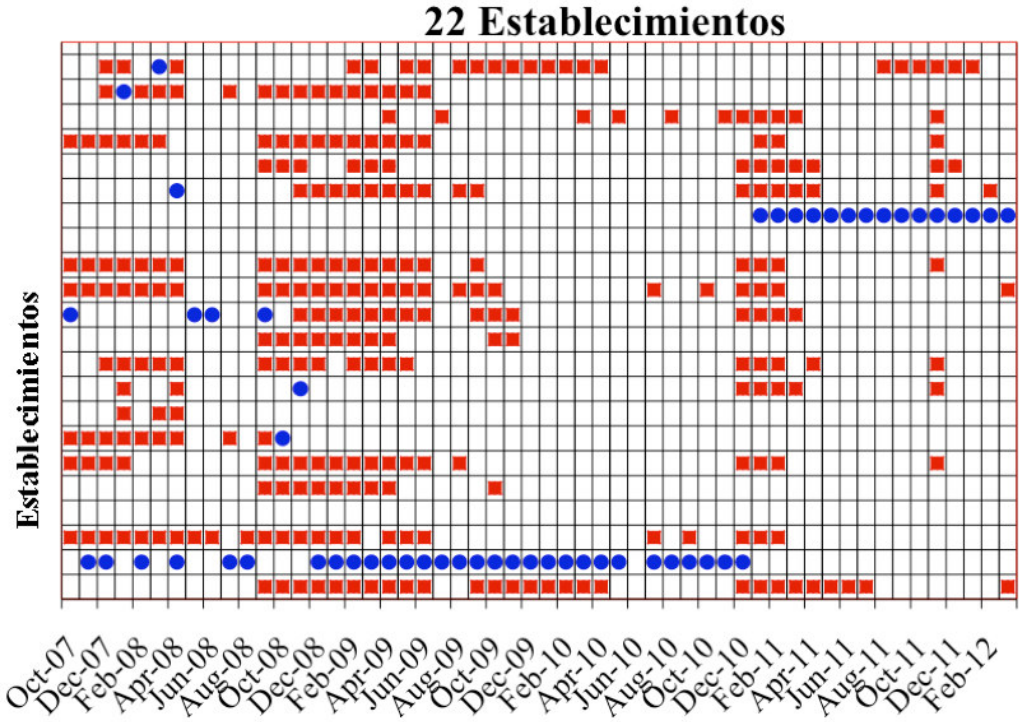
Anexo 6. Gráfico de precios máximos (●) y mínimos (■). Isla 2. Producto 2



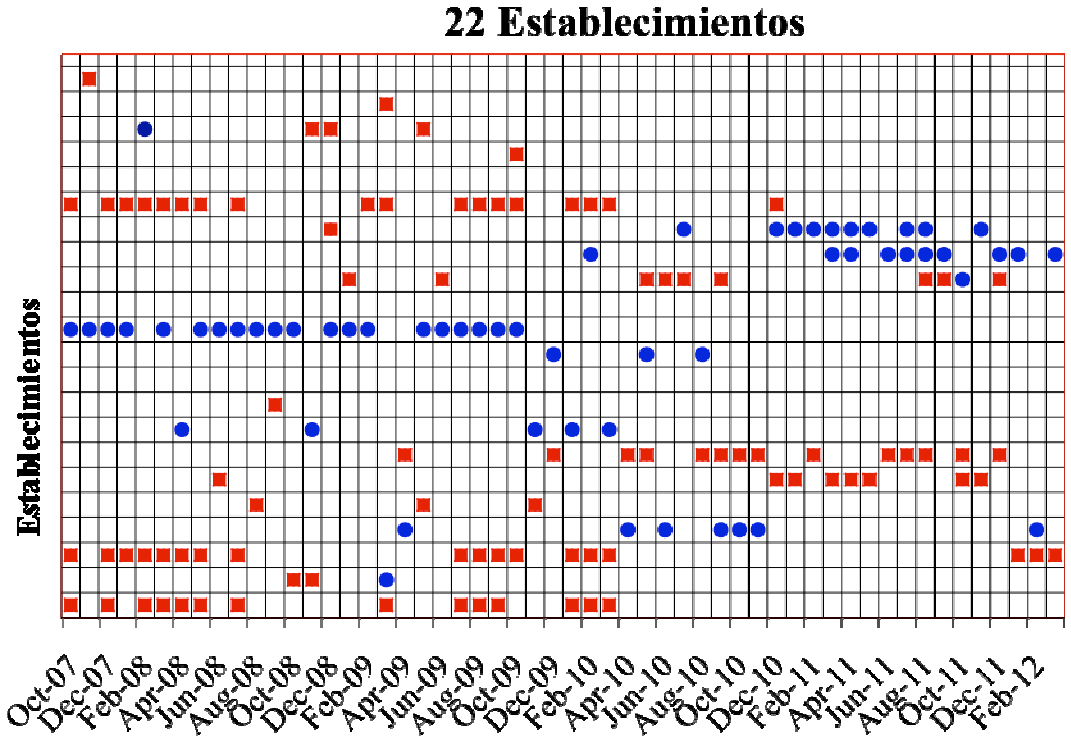
Anexo 7. Gráfico de precios máximos (●) y mínimos (■). Isla 3. Producto 1



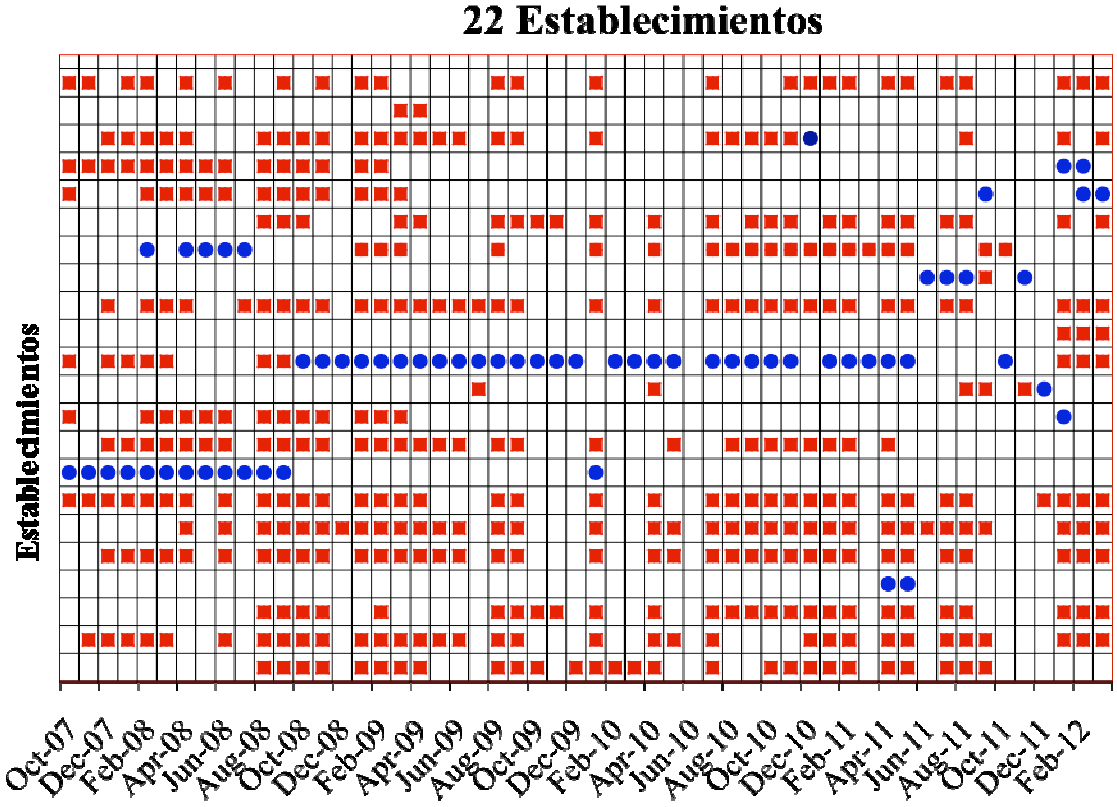
Anexo 8. Gráfico de precios máximos (●) y mínimos (■). Isla 3. Producto 2



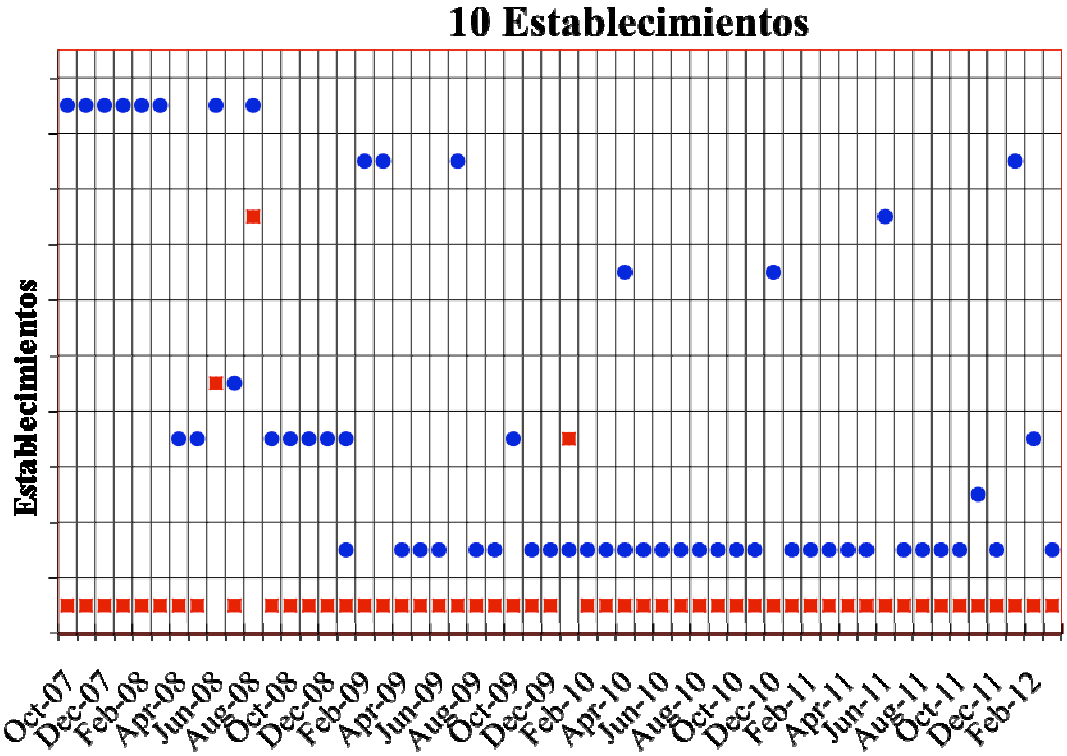
Anexo 9. Gráfico de precios máximos (●) y mínimos (■). Isla 4. Producto 1



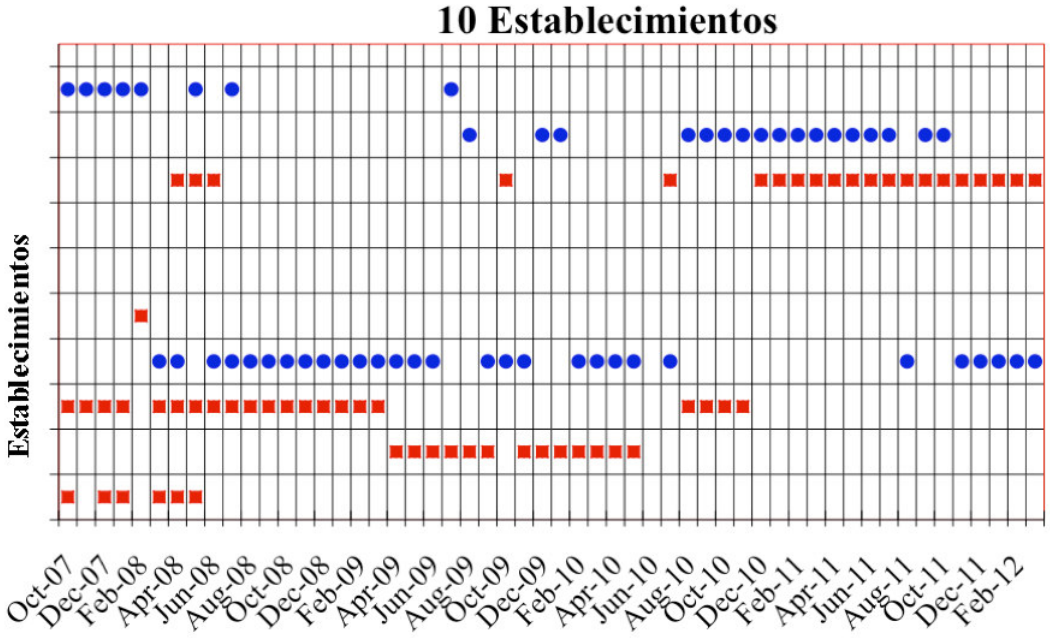
Anexo 10. Gráfico de precios máximos (●) y mínimos (■). Isla 4. Producto 2



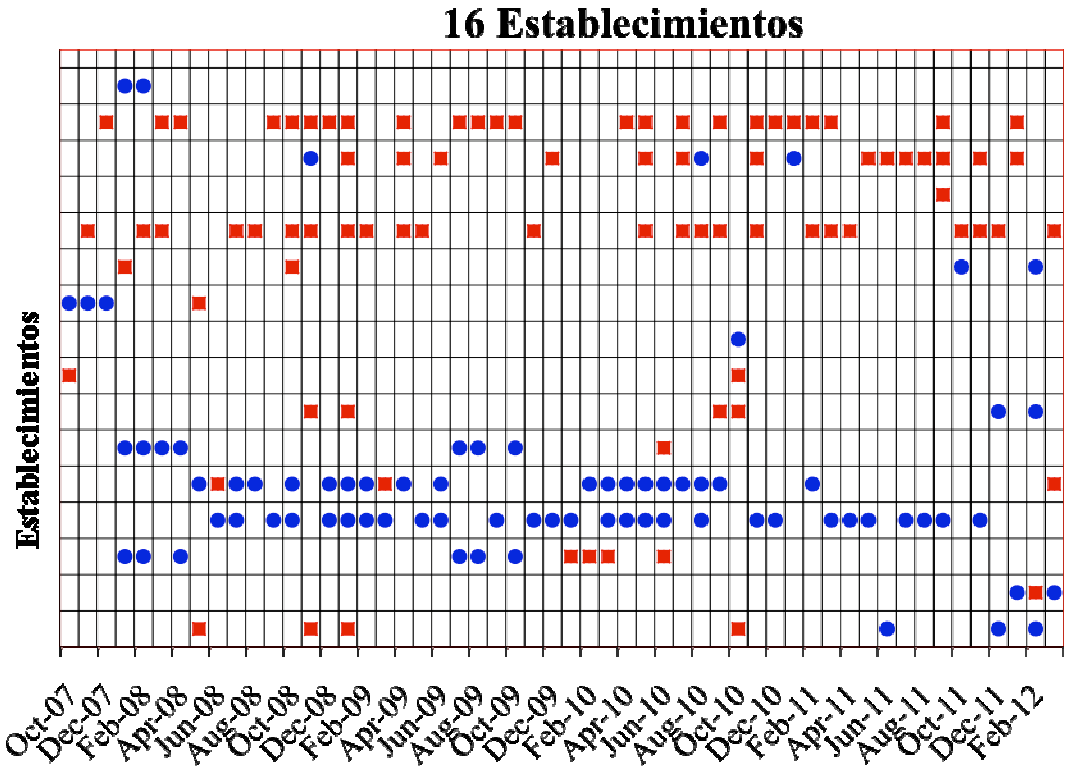
Anexo 11. Gráfico de precios máximos (●) y mínimos (■). Isla 5. Producto 1



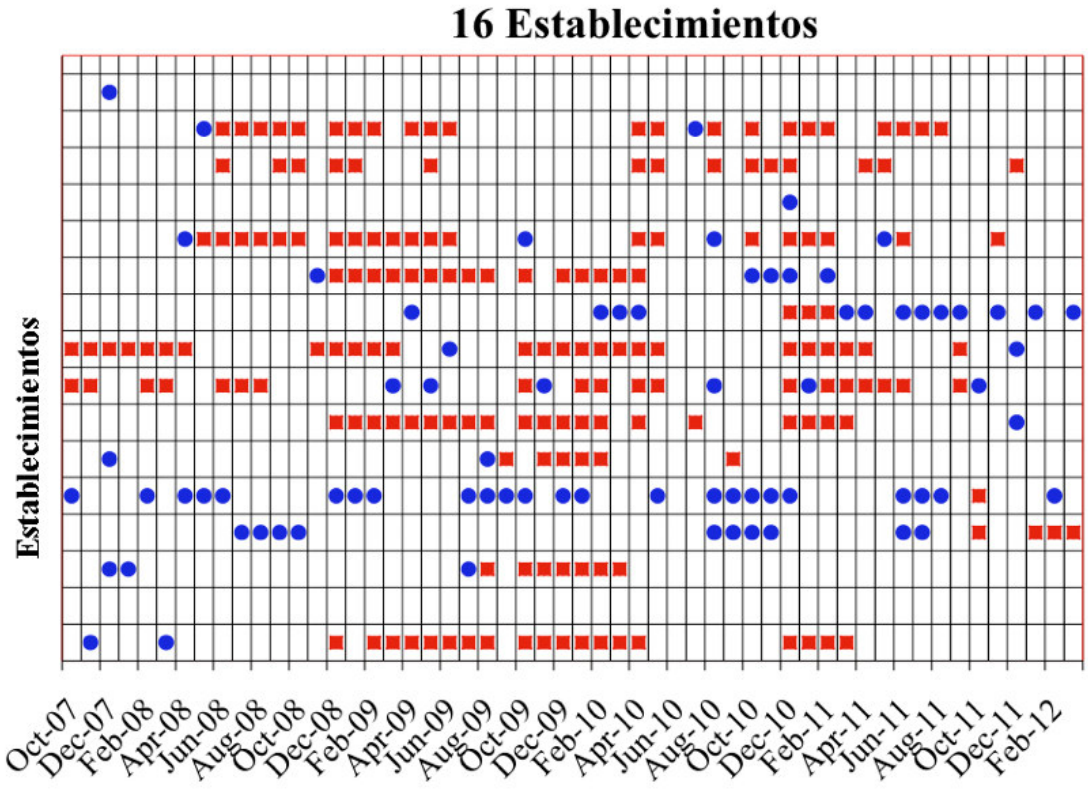
Anexo 12. Gráfico de precios máximos (●) y mínimos (■). Isla 5. Producto 2



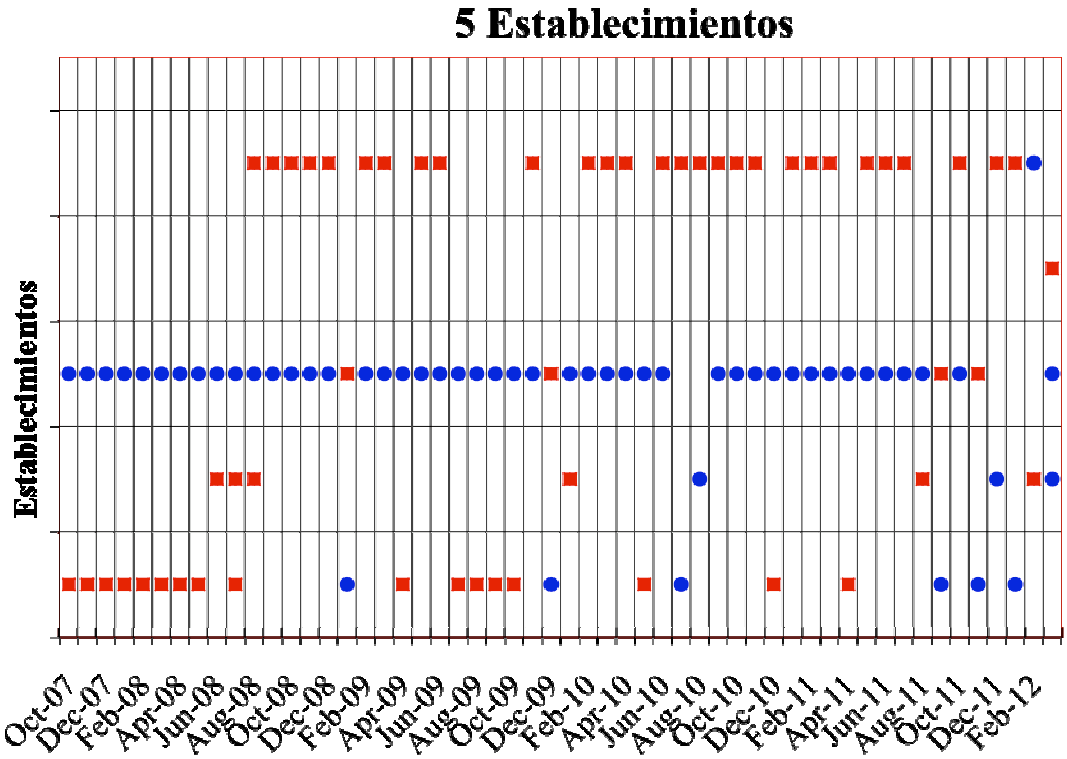
Anexo 13. Gráfico de precios máximos (●) y mínimos (■). Isla 6. Producto 1



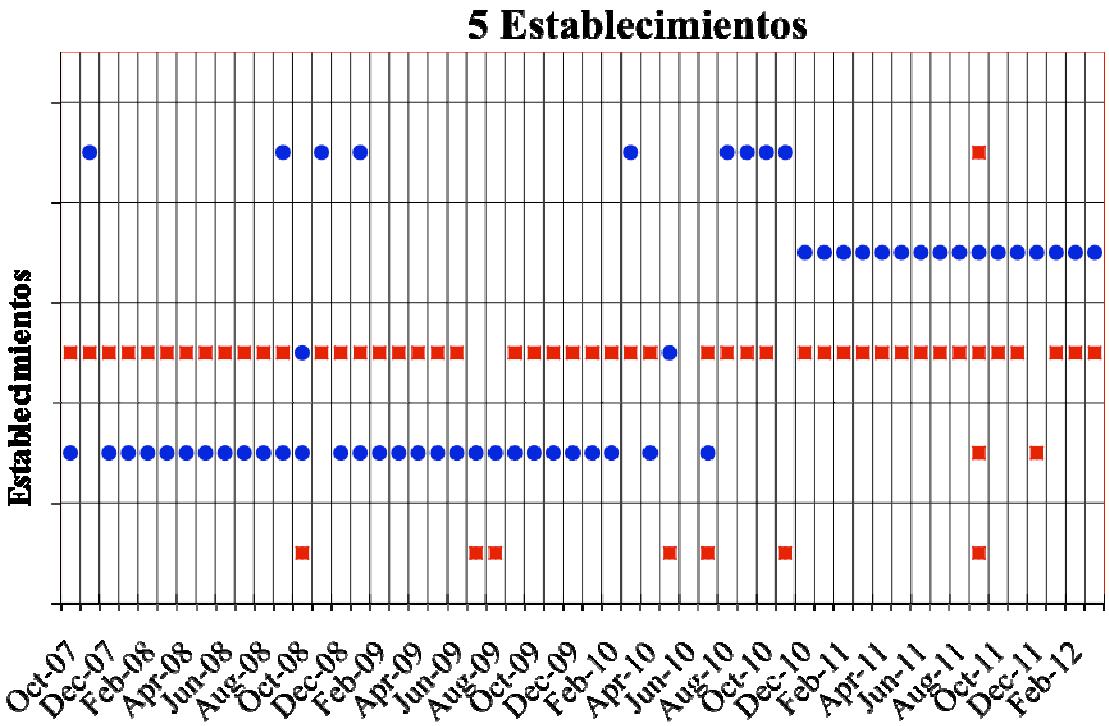
Anexo 14. Gráfico de precios máximos (●) y mínimos (■). Isla 6. Producto 2



Anexo 15. Gráfico de precios máximos (●) y mínimos (■). Isla 7. Producto 1



Anexo 16. Gráfico de precios máximos (●) y mínimos (■). Isla 7. Producto 2



THE PROGRESS OF QUALITY MANAGEMENT IN A TOURISM DESTINATION: A CASE STUDY OF THE BALEARIC ISLANDS.

SIDONIE DJOFACK
djsidonie@yahoo.fr
MARCO ANTONIO CAMACHO ROBLEDO
Departamento de Economía de la empresa
Universidad de las Ilas Baleares

Resumen

The Balearic Islands has quality as one of the key elements of its strategy to maintain its position as a leading international tourism destination. This strategy of excellence through tourism quality can be achieved through the implementation of different models, including the international ISO 9001:2000 quality standards, the European Model for Business Excellence, the Spanish tourism quality plan.
We first aim to investigate the evolution through time of each model in order to demonstrate the increasing importance of quality systems in the Balearic tourism industry (ISO 9000, EFQM, and the specific quality system of Spain). Secondly, we compare the efficiency among these models. The methodology is based on literature review and data analysis of the certified companies. Our study focuses on the importance of the expansion of quality systems in the Balearics and its effect on the rejuvenation of a mature tourism destination.

Palabras clave: quality management, quality systems, Balearic Islands, tourism destination.
Área Temática: Economía del Crecimiento económico.

INTRODUCTION

The competitive environment in the tourism industry is increasingly hostile and unforgiving (Robledo and Garau, 2001). Companies should look for strategies that allow them to confront successfully the challenges. The quality has become one of the key variables of competitiveness that tourism companies should incorporate to their corporate strategy. Changes in tourism demand and supply are forcing tourism destinations, especially mature ones, to re-think their production and organisation processes. Improving quality in tourist destinations can be taken as a basic to achieve an acceptable level of competitiveness. To improve the image of the product, enrich and diversify the supply so as to achieve a high degree of product differentiation are some of the aspects relating to the qualification of tourist destinations. It will also enable reaching benefits to face up to competitors and therefore achieve greater market share (Solsona, 2000). In order to protect the long-term interest of tourism company, they are seeking strategies to forge and maintain an ongoing relationship with their customers (Behara and Gundersen, 2001). In other words, Robledo and Garau think that many organisations have found that having a Quality Management System built around national or international standards or awards provides a framework upon which the organisation can build its quality foundations.. There are different models or frameworks used to implement Quality Management Systems. Among the available alternatives, we focus on the model proposed by the European Foundation for Quality Management (EFQM), the international standardization organisation (ISO 9000) and the Spanish Tourism Quality Plan. ISO and the Spanish Tourism Quality Plan are basically quality

assurance systems. The next step is Total Quality Management (TQM), which is a more global approach than the earlier and more oriented towards the continuous improvement and the processes in organisation. The importance of TQM in the tourist industry has risen to an extraordinary level because of the change in preferences of tourist's behaviour and the growth of competitiveness of new tourist destinations. Several authors agree that the EFQM model and ISO 9000:2000 are no alternatives as the EFQM model implies a greater commitment and a step towards total quality (Boulter and Bendell, 2002; Camisón et al., 2007; Nieto and Ros, 2006).

The aim of this article is to investigate the evolution through time of each model in order to demonstrate the increasing importance of quality systems in the Balearic tourism industry (ISO 9000, EFQM, and the specific quality system of Spain). Secondly, we compare the efficiency among these models. Our study focuses on the importance of the expansion of quality systems in the Balearics and its effect on the rejuvenation of a mature tourism destination.

The methodology is based on a wide review of specified literature in these topics. In this paper, many of these works are reported in the first section devoted to theoretical analysis and secondly, we deal with the research methodology which highlights data analysis of the certified companies.

1. THEORETICAL FRAMEWORK

1.1. Tourism quality as a strategy for Spain

The main role that quality plays in the Spanish tourism industry is highlighted in many strategic and marketing plans at both regional and national levels. Among such plans, the most important one is "FUTURES" (Secretaria General de Turismo 1992, 1996) which is devised to promote the competitiveness of the Spanish tourism (Robledo and Garau 2001). The "Plans of Tourist Excellence" arose more than ten years ago to contribute to the recovery and regeneration of certain tourist destinations (Lardiés, 2004). The plan included a diagnostic section where the major weaknesses of the Spanish tourism model were examined. According to the plan "FUTURES", an appropriate response to these conditions should be based primarily on improving the quality and applying other measures designed to develop competitive advantages (Zabalza, 1999; Robledo and Vadell, 2001). As an outcome of this process the Spanish Tourism Quality Plan (Plan de Calidad Turística Española) was developed.

More recently, the Horizon 2020 plan developed by the General Secretariat of Tourism and approved by the Cabinet meeting on November 8, 2007 has diagnosed a number of critical factors which will lead to the tourism strategy for the future of Spain. The following aspects are highlighted (Secretaría General de Turismo, 2007):

a. The excessive dependence of the product "sun and beach"

b. The tourism indicators are stable while the competition increases;

c. Productivity grows more slowly than the delicate economy sector;

d. There are some highly saturated places that represent a risk to environmental and urban sustainability;

e. a small degree of consciousness in comparison to the importance of tourism; f. There is a general deficiency in the areas of vocational training that hinders the improvement of quality tourism;

g. Tourism policies have been too focused on promoting the destination neglecting the aspects of the tourism product.

h. There is a need for greater dialogue between the public and private sector for the development of quality processes, innovation and continuous improvement of the sector.

As a result of this situation, the Tourism 2020 plan proposes six areas of work which enable the development of strategic in tourism industry. The overall objective is based on the Spanish model of tourism in sustainable development and competitiveness in the medium and long term. Consequently, quality appears to be an underlying issue in Tourism plan 2020. It will be necessary to emphasize upon vocational training of people, a belief in continuous

improvement, better processes, and greater efficiency in product development. These aspects are in relation with quality and thereby imply a willingness to pursue the strategic concept in order to maintain the tourist industry levels in the one hand and not succumb to the high competition in the sector in other hand, that appears to the new world emerging destinations.

1.2. Quality Management Systems

The development of approaches to managing quality (Froman 1994, Lockwood, Baker and Ghillyer, 1996) shows an evolution of the principles that relied on inspection of the finished product through the quality control and quality assurance through the Total Quality Management. These approaches are not mutually exclusive as one approach builds towards the next, because the Total Quality Management approach is introduced. It does not mean that quality inspection or quality control has to be ignored (Robeldo and Garau,, 2001). As indicated earlier, there are four kinds of quality systems (Lockwood, 1996, Robledo and Garau, 2001, Eraqi, 2008):

- **Quality inspection systems:** The earliest and probably the easiest approach to monitor, the quality is the inspection approach. This is a simple approach that relies on identifying defectsÿÿn a product or service before they rÿÿch the customer by introducing inspection procedures. It is based on having specification of what the product should be liked and checking the product against that standard. If defects are discovered then that product will be sent back for rework or scrap. The floor housekeeper checks rooms after servicing by a chambermaid offers a good example of inspection. This approach only works for tangible part of the operation.
- **Quality control systems:** The quality control system recognises that quality should be designed into a detailed specification and that quality is made throughout the production process. Quality control will not improve product or service quality, it will only highlight when it is not present. As a result, inefficiency and waste are not reduced. Indeed, the whole focus of quality control is on mistakes.
- **Quality Assurance Systems:** The quality assurance system makes attempts to design quality into the process so that things cannot go wrong. If they are wrong, they are identified and corrected. The emphasis here is on planning and preventing problems from arising at the source. Quality assurance must also involve development of a new operating philosophy that is proactive rather reactive. The ISO 9000 series is the most extended model of comprehensive application of the quality assurance philosophy. We can also add in this quality assurance, the specific quality system undertaken within the framework of the Spanish Tourism Quality Plan.
- **Total Quality Management Systems:** TQM is a system designed by Edward Deming and Joseph Juran and aims to use Statistical Control Process (SCP) in stages of manufacturing (Gilbert, 1992; Hyde, 1992; Martin, 1993). Of course, the implementation of these stages needs to be guided by TQM principles to guarantee sustainability in the tourism business. The key differences between quality assurance and total quality management are in the TQM focus on the customer (both internal and external), the nature of involvement within and outside the organisation, and the inclusion of a continuous improvement philosophy. The driving force of any TQM approach is its focus on the satisfaction of customer needs. The whole system of operation is directed at customer satisfaction. Therefore, it involves the whole organisation including suppliers looking for ways of continuously improving products or delivered service. The TQM system places more emphasize on the people in the organisation and their role. There is a move towards teamwork and employee participation, empowerment and motivation. It is more a culture than a system.

1.3. Quality Management Models

Quality is difficult to define. The decision of how it is implemented depends on the business strategy adopted by each organisation. Thus, some organisations choose to implement the

quality in order to get a quality price. On the contrary, others decide to introduce and certify a system of quality management, or simply certify the quality of a product or a service, etc.

1.3.1. Quality assurance according to ISO 9000 standards

In 1987, the International Standardization Organisation ISO 9000 was adopted by 101 countries as a national standard of quality assurance (Anderson, Daly and Jonson, 1999). These so-called ISO 9000 standards were originally designed for manufacturing of products, but they are now used in a variety of organisations and industries. The appearance and acceptance of ISO 9000 series as a reference standard has facilitated the harmonization of quality system approaches within the European Union (Dale, 1999). The goal of the application of such standard is twofold (Padron, 1996): on the one hand, it appears as a guide for the development, implementation, and internal assessment of a quality management system; on the other hand, as a framework to achieve certificate that visibly demonstrates a firm's commitment to the quality. Therefore, the implementation of quality system based on ISO 9000 standards and its posterior certification are going to redound, at least in theory, upon a higher confidence of clients with respect to the company operations and products/services. Most of the companies overlook the process of interpreting and adapting these standards to their inherent characteristics and capabilities. As a result, there is a misunderstanding with respect to the objectives of the standards. Nevertheless, the number of companies which use these standards in order to implement and obtain the certification of their quality system is continuously increasing (Casadesús and Heras, 1999; Durand et al., 1993; Forum Calidad, 1998, 2001; Iracheta et al., 2000; ISO, 2000b). ISO 9000 standards identify general criteria that help companies to articulate quality management system. These standards are generic, either in contents or in scope of application (Besterfield and al., 1995), because they may be applied to any organisation independently of its size or activity sector. The generic and flexible nature of ISO 9000 standards allow company managers to understand them, to interpret and to decide the best way to implement the framework provided by the company's strategy and its organisational characteristics (Kaye and Dyason, 1998). Thus, this matter brought up the researchers to investigate on the reasons for implementation, details of the process, and results. The phenomenon of its quick development led to a belief that it was a great advantage for a company to obtain registration. Along the same line, many researchers tried to identify the impact of ISO 9000 certification on the results and management of companies. Researcher group has not found any strong positive relationship between certification and results (Terziovski, Samson, and Dow 1997; Hua et al. 2000; Lima, Resende, and Hosenclever 2000; Singels, Ruel, and Van der Water 2001; Wayhan, Kirche, and Khumawala 2002). However, there is another group which points out an influence of registration on product quality (Romano 2000; Sun 2000; Withers and Ebrahimpour 2001). Taking out any influence derived from implementation, some authors advocated that certification could be a good first step toward a total quality management (TQM) system raising awareness of quality among workers and a good climate in order to implement it (Taylor 1995; Tummala, Rao, and Tang 1996; Skrabec 1999; Escanciano, Fernandez, and Vazquez 2001). As regards this point, the implementation of the standard was advised with the aim of implementing TQM in order to obtain maximum benefits from the registration (Brecka, 1994; Meegan and Taylor 1997; Huarng, Horng, and Chen 1999; Hughes, Williams, and Ryall 2000; Sun 2000; Gotzamani and Tsiotras 2002). Recently, the new version ISO 9001 has been published in 2008. The final draft (FDIS) has recently been approved and distributed, and according to Reyero(2008) , this could be the final form of the official edition of the standard. This standard is the well-known and probably the most used in the story because it is implemented in about a million companies distributed in 170 countries. The ISO requires that every five years, one draws out the revision of the standards, approximately 17000 edited norms. However, taking into account the changes between versions, 1994 and 2000 ones were so radical. It was decided to leave "rolling" for three years the implementation of former version and start the revision in 2003 instead of 2000. That is why the new version appeared later. Certainly, what is involved now is series of amendments and not a

real revision, because ISO was very clear about how ISO 9001:2008 will be. The new edition does not contain different or new requirements. However, the new version has received a lot of comments about urgent changes in the rule of norms, which by now, has been probably "parked" until the next thorough revision, presumably in 2013. Fairly, there is still a slight incompatibility between these management standards. Thefore, one will have to wait the 2013 version to achieve and solve the integration between these standards. The publication of the new version of ISO 9001 in Spain took place on 17 November 2008, and then it appeared in English and French two days later. ISO and the International Accreditation Committee have already announced the transition period of one year. After this time, ISO 9001:2000 certificate will not be accepted anymore. In summary, the new version of the standard is closer to a TQM system (Reyero, 2008). In particular, it tries to improve human resource management, customer focus, and leadership. According to managers, and supported by the conclusions of previous authors, the "soft" variables of TQM are responsible for the improvement in results derived from TQM implementation. It is an opportunity for companies to achieve better performance by applying adequately the new version. This new standard could also be a good first step toward a TQM system.

1.3.2. Quality assurance according to the Spanish Tourism Quality Plan: a specific quality standard.

The Spanish Tourism Quality Plan[1] was an initiative of the secretary of trade, tourism and SMEs of Spanish government aimed to improving quality in the Spanish tourism industry. For that purpose, the plan included two basic objectives. Firstly, its aim was to develop a competitive advantage through the implementation of the quality system that is specific and customised for each subsector of the tourism industry. The quality system would be a competitive tool for the promotion of Spain as quality destination. Its second aim was to implement a quality brand system developed through a certification process based on the compliance with certain quality standards. The brand is considered to be a commercial and promotional tool for companies which obtain it. The approach taken was to give a differentiated treatment to each subsector of the tourism industry in recognition of the variety and complexity of the industry. As a result, the implementation process was followed step-by-step where sector-specific institutes, responsible quality systems and standards for every sector were established. In addition, the private sector was heavily involved in the planning process. As a result of this plan, fours institutes have been created so far under the initiative of the Secretary of Trade, Tourism and SMEs. The central institute is the ICTE or the Institute for Spanish Tourism Quality (Instituto para la Calidad Turística Española), based in Madrid. Its task is to be responsible for the general management of the quality plan and for co-ordination of all the different sector-specific institutes. It is credited as a certification entity by the relevant authority. The other three institutes are sector- specific, The ICHE or the Institute for Spanish Hotel Quality (instituto para la Calidad Hotelera Española), the INCAVE or the Institute for Spanish Travel Agencies Quality (Instituto para la Calidad de las Agencias de viaje Española) and the ICRE or the Institute for Spanish Restaurants Quality(Instituto para la Calidad en Restaurantes Españoles). They had previously designed their own quality system, which was revised to make it convergent with the philosophy of the ICTE. Although the later three institutes (ICHE, INCAVE and ICRE) have been responsible for designing a quality system specific to hotels, travel and agencies and restaurants respectively, they share the same overall philosophy of the quality assurance(Robledo and Garau, 2001). The assess whether a particular company satisfies the requirements of the system; if they do so, these institutes award the company with the brand "Q" , which is the same brand regardless of the sector that the company belongs to(Robledo and Vadell 2001;Herrero de Laorden, 2002;Barba, 2005;Gomez, 2007). The use of a single brand has the purpose of creating synergies and it would become the unique brand of the

[1] Robledo and Garau (2001), The Spanish Tourism Quality Plan: A pioneer initiative in the development of industry-specific quality standards, vol.49,n 1, Tourism.

Spanish tourism Quality[2]. The Standards developed by the tourist quality system (SCTE) are between the ISO 9000 and the Business Excellence Model (EFQM)(Hernandez and González, 2003).These standards represent a philosophy of quality management and continuous improvement more extend than that offered by the ISO 9000 series. These goals by 2010 are in our agendas. The standardization of new sectors related to tourism, which are showing great interest in participating in the processes of quality assurance, sustainable growth in the number of companies, the continuous improvement of our content standards, will be treated by the technical standards committees. (info@icte.es or in www.calidadturistica.es) .

Finally, the viability of the ICTE also depends on its competition with others systems. Now, it is primarily the ISO 9000. Although the ISO 9000 standards in its 1994 version were compatible with the ICTE standards, the important changes in its 2000 version (Larsen and Häversjö, 2000) are not so clearly convergent[3]. However, Robledo and Garau (2001) concluded that convergent or not, as there is a cost attached to every certification, each company has to decide what certification does it want-ISO, ICTE or both.

In the following table we expose the different models of quality management and the quality system that belong to.

Table1: membership of quality system to quality model.

Quality system	Quality model
Inspection	
Control	
Assurance	-ISO 9000 -"Q" Spain Quality
Total quality management	EFQM

1.3.3. Total Quality Management and the EFQM Model

Many tools, methods and models have been developed throughout the world in order to give substance to the concept of TQM. In Europe, the Model for Business Excellence, issued by the European Foundation for Quality Management (EFQM) stem from the different quality prizes which was started to encourage the implementation of total quality plans in companies. These plans are used with increasing frequency as tools for evaluating and improving the quality system of a company. The first award was established in 1951 in Japan. In 1987, the 'Malcolm Baldrige' award was created in the USA. This initiative was followed in Europe in 1991 when the European Foundation for Quality Management (EFQM), together with the European Organisation for Quality (EOQ) and the European Commission created the 'European Quality

[2] Under the auspice of the ICTE , other sub-sectors of the Spanish tourism industry , such as campsites, ski resorts, bars and nightclubs, are developing their own quality systems in accordance to the principles of the Spanish Tourism Quality Plan(Robledo and Vadell, 2001).

[3] Camisón et al. (2007) identified a number of differences between the two:
1) a The ICTE standard considers the requirements and recommendations of ISO 9000. In fact, without any direct indication, ICTE standard specifies that, must be determined self quality system to tourism industry
2) ISO 9000 is specific to the implementation of a System of Quality Management, while ICTE standard include the specification of the service that must implement the company which adheres to the system.
3) From a generic point, ISO 9000 is more versatile and applicable to any organization while ICTE standard is applicable only in the tourism sector.
4) On return to the previous, from the customer point of view, the ISO 9000 certification does not guarantee a specific level of quality but the service will meet the specifications set by the establishment. The quality system of ICTE marks a minimum quality of service that ought to complete the organisation. Both systems are compatible and complementary; to be desirable for the tour company that chooses the philosophy of quality coverage both certificates.

Award' in order to improve the quality and reliability of products and services. It took two forms:

• The 'European Quality Award' which rewards the company that is the maximum exponent of TQM in Western Europe;

• The 'European Quality Prize' which rewards some companies that show excellence of quality management as a basic process of continuous improvement.

The European quality model allows strengths and weaknesses to be clearly distinguished focusing on the relationship among the staff, process and results. Processes are the mean through which a company guides and liberates its staff aptitudes so as to get results. Therefore, processes and people are the 'agents' that provide 'results'. The model that serves as a basic for the European Quality Award forms a management system which emphasizes upon maintaining leadership for achieving quality, formulating a policy and strategy to follow, developing the appropriate staff and resource management and guiding the design of all the processes in the company towards customers, in order to reach excellence in results: customer satisfaction, personnel satisfaction, a positive social impact and some economic results which allow the achievement of a competitive advantage. Each of the nine elements that appears in the model matches the criteria used by the EFQM to evaluate the level of excellence of company, and the percentage is the weight given by the element in the self-evaluation (on a total of 1.000 points).

Desiring to improve the competitiveness of tourism in destinations and thereby to enhance employment prospects, the European commission identified the quality award criteria framework that was developed by the European Foundation for Quality Management (EFQM) as a model in which one can refer to improve the performance of tourism destinations in Europe. As a non-prescriptive framework, the Quality Award can be used to assess an organisation progress towards achieving sustainable excellence of all aspects of performance. It is based on the premise that `Customer satisfaction, employee satisfaction and society impact are achieved through leadership, leading ultimately to excellence in business results. The objective of the EC research was to develop the EFQM model to manage quality in an integrated manner in tourist destinations. The diffusion of these ideas and methods of tested efficiency in the tourist firm were scarce. Although the current scoring of the EFQM model has been widely accepted by both academic and practitioner as a sound approach for self-assessment, writers such as Porter and Tanner (1996) and Siow *et al.* (2001) are in agreement that organisations have encountered problems applying them because the scoring criteria are too generally defined. As a consequence, large scoring variations are common, especially with inexperienced assessors (Yang *et al.*, 2001). In a survey of the British part of a major European project on the use and benefits of self-assessment, Coulambidou and Dale (1995) found that the majority of the companies experienced problems with measurement, including variations in scoring. Other writers such as Lascelles and Peacock (1996), Teo and Dale (1997), Schmidt and Zink (1998) have also identified and criticised the scoring system of the EFQM model. However, since the EFQM model was launched in 1991, little research has been undertaken into the scoring system. Total Quality Management is an excellent model to use after successful implementing an ISO 9000 Quality system. The EFQM model makes it easy to enlarge the quality management viewpoint and to take an objective looking at the organisation and the results is achieved. It helps to relate the different aspects of the organisation to the balanced viewpoint. Furthermore it helps to set priorities in the improvement projects, and it is an adequate model for evaluation of the achieved improvements. The combination of ISO 9000 and the EFQM-model, within the framework of Total Quality Management has proven to be very useful and effective.

1.4. A Comparison between ISO 9000, EFQM and the Spanish PCTE-

In the existing literature, there are many comparisons between ISO 9000 and EFQM (Santoma,1990;Cruz,2001;Karapetrovic and willborn,2001; Boulter y Bendell, 2002; Magd y Curry, 2003;Marimon,2004; Sun *et al.,2004;* Alonso *et al.*, 2005; Corma, 2005; Nieto y Ros,

2006;Camisón *et al.*, 2007. But by now, we have not seen any comparison between ISO 9000 and PCTE or EFQM and PCTE. Therefore, we propose this table among these models.

Table2: comparison of quality models

Item	ISO 9000	EFQM	PCTE -Spain
Organisation	I.S.O	E.F.Q.M.	P.C.T.E
First version	1987	1990	1998
Last version	December 2000	1999	2005 (transformation en public standard, UNE (Una Norma Espanola).
Geographical scope	World	Europe	Spain[4]
Name in the first publication	ISO 9000 Quality Assurance Standards	Business model EFQM of excellence	Quality ICTE
Title of the actual Model	ISO 9001:2000, Quality Management Systems, Requirements.	Excellence EFQM model	The Spanish Tourism Quality Plan
Applicability and sector	All kinds of private or public organisation	All kinds of private or public organisation	Tourism sector
Recognition	ISO certificate	European Quality Price	ICTE Quality certificate
Criteria	Basic elements of total quality system	Common aspects of the general direction of the organisation	Basic elements of a specific quality system in tourism
Emphasis	Effective direct practice	Business Excellence	Differentiation strategies of the tourist product o service.
Basic principles	-customer focus -leadership -staff involvement -approach to the process -system approach to management -continuous improvement -objective approach to decision making -mutual beneficial relationship with the supplier	-Results Orientation -Customer Focus -Leadership & Constancy of Purpose -Management by Processes & Facts -People Development & Involvement -Continuous Learning, Innovation & Improvement -Partnership Development -Public Responsibility	
Purpose	Guidance and guidelines for internal assessment of the quality system	Continuous improvement in all management aspects and organisational culture	Implement specific quality system for the different subsectors of tourism.
Methodology of improvement	Deming cycle PDCA(Plan ,Do, Check, Act)	RADAR(Results, Approach, Deployment, Assessment and Review , is a method for scoring)	Specific quality management system ,continuous improvement and satisfaction of the customer
Contribution	collect the structured way of the key elements of quality	Introduce explicitly the financial and non financial results	Determine the common quality system to tourist industry including the specificity of the service with a level of minimum quality.

The late 1980s and early 1990s saw a small revolution in quality management, with the introduction of the ISO 9000 international standards in 1987, the Spanish Tourism Quality Plan

[4]Outside branches of Spanish companies can get certification from PCTE

in 1998 and business excellence models in 1990. These models, representing criteria for sound quality assurance and management systems, respectively, demanded methodologies for evaluation of the actual versus the desirable system status and quality audits and self-assessments. In order to increase competitiveness, organisations must use all of these methodologies. The models are required for examination of the suitability and effectiveness of the existing quality system, and possibly for external recognition of the system via ISO 9000 registration and Spanish PCTE. Both models are necessary for identifying the enablers of continuous improvement and possibly for recognition of business excellence through attainment of a quality award (EFQM model). However, there is a need to analyze the commonalities and differences between these three methodologies in order to provide a theoretical background for their alignment and possible further integration. The three models have the same aim: develop the quality systems in the organisation in order to achieve more benefits although the way to implement is proper to each model. The ISO 9000 family of standards represents an international consensus on good management practices with the aim of ensuring that the organisation can continually deliver the product or services. These practices meet the customer's quality requirements and applicable regulatory requirements, while aiming to enhance customer satisfaction, and achieve continual improvement of its performance in pursuit of these objectives. These good practices have been distilled into a set of standardized requirements for a quality management system, regardless of what your organisation does, its size, or whether it is in the private, or public sector. Considering the idea developed above, the Spanish PCTE is approximately the ISO model. But it is important to note that PCTE is a specific model in only tourism organisation. ISO 9000 standards and PCTE are first step on the road to excellence and total quality management (Dale, 1999; van der Wiele *et al*., 2000a).The ISO 9001:2000 model provides only the minimum requirements for policies, processes and resources that must be in place for an organisation to have a quality system and, by virtue of its existence, to assure customer satisfaction. The PCTE assures the good implantation of quality in the organisation, in other words it helps to determine the common quality system to tourist industry including the specificity of the service with a level of minimum quality. Dale (1999) and van der Wiele *et al*. (2000a) think that ISO 9000 standards and business excellence models really reflect different levels of maturity in performance excellence. Consequently, corresponding evaluation methodologies are also able to measure differing levels of such excellence.

2. RESEARCH METHODOLOGY

We have developed a study focused on tourist companies certified by ISO 9001, EFQM and ICTE –Spain. According to the ISO 9001 population, we contact each certified organism as AENOR, Bureau Veritas, Der Nordske Veritas and Similar, Tüv, Lloyd's Register Quality Assurance, and successful, we gather 72 companies which represent the whole population. According to the EFQM, we contact the IDI (Instituto de Innovacion Empresariales), all the population has been studied because of its small size (11 companies). As the last model is concerned, the population of ICTE is 87 companies certified from the Institute for Spanish Tourism Quality (Instituto para la Calidad Turística Española). This was the method employed to have our data-gathering tool. Our purpose is to investigate the evolution through time of each model in order to demonstrate the increasing importance of quality systems in the Balearic tourism industry (ISO 9000, EFQM, and the specific quality system of Spain)

2.1. Results

The surveyed companies certificated by ISO 9001 standards are 72 in the Balearic Islands in which the evolution is shown as follows (Fig. 1 & 2)

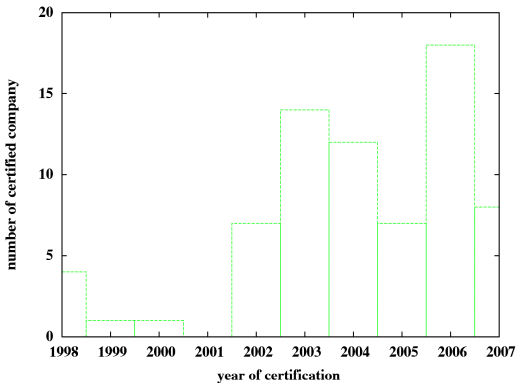


Figure 1: evolution through time of certified tourist companies in ISO 9001

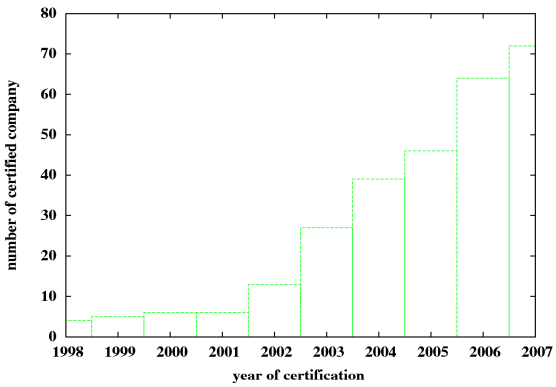


Figure 2: accumulated number of certified tourist companies in ISO 9001

The histogram plotted in Fig.1 shows the number of certified company versus the year of certification. The companies have been chosen according to the year of the publication of standards which was in 1987, and in 1990 the norms were setting in the companies. As it is shown in the histogram, there are few companies which were certified in Balearic before the year 2001. This proves that tourist industry in Balearic started the certification very late. But the number of the certified companies has rapidly increased probably because of the benefits obtained in the companies that have earlier implemented it. Fig.2 shows continuous evolution which can be interpreted that ISO 9000 system has advantages in tourist company. It is also important to notice that the tourist subsector that has the higher certificated number is the hotels followed by travels agencies. According to Salas and Torres (2007) investigations about ranking of communities in Spain, Balearic Islands is the first tourism destination. Thus, we think that this number of certified companies by ISO 9001 is not enough.

We can conclude by saying that ISO 9000 has a future. Rather than as a standard with enforced international compliance, it should become an internationally recognized tool for quality assurance and improvement. In a level playing field with other available tools, ISO 9000 will live and not perish.

Fig. 3 & 4 show the histogram (evolution though time and accumulated, respectively) of certified companies in the specific model of Spain. On the contrary to ISO 9000 model, the histogram exhibited in Fig 3 & 4 show that some companies have been certified by ICTE from the same year it was created. This acceptance of this system can be explained by the fact that it is specific for tourist sector. We see that the year 2002 was the glory one with more than 20 certifications. For the ICTE system, the most certified companies are hotels and travel agencies and restaurants. Although the majority group is constituted by the aforementioned subsector, we can notice a very few number of certification in the other subsector of tourism.

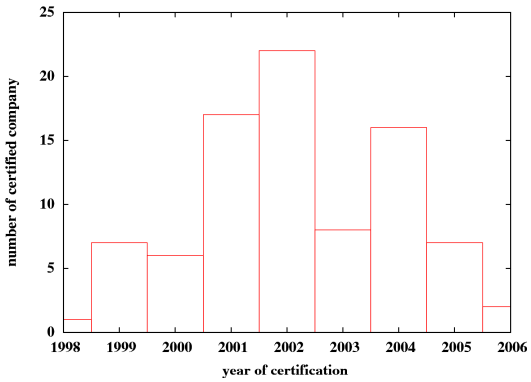


Figure 3: evolution through time of certified tourist companies in ICTE-Spain

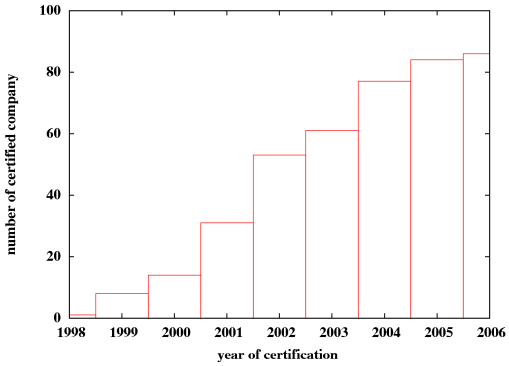


Figure 4: accumulated number of certified tourist companies in ICTE-Spain

Finally we take the privilege to compare the evolution of certified companies between ISO 9001 and ICTE. The result is surprising, the evolution of the curve is approximately the same; but finally we found an important evolution of ICTE than ISO 9001. This result can be interpreted as an important model for tourism industry, and also an appropriate system in tourism destination.

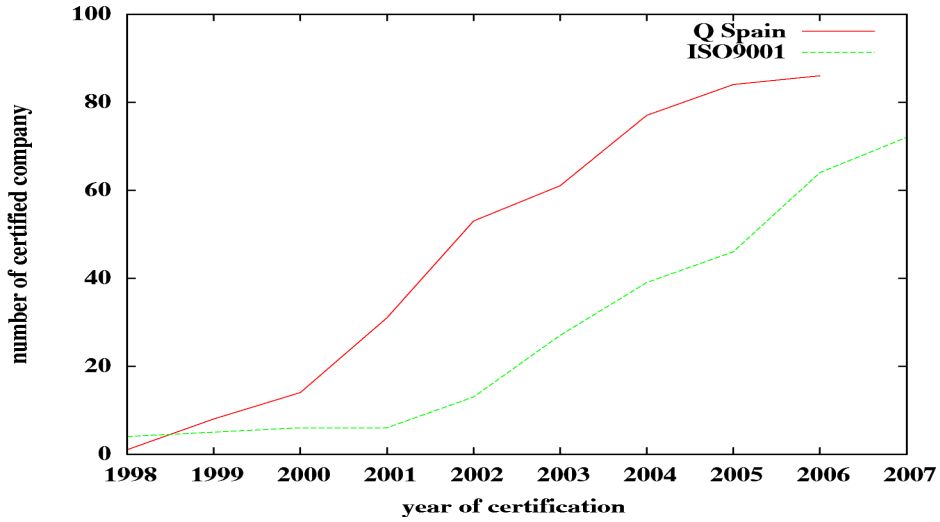


Figure 5: comparing the evolution of accumulated numbers of certified tourist companies between ISO 9001 and ICTE-Spain.

Finally, the last model of quality management is the EFQM, which contains only 11 companies. The main objective of this model is the Self-Assessment which is a comprehensive, systematic and regular review of organisation activities and results measured through the Excellence Model. RADAR is the method for scoring when using the excellence model results, approach, deployment, assessment and review. Considering the existing literature review on the step by step towards the total quality management, the EFQM is also situated at the end of the quality models. We also think ISO 9001 and ICTE models have predominated in Balearic in such a way that companies do not focus more on the EFQM model.

Fig 6 shows the evolution of the EFQM in tourism Balearic Islands mostly constituted by hotels, airport and tourist port.

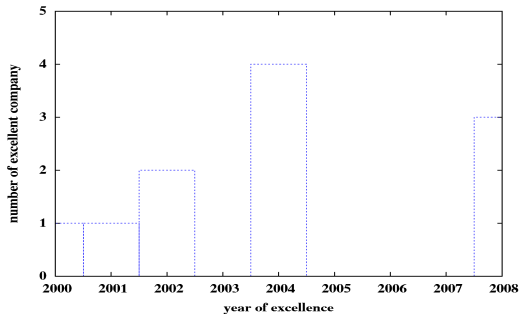


Figure 6: evolution through time of certified tourist companies in EFOM

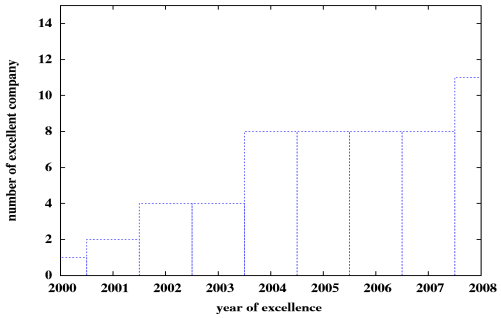


Figure 7: accumulated number of certified tourist companies in EFOM

In Fig 8, we also depict the histogram of scoring companies. We have 5 companies gathering more than 250 points and 4 companies with more than 400 points. One remarks the renunciation of 2 companies because of cost attached to the implementation of this model.

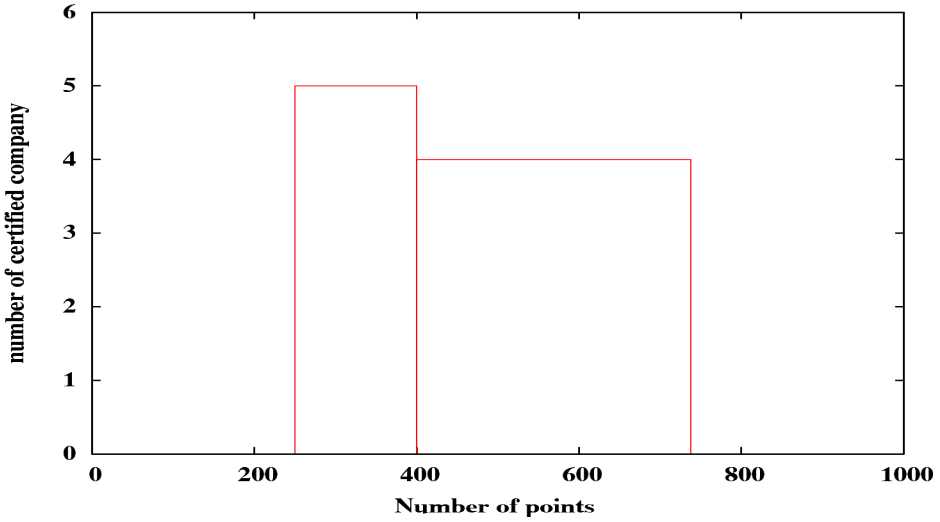


Figure 8: scoring certified tourist companies in EFQM

CONCLUSION

The Balearic Islands has quality as one of the key elements of its strategy to maintain its position as a leading international tourism destination. Our aim was to investigate the evolution through time of each model in order to demonstrate the increasing importance of quality systems in the Balearic tourism industry (ISO 9000, EFQM, and the specific quality system of Spain), and compare the efficiency among these models. We have understood that:

- The results show that there has been a significant increase in the implementation of quality systems in the Balearic tourism companies in the period.
- Increasing competition, a widespread of the quality culture and the promotion of that culture by institutions like ICTE, IQT and IDI are the reasons for that.
- That fact has resulted in an upgrading of the standards of those companies both at infrastructural level and systems and procedures.
- The tourism companies prefer to implement certifiable systems like ISO and Q (PCTE) for marketing reasons.
- ISO and Q are the preferred systems and there is not a clear preference for one or the other. Each one has its strengths and weaknesses.

- However, many companies might be using EFQM as a model, but only the ones that have obtained a prize or are registered in the catalogue of excellent companies published by IDI are easily traceable.

According to the data, we notice a competition between the models (87 companies certified by ICTE and 72 by ISO 9001). The outlook of ICTE institute is to introduce the specific system of Spain in the ISO 9000 series that can create more convergence between both models. However, convergent or not, as there is a cost attached to every certification, each company has to decide what certification does it want- ISO, ICTE or both (Robledo and Garau, 2001). But in the case of Balearic Islands, ICTE is currently most used because it responds perhaps to the specific problem face in tourist sector, and also the system can be used as a new strategic of rejuvenation of a mature tourism destination.

Typically, most organisations develop their strategy around a well-recognized business model. The ISO 9000 is geographically widespread and better known, but it is general and difficult to apply in comparison to the ICTE standards. Moreover, the ICTE is local, but it is specific, more detailed and easier to implement. As for the ISO 9000, It is more flexible and does not specify what levels a company must conform to.

BIBLIOGRAPHY

ALONSO, M; BARCOS, L. Y MARTÍN, J.I. (2006), "*Gestión de la calidad de los procesos Turísticos"*,Madrid: Síntesis.

BEAUMONT, N AND SOHAL, A (1999) 'Quality management in Australian service industries', *Benchmarking: An International Journal*, Vol. 6, No. 2, pp.107–123.

BUTTLE, F. (1996) "An investigation of willingness of UK certificated firms to recommend ISO 9000". *International Journal of Quality Science*, 1(2), pp. 40-50.

CAMISÓN, C. (1996),"Total quality management in hospitality: an application of the EFQM model", *Tourism Management*, 17(3), pp. 191-201.

CAMISÓN, C. (2004)," Estrategias de calidad turística. El papel de los sistemas de acredicación y certificados de calidad" *Mediterráneo económico. Las nuevas formas de hacer turismo*, 5, pp. 118-168.

CAMISÓN, C. Y BIGNÉ, E. (1995)," El enfoque de calidad en el turismo", En AECIT (1995), *¿España, un país turísticamente avanzado?* Madrid, Instituto de Estudios Turísticos.

CAMISÓN C; CRUZ S Y GONZÁLEZ T. (2007). "*Gestión de la calidad: conceptos, enfoques y sistemas"*, Madrid: Pearson – Prentice Hall.

Claver C.E.; Andreu G.R. and Quer R.D.(2006),"las ventajas de la diversificación estratégica para las empresas turísticas españolas. Una visión desde la teoría de recursos". Cuadernos de Turismo, nº 17, (2006); pp. 51-73 Universidad de Murcia

ISSN: 1139-7861

CONTI, T. (2007), "A history and review of the European Quality Award Model", The *TQM Magazine*, 19(2), pp. 112-128.

ERAQI M.L.(2008),"Quality management systems in tourism companies: the case of travel agencies in Egypt", Int. J. Business Excellence, Vol. 1, No. 3.

FÓRUM CALIDAD (2007),"Certificación en España según normas UNE-EN-ISO 9001 y 14001", XII Informe Fórum Calidad. *Fórum Calidad,* 180, pp. 22-27

FORUM CALIDAD (2008), "Revision de la norma ISO 9001, la nueva version : 2008", XX Informe Fórum Calidad. *Fórum Calidad,* 180.

GO, F. M. Y GOVERS, R. (2000), «Integrated quality management for tourist destinations: a European perspective on achieving competitiveness», *Tourism Management*, nº. 21 (1), pp. 79-88.

HERAS, I.; ARANA, G. Y CASADESÚS, M. (2006)," A delphi study on motivation for ISO 9000 and EFQM", *International Journal of Quality and Reliability Management*, 23(7), pp.807-827.

LANCELLES,D.M. Y DALE, B.G.(1990), "*The use of Quality Management Techniques"*,Quality Forum, Vol.16, No.4, pp.188-192.

LOCKWOOD, A. (1994)"Using Service Incidents to identify quality improvement points", *International Journal of Contemporary Hospitality Management*, 6 (1/2), pp. 75-85.

LOCKWOOD, A. (1996) "a systematic approach to quality . Quality management in hospitality", Lockwood,A. Baker,M.,and Ghillyer,A.(eds.)Quality management in hospitality.(14-32), cassell, London.

MADERUELO F.J.A. (2002)," Gestión de la calidad total. El modelo EFQM de excelencia", Vol. 12 – Núm. 10. MEDIFAM 2002; 12: 631-640

MARIMON, F. (2004),*La consultoría especializada en ISO 9000 en Cataluña: calidad del servicio y beneficios*,Tesis Doctoral. Departament d'organització, gestió empresarial i disseny de producte. Universitat de Girona.

MARTÍNEZ C.M. AND MARTÍNEZ L.A.R. .(2004),"ISO 9000 as a Tool for TQM: A Spanish Case Study", 20 QMJ VOL. 11, NO. 4, ASQ

NIETO, C.N. Y ROS, L. (2006)," Comparación entre los Modelos de Gestión de Calidad Total: EFQM,Gerencial de Deming, Iberoamericano para la Excelencia y Malcolm Baldrige. Situación frente a la ISO 9000", Presentado en *X Congreso de Ingeniería y Organización*. Valencia.

SALA R.M. AND TORRES S.T. (2007),"El turismo como elemento de crecimiento económico en el ámbito español" IX Reunión de Economía Mundial ,Madrid.20P

SERRANO B.A.M; LÓPEZ F.M. AND GÓMEZ L.R.(2007),"gestión de calidad y turismo: revisión e implicaciones para futuras investigaciones", Cuadernos de turismo, nº 20, (2007); pp. 251-266. issn: 1139-7861

SOLSONA, J.(2003),"gestión de calidad en espacio turístico rural, Gestión Turística", Edición Especial ISSN 0717 – 1811 Pág. 7 - 17

VAR DER WIELE, A.; DALE, B. Y WILLIAMS, R. (2000),"ISO 9000 series and excellence model: fad to fashion to fit", *Journal of General Management*, 25 (3) Spring, pp. 50-66.

CHANNELS OF TRANSMISSION OF INEQUALITY TO GROWTH: A SURVEY OF THE THEORY AND EVIDENCE FROM A PORTUGUESE PERSPECTIVE

ADELAIDE P. S. DUARTE
E-mail: maduarte@fe.uc.pt
MARTA C. N. SIMÕES
E-mail: mcsimoes@fe.uc.pt
Faculdade de Economia
Universidade de Coimbra (Portugal)

Resumo

Através de uma recensão da literatura teórica e empírica sobre a relação entre desigualdade e crescimento económico do ponto de vista da economia portuguesa, este trabalho identifica o sinal (previsível) da relação para este país em particular, bem como os mecanismos que permitem justificar e explicar o mesmo. A relação entre desigualdade e crescimento económico pode ser justificada com base em diversos mecanismos, o que explica a obtenção de resultados diferentes para a relação quando se analisam diferentes países durante o mesmo período de tempo, ou para o mesmo país analisado ao longo de vários períodos. Torna-se assim fundamental seleccionar os mecanismos apropriados para a análise da situação portuguesa. Idealmente, todos os mecanismos seleccionados deveriam ser testados, mas não sem antes se proceder à sua hierarquização de acordo com a sua relevância para a economia em concreto em análise. Uma correcta identificação e hierarquização permitirá derivar implicações mais correctas e concretas no que respeita às políticas de redistribuição do rendimento em Portugal.

Palavras-chave: desigualdade, crescimento económico, Portugal.
Área Temática: Crescimento Económico.

Abstract

We review the theoretical and empirical literature on the relationship between inequality and economic growth from the perspective of the Portuguese economy in order to identify the correct (predicted) sign for the relationship in this particular country and the underlying mechanisms that explain it. Different mechanisms relating inequality and economic growth can be at stake explaining why different countries can experience different outcomes in the same period of time or why the same country experiences different outcomes in different periods of time. It is thus fundamental to correctly identify the channels of transmission of inequality to growth in the Portuguese economy. Ideally, all the mechanisms selected should be tested, but prior judgments should also lead us to produce a ranking of the mechanisms according to its relevance for the economy under analysis. This correct identification and ranking leads to more accurate policy recommendations as far as redistributive policies for the Portuguese economy are concerned.

Keywords: inequality, growth, Portugal.
Thematic Area: Growth Economics.

1. INTRODUCTION

Why should we care about income (and capital) inequality in the Portuguese economy? Since the 1980's, Portugal has experienced a decrease in income inequality but this decrease has not been uniform across the income distribution, especially towards the end of the 20th century and the beginning of the 21st. According to recent European Commission reports on the social situation in the EU, Portugal was the EU_25 member state that presented the most unequal income distribution in 2001 and again in 2004, with a Gini coefficient of 41% and a rate of population at risk of poverty of 20.3% in the last year under analysis.

Income inequality *per se* is only relevant to the well-being of individuals if it is accompanied by a decrease in the rate of economic growth and by an increase in poverty. If inequality in the Portuguese economy is negatively associated to economic growth and positively associated to poverty this implies a reduction in the well-being of the Portuguese citizens. Furthermore, an economy experiencing high levels of income (and capital) inequality at medium levels of poverty will probably become more vulnerable to adverse shocks on economic growth, which in turn can lead to more poverty in the present, and most certainly in the future too if intergenerational mechanisms perpetuating poverty are in action.

The aim of this paper is to review the theoretical and empirical literature on the relationship between inequality and economic growth from the perspective of the Portuguese economy in order to identify the correct (predicted) sign for the relationship in this particular country and the underlying mechanisms that explain that sign. Different mechanisms relating inequality and economic growth can be at stake explaining why different countries can experience different outcomes in the same period of time or why the same country experiences different outcomes in different periods of time. It is thus fundamental to correctly identify the channels of transmission of inequality to growth in the Portuguese economy. Ideally, all the mechanisms selected as crucial for the Portuguese economy should be tested, provided data is available, but even if this is the case some prior judgments should lead us to a ranking of the mechanisms according to its relevance for the economy under research. This correct identification and ranking will enable us to derive more accurate policy recommendations as far as redistributive policies for the Portuguese economy are concerned.

The paper is organized as follows. In section 2 we provide a brief description of the Portuguese economy in terms of inequality and poverty in order to identify the mechanisms relating inequality to economic growth relevant for this particular country. In section 3, we analyze how these mechanisms work based on a survey of the theoretical literature on the subject. Section 4 is devoted to a systematization of the empirical evidence already collected on the theme. Finally, in section 5 we conclude and identify main avenues for future research on the theme applied to the Portuguese economy.

2. INEQUALITY AND POVERTY IN THE PORTUGUESE ECONOMY

In this section we start by characterizing inequality and poverty in the Portuguese economy during the 1990's and then we proceed to its characterization within the context of the 25 member-states of the European Union (EU_25) in order to identify its major determinants.

2.1. The 1990's

Rodrigues (1996; 1999a,b; 2007) analyses at deep extent the main characteristics of inequality and poverty in the Portuguese economy during the 1990's.

1989-2000 was a period of vigorous growth for the Portuguese economy, with growth accelerating in the last five years of the period under analysis[1]. In the this period, the increase in real income occurred at all levels of income translating into a decrease in absolute poverty[2]. But

[1] The real GDP per capita average annual growth rate was 2.8% for the whole period, 2.6% in 1989-1995 and 3.7% in 1995-2000.
[2] The poverty line is 6500 Euros per year.

the increase in income was uneven, with inequality rising strongly in the first half of the period. Based on the distribution of income by deciles per equivalent adult in the years 1989, 1995 and 2000 the growth rates of equivalent income for the whole distribution, as well as per decile, were 2.64% for the whole period, corresponding to a growth rate of 2,18% in the 1st decile, against 3.69% in the 10th decile. Within the period income inequality augmented at different speeds: in 1989-1995 the growth rate was 4.3% for the 10th decile against 1.41% for the 1st decile. This trend was reversed in the second period, 1995-2000: the growth rate for the 10th decile decreased to 2.6% and in the 1st decile it increased to 3.12%. But this change in trend was not sufficiently strong to determine a trend for the whole period similar to the one that characterized the 1980's, when changes in income and inequality benefitted the most those individuals at the bottom or at low levels of the income distribution.

2.2. Relative to the EU_25

Since the setting-up of the Lisbon Agenda, a renewed interest on Social Policies by the European authorities has taken place resulting in four important reports on the social situation in the EU_25 (see EC (2007; 2008a,b,c). We will rely on these reports to characterize inequality and poverty in the Portuguese economy relative to the EU_25 since they use comparable data and indicators[3]. The data used in the reports refers to the years 2001 (ECHP) and 2004 (EU_SILC) and the analysis is conducted at country level as well as for the average EU_25 economy.

The Portuguese economy was the most unequal economy in 2001 and maintained thÿÿÿÿlaceÿÿn ÿÿÿÿ: ÿÿ pÿÿsented a Gini coefficient of 41% and a rate of population at risk of poverty of 20.3%, against 20.6% for Lithuania and Poland and 9.2% for Sweden. For the group of Mediterranean Countries, this rate varies between 19% and 20.6%, with Portugal exhibiting the highest value, which corresponds to 2 135 000 of poor people. By age composition this aggregate rate corresponds to 19% of children, 57% of people at working age and 24% of elderly people, which translates into 405650 poor children, 1216950 poor people at working age and 512400 poor elderly people.

How does the composition of the population at risk of poverty vary according to household type[4] and work intensity? Households with couples with less than 65 years of age without children and work intensity less than unity correspond to 5% of the poor population (in levels, 106750 poor people); households with lone women with more than 65 years of age represent approximately 6% of the poor population (128100 poor people); households with couples without children with 65 years of age or more registered a rate of 14% (298900 poor people); households with couples with 1 or 2 children with work intensity of unity correspond to 12% of the poor population (256200 poor people); for the households with couples with 1 or 2 children with work intensity less than unity and at least equal to 0.5, the rate was 18% (384300 poor people); and for other household types the rate was 45% (960750 poor people)[5].

The age composition of the Portuguese population classified as poor shows that a huge percentage are at working age, which means that in most cases their work is not productive enough to ensure decent living standards, due most certainly to their low levels of educational attainment and/or training. Additionally, this situation has strong negative consequences upon children that live in households whose parents are at working age since it makes them more exposed to the risk of poverty. It is thus not surprising that the percentage of poor children in the total poor population amounts to 19%. Analyzing the percentage of the poor population by household type, we can conclude that 30% of the Portuguese considered poor live in households with couples with 1 or 2 children and work intensity between 0.5 and 1. This is an impressive number because a large number of poor children are raised in these types of households and

[3] Starting in 1994, for a period of 10 years, the European Community Household Panel Survey (ECHP) produced annual data on social conditions. In 2005 it was replaced by a new instrument, Statistics on Income and Living Conditions (EU-SILC), which is based on larger samples and allows improved characterization of certain households.
[4] See EC (2007), Chapter 2.
[5] See EC (2007), table 3 p.52 and first pie chart on p. 67.

their parents' work is not sufficiently productive to enable them to overcome the poverty threshold.

According to household type and the rate of children at risk of poverty, the Portuguese economy belongs, along with Spain and Italy, to the group of EU countries for which the prevalent household types are: i) households with couples with 2 children and; ii) complex households, for which the rate of children at risk of poverty is less or equal to 22% and not higher than 29%.

According to the national and EU rates of poverty among children and the overall national rate of poverty, Portugal belongs to the group of countries for which the rate of poverty of children is above (or equal to) to the EU average and the national rate of children at risk of poverty is higher than the national rate of risk of poverty for the overall population, but doesn't exceed the latter in more than 5 p.p.. For the Portuguese economy the rates are, respectively, 24% and 19%, a difference of exactly 5 p.p., so the Portuguese economy is not very far from the countries that belong to the worst group, those that present a differential between the two rates that exceeds 5 p.p..

How is inequality related to age, education and employment status[6]? Based on a static decomposition of the Mean Log Deviation (MLD) index it is possible to identify five groups of countries within the EU_25: the Anglo-Saxon, the Baltic, the Central European Countries, the Continental, the Nordic and the Mediterranean groups. Portugal belongs to the last group, for which the percentage of inequality explained by the different factors was, in 2004, 18% due to education, 4% attributable to employment status and 3% to age, that is, in the Portuguese economy education is the major factor explaining the observed inequality[7]. Compared to the other country groups, the Mediterranean group registered the highest percentage of inequality explained by education, while the Nordic group experienced the lowest value. The employment status has the highest influence on inequality in the Anglo-Saxon group, 19%, while the lowest value is registered in the Mediterranean group, 4%. Age accounts as the major factor determining inequality in the Nordic group, exhibiting the highest value, 8%, against only 2% for the Central European Countries group.

As for the dynamic decomposition of the MLD index, since the period under analysis is very short the differences between the country groups due to the three factors mentioned above are expected to be negligible. For the Portuguese economy, in the period 2001-2004 the increase in income dispersion within groups explains the increase in inequality, while age, education and employment status are relatively unimportant in the explanation of the increase in inequality in Portugal in this period.

Let us now focus on the characterization of households according to work intensity, since work is the most important (proximate) factor in preventing individuals at working age and their dependents from the risk of poverty, as it is the most important source of income to the household budget. Labour services should be productive enough in order to allow for earnings levels sufficiently high that, when translated into units of equalized income, should surpass, at least to some extent, the income poverty threshold.

For the EU_25, in households with work intensity higher or equal to 0.5, 13% of children are at risk of poverty, while in the Portuguese economy that rate is 21%. In households with parents with full time work the rate is 7% for the EU_25 and 14% for Portugal. As for households with couples with 1 or more children, the rate of children at risk of poverty differs according to the respective work intensities. Comparing the EU_25 and the Portuguese rates these are, respectively: for the jobless, 68% against 47%; for part-time only, 32% and 59%; for one full time work, 21% and 44%; for one full time work plus one part-time work, 6% and 26%; and for two full time work, 5% and 10%. Except for the jobless, the rates are higher for Portugal than for the EU_25. Additionally, the rates for Portugal in the cases of one full time work and one full time work plus part-time work are the highest among the EU_25 and for the two full time works it is the second highest and the double of the EU_25 rate. It is thus possible to conclude that educational levels of the Portuguese working population are a very important

[6] See EC (2007), pp. 19 -27 and pp.34-42.

[7] In fact, the share of between group inequality in total inequality by education of the household head is 24%, while by age is 1% and by employment status is 2%.

(ultimate) factor in explaining child poverty, since low levels of education are associated with lower labour productivity and lower wages[8].

As predicted by human capital theories (see e.g. Simões (2006)), employment rates are positively associated with educational attainment of individuals aged 15-64[9] (see European (2008c), chart 7.4, p. 159). For the EU_27, the employment rate of individuals with low educational attainment is less than 50% and for the majority of the EU member states for the individuals with high educational attainment the employment rate is more than 85%, although important differences within countries can be observed. For example, Portugal has the highest employment rate for individuals with low educational attainment, 66%, similar to the employment rate for individuals with medium educational attainment, and the employment rate for individuals with high educational attainment is approximately 85%. Unemployment rates by levels of educational attainment for the EU_27 over the period 2000-2007, seem to confirm the results above since for all member states, with the exception of Greece, the unemployment rate is higher the lower the educational attainment of individuals.

The educational levels of the Portuguese population aged 15-64[10], are lagging from the EU_27 values. The percentage of the population with low, medium and high educational attainment in the EU-27 for the years 2000 and 2007 were, respectively, 38%, 45% and 17%, and 32,7%, 46,7% and 20,6%, against 79%, 13,4% and 7,6%, and 71,3%, 16,7% and 12% in the Portuguese economy for the same years[11].

Analyzing the educational outcomes of children it is possible to understand if the intergenerational mechanism of transmission of disadvantages/inequality acts through the influence of the parents' educational background upon the children educational attainment. According to the educational level of the father, the probability of attaining high school for women and men aged 25-65 in the EU_25 and Portugal is, 0.18 and 0.11 when the father has a low level of education; 0.33 and 0.58 when the father has a medium level of education, and 0.63 and 0.65 when the father has high level of education, respectively. For the Portuguese economy, the influence of the father's educational background is weaker for women aged 25-65 than for men[12]. In younger cohorts that dependency is also lower when the father's educational attainment is low. The probability of attaining higher education for men and women by age (25-34, 35-44 and 45-54) when the father's level of education is low are: for the EU_25, 0.25, 0.19 and 0.15, respectively, and for Portugal 0.17, 0.09 and 0.09. Although the EU_25 has experienced improvements in educational attainment levels, the fact is that, in 2006. 15% of the population between 18 and 24 years old dropped out the education or training systems before completing the higher level of secondary schooling. In Portugal the situation is even worst with drop-out rates of more than 30%[13]. The European benchmark for the year 2010 of no more than 10% of the population aged 18-24 with less than upper secondary schooling not attending the education or the training systems seems thus difficult to be reached within a year.

The numbers presented above seem to confirm that the intergenerational mechanism of transmission of disadvantages/inequality acts primarily through the influence of the parents' educational background upon the educational attainment of children in the whole EU_25. Furthermore, they suggest that this is a very powerful mechanism in Portugal, with related detrimental effects upon the current and future well being of children, as well as upon economic growth (Carneiro (2007)).

Budria (2007) uses data from the ECHP to study economic inequality in Portugal for the period 1994-2001. He is the first to study the wealth distribution for the Portuguese households, namely its capital income distribution, characterizing it for the period under analysis as well as describing its evolution over time[14]. Budria (2007) shows that the distribution of capital is very

[8] See EC (2008a), pp.34 and 35, table A13.c on p. 165 and tables A14 bis on p. 169.
[9] See also EC (2008b), chart 7.5, p. 15
[10] Low level, ll; medium level, ml; and high level, hl.
[11] See EC (2008b), table Ann B.7.1, p.218.
[12] For women the probabilities are 0.14(ll), 0.64(ml) and 0.67(hl), against 0.08(ll), 0.52(ml) and 0.62(hl) for men.
[13] See EC (2008a), p. 99.
[14] On this topic, it is also worth mentioning a previous study by Cardoso and Cunha (2005) since it constitutes an attempt to estimate capital assets of the Portuguese households. However the authors do not analyze the wealth distribution of the Portuguese households in the period under analysis, 1980-2004.

uneven and that this situation increased over the whole period, contrary to income inequality and earnings inequality that registered a decrease over the same period. These findings suggest that we must also consider seriously the fiscal approach as a possible candidate to the explanation of the relationship between inequality and economic growth in the Portuguese economy. In fact, the capital distribution suggests the existence of a political mechanism dictated by the median voter that prefers more redistribution than the average owner of capital, which in turn entails a tax policy that benefits the median voter and decreases the after-tax marginal product of capital of the Cardoso and Cunha (2005) average owner of capital, hindering investment and thus economic growth. However, we have to carefully consider possible obstacles to the application of this approach to the Portuguese economy on the tax policy side, since in Portugal capital income at the top 1% or 5% of the distribution is not heavily taxed.

The previous picture on inequality and poverty for the Portuguese households suggests that the uneven distribution of human capital is a decisive factor in explaining inequality in the Portuguese economy. This uneven distribution of human capital is in turn the result of past investment decisions on human capital. Thus the borrowing constraints on human capital investment is a serious candidate approach to the explanation of the relationship between inequality and economic growth for the Portuguese economy.

The findings that the distribution of capital income in Portugal is very uneven and has increased over time also suggest that we must consider the fiscal approach as another serious candidate to the explanation of the relationship between inequality and economic growth in the Portuguese economy.

3. CHANNELS OF TRANSMISSION OF INEQUALITY TO GROWTH IN THE PORTUGUESE ECONOMY: A SELECTIVE SURVEY OF THE THEORETICAL LITERATURE

There is by now an extensive and rich literature on the relationship between inequality and economic growth that has been carefully and thoroughly surveyed by Aghion, Caroli and García-Penalosa (1999), Perotti (1994), Alesina and Perotti (1994), Benabou (1996), and Deininger and Squire (1996), among others. In the nineties of the 20th century a renewed interest on the subject took place that lead mostly to the development of theoretical models on the subject. Due to the scarcity of reliable and comparable data on inequality and due to the existence of poor measures of equality (inequality), empirical literature on the subject lagged behind. The theoretical literature relies on common sources despite the different approaches within it. Endogenous growth theory, theory of political economy, microeconomics of information are key sources for this literature, allowing it to explore, in the framework of optimal growth models or overlapping-generations models, for instance, stationary solutions in models with distribution effects, or richer dynamics beyond steady state equilibrium solutions. This literature has also challenged key predictions from a former literature on the subject, namely those derived from the Kuznets inverted-U curve relating economic growth to income inequality.

Our main purpose in this section is to briefly survey the theoretical literature that highlights the mechanisms relating inequality to economic growth that we have considered relevant for the Portuguese economy. We will consider in this paper only two approaches from which those mechanisms can be derived. The first approach we deal with, sometimes called borrowing constraints in human capital investment, explains the relationship through investments in human capital under imperfect capital markets. The second approach that we will survey is the fiscal policy approach emerging from the political links that relate income distribution to economic growth. Both approaches predict a positive relationship between equality in the distribution of human capital and economic growth.

3.1. Borrowing constraints to investments in Human Capital

In a seminal paper, Galor and Zeira (1993) analysed the theoretical link between income distribution and economic growth through human capital investments in the framework of an overlapping generations with individuals living for two periods and with intergenerational altruism. In the model, the economy produces a single good using skilled or unskilled labor. Individuals have identical potential skills and preferences; the only difference concerns their inherited wealth. In the first period they work as unskilled or invest in human capital. In the second period they work as unskilled or skilled workers depending on their level of education, they consume and leave bequests to the younger generation. The initial wealth distribution determines the short run equilibrium due to the existence of capital market imperfections to the borrowers. Only those individuals that have an initial level of wealth high enough are able to invest in human capital, those who have a poor initial wealth endowment are not able to invest in human capital because borrowing is costly and difficult.

But the initial level of wealth distribution will determine also the long-run equilibrium due to the indivisibility of individual investments in human capital. This last equilibrium is characterized by the existence of two groups of individuals, unskilled and skilled, whose relative size is determined by the initial wealth distribution. The level of income and wealth is positively related to the number of individuals that invest more than the critical point below which a skilled dynasty cannot be sustained[15]. The long-run equilibrium is path dependent: multiple equilibriums exist and the economy will approach one of the possible equilibriums depending on the initial wealth distribution. The main prediction of the model is: "*If we would like to describe these results in more popular terms, we could say that a country has better growth prospects, if it has a relatively larger middle class*"[16]. So an economy with a more equal wealth distribution will be growth enhancing because it invests more in human capital.

Galor and Moav (2004) aim at reconciling the classical view on inequality and growth with the modern view within the framework of a growth model with different accumulation *regimes*. At earlier stages of capitalism, the engine of growth is the accumulation of physical capital and a more unequal wealth distribution is more conductive to economic growth because the owners of capital have a higher propensity to save. At later stages of capitalism the engine of growth is human capital accumulation. There are borrowing constraints on individual's human capital investment due to the existence of capital market imperfections to the borrowers. So at later stages of capitalism, a more equal wealth distribution enhances economic growth because individuals decide to invest more in human capital. Finally, the influence of inequality on economic growth will depend on the relative return to physical and human capital. These predictions are obtained in the framework of an overlapping-generations model with two periods. The economy produces a single good and accumulates physical and human capital. As in the Galor and Zeira (1993) model, individual decisions about their investment in human capital depend on their inherited wealth because they face borrowing constraints.

Chiu (1998) builds an overlapping-generations model with income and talent heterogeneity and intergenerational altruism. Individuals face liquidity constraints but, contrary to Galor and Zeira (1993), they cannot save nor borrow against future income. Individuals live for two periods and in the first period they receive from their parents a bequest and decide whether to work or attend school. In the second period they work and receive an income that is proportional to their investment in human capital. Individuals with talent above the average will receive only in accordance to that rule if they have invested in human capital. The main predictions of the model are: the higher the initial income equality for one generation, the higher will be the aggregate human capital that is accumulated by that generation and the higher will be the initial income distribution of future generations. The steady state growth rate is independent from initial conditions, namely, from initial income distribution; it depends on the aggregate level of human capital. The convergence result for the growth rate implied by the model

[15] Galor and Zeira (1993), p. 42, equation (26).
[16] Galor and Zeira (1993), p. 42.

however has different implications in terms of output levels which depend always on the initial income distribution.

The important contributions derived from the borrowing constraints to investments in human capital approach shortly reviewed above all imply the same main prediction: a higher initial income (wealth) distribution implies higher economic growth through higher investment in human capital. This prediction also applies to a growth model with different regimes of economic development if human capital accumulation is the main engine of economic growth.

3.2. The fiscal policy approach

The fiscal policy approach is anchored on two recent strands of economic literature, models of political economy and endogenous growth models. This literature was extensively reviewed by Alesina and Perotti (1996), Perotti (1996) and Aghion, Caroli and García-Penalosa (1999), among others.

The fiscal policy approach results from the political links between income distribution and economic growth. The median voter's preference for income distribution determines income redistribution through voting, which in turn translates into government expenditures and taxation. The higher is the preference of the median voter for income redistribution relative to the average voter, the higher the level of taxation he will vote for. This in turn will discourage private investment, hindering in this way economic growth. The fiscal approach thus predicts a positive relationship between equality and economic growth, no matter the specific redistribution mechanisms that might be used to achieve that redistribution.

The fiscal approach literature is based on the interplay of two mechanisms, the political mechanism and the economic mechanism. The former determines income redistribution and government's expenditures and taxes, while the latter determines the influence of taxation on economic growth. Seminal references on the fiscal approach are Alesina and Rodrik (1991; 1994), Bertola (1993) and Persson and Tabellini (1991; 1994). As far as the political mechanism is concerned these three contributions consider different types of government expenditures in order to achieve the desired income redistribution: public investment in Alesina and Rodrik (1994), redistribution from capital to labour in Bertola (1993) and pure redistribution transfers in Alesina and Rodrik (1991; 1994) and Persson and Tabellini (1991; 1994). In what concerns the political mechanism they all share the conclusion that no matter the type of government's expenditures these imply a higher level of capital taxation that decreases the after-tax marginal revenue to capital thus harming private investment and consequently economic growth.

Alesina and Rodrik (1994) develop two endogenous growth models with wealth distribution effects. The first model considers the distribution conflict between two classes, workers and capitalists, that derive their revenues from different sources, labor and capital, respectively. It is an one sector AK optimal growth model with a production function *à la* Barro (Barro (1990)), with the government's budget constraint indicating that redistribution of income to workers has its source exclusively on capital income taxes. The model is completed by adding intertemporal utility functions for each class and for the government, with the associated budget constraints. Several important predictions are derived from the model. From the solution of the dynamic problem faced by the capitalists, the respective consumption growth rate is maximized if the share of government revenues transferred to workers is zero and the capital tax is associated with the growth maximizing level.

The problem faced by the government is thus to choose taxes by solving the dynamic problem of maximizing its utility function, which is a combination of the utilities functions of both classes and whose weights are exogenously determined and subject to the usual constraints. The main prediction is that if the government attributes importance to workers' consumption through income redistribution, the capital tax level chosen will surpass the level associated to the growth maximizing level and in consequence the growth rate will be inferior to the maximum growth rate. Considering that the time preference parameter is the same for both classes, the solutions are time invariant. Two regions for the solutions can be considered depending on the value of the parameter that measures the weight of the workers' welfare

function in the social welfare function: a region of low values for the parameter and a region of high values. It is proven that growth is inversely related to that parameter. In the former region, there are possible solutions without income transfers to workers if their time preference is low.

Considering that the workers' time preference is higher than the capitalist's time preference, the solutions to the government's dynamic problem are no longer time invariant, which leads to a problem of time inconsistency. The solution to the problem mentioned above under an optimal policy with commitment leads to a growth rate that increases over time. Correcting for the problem of time inconsistency implies that governments' parameters are time invariant and so will be the growth rate. The authors simulate the time consistent solutions and the optimal solutions (at time zero and at time=∞) for this version of the model and arrive to several important conclusions. For all values of the parameter that measures the weight of workers' welfare in the social welfare function, the solutions for both policies are associated with no redistribution of income to workers. This means that public expenditures are productive enough and workers will be in a better situation without income redistribution. Comparing the time consistent solutions with the optimal solutions (at time zero), for the same parameters, the latter ones lead to a higher capital income tax and a lower growth rate. Assuming that optimal policies with commitment are possible, under these policies at time zero government rewards more workers' preference for present consumption through public expenditures at the expense of higher taxes in capital income. The fundamental predictions of this two class model are restated in a new model where individuals are not split into workers and capitalists but instead differ by their initial relative capital endowment. Additionally, individuals will vote by majority rule only for the level of capital tax, so pure income transfers from the government to individuals are not considered. The main prediction of the model is: "in a democracy the more unequal is the distribution of wealth, (…), the lower is the rate of growth of the economy"[17].

Bertola (1993) develops an optimal one sector AK growth model with distributional effects derived from individuals' heterogeneity due to differences in initial endowment shares of accumulated to non-accumulated factors (capital to labor, respectively). Under the political mechanism of redistribution of income from capital to labor, a median voter will vote for a tax on capital that is inversely related to its capital-labor ratio. The economic mechanism will act through a lower after tax capital income appropriated by the private investor, which discourages private investment, jeopardizing economic growth.

Persson and Tabellini (1994) build an overlapping-generations model with non altruistic individuals living for two periods and constant population. Individual preferences are the same and the income of the young generation depends on an average endowment of basic skills and an endowment of individual-specific basic skills, as well as on the capital stock accumulated by the older generation. A pure redistributive income policy takes the form of taxes on profits: profit income from those individuals that invested above the average are transferred to those that have invested below the average through lump-sum transfers, while intergenerational redistribution is excluded. The redistributive income policy is politically determined by the median voter. If the individual skills of the median voter are lower (higher) he will prefer a tax (subsidy) on investment if he is poorer (richer) than the average investor. Two main predictions can de derived from the model in the context of a stationary solution: a more equal distribution of income increases growth and a higher level of basic skills is growth enhancing.

Persson and Tabellini (1991), discuss at full length more complex dynamics for the growth rate and inequality that can be path dependent and allow to reconcile, under certain conditions, their findings with the Kuznets curve. The same is true for Perotti (1993) that in the framework of an overlapping-generations model and of an infinite-horizon model analyses the relationship between income distribution and economic growth in a political model where individuals vote for income redistribution and invest in human capital. Multiple equilibriums are allowed, the initial income distribution along with the political equilibrium determine the growth rate. One of the main predictions of the model is the existence of an inverted U-shaped relationship between inequality and income levels in cross-section analyses, which does not necessarily stand for time series analyses. In a different framework of analysis, Aghion and Bolton (1997) use also an

[17] See Alesina and Rodrik (1994), p.23.

overlapping-generations model with capital market imperfections due to moral hazard, and they analyze the relationship between wealth distribution and economic growth in an economy where individuals invest in physical capital. Like Perotti (1993), the authors consider also a trickle down growth mechanism. The existence of an inverted U-shaped relationship, similar to the Kuznets curve, is one of their main predictions.

The main results that follow from the contributions shortly reviewed above and that stem from the fiscal approach to the relationship between income distribution and economic growth allow us to conclude that, in spite of the different assumptions adopted in the underlying models, a common prediction is obtained: initial higher equality in income (wealth) distribution is growth enhancing no matter the political and the economic mechanisms at stake. Models that allow for multiple equilibrium solutions are able to predict an inverted U-shaped relationship between inequality and growth similar to the Kuznets curve and are also able to predict that permanent redistribution policies can foster economic growth.

4. EMPIRICAL EVIDENCE

The aim of this section is to offer a summary and a critical discussion of the empirical evidence on the impact of inequality on economic growth with a focus on the evidence concerning the mechanisms identified as potentially more important for the Portuguese economy.

4.1. Main empirical findings

In this review we focus on a number of empirical studies that try to shed some light on the impact of inequality on economic growth. We start our empirical literature review with studies from the 1990's[18]. After analysing the main conclusions associated with the estimation of reduced-form equations we then proceed to the analysis of the main results concerning the clarification of the mechanism through which inequality influences growth.

4.1.1. Reduced-form equation

The studies reviewed in this section share a common regression specification of the general form:

$$GR_{it} = f(INEQ_{it}) + \beta_i X_{it} + \mu_{it}$$

where a measure of output growth, GR_{it}, is regressed on a measure of (initial) inequality/equality, $INEQ_{it}$, where the function f might be non-linear; and a set of other explanatory variables, X, that include initial output, some measure of human capital, physical capital investment ratio, and regional dummies, among others. This equation can thus also be classified as a "Barro" or *ad hoc* growth regression that includes inequality as a regressor along another set of independent variables that are believed to influence economic growth[19].

The aim of these regressions is to identify the sign of the relationship between inequality and growth in order to determine whether redistributive policies enhance growth. Until the late 1990's the identification of the sign of this relationship was based on the exploitation of cross-country variation of the series under analysis, but recent studies have been using pooled cross-country time series data[20].

A first set of studies (see Persson and Tabellini (1994), Alesina and Rodrik (1994), Alesina and Perotti (1996), Clarke (1995), Perotti (1996), Birdsall and Londono (1997), Deininger and Squire (1998), Castelló-Climent and Doménech (2002), Chen (2003), Balisacan and Fuwa (2003), Bleaney and Nishiyama (2004)) tries to assess the contribution of inequality to cross-

[18] Benabou (1996) presents a summary of the results of 23 previous cross-country studies, some prior to 1990, that reveal that inequality is unambiguously detrimental to growth.
[19] See for instance Barro and Sala-i-Martin (2004).
[20] This was made possible to a great extent by the work of Deininger and Squire (1996) that assembled a more reliable inequality dataset with time series information for a large enough group of countries. See section 4.2.

country variation in growth rates, after controlling for a number of variables that have been found relevant in the explanation of cross-country growth performance in a large number of empirical growth studies aimed at identifying the importance of some other factor for growth.

The estimates of the impact of inequality on economic performance of the various cross-country studies reviewed here reveal a negative long-run relationship between the two variables[21], a result that is usually robust to different sensitivity analyses such as controlling for different inequality measures, different samples and time periods, the presence of outliers, model uncertainty, measurement error, reverse causation, and heteroscedasticity. In what follows we identify the main differences in the methodology used and results achieved by these cross-country studies and aspects that hinder their comparability and robustness of the results.

All the studies focus on clarifying the importance of some measure of inequality in explaining cross-country differences in real per capita GDP growth rates averaged for 20-25 years periods, although this measure comes from different sources (e.g. Summers and Heston (1991) and the World Bank) and none tests the robustness of the results to the use of other proxies of macroeconomic performance, such as real GDP per worker or per hour worked, probably due to limited data availability at the time or for comparability reasons with previous studies.

A more fundamental issue, dealt with in several of the reviewed studies, concerns the sensitivity of the results to the use of different proxies of inequality. When the studies measure inequality solely as income inequality they tend to use different statistical measures of this indicator with a prevalence for the Gini coefficient of income, but also some measure of equality in the distribution of income aimed at capturing the size of the middle class, such as the share of income of the middle quintile(s) of the population or inequality measures based on the ratios between different income percentiles[22].

But a preliminary problem arises when deciding which measure of inequality to use in empirical studies since the relevant distribution in explaining the relationship between inequality and growth in many theoretical analyses is that of wealth, not income. Due to limited data availability on the distribution of wealth and to the fact that both distributions are often highly correlated, only a limited number of cross-country studies deals with this issue. The exceptions are Alesina and Rodrik (1994), Alesina and Perotti (1996), and Deininger and Squire (1998) that proxy wealth inequality with the Gini coefficient of land distribution and get more robust results with this inequality measure than with income inequality measures. For instance, in Alesina and Rodrik (1994), when the Gini coefficient of income and the Gini coefficient of land distribution are introduced simultaneously in the regressions the former looses statistical significance, while in Deininger and Squire (1998) the estimated coefficient on initial Gini coefficient of land distribution does not loose significance when regional dummies are introduced, contrary to what happens to the estimated coefficient on initial Gini coefficient of income.

Another interesting strand of the empirical literature defends that in important theoretical models that analyze the relationship between inequality and growth the distribution of education/human capital shapes the distribution of income and wealth and thus the relevant inequality measure refers to human capital inequality. At the cross-country level, Birdsall and Londono (1997) and Castelló-Climent and Doménech (2002) reach more robust results when using a human capital inequality measure that an income inequality measure.

The samples used in most studies integrate developing and developed countries and some split their samples into sub-samples according to the countries' political regime (democratic or non-democratic) or level of development (poor and rich countries) to test the robustness of the results. For instance, Alesina and Perotti (1996), Clarke (1995), and Perotti (1996) do not find any significant difference between democratic and non-democratic countries but when the sample is split according to the countries' level of development the same conclusion does not apply. Perotti (1996) estimates reveal a high (negative) and statistically significant coefficient in

[21] See also Benabou (1996).

[22] For instance, Clarke (1995) analyses the sensitivity of the results to the use of four different income inequality measures: the coefficient of variation, the Theil's index, the Gini coefficient, and the share of income earned by the 40 percent poorest relative to the richest 20 percent of the population.

the sample of rich countries and a low (negative) and statistically insignificant coefficient in the sample of poor countries. Chen (2003) identifies an inverted-U relationship between initial income inequality and long-run growth in a sample of 43 countries between 1970 and 1992, while Bleaney and Nishiyama (2004) detect no differences between rich and poor countries in two out of three estimated models, but significant differences in the sign and statistical significance of the estimated coefficient on inequality between models.

Despite the differences we have just highlighted between the reviewed cross-country studies, the broad picture is that: a) there is a significant negative correlation between inequality and growth; b) this relationship remains significant and of the same sign when using slightly different inequality measures, samples, data, sets of regressors and estimators (OLS, WLS, 2SLS); c) in most studies the relative impact of inequality is lower than that of other growth determinants.

Following the release of the Deininger and Squire inequality dataset that assembled more reliable data[23] with time series information for a large enough group of countries, a number of studies estimated the inequality and growth relationship using panel data techniques (see Persson and Tabellini (1994), Partridge (1997), Li and Zou (1998), López, Thomas and Wang (1998), Forbes (2000), Barro (2000), Panizza (2002), Banerjee and Duflo (2003), Voitchovsky (2005)), trying to uncover in this way a short to medium-term relationship. Contrary to the common message conveyed by most cross-country studies, the panel data evidence is quite diverse, finding either a positive, negative or non-existent correlation between inequality and growth.

Starting with the studies that consider wider samples of countries with both developing and developed countries, Li and Zou (1998) find that inequality is not necessarily harmful for growth, i.e. the sign of the relationship is ambiguous and positive in some cases; Forbes (2000) detects a positive relationship that persists across different samples, variables definitions, and model specifications but not the length of period under consideration; Barro (2000) uncovers a negative relationship for poor countries, a positive relationship for rich countries, and an insignificant one for the whole sample; and Banerjee and Duflo (2003) present evidence that it is a change in any direction, not the initial level of inequality that leads to slower future growth and also that there is no reason to impose a linear relationship between inequality and growth. With the exception of Barro (2000) that considers 10-year growth episodes, these studies assess the impact of inequality on growth over 5-year periods, all take the Gini coefficient of income as the reference inequality measure and estimate the relationship using different panel data techniques (pooled OLS, fixed and random effects, GMM, 3SLS, Kernel regression, series estimator). All of them also include some form of sensitivity analysis of the results such as the consideration of different inequality measures, control variables, or samples of countries. Forbes (2000) suggests that her results are not incompatible with the results from the cross-country studies since the former explore within-country information variation in a short period of time identifying the sign of the relationship between changes in inequality and changes in growth performance, while the latter explore between-countries information variation in a longer period of time between the initial level of inequality and the growth rate.

Considering a more restricted sample that includes only U.S. states, Partridge (1997) and Panizza (2002) also explore the panel structure of the data to uncover the sign of the relationship between inequality and growth. The first study considers ten-years growth episodes during the period 1960-90, finding a positive correlation between the Gini coefficient of income and growth, indicating a positive relationship between inequality and growth, but also a positive correlation between the income share of the middle quintile and growth, indicating a positive relationship between equality, not inequality, and growth. The second study uses state data from different sources covering a wider period (1940-80), Gini coefficients of income and the income share of the middle quintile, analyzes ten and twenty-years growth episodes, and uses different panel data techniques (pooled OLS, fixed effects, GMM), reaching the conclusion that the sign and significance of the estimated relationship

[23] See section 4.2.1.

changes when different measures of inequality are used and that controlling for outliers, serial correlation and structural breaks substantially changes the results.

In face of the ambiguous results from the panel data studies, Voitchovsky (2005) explores the idea that aggregate measures of income distribution such as the Gini coefficient of income are not appropriate to test the importance of inequality for economic growth since the relationship might depend on the whole shape of the income distribution, i.e. may hide movements in the incomes of different groups. For instance, overall inequality may remain relatively stable over time but there can be considerable change in the shares of total income received by individual groups. Using comparable cross-country time series data on inequality for twenty-one industrialized countries surveyed by the Luxembourg Income Study (LIS) between 1975 and the year 2000, she constructs measures of inequality in the bottom and top ends of the income distribution to assess whether top and bottom end inequality have different implications for growth. The results reveal that top end inequality positively influences growth while the influence of bottom end inequality is negative. These results are robust to the use of different estimators (pooled OLS, within-groups, first-differenced GMM and system GMM).

4.1.2. Transmission mechanisms specifications

The inconsistency of the results provided by the estimation of reduced-form equations may just be the consequence of the complexity of the effects of inequality on growth predicted by the theoretical literature that does not necessarily imply that more inequality decreases growth, as initially put forwarded by new growth theories. Since the reduced form analysis is not very informative regarding the different channels through which inequality might affect growth it is not surprising that a number of empirical studies has tried to examine in more detail the importance of some mechanism in the explanation of the inequality-growth relationship.

This approach to the empirical study of the relationship between inequality and growth is designated by the transmission mechanisms specifications and usually consists in estimating a system of equations with two dependent variables, output growth, *GR*, and a proxy for the mechanisms under analysis, *MECH*, where the former depends on the latter, and in turn *MECH* depends on a measure of inequality/equality, $INEQ_{it}$, such as the system described below:

$$GR_{it} = aMECH_{it} + \alpha_i W_{it} + \mu_{it}$$
$$MECH_{it} = bINEQ_{it} + \beta_i Z_{it} + \varepsilon_{it}$$

where *W* and *Z* are vectors of additional explanatory variables. There is evidence supporting the influence of inequality on economic growth through the mechanism under analysis when both *a* and *b* are statistically different from zero. In face of the different theoretical predictions on the channel through which inequality influences growth (see section 3) ideally the different mechanisms and their potential interactions should be tested simultaneously but this is not possible due to the limited inequality and other growth variables data availability.

Perhaps the most often analysed mechanism of transmission is the fiscal channel. Testing the fiscal channel (see Perotti (1996), Partridge (1997), Sylwester (2000)) involves choosing the appropriate fiscal variable, that can range from public expenditures that have an explicit redistributive nature (e.g. public expenditure on education and health, social security) to taxation needed to finance the redistributive expenses[24]. Perotti (1996) considers cross-section information from developed and developing countries during the period 1960-85 and uses the average marginal tax rate (MTAX) as the reference fiscal policy variable, obtaining a positive rather than a negative, as expected, estimated coefficient on MTAX in the growth equation, while income distribution plays no role in determining MTAX, except in the sample of democratic countries. These results survive a number of sensitivity checks, namely the

[24] An alternative way that has been followed to test the fiscal channel, justified in most cases by data availability constraints, consists in estimating a reduced-form equation for samples of democratic and non-democratic countries separately or introducing an interaction term between the inequality proxy and a dummy variable for democratic regimes. Support for this transmission mechanism implies a negative statistically significant coefficient for the inequality proxy for the sample of democratic countries but not necessarily for the sample of non-democratic countries or the interaction term for the whole sample (see e.g. Alesina and Rodrik (1994), Alesina and Perotti (1994), Clarke (1995), Perotti (1996), Deininger and Squire (1998)). The evidence from these studies is mixed but in most studies the results point to no significant differences between samples of democratic and non-democratic countries.

consideration of different fiscal policy variables such as the average share of labour taxation in GDP, the average share of personal income taxes in personal income, average share of government expenditure in social security and welfare in GDP, average share of government expenditure in health and housing in GDP and average share of government expenditure on education in GDP. Partridge (1997) focus in panel data for a sample of U.S. states and tests this channel using a slightly different approach. First, he regresses growth on inequality and equality measures and government intervention variables (percent state and local taxes, percent public welfare expenditures, and the government employment share), obtaining positive estimated coefficients for both income distribution variables and negative but usually insignificant ones for the government intervention variables. Afterwards, the author estimates the impact of the income distribution variables on the different government intervention variables finding weak evidence that the former influence the latter. Sylwester (2000) analyses the fiscal channel in a sample of developed and developing countries between 1970 and 1985 concentrating on the influence of public expenditures on education. From the results, the author concludes that the influence of contemporaneous expenditures on growth is negative, probably due to the necessity to increase distortionary taxation to finance these expenditures, but that past education expenditures have a positive impact because of the associated rise in human capital. Additionally, there is evidence that this is a relatively important channel since controlling for the influence of public education expenditures on growth renders the independent impact of income inequality insignificant. Also, the influence of this fiscal policy variable is quantitatively more important than that of other government policy variables (average government transfers as a fraction of GDP, average public investment as a fraction of GDP, the average of defence expenditures as a fraction of GDP, and the average of the government share of GDP minus education expenditures).

Perotti (1996) and Deininger and Squire (1998) test the validity of the credit market imperfections that lead to borrowing constraints channel. In Perotti (1996) the mechanism is tested considering only its action through investments in education and associated endogenous fertility decisions. Using data on female and male secondary school and fertility rates, the author concludes that there is evidence supporting that borrowing constraints limit investments in education by poor people. However, more direct tests of this approach using measures of the degree of credit market imperfections (e.g. the loan-to-value ratio and the ratio of domestic credit to GDP) interacted with the income distribution variable in the equation where education is the dependent variable lead to inconclusive results. Deininger and Squire (1998), on the other hand, test the impact through credit market imperfections that lead to borrowing constraints impending over both physical and human capital investment under the assumption that the effects over human capital investment will be greater due to the greater reluctance of lenders to accept the future higher stream of earnings as collateral. After estimating a reduced-form equation for high and low-income countries separately and confirming that inequality is only relevant in poor countries and thus for a tentative support of the credit markets imperfections channel, the authors then regress proxies of human and physical on inequality and confirm that this is a relevant variable in explaining investments in human capital but not investment in physical capital, concluding that the main channel through which inequality affects growth is schooling.

4.2. Methodological issues

In this section we pay particular attention to some of the most important empirical difficulties surrounding the estimation of the inequality and growth relationship in order to try to understand some of the possible causes for the lack of consensus on the empirical assessment of the sign of the relationship patent in the literature review from the previous section.

4.2.1. Alternative measures of distribution

The empirical assessment of the importance of inequality for economic growth implies the use of inequality measures suitable on theoretical grounds for the estimation of reduced-form

equations or transmission mechanisms specifications. The most widely used measures of inequality relate to the income distribution but most theoretical studies explain the relationship between inequality and growth based on the distribution of wealth. However, few empirical studies use wealth inequality as an explanatory variable due to data constraints. Despite the fact that the shape of wealth and income distributions usually vary together in cross-sections of countries, Alesina and Rodrik (1994), Alesina and Perotti (1994) and Deininger and Squire (1998) use the Gini coefficient of land distribution as a proxy for wealth inequality in their empirical studies and present better results with this measure than with the income distribution measure[25]. Human capital inequality measures have also been shown to provide more robust results than income inequality measures in the estimation of standard growth and investment equations (see Castelló-Climent and Doménech (2002)), a result attributed to the importance of human capital in the explanation of the distribution of wealth present in many theoretical models.

The analysis of the impact of inequality on economic growth also implies the availability of comparable inequality measures for a relatively large cross section of countries, ideally over time[26]. Comparability of the data on inequality is usually hindered by: (i) differences in construction methods, e.g. it is not always based on nationally representative surveys but inferred from national accounts data by making assumptions that should not be imposed but tested in the empirical analysis; (ii) differences in data coverage, i.e. it does not cover the entire population but only a subset (e.g. urban population); and (iii) differences in income definitions, e.g. it does not encompass all types of income such as nonwage income and income from household production. By collecting data from primary survey data, official statistical publications and research papers, Deininger and Squire (1996) overcame the previous problems for a large enough group of countries, constructing a new and readily available data set of higher quality income distribution data that has also the advantage of providing information over time for a sub-sample of countries, thus allowing to perform panel data studies. However, Atkinson and Brandolini (2001) show, for a sample of OECD countries, that major issues concerning data quality and data consistency[27] (across countries and across time) remain[28], which advises caution and sensitivity analysis of the results to measurement error even when using this much improved data set.

A final problem comes from the fact that the aggregate measures of distribution, such as the Gini coefficient, are not appropriate to test the importance of inequality for economic growth since the relationship might depend on the whole shape of the distribution, i.e. may hide movements in the incomes/wealth/human capital of different groups. This issue has been addressed in empirical studies by using alternative measures of inequality, e.g. the share or ratios of certain quintiles or percentiles, and comparing the results with the ones from the estimations that consider the aggregate measure, arriving at some important differences in some cases (see Barro (2000), Panizza (2002), Voitchovsky (2005)).

4.2.2. Specification and functional form

The evidence on the inequality and growth relationship comes from the estimation of reduced-form equations or more structural models that try to investigate the mechanism through

[25] Deininger and Squire (1997) call our attention to the fact that, in general, the Gini coefficient of land distribution presents higher cross-section variability than the Gini coefficient of income which can lead to different estimated coefficients (in sign and importance).

[26] Deininger and Squire (1998), p.260: "the inequality data used are deficient with respect to their quality, their comparability over time and across countries, and their geographical and temporal coverage."

[27] In the definition of income and in the unit of account. For instance, Knowles (2005) argues that the results from existing cross-country studies on inequality and growth have to be interpreted with great caution since the majority does not use consistent data in the sense that the inequality data for some countries is based on gross income, while for others is based on net income or expenditure. Also there are differences concerning the unit of measurement: in some cases it is the individual and in others the household. Using inequality data from the World Development Indicators the author shows that when the sample is restricted to countries with consistent data in the sense described above (in this case using only gross individual income distribution data which reduces the sample to 27 countries) the usual negative cross-country relationship between inequality and growth does not hold. However, he still finds a negative relationship between inequality of expenditure and growth for a sample of developing countries.

[28] For instance, Atkinson and Brandolini (2001) show that the ranking of OECD countries according to the Gini coefficient of income is quite different using the Deininger and Squire (1996) data or the alternative Gottschalk and Smeeding (1997) data.

which the influence is transmitted by adding inequality variables to an otherwise standard growth regression with a set of other control variables commonly found in empirical growth studies, such as initial income, and measures of human capital, financial development, openness, or physical capital investment ratios. For instance, Perotti (1996) considers a relatively parsimonious number of control variables, four, while Barro (2000) introduces around ten control variables. This raises the question of which growth determinants should be included in the model since different empirical models have been shown to lead to different conclusions concerning the significance of the different independent variables[29].

Another empirical problem with studies on inequality and growth is that, with the exception of a few examples in the literature, most studies do not consider whether the way in which the proxy for inequality enters the empirical model, linear or non-linear, has any substantial impact on the estimated inequality coefficient. Using non-parametric methods, Banerjee and Duflo (2003) show that the growth rate is an inverted-U function of net changes in inequality, while Chen (2003), based on the more heterodox cross-section estimation of reduced-form equations, also finds evidence that there is an inverted-U relationship between initial income inequality and long-run economic growth.

4.2.3. Structure of the data and estimation procedures

The methodology followed by the studies under review in this paper to examine the importance of inequality for macroeconomic performance is that of multivariate regression analysis to derive meaningful estimates of the effect on growth of several factors, in particular inequality.

Empirical studies on inequality and growth are often constrained by the available data. Early studies until the late 1990's applied cross-section estimation procedures to a sample of countries, averaging growth over 20-25 years and estimating the impact of initial levels of inequality, due mainly to the lack of availability of inequality data across time. The results from these studies should thus be interpreted as the long- run impact of inequality on economic growth across countries but can not explain how a change in a given country inequality level may affect growth within that country. The release of the Deininger and Squire (1996) data set represented a major improvement with respect to, among other things, the structure of the inequality data, providing a reasonable number of observations for a large enough group of countries across time, i.e. with a panel data structure. By exploring the information across countries and across time, panel data techniques not only increase the efficiency of the estimators since they allow for more degrees of freedom, but are also more able to deal with some of the econometric problems that confront the estimation of the inequality-growth relationship, namely the omitted variable bias problem that occurs when a variable that is relevant in explaining growth is correlated with any of the regressors but is not included in the regression, thus biasing the coefficient estimates and respective standard errors[30]. For instance, fixed effects panel data techniques control for this bias in respect to unobservable country characteristics that remain constant over time. The interpretation of the coefficient estimates from panel data studies is thus different from that concerning cross-section studies: they represent the impact of inequality on growth within a country over the short to medium term and not over the long-run across countries. It is thus not surprising that some panel data studies show that the sign and significance of the relationship changes as the length of the period under analysis changes (e.g. from 5-year to 10-year periods; see Forbes (2000)).

Since the work of Kuznets (1955) that postulated an inverted-U relationship between growth and inequality, it is consensual that there is a two-way relation between inequality and macroeconomic performance and that growth and inequality evolve simultaneously, which means that endogeneity bias is present in the estimation of the relationship between inequality and growth, implying that the inequality variables in the growth regressions are correlated with the error term. This correlation might be due to reverse causation (growth determines the

[29] See e.g. Levine and Renelt (1992) and Sala-i-Martin (1997). This problem is also referred to as model uncertainty.

[30] Evidence of this problem in cross-section studies is patent for instance in the fact that in some of them the consideration of regional dummies renders the coefficient estimates of the inequality variables not significant (see e.g. Deininger and Squire (1998)).

evolution of inequality) or to the omission of variables that jointly determine both growth and inequality. One way to overcome this problem is to use initial values of the inequality proxies in the regressions but this might not be sufficient especially in panel data studies that analyse growth over short periods of time and since there might be other factors that jointly determine growth and initial inequality. A second solution is to use instrumental variables estimators such as the 2SLS or the 3SLS but these imply selecting appropriate outside instruments that are usually not readily available or easy to select. Panel studies can also make use of GMM techniques that have the advantage of using lagged values of the endogenous variables as instruments.

Empirical studies on the importance of inequality for growth are based on information from a cross-section of countries at very different stages of development and often constrain the impact of inequality to be the same across countries. This raises another important empirical problem, that of parameter heterogeneity: if the latter applies then policy implications derived from empirical models that impose restrictions about the equality of parameters that do not apply are invalid. Especially, it is quite risky to make inferences on the impact of inequality in Portugal using coefficient estimates from such wide samples of countries. In fact, splitting the sample according to the level of development, Perotti (1996) concludes, by exploring cross-country information, that higher equality leads to faster future growth in rich countries but not in poor countries. A similar conclusion is reached by Barro (2000) using data with a panel structure: the effect is negative for low levels of development and becomes positive for higher levels Bleaney and Nishiyama (2004), on the other hand, find no significant differences between rich and poor countries.

4.3. Comments

The two main empirical approaches that have been followed to assess the importance of inequality for growth are the reduced-form equation that relates inequality to output growth, and the transmission mechanisms specifications that shed additional light on the channels that explain the sign of the relationship between inequality and growth by exploring different theoretical foundations. In fact, it was the estimation of the reduced-form equation that motivated a number of papers trying to explain the mechanisms through which inequality influences economic growth (see e.g. Aghion, Caroli and García-Penalosa (1999)). After reviewing the main empirical findings from each of these approaches we identified some methodological issues that hinder the comparability of the results, which are common to most empirical growth studies[31].

Taking the studies as a whole, there is no compelling evidence that higher inequality decreases growth. The evidence concerning the existence of a relationship between inequality and growth and the identification of its sign is thus weak, which can be due to a whole host of problems, namely data problems and econometric issues. For instance, the effect of inequality on growth is sensitive to the econometric technique used. While the cross-country empirical literature that estimates reduced-form equations is replete with studies that find that higher inequality slows future growth, the evidence from the studies that explore cross-country time series information is not compelling, pointing to an ambiguous effect of income distribution on growth. In any case, the size of the effect does not seem large relative to that of alternative growth determinants. But the inconsistency of the results may also be due to the complexity of the effects of inequality on growth predicted by the theoretical literature that does not necessarily imply that more inequality decreases growth as initially put forwarded by new growth theories (see e.g. Aghion, Caroli and García-Penalosa (1999) and section 3 above). In fact, many of the new growth theory models aimed at explaining the relationship between inequality and growth predict multiple equilibria so that, depending on the initial conditions, inequality can have a positive or a negative effect on growth.

[31] For a more technical discussion on methodological issues that confront empirical growth studies see Temple (1999) and Durlauf, Johnson and Temple (2004).

The picture conveyed by the studies that explore some mechanism through which the effect of initial inequality may be transmitted is also somewhat blurred. The results from the studies that explore the fiscal policy channel depend to a great extent on the fiscal variable used and the credit markets imperfection mechanism demands data that is not yet readily available to be properly tested. Additionally, none of these transmission mechanisms studies[32] attempts to explore the panel structure of the data.

5. CONCLUSIONS AND AREAS OF FURTHER RESEARCH APPLIED TO THE PORTUGUESE ECONOMY

Since the eighties of the last century a renewed interest on the relationship between economic growth and inequality has occurred, due probably to the increase of within income inequality in advanced economies, with the USA and the UK standing as examples of economies of innovation that operate in a globalised environment. Additionally, the persistent worldwide cross-country differences in real income per capita led to an inquiry into the main factors and mechanisms explaining those differences. The slowdown of worldwide economic growth after the end of the thirty glorious years, the fall of the Berlin Wall in 1989, the economic growth disaster experienced by the African continent in the last forty years of the Twentieth Century, as well as the growth experiences of emergent economies like the BRICs, implied a deeper scientific curiosity for subjects such as inequality and poverty, as well as to their links to economic growth.

Meanwhile, in the seventies and eighties of the last century important economic literature strands emerged as suitable to be used, later on, as a solid anchorage to new avenues of research on the relationship between income (wealth) distribution and economic growth. Microeconomics of information (Stiglitz (2002))[33], the economics of inequality and poverty (Silber (1999), Bourguignon and Verdier (2000), Sala-I-Martin (2006))[34]; economics of income distribution (Atkinson and Bourguignon (1999)); endogenous growth theories (Barro and Sala-i-Martin (2004), Aghion and Howitt (1998) and Azariadis (1993)); theories of Political Economy (Drazen (2000), Persson and Tabellini (2000)) are important strands of the economic literature mentioned above. The growing importance of this avenue of research on the relationship between income distribution and economic growth is well documented by Bertola, Foellmi and Zweimuller (2006) handbook on the subject.

It is thus not surprising that there is by now an extensive and rich theoretical and empirical literature that examines the role of inequality in economic growth. This paper surveys the literature produced on the theme within the boundaries of new growth theory or endogenous growth models, from the perspective of the Portuguese economy.

Based on the characterization of the Portuguese economy in terms of inequality and poverty that points to education (see Rodrigues (1996; 1999a; 2007) and European (2007; 2008a,b,c) and fiscal choices (see Budria (2007)) as main determinants of inequality in Portugal, we conduct a review of some theoretical models that aim at explaining the link between education/fiscal choices and income distribution and between this and the rate of economic growth. These are known as, respectively, the borrowing constraints on human capital and the fiscal policy approaches. Both approaches predict a positive relationship between equality in the distribution of human capital and income and economic growth. The first approach, concludes that a higher initial income (wealth) distribution implies higher economic growth through higher investments in human capital (see Galor and Zeira (1993)). This prediction also applies to a growth model with different regimes of economic development if human capital accumulation is the main engine of economic growth (see e.g. Chiu (1998), Galor and Moav (2004)). The second approach, is anchored on two recent strands of economic literature, political economy

[32] The exception is Partridge (1997).
[33] For a wider coverage of the subject see the seminal references included in the paper.
[34] The author estimates the world distribution of income and for the period under inspection,1970-2000, and using several measures of poverty and of income inequality, he concludes that poverty decreased over the period and that worldwide inequality has also decreased irrespective to the inequality measures used.

models and endogenous growth models, and results from the political links between income distribution and economic growth (see Alesina and Rodrik (1991; 1994), Bertola (1993) and Persson and Tabellini (1991; 1994). The median voter's preference for income distribution determines income redistribution through voting, which in turn translates into government expenditures and taxation. The higher the preference of the median voter for income redistribution relative to the average voter, the higher the level of taxation he will vote for. This in turn will discourage private investment, hindering in this way economic growth. The fiscal approach thus predicts a positive relationship between equality and economic growth, no matter the specific redistribution mechanisms that might be used to achieve it. Additionally, models that allow for multiple equilibrium solutions are able to predict an inverted U-shaped relationship between inequality and growth similar to the Kuznets curve and are also able to predict that permanent redistribution policies can foster economic growth (see e.g. Persson and Tabellini (1991), Perotti (1993), Aghion and Bolton (1997)).

The two main empirical approaches that have been followed to assess the importance of inequality for growth are the reduced-form equation, with the main objective of identifying the sign of the relationship, and the transmission mechanisms specifications, with the main objective of clarifying the channels of influence that explain that sign (see e.g. Perotti (1996), Aghion, Caroli and García-Penalosa (1999), Barro (2000)). The evidence shows that sign of the relationship is sensitive to a number of criteria, and only the socio-political instability channel finds clear support in the data, while the empirical relevance of the fiscal channel and the borrowing constraints on investments in human capital channel remains the subject of further research (see Perotti (1996), Castelló-Climent and Doménech (2002), Voitchovsky (2005)). Thus, taking the studies as a whole, there is no compelling evidence that higher inequality decreases growth, a result that lead us to the identification of some methodological issues that confront the empirical studies on inequality and growth. For instance, while the cross-country empirical literature that estimates reduced-form equations is replete with studies that find that higher inequality slows future growth, the evidence from the studies that explore cross-country time series information is not compelling, pointing to an ambiguous effect of income distribution on growth. The results are also sensitive to the group of countries under analysis (rich vs. poor countries) and the inequality proxy used (wealth vs. human capital vs. income distribution based proxies). In any case, the size of the effect does not seem large relative to that of alternative growth determinants.

Despite the considerable theoretical and empirical literature that has already been produced on this subject, it is the authors' opinion that there is a great potential for research on this theme applied to the Portuguese economy. Existing work on inequality in the Portuguese economy (see e.g. Pereirinha (1988), Cardoso (1991), Gouveia and Tavares (1995), Rodrigues (1996), Cardoso (1998), Rodrigues (1999a), Cardoso and Cunha (2005), Vieira, Couto and Tiago (2005), Carneiro (2007) and section 2 above) has concentrated on providing a description of this phenomenon without relating it with macroeconomic performance, namely, economic growth in Portugal.

The main contribution of the paper within the analysis of inequality and economic growth is to provide a selective review of the role of inequality in economic growth using as our main criteria the evidence on this relationship and its main determinants for a particular country, Portugal, following the suggestion of recent empirical growth studies that highlight the need for country studies in order to shed additional light on the conclusions of the studies that have been undertaken at the cross-country level (see e.g., Rodrik (2003; 2005; 2007). Country studies are fundamental for a better understanding of the relationship between inequality and growth and present possible ways to overcome data constraints and empirical methodological issues.

As far as Portugal is concerned, there is first the need to identify the sign of the relationship between inequality and growth, followed by the need to clarify the (relative) contribution of the two different mechanisms selected as crucial for the Portuguese economy, the borrowing constraints on investments in human capital and the fiscal policy channel, in explaining this link.

Additionally, the inconsistency of the results at the cross-country level may lie on the complexity of the effects of inequality on growth predicted by the theoretical literature. In fact,

many new growth theory models aimed at explaining the relationship between inequality and growth predict multiple equilibria (see e.g. Perotti (1993), Aghion and Bolton (1997)) so that, depending on the initial conditions, inequality can have a positive or a negative effect on growth. This suggests regional level analysis as an interesting avenue of research as far the Portuguese economy is concerned since Portuguese regions present different levels of output and human capital and different initial patterns of income and human capital distribution, possibly corresponding to different stages of influences as regards the impact of inequality on economic growth. Differences at the regional level may also point to variations in the extent to which countries, and especially Portugal, are benefiting from broader economic changes such as skill-biased technological change, with its consequences on the shape of the income distribution and inequality, given human capital distribution in Portugal, especially at the regional level. Furthermore, regions experiencing different (high) levels of inequality will probably be more vulnerable to adverse shocks on economic growth, which in turn can lead to more inequality in the present and hinder further future growth.

REFERENCES

AGHION, P. and P. BOLTON (1997), "A theory of Trickle-Down Growth and development." *Review of Economic Studies*, Vol. 64, pp. 151-172.

AGHION, P., E. CAROLI and C. GARCÍA-PENALOSA (1999), "Inequality and economic growth: the perspective of the new growth theories." *Journal of Economic Literature*, Vol. 37(4), pp. 1615-1660.

AGHION, P. and P. HOWITT (1998). *Endogenous Growth Theory*. Cambridge, Massuchusets, MIT Press.

ALESINA, A. and R. PEROTTI (1994), "The Political Economy of Growth: A Critical Survey of the Recent Literature." *The World Economic Review*, Vol. 8(3), pp. 351-371.

ALESINA, A. and R. PEROTTI (1996), "Income distribution, political instability, and investment." *European Economic Review*, Vol. 40(6), pp. 1203-1228.

ALESINA, A. and D. RODRIK (1991), "Distributive Politics and Economic Growth ", *NBER* WPnº 3668, pp. 42.

ALESINA, A. and D. RODRIK (1994), "Distributive Politics and Economic Growth." *Quarterly Journal of Economics,* Vol. 109(2), pp. 465-490.

ATKINSON, A. B. and F. BOURGUIGNON (1999). *Handbook of Income Distribution*. Amsterdam, North-Holland.

ATKINSON, A. B. and A. BRANDOLINI (2001), "Promises and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries as a Case Study." *Journal of Economic Literature*, Vol. 39(3), pp. 771–799.

AZARIADIS, C. (1993). *Intertemporal Macroeconomics*. London, Wiley-Blackwell.

BALISACAN, A. M. and N. FUWA (2003), "Growth, inequality and politics revisited: a developing-country case." *Economics Letters*, Vol. 79, pp. 53-58.

BANERJEE, A. V. and E. DUFLO (2003), "Inequality and Growth: What Can the Data Say?" *Journal of Economic Growth*, Vol. 8(3), pp. 267-99.

BARRO, R. (1990), "Government spending in a simple model of endogenous growth." *Journal of Political Economy*, Vol. 98(5, part II), pp. S103-S125.

BARRO, R. J. (2000), "Inequality and growth in a panel of countries." *Journal of Economic Growth*, Vol. 5(1), pp. 87-120.

BARRO, R. J. and X. SALA-I-MARTIN (2004). *Economic Growth*. New York, MacGraw-Hill.

BENABOU, R. (1996). *Inequality and economic growth.* NBER Macroeconomics Annual. Cambridge, Massachusetts**,** pp. 11-76.

BERTOLA, G. (1993), "Factor Shares and Savings in Endogenous Growth." *American Economic Review* Vol. 83(5), pp. 1184-1198.

BERTOLA, G., R. FOELLMI and J. ZWEIMULLER (2006). *Income distribution in macroeconomic models*. Princeton, Princeton University Press.

BIRDSALL, N. and J. L. LONDONO (1997), "Asset inequality matters: An assessment of the World Bank's approach to poverty reduction." *American Economic Review*, Vol. 87(2), pp. 32-37.

BLEANEY, M. and A. NISHIYAMA (2004), "Income Inequality and Growth: Does the Relationship Vary with the Income Level?" *Economics Letters*, Vol. 84(3), pp. 349-55.

BOURGUIGNON, F. and T. VERDIER (2000), "Oligarchy, democracry, inequality and growth." *Journal of Development Economics*, Vol. 62(2), pp. 285-313.

BUDRIA, S. (2007), "Economic Inequality in Portugal: A Picture in the Beginnings of the 21st century." *Centro de Estudos de Economia Aplicada do Atlântico Working Paper*, pp. 43.

CARDOSO, A. R. (1991), "Desigualdades salariais em Portugal: uma perspectiva regional, 1982-1989." *Masters Dissertation, ISEG, Lisbon*, pp. 143.

CARDOSO, A. R. (1998), "Earnings Inequality in Portugal: High and Rising?" *Review of Income and Wealth*, Vol. 44(3), pp. 325-43.

CARDOSO, F. and V. G. CUNHA (2005), "Household wealth in Portugal: 1980-2004." *Banco de Portugal WP Nº4*, pp. 47.

CARNEIRO, P. (2007), "Equality of Opportunity and Educational Achievement in Portugal." *Portuguese Economic Journal*, Vol. (forthcoming), pp.

CASTELLÓ-CLIMENT, A. and R. DOMÉNECH (2002), "Human Capital Inequality and Economic Growth: Some New Evidence." *The Economic Journal*, Vol. 112, pp. 187-200.

CHEN, B. L. (2003), "An Inverted-U Relationship between Inequality and Long-Run Growth." *Economics Letters*, Vol. 78(2), pp. 205-12.

CHIU, W. (1998), "Income inequality, human capital accumulation and economic performance." *Economic Journal*, Vol. 108, pp. 44-59.

CLARKE, G. R. G. (1995), "More evidence on income distribution and growth." *Journal of Development Economics*, Vol. 47, pp. 403-427.

DEININGER, K. and L. SQUIRE (1996), "A new data set measuring income inequality." *World Bank Economic Review*, Vol. 10(3), pp. 565-91.

DEININGER, K. and L. SQUIRE (1997), "Economic growth and income inequality: reexamining the links." *Finance and Development*, pp. 38-41.

DEININGER, K. and L. SQUIRE (1998), "New ways of looking at old issues: inequality and growth." *Journal of Development Economics*, Vol. 57(2), pp. 259-287.

DRAZEN, A. (2000). *Political Economy in Macroeconomics*. Princeton, Princeton University Press.

DURLAUF, S., P. A. JOHNSON and J. TEMPLE (2004). Growth Econometrics. Handbook of Economic Growth. P. Aghion and S. Durlauf (eds). Amsterdam, North-Holland.

EC (2007), Social Inclusion and Income Distribution in the European Union 2007. *Brussels, [report 1]: European Observatory on the Social Situation - Social Inclusion and Income Distribution Network*. Brussels, European Commission, pp. 312.

EC (2008a), Child Poverty and Well-Being in the EU, Current Status and Way Forward. *Brussels, [report 3]: D-G for Employment, Social Affairs and Equal Opportunities, Unit E.2*. Brussels, European Comission pp. 246

EC (2008b), Progress Towards the Lisbon Objectives in Education and Training, Indicators and Benchmarks. *Brussels, [report 4]: D-G Education and Culture, Unit A4* Brussels, European Comission pp. 232.

EC (2008c), The Social Situation in the European Union 2007. *Brussels, [report 2]: D-G for Employment, Social Affairs and Equal Opportunities - Unit E.1 and Eurostat - Unit F.3*. Brussels, European Comission pp. 199.

FORBES, K. (2000), "A reassessment of the relationship between inequality and growth." *American Economic Review*, Vol. 90(4), pp. 869-887.

GALOR, O. and O. MOAV (2004), "From Physical to Human Capital Accumulation: Inequality and the Process of Development." *Review of Economic Studies*, Vol. 71(4), pp. 1001-26.

GALOR, O. and J. ZEIRA (1993), "Income distribution and macroeconomics." *Review of Economic Studies*, Vol. 60, pp. 35-52.

GOTTSCHALK, P. and T. M. SMEEDING (1997), "Cross-national comparisons of earnings and income inequality." *Journal of Economic Literature*, Vol. 35(2), pp. 633-87.

GOUVEIA, M. and J. TAVARES (1995), "The Distribution of Household Income and Expenditure in Portugal: 1980 and 1990." *Review of Income and Wealth*, Vol. 41(1), pp. 1-17.

KNOWLES, S. (2005), "Inequality and Economic Growth: The Empirical Relationship Reconsidered in the Light of Comparable Data." *Journal of Development Studies*, Vol. 41(1), pp. 135-159.

KUZNETS, S. (1955), "Economic growth and income inequality." *American Economic Review*, Vol. 45, pp. 1–28.

LEVINE, R. and RENELT (1992), "A sensitivity analysis of cross-country growth regressions." *American Economic Review*, Vol. 82(4), pp. 942-963.

LI, H. and H.-F. ZOU (1998), "Income inequality is not harmful for growth: Theory and evidence." *Review of Development Economics*, Vol. 2(3), pp. 318-34.

LÓPEZ, R., V. Thomas and Y. Wang (1998), "Addressing the Education Puzzle: The Distribution of Education and Economic Reforms." *World Bank Working Paper* nº 2031.

PANIZZA, U. (2002), "Income Inequality and Economic Growth: Evidence from American Data." *Journal of Economic Growth*, Vol. 7(1), pp. 25-41.

PARTRIDGE, M. D. (1997), "Is inequality harmful for growth? Comment." *American Economic Review*, Vol. 87(5), pp. 1019-1032.

PEREIRINHA, J. A. (1988), "Inequalities, household income distribution and development in Portugal." *PhD Dissertation, Institute of Social Studies, The Hague, The Netherlands*, pp. 684.

PEROTTI, R. (1993), "Political Equilibrium, Income Distribution, and Growth ", *Review of Economic Studies*, Vol. 60(4), pp. 755-776.

PEROTTI, R. (1994), "Growth, Income Distribution, and Democracy." *Journal of Economic Growth*, Vol. 1(2), pp. 149-187.

PEROTTI, R. (1996), "Democracy, income distribution and growth: What the data say." *Journal of Economic Growth*, Vol. 1, pp. 149-187.

PERSSON, T. and GUIDOTABELLINI (1991), "Is inequality harmful for growth?: theory and evidence." *CEP, Discussion Papers nº581*.

PERSSON, T. and GUIDOTABELLINI (1994), "Is Inequality Harmful to Growth?" *American Economic Review*, Vol. 84(3), pp. 600-621.

PERSSON, T. and G. TABELLINI (2000). *Political Economics. Explaining Economic Policy*. Cambridge, Massuchussets, The MIT Press.

RODRIGUES, C. F. (1996), "Medição e Decomposição da Desigualdade em Portugal [1980/81-1989/90]." *Revista de Estatística*, Vol. 3(3), pp. 49-70.

RODRIGUES, C. F. (1999a). Income Distribution and Poverty in Portugal [1994/95], a comparison between the European Community Household Panel and the Houshold Budget Survey, CISEP, ISEG/UTL WP nº24.

RODRIGUES, C. F. (1999b), "REPARTIÇÃO DO RENDIMENTO E POBREZA EM PORTUGAL [1994/95] [UMA COMPARAÇÃO ENTRE O PAINEL DE AGREGADOS FAMILIARES E O INQUÉRITO AOS ORÇAMENTOS FAMILIARES]." *INE*.

RODRIGUES, C. F. (2007). *Distribuição Do Rendimento, Desigualdade E Pobreza*. Coimbra, Almedina.

RODRIK, D., Ed. (2003). In Search of Prosperity: Analytic Narratives on Economic Growth. Princeton, Princeton University Press.

RODRIK, D. (2005). Growth Strategies. *Handbook of Economic Growth*. P. Aghion and S. Durlauf (eds). Amsterdam, North-Holland, Vol. 1A**,** pp. 967-1014.

RODRIK, D. (2007). *One Economics, Many Recipes: Globalization, Institutions, and Economic Growth*. Princeton, Princeton University Press.

SALA-I-MARTIN, X. (1997), "I just run two million regressions." *American Economic Review*, Vol. 87(2), pp. 178-183.

SALA-I-MARTIN, X. (2006), "The World Distribution of Income: Falling Poverty And ...Convergence, Period." *Quarterly Journal of Economics*, Vol. 121(2), pp. 351-97.

SILBER, J., Ed. (1999). Handbook of Income Inequality Measurement. New-York, Springer

SIMÕES, M. (2006). Levels of Education, tecnhology and growth: the OECD evidence from a country and industry-level perspective. Faculty of Economics. Coimbra, University of Coimbra. Ph.D**,** pp. 287.

STIGLITZ, J. (2002), "Information and the change in the paradigm in economics." *American Economic Review*, Vol. 92(3), pp. 460-501.

SUMMERS, R. and A. HESTON (1991), "The Penn World Table (Mark V): an expanded set of international comparisons." *Quarterly Journal of Economics*, Vol. 106(2), pp. 327-368.

SYLWESTER, K. (2000), "Income inequality, education expenditures, and growth." *Journal of Development Economics*, Vol. 63(2), pp. 379-398.

TEMPLE, J. (1999), "The new growth evidence." *Journal of Economic Literature*, Vol. 37(1), pp. 112-156.

VIEIRA, J. C., J.P.COUTO and M. T. Tiago (2005), "Inter-Regional Wage Dispersion in Portugal." *IZA Discussion Paper* nº 1664.

VOITCHOVSKY, S. (2005), "Does the profile of income inequality matter for economic growth? Distinguishing between the effects of inequality in different parts of the income distribution." *Journal of Economic Growth*, Vol. 10, pp. 273-296.

APROXIMACIÓN AL COMPORTAMIENTO DE LOS PRECIOS EN UNA ECONOMÍA INSULAR. EL CASO DE CANARIAS

SANTIAGO RODRÍGUEZ FEIJOÓ
E-mail: srfeijoo@dmc.ulpgc.es
CARLOS GONZÁLEZ CORREA
ALEJANDRO RODRÍGUEZ CARO
Dpto. de Métodos Cuantitativos en Economía y Gestión
Universidad de Las Palmas de Gran Canaria

Resumen

El objetivo de este trabajo es aportar evidencia empírica a nivel microeconómico sobre el comportamiento de los precios en una economía pequeña, que presenta un mercado fraccionado, con el fin de buscar diferencias con respecto al comportamiento de los precios en mercados más amplios que puedan incidir en el ciclo de negocio de las empresas que están asentadas en su territorio y en los efectos sobre la economía del mismo por la aplicación de políticas económicas. Para ello, con datos de Canarias estudiamos el precio y su rigidez, caracterizada esta última a través de la frecuencia, la duración y el tamaño de los cambios de precios. Los resultados del análisis realizado muestran unos precios que cambian con mayor frecuencia en el mercado canario que en el conjunto de los países europeos y son similares en cuanto al tamaño de dicho cambio y a la falta de rigidez a la baja de los precios.

Palabras clave: Precios, micro datos, rigidez, dispersión, pequeñas economías
Área Temática: Economía del Crecimiento económico.

Abstract

The objective of this work is to provide micro-level empirical evidence on the behavior of prices in a small economy, which has a fractional market, in order to find differences in the behavior of prices in larger markets that may affect the business cycle of the established companies and the impact thereof on the economy by implementing economic policies. We studied, using data from the Canary Islands, price and its rigidity, the latter characterized by frequency, duration and size of price changes. The results of the analysis show that prices change more frequently in the Canaries than in all European countries and changes are similar in size and the lack of a downward rigidity of prices.

Thematic Area: Growth Economics.

1. INTRODUCCIÓN

En los últimos 30 años han aparecido múltiples trabajos que tienen como objetivo el estudio de la rigidez de los precios: Carlton (1996), Cecchetti (1986), Lach and Tsiddon (1992), Eden (2001), Bils and Klenow (2004), Baharad and Eden (2004), Dhyne et al.(2006), Álvarez y Hernando (2006) y Baudry et al. (2007), etc.

Bils y Klenow (2004) justifican el interés por el tema diciendo que la rigidez es una cuestión fundamental cuando se pretende modelizar las fluctuaciones de los ciclos de negocio o cuando se estudian políticas monetarias en las cuales se asume que las empresas ajustan precios con muy poca frecuencia. Trabajos como los de Taylor (1999) o Calvo (1983) incorporan rigidez en los precios, sin embargo Bils y Klenow (2004) sugieren que los modelos que se usan habitualmente para explicar el cambio temporal de precios debieran permitir cambios de precios con mayor frecuencia. Por otra parte, la conducta de los precios es una parte fundamental al abordar cualquier aspecto relacionado con la Teoría del Consumidor.

La literatura actual se ha centrado en estudios microeconómicos de ámbito nacional o supranacional y ha prestado poca importancia al entorno regional, en gran medida por la menor disponibilidad de datos para su análisis y bajo el supuesto que las regiones en mercados integrados deben tener un comportamiento de los precios similares a los mercados que las integran.

Sin embargo, existen unas regiones que pueden ser especialmente sensibles a soportar conductas de los precios específicas, son las regiones insulares. Estas regiones, como por ejemplo el caso de la región canaria, se caracterizan por distanciarse geográficamente del centro del mercado nacional y por tener un mercado propio fraccionado. Todo ello puede provocar que la conducta de sus precios difiera del que se presenta a nivel español y europeo, lo cual podría tener como consecuencia una falta de eficacia y eficiencia en las medidas que pretenden organizar y estructurar su economía.

El objetivo de este trabajo es estudiar a nivel de microdatos el comportamiento de los precios en Canarias y compararlo con el comportamiento que se describe en la literatura para otras zonas de ámbito geográfico más amplio. La comparación tendrá limitaciones tanto por el nivel de representación de los datos utilizados en el conjunto del consumo, sólo se utilizaran datos de cuatro productos del grupo de alimentación y bebidas, como por el número de registros disponibles, en torno a los 17.000 registros, y por el período muestral, desde el año 2003 al año 2008, por tanto utilizando como unidad de medida el euro, frente al resto de trabajos que finalizan el período muestral justo antes de la entrada en vigor de la moneda única.

En lo que resta el trabajo se estructura de la siguiente manera: en el epígrafe siguiente se resumen los principales resultados de los trabajos previos, en el punto tercero se describen los datos disponibles, en el punto cuarto se estudia el comportamiento de los precios y en el quinto se presentan los principales resultados que caracterizan el cambio de los precios en Canarias, comparándolos con el contexto español, europeo y americano. El último epígrafe contiene las principales conclusiones.

2. EVIDENCIA EMPÍRICA SOBRE EL COMPORTAMIENTO MICRO DE LOS PRECIOS

Bils y Klenow (2004), usando datos del U.S. Bureau of Labor Statistics para el período 1995-1997, obtienen una frecuencia media de cambio de 26,1% lo que les lleva a estimar una duración media de los precios de 3,3 meses, obteniendo una fuerte dispersión en función del tipo del bien considerado. Para estimar la duración media a partir de la frecuencia del cambio (λ) lo hacen como $\frac{-1}{\ln(1-\lambda)}$. Sin embargo, la información con la que cuentan Bils y Klenow (2004) no le permite medir empíricamente la duración media de los precios ni el tamaño de su cambio.

Dhyne et al. (2006) se centran en caracterizar la frecuencia media y el tamaño del cambio en precios en la zona euro, comparando sus resultados, en aquello que es posible, con los resultados obtenidos por Bils y Klenow (2004) y Klenow y Kryvtsov (2008). Los datos que utilizan Dhyne et al. (2006) son secuencias de precios para un producto específico en un establecimiento específico. Trabajan con un total de 50 productos o servicios con precios que abarcan el período temporal enero de 1996 a enero de 2001 y analizan siete indicadores, a saber: la frecuencia con que cambian los precios, la duración media del precio, la frecuencia de los incrementos de precios, la frecuencia de los decrecimientos de precios, el tamaño medio del incremento de precios, el tamaño medio del decrecimiento de precios y el grado de sincronización del cambio de precios. Sus principales conclusiones son: A) La duración media de los precios en la zona euro duplica a la observada por Bils y Klenow (2004). Dhyne et al. (2006) utilizan el inverso de la frecuencia del cambio para estimar la duración de los precios y partiendo de que la frecuencia del cambio es 15,1% y 24,8% para la zona euro y para US, respectivamente, obtienen una duración media de 13 meses para la zona euro y de 6,7 meses

para US (para las 50 categorías de productos o servicios)[1]. B) El tipo de producto es un factor determinante en la frecuencia del cambio y, por tanto, en la estimación de la duración media, siendo mucho más heterogéneo el comportamiento en la zona euro que en US. C) Existe heterogeneidad entre países en el cambio en precios, sin embargo, es menos relevante que la heterogeneidad entre sectores. D) No existe rigidez a la baja de los precios. En la zona euro, el 40% de los cambios en precios se corresponden a reducciones, en US la estimación es del 45%. Profundizando en este aspecto, Dhyne et al. (2006) también encuentran que: los productos con mayor incidencia en el crecimiento de precios también son los que tienen mayor incidencia en su decrecimiento justificándolo en las diferencias en las condiciones de oferta y demanda entre sectores; el sector servicios presenta una clara rigidez a la baja, reduciendo únicamente al 20% las caídas de precios con respecto a los cambios totales, justificando dicha especificidad por el alto peso de la mano de obra en los servicios y la rigidez a la baja de los salarios. E) La reducción media de precios es mayor que la subida media de precios. En concreto la reducción media es del 10% mientras que la subida media es del 8,2% en la zona euro mientras que en U.S. la reducción media es del 14,1% frente a un crecimiento medio del 12,7%. Este comportamiento es justificado por los autores por la reticencia que tienen las empresas a bajar los precios, haciéndolo solamente cuando los shock negativos son grandes. F) Midiendo la sincronización con el índice propuesto por Fisher y Konieczny (2000), la conclusión de Dhyne et al. (2006) es que existe poca sincronización en los cambios de precios a través de los productos.

Álvarez y Hernando (2006) estudian el caso español con datos para el período 1993-2001, destacando los siguientes resultados: A) La frecuencia del cambio de los precios en España es ligeramente más baja que en la mayoría de los países europeos. Los autores justifican este resultado por tres factores relacionados con la base de datos con la que trabajan. A saber, no incluye productos energéticos, no incluyen promociones y sí incluyen productos con precios controlados. B) La frecuencia del cambio es muy heterogénea entre productos destacando la elevada frecuencia del cambio de los productos de alimentación no elaborados. C) No existe evidencia de rigidez en los precios. Para el conjunto de precios, el 40% de los cambios de los precios son reducciones, siendo menos frecuentes las reducciones en los servicios y en los bienes industriales no energéticos. Ahora bien, para todos los tipos de bienes, la frecuencia del cambio es mayor en los incrementos que en las reducciones y a mayor frecuencia de cambio, mayor frecuencia de las reducciones de precios. D) La duración media de los precios es de 15,1 meses, presentando una alta heterogeneidad por tipo de producto, con un recorrido que va desde los tres meses para los productos de alimentación sin elaboración hasta los 20,1 de los servicios. E) Por último, en cuanto al tamaño de los cambios, tanto positivos como negativos, los resultados para España son muy similares a los encontrados para la zona euro, descritos en Dhyne et al. (2006).

3. LOS DATOS

Los datos que se usan son observaciones mensuales de precios que recoge el Instituto Canario de Estadística (ISTAC) con el fin de elaborar la operación estadística Cesta de la Compra. Encuesta de Precios-Canarias. La metodología puede consultarse en ISTAC(2008).

La muestra está formada por registros de precios para cuatro productos para un período máximo comprendido entre septiembre del año 2003 y febrero de 2008. Para cada precio individual se dispone de los siguientes datos: la identificación del producto, etiquetado como producto 1 a producto 4, el año y el mes en que fue medido el precio y el código identificativo del establecimiento en el que se observó el precio. Además, los productos 1 y 2 pertenecen al grupo de alimentación, en concreto pertenecen al grupo COICOP 01.1, el producto 3 al grupo de bebidas no alcohólicas (COICOP 01.2) y el producto 4 al grupo de bebidas alcohólicas (COICOP 02.1). Para los dos productos del grupo COICOP 01.1 se conoce la isla en la cual se

[1] Los autores justifican esta diferencia por la elevada rotación de establecimientos en la recogida de precios en US, procedimiento que no siguen las agencias europeas de estadística.

obtuvo la información y para los otros dos productos se conoce el tipo de establecimiento, que puede ser bien grandes superficies/hipermercados o bien supermercados grandes, pequeños y medianos. Los establecimientos en los que se recogen precios de los productos 1 y 2 son los mismos pero distintos a los que se observó el precio para los productos 3 y 4. El total de establecimientos muestreados han sido 197, de los cuales en 99 se recogieron precios de los productos 1 y 2 y en otros 98 precios de los productos 3 y 4.

Los productos están perfectamente definidos incluyendo la marca y el empaquetado. Sin embargo, también se dispone de una variable que nos indica la cantidad que implica el precio medido. Mayoritariamente la cantidad es igual a 1 (en un 89,82% de los casos). La distribución del 10,18% restante entre los cuatro productos es muy heterogénea. En el producto 4 no hay ningún caso con cantidad distinta de 1, en el dos sólo un caso, en el producto 1 hay 208 casos y la mayor concentración se da en el producto 3 con 1.576 casos, lo cual supone el 88,29% del total de valores de cantidad distintos de 1 y el 33,76% de los precios de dicho producto.

Con estos datos se dispone de 197 series temporales con una duración máxima que abarca desde septiembre de 2003 hasta febrero de 2008. Cada una de estas series mide, a lo largo del tiempo, el precio de uno de los cuatro productos en un establecimiento concreto, que a su vez pertenece bien a una determinada isla (productos 1 y 2) o a un tipo de establecimiento (productos 3 y 4). El tamaño máximo de estas series es de 54 datos, pero existe un importante número de datos faltantes. De hecho, para el producto 1 únicamente se dispone de 35 series completas, lo cual supone un el 35,71% del total de rachas completas. Para el producto 2 no hay ninguna serie completa debido que a que no se dispone del dato del mes de mayo del año 2005, pero se dispone de 55 series (56,12%) a las cuales únicamente les falta dicho dato. El producto 3 es el que más series tiene completas, un total de 59 sobre 99. Para estos tres productos las series más pequeñas son de 5 datos. El producto 4 es el que más datos faltantes tiene, además sólo tiene una serie completa y la moda para el tamaño de las series es 45. Con esta estructura, finalmente se dispone de un total de 17.470 precios.

4. ANÁLISIS DE LOS PRECIOS

El primer análisis consiste en una descripción de los precios tomando como referencia bien el tiempo. La tabla 1 muestra la una descriptiva de los precios de los cuatro productos.

Tabla 1. Estadística descriptiva de los precios

Producto	Media	Mediana	Tercer cuartil	Desv. Típica	Asimetría	Curtosis	Obs.
1	0,91	0,88	1,0	0,18	1,94	10,33	4.384
2	0,79	0,75	0,84	0,13	1,54	5,95	4.476
3	1,24	1,19	1,31	1,18	1,85	10,83	4.577
4	0,60	0,59	0,63	0,08	1,35	5,89	4.033
Todos	0,89	0,80	1,1	0,28	0,81	3,71	17.470

Como se puede observar los cuatro productos analizados tienen un precio moderado, presentando todos ellos una clara asimetría positiva y un comportamiento leptocúrtico muy acentuado, con una fuerte correlación entre el nivel de asimetría y el nivel de curtosis (0,88). Es decir, la gran mayoría de los precios se sitúan muy cerca de su media, pero existen algunos precios especialmente grandes con respecto al resto.

En el análisis temporal destaca en los cuatro productos los grandes cambios que se produce en la curtosis y la relación cuadrática que mantiene con la asimetría, tal y como se puede comprobar en la salida de regresión estimada que se muestra en la ecuación (1).

$$Curtosis = 2,78^{**} + 1,41^{**} * Asimetria^2$$

$$R^2 = 0,90 \qquad (1)$$

** Significativo al menos al 1%

Este resultado es consistente con un modelo de difusión escalar para los precios [Schopflocher and Sullivan (2005), página 343].

El siguiente análisis se centra en los momentos de primer y segundo orden. Del análisis de estos dos momentos se puede concluir la forma en la que se desplazan las distribuciones de precios, lo cual tiene consecuencias tanto sobre la fórmula qué se debe utilizar para el cálculo del índice elemental, como para el comportamiento del consumidor. Por ejemplo, un crecimiento de los precios medios acompañados de una estabilidad de los coeficientes de variación, nos indicaría que las distribuciones de precios se desplazan de forma multiplicativa, lo cual implica: a) las fórmulas de Dutot, Jevons y Bergson (elasticidad de sustitución igual a 2) son igualmente válidas para el agregado elemental, y b) el consumidor puede optimizar su gasto seleccionando el punto de compra.

Para corroborar esta hipótesis, a continuación analizamos la evolución de los momentos de primer y segundo orden, cuya gráfica se muestra en la ilustración 1. Como se puede observar, los patrones de cambio no son homogéneos entre los productos, ni en valores medios ni en su dispersión. Los tres primeros productos muestran comportamientos fuera de lo normal en su media al final del período muestral. Los dos primeros con fuertes crecimientos al final de la serie y el tercero con una caída de los precios muy grande desde comienzos del año 2007. Siguiendo con la media, los productos 2 y 4 presentan un claro comportamiento creciente, como se esperaba, mientras que el 1 y el 3 nos induce a pensar que se ajustan a formas cuadráticas. Si nos fijamos en la dispersión, tanto absoluta como relativa, se observan crecimientos en los productos 1, 2 y 3 y estabilidad en el 4. En los anexos 1 y 2 se muestran los resultados de la modelización de la media y del coeficiente de variación, respectivamente, en función del tiempo y de outliers impulso. Como puede verse, el nivel de ajuste para los valores medios es muy elevado, presentando un coeficiente de determinación de 0,997, mientras que para el coeficiente de variación este estadístico es de 0,77. Sin embargo, la evolución de la media con respecto al tiempo tiene una estructura de polinomio de grado 2 para los productos 1, 2 y 3, cóncava ($\cup$) para los productos de alimentación y convexa ($\cap$) para el producto 3. Estos resultados contradicen la idea de que la tendencia de los pecios es creciente. Para el producto 4 la evolución es lineal y para el conjunto de los 4 productos la media está sujeta a un efecto externo que se concentra fundamentalmente a finales del año 2007 y principios del 2008, período que coincide con un período fuertemente inflacionista, explicado en gran medida por el crecimiento del precio del petróleo y de los cereales.

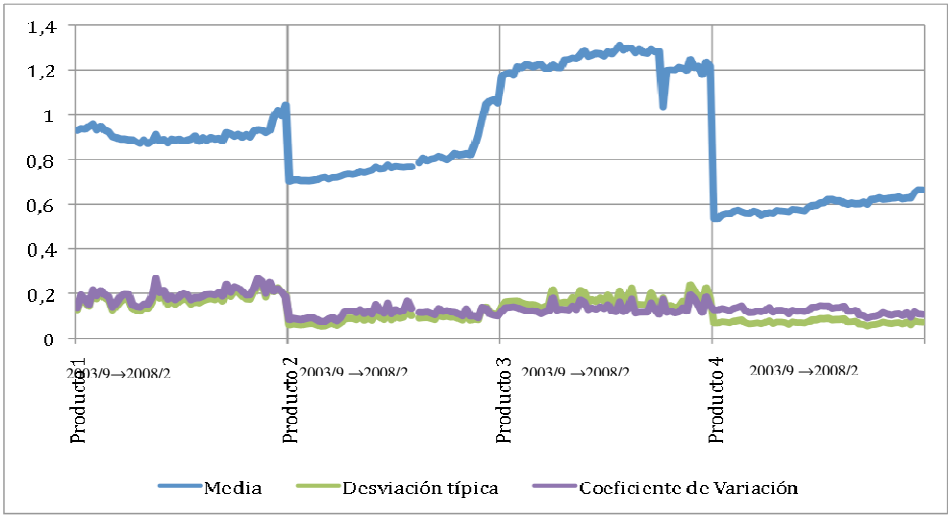
El comportamiento descrito para la media no se mantiene para el coeficiente de variación. En este último caso, los outliers impulso no son genéricamente significativos (han sido eliminados de la regresión) y con la excepción del producto 2, que tiene un cambio de dispersión convexo ($\cap$), los otros tres producto no presentan cambios significativos de su dispersión relativa a lo largo del tiempo.

Por tanto, existe evidencia que apoya la hipótesis de que las distribuciones de precios se desplazan de forma multiplicativa, lo que implica que el consumidor puede optimizar su gasto en el sentido de que el rango de los precios es mayor, pero necesita mayor información para ello.

Evidentemente el análisis realizado no permite identificar los movimientos que se producen dentro de las distribuciones de precios. Sí es verdad que el comportamiento poco esperado en las medias, junto con la estabilidad en la dispersión relativa y los fuertes cambios en los momentos de tercero y cuarto orden, nos indican que las distribuciones de precios sufren fuertes cambios.

Con el objeto de identificar posibles causas de estos cambios, se ha estimado un modelo para la variable precio utilizando como variables explicativas el establecimiento, el mes, el tiempo y el tiempo al cuadrado. Además, para los productos 1 y 2 se ha introducido la variable isla y para el 3 y 4 la variable tipo de establecimiento. El objetivo es doble. Por un lado comprobar si la forma cuadrática se mantiene a nivel individual y, por otro lado, buscar evidencia empírica sobre el efecto establecimiento, isla y tipo de establecimiento. Para ninguno de los productos se ha obtenido un efecto mes estadísticamente significativo, por lo que finalmente no se han introducido en el modelo.

Ilustración 1. Evolución temporal de los momentos de orden 1 y 2 de los precios



La tabla 2 muestra los principales resultados obtenidos con estos modelos. Los datos que se muestran en la tabla se corresponden con los resultados obtenidos al estimar 4 modelos, uno para cada producto, teniendo todos ellos la misma variable endógena: el precio unitario del producto. La ecuación del producto 1 se estimó con 4.384 datos y la del producto 2 con 4.476 datos de precios unitarios. En ambos casos las variables explicativas fueron los establecimientos (98 establecimientos), la variable tiempo, que toma valor 1 para el mes de noviembre de 2003 hasta 54 para el mes de febrero de 2008 que se usa iterando, tanto en forma lineal como cuadrática con cada una de las islas (7 islas). El mismo esquema se mantiene para las ecuaciones de los productos 3 y 4, pero en este caso se dispone de información sobre el tipo de establecimiento en el que se observó el precio y no de la isla, con lo cual sólo se dispone de dos posibles valores.

Tabla2. Modelos explicativos de los cambios en el precio unitario

Variable endógena: Precio unitario						
Exógenas (nº de variables)	Datos muestrales	Significatividad (al menos 5%)	R^2	Efecto Establecimiento*	Tiempo	
					Lineal-Isla*	Cuadrática-Isla*
Establecimientos (98) Islas*tiempo (7) Islas*tiempo^2 (7)	Producto 1 (4.384)	Excepto forma lineal y cuadrática para El Hierro	59,4%	71,8% (4.753 contrastes)	38,2% (21 contrastes)	28,6% (21 contrastes)
	Producto 2 (4.476)	Todas	69,9%	57,8% (4.753 contrastes)	9,5% (21 contrastes)	4,8% (21 contrastes)
Establecimientos (99) Tipo_establecimiento*tiempo (2) Tipo_establecimiento*tiempo^2 (2)					Lineal-Tipo establecimiento	Cuadrática-Tipo establecimiento
	Producto 3 (4.577)	Todas	59,1%	72,5% (4.851 contrastes)	0% (1 contraste)	100% (1 contraste)
	Producto 4 (4.033)	Excepto forma cuadrática para tipo de establecimiento 2	61,8%	65,8% (4.851 contrastes)	0% (1 contraste)	100% (1 contraste)

* Porcentaje de rechazo de los contrastes de Wald sobre las igualdad entre cada dos parámetros de cada efecto

La tercera columna aporta más evidencia de que los precios para el período analizado presentan una estructura cuadrática con respecto al tiempo, con la excepción del la isla de El Hierro para el producto 1 y para el tipo de establecimiento 2 (supermercado grandes, medianos y pequeños) para el producto 4.

Los valores de los coeficientes de determinación (comprendidos entre el 59,1% y el 69,9%) son lo suficientemente elevados como para afirmar la relevancia del establecimiento, el tiempo, la isla y el tipo de establecimiento a la hora de fijar el precio de un producto. Sin embargo, los resultados hasta ahora comentados no nos permiten analizar el efecto diferencial de estos factores sobre el precio. Es decir, ¿el efecto establecimiento se debe a que existen unos pocos

establecimientos que son más caros/baratos que el resto? o, por el contrario, ¿cada establecimiento tiene un una forma de fijar precios propia?. Las tres últimas columnas de la tabla 2 contienen información que nos permite contestar a esta pregunta y resumen los resultados obtenidos con la realización de 19.296 contrastes de Wald de igualdad entre los distintos coeficientes de regresión, dos a dos. De estos contrastes, 4.753 se corresponden con contrastes de igualdad para establecimientos del producto 1, igual cantidad para el producto 2 y 4.851 para los productos 3 y 4 respectivamente. En el 71,8% de estos contrastes realizados entre los establecimientos que venden el producto 1 rechazan la hipótesis de igualdad. Es decir, en el 71,8% de los casos existe un efecto establecimiento estadísticamente distinto de cero y distinto de otros establecimientos. Una cifra ligeramente superior se obtiene para el producto 3 y ligeramente inferior para los productos 2 y 4. En consecuencia, se pude concluir que el establecimiento (y todo lo que ello conlleva: localización, accesibilidad, calidad de servicio, barreras de entrada, etc…) es un elemento fundamental en la determinación del precio de venta.

En las dos últimas columnas de la tabla 2 se muestran los mismos resultados con el objeto de analizar el efecto diferencial del tiempo, forma lineal y forma cuadrática. De su análisis se puede concluir que el factor tiempo afecta a los precios de una forma mucho más homogénea, aunque significativa, concentrándose las diferencias en las islas de El Hierro y fundamentalmente La Palma.

5. ANÁLISIS DEL CAMBIO DE LOS PRECIOS

En el estudio del cambio en precios tres son los datos fundamentales: la frecuencia con la que cambian los precios, la duración de los precios y el tamaño de los cambios.

5.1 La frecuencia del cambio

En la Tabla 3 se muestran las frecuencias de cambio de los precios por producto, distinguiendo crecimientos y decrecimientos. Cada celda de la tabla contiene el número de cambios y su importancia relativa con respecto al total por producto.

El elemento que más destaca es la alta frecuencia con la que se producen reducciones de precios. Sin embargo, esta falta de rigidez a la reducción de precios es un resultado que se observa en todos los trabajos referenciados y compatible con los resultados del apartado anterior. Como se puede observar, por término medio, por cada 10 incrementos de precios se producen 7,6 decrecimientos, existiendo diferencias en función del producto, con una orquilla que va desde el ratio 19/9,4 para el producto 1 hasta 10/6,2 para el producto 2. Sin embargo, lo que sorprende al lector especializado es la alta frecuencia con la que se producen los cambios. Como ya hemos dicho, los productos 1 y 2 se corresponden con alimentos elaborados y el 3 y 4 con bebidas pertenecientes a al grupo de bienes industriales no energéticos. Álvarez y Hernando (2006) obtienen una frecuencia de cambio del 18,6% (11% de incrementos y un 7% de decrecimientos) para los primeros y del 7% (5% de incrementos y un 2% de decrecimientos) para los segundos, muy inferiores a los datos que se muestran en la tabla 3.

Tabla 3. La frecuencia en los cambios de precios

Producto	Decrecimiento de precios	Crecimiento de precios	Estabilidad de precios	Total cambios
Producto 1	795 (19,11%)	847 (20,36%)	2.518 (60,53%)	4.160
Producto 2	559 (13,13%)	903 (21,21%)	2.796 (65,66%)	4.258
Producto 3	703 (15,82%)	910 (20,47%)	2.832 (63,71%)	4.445
Producto 4	603 (15,86%)	824 (21,67%)	2.376 (62,48%)	3.803
Total	2.660 (15,96%)	3.484 (20,90%)	10.522 (63,13%)	16.666

¿Es este resultado una consecuencia de que los datos provienen de una economía insular con un mercado fragmentado, estando afectado por el factor isla o tipo de establecimiento? Hemos calculado la tabla 3 para islas en vez de para productos y, aunque existen diferencias significativas entre ellas, las cifras son extremadamente altas para todas ellas. La frecuencia más baja la presenta la isla de La Gomera, con una frecuencia del 29% (16,8% de incrementos y un 12,3% de decrecimientos) y la más alta Lanzarote con un 44% (24,0% de crecimientos y un 20,6% de decrecimientos), en todos los caso muy por encima de los referenciados por Álvarez y Hernando (2006) y más cercanas a las obtenidas por Dhyne et al. (2006) para los Estados Unidos (frecuencia del cambio en precios del 27,1% para alimentos procesados). Repetimos el cálculo para el tipo de establecimiento y no se ha observado diferencias significativos en este factor.

Es evidente que estamos comparando cifras que no son homogéneas, dado que en nuestro caso únicamente disponemos de cuatro productos, mientras que en los trabajos con los que comparamos nuestros resultados son resultados medios que incorporan muchos más. El trabajo de Bils y Klenow (2004) nos permite ser más precisos en la comparación, con respecto a Estados Unidos. De los niveles elementales que se definen en la tabla A1 del trabajo referenciado, el más relacionado con el producto 1 es el 10021 y presenta una frecuencia de cambio del 32,9% frente al 39,47% que se observa con nuestros datos. El producto 2 no se encaja en ninguno de los agregados elementales que presenta la tabla A1, siendo los más próximos los agregados elementales 9021 y 9011, que presentan una frecuencia de cambio de 31,6% y 34,4% respectivamente, muy similar esta última a la que se observa para el producto 2 en Canarias. El producto 3 pertenece al agregado elemental 17011 observándose una frecuencia muy similar entre ambos: un 36,29% según la tabla 3 y un 38,8% para el agregado 17011. Por último el producto 4 pertenece al agregado elemental 20051. Dicho agregado tiene una frecuencia de cambio de 6,4% mientras que el producto 4 se eleva hasta el 37,53%. Como se puede observar, con la excepción del producto 4, la frecuencia del cambio en precios en Canarias es mucho más similar a la de los Estados Unidos que a la de España o a la de la Unión europea.

Si hemos de hacer caso a este último resultado, no se justifica la alta frecuencia de los cambios de los precios en Canarias por razones de insularidad o de mercado fraccionado.

5.2. La duración de los precios

La frecuencia y la duración son dos medidas mutuamente relacionadas. En datos transversales, si los cambios en precios son estacionarios y homogéneos y el tiempo es discreto, Baudry et al. (2005) demuestran que la inversa de la frecuencia del cambio en precios converge en muestran grandes a la duración media de los precios. Si el tiempo es continuo, se admiten varios cambios dentro de un mismo período de observación, y el ratio de Hazard es constante, la duración media se calcula como el valor negativo del inverso del logaritmo neperiano de 1 menos la frecuencia. Baharad y Eden (2004) muestran las grandes disparidades que se pueden producir al calcular la duración media cuando se parte del dato de la frecuencia.

Por otra parte, si se dispone de las series temporales para cada establecimiento, se pude calcular el tiempo para el cual permanece constante un cierto precio. El cálculo no es complicado, pero puede ser extremadamente engorroso al no disponer de series completas y siempre está sujeto a un sesgo potencial por truncamiento de los datos al principio y al final de cada tramo temporal completo[2]. Como consecuencia de esta medición se obtiene una variable de recuento que toma valores enteros desde 1, cuando el precio sólo dura un mes, hasta S, que indica que el precio se ha mantenido constante durante S meses. La media de esta serie nos daría la duración media del precio. Ahora bien, esta media está sometida a un posible sesgo, dado que la primera y la última duración que se observa en cada una de los tramos de series temporales completas infravaloran el verdadero tamaño de la duración.

[2] Para problemas concernientes a la relación entre la frecuencia y la duración y la medición de está última, también se puede consultar Baudry et al. (2005) y Baudry et al. (2007).

En la tabla 4 se muestran los valores medios de la duración obtenidos con las distintas fórmulas comentadas.

Tabla 4. La duración de los precios

Producto	Enfoque de la frecuencia (F*)				Enfoque de la duración (D**)			
	Enfoque discreto		Enfoque continuo		Con todas las duraciones		Excluyendo las duraciones incompletas	
	Media $1/F$	Mediana $\frac{-\ln(0,5)}{F}$	Media $\frac{-1}{\ln(1-F)}$	Mediana $\frac{\ln(0,5)}{\ln(1-F)}$	Media (observaciones) $\frac{1}{n}\sum_{i=1}^{n} D_i$	Frecuencia Duración=1	Media (observaciones) $\frac{1}{n}\sum_{i=1}^{n} D_i$	Frecuencia Duración=1
Producto 1	2,53	1,76	1,99	1,38	2,35 (1.866)	59,64%	2,26 (1.675)	61,31%
Producto 2	2,91	2,02	2,38	1,65	2,66 (1.680)	50,36%	2,60 (1.488)	52,22%
Producto 3	2,76	1,91	2,22	1,54	2,62 (1.745)	57,25%	2,59 (1.548)	59,43%
Producto 4	2,66	1,85	2,13	1,47	2,44 (1.652)	50,12%	2,38 (1.461)	52,84%
Total	2,71	1,88	2,17	1,51	2,52 (6.943)	54,50%	2,46 (6.172)	56,64%

* Frecuencia del cambio en precios
** Duración de cada precio

La tabla recoge distintas estimaciones de la duración de los precios para el total de los precios y por productos. En las columnas 2 a 5 se recogen las estimaciones para la duración media y mediana obtenidas a partir del enfoque de la frecuencia. Bajo este enfoque, la duración media de los precios se sitúa entre los 2,71 meses y los 2,17 meses, con una orquilla por producto entre los 2,53 meses para el producto 1 hasta los 2,91 meses para el producto 2, para el caso discreto y 1,38 meses y 2,38 meses para el enfoque continuo. Como se puede observar, la distancia entre los valores medios entre el enfoque continuo y discreto si sitúa en torno al medio mes para el caso discreto, como era de esperar. El análisis de la mediana nos indica que la distribución de la duración es claramente asimétrica positiva, resultado también esperado.

En las cuatro últimas columnas de la tabla 4 se muestran las estimaciones de la media y la frecuencia de precios que únicamente duran un mes, bajo el enfoque de la duración, en las columnas 6 y 7 usando la duración con todos los datos disponibles y en las columnas 8 y 9 eliminando de las anteriores las duraciones que pueden presentar problemas de truncamiento o censura. En las columnas de las medias, entre paréntesis, se indican los tamaños muestrales con las que se calculó el valor promedio. Según el enfoque de la duración, la duración media de los precios se sitúa entre los 2,52 meses y los 2,46 meses, según la variable de duración que es utilizada, valor que es ligeramente inferior al enfoque de la frecuencia discreto y superior al enfoque de la frecuencia continuo. Los resultados por productos mantienen la misma pauta que el total y, además, el orden de las medias por productos es el mismo para el enfoque de la frecuencia que para el enfoque de la duración. Esto es, la duración más corta la presenta el producto 1, seguido del 4, el 3 y por último el que presenta la media más grande es el 2. Un dato que sorprende es el de la frecuencia con la que aparecen precios que duran como máximo un mes. Más de la mitad de los precios sólo duran un mes en el mercado, con lo cual la mediana es para los cuatro bienes igual a 1. Este dato es claramente inferior al que se obtiene en otros trabajos [Baudry et al. (2005), Baudry et al. (2007)].

El uso del enfoque de duración nos permite realizar de forma sencilla distintos contrastes de hipótesis, mediante los cuales podemos concluir: a) no existe diferencias significativas entre los valores medios ni entre las varianzas en función de la variable duración que se utiliza y b) sí existen diferencias significativas en la duración media, mediana y varianza en función del producto considerado, independientemente de la variable duración utilizada[3].

Para los productos 1 y 2 se ha realizado un análisis similar por islas y para los productos 3 y 4 por tipo de establecimiento. Por islas, tanto la media, como la mediana y la varianza muestra

[3] Se han realizado los siguientes contrastes: para la media mediante un contraste de Welch, para la mediana el contraste de van den Vaerden y para la varianza el test de Levene.

diferencias estadísticamente significativas, mientras que por tipo de establecimiento se acepta para los tres parámetros su igualdad.

Por último, se han estimado dos modelos de recuento[4], en ambos la variable endógena es la variable duración[5], uno para los productos 1 y 2 y otro para los productos 3 y 4, usando como variables explicativas el año de observación, el mes, el producto, el precio unitario, la interacción entre el producto y el precio unitario y el establecimiento. Los niveles de ajuste son muy similares para ambos modelos (L-likelihood -5.982,4 para el modelo con los productos 1 y 2 y -5.807,4 para el modelo con los productos 3 y 4), pero el efecto de las variables es distinto. La variable duración para los producto 1 y 2 depende negativamente del precio unitario y tiene un efecto estadísticamente distinto para los dos productos. El paso de los años incrementa la probabilidad de que la duración sea mayor y existe un claro efecto estacional, siendo los meses de enero febrero, marzo, abril, mayo y octubre los meses que presentan una duración mayor. Además, sólo la mitad de los establecimientos muestran efectos estadísticamente significativos entre ellos, siendo estos fundamentalmente pertenecientes a la isla de La Gomera para lo cuales se observa mayor duración. Sin embargo, en el modelo para los productos 3 y 4, ni el establecimiento ni el precio unitario afectan significativamente a la duración de sus precios. Son el año y el mes los factores que afectan a la duración, pero de una forma distinta a la de los productos 1 y 2. Hasta el año 2006 la duración de los precios tiene una tendencia creciente, para a partir del 2007 reducirse. En cuanto al efecto estacional, el mes de diciembre es el que mayor duración presenta, mientras que febrero, marzo y abril son los que se asocian con menor duración. Este resultado puede explicarse por la demanda esperada para dichos productos, al coincidir los mayores niveles de demanda con los meses en los que se obtiene una menor duración de los precios.

Es decir, al contrario que el nivel de precios, su duración no está significativamente afectada por la isla, ni por el tipo de establecimiento y sólo en algún caso y de forma muy poco significativa por el establecimiento. Una posible justificación a este resultado la podemos encontrar en la gran dependencia del exterior en el suministro de bienes, con una alta concentración en las importaciones. Sin embargo, la significatividad del factor estacional nos hace pensar en que también existen factores relacionados con los cambios de demanda. Por otra parte, el reducido nivel de ajuste de los modelos utilizados ponen de manifiesto la existencia de una alta volatilidad cuyo origen no puede ser explicado.

5.3. El tamaño del cambio

En este aparado nos centramos en el tamaño de los cambios de precios. Los resultados obtenidos hasta este momento ponen de manifiesto la poca estabilidad que presentan los precios en Canarias y la aparente inexistencia de rigideces a la baja. En la tabla 5 se resumen las características descriptivas de las reducciones y de los incrementos de precios, diferenciado por productos. En su conjunto, la reducción media es del -9,70% y el incremento medio del 11,06%, con una dispersión relativa del 86% para la reducción y del 118% para los incrementos. Además, el 50% de las reducciones más pequeñas son más pequeñas que -7,69 y el 50% de los incrementos más pequeños son menores que 7,69. Tanto las reducciones de precios como los incrementos son leptocúrticos, con un apuntalamiento mucho mayor en el caso de los incrementos, y mientras las reducciones son asimétricamente negativas, los incrementos de precios son asimétricas positivas. Este patrón se repite para cada uno de los productos con ligeras modificaciones en los valores numéricos.

[4] Se ha optado por el modelo binomial negativo al obtenerse un parámetro de precisión significativo.
[5] No se observan diferencias por el hecho de utilizar cualquiera de las dos variables duración obtenidas con dicho enfoque.

Tabla 5. Tamaño del cambio en los precios

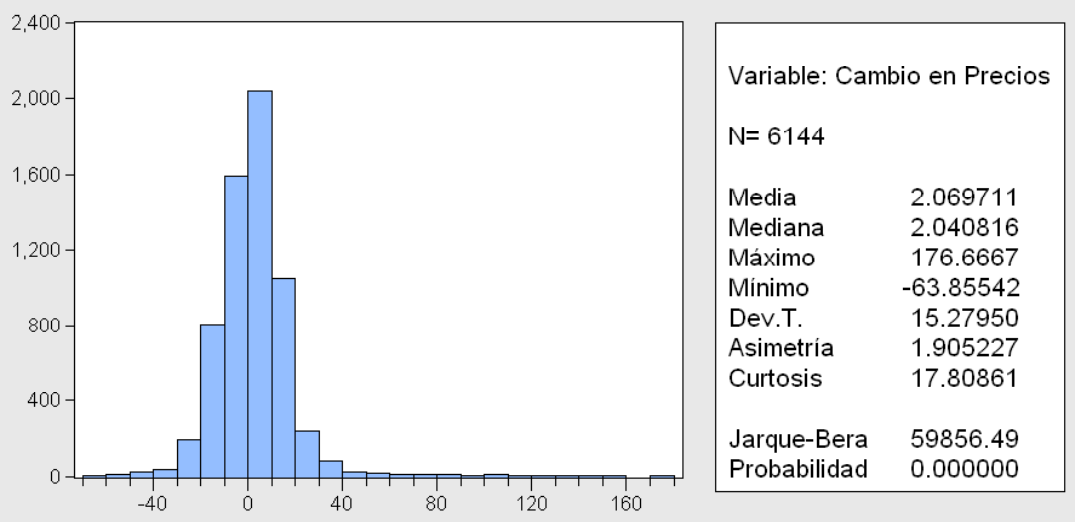
		N	Media	Coeficiente Variación	Mediana	Asimetría	Curtosis
Producto 1	Reducción	795	-10,81	-0,84	-8,70	-1,92	8,92
	Incremento	847	14,06	1,20	10,67	4,11	25,18
Producto 2	Reducción	559	-8,61	-1,23	-7,04	-2,18	12,46
	Incremento	903	10,15	0,71	7,22	5,79	74,16
Producto 3	Reducción	703	-10,33	-0,93	-7,58	-1,61	6,03
	Incremento	910	10,55	0,55	5,79	4,21	26,96
Producto 4	Reducción	603	-8,51	-0,72	-7,24	-1,08	4,25
	Incremento	824	9,51	0,85	7,81	2,37	14,52
Total	Reducción	2.660	-9,70	-0,86	-7,69	-1,89	8,75
	Incremento	3.484	11,06	1,18	7,69	4,72	37,64

Además, la tabla 5 muestra que el tamaño del cambio es importante, tanto para los incrementos como para los decrecimientos. En términos medios el ratio entre ambos tamaños se encuentra ligeramente por encima de la unidad, mientras que en términos de mediana es la unidad para el total, destacando el bajo valor del crecimiento mediano para el producto 3 con respecto a su crecimiento, siendo el único producto para el cual el crecimiento mediano es inferior a su decrecimiento.

La dispersión en los cambios es importante, pero dados los elevados niveles de apuntalamiento que presentan los cambios, sobre todo en los incrementos, junto con la distancia entre la media aritmética y la mediana, nos indica que existen valores muy extremos, tal y como se puede ver en la ilustración 2.

Las cifras de la tabla 5 son muy similares a las que obtiene Baudry et al. (2007) para el grupo de alimentación (incremento medio del 10,1% y decrecimiento medio del -9,7%) para Francia, sin embargo, los resultados que obtienen Álvarez y Hernando (2006) para España son ligeramente diferentes. En primer lugar, por el propio tamaño. Para el grupo de alimentos procesados los autores obtienen un tamaño medio del cambio en torno a ±8%. En segundo lugar, porque la cifra de los crecimientos es ligeramente inferior, en términos absolutos, a la de decrecimiento.

Ilustración 2. Distribución del cambio en precios



El tamaño de los cambios calculado es el que observa el consumidor, sin embargo en su cálculo se están tratando de forma igual a cifras diferentes. Es decir, no es lo mismo un

crecimiento del 10% después de llevar 10 meses con el mismo precio que si el precio sólo lleva en el mercado un mes. Baudry et al. (2005) calculan el cambio mensual medio como la media del cociente entre el tamaño del cambio y la duración del precio. La descripción de esta nueva variable se recoge en la tabla 6. Comparando los resultados de esta tabla con los de la tabla 5, se observa un mayor apuntalamiento, asimetría y dispersión relativa, frente a unos valores más pequeños en las medidas de posición central. En cambio medio mensual de precios es del -7,56% y del 8,19% para decrecimientos y crecimientos respectivamente, mientras que los valores medianos se reducen hasta el -4,76% y 4,46%. Estos valores son significativamente mayores a los que encuentran Baudry et al. (2005) para el grupo de alimentación con precios de consumo en Francia.

Por tanto, estas cifras vuelven a mostrar el comportamiento diferencial de los precios en Canarias al presentar unos cambios más grandes, tanto si se trata de una reducción de precios como si se trata de su incremento. En un mercado en donde el consumidor está perfectamente informado de la distribución de precios, esto le permitiría optimizar su gasto, pero dado que la información sobre pÿÿcios es muy poca y su búsqueda se obtiene con unos costes considerables, este comportamiento de los precios acentúa la sensación de inflación para los consumidores en Canarias e impide que éste pueda actual de forma eficiente.

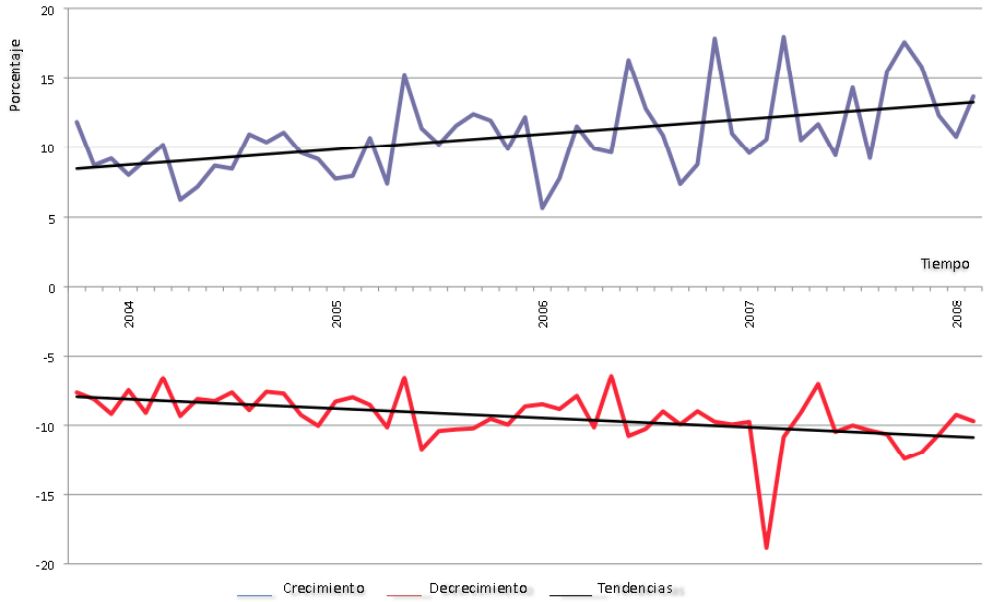
Tabla 6. Tamaño del cambio mensual en los precios

		N	Media	Coeficiente Variación	Mediana	Asimetría	Curtosis
Producto 1	Reducción	752	-8,69	-1,06	-5,42	-2,23	10,05
	Incremento	798	10,48	1,40	6,67	5,00	37,57
Producto 2	Reducción	532	-7,04	-1,04	-4,71	-2,56	14,59
	Incremento	838	6,93	1,44	4,00	7,17	104,42
Producto 3	Reducción	665	-7,86	-1,20	-3,85	-2,12	8,01
	Incremento	852	8,46	1,58	3,71	4,25	28,00
Producto 4	Reducción	578	-6,37	-0,93	-4,47	-1,50	5,52
	Incremento	779	6,90	1,10	3,92	2,63	16,81
Total	Reducción	2.527	-7,59	-1,09	-4,76	-2,32	10,77
	Incremento	3.267	8,19	1,45	4,46	5,31	48,88

Se han realizado contrastes Anova tanto para los cambios positivos como para los negativos y se concluye que no existe un efecto isla. Si existe un comportamiento diferencial en función del tipo de establecimiento, al 2,5% para los crecimientos de precios y al 1% para los decrecimientos, siendo los establecimientos pequeños los que mayores cambios presentan tanto para crecimientos como para decrecimientos. El efecto estacional también es significativo, produciéndose los mayores decrecimientos en el mes de febrero y, en menor medida, en junio, claramente relacionado con las rebajas, a pesar de que son productos de alimentación y que no deberían verse afectados por estas. Por el contrario, los mayores crecimientos tienen lugar en el mes de marzo y en los meses de septiembre, octubre y noviembre.

El tiempo es también un factor significativo en el tamaño de los cambios presentando tendencias contrarias para los crecimientos y para los decrecimientos. Esto es, cada vez se producen cambios de precios mayores en términos de crecimientos y más pequeños en términos de decrecimientos, con lo cual cada vez la brecha entre los crecimientos y decrecimientos medios es más amplia, tal y como se puede ver en la ilustración 3. En esta ilustración se representan en el eje de ordenadas los cambios medios, tanto positivos como negativos y en el eje de abscisas el tiempo en meses. El mismo análisis realizado por productos mantiene para todos ellos una pauta similar a la descrita para su conjunto, siendo únicamente de destacar la mayor intensidad del fenómeno para el producto 1.

Ilustración 3. Crecimientos y decrecimientos medios



Con el objeto de profundizar en esta divergencia entre los crecimientos y los decrecimientos se ha creado una variable precio tipificada y se ha repetido el análisis anterior, pero diferenciando el comportamiento para el primero, tercero y quinto quintil de la variable precio tipificada. La tabla 7 recoge el tamaño de las muestras con las que se trabaja en cada quintil y en la ilustración 4 se representa la evolución de los cambios medios en los precios a lo largo del tiempo.

La tabla 7 refleja un comportamiento esperado: a mayor precio de partida, los descrecimientos son más frecuentes y los crecimientos menos frecuentes; a menor precio de partida, los decrecimientos son menos frecuentes y los crecimientos más frecuentes. Sin embargo, existe una fuerte discrepancia entre la rigidez a la baja de los precios más bajos y la rigidez al alta de los precios más altos. Mientras que únicamente el 5% de los precios más bajos sigue bajando, el 14,9% de los precios más altos siguen subiendo.

Tabla 7. Cambio en precios por quintiles

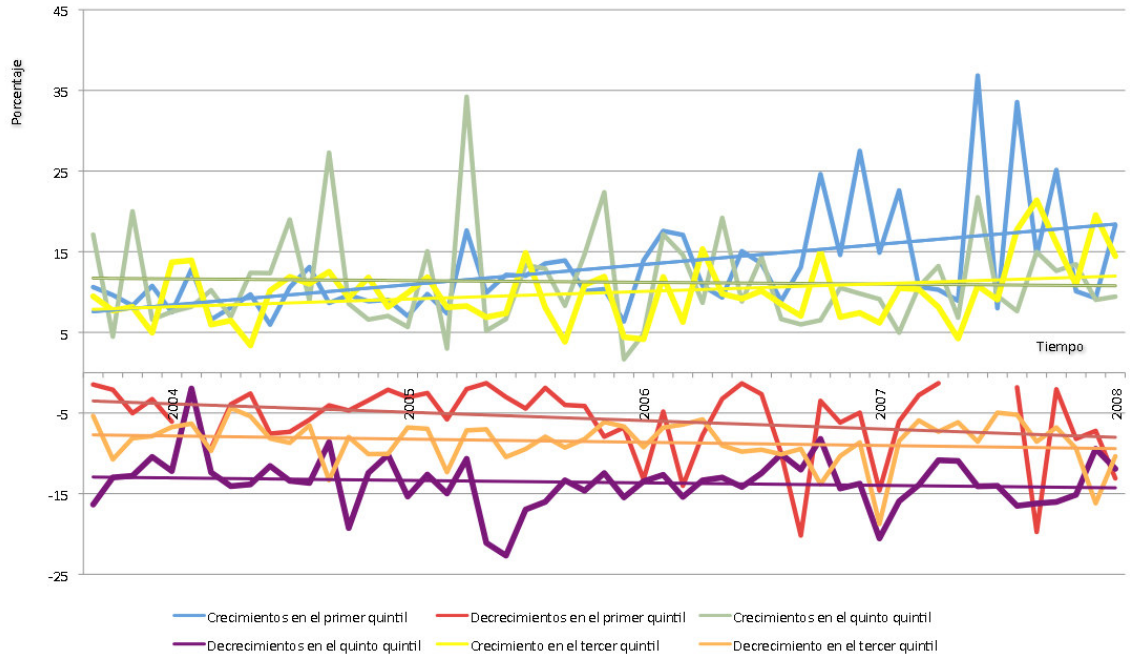
Quintil	Número de precios (Intervalo de precio)	Número de cambios de precios		Peso sobre el total del quintil
Primero	3.545 ($-\infty <$ Precio Unitario $\leq$ -0,845468)	Decrecimiento	172	5,0%
		Sin cambio	2.262	66,0%
		Crecimiento	993	29,0%
		Total	3.427	100%
Tercero	3.598 (−0,447341 < Precio Unitario $\leq$ -0,008373)	Decrecimiento	587	16,8%
		Sin cambio	2.321	66,5%
		Crecimiento	582	16,7%
		Total	3.490	100%
Quinto	3.549 (0,66790 < Precio Unitario $\leq \infty$)	Decrecimiento	808	24,8%
		Sin cambio	1.966	60,3%
		Crecimiento	489	14,9%
		Total	3.263	100%

Por otra parte, mientras que el 29% de los precios más bajos tienden a subir, solamente el 24,8% de los precios más altos tienden a bajar. Eso se traduce en que en los precios más bajos, por cada decrecimiento de precios se producen 5,6 subidas, mientras en los precios más altos por cada crecimiento de precios se produce únicamente 1,66 reducciones. Como se puede deducir de la tabla 7, la relación entre crecimientos y decrecimientos en el tercer quintil es 1. Si además se tiene en cuenta que el porcentaje de crecimientos es muy similar entre el tercer y

quinto quintil y que el peso de los que no cambian de precios se reduce en 6 puntos porcentuales entre ambos casos, ratifica la existencia de unos un grupo de establecimientos, que podemos cifrar en torno al 15%, que presentan precios altos con crecimientos, mientras que un porcentaje cercano al 6% tienden a reducir sus precios cuando observan que se sitúan entre los establecimientos de mayor peso.

La ilustración 4 nos permite analizar la evolución temporal del tamaño medio de los cambios diferenciando por quintiles (primero, tercero y quinto) y por decrecimientos y crecimientos. De su análisis, destacamos la homogeneidad de existe para el crecimiento en los tres quintiles representados, que únicamente se rompe para el primer quintil a partir de finales del año 2007. Este resultado tiene implicaciones muy negativas sobre la percepción del consumidor sobre el cambio de los precios, puesto que los establecimientos más baratos se encarecen relativamente más rápido que el resto, con lo cual el consumidor ve reducido su capacidad para buscar un establecimiento que le permita no asumir la oferta de inflación de su establecimiento habitual.

Ilustración 4. Crecimientos y decrecimientos medios en los quintiles primero, tercero y quinto



Por el contario, el comportamiento de los decrecimientos de precios sí muestran diferencias cuantitativas claras, siendo los establecimientos con precios más bajos los que presentan reducciones de precios más pequeñas y los más caros decrecimientos más fuertes. A pesar de ello, a partir del año 2006 se produce una mayor homogeneidad también en las reducciones de precios siendo menos relevante el quintil en el cual se encuentra el precio.

En consecuencia, podemos afirmar que el comportamiento que describimos sobre la base de la ilustración 3 se cumple también por quintiles, pero se debe destacar la homogeneidad en el tamaño de los crecimientos frente a la mayor heterogeneidad de los decrecimientos.

6. CONCLUSIONES

La frecuencia con la que cambian los precios en Canarias es mucho más elevada que la que se ha observado en cualquier otra región o economía.

La duración media de los precios medida mediante el enfoque de la duración se sitúa entre los valores medidos con el enfoque de la frecuencia discreto y continuo. El enfoque de la

duración permite la realización de distintos contrastes de hipótesis, los cuales nos permiten concluir que existen diferencias estadísticamente significativas por productos. Sin embargo, el dato más relevante es la presencia en los cuatro productos analizados de un porcentaje muy elevado de precios (siempre por encima del 50%) que presentan una duración no más allá de un mes.

La duración de los precios no está significativamente afectada por la isla, ni por el tipo de establecimiento y sólo en algún caso y de forma muy poco significativa por el establecimiento.

El tamaño de los cambios en precios, tanto de los incrementos como de los decrecimientos, se sitúa en torno al 10%, cifras similares a la de otros países europeos y ligeramente superiores a las de España.

Según va pasando el tiempo, en el mismo instante de tiempo conviven crecimientos medios más grandes con decrecimientos medios también más grandes.

Un importante número de establecimientos presentan precios altos y, a la vez, crecimiento de los mismos. Sin embargo, un 6% de establecimientos, cuando observan que se sitúan entre los establecimientos caros, tienden a bajar sus precios, mientras que aquellos establecimientos cuyos precios se sitúan en tono a la mediana de la distribución de precios bajan y suben los precios con la misma frecuencia.

El estudio realizado aporta evidencia de que todos los establecimientos incrementan sus precios en una cantidad muy similar y en una tasa bastante estable a lo largo del período analizado, con la única excepción de los precios más bajos, los cuales presentan una clara tendencia positiva. Es decir, los establecimientos más bajos tienden a incrementar sus precios cada vez en una cantidad mayor. Sin embargo, esta homogeneidad se convierte en heterogeneidad cuando hablamos de decrecimientos. Los establecimientos más caros tienen más margen para reducir sus precios y así lo hacen con unas reducciones medias significativamente superiores a la que presentan los establecimientos más baratos, aunque tanto unos como otros tienden a incrementar el tamaño medio de sus reducciones.

En definitiva, los establecimientos parece que juegan cada vez más con los cambios en precios en su búsqueda de una mayor cuota de mercado y los consumidores se encuentran con mayores dificultades para tomar sus decisiones de forma racional al encontrarse en un mar de precios cada vez más grande.

BIBLIOGRAFÍA

ÁLVAREZ, L. y HERNANDO, I. (2006). "Price setting behavior in Spain: Evidence from consumer price micro-data", *Economic Modelling*, Vol. 23, pp. 699-716.

BAHARAD, E. y EDEN, B. (2004). "Price rigidity and price dispersion: evidence from micro data", *Review of Economic Dynamics*, Vol. 7, pp. 613-641.

BAUDRY, L., BIHAN, H., SEVESTRE, P. Y TARRIEU, S. (2005). "La rigidité des prix en France: quelques enseignements des relevés de prix à la consommation", *Économie et Statistiques*, nº 386, pp. 37-57.

BAUDRY, L., BIHAN, H., SEVESTRE, P. Y TARRIEU, S. (2007). "What do thirteen million price records have to say about consumer price rigidity", *Oxford Bulletin of Economics and Statistics*, Vol. 69, nº 2, pp. 139-183.

BILS, M. and KLENOW, P. (2004). "Some evidence on the importance of sticky prices", *Journal of Political Economy*, Vol. 112, nº 5, pp. 947-985.

CALVO, G. (1983). "Staggered prices in a utility-maximizing framework". *Journal of Monetary Economics*, Vol. 12, pp. 383-398.

CARLTON, D. (1996). "The rigidity of prices". *American Economic Review*, Vol. 76, pp. 637-658.

CECCHETTI, S. (1986). "The frequency of price adjustment - A study of the newsstand prices of magazines", *Journal of Econometrics*, Vol. 31, pp. 255-274.

DHYNE, E., et al. (2006). "Price Changes in the Euro Area and the United States: Some Facts from Individual Consumer price data", *Journal of Economic Perspectives*, Vol. 20, nº 2, pp. 171-192.

EDEN, B. (2001). "Inflation and price adjustment: an analysis of micro data", *Review of Economic Dynamics*, Vol. 4, pp. 607-636.

FISHER, T. y Konieczny, J. (2000). "Synchronization of price changes by multiproduct firms: evidence from Canadian newspaper prices". *Economics Letters*, Vol. 68, pp. 271-277.

ISTAC (2008). Cesta de la Compra. Encuesta de Precios-Canarias. http:// www.gobiernodecanarias.org/istac.

KLENOW, P. y KRYVTOV, O. (2008). "State-dependent or time-dependent pricing: Does is matter for recent U.S. inflation?". *The Quarterly Journal of Economics*, Vol. CXXIII, nº 3, pp. 863-904.

LACH, S. and TSIDDON, D. (1992). "The behavior of prices and inflation: an empirical analysis of disaggregated price data". *Journal of Political Economy*, Vol. 100, nº 2, pp. 349-389.

SCHOPFLOCHER, T. and SULLIVAN, P. (2005). "The Relationship Between Skewness and Kurtosis of a Diffusing Scalar". *Boundary-Layer Meteorology*, Vol. 115, páginas: 341-358.

TAYLOR, J. (1999). "Staggered price and wage setting in macroeconomics". In *Handbook of Macroeconomics*, Vol. 1B, editado por Jhon B. Taylor y Michael Woodford. New York: Elservier.

Anexo 1. Modelo para la media de los precios (Media) en función del producto (PROD1 a PROD4), del tiempo (TIEMPO) y outliers impulso (añomes_producto).

Dependent Variable: Media
Method: Least Squares
Sample: 1 215
Included observations: 215
Newey-West HAC Standard Errors & Covariance (lag truncation=4)

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.542795	0.006923	78.40390	0.0000
PROD1	0.409534	0.011436	35.81190	0.0000
PROD2	0.154005	0.007780	19.79392	0.0000
PROD3	0.607477	0.013229	45.91880	0.0000
TIEMPO*PROD1	-0.005191	0.000645	-8.043904	0.0000
TIEMPO*PROD2	0.001902	0.000389	4.886618	0.0000
TIEMPO*PROD3	0.008684	0.000975	8.908661	0.0000
TIEMPO*PROD4	0.001783	0.000653	2.727944	0.0070
TIEMPO^2*PROD1	9.76E-05	1.10E-05	8.891631	0.0000
TIEMPO^2*PROD2	2.18E-05	7.70E-06	2.826533	0.0052
TIEMPO^2*PROD3	-0.000151	1.77E-05	-8.553803	0.0000
TIEMPO^2*PROD4	5.52E-07	1.14E-05	0.048618	0.9613
200711_PROD1	0.047784	0.004517	10.57761	0.0000
200712_PROD1	0.072644	0.004996	14.53974	0.0000
20081_PROD1	0.045637	0.005503	8.293664	0.0000
20082_PROD1	0.084613	0.006035	14.01971	0.0000
20079_PROD2	0.067616	0.003628	18.63502	0.0000
200710_PROD2	0.133903	0.003963	33.78742	0.0000
200711_PROD2	0.194984	0.004319	45.14182	0.0000
200712_PROD2	0.204962	0.004696	43.64436	0.0000
20081_PROD2	0.206181	0.005093	40.48516	0.0000
20082_PROD2	0.188301	0.005508	34.18428	0.0000
20072_PRDO3	-0.212563	0.008100	-26.24194	0.0000
20081_PROD3	0.048605	0.013630	3.565973	0.0005
20082_PROD3	0.037350	0.014445	2.585777	0.0105
20081_PROD4	0.025587	0.003905	6.552853	0.0000
20082_PROD4	0.023105	0.004412	5.237045	0.0000

R-squared	0.997055	Mean dependent var	0.882979
Adjusted R-squared	0.996648	S.D. dependent var	0.241284
S.E. of regression	0.013970	Akaike info criterion	-5.586825
Sum squared resid	0.036691	Schwarz criterion	-5.163536
Log likelihood	627.5837	Hannan-Quinn criter.	-5.415797
F-statistic	2448.036	Durbin-Watson stat	0.927293
Prob(F-statistic)	0.000000		

Anexo 2. Modelo para el Coeficiente de Variación (CV) en función del producto (PROD1 a PROD 4 y del tiempo (TIEMPO)

Dependent Variable: CV
Method: Least Squares
Sample: 1 215
Included observations: 215
Newey-West HAC Standard Errors & Covariance (lag truncation=4)

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.125941	0.005183	24.30000	0.0000
PROD1	0.047702	0.012265	3.889377	0.0001
PROD2	-0.055749	0.009115	-6.115988	0.0000
PROD3	0.007484	0.007202	1.039136	0.3000
TIEMPO*PROD1	0.000247	0.000973	0.254067	0.7997
TIEMPO*PROD2	0.003346	0.000543	6.158913	0.0000
TIEMPO*PROD3	-0.000271	0.000510	-0.531980	0.5953
TIEMPO*PROD4	0.000488	0.000623	0.783840	0.4340
TIEMPO^2*PROD1	1.41E-05	1.86E-05	0.758246	0.4492
TIEMPO^2*PROD2	-5.06E-05	9.16E-06	-5.523606	0.0000
TIEMPO^2*PROD3	9.58E-06	1.04E-05	0.920924	0.3582
TIEMPO^2*PROD4	-1.74E-05	1.08E-05	-1.617693	0.1073

R-squared	0.771756	Mean dependent var	0.140973
Adjusted R-squared	0.759388	S.D. dependent var	0.039051
S.E. of regression	0.019155	Akaike info criterion	-5.018270
Sum squared resid	0.074487	Schwarz criterion	-4.830142
Log likelihood	551.4641	Hannan-Quinn criter.	-4.942258
F-statistic	62.39987	Durbin-Watson stat	1.603225
Prob(F-statistic)	0.000000		

CAPITAL PÚBLICO Y PRODUCCIÓN INDUSTRIAL EN COLOMBIA (1990-2005)

SERGIO JIMÉNEZ RAMÍREZ
E-mail: sjimenez@unipamplona.edu.co
Universidad de Pamplona (Colombia)
JAIME SANAÚ VILLARROYA
E-mail: jsanau@unizar.es
Universidad de Zaragoza

Resumen

El objetivo de este trabajo es analizar el impacto del capital público sobre la producción de las industrias manufactureras en Colombia entre 1990 y 2005 (único periodo para el que hay información estadística de todas las variables requeridas), mediante el enfoque basado en la teoría de la dualidad. Concretamente, estudia los efectos que la inversión en capital público tiene sobre la estructura de costes de la industria, a través de la interrelación entre los diferentes factores de producción privados y aquél.

Palabras Clave: Capital público, crecimiento económico, teoría de la dualidad.
Área Temática: Economía del Crecimiento Económico.

Abstract

The aim of this paper is to analyze the impact of public capital on the production of manufacturing industries in Colombia between 1990 and 2005 (the only period for which there is statistical information for all required variables) using an approach based on the duality theory. This piece of research focuses on the effects that investment in public capital has on the industry cost structure, through the interplay between different production factors and public capital.

Keywords: Public capital, economic growth, duality theory.
Thematic Area: Growth Economics.

1. INTRODUCCIÓN

Durante los últimos decenios del siglo XX surgió entre políticos y economistas una creciente preocupación por averiguar el origen de la desaceleración del crecimiento de la productividad ocurrida en el decenio de 1970 en los Estados Unidos, tras mantener un elevado crecimiento en los diez años anteriores. Esta preocupación renovó el interés de los investigadores por los factores que dinamizan la productividad y el crecimiento económico dando paso a la publicación de algunos artículos seminales como fue el caso de Ratner (1983) y Aschauer (1989). En su trabajo publicado en 1989, aplicado a la economía de los Estados Unidos, David Aschauer presentó unos resultados que otorgaban al capital público el rol de "factor relevante" en el proceso de producción con una elasticidad del producto con respecto a esta variable de 0,39. A partir de este trabajo se desencadenó una copiosa literatura, con una notoria variedad de resultados.

Para agrupar esta literatura que contrasta la denominada hipótesis del capital público Jiménez (2009) distingue tres escenarios generales. El primero de ellos, llamado *La primera generación*, relaciona aquellos trabajos iniciales que indagaron empíricamente sobre el potencial vínculo entre inversión en infraestructuras y crecimiento de la productividad del sector privado por medio de una función de producción agregada ampliada con el capital público. Entre otros, se puede contar a Aschauer (1989), Munnell (1990a, 1990b), García-Milá y

McGuire (1992) para los Estados Unidos, y a Arguimón, González-Páramo, Martín y Roldán (1994), Mas, Maudos, Pérez y Uriel (1996), Sanaú (1998), Fernández y Polo (2002) y Álvarez, Orea y Fernández (2003) para España.

Con el transcurrir de los primeros años de la década de 1990 el análisis de la relación entre capital público y productividad cambió de enfoque metodológico como consecuencia de las fallas (o debilidades) del esquema de la función de producción agregada, dando paso a un nuevo escenario de análisis. Así, el enfoque dual basado en la función de costes comenzó a copar el interés de los investigadores para acercarse a la comprensión del vínculo entre esas dos variables económicas. Si bien, esta literatura ha resultado menos prolífica que la de funciones de producción.

Del interior de ésta nueva generación de trabajos surgió una nueva subdivisión dando origen a los otros dos escenarios que completan la revisión de la literatura: los trabajos que utilizan una *función de costes generalizada de Leontief* y los que utilizan una *función de costes translog*. Entre los primeros puede incluirse a Morrison y Schwartz (1996) para los Estados Unidos y Boscá *et al*. (1999) para España; y entre los segundos a Lynde y Richmond (1992) y Nadiri y Mamuneas (1994) para los Estados Unidos, Conrad y Seitz (1994) y Seitz y Licht (1995) para Alemania, Avilés, *et al.* (2001), y Moreno, *et al.* (2002) para España. No obstante, trabajos como el de Brox y Fader (2005) para Canadá, que utilizan un modelo de costes de elasticidad de sustitución constante-translog (CES-TL), representan a los que han optado por utilizar formas funcionales menos comunes.

El objetivo de este trabajo es analizar –por primera vez en la literatura aplicada- el impacto de las infraestructuras públicas sobre la productividad en las industrias manufactureras en Colombia entre 1990 y 2005 (único periodo para el que hay información estadística de todas las variables requeridas) mediante el la estimación de una función de costes translogarítmica. Se trata de estudiar los efectos que la inversión en capital público tiene sobre la estructura de costes de la industria, a través de la interrelación entre los diferentes factores de producción privados y aquel.

La relevancia de la aplicación a una economía como la colombiana descansa básicamente sobre dos pilares. El primer sustento proviene del hecho de realizar este análisis en un país en vías de desarrollo como es Colombia, para los que no se cuenta con estudios de este tipo, los cuales en su gran mayoría han privilegiado a países de la OCDE, y especialmente a España y a los Estados Unidos. Por tanto, la falta de evidencia empírica acerca de la importancia que tienen las infraestructuras públicas en el proceso de desarrollo económico de los países de renta media y baja, como es el caso de los latinoamericanos, también da apoyo a la realización de este trabajo. El segundo pilar que apoya este trabajo resulta, por una parte, de evaluar el efecto del capital público sobre la industria manufacturera en un país que durante el periodo de estudio fue objeto de reformas estructurales muy importantes a nivel político y económico. Y, por otra parte, de intentar comprender, a la luz de los resultados obtenidos, el proceso de desarrollo de la economía colombiana en ese periodo.

La estructura del trabajo es la siguiente. En el próximo apartado, se introduce el modelo teórico basado en la teoría de la dualidad, de donde se derivan las expresiones para las contribuciones marginales del capital público y del capital privado a la reducción de los costes variables, es decir, los precios sombra, así como para las elasticidades del *output* y de los costes con respecto a ambos capitales. Seguidamente se describen las variables y los datos usados y se introduce el modelo econométrico que se usará para llevar a cabo la contratación empírica. En el epígrafe final, se plasman las conclusiones más importantes a las que se llega a partir de los resultados obtenidos.

2. MODELO, ESTIMACIÓN Y RESULTADOS

2.1. Modelo

El enfoque de la teoría de la dualidad que utiliza una función de costes para representar los rendimientos de la inversión en infraestructuras proporciona una perspectiva un tanto diferente

de la acostumbrada con la estructura de la función de producción de la que numerosos trabajos dan cuenta. Según lo destacan Morrison y Schwartz (1996), una característica útil del enfoque de la función de costes es la representación que hace de la reacción conductual así como de las relaciones tecnológicas, asumiendo que la minimización de costes es un supuesto apropiado.

Otro aporte valioso de la estructura de la función de costes es que produce ecuaciones de demanda de factores con variables dependientes endógenas, en contraste con las ecuaciones de estimación derivadas al utilizar el enfoque de la función de producción (muy a menudo la función de producción en si misma), donde los niveles de los *inputs* son los argumentos de la función.

El punto de partida, siguiendo a Boscá *et al*. (1999) y Moreno *et al*. (2002), será una función de producción donde Y es el *output* y X_i (i = 1,…,s) es el factor i-ésimo:

$$Y = F(X_1,\ldots,X_s) \qquad (1)$$

Se supone que las empresas deben aceptar un vector de precios de los factores de producción, P_1,…, P_s, así que el problema de optimización radica en elegir la cantidad de factores que minimizan el coste de producir un nivel de producto dado, Y. Entonces, puede conseguirse un grupo de funciones de demanda para los factores privados:

$$X_i = X_i(P_1,\ldots,P_s,Y) \qquad (2)$$

donde X_i es la cantidad óptima del factor i-ésimo. En este caso, el nivel de costes óptimo (C) produce una función de costes que es *dual* a la función de producción, siendo dependiente de los precios de los factores y del producto:

$$C = C(P_1,\ldots,P_s,Y) \qquad (3)$$

Por lo tanto, se asume que todos los factores de producción pueden ajustarse dentro de un periodo de tiempo, de manera que la empresa determina instantáneamente las demandas de factores a largo plazo.

Existen razones que apuntan que ciertos factores no se ajustan instantáneamente a sus valores de equilibrio de largo plazo. Entre estas razones pueden contarse los controles de precios y regulaciones, los costes de inversión y desinversión, y restricciones institucionales que están más allá del control de una empresa individual en el corto plazo. De ahí que se distinga entre *inputs* que están en equilibrio, los llamados *inputs* variables, y aquellos que no lo están, los llamados *inputs* fijos. A esta situación se le conoce como *equilibrio estático parcial*. La estructura adoptada aquí distingue entre *inputs* variables e *inputs* fijos. El objetivo de la empresa es minimizar el coste de los factores variables, condicionado al *stock* dado de factores fijos.

Dado que uno de los fines del trabajo empírico es obtener elasticidades del capital público, el enfoque parte de una función de producción ampliada con el capital público como un factor no remunerado, lo que debe tenerse en cuenta al obtener la correspondiente función de costes. Por lo tanto, la función de costes variables utilizada incluye al capital público como un *input* fijo externo:

$$CV = CV(P_L, P_M, Y, K_P, K_G) \qquad (4)$$

en la que se consideran dos *inputs* privados variables, trabajo (L) y materiales intermedios (M), los cuales aparecen en la función de costes representados por sus precios, P_L y P_M, respectivamente; un factor cuasi fijo, capital privado, K_P; Y es el *output* y K_G el capital público, que actúa como factor externo. En consecuencia, las infraestructuras públicas son consideradas un factor fijo no remunerado en el proceso de producción y del que las empresas no tienen casi ningún control.

La función de costes totales de corto plazo será la suma de los costes variables y del coste del capital privado existente:

$$C = CV(\cdot) + P_{KP} \cdot K_P \qquad (5)$$

donde P_{KP} es el precio del capital privado.

El efecto de corto plazo de la inversión en infraestructuras sobre el proceso de producción consiste en que las empresas ajustan sus decisiones sobre las cantidades de los diferentes *inputs*

privados variables usados en el proceso de producción de acuerdo a sus relaciones de complementariedad o sustitución con las infraestructuras después de que éstas hayan sido aumentadas o mejoradas, dada la cantidad existente de factores fijos como el capital privado.

Diferenciando la función de costes variables, $CV(.)$, con respecto a K_G se obtiene el precio sombra, Z_{KG}, asociado con el capital público, que se define como:

$$Z_{KG} \equiv -\frac{\partial CV(\cdot)}{\partial K_G} \qquad (6a)$$

El mismo procedimiento puede seguirse en el caso del capital privado, K_P, y definir su precio sombra, Z_{KP}, como

$$Z_{KP} \equiv -\frac{\partial CV(\cdot)}{\partial K_P} \qquad (6b)$$

Estos precios sombra muestran los beneficios marginales de las empresas ocasionados por un incremento en el *stock* de capital tanto público como privado. Son una medida de la disponibilidad implícita en el corto plazo de los empresarios privados a pagar por capital público o privado. Concretamente, se definen los precios sombra como la reducción en los costes variables debido a un incremento marginal en los *stocks* de capital público o privado. Mientras el valor del precio sombra sea positivo las empresas se beneficiarán de contar con infraestructuras adicionales, ya que esto les permite lograr ahorros en los costes variables[1].

Suponiendo que los precios de los factores variables son exógenos al productor, puede aplicarse el lema de Shephard, y obtener el vector de los diferentes factores variables que minimiza los costes, es decir, las demandas minimizadoras de costes[2]:

$$X_i = X_i(P_L, P_M, Y, K_P, K_G) = \frac{\partial CV}{\partial P_i} \qquad i = L, M \qquad (7)$$

Las funciones de demanda condicionadas de factores que minimizan los costes pueden tomar la siguiente forma específica:

$$L(P_L, P_M, Y, K_G, K_P) = \frac{\partial CV(\cdot)}{\partial P_L} \qquad M(P_L, P_M, Y, K_G, K_P) = \frac{\partial CV(\cdot)}{\partial P_M} \qquad (8)$$

A partir de las demandas condicionadas de factores es posible reescribir la función de costes variables como:

$$CV(P_L, P_M, Y, K_G, K_P) = P_L L(\cdot) + P_M M(\cdot) \qquad (9)$$

La expresión (9) resulta útil para obtener las relaciones de complementariedad o sustitución entre el factor fijo considerado y cada uno de los factores variables. Haciendo uso de las definiciones en (6a) y (6b) se tiene que

$$Z_{KG} \equiv -\frac{\partial CV}{\partial K_G} = -P_L \frac{\partial L(\cdot)}{\partial K_G} - P_M \frac{\partial M(\cdot)}{\partial K_G} = L_{KG} + M_{KG}$$

$$Z_{KP} \equiv -\frac{\partial CV}{\partial K_P} = -P_L \frac{\partial L(\cdot)}{K_P} - P_M \frac{\partial M(\cdot)}{\partial K_P} = L_{KP} + M_{KP} \qquad (10)$$

[1] En este caso sólo es necesario que el precio sombra sea positivo puesto que la estructura plateada aquí considera que las empresas no pagan por el capital público, ya que se asume que es un factor exógeno. Sin embargo, aunque las empresas no perciban de manera directa los costes ocasionados por la acumulación de este factor, ellas pagan por las infraestructuras indirectamente vía impuestos. Como los impuestos no están directamente vinculados con los costes en los que incurre el gobierno para incrementar el *stock* de capital público, el precio de éste puede considerarse como cero para la empresa. Ésta perspectiva es la habitual en trabajos precedentes y, por lo tanto, la que se aplica en este estudio.

[2] Como es sabido, el lema de Shephard se usa para generar funciones de demanda de factores minimizadoras de costes. De este modo, se obtienen tantas ecuaciones adicionales a la función de costes como factores productivos intervengan en el proceso de producción. La estimación del sistema formado por la función de costes y las funciones de demanda derivadas por factores permite obtener estimaciones más eficientes de los parámetros que las que se obtendrían si se estimase sólo la función de costes, tal como señala McFadden (1978).

los cuales descomponen el efecto que un incremento en K_G (K_P) tiene sobre los costes, en los efectos de ajuste sobre el trabajo y sobre los consumos intermedios. Si L_{KG} (M_{KG}) es menor que cero, el trabajo (los consumos intermedios) es (son) un factor complementario del capital público. Si L_{KG} (M_{KG}) es mayor que cero, el trabajo (los consumos intermedios) es (son) un factor sustitutivo del capital público. Las mismas conclusiones son aplicables para el caso del capital privado.

Además, se puede definir cada participación del factor (S_i), esto es, el porcentaje del coste supuesto por el factor *i*-ésimo:

$$S_i = \frac{P_i \cdot X_i}{CV} = \frac{\partial \ln CV}{\partial \ln P_i} = \frac{\partial CV}{\partial P_i}\frac{P_i}{CV} \qquad i = L, M \qquad (11)$$

El conjunto de ecuaciones (4) y (11) constituye la solución a lo que puede ser definido como el equilibrio de corto plazo relacionado con los factores variables. También pueden usarse las funciones de demanda; alternativamente se hablaría del conjunto de ecuaciones (4) y (8).

Seguidamente, se van a definir algunas elasticidades de interés de la función de costes totales respecto al capital público y al capital privado. En primer lugar, dado el interés en evaluar el cambio en los costes totales de corto plazo ocasionado por un aumento marginal en el *stock* de infraestructuras, es necesario calcular la elasticidad de los costes de corto plazo respecto al capital público:

$$\varepsilon_{CK_G} = \frac{\partial \ln C}{\partial \ln K_G} = \frac{\partial C}{\partial K_G}\frac{K_G}{C} = \frac{\partial CV}{\partial K_G}\frac{K_G}{C} \qquad (12)$$

A partir de (12) y (6a, 6b) puede obtenerse la elasticidad de los costes variables respecto al capital público

$$Z_{K_G} \equiv -\frac{\partial CV}{\partial K_G} = -\varepsilon_{CVK_G}\left(\frac{CV}{K_G}\right) \qquad (13)$$

de donde se obtiene que

$$\varepsilon_{CVK_G} = \frac{\partial \ln CV}{\partial \ln K_G} = \frac{\partial CV}{\partial K_G}\frac{K_G}{CV} \qquad (14)$$

Dado que las empresas no pagan de manera directa por las infraestructuras, entonces puede afirmarse que $\varepsilon_{CK_G} = -Z_{K_G}\left(\frac{K_G}{C}\right)$, por lo que la única condición que debe satisfacerse para que la inversión en capital público genere un efecto positivo sobre la producción es que $Z_{K_G} > 0$. Si $Z_{K_G} > 0$, entonces $\varepsilon_{CK_G} < 0$. Esto ocurre en la medida en que el capital público tenga una relación de sustitución con los *inputs* variables; es decir, mientras las infraestructuras públicas incrementen la eficiencia como resultado de la disminución en la utilización de *inputs* variables y, con esto, de los costes variables.

Puede decirse que las empresas ajustarán sus decisiones de producción respecto a sus propios factores variables de acuerdo a la relación entre ellos y el capital público. Este efecto puede calcularse como la elasticidad (de corto plazo) de la demanda condicionada de *inputs* variables a las infraestructuras:

$$\varepsilon_{X_i K_G} = \frac{\partial \ln X_i}{\partial \ln K_G} = \frac{\partial X_i}{\partial K_G}\frac{K_G}{X_i} \quad i = L, M \qquad (15)$$

En segundo lugar, si el *stock* de capital privado no se encuentra en su nivel de equilibrio de largo plazo, los mismos efectos que se describieron para el capital público pueden obtenerse para el capital privado. Puede calcularse la elasticidad $\varepsilon_{X_i K_P}$, también como el precio sombra, Z_{K_P}, de la misma manera que para el capital público. En este caso, como las empresas pagan por el capital privado, la elasticidad coste incluye este efecto precio. Entonces:

$$\varepsilon_{CK_P} = \frac{\partial \ln C}{\partial \ln K_P} = \left(P_{K_P} - Z_{K_P}\right)\frac{K_P}{C} \qquad (16)$$

Cuando $K_P = K_P{}^*$, es porque $P_{K_P} = Z_{K_P}$, así $\varepsilon_{CK_P} = 0$. Sin embargo, fuera del equilibrio del estado estacionario, esto es, si las empresas no son capaces de ajustar K_P instantáneamente, entonces $\varepsilon_{CK_P} \neq 0$.

Se puede comprobar también, que algunas de las variables definidas a partir de la función de costes están muy relacionadas con las habituales medidas de elasticidades de la función de producción. Como se señala en Boscá *et al*. (1999), utilizando las expresiones derivadas anteriormente es posible relacionar las elasticidades producto con respecto a los *stocks* de capital con las participaciones sombra de éstos factores en el coste total. Las elasticidades del *output* respecto a los factores fijos se obtienen a partir de (6) y (11):

$$\varepsilon_{Y,KG} \equiv \frac{\partial \ln Y}{\partial \ln K_G} = \frac{\partial Y}{\partial K_G}\cdot\frac{K_G}{Y} = \frac{\partial Y}{\partial C}\cdot\frac{\partial C}{\partial K_G}\cdot\frac{K_G}{Y} =$$

$$= \frac{1}{CMa}\cdot Z_{KG}\cdot\frac{K_G}{Y} \equiv \frac{S^*_{KG}}{\varepsilon_{C,Y}} \qquad (17)$$

$$\varepsilon_{Y,KP} \equiv \frac{\partial \ln Y}{\partial \ln K_P} = \frac{\partial Y}{\partial K_P}\cdot\frac{K_P}{Y} = \frac{\partial Y}{\partial C}\cdot\frac{\partial C}{\partial K_P}\cdot\frac{K_P}{Y} =$$

$$= \frac{1}{CMa}\cdot Z_{KP}\cdot\frac{K_P}{Y} \equiv \frac{S^*_{KP}}{\varepsilon_{C,Y}} \qquad (18)$$

donde

$$\varepsilon_{C,Y} \equiv \frac{\partial \ln C}{\partial \ln Y} = \frac{\partial C}{\partial Y}\cdot\frac{Y}{C} = \frac{CMa}{C/Y} \qquad (19)$$

lo que muestra que el cociente entre el coste marginal y el coste medio determina la elasticidad de los costes al *output* a corto plazo, $\varepsilon_{C,Y}$, que está relacionada con la elasticidad de los costes variables al *output*, $\varepsilon_{CV,Y}$.

2.2 Contrastación empírica

Para la contrastación empírica se han usado datos anuales sobre precios y cantidades de los *inputs* y el *output* de los diferentes sectores de la industria manufacturera colombiana para el periodo 1990-2005 (único periodo para el que hay información de todas las variables requeridas), compilados desde varias fuentes.

Los datos sobre *output*, consumo de *inputs* intermedios, número de trabajadores y salarios de los sectores industriales se tomaron de la *Encuesta Anual Manufacturera*, *EAM*, producida y publicada por el DANE. Los datos de formación bruta de capital fijo así como los de inversión en infraestructuras con los cuales se construyeron el *stock* de capital privado y el *stock* de capital público se obtuvieron de las *Cuentas Nacionales* colombianas (DANE) y de la *Dirección de Infraestructura y Energía Sostenible* del Departamento Nacional de Planeación, DNP, respectivamente. Todos los datos en cantidades monetarias se expresan en precios constantes (pesos colombianos) de 1994, que es el año base vigente de las *Cuentas Nacionales* colombianas. Una explicación más detallada de la construcción de las variables puede consultarse en Jiménez (2009).

El trabajo empírico realizado en este apartado, dedicado a probar el efecto de las dotaciones de capital público sobre los costes de fabricación de la industria manufacturera privada, está

basado en una función de costes translogarítmica, una polinomial general de segundo grado en logaritmos, como la mostrada en la expresión (19) y muy similar a la desarrollada por Moreno *et al*. (2002), donde *t* es una tendencia temporal que resume el cambio tecnológico, como, por ejemplo, en Morrison y Schwartz (1996).

Esta forma funcional permite la consideración de un amplio rango de posibilidades de sustitución y puede contemplarse dentro de cualquier tecnología de producción sin la necesidad de imponer restricciones a priori sobre los rendimientos a escala.

La forma funcional translogarítmica, que se compone básicamente de dos ecuaciones, puede tomar la siguiente forma general incluyendo dos *inputs* variables (empleo, *L*, y consumos intermedios, *M*) y dos fijos (capital privado, K_P, y capital público, K_G):

- Función de costes variables:

$$\begin{aligned}\ln(CV/P_M) &= \beta_0 + \beta_L \ln\frac{P_L}{P_M} + \beta_Y \ln Y + \beta_{K_P} \ln K_P + \beta_{K_G} \ln K_G + \beta_T t \\ &+ \tfrac{1}{2}\left[\beta_{LL} \ln^2 \frac{P_L}{P_M} + \beta_{YY} \ln^2 Y + \beta_{K_PK_P} \ln^2 K_P + \beta_{K_GK_G} \ln^2 K_G + \beta_{TT} t^2\right] \\ &+ \beta_{LY} \ln\frac{P_L}{P_M} \ln Y + \beta_{LK_P} \ln\frac{P_L}{P_M} \ln K_P + \beta_{LK_G} \ln\frac{P_L}{P_M} \ln K_G + \beta_{LT} \ln\frac{P_L}{P_M} t \\ &+ \beta_{YK_P} \ln Y \ln K_P + \beta_{YK_G} \ln Y \ln K_G + \beta_{YT} \ln Y t + \beta_{K_PK_G} \ln K_P \ln K_G \\ &+ \beta_{K_PT} \ln K_P t + \beta_{K_GT} \ln K_G t \end{aligned} \quad (19)$$

P_L *y* P_M son los precios del factor trabajo y de los *inputs* intermedios, respectivamente; *Y* es el *output*. El precio de los consumos intermedios se incluyó como un factor relativo para asegurar el cumplimiento de la condición de que la función de costes es homogénea de grado uno en los precios de los factores.

Para obtener las ecuaciones de participación de los *inputs* variables en los costes variables en necesario aplicar el lema de Shephard. Dado que la suma de las participaciones de los factores es uno, sólo una ecuación es independiente, por lo que la participación de los consumos intermedios se obtiene a partir de la participación del trabajo:

- Ecuaciones de participación de los factores variables:

$$S_L = \frac{P_L \cdot L}{CV} = \frac{\partial \ln CV}{\partial \ln P_L} = \beta_L + \beta_{LL} \ln\frac{P_L}{P_M} + \beta_{LY} \ln Y + \beta_{LK_P} \ln K_P + \beta_{LK_G} \ln K_G + \beta_{LT} t$$

$$S_M = 1 - S_L \quad (20)$$

Así, el equilibrio de corto plazo sería el descrito por las ecuaciones (19) y (20).

Finalmente, es posible dar paso a la estimación de este sistema de dos ecuaciones, para el análisis de corto plazo, con el fin de obtener los parámetros relevantes de la función de costes, los que se usarán para calcular los precios sombra y las elasticidades que servirán para analizar los efectos de las infraestructuras y el capital privado. Esto se mostrará a continuación.

2.3. Resultados

Este modelo se incluyó dentro de un marco estocástico, al igual que el del apartado anterior, considerando que los términos de error añadidos a las ecuaciones de los modelos obedecen a errores de optimización. El modelo basado en la función de costes translogarítmica se estimó como un sistema de ecuaciones de regresión aparentemente no relacionados (*SURE*), que le añadió estructura y robustez, así como un aumento de eficiencia, evidenciado en valores de los

t–estadísticos más altos de los estimadores. Este método también permite imponer restricciones de igualdad entre parámetros a través de las ecuaciones para ajustarlas a los modelos teóricos. Por lo tanto, las Ecuaciones (19) y (20) se estimaron por el método *SURE*, utilizando el STATA 9.0.

En la Tabla 1 se presentan los resultados de la estimación en donde se muestran los parámetros y los *t*-estadísticos obtenidos. Aunque, nuevamente, el signo y la magnitud de los coeficientes no pueden interpretarse directamente, se observa que la mayoría de ellos son estadísticamente significativos a los niveles habituales.

Los valores de los parámetros se utilizaron para calcular los precios sombra y las elasticidades que reflejan los efectos de corto plazo del capital público y el capital privado sobre el desempeño económico de la industria manufacturera en Colombia. Estos indicadores se presentarán mediante tres promedios: sectoriales, temporales y globales.

Las Tablas 2 y 3 resumen los efectos de corto plazo más importantes de la inversión en capital público y capital privado evaluados mediante la forma funcional de costes translogarítmica. Las Tablas 2(a) y 2(b) reportan las magnitudes referidas a los valores sombra de los dos tipos de capital y las relaciones sustitutivas o complementarias existentes entre cada uno de los capitales y cada uno de los *inputs* variables, mientras que las Tablas 3(a) y 3(b) muestran los resultados referentes a las elasticidades del *output* con respecto a los capitales público y privado, así como las elasticidades de los costes con respecto a los mismos dos capitales. Todas estas magnitudes se calcularon para cada una de las doce agrupaciones industriales como los valores medios para el periodo 1990-2005 de la correspondiente magnitud sectorial

Tabla 1. Coeficientes estimados. Función de costes translogarítmica

Parámetro	Coeficiente	*t*-estadístico	Parámetro	Coeficiente	*t*-estadístico
β_L	0,000049	0,03	β_{LY}	-0,055818	-13,64
β_Y	-10,612800	-3,81	β_{LKP}	0,040955	10,57
β_{KP}	8,683992	3,20	β_{LKG}	0,030712	7,23
β_{KG}	26,865110	1,11	β_{LT}	-0,015165	-17,32
β_T	-2,957329	-1,20	β_{YKP}	-0,238457	-3,56
β_{LL}	-0,080911	-14,85	β_{YKG}	0,621945	3,65
β_{YY}	0,347599	5,07	β_{YT}	-0,064577	-3,71
β_{KPKP}	0,215252	3,25	β_{KPKG}	-0,515746	-3,12
β_{KGKG}	-1,761365	-1,18	β_{KPT}	0,050249	2,96
β_{TT}	-0,009700	-0,62	β_{KGT}	0,184284	1,21

Función de costes variables: $R^2 = 0,9975$
Función de participación del empleo: $R^2 = 0,7673$

Nota: Periodo muestral 1990-2005. Doce sectores de la industria. No. de observaciones 192.

También se calcularon para cada uno de los dieciséis años del periodo estudiado como los valores medios para el conjunto de los doce sectores de la industria.

En las dos primeras columnas de las Tablas 2(a) y 2(b) se muestran los precios sombra del capital público y del capital privado respectivamente. La Tabla 2(a) sintetiza los resultados de una forma que permite mostrar los promedios de los dieciséis años de la muestra para cada uno de sectores industriales, mientras que la Tabla 2(b) presenta los promedios de los doce sectores para cada uno de los años estudiados.

Como puede verse, el precio sombra del capital público, Z_{KG}, es positivo en todos los sectores industriales estudiados y para todos los años de la muestra, registrando un valor promedio de 0,0734. Esta disponibilidad implícita del sector privado industrial a pagar por unidades adicionales de capital público en el corto plazo en todas las industrias y durante todo el periodo de estudio implica una sustitución neta entre el capital público y los dos *inputs*

variables. El valor positivo del precio sombra puede interpretarse como que la inversión adicional de una (1) unidad monetaria colombiana en infraestructuras genera un ahorro en los costes variables de las empresas de aproximadamente 7,34 céntimos.

Tabla 2(a). Precios sombra y relaciones de sustituibilidad y complementariedad. Función de costes translogarítmica. Promedios sectoriales

	Z_{KG}	Z_{KP}	ε_{LKG}	ε_{MKG}	ε_{LKP}	ε_{MKP}
Alimentos	0,2320	0,2955	0,1292	-0,5124	-1,4343	-1,4695
Textiles	0,1098	0,1114	-0,0656	-0,3668	-1,8976	-1,9396
Madera	0,0140	0,1168	-0,0593	-0,3451	-2,0710	-2,1137
Papel	0,0763	0,0950	0,0260	-0,3471	-1,8947	-1,9339
Petroquímica	0,0725	0,1215	1,0850	-0,3337	-1,8385	-1,8718
Química	0,1052	0,1860	-0,0048	-0,4405	-1,6017	-1,6395
Caucho-plástico	0,0539	0,0960	0,0353	-0,3401	-1,9508	-1,9899
No-metálicos	0,0539	0,0220	0,0977	-0,2284	-2,2104	-2,2512
Metálicos	0,0794	0,0831	0,0772	-0,3250	-1,9718	-2,0107
Maquinaria	0,0162	0,2841	-0,1678	-0,4723	-1,7136	-1,7555
Electrónica	0,0282	0,0817	0,0091	-0,3145	-2,0451	-2,0861
Transporte	0,0393	0,5909	0,1174	-0,5350	-1,4636	-1,4988
Promedio global	**0,0734**	**0,1737**	**0,1066**	**-0,3801**	**-1,8411**	**-1,8800**

Fuente: Elaboración propia.

Nota: Z_{KG} es el precio sombra del capital público, Z_{KP} el precio sombra del capital privado, L_{KG} el efecto directo del capital público sobre el empleo, M_{KG} el efecto directo del capital público sobre los *inputs* intermedios, L_{KP} el efecto directo del capital privado sobre el empleo y M_{KP} el efecto directo del capital privado sobre los *inputs* intermedios.

El comportamiento de los valores sombra del capital público por industrias, por su parte, exhibe un patrón bastante irregular, aunque siempre con magnitudes positivas. Industrias tales como Alimentos, bebidas y tabaco (Sector 1), Textiles, prendas de vestir, cuero y calzado (Sector 2), Sustancias químicas y otros productos químicos (Sector 6) y Metales comunes y productos metálicos básicos (Sector 9), fueron las que sobresalieron, en este orden, por registrar las mayores contribuciones marginales del capital público a la reducción de los costes variables. Sin embargo, algunas industrias registraron valores de este precio sombra muy inferiores a la media como fue el caso de Productos de madera, corcho y accesorios derivados (Sector 3), Maquinaria de uso general (Sector 10) y Maquinaria y equipo eléctrico, electrónico y científico (Sector 11).

La considerable variación en el valor de los precios sombra, también experimentada bajo esta especificación de la función de costes, confirma nuevamente la necesidad de llevar a cabo este tipo de análisis a un nivel desagregado de la industria manufacturera, pues no todos los sectores se benefician en la misma medida de las inversiones en infraestructuras públicas.

Llevando el análisis a lo largo del periodo 1990-2005, puede destacarse el patrón ascendente del precio sombra del capital público. El valor de 2005 es el doble del registrado en 1990. Estos resultados implican que, conforme transcurre el tiempo, se hace más latente la necesidad de mayores inversiones en infraestructuras públicas, pues la disponibilidad a pagar de los industriales por éstas, así como el ahorro de costes variables que ocasionan crece con el paso de los años de forma importante.

Tabla 2(b). Precios sombra y relaciones de sustituibilidad y complementariedad.
Función de costes translogarítmica. Promedios temporales.

	Z_{KG}	Z_{KP}	ε_{LKG}	ε_{MKG}	ε_{LKP}	ε_{MKP}
1990	-0,0024	0,0152	0,5119	-0,0041	0,1792	-0,2078
1991	0,0121	0,0229	0,4740	-0,0224	0,0387	-0,3336
1992	0,0331	0,0463	0,3226	-0,0680	-0,2537	-0,5466
1993	0,0583	0,0787	0,2762	-0,1156	-0,5285	-0,8224
1994	0,0715	0,0910	0,2131	-0,1626	-0,8318	-1,1135
1995	0,0824	0,1095	0,1669	-0,2258	-1,0718	-1,3663
1996	0,0880	0,1226	0,1007	-0,2883	-1,3999	-1,6916
1997	0,0917	0,1517	0,0353	-0,3679	-1,7047	-2,0071
1998	0,0900	0,1585	-0,0470	-0,4246	-2,0472	-2,3305
1999	0,0818	0,1331	-0,0364	-0,4398	-2,2888	-2,5913
2000	0,0890	0,2080	-0,0235	-0,5459	-2,2576	-2,6494
2001	0,0908	0,2618	-0,0797	-0,6019	-2,3437	-2,7353
2002	0,0909	0,2891	-0,1079	-0,6475	-2,4672	-2,8719
2003	0,0961	0,3231	-0,0574	-0,6940	-2,4221	-2,8995
2004	0,0998	0,3795	-0,0282	-0,7301	-2,3938	-2,9201
2005	0,1010	0,3876	-0,0150	-0,7425	-2,4480	-2,9935
Promedio global	**0,0734**	**0,1737**	**0,1066**	**-0,3801**	**-1,5151**	**-1,8800**

Fuente: Elaboración propia.

Nota: Z_{KG} es el precio sombra del capital público, Z_{KP} el precio sombra del capital privado, L_{KG} el efecto directo del capital público sobre el empleo, M_{KG} el efecto directo del capital público sobre los *inputs* intermedios, L_{KP} el efecto directo del capital privado sobre el empleo y M_{KP} el efecto directo del capital privado sobre los *inputs* intermedios.

Por su parte, los valores del precio sombra del capital privado, Z_{KP}, reportan mayor variación entre sectores y a lo largo de los años que el del capital público. También su valor promedio global es bastante más alto que el del capital público, llegando a ser más del doble. Sin embargo, son muy similares en la tendencia creciente que presenta tanto el uno como el otro, si bien es más aguda la del precio sombra del capital privado.

En cuanto a los valores promedios de Z_{KP} por sectores, pueden se destacarse cuatro industrias por haber experimentado las mayores contribuciones marginales del capital privado a la reducción de los costes variables: Equipo de transporte (Sector 12), Alimentos, bebidas y tabaco (Sector 1), Maquinaria de uso general (Sector 10) y Sustancias químicas y otros productos químicos (Sector 6). Las tres primeras presentan cifras bastante elevadas, indicando que el *stock* de capital privado con que cuentan está muy por debajo de su nivel óptimo y que se beneficiarían en gran manera de nuevas inversiones en capital fijo. No obstante, en el otro extremo y registrando los precios sombra más bajos aparecen los sectores papel, cartón e

impresos; productos minerales no metálicos; metales comunes y productos metálicos básicos y maquinaria y equipo eléctrico, electrónico y científico.

La desagregación del valor de Z_{KP} entre las relaciones de sustitución y/o complementariedad entre los dos capitales y los dos factores variables permitirá vislumbrar de una mejor forma este comportamiento.

En este sentido, las Tablas 2(a) y 2(b) también muestran las relaciones que mantienen el capital público y el capital privado con cada uno de los *inputs* variables. El *stock* de capital privado mostró, en promedio, una relación de sustitución tanto con el factor trabajo como con los consumos intermedios. Las magnitudes de ambas elasticidades de sustitución fueron muy similares y mayores que la unidad (1), indicando que los dos factores variables fueron muy elásticos a los aumentos en el *stock* de capital privado durante el periodo estudiado.

Los sectores que se sobresalieron por registrar las mayores elasticidades de sustitución del trabajo con respecto al *stock* de capital privado fueron los mismos que presentaron las mayores elasticidades de sustitución de los *inputs* intermedios con respecto al *stock* de capital privado. Lo mismo ocurrió en el caso de los sectores que registraron los valores más bajos de las elasticidades. Las industrias con la mayor elasticidad de sustitución fueron Productos de madera, corcho y accesorios derivados (Sector 3), Productos minerales no metálicos (Sector 8) y Maquinaria y equipo eléctrico, electrónico y científico (Sector 11). Mientras que las de menores elasticidades fueron Alimentos, bebidas y tabaco (Sector 1), Sustancias químicas y otros productos químicos (Sector 6) y Equipo de transporte (Sector 12).

En cuanto a lo ocurrido entre el capital público y los dos factores variables, los resultados promedios muestran que el capital público sostuvo una relación de sustitución con los *inputs* intermedios en tanto que una relación de complementariedad con el factor trabajo. No obstante, cuatro sectores -textiles, prendas de vestir, cuero y calzado; productos de madera, corcho y accesorios derivados; sustancias químicas y otros productos químicos y maquinaria de uso general -presentaron signo contrario, evidenciando una relación sustitutiva entre capital público y empleo. En cambio, en todos los sectores la relación entre el capital público y los *inputs* intermedios fue complementaria. En general, la fuerza de la relación sustitutiva en favor de los consumos intermedios resultó bastante mayor (-0,3801), que la de la relación complementaria en favor del empleo (0,1066), lo que debe reflejarse en una relación sustitutiva neta entre el capital público y los factores variables o, un ahorro de costes variables, es decir, lo que mostrado por la magnitud y el signo del precio sombra del capital público en la primera columna de las tablas 2(a) y 2(b).

Trasladar los precios sombra a las elasticidades sombra de los costes con respecto a los capitales público y privado, así como trasladarlos a las elasticidades del *output* con respecto a esos dos *stocks* de capital, ofrece una visión aumentada de los efectos de las inversiones en activos fijos privados y públicos sobre la evolución de los diferentes sectores de la industria manufacturera. En la Tabla 3(b) se presentan medidas de estas dos diferentes elasticidades que servirán para complementar el análisis.

Las dos primeras columnas de la Tabla 3(a) reportan las elasticidades por sectores de la producción del *output* con respecto al capital público y al capital privado, respectivamente. La elasticidad promedio del *output* con respecto al capital público resulta mayor que uno (1) (1,2346), lo que supone una cifra bastante alta. En general, este resultado indica que la expansión del *output* requiere un incremento menos que proporcional en el *stock* de capital público, lo que le otorgaría un papel crucial en la determinación del crecimiento del *output* industrial y de la economía en su conjunto.

Tabla 3(a). Elasticidades del *output* y de los costes respecto a los dos capitales. Función de costes translogarítmica. Promedios sectoriales

	ε_{YKG}	ε_{YKP}	ε_{CKG}	ε_{CKP}
Alimentos	0,7548	0,2595	-0,7137	0,1511
Textiles	1,2185	0,1853	-0,7615	0,4291
Madera	1,5749	0,1880	-0,9137	0,3907
Papel	1,2536	0,1834	-0,7730	0,4140
Petroquímica	1,2719	0,1506	-0,8848	0,3538
Química	0,9365	0,2322	-0,7397	0,2723
Caucho-plástico	1,3323	0,1834	-0,8356	0,3958
No-metálicos	1,6847	0,0917	-0,5462	0,6793
Metálicos	1,3417	0,1668	-0,7661	0,4565
Maquinaria	1,1034	0,2766	-0,8963	0,1672
Electrónica	1,4964	0,1739	-0,7928	0,4616
Transporte	0,8462	0,2886	-0,9509	-0,0328
Promedio global	**1,2346**	**0,1983**	**-0,7979**	**0,3449**

Fuente: Elaboración propia.

Nota: ε_{YKG} es la elasticidad del *output* con respecto al capital público, ε_{YKP} la elasticidad del *output* respecto al capital privado, ε_{CKG} la elasticidad de los costes respecto al capital público y ε_{CKP} la elasticidad de los costes respecto al capital privado.

Por su parte, las variaciones entre sectores también están presentes en términos de ε_{YKG}. Los sectores que percibieron el mayor efecto productividad desde el *stock* de capital público, con valores por encima de la media, fueron "Productos de madera, corcho y accesorios derivados" (Sector 3), "Productos de caucho y de plástico" (Sector 7), "Productos minerales no metálicos" (Sector 8), "Metales comunes y productos metálicos básicos" (Sector 9) y, "Maquinaria y equipo eléctrico, electrónico y científico" (Sector 11).

En el caso de la elasticidad del *output* con respecto al capital privado, puede apreciarse que, en términos de su promedio global, su magnitud es positiva, pero muy inferior a la elasticidad del *output* con respecto al capital público. Sin embargo, el valor de dicha elasticidad de 0,198 demuestra un importante efecto productividad por parte del capital privado en la industria manufacturera colombiana. Además, como cabría esperar, las variaciones de ε_{YKP} entre industrias también son importantes, y puede destacarse las mayores elasticidades en los sectores de "Alimentos, bebidas y tabaco" (Sector 1), "Sustancias Químicas y otros productos químicos" (Sector 6), "Maquinaria de uso general" (Sector 10) y "Equipo de transporte" (Sector 12).

Si se observa en la Tabla 3(b), donde se presentan los resultados promedio de la industria manufacturera para cada uno de los 16 años estudiados, puede notarse que la elasticidad del *output* con respecto al capital privado sigue el mismo patrón que su precio sombra, es decir, aumenta de forma ininterrumpida a lo largo de todo el periodo. Por su parte, la elasticidad del *output* respecto al capital público también mantiene una tendencia creciente a lo largo de los años, aunque se ralentizó un poco entre 2002 y 2005. Estos hechos apuntan a que la industria manufacturera colombiana sufrió durante la década de 1990 y hasta mediados de la siguiente cierta descapitalización. En otras palabras, las cantidades de capital privado y de capital público

fueron inferiores a las que habrían sido necesarias para obtener mejores resultados en cuanto a la optimización de la producción agregada.

En términos generales, los resultados promedios indican que la productividad marginal de los dos tipos de capital es positiva, por lo que el crecimiento del *output* debería motivar incrementos adicionales en ambos *stocks* de capital, y más aún en infraestructuras públicas que, según este modelo basado en la función de costes translogarítmica, presenta un precio sombra y unas elasticidades mucho muy altas, si bien la elasticidad del *output* con respecto al capital público resultó mayor que la elasticidad del *output* con respecto al capital privado.

Las últimas dos columnas de las Tablas 3(a) y 3(b) muestran los valores obtenidos para las elasticidades de los costes con respecto al capital público y al capital privado, otro de los importantes resultados que se obtienen mediante la aplicación del enfoque dual.

En primer lugar, puede observarse que la elasticidad de los costes con respecto al capital privado, ε_{CKP} (donde influye el coste de uso de éste) presentó signo positivo, indicando un aumento de los costes totales de corto plazo ocasionado por un incremento del *stock* de capital privado de la industria manufacturera. Esto pudo presentarse porque la disminución en los costes variables causada por la inversión en capital privado (advertido en el valor positivo de su precio sombra) se contrarrestó por el incremento en los costes totales de corto plazo ocasionado por los pagos que deben hacer las empresas por unidades adicionales de este capital. Dicho comportamiento puede explicarse a la luz del alto valor del coste de uso del capital privado en todo el periodo de estudio, el cual fue muy superior al precio sombra de los doce sectores y para todos los años.

La principal causa del elevado coste de uso del capital privado fue, sin duda, una de las variables clave usada en su cálculo, como es el tipo de interés y su extraordinario nivel alcanzado durante toda la década de 1990, pero principalmente entre 1994 y 1998, tal como ya se comentó en su momento, a partir de los resultados obtenidos en el trabajo de Botero *et al*. (2007). Además, un ejercicio de simulación donde se tomaron en cuenta unos tipos de interés similares a los de algunos países europeos en esos años, reportó resultados más acordes con los valores esperados, es decir, los precios sombra del capital privado para todos los doce sectores llegaron a ser superiores al coste de uso del mismo (diferente a lo experimentado aquí) y, por supuesto, la elasticidad de los costes totales con respecto al capital privado también llegó a ser negativa.

En lo referente a la elasticidad de los costes con respecto al capital público, ésta registró un alto valor y su signo estuvo acorde a lo esperado (negativo en todos los sectores), destacando como ramas que experimentaron una mayor reducción en sus costes totales de corto plazo derivado de un aumento del *stock* de capital público productos de madera, corcho y accesorios derivados; petróleo refinado, combustibles y derivados; productos de caucho y de plástico; maquinaria de uso general y equipo de transporte.

En promedio, la industria manufacturera colombiana exhibió una ε_{CKG} de -0,7979, indicando que un aumento de 1% en el *stock* de infraestructuras públicas reduciría los costes totales de las empresas manufactureras en 0,8% aproximadamente. A primera vista parece un resultado algo elevado, si bien puede decirse, sin mucho temor a incurrir en equivocaciones, que el *stock* de capital público de Colombia, al igual que el de la mayoría de países en vías de desarrollo del mundo, se encuentra por debajo de su nivel óptimo, y que, por lo tanto, puede ser social y económicamente eficiente aumentarlo, especialmente en regiones en donde se concentren empresas pertenecientes a las agrupaciones industriales productos de madera, corcho y accesorios derivados; petróleo refinado, combustibles y derivados; productos de caucho y de plástico; maquinaria de uso general y equipo de transporte.

Tabla 3(b). Elasticidades del *output* y de los costes respecto a los dos capitales.
Función de costes translogarítmica. Promedios temporales.

	ε_{YKG}	ε_{YKP}	ε_{CKG}	ε_{CKP}
1990	0,1872	-0,0656	-0,1042	0,2665
1991	0,3104	-0,0439	-0,1822	0,3030
1992	0,4934	0,0073	-0,3393	0,2710
1993	0,6893	0,0521	-0,4931	0,2892
1994	0,8672	0,0924	-0,5760	0,3694
1995	0,9911	0,1404	-0,6197	0,4241
1996	1,1402	0,1811	-0,6860	0,4637
1997	1,2562	0,2253	-0,8108	0,4343
1998	1,3736	0,2539	-0,7122	0,5529
1999	1,4768	0,2626	-0,8097	0,5433
2000	1,4477	0,3067	-1,0977	0,3526
2001	1,4556	0,3267	-1,1109	0,3376
2002	1,4863	0,3425	-1,2742	0,2599
2003	1,4737	0,3567	-1,3148	0,2239
2004	1,4614	0,3663	-1,3002	0,2177
2005	1,4753	0,3690	-1,3348	0,2090
Promedio global	**1,2346**	**0,1983**	**-0,7979**	**0,3449**

Fuente: Elaboración propia.

Nota: ε_{YKG} es la elasticidad del *output* con respecto al capital público, ε_{YKP} la elasticidad del *output* respecto al capital privado, ε_{CKG} la elasticidad de los costes respecto al capital público y ε_{CKP} la elasticidad de los costes respecto al capital privado.

Recapitulando, el modelo evaluado bajo la estructura de la función de costes translogarítmica, le otorgó a la inversión en capital público, en términos generales, un efecto productividad (ε_{YKG}) y un efecto ahorro de costes (ε_{CKG}) mayores que los obtenidos para la inversión en capital privado. También puede afirmarse que, en términos globales, el capital privado actuó como sustitutivo tanto del trabajo como de los *inputs* intermedios, mientras que el capital público fue sustitutivo de los consumos de materiales intermedios, pero fue complementario del factor trabajo. Sin embargo, todas estas medidas presentaron una considerable variación entre agrupaciones industriales, destacando como sectores que percibieron mayor impacto positivo del capital público productos de madera, corcho y accesorios derivados; petróleo refinado, combustibles y derivados; productos de caucho y de plástico; maquinaria de uso general y equipo de transporte. Por su parte, los que más se beneficiaron de las inversiones en capital privado fueron alimentos, bebidas y tabaco; sustancias químicas y otros productos químicos; maquinaria de uso general y equipo de transporte.

Comparando algunos de los resultados de este trabajo con los obtenidos, por ejemplo, por Moreno *et al*. (2002) y Avilés *et al*. (2001) aplicados a la industria española en el periodo 1980-1991, los cuales también emplearon para la estimación una función de costes translogarítmica muy similar a la estimada en este apartado, pueden resaltarse algunos aspectos. En cuanto al capital público, este se mostró como un sustituto de los *inputs* intermedios en Moreno *et al*. Avilés *et al*. y este trabajo. Por su parte, la relación entre este mismo tipo de capital y el empleo fue complementaria en Moreno *et al*. y este trabajo, pero sustitutiva en Avilés *et al*. Asimismo, el capital privado mantuvo una relación sustitutiva con el consumo de *inputs* intermedios en los tres trabajos mencionados, mientras que con el factor trabajo la relación fue complementaria en

Moreno *et al*., así como lo mostrado en esta apartado para el caso colombiano, pero sustitutiva en Avilés *et al*.

En términos generales, tomando en cuenta las diferencias entre los dos países, los resultados hallados en el presente trabajo con aplicación a la manufactura colombiana son muy similares a los obtenidos para la manufactura española por Moreno *et al*. (2002) en el periodo 1980-1991.

IV. CONCLUSIONES

La literatura empírica existente sobre la relación entre las infraestructuras públicas y el aumento en la productividad de la industria manufacturera se ha centrado principalmente en dos enfoques metodológicos: el uso de funciones de producción y el uso de funciones duales de costes. Las desventajas más importantes evidenciadas con el uso de funciones de producción radican en la utilización de supuestos muy restrictivos como pueden ser la imposición de la tecnología (por lo general Cobb-Douglas), así como la no consideración de los precios de los *inputs* privados, los cuáles pueden afectar la intensidad de su uso. Asimismo, puede contarse entre ellas la imposición de rendimientos constantes, el ajuste instantáneo, y por ende, la no distinción entre corto y largo plazo, entre otras.

El enfoque dual basado en la función de costes, en cambio, aproxima de forma más integral los determinantes que actúan sobre el comportamiento de la empresa optimizadora. Esta metodología también permite examinar las relaciones de complementariedad o sustitución entre los factores privados y públicos, así como el efecto marginal del capital público sobre la estructura de costes de las empresas. Dentro de esta corriente metodológica también se pueden identificar dos vertientes predominantes: los trabajos que emplean una función de costes generalizada de Leontief y los trabajos que emplean una función de costes translogarítmica.

El modelo evaluado bajo la estructura de la función de costes translogarítmica concedió a la inversión en capital público, en promedio, unos efectos productividad y de ahorro de costes mayores que los obtenidos para la inversión en capital privado. No obstante, las magnitudes positivas de los precios sombra fueron mayores para este último. También puede notarse que el capital privado se presentó como sustitutivo tanto del factor trabajo como de los *inputs* intermedios, mientras que el capital público se comportó como sustitutivo de los consumos de materiales intermedios y complementario del trabajo.

La variabilidad entre industrias también fue notoria dentro de la estimación de la función de costes translogarítmica. En lo referente a las contribuciones marginales a la reducción de los costes variables (precios sombra), pueden destacarse los sectores 1 (alimentos, bebidas y tabaco), 2 (textiles, prendas de vestir, cuero y calzado), 6 (sustancias químicas y otros productos químicos) y 9 (metales comunes y productos metálicos básicos), como los que registraron el mayor impacto positivo de las inversiones en capital público, y los sectores 1 (alimentos, bebidas y tabaco), 6 (sustancias químicas y otros productos químicos), 10 (maquinaria de uso general) y 12 (equipo de transporte) como los que más se beneficiaron de las inversiones en capital privado.

Los resultados de cada una de las estructuras consideradas coincidieron en otorgar a los sectores 1 (alimentos, bebidas y tabaco), 2 (textiles, prendas de vestir, cuero y calzado), 6 (sustancias químicas y otros productos químicos) y 9 (metales comunes y productos metálicos básicos) las mayores contribuciones marginales del capital público a las reducciones de los costes variables, mientras que en lo referente a la inversión en capital privado las dos funciones de costes no estuvieron de acuerdo en registrar los mayores efectos en los mismos sectores.

En cuanto a la evolución en el tiempo, bajo la función translogarítmica los efectos positivos del aumento de este tipo de capital sobre la industria manufacturera aumentaron en magnitud a lo largo de todo el periodo. En cambio, los efectos positivos marginales de aumentar el stock de capital público aumentaron año tras año durante el periodo analizado.

De la comparación entre algunos de los resultados de este trabajo y los obtenidos, por ejemplo, por Moreno *et al*. (2002) y Avilés *et al*. (2001) aplicados a la industria española en el periodo 1980-1991, empleando en la estimación una función de costes translogarítmica muy similar a la estimada en este trabajo, pueden resaltarse algunos aspectos. En cuanto al capital

público, este se mostró como un sustituto de los *inputs* intermedios en Moreno *et al*., Avilés *et al*. y este trabajo. Por su parte, la relación entre este mismo tipo de capital y el empleo fue complementaria en Moreno *et al*. y esta aplicación a Colombia, pero sustitutiva en Avilés *et al*. Asimismo, el capital privado mantuvo una relación sustitutiva con el consumo de *inputs* intermedios en los tres trabajos mencionados, mientras que con el factor trabajo la relación fue complementaria en Moreno *et al*., así como lo mostrado aquí para el caso colombiano, pero sustitutiva en Avilés *et al*.

En términos generales, tomando en cuenta las diferencias entre las industrias de manufacturas de los dos países, los resultados hallados en el presente trabajo con aplicación a la manufactura colombiana son muy similares a los obtenidos para la manufactura española por Moreno *et al*. (2002) en el periodo 1980-1991.

Desde el punto de vista de los efectos del capital privado, los resultados obtenidos en las estimaciones de la función de costes translogarítmica mostraron que el aumento de la inversión en maquinaria, equipo y edificios generó reducciones marginales en los costes variables de la industria manufacturera de Colombia, permitiendo aumentos en la productividad de los diferentes sectores de la manufactura que, a la postre, se reflejaron en un mayor crecimiento del PIB industrial y del PIB total.

Por lo tanto, dado que los resultados de la estimación establecieron que, la relación entre el capital privado y el factor trabajo fue complementaria, mientras que entre éste y los *inputs* intermedios fue sustitutiva, por una parte, pero también, si se tiene en cuenta que la tendencia de la magnitud del precio sombra del capital privado observada a lo largo del periodo fue creciente, puede concluirse que hubo una notable escasez de inversión en capital privado que, al igual que en el caso del capital público, redujo el potencial de crecimiento y generación de empleo de la industria y la economía colombiana desde mediados de los años noventa y pudo haber demorado la recuperación tras la crisis.

Aparte de la inseguridad causada por los fenómenos de la guerrilla, los paramilitares y los cárteles de las drogas, el factor que determinó aquella aversión por parte de las empresas a invertir al ritmo que hubiera sido óptimo, fue el alto nivel de los tipos de interés que soportó la economía colombiana en la década de 1990. Sin duda, ésta pudo ser una de las principales fuerzas que llevaron a la desaceleración y posterior recesión de la economía colombiana experimentada entre finales de 1998 y comienzos de 2000.

En un estudio sobre la relación entre el coste de uso del capital privado y la inversión en Colombia para el periodo comprendido entre 1990 y 2007, Botero *et al*. (2007), indagando acerca de las causas más importantes del acusado descenso del índice del coste de uso del capital en el periodo comprendido entre 2001 y 2007 establecieron que dicha caída se asoció principalmente con tres factores: el paulatino descenso de los tipos de interés real, la estabilización y posterior caída de los precios de los bienes de capital (incluyendo el efecto revaluación) y, el factor tributario (disminución del impuesto de renta y deducción inmediata del 40% de éste causado por la inversión en activos fijos).

Sin embargo, estos autores quisieron profundizar más y determinar la importancia relativa que tuvo cada uno de estos tres componentes. En tal sentido, realizaron un ejercicio de simulación en el que encontraron que el efecto del tipo de interés real fue el más importante de los tres factores en la explicación del descenso del coste de uso del capital privado, puesto que explicaba casi en su totalidad la disminución de éste y, por lo tanto, que la caída de los tipos de interés es el principal determinante del aumento de la inversión en activos fijos acaecida en el país en el periodo por ellos estudiado.

En este sentido, si el Banco de la República hubiera podido intervenir a tiempo, y con los instrumentos precisos, para cambiar el sentido de la política monetaria restrictiva que aplicaba entonces, tal como lo exigía la situación de desaceleración económica que vivía Colombia durante 1997 y 1998, *ceteris paribus*, el desenlace podría haber sido otro. Pero la culpa no fue sólo de la rígida política monetaria. También faltó gestión por parte del gobierno para facilitar el acceso a un crédito más barato. No ha de olvidarse que, tanto el gobierno como el Banco de la República dedicaron toda su atención a mantener el nivel del tipo de cambio y a reducir la tasa de inflación, dejando en un segundo plano la recesión.

Si ante los signos de desaceleración se hubiera podido aplicar una política monetaria expansiva, se habría posibilitado un mayor nivel de inversión privada que, dados los resultados, habría aumentado el crecimiento de la productividad y, puesto que la relación observada entre el capital privado y el empleo fue complementaria, se habrían generado los empleos necesarios para disminuir la recesión o quizá evitarla.

Si esta medida de política se hubiera acompañado de otra que fomentara la inversión pública en infraestructuras por parte del gobierno (o la participación de inversionistas privados en la ampliación del stock de capital público), la unión de los efectos positivos de los capitales público y privado sobre la productividad y el empleo de la industria, sin duda, habrían evitado, *ceteris paribus*, la involución en el proceso de desarrollo económico por el que avanzaba Colombia entre finales del siglo pasado y comienzos de éste.

En definitiva, puede concluirse, que mayores niveles de inversión en capital público y en capital privado fueron requeridos por la industria manufacturera colombiana para optimizar la producción entre 1990 y 2005. Sin embargo, dada la tendencia creciente de las contribuciones marginales de los capitales público y privado a la reducción de los costes durante estos años, debe entenderse que la escasez de estos dos tipos de activos persiste en la actualidad y que las anteriores recomendaciones tienen plena vigencia de cara a la actual crisis de la mayoría de las economías del mundo, incluida la colombiana.

Finalmente, cabe señalar que ésta es una primera investigación acerca de la relación entre infraestructuras públicas y productividad de la industria manufacturera en países latinoamericanos, que irá seguida de otros trabajos para países como Brasil y México para los que, hasta el momento, se ha logrado recopilar gran parte de la información estadística necesaria.

Asimismo, se espera que los problemas de escasez de datos puedan resolverse, lo que en el futuro permitirá mejorar la investigación económica y, en particular, las estimaciones de los stocks de capital que aquí se presentan y, acometer la tarea pendiente de aplicar la nueva metodología de la OCDE a la estimación de los servicios del capital de algunos de los países latinoamericanos.

REFERENCIAS BIBLIOGRÁFICAS

ÁLVAREZ, A., L. OREA y J. FERNÁNDEZ (2003): "La productividad de las infraestructuras en España", *Papeles de Economía Española*, No. 95, págs. 125-136.

ARGUIMÓN, I., J.M. GONZÁLEZ-PÁRAMO, M.J. MARTÍN y J.M. ROLDÁN (1994): "Productividad e infraestructuras en la economía española", *Moneda y Crédito*, 198, págs. 207-252.

ASCHAUER, D.A. (1989): "Is Public Expenditure Productive?" *Journal of Monetary Economics*, 23, págs. 177-200.

AVILÉS, C.A., R. GÓMEZ y J. SÁNCHEZ (2001): "The Effects of Public Infrastructure on Cost Structure of Spanish Industries". *Spanish Economic Review*, 3, págs. 131-150.

BOSCÁ, J. E., J. ESCRIBÁ y T. DABÁN (1999): "Capital privado e infraestructuras en la producción industrial regional". *Revista de Economía Aplicada*, Vol. VII, No. 21.

BOTERO, J.A., A. RAMÍREZ, J.F. PALACIO (2007): "El costo de uso del capital y la inversión en Colombia 1990-2007". *Working Papers de Economía EAFIT*, No. 1. Universidad EAFIT, Medellín, Colombia.

BROX, J.A. y C.A. FADER (2005): "Infrastructure Investment and Canadian Manufacturing Productivity", *Applied Economics*, 37, págs. 1247-1256.

CONRAD, K. y H. SEITZ (1994): "The Economic Benefits of Public Infrastructure". *Applied Economics*, 26, págs. 303-311.

DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA, DANE, (Varios años): "Encuesta Anual Manufacturera". Carrera 59 No. 26-70 Interior 1 – CAN, Bogotá, Colombia.

DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA, DANE, (Varios años): "Cuentas Nacionales Anuales". Carrera 59 No. 26-70 Interior 1 – CAN, Bogotá, Colombia.

DEPARTAMENTO NACIONAL DE PLANEACIÓN, DNP. "Estadísticas de Inversión en Infraestructuras". Dirección de Infraestructura y Energía Sostenible, en: www.dnp.gov.co.

FERNÁNDEZ, M. y C. POLO, (2002): "Productividad del capital público en presencia de capital tecnológico y humano", *Revista de Economía Aplicada*, Vol. X, No. 29, págs. 151-161.

GARCÍA-MILÁ, T. y T. MCGUIRE (1992): "The Contribution of Publicly Provided *Inputs* to States' Economies". *Regional Science and Urban economics*, 22, págs. 229-241.

JÍMENEZ RAMÍREZ, S. (2009): I*nfraestructuras públicas y productividad en la industria manufacturera de Colombia,* Memoria de Tesis doctoral, Zaragoza: Universidad de Zaragoza.

LYNDE, C. y J. RICHMOND (1992): "The Role of Public Capital in Production", *The Review of Economics and Statistics*, Vol. 74 (1), pp. 37-44.

McFADDEN, D. (1978): "Production Economics: a dual approach to theory dual applications". *Amsterdam: North-Holand Publishy Company*.

MAS, M., J. MAUDOS, F. PÉREZ y E. URIEL. (1996): "Infrastructures and Productivity in the Spanish Regions", *Regional Studies*, Vol. 30 (7), págs. 641–649.

MAS, M., F. PÉREZ y E. URIEL. (2005): "El *Stock* y los servicios del capital en España, (1964-2002), Nueva Metodología". *Fundación BBVA*, Bilbao.

MORRISON, C.J. y A.E. SCHWARTZ (1996): "State Infrastructure and Productive Performance". *The American Economic Review*, 86 (5), págs. 1095-1111.

MORENO, R., E. LÓPEZ-BAZO y M. ARTÍS (2002): "Public Infrastructure and the Performance of Manufacturing Industries: short and long-run effects". *Regional Science and Urban Economics*, 32:1, págs. 97-121.

MUNNELL, A. (1990a): "Why Has Productivity Declined?" Productivity and Public Investment *New England Economic Review*, Federal Reserve Bank of Boston, January/February, págs. 3-22.

MUNNELL, A. (1990b): "How Does Public Infrastructure Affect Regional Economic Performance?" *New England Economic Review*, Federal Reserve Bank of Boston, September/October, págs. 11-32.

NADIRI, I. y T. MAMUNEAS (1994): "The Effects of Public Infrastructure and R&D capital on the Cost Structure and Performance of US Manufacturing Industries". *The Review of Economics and Statistics*, LXXVI (1), 189-198.

OCDE (2001a): "Measuring Productivity – OECD Manual: Measurement of Aggregate and Industry-level Productivity Growth". Paris: Organization for Economic Co-operation and Development.

OCDE (2001b): "Measuring Capital–OECD Manual: Measurement of Capital *Stocks*, Comsumption of Fixed Capital and Capital Services". Paris: Organization for Economic Co-operation and Development.

RATNER, J. (1983): "Government Capital and the Production Function of US Private *Output*". *Economics Letters*, 13: 213-217.

SANAÚ, J. (1998): "Telecommunications Infrastructure and Economic Growth: An Analysis of Spanish Manufacturing Industry". En: S. Macdonald y G. Madden (Eds.): "*Telecommunications and Socio-Economic Development*". Elsevier Science B. V., págs. 39-54.

SEITZ, H. y G. LICHT (1995): "The Impact of Public Infrastructure Capital on Regional Manufacturing Production Cost". *Regional Studies*, 29:3, págs. 231-240.

FOREIGN DIRECT INVESTMENT, ECONOMIC GROWTH AND THE ABSORPTIVE CAPACITY OF HOST ECONOMIES

GONÇALO ALVES
E-mail: gfalves@ua.pt
VERA CATARINA ROCHA
E-mail: verarocha@ua.pt
Department of Economics, Management and Industrial Engineering
University of Aveiro

Abstract

This paper bears evidence to the widespread belief that Foreign Direct Investment (FDI) impact on economic growth is enhanced by host economy absorptive capacity. The study is conducted for OECD countries in the period 1996-2006. Our panel data analysis reveals that FDI has a positive and significant effect on economic growth. However, the quality of the host environment -in human capital and technological competencies - affects significantly a country's ability to benefit from FDI. An active policy towards FDI implies therefore, at the same time, to support human capital development, learning and investment by local firms, as a way to attract high quality FDI and to enhance the potential benefits arising from foreign presence. This paper suggests that the use of investment policies focusing exclusively on foreign firms, although motivated in some cases from a theoretical point of view, is generally not an efficient way to raise national welfare.

Keywords: Foreign Direct Investment (FDI), Economic Growth, Human Capital, Technological Capacity, Absorptive Capability
Thematic Area: Economic Growth

1. INTRODUCTION

Developing countries have been the main focus of empirical research on FDI, since foreign investments are expected to bring a bonus effect on economic growth of host countries. Over the last decades, governments have been promoting policies aiming the attraction of foreign firms, taking for granted the benefits arising from their presence. However, recent studies have been highlighting the importance of the absorptive capacity of host countries, as a necessary condition to learn with foreign presence and to take advantage from FDI.

In this paper we study the potential growth-enhancing effect of FDI on developed countries. Based on a sample of 30 countries of OECD, for the period 1996-2006, we use panel data estimation techniques to prove that FDI has a positive and significant effect on the acceleration of economic growth. We contribute to the existent literature by quantifying the minimum threshold of human capital and technological competencies that developed countries must attain to absorb the benefits arising from foreign investments. Our results support the need for policies aiming to improve the quality of the host environment, encouraging human capital formation and private R&D expenditures, in order to attract high quality FDI and to enhance the potential benefits on economic growth.

The remainder of this paper is structured as follows: in section 2 we present a theoretical framework and a brief literature review of the main empirical studies on this subject, in section 3 we describe the data and the methodology used, in section 4 we discuss the empirical results and section 5 concludes.

2. FDI, ECONOMIC GROWTH AND ABSORPTIVE CAPACITY: A THEORETICAL FRAMEWORK

FDI has been absorbing a substantial body of economic literature. Studies on FDI are clearly divided into two major topics: those dealing with the determinants of FDI and those ones concerning the impacts of FDI on economic performance. Developing countries have been the main focus of empirical research on FDI, since foreign investments appear to be a sort of panacea for every economic problem in the emerging market economies (Mencinger, 2003). Consequently, the respective governments have been changing their attitude towards FDI, promoting policies aiming the attraction of foreign firms, expecting large benefits as an enhancing effect on economic growth for host economies. For this reason, there is still little empirical evidence about those topics on developed countries. This paper contributes to the scarce literature on the potential effects of FDI for the economic growth of developed countries of OECD.

A vast literature highlights the potential benefits arising from FDI for host economies. De Mello (1997) and Blomström and Kokko (1997) survey the main developments in the literature on this issue, focusing on developing countries. From the empirical studies reviewed, FDI is believed to be a source of human capital augmentation and technological change, as well as a potential channel for knowledge transfers and productivity gains through spillovers to domestic firms. They also emphasize that MNCs, as agents of FDI, play an important role on trade performance and export growth of host countries. More recent empirical studies have been stressed that FDI provides a major source of saving and investment for host countries, particularly when domestic savings are low and domestic investment efforts are weak. Additionally, FDI improves the balance of payments and the value added of host production, as well as helps to intensify the competition in domestic markets, propelling domestic firms to increase their performance (Ghatak and Halicioglu, 2006; Mingyong et al., 2006; Duttaray et al., 2008; Xu et al., 2008). Finally, FDI may supply employment training, improving the human capital stock on host countries, in addition to better management practices, which can be imitated by local firms (Urata and Kawai, 2000; Chudnovsky et al., 2008; Fu, 2008).

On the other hand, some drawbacks may arise from the foreign presence on host markets. Ford et al. (2008) refer that, sometimes, the incentives to attract foreign firms result in decreased public education expenditures by host governments. Herzer et al. (2008) point that foreign firms may reduce the productivity of local firms through competition effects. In addition to this fact, Ozturk (2007) stresses the possible deterioration of the host balance of payments, as profits are repatriated. Xu et al. (2008) also argue that FDI may result in a crowding-out effect, reducing domestic investment.

So, the channels through which all those FDI positive and negative spillovers may occur can be summarized in some standard topics: i) demonstration effects, ii) human capital augmentation and skilled labour turnover, iii) higher exposure to fierce competition and iv) international trade (Urata and Kawai, 2000; Barrios et al., 2002; Mencinger, 2003; Akinlo, 2004; Ozturk, 2007; Fu, 2008).

However, literature proves that the benefits of FDI should not be taken for granted. The bonus arising from inward FDI not only depends on the motivations and the quality of foreign investments (Akinlo, 2004; Tytell and Yudaeva, 2006), but foremost on two main conditions: 1) sufficient effective linkages between foreign and domestic firms and 2) the absorptive capacity of the host economy to capture the potential spillovers from FDI (Fu, 2008). In this paper, we focus on the second condition, by stressing the importance of absorptive capacity, even in developed economies, since the major existent studies on this issue are applied on developing countries.

The study of Borensztein et al. (1998) was one of the pioneer studies that have stressed the importance of the absorptive capacity of host economies, usually proxied by human capital measures. Their study, applied to 69 developing countries for the period of 1970-1989, was the first to quantify a minimum threshold of human capital above which host countries would positively benefit from FDI. Since then, the literature on FDI has gained a new line of research, where the benefits of FDI on economic growth are contingent on the ability of host countries to

learn and absorb new technologies from foreign investments, in order to raise skills and local efficiency. Table 1 presents some of the main empirical studies on the impact of FDI on economic growth of host economies, highlighting the importance of the local absorptive capacity to benefit from foreign presence. Table 2 summarizes the reviewed studies that had stressed the role of absorptive capacity of host economies, presenting also the proxies used to measure it and the obtained results for the cases where any threshold level was quantified. From our knowledge, there are still few studies that have quantified the minimum threshold of variables determining the level of absorptive capacity of host countries. This paper contributes to the existing results, by providing new insights about this question.

From the reviewed empirical studies on the effect of FDI on host countries' economic growth, only three studies have found no effect from FDI on economic performance (Carkovic and Levine, 2002; Akinlo, 2004; Herzer et al., 2008). According to Ford et al. (2008), many studies fail to find a positive effect from FDI because they do not control for the ability of local firms to absorb foreign technology. In fact, Herzer et al. (2008) did not account for the absorptive capacity of host countries. On the other hand, Carkovic and Levine (2002) have not found any positive benefit arising from FDI, even after controlling for the existence of a possible interaction between FDI and human capital. Inversely, Olofsdotter (1998), who has found a positive effect of FDI irrespective of the capability level, argues that results may be sensitive to the empirical method, which can explain the above conclusions. At last, the results obtained by Akinlo (2004) for Nigeria are explained by the type of FDI that is mainly channelled to extractive activities.

Mencinger (2003), in a study applied to 8 Central and Eastern European Countries (CEECs), before their accession into the EU, was the only one who found a negative effect of FDI on economic growth of the receptive countries. The reasons advanced by the author are the characteristics of FDI – predominantly through acquisitions and strongly concentrated in trade and finance – as well as the characteristics of host countries, namely their small size.

Within the group of studies that have proved the existence of a positive benefit arising from FDI, some did not found a contingent relation on absorptive capacity of host countries. Zhang (2001), in a study applied to 11 economies from East Asia and Latin America, ascertained a positive effect from foreign investments for 5 of those countries. Although the author refers that spillovers efficiency do not appear automatically, but instead depends on host country's absorptive capacity, he did not check the conditional effect of FDI on that country-specific characteristic. Similar conclusions were obtained from the studies of Yao and Wei (2007) and Xu et al. (2008), both applied to 29 Chinese provinces, and Baharumshah and Almasaied (2009) for Malaysia. Only Ghatak and Halicioglu (2006) and Duttaray et al. (2008) found a positive effect of FDI on economic growth without controlling for any absorptive capacity.

However, since the results obtained by Borensztein et al. (1998), the absorptive capacity of host countries has been gaining an increasing concern from researchers, who have been trying to quantify the minimum threshold of variables related to human capital, trade openness, initial GDP and R&D expenditures, given that higher values for these variables are recognized as potential enhancers of host capabilities to learn from foreign presence. Borensztein et al. (1998) opened this new line of research by proving that only countries with a secondary school attainment above 0,52 years would benefit positively from FDI. Balasubramanyam et al. (1999) have also attempted to quantify the minimum threshold of human capital, through the real wage level, and found a threshold a little below the second quartile. Additionally, they have also proved that the efficiency of FDI in promoting growth is likely to be higher in countries pursuing an export promotion strategy than in countries pursuing an import substitution strategy. Xu (2000), using the same proxy for human capital applied by Borensztein et al. (1998), in a study conducted for 20 developed countries and 20 developing countries, revealed that the minimum necessary human capital threshold was somewhere between 1,4 and 2,4 years.

Many studies have followed this line of research, but few have reached a precise estimation for the minimum threshold of host absorptive capacity. The studies of Barrios and Strobl (2002) and Barrios et al. (2002) have purposed two new measures for the absorptive capacity of host countries: R&D expenditures and the exporting behaviour of domestic firms. The authors point that a firm's R&D activity may suggest its need for highly skilled workers. On the other hand,

export activities may also be important as exposure to foreign markets is likely to have made participating firms more likely to already have higher levels of technology. Mingyong et al. (2006), in addition to human capital, demonstrated that the degree of openness to trade determine the host country's capacity to absorb technology spillovers arising from FDI. Tytell and Yudaeva (2006) have also evidenced that not all FDI has positive spillover effects on host countries, by proving that spillovers are positive only in the case of export-oriented FDI, when FDI is driven by the more productivity foreign firms, when the level of corruption of host regions is low and the host education level is sufficiently high. Chudnovsky et al. (2008), in a study conducted to Argentine manufacturing sector with firm-level data, showed that positive spillovers from foreign presence were only captured by those firms with high absorptive capacities, which were measured through an index constructed by the authors, including quantitative and qualitative variables, with R&D activities and human capital formation assuming a relevant weight. Fu (2008) finds a positive, but indirect, effect of FDI on economic growth of Chinese provinces, through innovation capacities of host regions. However, a minimum level of regional R&D intensity and labour force quality are needed to benefit from this indirect spillover. Nevertheless, those benefits are contingent on geographic location of host provinces: only coastal regions profit from foreign investments, whereas in inland regions, even if the absorptive capacity is sufficient, FDI is not a significant driver of innovation capability nor economic growth.

Although all these last studies have highlighted the importance of absorptive capacity of host economies, as a necessary condition for a significant and positive effect of FDI on growth, no one have quantified the minimum threshold of the variables presented as proxies for the absorptive capacity. Only the recent studies of Ford et al. (2008) and Jyun-Yi and Chih-Chiang (2008) have contributed to the sparse existing results. Ford et al. (2008) have tested the effects of FDI on a developed country (U.S.A.) and proved that only the U.S. States that meet a level of population with a college degree between 12% and 16% receive a great impact on per capita output growth from foreign investments. Jyun-Yi and Chih-Chiang (2008), in a study conducted to 62 countries, demonstrated that FDI is found to have a positive and significant effect on economic growth when the host countries present a population with an average secondary school attainment of, at least, 2,108 years, a trade openness above -0,813 and an initial log of real GDP per capita over 8,011.

From all the reviewed studies, only Olofsdotter (1998) and Carkovic and Levine (2002) have considered the absorptive capacity not important to the potential benefit of FDI on economic growth. The first study considers the institutional capacity of host country, measured by the degree of property right protection and the bureaucratic efficiency, more important than absorptive capacity. The second study did not find evidence for a positive relation between FDI and growth, even after controlling for absorptive capacity.

Tables 1 and 2 summarize the empirical evidence presented above on the effect of FDI on economic growth and the importance of the absorptive capacity of host countries.

Table 1. Empirical Studies on the Effect of FDI on Host Countries' Economic Growth

Study	Countries	Time Period of Data	Methodology	Growth Proxy	Major Findings on the effect of FDI
Borensztein et al. (1998)	69 Developing Countries	1970-1989	SUR Techniques	Growth of real GDP pc	Positive (AC)
Olofsdotter (1998)	50 Developed and Developing Countries	1980-1990	OLS	Growth of real GDP pc	Positive (IC)
Balasubramanyam et al. (1999)	46 Countries	1970-1985	OLS and GIVE	Log(Real GDP)	Positive (AC; EPS)
Xu (2000)	20 Developing and 20 Developed Countries	1966-1994	2SLS and IVM	Growth Rate of Total Factor Productivity	Positive (AC)
Zhang (2001)	11 Countries from East Asia and Latin America	At least, 30 years, up to 1997	Causality Tests	Real GDP	Positive (in 5 countries)
Barrios and Strobl (2002)	Spain	1990-1998	OLS and FE Panel Regression	Total Factor Productivity	Positive (AC)
Barrios et al. (2002)	Greece, Ireland and Spain	1992-1997	OLS	Labour Productivity of Domestic Firms	Positive † (AC)
Carkovic and Levine (2002)	72 Countries	1960-1995	OLS and GMM	Growth of real GDP pc	No effect
Mencinger (2003)	8 CEECs	1994-2001	Causality Tests and OLS	Real GDP	Negative
Akinlo (2004)	Nigeria	1970-2001	ECM and OLS	Log (Real GDP)	No effect
Ghatak and Halicioglu (2006)	140 Countries across the world	1991-2001	OLS and IVM	Real GDP pc, Growth Rate of real GDP pc	Positive
Mingyong et al. (2006)	30 Chinese Provinces	1996-2002	Pooled OLS, FGLS, FE and RE	Real GDP Growth	Positive (AC)
Tytell and Yudaeva (2006)	Poland, Romania, Russia and Ukraine	1998-2003	OLS, FE, GMM	Log (Value Added), Total Factor Productivity	Positive (AC; EO; LC)
Yao and Wei (2007)	29 Chinese Provinces	1979-2003	RE, FE, GMM	Log(Real GDP)	Positive
Chudnovsky et al. (2008)	Argentina	1992-2001	FE Panel Regression	Log (Production of Firm)	Positive (AC)
Duttaray et al. (2008)	66 Developing Countries	1970-1996	Causality Tests	Real GDP Growth	Positive (in 29 countries)
Ford et al. (2008)	48 contiguous states - USA	1978-1997	LSDV, SUR Techniques and OLS	Growth Rate of GDP Per Worker	Positive (AC)
Fu (2008)	31 Chinese Provinces	1998-2004	RE and FE	Real GDP Growth	Positive * (AC; CR)
Herzer et al. (2008)	28 Developing Countries	1970-2003	Causality Tests	Log(Real GDP)	No effect
Jyun-Yi and Chih-Chiang (2008)	62 Countries	1975-2000	IVM, 2SLS, GMM	Growth of real GDP pc	Positive (AC)
Xu et al. (2008)	29 Chinese Provinces	1994-2006	FE Panel Regression	Total Factor Productivity	Positive
Baharumshah and Almasaied (2009)	Malaysia	1974-2004	ARDL models	Growth of real GDP pc	Positive

AC: Absorptive Capability; IC: Institutional Capability; EPS: Export Promotion Strategy; EO: Export-Oriented FDI; LC: Level of Corruption; CR: Coastal Regions; † Only for Ireland and Spain; * FDI affects indirectly economic growth, through innovation efficiency.

Table 2. Studies on the Importance of Absorptive Capacity

Study	Proxies for Absorptive Capacity	Threshold	Is AC Important ?
Borensztein et al. (1998)	Human Capital: Initial-year level of average years of the male secondary schooling	0.52 years	Yes
Olofsdotter (1998)	Human Capital (Average Mean Years of Schooling) and Openess to Trade	-	No
Balasubramanyam et al. (1999)	Human Capital (Real Wage Level)	A little below the 2nd quartile of wage	Yes
Xu (2000)	Human Capital (Years of Secondary School Attainment)	[1.4; 2.4] years	Yes
Barrios and Strobl (2002)	R&D expenditures and exporting behaviour	-	Yes
Barrios et al. (2002)	R&D expenditures and exporting behaviour	-	Yes
Carkovic and Levine (2002)	Human Capital (Average Years of Schooling)	-	No
Mingyong et al. (2006)	Human Capital (Average Educational Attainment per capita) and Openess to Trade	-	Yes
Tytell and Yudaeva (2006)	Human Capital (Share of Population with, at least, secondary school)	-	Yes
Chudnovsky et al. (2008)	Index of Absorptive Capabilities (includes R&D expenditures, innovation activities…) ranging between 1 and 5	-	Yes
Ford et al. (2008)	Human Capital (Proportion of the Population with a College Degree)	[12.04%; 15.56%] of population with a college degree	Yes
Fu (2008)	Regional R&D Intensity and Human Capital (Proportion of Population with 15 years' schooling)	-	Yes
Jyun-Yi and Chih-Chiang (2008)	Initial GDP (Log real GDP pc), Human Capital (Average Years of Secondary School) and Volume of Trade ((X+M)/GDP)	Initial GDP: 8.011; HC: 2.108; Trade Volume: -0.813	Yes

3. DATA AND METHODOLOGY

The data sources used in this paper are mainly OECD.Stat (Country Statistical Profiles 2008) and, additionally, UNESCO Custom Tables. We cover all the OECD countries, for the period of 1996-2006. The choice between those databases depended on the easy and free access to obtain the observations. However, the time span of analysis was limited by the existence of missing data, specially on human capital and innovation proxies.

In the first place, it was useful to use gross data as we are interested in the effects of FDI on host countries, particularly their economic growth. As a result, our dependent variable will be the economic growth of host countries, measured by the natural log of real GDP per capita (2000 US thousand dollars constant prices). With this dependent variable we can quantify the effect of the explanatory variables on the acceleration of host countries' economic growth, as argued by Balasubramanyam et al. (1999). Our base empirical specification for measuring the impact of inward FDI on economic growth of host OECD countries is as follows:

$$Log(Y_i) = \beta_0 + \beta_1 log(FDI_i) + \beta_2 HE_i + \beta_3 Pat_i + \beta_4 F_R\&D_i + \beta_5 B_R\&D_i + \beta_6 GDP_i(0), \quad (1)$$

where FDI is the inflows FDI to GDP ratio, HE is the proportion of population, with age between 25-34 years old, with higher education, Pat is the number of patents per capita, F_R&D is the proportion of GDP expenditures on R&D financed by foreigners, B_R&D is the proportion of GDP expenditures on R&D financed by private firms and GDP(0) is the initial real GDP per capita level. The chosen variable as proxy for human capital (HE) seems to be appropriate for our study, as we are analysing a sample of developed countries, so it is reasonable for our benchmark a proxy that captures a higher education level. All the variables were extracted from OECD.Stat, except F_R&D and B_R&D that were collected from UNESCO Custom Tables. Table 3 presents the descriptions of the variables, some statistics and expected signals.

Table 3. Description of variables and expected signals

Variable	Definition	Description	Obser v.	Mean	Expected Signal
Y	Real GDP per capita	Real GDP per capita (2000 US thousand dollars constant prices)	330	24.502	N.A.
FDI	FDI/GDP	Inflows of FDI to GDP Ratio (includes disinvestments)	318	0.313*	+
HE	Higher Education	Percentage of population aged between 25-34 years with higher education	289	27.523	+
F_R&D	Foreign R&D/R&D Expenditures	Proportion of GDP Expenditures on R&D financed by foreigners	223	0.071	+
B_R&D	Business R&D/ R&D Expenditures	Proportion of GDP Expenditures on R&D financed by private firms	233	0.526	+
Pat	Patents per capita	Number of Patents to Population Ratio	300	0.031	+
GDP(0)	Initial GDP per capita	Real GDP per capita in 1996 (US thousand dollars)	330	23.181	-

*Excluding Luxembourg, detected as an outlier in this variable. Y and FDI variables are used in natural logs in our model.

According to the reviewed literature, we expect a positive effect arising from FDI on economic growth. The variables HE, B_R&D, Pat and GDP(0) are used as proxies for the absorptive capacity (HE is used as proxy for human capital, B_R&D and Pat are proxies for technological competencies) so we expect that higher values of these variables will be associated with higher abilities to learn from foreign presence. Consequently, they are expected to exert a positive influence on host economic growth, except GDP(0), from which we expect a negative sign, consistent with conditional convergence theory. F_R&D, in conjunction with FDI, is another proxy for potential spillovers arising from foreign investments, then we also expect a positive effect from this variable.

We first estimated (1) for all countries, using FGLS (Feasible Generalized Least Squares) technique, applied to panel data. FGLS estimation is very useful,ÿÿllÿÿÿÿg for the correction of any prÿÿlems related to heteroskedasticity between groups and autocorrelation. Subsequently, to search for eventual differences on the effect of FDI between the countries, according to their level of human capital, we separated the observations in two samples – countries with HE above and below the mean – and ran the regression (1). All the results are presented in the next section.

4. RESULTS AND DISCUSSION

The findings from the application of the methodology presented in the previous section are summarized in Table 4.

Table 4. Regression results for Log (GDPpc), Panel Data, FGLS[1]

Dependent Variable: Log (GDPpc)	Cross-Sectional Time-Series FGLS Regression		
	All Countries	HE above the mean	HE below the mean
FDI	0.0935 ***	0.0310 ***	0.0498
	(0.0184)	(0.0089)	(0.0380)
HE	0.0121 ***	0.0037 **	0.0014
	(0.0023)	(0.0015)	(0.0089)
Pat	4.5585 ***	1.0605 **	8.4220 ***
	(1.0038)	(0.4202)	(2.1944)
F_R&D	3.7120 ***	0.5059 **	4.2613 ***
	(0.4275)	(0.2506)	(0.6152)
B_R&D	0.7969 ***	-0.1543	1.1357 ***
	(0.2490)	(0.1324)	(0.4090)
GDP(0)	-0.0059 **	0.0401 ***	-0.0153 ***
	(0.0029)	(0.0027)	(0.0040)
Constant	2.3694 ***	2.3359 ***	2.3179 ***
	(0.1552)	(0.1157)	(0.3237)
Number of Observations	198	98	99
Number of Groups	29	18	18

Notes: *Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level. Standard errors within parentheses.

From the estimation results of regression (1) for all countries, we can conclude that all the variables are significant at 1% level, except GDP(0) which is significant at 5% level, with the expected signals. We confirm that FDI inflows contribute positively to the economic growth of host countries, although the marginal effect of FDI on growth appears to be decreasing, since its coefficient, corresponding to elasticity, is lower than one. All the variables used as proxies to absorptive capacity, except the initial GDP level, have an enhancing effect on growth. The negative signal obtained for GDP(0) confirms the conditional convergence, according to which countries with lower GDP levels tend to grow faster than countries with higher initial GDP levels, demonstrating the catching-up effect.

Separating the full sample in two sub-samples, simply by those with HE above and below the mean, we obtained interesting results by detecting a significant difference on the FDI effect, as well as on the higher education effect. The second and third columns of Table 4 show that only the countries with higher education levels above the mean can benefit from FDI, whereas those with low higher education levels do not profit from foreign investments. Additionally, improving the proportion of population with higher education seems only to be significant to economic growth for those countries with a minimum level on that variable. These first results highlight the possible existence of a minimum threshold that even developed countries must attain to gain with foreign presence. Another result that deserves attention refers to the differences found on the signal and significance of GDP(0). The catching-up effect only binds for countries with higher education levels below the mean, suggesting that these countries are also those which lag behind on the economic growth acceleration.

These first results were the main attraction for us to search for a minimum threshold of HE level, in order to quantify the stage above which FDI turns to be significant and growth enhancing. To achieve this purpose, we adopted the methodology of Xu (2000). We ran the

[1] After the detection of some particular outliers, we run the above regressions with and without them, but the results were not significantly different.

regression (1) using samples selected according to different HE thresholds. The objective consists on, iteratively, find the threshold above which the FDI coefficient becomes positive and, then, the threshold above which the FDI coefficient turns out to be positive and significant for the acceleration of economic growth. The minimum threshold that host countries must achieve to benefit from FDI would be between those two values. The results are presented in Table 5. The coefficient of FDI is negative and statistically insignificant when the sample includes only those countries with HE values below 18%, corresponding to the percentile 25 of the variable. When the sample includes countries with HE values between 18% and about 33,5%, FDI coefficient is positive but statistically insignificant on the economic growth acceleration. Finally, when the sample includes countries with, at least, 34% of population aged between 25-34 years old with higher education, FDI becomes statistically significant. Curiously, FDI coefficient became positive and significant for economic growth only after the HE variable become also significant, confirming the need for an absorptive capacity in order to capture FDI spillovers. Although these regressions do not provide a precise estimate of HE minimum threshold, the method suggests that such a thresholds exists and it is somewhere between 18% and 34%. Alternatively, as argued by Xu (2000), we can say that the estimated HE threshold value is around 26% (the middle point of [18%; 34%] interval).

Table 5. Regression results for identifying human capital threshold, panel data, FGLS
All equations use the specification of regression (1). Sample is selected based on the HE threshold specified in the first column[2].

Sample (HE)	FDI Coefficient	FDI Z-statistic	Number of Observations
< 18[1]	-0.0224	-0.34	58
< 19	0.0516	0.89	65
< 20	0.0493	0.89	69
< 21	0.0440	0.82	71
…	…	…	…
< 28[2]	0.0439	1.18	102
< 29	0.0388	1.07	108
< 30	0.0407	1.14	113
< 31	0.0459	1.42	118
< 32	0.0434	1.36	122
< 33	0.0430	1.38	128
< 33.5	0.0450	1.47	130
< 34	0.0503	1.67*	132
< 35	0.0567	1.93*	137
< 36[3]	0.0575	2.07**	143
< 37	0.0650	2.39**	150
< 38	0.0799	3.18***	160

Notes: *Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level. [1]Percentil 25; [2]Percentil 50; [3]Percentil 75; HE values in percentage.

Ford et al. (2008) was the only study, from our knowledge, that has quantified a minimum threshold for this variable. Our result is a little above the threshold of [12.04%; 15.56%] estimated by the authors. However, as our proxy for human capital is restricted to the population aged between 25 and 34 years old, the estimation obtained for HE threshold seems to be

[2] Although all the results are not presented in the table, every intermediate values of HE were controlled. The signals and the significance level did not suffer any change. The same applies to Tables 6 and 7.

reasonable for the sample of countries under analysis. The previous methodology was replicated for the variables B_R&D and GDP(0), since the literature suggests that both may have a positive effect on the absorptive capacity of host countries. A minimum threshold for the variable Pat has also been estimated, however, due to the infinitesimal values obtained, we did not consider it so important as B_R&D and GDP(0). Although, the results are available upon request. Tables 6 and 7 present the results for B_R&D and GDP(0) thresholds.

Table 6. Regression results for identifying B_R&D threshold, panel data, FGLS
All equations use the specification of regression (1). Sample is selected based on the B_R&D threshold specified in the first column.

Sample (B_R&D)	FDI Coefficient	FDI Z-statistic	Number of Observations
< 0.60	-0.0105	-0.28	142
< 0.61	-0.0126	-0.35	146
< 0.62	-0.0098	-0.29	148
< 0.63	0.0039	0.12	150
…	…	…	…
< 0.70	0.0122	0.43	171
< 0.71	0.0129	0.46	174
< 0.72	0.0129	0.46	175
< 0.73	0.0440	1.81 *	182
< 0.74[1]	0.0484	2.1 **	184
< 0.75	0.0617	2.78 ***	190

Notes: *Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level. [1]Percentil 95; B_R&D values correspond to the share of GDP expenditures on R&D financed by private firms.

Table 7. Regression results for identifying GDP(0) threshold, panel data, FGLS
All equations use the specification of regression (1). Sample is selected based on the GDP(0) threshold specified in the first column.

Sample (GDP(0))	FDI Coefficient	FDI Z-statistic	Number of Observations
< 23.5	-0.0030	-0.21	120
< 23.6	-0.0030	-0.21	120
< 23.7	0.0017	0.12	121
< 23.8	0.0017	0.12	121
< 23.85	0.0017	0.12	121
< 23.89	0.0223	2.11 **	130
…	…	…	…
< 24.5	0.0171	1.72 *	146
< 25	0.0157	1.66 *	152
< 25.5[1]	0.0194	2 **	161
< 26	0.0229	2.61 ***	170
< 27	0.0215	2.46 ***	174

Notes: *Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level. [1]Percentil 75; GDP(0) values in 2000 US thousand dollars constant prices.

The interpretation of the results presented in previous tables is similar to that made for HE threshold. We can conclude that the B_R&D threshold is located somewhere between 62% and

73% or, alternatively, around 67,5%. The regression results suggest that the business R&D of host countries should be, at least, 67,5% of the total R&D expenditures of the country, in order to obtain a sufficient absorptive capacity to benefit for FDI. Similarly, the minimum threshold for initial GDP seems to be around US $23.750.

The complementarity of FDI and HE is illustrated in Figure 1. The sample was divided into 9 (3x3) groups according to the percentiles of FDI and HE variables. The figure shows that countries with high levels of FDI tend to have an accelerating effect on economic growth, *ceteris paribus*. The same conclusion can be taken for countries with high levels of HE. However, countries with high levels of both variables tend to benefit from a higher accelerating effect on economic growth. With this result, we demonstrate that absorptive capacity is as important to developed economies as to developing ones.

Figure 1. The Importance of Absorptive Capacity on Economic Growth Acceleration[3]

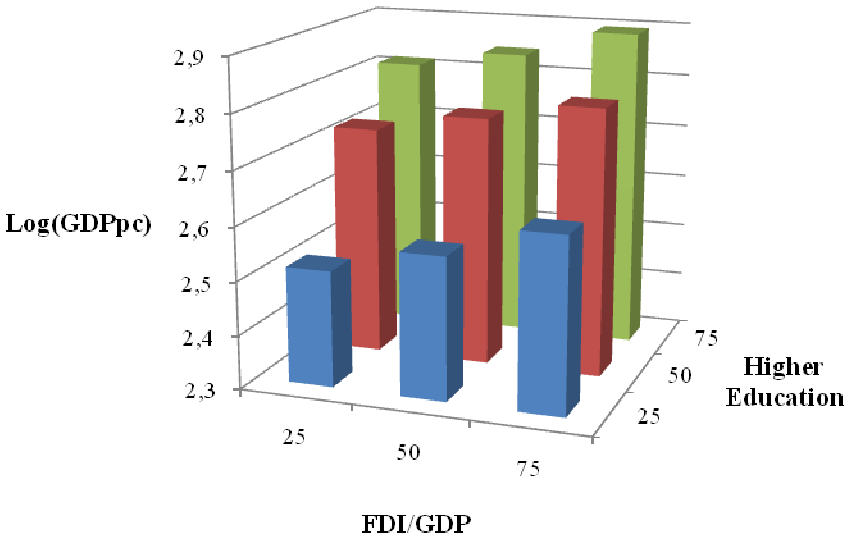
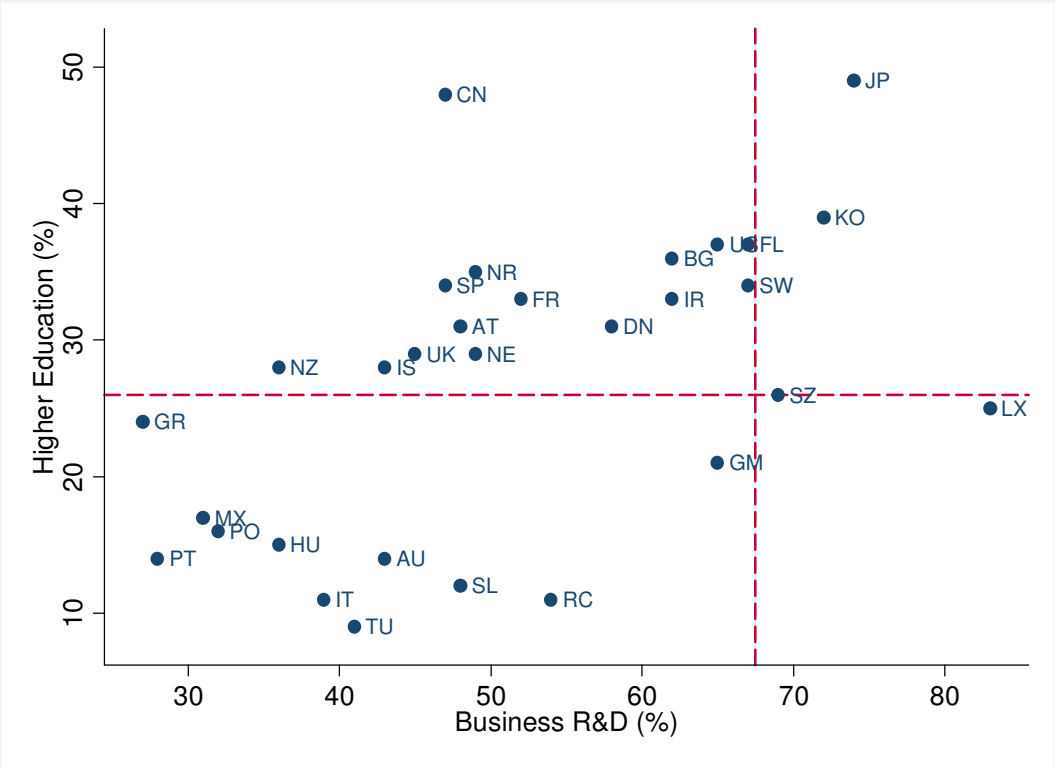


Figure 2 shows the position of the 30 countries of OECD relating to the minimum absorptive capacity estimated. Within the three thresholds estimated, we consider HE and B_R&D thresholds the most important ones. Whereas host governments can, through their policies, manipulate the position of their countries regarding HE and B_R&D levels, initial GDP level can not be influenced by future policy measures.

Since a minimum threshold for HE and B_R&D variables was found, it is interesting to verify where OECD countries stand. From the next figure, we can assess that, although we are facing a sample of developed countries, not all of them present the minimum absorptive capacity required. Only Japan, Korea, Finland, Switzerland and Sweden can be considered sufficiently prepared to benefit from foreign investments received, corresponding to less than 20% of OECD countries. On the reverse side, there are about ten countries in a critical position, still laying far from the both thresholds. We can also observe that, while more than an half of countries has already overcame the minimum threshold for HE, more than 80% of OECD countries are still behind the minimum level of B_R&D required to absorb the potential spillovers from FDI. The scenario suggests that policies more concerned with the improving of private R&D share in total R&D expenditures, are needed to boost the absorptive capacity and, consequently, the acceleration of economic growth.

[3] Note: Luxembourg excluded

Figure 2. HE and B_R&D Thresholds – The Position of OECD Countries



5. CONCLUSIONS

The main aim of this study was to find whether developed countries need to attain a minimum absorptive capacity to benefit from foreign investments received. Using panel data techniques on OECD countries data, for the period 1996-2006, we have proved that even developed countries must achieve a minimum threshold of human capital and technological competencies, in order to capture the spillover effects of FDI.

From our knowledge, there is only one study that has quantified the minimum threshold of human capital for developed countries (Ford et al., 2008). We contribute to the scarce existent literature on this topic, by demonstrating that developed countries will only benefit from FDI if they retain a minimum level of human capital (about 26% of population, aged between 25-34 years old, with higher education), as well as a minimum threshold of private R&D share in total R&D expenditures (about 67,5%).

Since the acceleration of economic growth of developed countries seems to be sensitive to the quality of the host environment – in human capital and technological competencies – government policies aiming to encourage higher levels of education and higher private R&D shares will improve absorptive capacities and will accelerate economic growth. According to the position of OECD countries on the minimum absorptive capacity required, it is crucial to provide incentives to R&D investments by private firms, since more than 80% of our countries sample are still below the technological competencies required to profit from FDI spillovers.

This paper gives empirical evidence that a well educated and technologically competent workforce is important in realizing the potential growth effects of foreign investments for advanced countries. Consequently, policies aiming to fostering absorptive capacity systematically should be at the top of the policy agenda of host countries that want to attract high quality FDI.

REFERENCES

AKINLO, A. (2004), "Foreign Direct Investment and Growth in Nigeria – An Empirical Investigation", *Journal of Policy Modelling*, 26, 627-639.

BAHARUMSHAH, A. AND ALMASAIED, S. (2009), "Foreign Direct Investment and Economic Growth in Malaysia: Interactions with Human Capital and Financial Deepening", *Emerging Markets Finance & Trade*, 45, 1, 90-102.

BALASUBRAMANYAM, V., SALISU, M. AND SAPSFORD, D. (1999), "Foreign Direct Investment as an Engine of Growth", *The Journal of International Trade & Economic Development*, 8, 1, 27-40.

BARRIOS, S. AND STROBL, E. (2002), "Foreign Direct Investment and Productivity Spillovers: Evidence from the Spanish Experience", Weltwirtschaftliches Archiv 2002, 138, 3, 459-481.

BARRIOS, S., DIMELIS, S., LOURI, H. AND STROBL, E. (2002), "Efficiency Spillovers from Foreign Direct Investment in the EU Periphery: A Comparative Study of Greece, Ireland and Spain", *FEDEA*, Documento de Trabajo 2002-02.

BLOMSTRÖM, M. AND KOKKO, A. (1997), "The Impact of Foreign Investment on Host Countries: A Review of the Empirical Evidence", *World Bank Policy Research Working Paper No. 1745*.

BORENSZTEIN, E., DE GREGORIO, J. AND LEE, J-W. (1998), "How does foreign direct investment affect economic growth?", *Journal of International Economics*, 45, 115-135.

CARKOVIC, M. AND LEVINE, R. (2002), "Does Foreign Direct Investment Accelerate Economic Growth?", *University of Minnesota*, Department of Finance, Working Paper Series.

CHUDNOVSKY, D., LÓPEZ, A. AND ROSSI, G. (2008), "Foreign Direct Investment Spillovers and the Absorptive Capabilities of Domestic Firms in the Argentine Manufacturing Sector (1992-2001)", *Journal of Development Studies*, 44, 5, 645-677.

DE MELLO, L. (1997), "Foreign Direct Investment in Developing Countries and Growth: A Selective Survey", *Journal of Development Studies*, 34, 1, 1-34.

DUTTARAY, M., DUTT, A. AND MUKHOPADHYAY, K. (2008), "Foreign Direct Investment and Economic Growth in Less Developed Countries: An Empirical Study of Causality and Mechanisms", *Applied Economics*, 40, 1927-1939.

FORD, T., RORK, J. AND ELMSLIE, B. (2008), "Foreign Direct Investment, Economic Growth and the Human Capital Threshold: Evidence from US States", *Review of International Economics*, 16, 1, 96-113.

FU, X. (2008), "Foreign Direct Investment, Absorptive Capacity and Regional Innovation Capabilities: Evidence from China", *Oxford Development Studies*, 36, 1, 89-110.

GHATAK, A. AND HALICIOGLU, F. (2006), "Foreign Direct Investment and Economic Growth: Some Evidence from Across the World", *Munich Personal RePEc Archive*.

HERZER, D., KLASAN, S. AND NOWAK-LEHMANN, F. (2008), "In Search of FDI-led Growth in Developing Countries: the Way Forward", *Economic Modelling*, 25, 793-810.

JYUN-YI, W. AND CHIH-CHIANG, H. (2008), "Does Foreign Direct Investment Promote Econmic Growth? Evidence from a Threshold Analysis", *Economics Bulletin*, 15, 12, 1-10.

MENCINGER, J. (2003), "Does Foreign Direct Investment Always Enhance Economic Growth?", *KYKLOS*, 56, 4, 491-508.

MINGYONG, L., SHUIJUN, P. AND QUN, B. (2006), "Technology Spillovers, Absorptive Capacity and Economic Growth", *China Economic Review*, 17, 300-320.

OLOFSDOTTER, K. (1998), "Foreign Direct Investment, Country Capabilities and Economic Growth", *Weltwirtschaftliches Archiv 1998*, 134, 3, 534-547.

OZTURK, I. (2007), "Foreign Direct Investment – Growth Nexus: A Review of the Recent Literature", *International Journal of Applied Econometrics and Quantitative Studies*, 4, 2, 79-98.

TYTELL, I. AND YUDAEVA, K. (2006), "The Role of FDI in Eastern Europe and New Independent States: New Channels for the Spillover Effect", *Centro Studi Luca D'Agliano*, Development Studies Working Papers No. 217.

URATA, S. AND KAWAI, H. (2000), "Intra-firm Technology Transfer by Japanese Manufacturing Firms in Asia", in Takatoshi Ito and Anne Krueger (eds.), *The Role of Foreign Direct Investment in East Asian Economic Development*, Chicago: University of Chicago Press.

XU, B. (2000), "Multinational Enterprises, Technology Diffusion and Host Country Productivity Growth", *Journal of Development Economics*, 62, 477-493.

XU, H., LAI, M. AND QI, P. (2008), "Openess, Human Capital and Total Factor Productivity: Evidence from China", *Journal of Chinese Economic and Business Studies*, 6, 3, 279-289.

YAO, S. AND WEI, K. (2007), "Economic Growth in the Presence of FDI: The Perspective of Newly Industrialising Economies", *Journal of Comparative Economics*, 35, 211-234.

ZHANG, K. (2001), "Does Foreign Direct Investment Promote Economic Growth? Evidence from East Asia and Latin America", *Contemporary Economic Policy*, 19, 2, 175-185.

ANNEX I

Sample of Countries used in the Study

Australia	AT	Hungary	HU	Norway	NR
Austria	AU	Island	IS	Poland	PO
Belgium	BG	Ireland	IR	Portugal	PT
Canada	CN	Italy	IT	Slovak Republic	SL
Czech Republic	RC	Japan	JP	Spain	SP
Denmark	DN	Korea	KO	Sweden	SW
Finland	FL	Luxembourg	LX	Switzerland	SZ
France	FR	Mexico	MX	Turkey	TU
Germany	GM	Netherlands	NE	United Kingdom	UK
Greece	GR	New Zealand	NZ	USA	USA

HEALTH CONDITIONS AND ECONOMIC GROWTH: EVIDENCE FROM OECD COUNTRIES

ANA ISABEL DA SILVA ALVES POÇAS
E-mail: anapocas@ipg.pt
Escola Superior de Tecnologia e Gestão
Instituto Politécnico da Guarda

ELIAS SOUKIAZIS
E-mail: elias@fe.uc.pt
Faculdade de Economia
Universidade de Coimbra

Resumo

Este estudo tem por objectivo analisar o impacto do capital humano num sentido lato (incluindo quer o factor educação, quer o factor saúde) no crescimento económico dos países da OCDE. São estimadas regressões de crescimento económico do tipo Solow, recorrendo a um modelo dinâmico de dados em painel constituído por 22 países da OCDE (para os quais existe disponibilidade de dados) para o período 1980-2004. Os resultados obtidos permitem concluir que o capital humano ligado à inovação e às condições de saúde são factores importantes para explicar o crescimento e a convergência económica. Em particular, algumas das proxies utilizadas, como sejam a taxa de mortalidade causada por AVC e cancro, o número médio de dias de internamento, a densidade de médicos, número de camas, despesa per capita em produtos farmacêuticos, entre outros, afectam significativamente o crescimento económico (e a convergência) dos países desenvolvidos.

Palavras-chave: crescimento económico, capital humano, capital saúde, dados em painel.
Área Temática: Crescimento económico.

Abstract

The aim of this study is to analyse the impact of human capital in a broad sense (including both education and health) on economic growth for the OECD countries. Growth equations of the Solow type are estimated by using panel data estimation techniques. The sample is constituted by 22 OECD countries and the time spans from 1980 to 2004 where data is available. We find evidence that human capital related to innovation and health conditions are important conditioning factors to explain growth and convergence. In particular, some of the proxies used such as mortality rate caused by AVC and cancer, average length of stay for acute care, density of practising physicians, acute care beds, per capita pharmaceutical expenditure, among others, significantly affect economic growth (and convergence) of developed countries.

Keywords: economic growth, human capital, health capital, panel data.
Thematic Area: Growth Economics.

1. INTRODUCTION

It is now well recognized the important role that health plays on economic growth. Improved health cannot only contribute to a better economic performance at an individual level but also at aggregate level affecting the whole economy. On the other hand, economic growth also contributes to further accumulation of health capital, allowing higher investment on health sectors. There are several studies that corroborate the idea that there is a reverse causality between health and economic growth (Rivera and Currais, 1999). In this way, investment in health must be seen not only as a social benefit but also as a key factor to economic growth.

In fact, health can be seen as one component of human capital just as education is for improving labour productivity. Traditionally, human capital was associated with education and skills. The main idea is that more educated individuals are also more productive. Human capital is seen as the capacities/skills that are incorporated in workers, which are key factors of productive efficiency. Endogenous growth models usually consider human capital as the quantity of accumulated human capital stock, generally measured by years of education of the working population. In this sense human capital is the driving force of economic growth, with cumulative effects on physical capital and technical progress.

However, in recent years an increasing literature has emerged using a broader notion of human capital which considers that an increase in individual's stock of knowledge and health[1] raise his productivity. A seminal work in this area is due to Grossman's (1972) model for health demand that captures the interrelation among work time, wages and health.

Albeit the interest of the impact of health on economic growth (and convergence process), most of the economic literature in this area concerns differences between rich and poor countries while there is much less research on developed countries (European Commission, 2005; Suhrcke et al., 2006). Additionally, the results of some of these studies are not very clear. When Knowles and Owen (1997) estimated an extended Mankiw Romer Weil (MRW) model that includes health (given by life expectancy) for 22 developed countries, they found no statistical significance of the health variable[2]. Another study from Bhargava et al. (2001) identifies an important link between health (using adult survival rate as proxy of health status) and economic growth in poor countries, but is unable to replicate this result when only developed countries are considered.

According to Suhrcke and Urban (2006), these unclear results can be in part a consequence of the use of conventional health proxies such as life expectancy or infant mortality rates. If there is more consensus in the use of these proxies when analysing developing countries, other variables can be more relevant when developed countries are concerned. In this case, Tompa (2002) suggests the use of proper health proxies, such as cardiovascular disease mortality rate of active population, mental diseases or morbidity indicators. Between developed countries these indicators show greater variability and so they can better reflect the efficiency of the health system.

The aim of this study is therefore to analyse the contribution of the extended notion of human capital (including both education and health) on economic growth, considering a sample of 22 OECD countries. Following recent literature suggestions, we use different proxies to capture the efficiency of the health system and to differentiate better the health conditions among developed countries. With this purpose we examine the impact some health indicators have on growth, such as cerebrovascular disease mortality, life expectancy at 65 years old, average length of staying for acute care (days), practising physicians (density per 1 000 population), pharmaceutical expenditure per capita, among others.

The augmented Solow growth model is used that considers human and health capital as conditioning factors to growth. Panel data estimations techniques are applied to estimate growth equations, for a sample of 22 OECD countries from 1980 to 2004.

The paper is organized as follows. Section 2 describes the several links between health, human capital and income per capita, and reviews some of the existing literature. In section 3 we explain the growth model and the estimation approach suitable to panel data to obtain consistent estimates. The empirical results are presented and discussed in section 4. The final section concludes suggesting some policy implications.

---

[1] It is assumed that, like physical capital, health capital depreciates over time at an increasing rate with age.

[2] The authors' empirical results are reported for a full sample of 77 countries and also for a sub-sample of 55 less developed countries, over the period 1960 to 1985, and suggest a strong positive relationship between economic growth and health. Only when the model is estimated for the 22 high-income countries, the authors find that the health status parameter is not significant. They consider this result expectable, given the lack of variation in life expectancy over this sub-sample.

2. LINKS BETWEEN HEALTH, EDUCATION AND ECONOMIC GROWTH

There are several links through which health contributes to improve standards of living measured by income per capita.

The first link is through education. There is a direct relationship between health and education: a healthier childhood (through better nutrition) improves cognitive capacities and school attendance. So it is expectable that healthier children will have better education and will be more productive in the future. As Bloom (2007) points out, the health of other family members and maternal health (which is closely connected with child health) also affect educational performance.

On the other hand, good health also makes investment in education more likely. Improvement in health is associated with a raise in life expectancy[3], which makes investment in education more attractive, because individuals expect to have the return of this investment for a longer period of time. The increase in longevity may also influence expenditure/saving decisions over the life cycle, by decisions related to retirement planning.

The relationship between education and health has also been shown by several authors (Cutler and Lleras-Muney, 2006; Albert and Davia, 2007), usually exploring three channels: productive efficiency, allocation efficiency and time preference. The first approach states that more educated people are more efficient in the use of healthcare services. Under the second approach, education is considered the driving force of health improvement: more educated individuals are more conscious of the negative impact of their risk behavior and tend to invest more time and resources on healthcare. According to the time preference hypothesis, improvements on individuals' outlook on the future (which means an increase of the present discounted value of future lifetime utility) make people more likely to invest in protecting that future (Cutler and Lleras-Muney, 2006).

Health is also a determinant factor of the relative size of labour force. One important consequence of a better health system is a decline in infant mortality, which has positive effects over active population growth. However, as Weil (2006) points out, the effect of health improvements over population growth is ambiguous because in the long run lower infant and child mortality may lead to a decline in fertility.

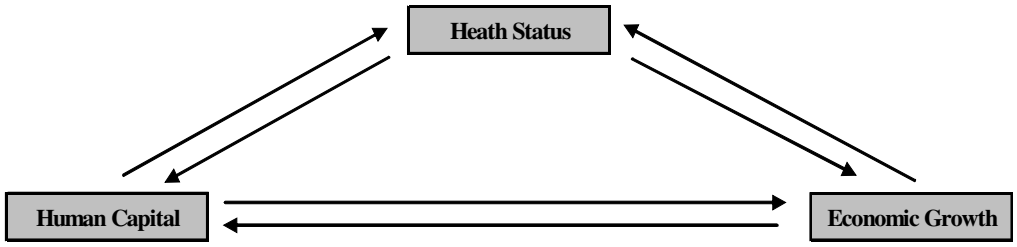
Labour productivity is another important link through which better health contributes to increased efficiency. At a microeconomic level it is intuitively understandable that individual's health affects labour productivity and labour supply. There is empirical evidence (Schultz, 2005; Cai and Kalb, 2006) that healthier workers have more physical and mental energy and they are more creative and productive. These workers earn better wages and have a lower probability of absenteeism at work (Bloom *et al.*, 2001). Additionally, they have higher chances of receiving skill upgrading investment from the part of the firms they work. Usually more educated individuals are the ones who have better jobs, with higher incomes and safer work conditions and so they are likely to invest more on health.

As Soukiazis (2008) explains there is a mutual causation tendency between income, health and human capital with feed-back and spillover effects. This reciprocal interrelation can give rise to a cumulative causation process, with health improvements leading to higher human capital accumulation and therefore higher economic growth, and the process continues to expand in a virtuous circle[4] as it is explained in Figure 1:

[3] While in less developed countries the raise in life expectancy is due mainly to the reduction of infant mortality (partly through improvements in child vaccination and nutrition), in rich countries it is driven by reductions in the mortality rates of the middle-aged and elderly (partly through improvements in medical prevention and treatment).

[4] Contrasting to situations characterised by poor health and poor economic growth leading to poverty trap.

Figure 1 - Causality effects between Heath, Human Capital and Economic Growth



Source: Soukiazis, Elias (2008).

As it can be seen, health has direct effects on human capital and economic growth due to better education and higher productive efficiency. On the other hand, better education contributes to improve health conditions through two channels: at the individual level more educated people invest more on health care and live in healthier conditions, at the aggregate level, societies with higher levels of educational standards invest more on research and development in the health sector. A trivial causation effect is from economic growth to health and human capital. As countries improve their economic performance they have the capacity to invest more on education and health services. This cumulative causation process is characterised by increasing returns to scale properties and positive externalities stemming from human capital and health sectors, that turn the process self expanding.

There are many studies that try to analyse the channels between health and economic performance at the individual or the aggregate level. The economic literature that studies the macroeconomic impact of health usually follows two approaches: the aggregate production function or the economic growth regressions. The first approach carries out an accounting decomposition of differences in output levels and was recently followed by Bloom and Canning (2005)[5]. Some of the limitations of this approach are that it imposes technology parameters based on microeconomic evidence and it assumes an aggregate production function that works in a similar way as the production function at the firm level.

The economic growth approach has a more solid theoretical background than the production function approach and it is in fact the most used in the broader literature of economic growth. Following this approach, Rivera and Currais (1999, 1999a) use health expenditure as a share of GDP and estimate its impact on growth of the OECD countries for the period 1960-90. Both studies show a significant positive effect of this expenditure on income per capita and growth.

Beraldo et al. (2005) also use health expenditure (as a share of GDP) to analyse its impact on growth applied to a panel of 19 OECD countries over the period 1971-1998. Using a production function approach, they show that the explaining power of health expenditure on growth (between 16 and 27%) is greater than the education expenditure (nearly 3%). Although they found statistically significant results, some authors (Tompa, 2002) argue that this proxy for health should be used with caution, since it is not a good measure of the efficiency of health systems.

Soukiazis and Cravo (2008) implement a dynamic panel data approach for 77 countries over the period 1980-2000 and they divide this sample in three subgroups according to their per capita income level. The authors show that for the whole sample both health (measured by infant mortality rate) and human capital (measured by article publication rate) as well as physical capital are important determinants of economic growth and convergence. The authors also analyse the convergence process in the different subgroups and they found that while in the

[5] In this study, based on a panel of countries for the period 1965-1995, the authors show that health, in the form of adult survival rates, has a positive and statistically significant contribution on aggregate output.

high-income countries human capital is more relevant, in the low-income countries health is more important to differentiate these economies.

Amaral (2007) studies the impact of human capital and health on the convergence process across the European Union (15) countries for the period 1980-2000. Using a panel data approach, he finds evidence that both factors have a positive impact on the convergence process (with fertility rate as a proxy for health), but spending on health and education are not relevant.

Last but not least, Suhrcke and Urban (2006) compare 26 rich countries for the 1960-2000 period, using cardiovascular disease mortality (CVD) rate as a proxy for health. Implementing a dynamic panel growth approach, the authors found that CVD mortality rate in working-age population is a robust predictor of subsequent five-year growth rates. Their empirical analysis shows that a reduction of CVD mortality by 10% was associated with an increase in the growth of per capita income by 1%.

3. THE MODEL AND ESTIMATION APPROACH

Our empirical model follows Mankiw, Romer and Weil (1992) (hereafter MRW), which is an augmented Solow model with human capital and health conditions. The MRW model improves the Solow model by including the accumulation of both human and physical capital. In our model human capital is intended in a broader sense that includes education as well as health conditions.

The growth equation to estimate is given by:

$$gy_{i,t} = b\,ln(y_{i,t-1}) + c_1\,ln(n_{i,t} + g + \delta) + c_2\,ln(s_{i,t}) + c_3\,ln(E_{i,t}) + c_4\,ln(H_{i,t}) + \varepsilon_{i,t}$$

where $\varepsilon_{i,t} = \alpha_i + u_{i,t}$, with α_i denoting the country-specific effects or measurement errors and $u_{i,t}$ refers to the idiosyncratic error term.

The dependent variable, $gy_{i,t}$, is the growth of per capita income considering five year intervals. We regress $gy_{i,t}$ on $y_{i,t-1}$, the initial per capita income of each period whose coefficient reflects the well known convergence hypothesis[6] when appears with negative sign; $n_{i,t}+g+\delta$ is the annual growth rate of population plus the rate of technological progress (g) and the rate of capital depreciation (δ); $s_{i,t}$ denotes the investment share, $E_{i,t}$ is human capital and $H_{i,t}$ health capital.

We estimate this equation using panel data for 22 OECD countries (given by the subscript i) over the period 1980-2004. With the aim of controlling economic cycle effects, we consider five year periods, so each country has five observations (given by the subscript t). Table 1 explains the set of variables considered in our empirical study and data sources:

[6] Barro and Sala-i- Martin (1992) developed this idea.

Table 1: Description of variables and data sources

Variable	Description	Source
y	Real GDP per capita (Laspeyres), dollars in 2000 constant price – RGDPL	Penn World Table 6.2.
n	Annual average growth rate of population	Penn World Table 6.2.
s	Investment share as a percentage of RGDPL in 2000 constant prices	Penn World Table 6.2.
pat	Number of patents per million of inhabitants aged 25 or over	USPTO (patents) and International Labour Office (population)
healthexpend	Total expenditure on health as a percentage of GDP	OECD, Health Data 2008
phexpepc	Pharmaceutical expenditure per capita, US$ PPC	OECD, Health Data 2008
infmort	Infant mortality rate	OECD, Health Data 2008
lifexpect	Life expectancy at birth	OECD, Health Data 2008
fert	Fertility rate	OECD Factbook 2008: Economic, Environmental and Social Statistics
avc	Cerebrovascular diseases: deaths per 100 000 total population	OECD, Health Data 2008
cancer	Cancer: deaths per 100 000 total population	OECD, Health Data 2008
respir	Diseases of respiratory system: deaths per 100 000 total population	OECD, Health Data 2008
myocard	Acute myocardial infarction: deaths per 100 000 total population	OECD, Health Data 2008
diabet	Diabetes: deaths per 100 000 total population	OECD, Health Data 2008
lifelostfem	Potential years of life lost, all causes, females (0-69)	OECD, Health Data 2008
lifelostmale	Potential years of life lost, all causes, males (0-69)	OECD, Health Data 2008
lifexpect65f	Life expectancy at 65 years old (females)	OECD, Health Data 2008
lifexpect65m	Life expectancy at 65 years old (males)	OECD, Health Data 2008
physicians	Practising physicians, density per 1 000 population	OECD, Health Data 2008
acbed	Acute care beds, density per 1 000 population	OECD, Health Data 2008
dconsult	Doctor consultations per capita	OECD, Health Data 2008
stay	Average length of stay for acute care, all conditions (days)	OECD, Health Data 2008

Notes: Practising physicians are those seeing patients either in a hospital, practice or elsewhere; acute care beds have been defined as beds accommodating patients in a hospital or hospital department whose average lenght of stay is 30 days or less until the 1980s and 18 days or less afterwards.
Data from Penn World Table was collected from Heston et al. (2006), it is available at http://pwt.econ.upenn.edu/.
Data from U.S. Patent and Trademark Office (USPTO) is available at http://www.uspto.gov.
Data from International Labour Office is available at http://laborstat.ilo.org.
OECD Health Data 2008 is available at http://www.oecd.org/topicstatsportal.
OECD Factbook 2008 is available at http://lysander.sourceoecd.org/vl=1275449/cl=20/nw=1/rpsv/factbook/.

Table 2 presents some descriptive statistics of the explanatory variables.

Table 2: Variables' descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
y	132	21.233,220060	5.286,957041	9.956,772	36.100,439
n	132	0,054452	0,003994	0,048	0,065
s	132	23,150184	3,588678	13,663	34,414
pat	132	178,552384	180,045890	0,949	988,495
healthexpend	130	8,086923	1,639833	5,100	15,200
phexppc	120	236,891667	140,880293	54,000	756,000
infmort	131	7,158779	3,343832	2,800	24,200
lifexpect	132	76,927273	2,179680	71,500	82,100
fert	132	1,707955	0,300927	1,180	3,250
avc	130	79,777692	38,954086	29,200	273,900
cancer	130	177,607692	20,768509	136,200	222,500
respir	130	59,646923	22,124665	28,500	139,300
myocard	116	80,106034	42,585493	18,700	214,100
diabet	130	13,237692	5,618053	2,000	28,800
lifelostfem	130	3.273,461538	838,273781	1.724,000	6.308,000
lifelostmale	130	6.089,992308	1.581,547993	2.946,000	11.987,000
lifexpect65f	132	18,919697	1,420109	15,700	23,300
lifexpect65m	132	15,240909	1,414697	12,500	18,200
physicians	117	2,664103	0,720328	1,300	4,900
acbed	109	4,527523	1,670378	2,200	12,000
dconsult	99	5,746465	2,490280	2,500	14,600
stay	114	8,639474	4,111878	3,400	33,200

As Table 2 shows, some of the proxies of health considered in our model have missing observations, so in those cases we have an unbalanced panel.

When we add health conditions in economic growth models, there are some additional difficulties we must attend for. One of the problems is related with unobservable heterogeneity. Countries have different economic, political and institutional characteristics and so the use of linear regressions that do not take into account these differences are not adequate. As Islam (1995) notes, panel data techniques are more suitable, since they take into account different structures across countries, in the form of fixed or random effects.

Another problem has to do with measurement error. When there are several badly measured variables or they depart from classical hypotheses, the bias can be in either direction (Temple, 1999). Concerning the use of health proxies, this problem can result from the choice of subjective measures of health that do not differentiate properly the countries in the sample.

Testing the effect of health, but also initial income and human capital on economic growth is a difficult task because of the problem of the endogeneity of regressors. Given this endogeneity problem, the use of Ordinary Least Squares (OLS) delivers biased and inconsistent estimates. To avoid this problem, instrumental variables techniques are often used (Bloom *et al.* 2001, Rivera and Currais, 2005). An instrumental variable must satisfy two conditions: first, it should be correlated with the endogenous explanatory variable and, second, it should not be correlated with the error term. Usually lags of the endogenous variables are considered as instruments. Yet, as Temple (1999: 129) notes, in what concerns human capital accumulation there may be some delay in the effects of this factor and the exogeneity of the instruments is not always clear. Another problem to consider is the omitted variable bias due to omission of relevant factors that explain growth.

Having in mind these problems, and in order to obtain consistent and efficient estimates, we use the Generalized Method of Moments (GMM) panel estimation technique designed by Arellano and Bond (1991)[7]. According to Arellano (2003), some of the reasons that make the

[7] This approach was first introduced in growth literature by Caselli et al. (1996).

use of GMM techniques in dynamic panel data popular are that the estimates are consistent in short panels, robust and have general applicability. On the other hand, GMM methods allow to control for measurement errors and omitted variables bias (Bond et al., 2001).

According to Bond et al. (2001), the use of difference GMM techniques in studying economic growth has important advantages over cross-section regressions or other estimation methods for dynamic panel data models. In fact, it avoids the problem of omitted variables that are constant over time (unobserved country-specific effects) and so estimates will no longer be biased. On the other hand, the use of instrumental variables allows parameters to be estimated consistently in models that include endogenous right-hand-side variables even in the presence of measurement error.

However, it must be noted that difference GMM has also some disadvantages. When the time series are persistent and the number of time series observations is small, the first difference GMM is poorly behaved because lagged levels of variables are weak instruments for subsequent first-differences.

There are two kinds of estimators obtained from one-step and two-step estimations. They differ in the errors assumptions. One-step considers the errors to be i.i.d., while two-step estimator assumes the heteroscedasticity of errors. According to Windmeijer (2005), two step GMM estimates are better than one step, because it has lower bias and smaller standard errors.

As already referred, some of the proxies of health considered in our model have missing observations, so in those cases we have an unbalanced panel. In order to get consistent estimates of the parameters of interest, and following Roodman (2006), we used two common transformations: the first-difference and the orthogonal deviations transformation. The first-difference transformation eliminates the country-specific effects, but has the disadvantage of rising gaps in unbalanced panels. In those cases, we used "forward orthogonal deviations" or "orthogonal deviations" transformation which subtracts the average of all subsequent available observations of a variable. This procedure is computable for all observations except the last for each country and so it minimizes the data loss[8]. On the other hand, both transformations allow the use of lagged variables as valid instruments and so they are consistent and comparable.

4. EMPIRICAL EVIDENCE

Since our main aim is to analyse the impact of health conditions on economic growth, we will give a special attention to the results related to health proxies used.

Table 3 shows the expected sign of health proxies as well as the signs obtained from the estimation. With the exception of "total expenditure on health", the estimated coefficients of all other health variables have the expected sign. With respect to "total expenditure on health", its negative impact on growth (but not statistically significant) can be interpreted as reinforcing the idea that is not the amount of expenditure on health that really matters but the efficiency of that expenditure. Amaral (2007) also found a negative impact of this variable with respect to EU countries.

In what concerns the variables fertility rate, diabetes, diseases of respiratory system and acute myocardial infarction, potential years of life lost (male and female) and doctor consultations per capita, although they have their expected sign, they are not statistically significant.

[8] We ran this transformation using the command "orthog" available on STATA version 9.2.

Table 3: Expected sign and estimation results for health proxies

Variable	Expected Sign	Estimation
healthexpend	Positive	Negative coefficient, without statistic significance
phexppc	Positive	Positive, statistically significant at 5% level
infmort	Negative	Negative, statistically significant at 5% level
lifexpect	Positive	Positive, statistically significant at 10% level
fert	Positive	Positive coefficient, without statistical significance
avc	Negative	Negative, statistically significant at 1% level
cancer	Negative	Negative, statistically significant at 5% level
respir	Negative	Negative coefficient, without statistical significance
myocard	Negative	Negative coefficient, without statistical significance
diabet	Negative	Negative coefficient, without statistical significance
lifelostfem	Negative	Negative coefficient, without statistical significance
lifelostmale	Negative	Negative coefficient, without statistical significance
lifexpect65f	Positive	Positive, statistically significant at 5% level
lifexpect65m	Positive	Positive coefficient, without statistical significance
physicians	Positive	Positive, statistically significant at 1% level
acbed	Negative	Negative, statistically significant at 1% level
dconsult	Positive	Positive coefficient, without statistical significance
stay	Negative	Negative coefficient, statistically significant at 1% level

The results from the dynamic panel estimation using difference GMM are reported in Table 4[9]:

[9] In this Table we only report the regressions with significant results. The 22 OECD countries considered in our sample (for which data was available) are: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom and United States of America.

Table 4: Growth regressions and the relevance of health variables. OECD countries, 1980-2004

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$\ln(y_{i,t-1})$	-0.0372*	-0.0987**	-0.1211***	-0.1781*	-0.1742***	-0.0872***	-0.1827**	-0.1543**	-0.1682***	-0.1469***	-0.1922***
	(-1.807)	(-2.312)	(-3.264)	(-2.065)	(-4.244)	(-2.937)	(-2.433)	(-2.172)	(-4.065)	(-7.705)	(-6.105)
$\ln(ni,t+g+\delta)$	0.0227	0.1849	-0.0583	0.0647	0.1207	0.1068	0.0955	0.0135	-0.0153	-0.0622	0.1085
	(0.149)	(1.628)	(-0.482)	(0.964)	(1.148)	(1.120)	(0.893)	(0.107)	(-0.229)	(-1.159)	(0.922)
$\ln(s_{i,t})$	0.1146***	0.0373	0.0273	0.0718*	0.0640***	0.0825**	0.0981***	0.0694*	0.0573*	0.0349	0.0805***
	(3.223)	(1.122)	(0.597)	(1.788)	(2.976)	(2.432)	(3.337)	(1.950)	(2.070)	(1.394)	(2.857)
$\ln(pat_{i,t})$		0.0404	0.0019	-0.0096	0.0030	0.0107	-0.0040	0.0087	-0.0021	0.0019	0.0408*
		(1.481)	(0.0910)	(-0.483)	(0.255)	(0.678)	(-0.190)	(0.495)	(-0.207)	(0.232)	(2.072)
$\ln(infmort_{i,t})$			-0.0553**								
			(-2.473)								
$\ln(lifexpect_{i,t})$				1.1653*							
				(1.822)							
$\ln(avc_{i,t})$					-0.0924***						
					(-2.906)						
$\ln(cancer_{i,t})$						-0.1177**					
						(-2.298)					
$\ln(lifexpect65f_{i,t})$							0.4293**				
							(2.237)				
$\ln(lifexpect65m_{i,t})$							0.0448				
							(0.246)				
$\ln(stay_{i,t})$								-0.0919***			
								(-3.490)			
$\ln(phexppc_{i,t})$									0.0457**		
									(2.631)		
$\ln(physicians_{i,t})$										0.1417***	
										(6.142)	
$\ln(acbed_{i,t})$											-0.0899***
											(-2.882)
Observations	88	88	86	88	86	86	88	77	81	79	74
Number of countries	22	22	22	22	22	22	22	21	21	22	21
Number of instruments	6	8	10	10	10	10	12	10	10	10	10
AR1	-0.686	-0.746	-1.746	0.417	-0.212	-1.681	0.895	0.396	-1.225	-1.460	0.0338
Ar1 p-value	0.493	0.456	0.0809	0.677	0.832	0.0927	0.371	0.692	0.221	0.144	0.973
AR2	0.138	0.605	1.641	0.874	0.0302	0.653	-0.492	0.800	0.0367	0.341	-0.239
AR2 p-value	0.890	0.546	0.101	0.382	0.976	0.514	0.623	0.424	0.971	0.733	0.811
Hansen test	12.20	9.564	8.262	6.263	6.588	9.072	7.179	5.250	11.88	3.139	3.279
Hansen p-value	0.00671	0.0485	0.142	0.281	0.253	0.106	0.305	0.386	0.0365	0.679	0.657

Notes:

The dependent variable is the annual average growth rate of per capita income considering five year intervals (with the exception of the last one which is a four year period), $\ln(y_{i,t})$ is the initial per capita income of each period, $\ln(n_{i,t})$ is the annual average growth rate of population considering five year intervals, the values of all the other variables refer to the initial year of each period. Data for 2004 refers to the year 2004 or the latest available.

Columns (3), (5), (6) and (11) are one-step difference GMM; all the other columns are two-step difference GMM.

Regressions (8) and (11) do not include New Zeland and regression (9) doesn't include Ireland due to lack of data.

AR1 is the Arellano and Bond test for first order autocorrelation; AR2 test for second order autocorrelation.

Hansen J-test tests for overidentifying restrictions.

t statistics in parentheses.

***** p<0.01, ** p<0.05, * p<0.1**

Because the two-step estimates of the standard errors tend to be downward biased (Arellano and Bond, 1991), the standard errors are corrected via a finite-sample correction to the two step covariance matrix derived by Windmeijer (2005)[10].

We used lagged variables as instruments. The choice of the number of the instruments must have in consideration that there is a trade-off between increased efficiency of additional instruments and an aggravation of the weak-instrument problem, if additional time lagged instrumental variables are only weakly correlated with the instrumented covariate (Suhrcke and Urban, 2006). As Roodman (2006) points out, the use of too many instruments, although does not compromise the coefficient estimates, can weaken the Sargan/Hansen test, so the number of instruments should be reported and check the robustness of the results. In order to reduce the number of instruments we use the collapse command[11].

The Hansen statistic of over-identifying restrictions tests the validity of the instruments for the GMM models, assuming in the null hypothesis that the instruments are not correlated with the residuals. In some cases (regressions (1), (2) and (9)) the Hansen test null hypothesis is rejected at 5% significance level, so we should look to the estimation results with some caution.

[10] The command xtabond2 designed by Roodman and available in STATA version 9.2 performs difference GMM (both one-step and two-step) and computes automatically this correction.

[11] This is available from the STATA command xtabond (Roodman, 2006).

The tests of Arellano-Bond (AR) indicate if there are problems with serial correlation of the error terms. Test for AR(2) in first differences assumes in the null hypothesis that the errors in the first difference regression exhibit no second order serial correlation.

Regression (1) of Table 4 shows the results from the growth regression of the Solow's type with physical capital and population growth, assuming that physical capital depreciation plus technical progress is 0.05 as in MRW. We found that the initial per capita income has a negative impact on economic growth, as expected, with 10% statistical significance, confirming therefore, the hypothesis of convergence. The annual growth rate of population has a positive sign but with no statistical significance. Physical capital is the only variable with a positive and significant impact (at 1% level) on economic growth. However, in this regression the omitted variable bias can be a problem since no human capital is considered in this specification.

Column (2) adds human capital to the previous model. Empirical evidence of recent studies (Soukiazis and Cravo, 2008; Soukiazis and Antunes, 2008) show that in the OECD sample the "average years of schooling" is not a suitable proxy for human capital, and that other proxies must be used to capture the efficiency of human capital, such as, scientific production (proxied by the rate of published articles) or activities related to research and innovation (proxied by the rate of registered patents). Therefore, to differentiate better the OECD countries with respect to human capital we considered the "number of patents per million of inhabitants aged 25 or over".

In columns (3) to (11) we analyse the effects of introducing different proxies for health, one at a time, to avoid possible colinearity. One important aspect to notice is that introducing human capital and health conditions in the growth model the convergence coefficient in most cases gains statistical significance, showing that human capital and health are important factors for explaining growth and convergence. On the other hand, most of the health conditions used are shown to be statistically relevant.

At 1% statistical significance level, the variables "avc" (cerebrovascular diseases mortality rate, column 5) and "stay" (average length of stay for acute care, column 8) have a negative impact on economic growth, as expected. Interpreting these results, we can predict that an increase of 1% in the mortality rate caused by cerebrovascular diseases has a negative impact of 0.0924% in the growth of income per capita, everything else remaining constant. This is an important result and of a major concern, reinforcing the alerts given by the World Health Organization (WHO) and other institutions. In fact, according to WHO (2008), stroke and other cerebrovascular diseases were the second main causes of death in 2004 in high-income countries, being responsible for 9.7% of deaths. It is also important to note that this kind of diseases doesn't kill just the elderly; it affects more and more low age groups, with all the drawbacks it has on labour supply, in a context of increasing ageing population in the OECD countries.

On the other hand, it is important to note that cerebrovascular diseases are also highly correlated with other pathology that also characterizes developed countries, diabetes. Although not significant (and so we do not report the results), the estimated coefficient of diabetes has a negative sign, as expected, affecting negatively economic growth. In fact, according to WHO (2008a), diabetes has become one of the major causes of premature illness and death in most countries, mainly through the increased risk of cardiovascular disease (almost 30% of people who die from cerebrovascular or heart diseases have developed these pathologies because they had diabetes). These kind of chronic diseases do represent a serious health problem that OECD countries have to face. As Anderson et al. (2007) point out, five of the most common chronic diseases (including diabetes[12], chronic lower respiratory disease, heart disease, hypertension, cancer, and HIV infection) are responsible for approximately half to two-thirds of deaths in most high-income countries. The combined share of the five diseases is highest in the U.S., accounting for about two-thirds of all deaths in the country. As the authors refer, "if the United States and other OECD countries do not address the growing prevalence of chronic disease, then

[12] In what concerns diabetes, there are projections that this disease will be responsible (for the first time) for a decrease of life expectancy in the next decade.

their health care systems will have difficulty in improving the quality of care and controlling health care spending". Using 2004 OECD health spending data, the authors analyse the burden of chronic disease, which makes up 80 percent of most OECD countries' health care use.

Another significant result from the estimation process (at 1% significance level) concerns the variable that measures the average length of stay for acute care. Our model predicts (from column 8) that a reduction in 1% of the average length of stay is responsible for an increase in per capita income by 0.0919%, everything else being constant. It is more plausible to think that this impact is attributable to the cost reduction policy due to reduction in the length of stay in hospitals. On the other hand, it reveals a kind of resource efficiency associated to the advances made on day surgery that avoids long stay in hospitals. According to OECD (2007), most of the countries of this group registered an important decline in this indicator (average length of stay for acute care has fallen from 8.7 days in 1990 to 6.3 days in 2005 - for the 25 countries for which consistent data over time was available), which can be explained by the advances made on day surgery that avoids long stay in hospitals and to the expansion of early discharge programmes that permits patients to return to go home, receiving follow-up care.

This downward trend in "stay" is also associated with the evolution of the number of acute care beds ("acbed"). In fact, according to OECD (2007), most of the countries of this group show a long-term trend towards a decline in the number of acute care beds: considering a consistent group of 24 countries, the average number of acute care hospital beds dropped from 5.1 per 1000 population in 1990 to 3.9 in 2005. Once more this has been driven, at least partly, by the progress made in medical technology. At the same time, health reforms in OECD countries have been characterised by cost-reducing strategies in the medical sector. At 1% statistical significance level, our results also show (column 11) that a raise (decrease) in "acbed" (acute care beds) has a negative (positive) impact on economic growth: it is predicted that 1% decrease in the number of acute care beds is responsible for 0.0899% increase in income per capita, everything else remained constant.

The variable "physicians" (number of physicians, from column 10) is also statistically significant at the 1% level, meaning that an increase of the number of practising physicians has a positive and significant impact on the standards of living, as expected. It is predicted that 1% increase in the number of practising physicians is responsible for 0.1417% increase in income per capita, everything else remained constant. This is also an encouraging result showing the contribution of labour force employed in the health sector for improving the standards of living of the whole population.

At 5% significance level, infant mortality (column 3) and deaths caused by cancer (column 6) have the predicted negative effect on income growth, while female life expectancy at 65 years old (column 7) and pharmaceutical expenditures (column 9) affect positively the growth of income as expected.

It should be noted that, although usually is not given too much attention to differential gains in what concerns infant mortality rate of developed countries, there are still some important differences: in 2004 the lowest infant mortality rates were reported in Nordic countries and Japan but in USA the same rate was relatively high (more than 6 deaths per 1 000 live births). These differences can be explained in part by the increasing number of premature births (leading to a rising number of babies born with low weight) which is linked to the delay of motherhood decision and to the rise in multiple births by fertility treatments. According to OECD (2008), there is a higher risk of neonatal deaths that has contributed to higher infant mortality rates in some developed countries (like USA) and that can lead to an inversion of the downward trend in infant mortality rates that characterised OECD over the past few years.

As expected cancer has an important negative impact on economic growth of OECD countries, being one of the most important causes of death. According to WHO (2009), cancer is one of the major causes of death worldwide, accounting for 7.4 million deaths (around 13% of all deaths) in 2004, and it is directly linked with lifestyles but also with ageing. Yet, according to the same organization, more than 30% of cancer could be prevented by modifying or avoiding key risk factors. So, it is of extreme importance the implementation of strategies for cancer prevention and early detection of cancer. At the same time, investment in laboratory

research, in technology progress and in treatment and care play a crucial role in controlling this disease.

In what concerns per capita pharmaceutical expenditure, it is important to note that, on average, this ratio has risen more than 50% in real terms since 1995 until 2005 (OECD, 2007). Increased consumption of pharmaceuticals is due mainly to the ageing of population and to the introduction of new drugs. Nevertheless, it should be noted that this variable shows considerable differences across OECD countries, the USA being the highest per capita spender on pharmaceuticals in 2004.

At 10% significance level, life expectancy at birth has a positive effect on income (column 4). This is an expected result since higher life expectancy at birth is an incentive for people to invest more in education and health care and save more for retirement plans.

6. CONCLUSIONS

The aim of this paper was to analyze the impact of health conditions on economic growth, using a growth regression framework. Given the relatively scarce literature that concerns specifically developed countries, we focused our attention on 22 OECD countries. This sample has the advantage of having more reliable available data.

The economic growth regression approach, although has many advantages over accounting approach, has also some limitations. Additional problems have to be taken into account when health factors are considered, mainly because of the endogeneity of the regressors. Following recent developments in economic growth literature, the methodology we used was based on panel data dynamic analysis and estimations were made for the period 1980-2004 using GMM methods.

In our study we used specific proxies for health, different than the traditional ones. Along with infant mortality rate, fertility rate or life expectancy, we also considered mortality rates caused by chronic diseases or variables that measure resources and the efficiency of health systems. These proxies can differentiate more properly health conditions of the developed countries and so they can be more relevant in measuring their impact on economic growth and convergence.

Our empirical analysis initially shows that human capital and health conditions are important conditioning factors for growth and convergence. When these factors are introduced into the growth equation the convergence coefficient gains statistical significance.

Our findings suggest that cerebrovascular and cancer mortality rates, average length of stay, per capita pharmaceutical expenditure, number of physicians and acute care beds are the most significant health factors affecting the standards of living of the developed countries. These results reinforce the idea that analysing the impact of health on economic growth needs to go beyond the use of the most common conventional factors.

Other health conditions, such as infant mortality or life expectancy, have their expected impact on income, but at a lower lever of statistical significance.

Our results suggest some policy implications. Investment in preventing and controlling chronic diseases seems to be of extreme importance. For instance, it is important the implementation of educational policies that may influence lifestyles and contribute to more conscious risk behaviour. On the other hand, given the cost reduction strategies that characterise rich countries' health systems and, at the same time, the need to allocate more resources to control this kind of diseases, it is necessary to evaluate the efficiency of resource allocation in the health sector.

REFERENCES

ALBERT, C. AND DAVIA, M. (2007), Understanding the Effect of Education on Health Across European Countries, available at http://www.ief.es/publicaciones/ PapelesDeTrabajo/pt2007_06.pdf

ANDERSON, G., FROGNER AND E. REINHARDT (2007), Health Care Spending in OECD Countries in 2004: An Update, *Health Affairs*, Sept./Oct.

AMARAL, E. (2007), O Impacto da Educação, da Inovação e da Saúde no Crescimento e na Convergência Económica dos Estados-Membros da União Europeia, *MA thesis*, Faculty of Economics University of Coimbra.

ARELLANO, M. AND BOND, S. (1991) Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies,* 58, 277-297.

ARELLANO, M. (2003), Modelling Optimal Instrumental Variables for Dynamic Panel Data Models, *CEMFI*, Madrid, available at http://www.econ.yale.edu/ seminars/ecm/ecm04/arellano-041005.pdf.

BARRO, R. AND X. SALA-I-MARTIN (1992), Convergence, *Journal of Political Economy* 100: 223-251.

BERALDO, S., MONTOLIO, D. AND G. TURATI (2005), Healthy, Educated and Wealthy: Is the Welfare State Really Harmful for Growth?, available at http://www-1.unipv.it/websiep/wp/245.pdf

BHARGAVA, A., JAMISON, D., LAU, L. AND C. MURRAY (2001), Modelling the Effects of Health on Economic Growth, *GPE Discussion Paper Series 33*, available at http://www.who.int/healthinfo/paper33.pdf.

BLOOM, D. ,CANNING, D. AND J. SEVILLA (2001), The Effect of Health on Economic Growth: Theory and Evidence, *NBER Working Paper 8587*.

BLOOM, D., CANNING, D. AND T. JAMISON (2004), Health, Wealth and Welfare, Finance & Development, 41-1, March, available at http://www.imf.org/external/ pubs/ft /fandd /2004/03/ pdf/ bloom.pdf.

BLOOM, D. AND CANNING, D. (2005), Health and Economic Growth: Reconciling the Micro and Macro Evidence, available at http://www.imf.org/ external/ pubs/ft /fandd/ 2004/03/pdf/bloom.pdf

BLOOM, D. (2007), Education, Health and Development, American Academy of Arts and Sciences, Project on Universal basic and Secondary Education, available at http://www.amacad.org/publications/ubase_edu_health_dev.pdf .

BOND, S. R., A. HOEFFLER, AND J. TEMPLE (2001), GMM Estimation of Empirical Growth Models, *CEPR Discussion Paper* 3048, London.

CAI, L. AND KALB, G. (2006), Health Status and Labour Force Participation: Evidence from Australia, *Health Economics* 15: 241-261.

CASELLI, F., G. ESQUIVEL AND F. LEFORT (1996), "Reopening the Convergence Debate: A New Look at Cross-country Growth Empirics", *Journal of Economic Growth*, 1, 363-389.

CUTLER, D. AND LLERAS-MUNEY, A. (2006), Education and Health: Evaluating Theories and Evidence, National Poverty Center Working Papers Series 06-19, available at http://www.npc.umich.edu/publications/working_papers/.

EUROPEAN COMMISSION (2005), The Contribution of Health to the Economy in the European Union, available at http://www2.eur.nl/bmg/ecuity/public_papers/ ECuity3wp20GambinGenderhealthonincome.pdf .

GROSSMAN, M. (1972), On the Concept of Health and the Demand for Health, *Journal of Political Economy* 80(2): 223-255.

HESTON, A., SUMMERS, R. AND ATEN, B. (2006), Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, available at http://pwt.econ.upenn.edu/php_site/pwt_index.php.

ISLAM, N. (1995), Growth Empirics: A Panel Data Approach, *Quaterly Journal of Economics*,110: 1127-1170.

KNOWLES AND OWEN (1997), Education and health in an Effective-Labour Empirical Growth Model, *The Economic Record*, vol. 73, 223: 314-328.

MANKIW, N., D. ROMER AND D. WEIL (1992), A Contribution to the Empirics of Economic Growth, *Quaterly Journal of Economics* 107: 407-437.

OECD (2007), *Health at a Glance: Health Indicators*, available at http://oberon.sourceoecd.org/vl=1424962/cl=19/nw=1/rpsv/health2007/index.htm (30/12/2008).

OECD (2008), *OECD Health Data 2008*, available at http://www.oecd.org/topicstatsportal.

OECD (2008a), *OECD Factbook 2008: Economic, Environmental and Social Statistics*, available at http://lysander.sourceoecd.org/vl=1275449/cl=20/nw=1/rpsv/factbook/.

OECD (2008b), Infant Mortality available at http://www.oecd.org/dataoecd/ 4/36/40321504.pdf.

RIVERA, B. AND CURRAIS, L. (1999), Economic Growth and Health: Direct Impact or Reverse Causation? *Applied Economic Letters*, 6, 761-764

RIVERA, B. AND CURRAIS, L. (1999a), Income Variation and Health Expenditure: Evidence for OECD Countries, *Review of Development Economics*, 3(3), 258-267.

RIVERA, B. AND CURRAIS, L. (2005), Individual Returns to Health in Brazil, *in* López-Casasnovas *et al.* (Eds.), *Health and Economic Growth: Findings and Policy Implications*, The MIT Press, 287-311

ROODMAN, DAVID (2006), How to Do xtabond2: An Introduction to "Difference" and "System" GMM in Stata, *Center for Global Development, Working Paper 103*.

SCHULTZ, T. PAUL (2005), Productive Benefits of Health: Evidence from Low-Income Countries, *in* López-Casasnovas *et al.* (Eds.), *Health and Economic Growth: Findings and Policy Implications*, The MIT Press, 287-311.

SOUKIAZIS, E. AND CRAVO, T. (2008), Human Capital and the Convergence Process among Countries, *Review of Development Economics*, 12(1), 124-142, May.

SOUKIAZIS, E. AND ANTUNES, M. (2008), Which Combination of Human Capital and Foreign Trade Matters for Growth? An Empirical Approach, *CEUNEUROP*, Discussion Paper 49.

SOUKIAZIS, E. (2008), Capital Humano e Saúde como Factores de Crescimento e de Convergência, Sumário da Lição Síntese no âmbito de Provas de Agregação, Faculdade de Economia da Universidade de Coimbra.

SUHRCKE, M. AND URBAN (2006), Are Cardiovascular Diseases Bad for Economic Growth? available at http://www.aussenwirtschaft.vwl.uni-mainz.de/CVD_paper-20-07-06-CESifo.pdf.

TEMPLE, J. (1999), The New Growth Evidence, *Journal of Economic Literature* 37: 112-156.

TOMPA, E. (2002), The impact of health on productivity: empirical evidence and policy implications, Centre for the Study of Living Standards, Economic Papers, available at http://www.csls.ca/repsp/2/emiletompa.pdf.

WEIL, D. (2006), Accounting for the Effect of Health on Economic Growth, available at http://weblamp.princeton.edu/chw/papers/Weil_David_Accounting_for_the_Effect_of_Health_on_Economic_Growth_Oct_2006.pdf

WHO (2007), Cardiovascular Diseases, *Fact Sheet nr 317*, February, available at http://www.who.int/mediacentre/factsheets/fs317/en/index.html.

WHO (2008), The Top Ten Causes of Death, *Fact Sheet nr 310*, November, available at http://www.who.int/mediacentre/factsheets/fs310/en/index.html.

WHO (2008a), Global Burden of Disease website http://www.who.int/dietphysicalactivity/publications/facts/diabetes/en/

WHO (2009), Cancer, *Fact Sheet nr 297*, February, available at http://www.who.int/mediacentre/factsheets/fs297/en/index.html.

WINDMEIJER (2005), A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators, *Journal of Econometrics* 126: 25-51.

LA ECONOMÍA ESPAÑOLA ANTE LA CRISIS

JUAN IGNACIO CÁCERES RUIZ
E-mail: cacrui@ceu.es
Departamento de Economía Aplicada
Universidad San Pablo Ceu
VIDAL FERNÁNDEZ RODRÍGUEZ
E.mail: vidal.fernandez@esic.es
Business & Marketing School. ESIC

Resumen

La economía Española está en crisis, como el resto de las economías de los países desarrollados. Todos los países se han lanzado a buscar soluciones y a aplicar estrategias, unas veces independientes y otras conjuntas, para aminorar los efectos de la misma. Pero en nuestro caso, se pueden encontrar algunos aspectos que la diferencian del resto y que determinan que los efectos negativos sobre el mercado de trabajo sean mayores y por tanto que la posible salida y recuperación sea más lenta que en el resto.
En el trabajo se analizan éstos aspectos diferenciadores, incidiendo de forma particular en el mercado laboral, revisando de manera crítica las soluciones aportadas y sus efectos para salir de la crisis.

Palabras clave: Crisis, empleo, medidas económicas, soluciones.
Área Temática: Economía del Crecimiento económico.

Abstract

The Spanish economy is under crisis, as the rest of development countries economy. But in our case, there are aspects who made it different, and deserves to revise with attention.

Thematic Area: Growth Economics.

INTRODUCCIÓN.

La economía Española está oficialmente en crisis, aunque el reconocimiento de tal situación no ha sido particularmente fácil. Desde el año 2007, se venían registrando señales, fundamentalmente en el sector financiero, que apuntaban a que la pretendida fortaleza de nuestra economía tal vez no se asentase sobre unas bases tan sólidas como se quería creer.

Al año siguiente, los síntomas se hicieron cada vez más evidentes y contrastaban con la sorprendente negativa de nuestras autoridades económicas a aceptar la realidad de la misma, empeñadas éstas, básicamente, en corregir el alza de precios en el sector de la construcción en lo que se denominó " La burbuja Inmobiliaria". No fue hasta el verano del 2008 cuando el Gobierno empezó a aceptar la idea de la crisis y en el Otoño, ante la inevitable tozudez de las cifras económicas, tuvo que rendirse a la evidencia. Pero para entonces, la situación había cambiado, el entorno financiero internacional entraba en barrena con la denominada " crisis de las subprime" en USA y sus efectos se dejaban sentir en las Bolsas con caídas generalizadas. La nuestra no fue una excepción y a principios de este año 2009 se registraban las mayores caídas en el mercado continuo de toda su historia, situando a la Bolsa en mínimos, afectando fundamentalmente a las empresas Inmobiliarias y a los Bancos, pero también y no en menor medida al sector industrial.

Las consecuencias son de todos conocidas, caída de la actividad, reducción del consumo de las familias, caída de la inversión y en consecuencia incremento de los niveles de desempleo. La

solución de nuestras autoridades económicas se orienta hacia una mayor intervención del Sector Público en la economía y el mantenimiento de los Gastos Sociales, pero la sostenibilidad de esta solución es discutible, tal y como ha puesto sobre el tapete el Gobernador del Banco de España en su comparecencia ante el congreso.

En el proceso, sorprende la enorme velocidad con la que se han ido encajando los distintos sucesos y la generalización de sus efectos. Como es bien conocido, el sector inmobiliario ha sido el gran afectado por la situación, pero no ha sido desde luego el único, puesto que tanto el sector industrial como el de servicios se han resentido claramente.

Hasta fechas muy recientes, uno de los mejores negocios que había en nuestro país era el de la Construcción y sus actividades anexas, las actividades Inmobiliarias y la concesión de créditos, tanto al promotor como hipotecarios, para los Bancos. Estas actividades presentaban además una inercia y un efecto de arrastre importantísimos que se basaban en mi opinión en tres factores :

.- En primer lugar en **Factores Demográficos**. Ha habido un importante cambio en el perfil demográfico de la población española que, además, se ha producido en muy pocos años. Por una parte, ha habido un importante incremento en el número de separaciones y divorcios que junto a la generalización de la soltería ha dado lugar al surgimiento de una nueva clase social que son los denominados " Singles" que han revitalizado la demanda de nuevos pisos. A ello se le ha sumado el empuje de la inmigración, en la que se pueden diferenciar claramente dos sectores. En primer lugar los que podríamos denominar inmigrantes de bajo perfil, básicamente Sudamericanos y Marroquíes, que han ido pasando de alquilar a comprar pisos, situados fundamentalmente en las ciudades aprovechando la bonanza económica generada, entre otros, por el Sector de Construcción. Y, en segundo lugar, los inmigrantes ricos, procedentes de Centro Europa e Inglaterra, que demandan una construcción de mayor nivel situada normalmente en la costa, con importantes repercusiones en los precios de las mismas.

.- En segundo lugar **Factores Financieros** , la evolución de los tipos de interés ha sido a la baja durante los últimos años, destacando especialmente el que los tipos de interés reales llegaron a ser negativos entre 2002 y 2006, los años de mayor expansión en la construcción. La consecuencia del abaratamiento de la financiación fue un crecimiento espectacular tanto de la demanda como de la oferta de nuevas construcciones.

.- En tercer lugar, lo que podríamos denominar el **Efecto Demostración** (J.Duesenberry) la combinación de una financiación barata con una demanda creciente de nuevas construcciones , atrajo al sector de construcción a una gran masa de empresas y particulares decididos a aprovechar la bonanza del sector en el que todo se vendía en poco tiempo y con grandes ganancias. Ello ayudó a generar una espiral de precios cuyas consecuencias se están pagando actualmente, aunque con matices.

Por una parte, las grandes ganancias que se obtenían en poco tiempo atrajo al Sector de Construcción capitales que antes estaban en activos y fondos financieros , animando a invertir en pisos tanto a grandes como a pequeños inversores , fundamentalmente familias. Ello indudablemente repercutía en los precios, pero como todo se vendía con facilidad, las ganancias se reinvertían rápidamente tirando incluso del crédito para ganar más en poco tiempo. Esta subida de precios presionaba a los indecisos que veían cómo mes a mes costaban más los pisos empujándoles a endeudarse cuanto antes, con especial incidencia en los inmigrantes de perfil bajo.

Pero por otra, las medidas impulsadas por el Gobierno para enfriar el sector de Construcción generaron una elevada incertidumbre en el mismo y los tímidos intentos de reorganización acabaron coincidieron con la crisis financiera internacional de tal manera que la suma de ambos efectos ha sido demoledora afectando al sector tanto por el lado de la Oferta como por el de la Demanda.

Por el lado de la demanda podemos diferenciar dos tipos de compradores.

En primer lugar , las familias y pequeños inversores que, a su vez, podemos dividir en dos : Los nuevos compradores que se incorporan ahora, a los que el efecto de la crisis les ha retraído la compra . Aquí se puede apreciar claramente el efecto, en este caso negativo, de las

expectativas, los compradores potenciales siguen necesitando pisos, pero no compran porque están esperando a que los precios bajen y éstos , aunque ya han bajado un 25% , siguen cayendo porque las noticias sobre la construcción siempre son negativas y apuntan hacia nuevas caídas.

Por otro lado, están aquellos que compraron hace algún tiempo y que son los que más están pagando los efectos de la crisis . En primer lugar porque algunos se han encontrado con que las constructoras han quebrado y no tienen su piso, pero tienen que hacer frente al crédito que pidieron al banco. Y por otra parte hay otros a los que les han entregado su piso ahora, en plena crisis, y se encuentran con un activo que vale menos de lo que han pagado por él y sin posibilidades de devolverlo y además tienen que pagar el crédito al banco. Como no pueden venderlo y tienen que pagarlo entran en bancarrota e incluso en este caso se puede hacer otra diferenciación, por una parte estarían la clase media y los pequeños ahorradores que tienen un sueldo fijo o alguna garantía adicional presentada al banco y que se ven ahogados por el crédito. Por otra estarían los inmigrantes de perfil bajo, sin más garantías que su trabajo, a los que la crisis les ha dejado en paro y que, en el mejor de los casos, devuelven las llaves al banco .

En segundo lugar, estarían los grandes inversores, a los que el efecto de la crisis les ha obligado a renegociar deudas e incluso les ha llevado al cierre. Dentro de estos, estarían por una parte las Inmobiliarias, algunas de las cuales han tenido que cerrar, pero que también aprovecharon la fase expansiva para ganar mucho dinero (Astroc, Fadesa, etc.) y por otra parte los Bancos, pero el efecto sobre estos (aparte de alguna Caja de Ahorros con alta exposición al sector) ha quedado limitado, en parte porque muchos de ellos titulizaron sus créditos hipotecarios minimizando sus riesgos y en parte porque algunos otros vendieron a tiempo sus activos inmobiliarios haciendo caja para posteriores operaciones (Santander).

Finalmente, por el lado de la oferta, ha habido una fuerte corrección dentro de las Constructoras. La mayor parte, aprovechó la caída de los tipos de interés reales para captar obra y construir, pero actualmente al entregar la obra y no existir carga de trabajo el sector se ha convertido en una máquina de producir desempleados que, además, debido a la baja cualificación de los mismos tienen muy difícil la recolocación, registrándose en algunos colectivos de inmigrantes bastantes casos de vuelta a sus países de origen. A ello se ha sumado el corte de la financiación por parte de la banca y la exigencia de garantías adicionales que han situado a muchas de las empresas constructoras en una delicada situación financiera (Sacyr) cuando no en la bancarrota.

En todo caso a lo largo de las siguientes páginas haremos un repaso de la situación económica de los países de nuestro entorno para comprender mejor el efecto de internacionalización de la crisis económica para pasar a analizar nuestra coyuntura e incidir sobre los aspectos diferenciadores en nuestro caso, especialmente en materia de desempleo y sus efectos sobre la afiliación a la Seguridad Social tanto en el régimen general como en el de autónomos.

COYUNTURA INTERNACIONAL.

Lo más destacado ha sido la entrada en recesión de la **economía Norteamericana** que por la globalización de los mercados, ha tenido su reflejo inmediato en el resto de economías. La Oficina Nacional de Investigaciones Económicas (NBER) de Estados Unidos ha hecho público la entrada en recesión de su economía desde enero de 2008, antes incluso de que técnicamente (hacen falta dos trimestres seguidos de caída de la producción nacional) entrara en recesión, debido a la notable reducción de la actividad y el empleo a lo largo del año pasado.

ESTADOS UNIDOS: PRINCIPALES INDICADORES ECONÓMICOS

Porcentaje de variación sobre el mismo periodo del año anterior, salvo indicación expresa

	2007	2008	2008						2009	
			I	II	III	Octubre	Noviembre	Diciembre	Enero	Febrero
PIB real	2,0	1,3	2,5	2,1	0,7	–	–0,2	–	–	...
Ventas al por menor	4,2	–0,5	3,1	2,6	0,6	–5,1	–8,4	–10,5	–9,7	...
Confianza del consumidor (1)	103,4	58,0	76,5	57,3	57,3	38,8	44,7	38,6	37,7	25,0
Producción industrial	1,7	–1,8	1,9	0,2	–3,0	–4,3	–5,9	–8,2	–10,0	...
Índice actividad manufacturera (ISM) (1)	51,1	45,5	49,5	49,1	47,4	38,7	36,6	32,9	35,6	...
Venta de viviendas unifamiliares	–26,8	–37,6	–33,1	–39,1	–36,7	–43,8	–38,3	–44,8	...	...
Tasa de paro (2)	4,6	5,8	4,9	5,4	6,1	6,6	6,8	7,2	7,6	...
Precios de consumo	2,9	3,8	4,1	4,4	5,3	3,7	1,1	0,1	0,0	...
Balanza comercial (3)	–700	–677	–699	–702	–712	–713	–695	–677	...	...
Tipo de interés interbancario 3 meses (1)	5,3	2,8	3,0	2,8	3,2	3,0	2,2	1,4	1,2	1,2
Tipo de cambio efectivo nominal (4)	77,9	74,4	72,0	70,9	73,5	80,4	82,7	80,7	81,0	82,4

NOTAS: (1) Valor.
(2) Porcentaje sobre población activa.
(3) Saldo acumulado de 12 meses de la balanza de bienes y servicios. Miles de millones de dólares.
(4) Tipo de cambio ponderado por los flujos del comercio exterior. Valores mayores implican apreciación de la divisa.

Fuente: Servicio de estudios de La Caixa.

En gran parte, esta situación se ha debido a la caída del consumo privado, derivado de la pérdida de más de dos millones de puestos de trabajo que ha empujado la tasa de paro hasta el 7%. Pero es que además, los hogares norteamericanos tenían un nivel de endeudamiento cercano al 100% del PIB invertido en unos activos tanto financieros como reales (casas y otros bienes) que han perdido la mayor parte de su valor, en un contexto financiero en el que los bancos no se prestan dinero entre sí y tampoco al consumidor final por la falta de confianza en su solvencia. Este hecho es muy significativo debido a que el consumo de las economías domésticas supone un 70% del PIB y si no hay consumo, las empresas no venden y reaccionan bajando los precios y despidiendo gente, con lo que el riesgo de deflación se dispara, en especial si tenemos en cuenta que el precio del petróleo se ha ido reduciendo drásticamente en los últimos meses.

La reacción de las autoridades no se ha hecho esperar, la Reserva Federal (FED) redujo hasta en ocho ocasiones el tipo de interés de referencia (al que presta el dinero a los demás bancos) a lo largo de 2008 dejándolo entre el 0 y el 0,25%, el mínimo desde que se creó el Banco Central Norteamericano (1913) y, además, dejó clara la orientación expansiva de su Política Monetaria:

. Que pondrá en juego todas sus herramientas para devolver la estabilidad al crecimiento económico.

. Que los tipos de interés seguirán bajos mientras se mantengan las condiciones actuales.

. Que para apoyar al mercado hipotecario y al sector de construcción, la FED adquirirá deuda y bonos de agencias, cuyas garantías colaterales sean activos hipotecarios.

La consecuencia ha sido la reducción inmediata de todos los tipos de interés (la deuda pública a diez años que estaba dando un 4,3% en Junio ahora está en el 2%) y la depreciación del dólar, lo que ha incrementado su ya abultado déficit exterior.

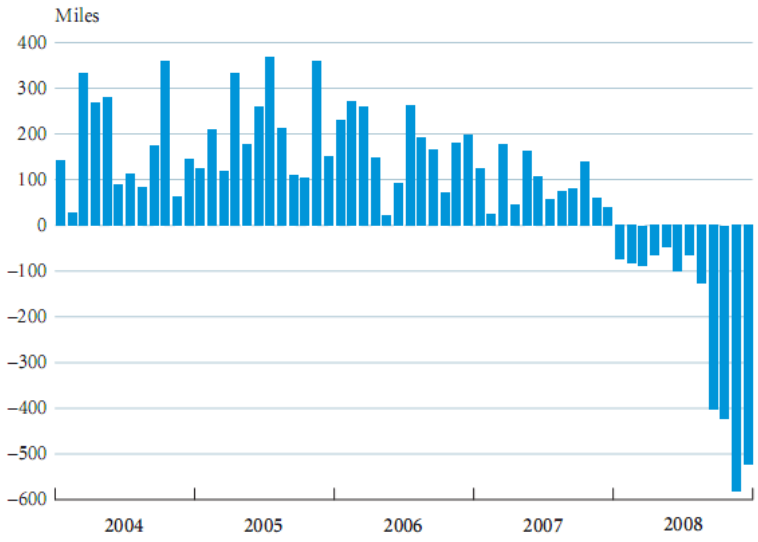
Uno de los temores más extendidos era que se repitiese de nuevo una crisis como la del 29 y si bien es cierto que existen algunas similitudes, como por ejemplo la sobrevaloración de algunos activos, la reacción de las autoridades económicas ha sido muy diferente, tanto en USA como en el resto de economías desarrolladas. En primer lugar, el reconocimiento público de los problemas financieros ha sido rápido y se ha tratado de evitar la quiebra de instituciones relevantes, sobre todo después de la de Lehman Brothers, afianzando la confianza de los depositantes mediante seguros de depósito y medidas monetarias y fiscales que eviten la caída del crédito. Por otro lado, ha habido una coordinación internacional de políticas monetarias que ha llevado a una reducción coordinada de los tipos de interés a ambos lados del Atlántico y a

incrementar la oferta monetaria mediante subastas extraordinarias de liquidez o facilitando el acceso al crédito ampliando los activos de garantía (aquellos que hay que presentar para que los bancos centrales concedan crédito a los bancos comerciales).

A pesar de estos esfuerzos, los consumidores y los empresarios han empeorado sus previsiones de forma drástica, lo que ha lastrado al mercado laboral que en el conjunto del año 2008 ha perdido nada menos que 2,6 millones de empleos (cuando venía de crear más de un millón de nuevos empleos en 2007) y lo más acuciante es que 1,9 millones los ha perdido en los últimos cuatro meses, situando la tasa de paro de Diciembre en el 7,2%, la más alta en dieciséis años. Sectorialmente es muy delicada la situación de las empresas del sector automovilístico, del que dependen más de 5 millones de empleos, el Senado ha concedido ya una ayuda de 14.000 millones de dólares, pero no ha sido suficiente y las tres grandes empresas General Motors, Ford y Chrysler acaban de solicitar una ayuda conjunta de 26.000 millones de Euros para salvar la caída de las ventas. La Construcción que sigue sin tocar fondo porque la pérdida de ingresos de las familias hunde la demanda y por el exceso de oferta existente, gran parte de la cual procede de las subastas de pisos embargados, a pesar de que las viviendas iniciadas están en mínimos históricos.

ESTADOS UNIDOS: DOS MILLONES Y MEDIO DE EMPLEOS PERDIDOS

Creación (+) o destrucción (–) mensual neta de empleo no agrícola



Fuente: La Caixa.

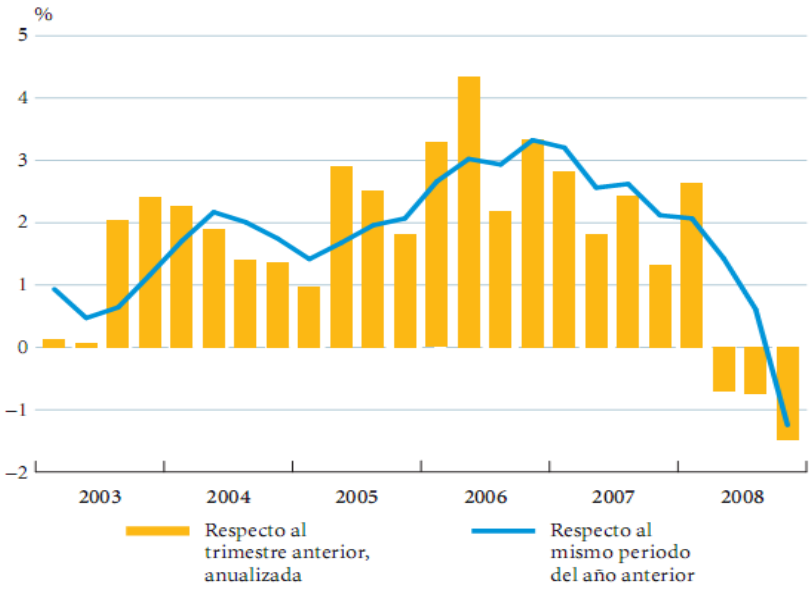
Las medidas monetarias, se han complementado con otros estímulos Fiscales para tratar de contrarrestar el deterioro económico. Así, el recién elegido presidente Obama ha previsto un paquete de ayudas de 825.000 millones de dólares para los años 2.009 -10 (un tercio serán reducciones de impuestos y el resto inversiones públicas). Estas medidas tendrán como consecuencia un incremento del déficit, desde el 5% del PIB con que se cerró 2008, hasta algo más de un 12% que se prevé para el 2009, lo que apunta hacia un ajuste lento de la economía Norteamericana, puesto que la financiación de tal déficit será difícil, dado el estancamiento del consumo interno y la debilidad de la economía global incluidos los países emergentes.

Los problemas de la economía norteamericana están afectando al resto de las economías, con una incidencia especial en **Japón**. Con una demanda interna muy limitada por factores sociológicos, (el ahorro es un casi un dogma) la salida está en el mercado exterior que ha tenido seis años de crecimiento constante, pero la crisis ha acabado en el 2008 arrastrando a la inversión en bienes de equipo, cuya demanda venía fundamentalmente de los bienes duraderos que compran en su mayoría los norteamericanos.

En **China** ocurre algo parecido, la ralentización de las exportaciones deprime la producción industrial que en seis meses ha pasado de crecer al 16% interanual a hacerlo sólo a algo menos del 6%. No obstante su economía sigue creciendo aunque con menos empuje pasando del 13% en 2007 al 6,8% en el último trimestre de 2008. El Gobierno está aplicando medidas fiscales y estructurales con el fin de dinamizar la demanda interna para compensar la reducción del sector exterior.

La Unión Europea ha entrado en recesión por vez primera en su historia. La tasa de crecimiento del PIB a lo largo de 2008 ha sido del 0,9% y se espera para este año una caída del 1,9%. La contracción del crédito a familias y empresas se espera que produzca una disminución de la demanda interna cuyo efecto final será negativo para las economías, pero todavía es pronto para poder estimarlo correctamente. Existe por tanto un elevado nivel de incertidumbre, entre otras cosas porque no se conoce con precisión el alcance de los planes de ayuda de los distintos estados miembros, pero sí que estos planes conllevarán un incumplimiento del Pacto de Estabilidad, es decir, los crecientes déficits públicos conllevarán aparejados un alto servicio de la deuda (los Gobiernos emitirán deuda para financiar el mayor Gasto Público, lo que endeudará a las generaciones futuras y reducirá la renta disponible) que retrasará la recuperación económica.

Variación del producto interior bruto en términos reales

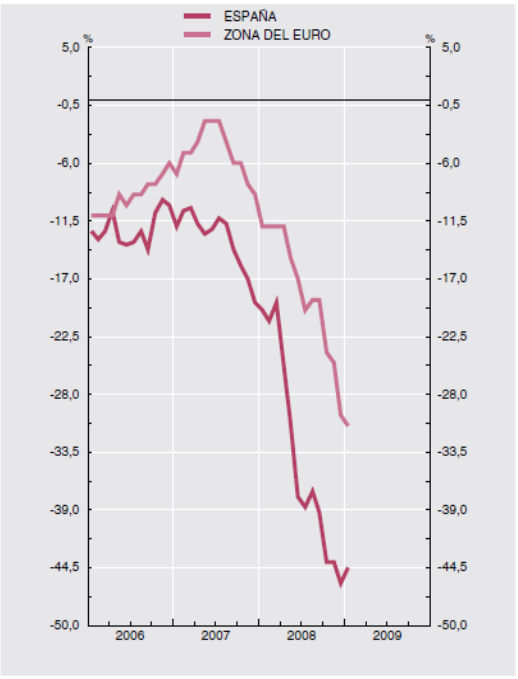


Fuente: La Caixa .

Los indicadores tanto de demanda como de oferta apuntan en esta dirección. Por el lado de la Demanda, en la interna destaca la contracción del consumo privado que pierde más de un tercio del crecimiento alcanzado en el 2007 para situarse en un reducido 0,5% de crecimiento interanual, compensado con la aceleración del consumo público que se situó en un 2,1%. Siendo la Inversión bruta el componente que registra el mayor ajuste situándose en un 0,6% en 2008 frente al 4,1% anterior. Mientras que la demanda externa registró una notable disminución derivada de la fuerte reducción de las exportaciones. En consecuencia, los índices cualitativos de sentimiento económico y de confianza del consumidor se sitúan en mínimos históricos.

Por el lado de la Oferta, los indicadores también reflejas el importante deterioro de las condiciones económicas. Así el Índice de Producción Industrial está acusando la recesión y la caída de la demanda interna, registrando una acelerada evolución negativa que lo sitúa al mes de Diciembre en el -12%. El resultado es que la utilización de la capacidad productiva ha pasado en el último trimestre al 75%, cuando siempre ha oscilado entre el 80-85%. Y lo mismo ocurre en los sectores de construcción y servicios.

ÍNDICE DE CONFIANZA CONSUMIDORES



ÍNDICE DE PRODUCCIÓN INDUSTRIAL, TOTAL
(Tendencia obtenida con TRAMO-SEATS)

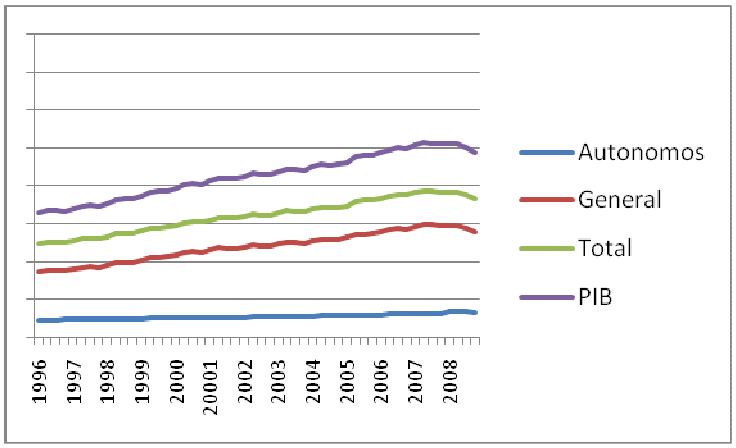


FUENTES: INE y BCE.

El deterioro de los mercados, está reflejándose en las cifras de desempleo que están teniendo un importante incremento situándose en un 7,8% de media, aunque se registran notables diferencias por países, teniendo, por desgracia, el nuestro un destacado papel, que analizaremos con mayor detalle más adelante.

Finalmente, los precios están son tal vez el único de los indicadores importantes que está teniendo un comportamiento positivo apoyado en la disminución del precio del petróleo, situándose la tasa de inflación armonizada en un 1,6% interanual en diciembre, 2,5 puntos porcentuales menos que seis meses antes. Este alivio inflacionista está permitiendo al Banco Central Europeo rebajar el tipo de interés al 2% y aplicar una política monetaria expansiva que apoye la recuperación de la demanda de las familias y empresas. Adicionalmente, se aprobó en diciembre un conjunto de medidas de carácter fiscal, sin perjuicio de las ayudas que decida conceder cada estado, por importe de 200.000 millones de euros de los que 170.000 corresponden a los países integrantes y el resto al presupuesto comunitario dirigido a Pymes e infraestructuras. A este respecto, se pueden diferenciar los planes de estímulo fiscal en función de la importancia del Sector de Construcción en su economía. Los países en los que éste ha tenido más importancia (España, Inglaterra o Irlanda) han aplicado Políticas Fiscales más expansivas, porque el efecto de la crisis financiera ha sido mayor, mientras que en el resto, sus medidas fiscales son menos activas.

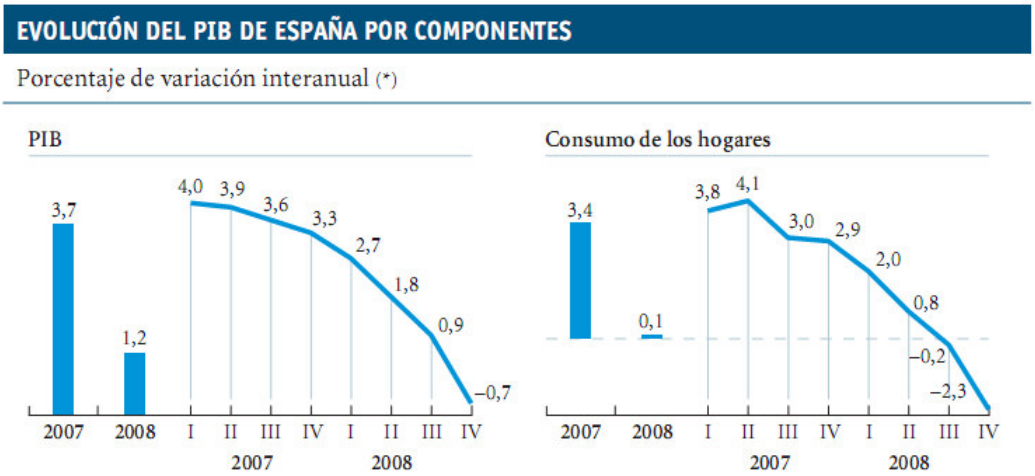
En **España,** el Gobierno se vio obligado a revisar las previsiones establecidas en los presupuestos generales pasando de un crecimiento estimado del 1% en el PIB a una disminución del 1,6% para el año 2009, previsión que ha sido ampliada por la Comisión Europea y otros organismos, como el FMI hasta el -2 % y el -1,7 %. Es decir, que ambos organismos prevén una contracción de la economía mayor y que en consecuencia el efecto sobre el empleo y el déficit público será mayor del estimado.



La evolución de los indicadores económicos no ha sido muy diferente de la comentada con anterioridad para el resto de países, salvo tal vez en el capítulo del empleo cuya evolución ha sido peor en nuestra economía. En concreto, los componentes del PIB han tenido una notable reducción tanto desde el punto de vista de la Demanda como desde el punto de vista de la Oferta en la que tanto la construcción como la industria y los servicios han registrado una evolución negativa.

A pesar de ello, la información recogida por el Banco de España para las empresas no financieras apunta hacia una situación económica buena aunque con claras señales de deterioro, de hecho los gastos de personal registraron una subida media del 5% con relación a 2007, que junto a los mayores gastos financieros determinaron una caída del resultado operativo ordinario compensado con un crecimiento de los resultados extraordinarios superior al 100% que, al final, permitieron una mejora del resultado neto superior al 25%. Es decir que para las empresas de la muestra el año 2008 ha sido malo, pero a pesar de todo la media ha ganado dinero.

Por el lado de la Demanda, el consumo las familias se redujo a mínimos históricos en especial el de bienes duraderos , como los automóviles que registraron una caída del 50 % respecto al año anterior, el peor dato en diez años. Ello se ha debido tanto al endurecimiento de las condiciones de financiación por parte de la banca como a pérdida de confianza de los consumidores por la crisis y la elevación del paro, como recoge el índice de confianza de los consumidores que se redujo más de treinta puntos porcentuales en el último trimestre.



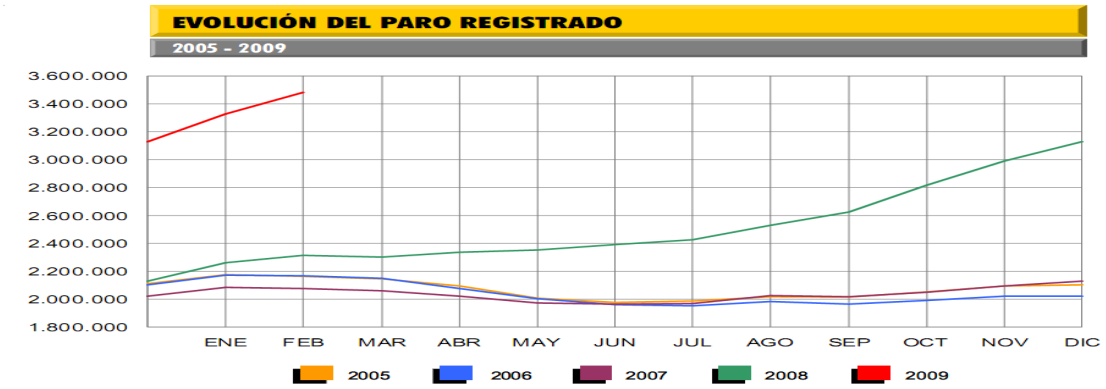
Fuente : La Caixa.

La producción y las ventas de bienes de equipo también registraron resultados negativos con caídas superiores al 15% en el índice de producción industrial, aunque sin alcanzar las habidas en el consumo. Pero el sector que ha sufrido el ajuste más riguroso ha sido sin duda la

Construcción, ya dentro de la Oferta productiva. Todos sus indicadores han registrado mínimos desde el consumo aparente de cemento, indicador adelantado del sector que cayó más del 40% en el último trimestre, a las ventas de viviendas tanto nuevas como de segunda mano con disminuciones en torno al 30% ; o los visados de obra nueva que se han reducido más del 50%. Incluso el Sector servicios ha registrado una disminución en torno al 15% en su cifra de negocio. Dentro de este, las mayores caídas se registraron en el comercio de vehículos y en las empresas de contratación de personal, aunque otros segmentos como el turismo o el tráfico de pasajeros registran disminuciones superiores al 20%.

Esta negativa evolución tiene su reflejo en el aumento de la morosidad que ha crecido más de un 35% en 2008 y en la creación de empresas que se ha reducido en torno al 40% con el efecto inducido que ello supone para el empleo y que analizaremos más en detalle a continuación.

El **Mercado de trabajo** ha sufrido una fuerte corrección en el ejercicio 2008 que se saldó con la pérdida de algo más de 841.000 empleos en el conjunto del año, lo que supuso una caída del 4,3 % en la tasa de ocupación interanual, la más alta desde que se disponen de estadísticas de empleo. Ello supone un promedio mensual de destrucción de empleo a lo largo del año superior a 70.000 empleos, cifra que en los dos primeros meses de este año 2009 se ha incrementado notablemente, perdiéndose 419.000 afiliados más . En concreto según el informe de Afiliados a la Seguridad Social de Febrero de 2009, a final de este mes se han perdido más de 1.139.000 afiliados respecto al año anterior, lo que supone un 5,9 % de caída en el total de filiados, su evolución gráfica se recoge a continuación.



Por Regímenes de afiliación, el General, arroja una disminución del 7,20% algo más de 1.000.000 de personas, seguido del régimen de Autónomos cuya afiliación se reduce en conjunto unas 135.000 personas, casi un 4%. El resto registran evoluciones dispares, creciendo la afiliación un 6% en los regímenes Agrario y en el de Hogar; mientras que se reduce en torno a un 4,5% en los de Mar y el de Carbón.

Por Género, la caída en la ocupación se concentra en los hombres con una reducción interanual media en el año 2008 cercana al 7%, que a Febrero se amplía hasta el 8,6%, posiblemente por su mayor presencia en la Construcción; mientras que la de las mujeres se sitúa en el 2,2%. Ello ha determinado que el paro registrado se haya concentrado mucho más en los hombres con un crecimiento interanual del 78% al mes de Diciembre, mientras que el de las mujeres ha sido bastante menor, cercano al 25%. En términos absolutos, la tasa de paro se situó al final de 2008 en algo más de 3.100.000 desempleados, cifra que en los dos primeros meses del 2009 ha subido a 3,5 millones, que se reparten casi al 50% entre hombres1,6 y mujeres 1,5 y que al mes de Febrero casi roza los 3,5 millones de parados, registrando la mayor subida histórica y comprometiendo el objetivo del Gobierno que ya se plantea que el año 2009 acabe por encima de los 4 millones de desempleados.

FEBRERO 2009

PARO REGISTRADO POR SEXO, EDAD Y SECTOR DE ACTIVIDAD ECONÓMICA

	TOTAL	VARIACIÓN			
		MENSUAL		ANUAL	
		ABSOLUTA	RELATIVA	ABSOLUTA	RELATIVA
TOTAL	3.481.859	154.058	4,63	1.166.528	50,38
SEXO:					
HOMBRES	1.755.969	82.734	4,94	791.736	82,11
MUJERES	1.725.890	71.324	4,31	374.792	27,74
MENORES DE 25 AÑOS:					
HOMBRES	259.578	14.707	6,01	120.674	86,88
MUJERES	182.925	12.955	7,62	50.506	38,14
TOTAL	442.503	27.662	6,67	171.180	63,09
MAYORES DE 25 AÑOS:					
HOMBRES	1.496.391	68.027	4,76	671.062	81,31
MUJERES	1.542.965	58.369	3,93	324.286	26,61
TOTAL	3.039.356	126.396	4,34	995.348	48,70
SECTOR:					
AGRICULTURA	114.215	7.339	6,87	29.996	35,62
INDUSTRIA	453.491	22.343	5,18	162.506	55,85
CONSTRUCCIÓN	632.505	24.600	4,05	327.242	107,20
SERVICIOS	2.001.314	88.654	4,64	595.200	42,33
SIN EMPLEO ANTERIOR	280.334	11.122	4,13	51.584	22,55

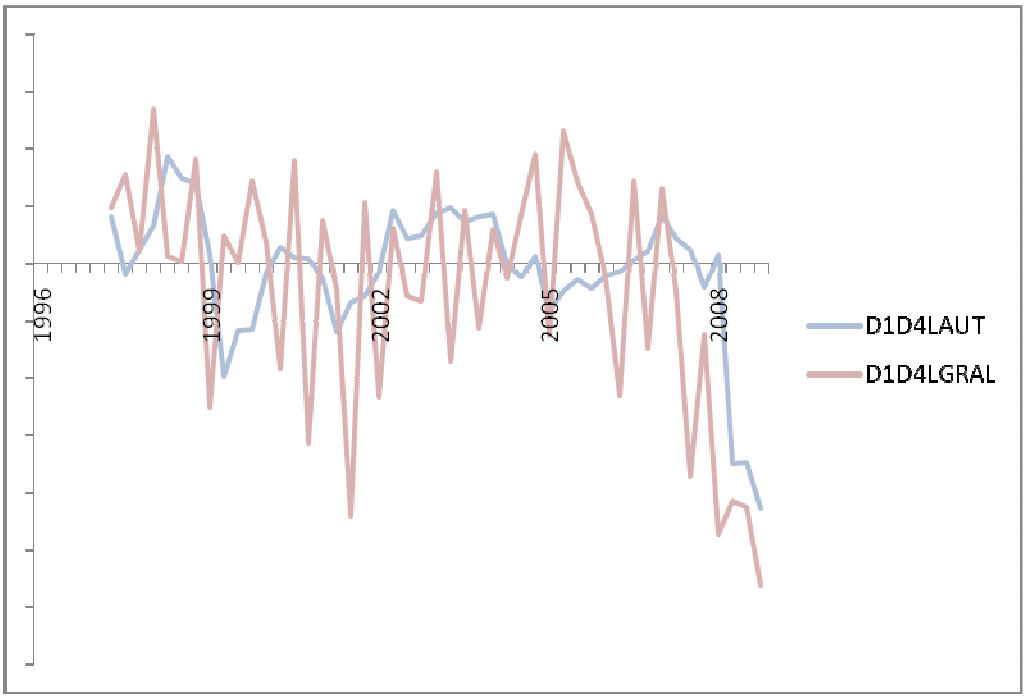
Fuente: Ministerio de Trabajo: Informe demandantes de empleo. Febrero 2009.

Por Sectores, aunque la crisis se hace notar en todos, las mayores reducciones se han registrado en la Construcción en donde la tasa de afiliación a la Seguridad Social ha pasado de un crecimiento algo superior al 1% en 2007 a una disminución superior al 10% en 2008, concentrada fundamentalmente en el último trimestre del año en que se redujo casi un 60%. En consecuencia, ha sido este sector el que ha concentrado el mayor aumento del desempleo, que se incrementó más del 100% (107%) a lo largo del ejercicio 2008, mientras que en el resto de sectores aunque también se incrementó, sus tasas de variación son sensiblemente menores ; 38% en los Servicios, 43% en la Industria y 47% en la Agricultura .

Evolución económica y afiliación a la Seguridad Social.

Ante la difícil situación que enfrenta la economía Española, hemos querido observar el comportamiento de la afiliación a la Seguridad Social a lo largo de los últimos años para ver si es posible determinar alguna relación entre ésta y la evolución de la economía, medida ésta por el PIB.

Para ello, hemos desarrollado dos modelos analíticos utilizando las series temporales publicadas por el INE correspondientes a la actividad de la economía (PIB) y las de afiliación a la Seguridad Social tanto en el régimen general como en el de autónomos entre los años 1996 y 2009. El primer modelo consiste en una regresión que nos permite estudiar el efecto que una variación en la tasa de actividad de la economía provoca en el número de afiliados, tanto en el régimen general como en el de autónomos.



En el caso del régimen general, el coeficiente asociado es significativamente menor que en el caso de los autónomos, indicando con ello que la elasticidad al crecimiento de la economía es menor en el régimen general que en el autónomo y por lo tanto al crecer la actividad en los últimos años la respuesta de los autónomos ha sido más elástica y ha permitido generar proporcionalmente más empleo en éste régimen que en el general por cada punto porcentual de crecimiento del PIB.

El modelo también pone de manifiesto lo contrario que, ante una caída del PIB, las variables cambian de signo, lo que indica que en épocas de recesión se reducen más empleos en el régimen autónomo que en el general, lo que es consecuente con su mayor flexibilidad. No obstante, el modelo no proporciona una información cualitativa del efecto que se produce en cada régimen por separado.

Por ello, se ha desarrollado un modelo tipo VAR con tres ecuaciones en las que las variables dependientes son los regímenes de afiliación, general y autónomo, y el Producto Interior Bruto, como medida del nivel de actividad de la economía. Como variables independientes se utilizan las mismas variables, pero con un periodo de retardo (véase anexo). Estimado el modelo conjuntamente, se observa un comportamiento más explicativo que el modelo anterior.

En efecto, al igual que en el primer modelo, cuando la tasa de variación interanual del PIB es positiva, el régimen de autónomos crece pero crece menos que el régimen general. Sin embargo, cuando se reduce la tasa de actividad, la afiliación en el régimen de autónomos cae menos que la del régimen general, lo que significa que el régimen general presenta una rigidez mayor a las variaciones de la actividad y por tanto en épocas expansivas crece más que el de autónomos, y en épocas recesivas cae más que éste, por tanto queda significativamente contrastado que la rigidez del régimen general expone más a los trabajadores a las oscilaciones de la actividad económica.

EVOLUCIÓN DE LA AFILIACIÓN POR REGÍMENES

	General	Autónomos	Agrario	Mar	Carbón	Hogar	Total
Medias Anuales							
Año 2001	11.697.069	2.614.044	1.127.395	78.370	16.607	156.019	**15.689.504**
Año 2002	12.126.322	2.654.749	1.123.529	76.577	14.917	175.361	**16.171.455**
Año 2003	12.540.535	2.730.806	1.136.210	76.146	13.493	185.250	**16.682.440**
Año 2004	12.948.693	2.837.533	1.089.774	74.974	12.030	181.214	**17.144.218**
Año 2005	13.561.695	2.933.628	1.046.062	73.289	10.512	279.066	**17.904.252**
Año 2006	14.226.490	3.017.742	1.003.215	72.079	9.405	338.668	**18.667.599**
Año 2007	14.783.305	3.119.916	971.531	71.183	8.683	277.369	**19.231.986**
Año 2008	14.650.273	3.383.685	744.353	69.740	7.988	280.018	**19.136.055**
Medias mensuales 2009							
enero	13.755.624	3.287.286	781.166	63.685	7.598	286.384	**18.181.743**
febrero	**13.700.543**	**3.263.872**	**787.313**	**65.443**	**7.631**	**287.808**	**18.112.611**

Fuente: Ministerio de trabajo: Afiliados a la Seguridad Social, Febrero 2009.

En resumen, los resultados del modelo demuestran empíricamente que el régimen de autónomos tiene un comportamiento más flexible que el general, en especial en las fases recesivas, en las que se muestra menos afectado en términos de destrucción de empleo. En concreto los datos del informe de Afiliados a la seguridad social del mes de febrero de 2009 ponen de manifiesto este hecho, desde Septiembre de 2007 la tasa de variación interanual de la afiliación al régimen general presenta una tendencia decreciente , al igual que la de los autónomos, pero el ritmo de caída de la primera es mucho más rápido hasta el extremo que desde Junio de 2008 el régimen general presenta ya una tasa de variación interanual negativa mientras que el de autónomos no registra variaciones negativas hasta enero de 2009 y además son la mitad de las que se registran en el régimen general.

Anexo.

En primer lugar, se ha llevado a cabo un contraste de raíces unitarias, el contraste ADF (Aumented Dick-Fuller), que nos permite estudiar la hipótesis de cointegración de las variables del modelo:

Unit-root tests
The sample is: 1996(1) - 2008(4)
RG: ADF tests (Constant+Trend; 5%=-3.42 1%=-3.98)

D-lag	t-adf	beta Y_1	sigma	t-DY_lag	t-prob	AIC	F-prob
4	-4.955**	0.89094	2.139	10.75	0.0000	1.535	0.0000
3	-2.973	0.92822	2.376	-5.580	0.0000	1.743	0.0000
2	-4.053**	0.90123	2.448	-4.275	0.0000	1.801	0.0000
1	-5.140**	0.87625	2.491	-2.437	0.0152	1.834	0.0000
0	-5.997**	0.86034	2.504			1.842	0.0000

Unit-root tests
The sample is: 1996(1) - 2008(4)
RAU: ADF tests (Constant+Trend; 5%=-3.45 1%=-4.05)

D-lag	t-adf	beta Y_1	sigma	t-DY_lag	t-prob	AIC	F-prob
4	-2.916	0.38383	3.904e+005	-7.518	0.0000	25.86	0.0000
3	-6.503**	-0.46147	4.905e+005	0.5044	0.6151	26.30	0.0000
2	-7.949**	-0.39072	4.886e+005	1.839	0.0690	26.29	0.0000
1	-8.924**	-0.17575	4.945e+005	2.058	0.0422	26.30	0.0000
0	-10.92**	0.026625	5.025e+005	6.015	0.0000	26.33	0.0000

Como se puede observar, el estadístico t-adf muestra un valor significativo en todos los casos, con valores críticos inferiores al 5% y al 1%, lo que muestra que las series están cointegradas y nos permite desarrollar los modelos con las variables en niveles.

Los modelos estimados que proponemos tienen como variable dependiente el número de afiliados a la Seguridad Social, en cada uno de los regímenes, RG_t y RA_t, y como variables independientes el Producto Interior Bruto, PIB_t, una variable artificial, I_t, que toma el valor 1 en las fases expansivas y 0 en los periodos de contracción, y los efectos de interacción. Los resultados obtenidos son:

$$\text{Modelo 1 : } RG_t = \beta_0 + \beta_1 PIB_t + \beta_2 I_t + \beta_3 PIB_t * I_t + \varepsilon_t$$

Estadísticas de la regresión	
Coeficiente de determinación R^2	0,8339906

	Coeficientes	*Error típico*	*Estadístico t*	*Probabilidad*
Constante	6,99028E-05	0,001421398	0,049178904	0,961004448
β_1	0,002088035	0,001904398	4,096428099	0,002789942
β_2	0,056673777	0,278608361	3,203417358	0,048397684
β_3	0,37077977	0,341975168	2,084230099	0,028430469

$$\text{Modelo 2 : } RA_t = \alpha_0 + \alpha_1 PIB_t + \alpha_2 I_t + \beta\alpha_3 PIB_t * I_t + a_t$$

Estadísticas de la regresión	
Coeficiente de determinación R^2	0,70012996

	Coeficientes	*Error típico*	*Estadístico t*	*Probabilidad*
Constante	-0,001167769	0,319789715	0,2043469	1,564935488
β_1	0,001278346	12,67032916	-16,1886448	2,66055E-21
β_2	0,141974391	1,16941E-05	8,028684853	1,72712E-10
β_3	0,144675871	0,2043469	1,564935488	0,12476262

El modelo tipo VAR nos permite expresar los efectos dinámicos de dos o más variables dependientes, respecto de ellas mismas en diferentes instantes del tiempo.

En nuestro caso, se ha estimado el siguiente modelo:

$$RG(t) = \boldsymbol{a_{10}} + \boldsymbol{a_{11}}RG(t-1) + \boldsymbol{a_{12}}RA(t-1) + \boldsymbol{a_{13}}PIB(t-1) + U_1(t)$$
$$RA(t) = \boldsymbol{a_{20}} + \boldsymbol{a_{21}}RA(t-1) + \boldsymbol{a_{22}}RG(t-1) + \boldsymbol{a_{23}}PIB(t-1) + U_2(t)$$
$$PIB(t) = \boldsymbol{a_{30}} + \boldsymbol{a_{31}}RA(t-1) + \boldsymbol{a_{32}}RG(t-1) + \boldsymbol{a_{33}}PIB(t-1) + U_3(t)$$

en el que:
RG (t), es la serie de afiliados al Régimen General de la Seguridad Social, en el periodo t.
RA (t), es la serie de afiliados al Régimen Autónomo de la Seguridad Social en el periodo t.
PIB (t), es la serie de Producto Interior Bruto, en el periodo t.

Todas las series se introducen en el modelo, una vez corregidas de variación estacional. La estimación del sistema nos proporciona los coeficientes a_{ij}, que nos indican el peso que tiene el PIB en cada uno de los regímenes de afiliación a la Seguridad Social.

Modelling VAR by Markquardt Algorithm
The estimation sample is: 1996(1) - 2008(4)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	10.1442	0.4341	23.4	0.0000	0.5230
RGt	0.0480852	0.001502	32.0	0.0000	0.6731
RGt-1		0,0461	0.001235		0.0032
RAt		0,0345	0.002342		0.0044
RAt-1		0,0546	0.001698		0.0078
PIBt		0.3723	0.01767		0.0000
PIBt-1	0.2344	0.01934		0.0000	

sigma	4.84658	RSS	11697.7124
R^2	0.673092	F(1,498) =	1025 [0.000]**
log-likelihood	-1497.6		
no. of observations	141	no. of parameters	9
mean(Y)	22.1896	var(Y)	2123333

Towards an Inclusive Model of Sustainable Growth: Externalities and Distortions*

Catarina Roseta-Palma† Alexandra Ferreira-Lopes‡
Tiago Neves Sequeira§

March, 2009

Abstract

Models of economic growth are typically based on the use of one or more stocks of productive assets to create goods for utility-generating consumption. The roles played by man-made capital, natural capital, and human capital have been explored, often separately, in the literature, and more recently the notion of social capital has been brought to the fore. This paper provides an attempt to construct an inclusive model of growth, analysing externalities and distortions due to market failures linked with social capital and environmental problems.

JEL Classification: O13, O15, O41, Q20.

Keywords: Human Capital, Social Capital, Natural Capital, Produced Capital, Economic Growth, Sustainability.

*The authors thank participants at the 2^{nd} Annual Meeting of the Portuguese Economic Journal and at the XVI Conference of the European Association of Environmental and Resource Economists, as well as at seminars at Instituto Superior de Economia e Gestão (ISEG-UTL) and Fundación de Estudios de Economia Aplicada (FEDEA), for helpful comments and discussions.

†corresponding author; ISCTE (Lisbon University Institute), Economics Department, ERC-UNIDE, and DINÂMIA. Av. Forças Armadas, 1649-026 Lisboa, Portugal. Tel.: +351 217903438; Fax: +351 217903933; e-mail: catarina.roseta@iscte.pt. Catarina Roseta-Palma acknowledges financial support from FCT, scholarship SFRH / BSAB / 861 / 2008.

‡ISCTE (Lisbon University Institute), Economics Department, ERC-UNIDE, and DINÂMIA. Av. Forças Armadas, 1649-026 Lisboa, Portugal. Tel.: +351 217903901; Fax: +351 217903933; e-mail: alexandra.ferreira.lopes@iscte.pt. Alexandra Ferreira-Lopes acknowledges the financial support from FCT.

§UBI and INOVA, UNL. E-mail: sequeira@ubi.pt. Management and Economics Department. Universidade da Beira Interior. Estrada do Sineiro, 6200-209 Covilhã, Portugal. Tiago Neves Sequeira acknowledges the financial support from POCI/FCT.

1 Introduction

Models of economic growth are typically based on the use of one or more stocks of productive assets that are drawn upon to create goods for utility-generating consumption. The role of natural resources or, more generally, environmental assets has been thoroughly analysed by many authors, of which Smulders [29] gives an excellent non-technical summary. Whether non-renewable or renewable resources are considered, the crucial factors initially seen to determine the possibility of growth (in the sense of a sustained increase in consumption) were the substitutability between natural resources and produced capital, and the drive provided by technical change. Additionally, environmental resources often have a public good character, so that without proper policies markets are unable to ensure their adequate allocation, thus jeopardizing growth opportunities.

Outside the realm of environmental economics, growth theory has emphasized the contribution of human capital, or knowledge, which unlike produced capital is not necessarily subject to diminishing returns. When such models do take natural capital into account, knowledge can supply the thrust to keep consumption rising in spite of the bounds placed on humanity by the natural environment. As noted by Rodrigues et al [25], the "material" side of the economy can stop growing, achieving a biophysical steady state, while the intellectual side continues to rise "at the pace allowed by knowledge formation and dematerialization".

More recently, the notion of social capital has been brought to the fore as a potential source of economic growth. The most usual definition is from Putnam [22], who presents social capital as the "features of social organization, such as trust, norms, and networks, that can improve the efficiency of society by facilitating co-ordinated actions." Several empirical studies attempt to quantify the contribution of social capital to growth. For example, Knack and Keefer [17] establish a causal relationship between trust and growth. Controlling for other variables, they find that a level of trust that is 10 pp higher is associated with an annual growth rate that is higher by 0.8 pp. Temple and Johnson [31] use an index composed of

several measures of social capital, and find it a useful predictor of economic growth. Other empirical studies estimate a robust relationship between social capital and growth (e.g. Beugelsdijk et al [3]; Whiteley [32]; and Rupasinga [26]), although the estimates vary widely.

A few theoretical developments have been proposed, namely Beugelsdijk and Smulders [4] where social capital is modelled as participation in two distinct networks, a closed one (family and friends) and an open one that bridges across different communities. The model is tested using data from the European Values Survey and the conclusion is that higher values attached to family life reduce output growth. Most authors model social capital as an asset that contributes to production (Glaeser et al [13], Bisin and Guaitoli [5]). Few papers explicitly model the links between social and human capital and, to our knowledge, none consider natural capital as well, even if there are many empirical studies analyzing the links between economic growth, natural resource abundance, education, and institutions (see, for example, Gylfason and Zoega [15], Brunnschweiler and Bulte [6], and Constantini and Monni [8]).[1]

The importance of developing a complete asset-based framework is highlighted when the focus is on achieving sustainable development. Pearce et al [20] argue intuitively that the value of changes in an economy's aggregate capital stock can be seen as a measure of sustainable development, since it represents the total wealth bequeathed to the next generations. Thus, a decreasing stock of assets (negative genuine savings) is a sign of unsustainability, although the converse isn't necessarily true (see Pezzey and Toman [21] for a more detailed exposition).

In spite of serious data shortcomings, the World Bank has attempted an empirical estimation of the value of wealth for a large set of countries [34]. The report clearly acknowledges the above-mentioned connection between wealth and sustainability, for instance noting in the Foreword that the estimation "yields important insights into the prospects for sustainable development in countries around the world". Direct estimates are provided

[1]Exceptions are Bisin and Guaitoli [5], Sequeira and Ferreira-Lopes [27], and Dinda [10].

for produced capital and natural capital (although only some components are included), whereas intangible capital, which includes human and institutional factors, is obtained indirectly. One of the main conclusions is that the intangible component almost always accounts for the largest share of countries' wealth.

In this paper we present a general model of the relationships between all relevant assets, i.e. natural capital, produced capital, social capital, and human capital, in an attempt to provide an inclusive theoretical view of their potential contributions to economic growth in an endogenous growth framework.[2] The goal is to contribute to the development of analytical tools that yield useful policy implications. Moreover, we calibrate the model using available empirical data and show the distortions arising in the decentralized market economy. We conclude that the presence of an environmental asset in the model implies non-negligible distortions in the human capital allocations between sectors. The following section presents the model, whereas section 3 shows the main relationships between variables from a social planner's point of view. Section 4 presents the decentralized equilibrium. Section 5 presents the allocations of human capital and other macroeconomic variables and shows the distortions in the market solution. In section 6 we calibrate the model and quantify the distortions presented in the previous section, whereas section 7 concludes.

2 Model

In this section the several types of capital mentioned in the Introduction are modelled, along with the assumptions on their evolution. The accumulation of produced capital, K_P, arises through production that is not consumed, and is, as usual, subject to depreciation:

$$\dot{K}_P = Y - C - \delta_P K_P \tag{1}$$

where Y denotes production of final goods, C is consumption, and δ_P represents depreciation.

[2] A recent short paper by Dasgupta [9] also uses the term inclusive wealth to denote the social value of a country's productive base.

We propose that human capital K_H is produced using human capital allocated to schooling as well as the total amount of social capital, Ks, according to:

$$\dot{K}_H = \xi H_H + \alpha K_S - \delta_H K_H \tag{2}$$

where H_H are school hours, $\xi > 0$ is a parameter that measures productivity inside schools, $\alpha \geq 0$ measures the sensitivity of human capital accumulation to the stock of social capital, and $\delta_H \geq 0$ is the depreciation of human capital. This expression captures the idea of Coleman [7] and Teachman et al [30] according to which social capital is important to the production of human capital. It also indicates that human and social capital are substitutes in the production of human capital.

Human capital is an essential part of the model as it is the main driver for growth. To highlight its many roles, we divide human capital into several components, thus indicating that the stock of knowledge and skills at any given time can be put to different uses, namely: final good production (H_Y), school attendance (H_H), social capital accumulation (H_S), and environmental research or pollution prevention (H_N). Assuming that the different human capital activities aren't done cumulatively, we have:

$$K_H = H_Y + H_H + H_S + H_N \tag{3}$$

Social capital accumulation requires human capital to be allocated to its production but at each point in time it will also depend on the current stock of social capital, i.e.:

$$\dot{K}_S = \omega H_S + \Omega K_S \tag{4}$$

where ω measures the productivity of human capital in the production of social capital and $\Omega \lesseqgtr 0$ measures the dynamic effect of social capital on its own production. If $\Omega > 0$ the existing stock of social capital is strong enough to keep growing without additional human capital. If $\Omega < 0$, on the other hand, there is a net depreciation effect.

Natural capital, K_N, can be thought of as an aggregate measure of natural amenities composed of all environmental assets, including traditional

natural resources, waste disposal, and environmental services. This stock variable behaves like a renewable resource, with:

$$\dot{K}_N = R(K_N) + P - N_Y \tag{5}$$

where $R(K_N)$ is the regeneration rate and P represents the positive effects of environmental protection and N_Y represents resource depletion due to its use in production. Regeneration is assumed to follow a logistic function, attaining a zero value when the stock is zero and when it approaches its carrying capacity, CC. This rules out the unrealistic possibility of infinitely growing natural capital. A possible functional form is:

$$R(K_N) = iK_N(CC - K_N) \tag{6}$$

where i is a growth parameter and CC is constant (for a more complex model where carrying capacity depends on stock size, see Rodrigues et al [25]). We also assume that P is endogenously determined. In particular, it is expected to depend negatively on the stock of produced capital and positively on the allocation of human capital to environmental protection (or pollution prevention). We also assume that improvements in social capital (K_S) increases environmental awareness, to capture the idea that higher levels of trust or institutional quality reinforce environmental protection.[3] Thus, environmental protection has the following expression:

$$P = m_0 \frac{H_N^{\epsilon} K_S^{\kappa}}{K_P^{\varphi}} \tag{7}$$

where m_0 is a scale factor, φ captures the elasticity of environmental deterioration with respect to phyisical capital, and ϵ and κ are parameters for the impact of environmental research and social capital, respectively. The total stock of natural capital can be divided into natural capital used in production (N_Y), which is subtracted in equation (5), and the remaining

[3] A similar effect is considered in Economides and Phillipopoulos ([11]), although in that case there is a public cleanup policy which depends on tax returns. Here, because we want to compare the decentralized equilibrium with the optimal solution in a model, we do not consider taxes to finance cleanup. However, and differently from those authors, we consider that the environmental protection is fostered by social capital and hindered by physical capital.

stock, which will be a source of utility to consumers (N_U):

$$K_N = N_Y + N_U \tag{8}$$

The production of final goods will draw upon man-made capital, use of natural resources, knowledge, and social capital (networks or institutions), so that all inputs are required. Assuming a Cobb-Douglas technology, so that each exponent is the production elasticity of the associated input, we have:

$$Y = K_P^{\beta} K_S^{\sigma} N_Y^{\nu} H_Y^{\eta} \qquad \beta + \sigma + \nu + \eta = 1 \tag{9}$$

While K_P and K_S fully contribute to final goods production, but only the part of natural capital that is removed from the stock variable is included. Thus, only the flows of matter that are being transformed contribute to production. For a more complete explanation of the distinction between the role of Nature as a transformative fund and as a transformed resource flow, see for example Kraev [18].[4] In a similar manner, only the portion of human capital that is specifically dedicated to production is considered, since the remainder is used for different activities, as stated in equation (3).

Finally, to capture the multi-faceted character of an inclusive model, household preferences specify social capital and the part of natural capital that is not used in production (N_U), along with consumption, as arguments of the intertemporal utility function:

$$U(C, N_U, K_S) = \frac{\tau}{\tau - 1} \int_0^{\infty} \left(C_t N_{Ut}^{\phi} K_{St}^{\psi} \right)^{\frac{\tau - 1}{\tau}} e^{-\rho t} dt \tag{10}$$

where τ is the intertemporal elasticity of substitution, ϕ represents the preference for Nature, ψ is the preference for social capital, and ρ is the utility discount rate, so that a higher ρ indicates more impatient consumers. The t subscripts are dropped in the remaining sections for ease of notation.

[4] However, both Kraev [18] and England [12] emphasize complementarity of natural capital in production.

3 Optimal Growth

It is clear that when assets directly provide utility, while simultaneously acting as inputs to the production function, the decentralized equilibrium will in general not maximize aggregate welfare. Thus we must solve a social planner's problem. In this section we derive the conditions associated with the maximization of (10) subject to the production function (9) as well as the transition equations for the different types of capital.

The problem gives rise to the following Hamiltonian function:

$$\begin{aligned}\mathcal{H} &= \frac{\tau}{\tau-1}\left(C\left(K_N-N_Y\right)^{\phi}K_S^{\psi}\right)^{\frac{\tau-1}{\tau}}+\lambda_P\left[K_P^{\beta}K_S^{\sigma}N_Y^{\nu}H_Y^{\eta}-C-\delta_P K_P\right] \\ &+\lambda_H\left[\xi(K_H-H_Y-H_S-H_N)+\alpha K_S-\delta_H K_H\right]+\lambda_S\left(\omega H_S+\Omega K_S\right)+ \\ &+\lambda_N\left[R(K_N)+m_0\frac{H_N^{\epsilon}K_S^{\kappa}}{K_P^{\varphi}}-N_Y\right]\end{aligned} \tag{11}$$

where the λ_j are the co-state variables for each stock K_j, with $j=P,H,S,N$. Considering choice variables C, H_Y, H_S, H_N, and N_Y, and equations (3) and (8) the first order conditions yield:

$$\frac{\partial U}{\partial C}=\lambda_P \tag{12}$$

$$\lambda_H=\lambda_P\frac{\eta Y}{H_Y\xi} \tag{13}$$

$$\lambda_H=\frac{\lambda_S\omega}{\xi} \tag{14}$$

$$\frac{\partial U}{\partial N_Y}+\lambda_P\frac{vY}{N_Y}=\lambda_N \tag{15}$$

$$\lambda_H=\lambda_N\frac{m_0\epsilon H_N^{\epsilon-1}K_S^{\kappa}}{K_P^{\varphi}\xi} \tag{16}$$

as well as:

$$\rho\lambda_P-\dot{\lambda}_P = \lambda_P\frac{\beta Y}{K_P}-\lambda_P\delta_P-\lambda_N\frac{m_0\varphi H_N^{\epsilon}K_S^{\kappa}}{K_P^{\varphi+1}} \tag{17}$$

$$\frac{\dot{\lambda}_H}{\lambda_H} = \rho+\delta_H-\xi \tag{18}$$

$$\rho\lambda_S-\dot{\lambda}_S = \left(\frac{\partial U}{\partial K_S}+\frac{\lambda_P\sigma Y}{K_S}+\lambda_H\alpha+\lambda_S\Omega+\lambda_N\frac{m_0\kappa H_N^{\epsilon}K_S^{\kappa-1}}{K_P^{\varphi}}\right) \tag{19}$$

$$\rho\lambda_N-\dot{\lambda}_N = \left(\frac{\partial U}{\partial K_N}+\lambda_N\frac{\partial R}{\partial K_N}\right) \tag{20}$$

and the required transversality conditions. As usual, $\frac{\partial U}{\partial C}$, $\frac{\partial U}{\partial K_S}$, $\frac{\partial U}{\partial K_N}$, represent the marginal utilities of consumption, social capital, and natural capital, respectively, and $\frac{\partial U}{\partial N_Y} = -\frac{\partial U}{\partial K_N}$. Also, $\frac{\partial R}{\partial K_N}$ is the impact of an additional unit of natural capital on its regeneration.

Conditions (12)-(16) tell us that for each control variable marginal benefits will have to be equated to marginal costs for efficiency to be achieved. For instance, condition (12) balances the marginal utility of consumption with the shadow price of produced capital (since one unit of production that is consumed is no longer available for capital accumulation); likewise, condition (13) equates the shadow price of human capital to its value in production, whereas conditions (14) and (16) equate the same shadow price (as human capital can be put to different uses) to its value in social capital accumulation and environmental protection, respectively. Finally, condition (15) balances the costs of using a unit of natural capital in production (its lost shadow value, reflecting future benefits, as well as its immediate impact on utility), with its benefit in terms of production.

On the other hand, conditions (17)-(20) show the factors influencing the dynamic evolution of the shadow prices for each one of the capital types. Namely, condition (17) reflects that giving up a unit of K_P yields a benefit (from the discount rate and the avoided depreciation) as well as a loss equal to the value of the marginal productivity of produced capital; condition (18) tells a similar story except the loss is in the accumulation of human capital; condition (19) shows that, for each unit of K_S that is relinquished, the value foregone includes the direct impact on utility, the value of the marginal productivity of social capital, its contribution to both human and social capital accumulation, and its benefits on environmental protection**).** Finally, condition (20) highlights the value of the natural capital stock to utility as well as its role in future regeneration.

3.1 Optimal Growth Rates

At the steady state of the model, growth rates will, by definition, be constant, so equation (1) tells us that K_P**,** Y, and C all grow at the same rate.

Furthermore, K_S and K_H will be growing at a different rate, as will be shown below, respecting equations (2) and (4). The growth rate of K_N is zero in the steady-state, as it is physically impossible for natural resources to grow without bounds, so that the growth rate of its components, N_Y and N_U will also be zero.

In the steady-state, we can obtain the human capital growth rate as follows. From (13) we find $\frac{\dot{\lambda}_P}{\lambda_P} = g_{K_H} + \frac{\dot{\lambda}_H}{\lambda_H} - g_Y$ and using the result from equation (18), we can then replace the previous two equations into $-\frac{1}{\tau} g_Y + \psi \left(1 - \frac{1}{\tau}\right) g_{K_H} = \frac{\dot{\lambda}_P}{\lambda_P}$, which we calculated from (12). The substitution yields:

$$\left(1 - \frac{1}{\tau}\right) g_Y + g_{K_H} \left(\psi \left(1 - \frac{1}{\tau}\right) - 1\right) = \rho + \delta_H - \xi \tag{21}$$

To simplify the above expression, we log-differentiate equation (9), yielding:

$$g_Y = g_{K_H} \left(\frac{\eta + \sigma}{1 - \beta}\right) \tag{22}$$

We can substitute the last expression in equation (21) to get the growth rate of human capital:

$$g^*_{K_H} = \frac{\xi - \delta_H - \rho}{1 + \left(\frac{\eta+\sigma}{1-\beta} + \psi\right)\left(\frac{1}{\tau} - 1\right)} \tag{23}$$

Substituting equation (23) in equation (22) we get the output growth rate:

$$g^*_Y = \frac{(\xi - \delta_H - \rho)\left(\frac{\eta+\sigma}{1-\beta}\right)}{1 + \left(\frac{\eta+\sigma}{1-\beta} + \psi\right)\left(\frac{1}{\tau} - 1\right)} \tag{24}$$

Since $1 - \beta = \sigma + \eta + \nu$, the growth of production (and subsequently, of consumption) is now lower than in models where natural capital is absent and where it would coincide with the growth of human capital. Also, the difference between the two increases for larger values of ν, as expected, since a higher share of natural capital in production implies a greater effort in terms of human and social capital to keep consumption growing.

4 Decentralized Equilibrium

In the decentralized equilibrium both consumers and firms have choices to make.

Consumers maximize the intertemporal utility function:

$$\frac{\tau}{\tau-1}\int_0^{\infty}\left(C_t N_{Ut}^{\phi} K_{St}^{\psi}\right)^{\frac{\tau-1}{\tau}} e^{-\rho t}dt$$

subject to the budget constraint:

$$\dot{a} = (r-\delta_p)a + W_H H_Y - C \tag{25}$$

where a represents the value of physical assets, r is the return on physical capital, and W_H is the market wage. The market price for the consumption good is normalized to 1. Since it is making an intertemporal choice, the family also takes into account equations (2) and (4), which represent human and social capital accumulation, respectively. Note that, although natural capital influences utility, there is nothing in the consumer choice affecting K_N directly. This will be a source of externalities from the production to the consumption side.[5]

The firm maximizes profit, π:

$$\pi = K_P^{\beta} K_S^{\sigma} N_Y^{v} H_Y^{\eta} - W_H H_Y - W_N N_Y - rK_P \tag{26}$$

where W_N is the private cost of using natural capital (we assume that firms extract natural resources directly). The markets for purchased production factors are assumed to be competitive. However, the firm cannot buy social capital, as there is, in effect, no market for it. Social capital is treated here as exogenous, although it affects the firm's production. Hence, consumer decisions will carry social capital externalities.

From this problem we know that returns on production factors equal marginal productivities, as follows:

$$W_H = \frac{\eta Y}{H_Y} \tag{27}$$

[5]First order conditions for the consumer problem in the decentralized equilibrium can be found in Appendix A.

$$r = \frac{\beta Y}{K_P} \tag{28}$$

$$W_N = \frac{vY}{N_Y} \tag{29}$$

In Appendix B we have calculated the growth rates of human capital and output in the decentralized equilibrium, and they are shown to be the same as those for the optimal solution.

5 Distortions

The shares of human capital allocated to the different sectors in the social planner are:[6]

$$u_Y^* = \frac{H_Y}{K_H} = \frac{\left(\xi - \delta_H - \frac{\omega\alpha}{\xi} - \Omega\right)\frac{K_S}{K_H} - \frac{\omega\kappa}{\epsilon}u_N^*}{\frac{\omega}{\eta}\left(\psi\left(\frac{C}{Y}\right)^* + \sigma\right)} \tag{30}$$

$$u_S^* = \frac{H_S}{K_H} = \left(\frac{g_{K_S}^* - \Omega}{\omega}\right)\left(\frac{K_S}{K_H}\right)^* \tag{31}$$

$$u_H^* = \frac{H_H}{K_H} = \frac{1}{\xi}\left(g_{K_H}^* + \delta_H\right) - \frac{\alpha}{\xi}\left(\frac{K_S}{K_H}\right)^* \tag{32}$$

$$u_N^* = \frac{H_N}{K_H} = \frac{\Phi}{\frac{\eta\varphi}{\epsilon}}u_Y^* \tag{33}$$

where $\Phi = \left[\beta - \frac{\rho+\delta_P+g_Y\left[\frac{1}{\tau}-\psi\left(1-\frac{1}{\tau}\right)\left(\frac{1-\beta}{\sigma+\eta}\right)\right]}{\frac{g_Y+\delta_P}{\left(1-\left(\frac{C}{Y}\right)^*\right)}}\right]$, and an expression for $\left(\frac{C}{Y}\right)^*$ is given in Appendix C. The equations that are presented in this section provide a basis for a complete analysis of all the relationships between the different capital stocks and also a basis for the comparison with the decentralized equilibrium solution. Using the restriction that $u_Y^* + u_S^* + u_H^* + u_N^* = 1$, we obtain the social to human capital ratio that the social planner would choose:

$$\left(\frac{K_S}{K_H}\right)^* = \frac{1 - \frac{1}{\xi}\left(g_{K_H}^* + \delta_H\right)}{\frac{\left(\xi - \delta_H - \frac{\omega\alpha}{\xi} - \Omega\right)}{\frac{\omega}{\eta}\left(\psi\left(\frac{C}{Y}\right)^* + \sigma\right) + \omega\kappa\frac{\Phi}{\eta\varphi}} + \left(\frac{g_{K_S}^* - \Omega}{\omega}\right) - \frac{\alpha}{\xi}} \tag{34}$$

[6] See Appendix D for the derivation of the following equations.

The shares of human capital allocated to the different sectors in the decentralized equilibrium are:[7]

$$u_Y^{DE} = \frac{H_Y}{K_H} = \frac{\xi - \delta_H - \frac{\omega\gamma}{\xi} - \Omega}{\frac{\omega}{\eta}\left[\psi\left(\frac{C}{Y}\right)^{DE}\right]}\left(\frac{K_S}{K_H}\right)^{DE} \tag{35}$$

$$u_S^{DE} = \frac{H_S}{K_H} = \left(\frac{g_{K_S}^{DE} - \Omega}{\omega}\right)\left(\frac{K_S}{K_H}\right)^{DE} \tag{36}$$

$$u_H^{DE} = \frac{H_H}{K_H} = \frac{1}{\xi}\left(g_{K_H}^{DE} + \delta_H\right) - \frac{\alpha}{\xi}\left(\frac{K_S}{K_H}\right)^{DE} \tag{37}$$

$$u_N^{DE} = \frac{H_N}{K_H} = 0 \tag{38}$$

where an expression for $\left(\frac{C}{Y}\right)^{DE}$ is given in Appendix C. Using the restriction that $u_Y^{DE} + u_S^{DE} + u_H^{DE} + u_N^{DE} = 1$, we obtain the decentralized equilibrium's social to human capital ratio:

$$\left(\frac{K_S}{K_H}\right)^{DE} = \frac{1 - \frac{1}{\xi}\left(g_{K_H}^{DE} + \delta_H\right)}{\frac{\xi - \delta_H - \frac{\omega\gamma}{\xi} - \Omega}{\frac{\omega}{\eta}\left[\psi\left(\frac{C}{Y}\right)^{DE}\right]} + \frac{g_{K_S}^{DE} - \Omega}{\omega} - \frac{\alpha}{\xi}} \tag{39}$$

As growth rates are equal in the decentralized equilibrium and in the optimal solution, the differences between the two solutions are given by the following externalities:

- The environment externality. This externality has two different sources. A direct source linked with $\Phi > 0$ forces $u_N^* > u_N^{DE}$, which implies that all other allocations in the market tend to be lower than the optimal ones. An indirect source linked with the effect that physical capital has in pollution, affecting the consumption-output ratio, which is lower in the optimal solution than it is in the decentralized equilibrium ($\left(\frac{C}{Y}\right)^* < \left(\frac{C}{Y}\right)^{DE}$). This indirect effect acts in order to $\left(\frac{K_S}{K_H}\right)^{DE} > \left(\frac{K_S}{K_H}\right)^*$, i.e., to overinvestment in social capital relative to human capital. This implies that the market economy tends to under-allocate resources to human capital and over-allocate resources to the final good and social capital sectors. This indirect effect also

[7] Again, see Appendix D.

means that the market economy is saving less than the optimal level, which is a new result in this type of models, where the saving rate used to be optimal.

- The social capital externality. This externality has also two sources. As there is not a market for social capital, the social planner would invest more in social capital than the decentralized equilibrium does. This happens due to the fact that the decentralized equilibrium does not internalize the positive effect social capital has in production ($\sigma > 0$). Thus, in the market equilibrium, one should encounter lower allocation to social capital and to the final good production and higher allocation to the human capital accumulation when compared to the respective optimal allocations. This externality may act in opposition to the direct source of the environment externality, as when pushing u_S^* up, the social planner faces a trade-off with his desire of also pushing u_N^* up. However, there is another source of externality that links social capital to natural capital ($\kappa > 0$). The social planner will desire more social capital than the decentralized acting agents, since social capital acts in order to increase environmental protection, a valuable effect from the social point of view.

One of the objectives of this analysis is to determine the optimal allocation of human capital through sectors of activity and what should be done in order to approximate the genuine savings from its optimal level. To summarize results we state two corollaries of what we have demonstrated previously.

Corollary 1 *The decentralized economy saves, as a percentage of income, less than the social planner would do (and hence consumes more), because the social planner wants to save more resources for the future and spend less in the present, as a means to protect the environment.*

Corollary 2 *The social planner would allocate human capital to environmental protection, while the decentralized economy would not. Specific sec-*

tors to which the social planner allocates less human capital than the decentralized economy depends on the specific parameter values.

6 Results and Calibration

6.1 Calibration

In this section, we present and justify the calibrated values for parameters. We use three main sets of parameter values: the first is set to fit the whole world, the second is set to fit Europe and Central Asia and the last is set to fit Low Income Countries (mostly African and Latin American countries). Most parameters come from the report "Where is the Wealth of Nations?" from World Bank [34]. This reference has values for the shares of produced capital, intangible capital, and natural capital for different regions in the world, which corresponds in the model to values for the following parameters: $\beta, \nu, \eta + \sigma$. Then, it remains to disentangle the share of social capital. For that, we use the fact that a level of trust that is 10

rate of human capital.[8] The discount rate is set to 0.01, a value in the range used in the literature. Depreciation rate for physical capital is also set to be in a lower bound of the interval usually found in the literature (0.01). It is usual to see endogenous growth models with human capital accumulation considering no depreciation of human capital, thus we also set $\delta_H = 0$. Small oscillations in these parameters are not of crucial importance for our results. There are not empirical references that can easily guide the calibrations of parameters that rule the mutual influences of social capital in and from human capital. Thus, we considered $\omega = \alpha = 0.01$. We note however that reasonable changes in these values do not produce significant changes in our results. We also consider that $\Omega = 0.01$ (allowing for exogenous social capital growth) and we set the weight of social capital in utility is $\psi = 0.2$ (allowing for a lower weight than that of consumer) and we also note that reasonable changes in these values do not produce significant changes in our results.

To calibrate the parameter that estimates the relationship between physical capital and environmental protection (φ), we must note that the growth rate of physical capital is equal to the per capita output growth rate, and thus we refer to the discussion on the Environmental Kuznets Curve (EKC). This concept predicts an inverted U-shaped relationship between the level of income and pollution (we can see pollution as an inverted function of environmental protection). However, recently, specific country studies have noted that the relationship between income and pollutants is almost linear or, for some pollutants, N-shaped. From Roca et al. [24] we see that the EKC is rejected and the effect of income in pollution is near 1.2. From Akbostanci et al. [1], we learn that this coefficient for Turkey is about 3.5. From Song et al. [28], we see coefficients from 1.5 to 3, in a work applied to China. Given these evidence we choose a parameter $\varphi = 2$ for the World, $\varphi = 1.2$ when applied to Europe, and $\varphi = 3$ when applied to Low Income Countries. For the influence of human capital in environmental protection (ϵ), we adopted the value of 1 from Economides

[8] $\xi = \left[\left(\frac{\eta+\sigma}{1-\beta} + \psi\right)\left(\frac{1}{\tau} - 1\right) + 1\right] g_Y + \rho + \delta_H$.

and Phillipopoulos [11]. For the effect of natural resources in the utility, we have adopted the mean value used also in [11], $\phi = 0.35$. Finally, for κ, we use the fact that pollution must be constant at the steady-state and thus $\kappa = \varphi\frac{1+\sigma}{\eta+\beta} - \epsilon$. The next tables summarize the calibration values for the three different scenarios we will work with.

Table 1 - Parameters Values - World

Production and Utility							
g_Y	β	δ_P	τ	ρ	C/Y^{DE}		
0.141	0.18	0.01	0.8	0.01	0.89		
Social and Human Capital							
ξ	δ_H	α	ω	Ω	ψ	σ	η
0.042	0	0.01	0.01	0.01	0.2	0.08	0.7
Natural Capital							
φ	ϵ	κ	ν	ϕ			
2	1	0.58	0.04	0.35			

Table 2 - Parameters Values - Europe and Central Asia

Production and Utility							
g_Y	β	δ_P	τ	ρ	C/Y^{DE}		
0.128	0.31	0.01	0.8	0.01	0.81		
Social and Human Capital							
ξ	δ_H	α	ω	Ω	ψ	σ	η
0.048	0	0.01	0.01	0.01	0.2	0.08	0.34
Natural Capital							
φ	ϵ	κ	ν	ϕ			
1.20	1	0.88	0.27	0.35			

Table 3 - Parameters Values - Low Income Countries

Production and Utility							
g_Y	β	δ_P	τ	ρ	C/Y^{DE}		
0.0055	0.16	0.01	0.8	0.01	0.90		
Social and Human Capital							
ξ	δ_H	α	ω	Ω	ψ	σ	η
0.0255	0	0.01	0.01	0.0055	0.2	0.08	0.47
Natural Capital							
φ	ϵ	κ	ν	ϕ			
3	1	0.54	0.29	0.35			

6.2 Quantifying Distortions

In this section, we present the main economic indicators for the Social Planner solution and for the decentralized equilibrium, for the different exercises we have described above. Tables 4 to 6 present results.

Table 4 - Results on Distortions - World

Social Planner					
K_S/K_H	C/Y	u_Y	u_H	u_S	u_N
0.08	0.887	62.6%	33.5%	3.8%	0.1%
Decentralized Equilibrium					
K_S/K_H	C/Y	u_Y	u_H	u_S	u_N
0.05	0.89	63.4%	34.1%	2.6%	0%

Table 5 - Results on Distortions - Europe and Central Asia

Social Planner					
K_S/K_H	C/Y	u_Y	u_H	u_S	u_N
0.095	0.809	47.7%	41.8%	10.4%	0.04%
Decentralized Equilibrium					
K_S/K_H	C/Y	u_Y	u_H	u_S	u_N
0.067	0.810	50.3%	42.4%	7.3%	0%

Table 6 - Results on Distortions - Low Income Countries

Social Planner					
K_S/K_H	C/Y	u_Y	u_H	u_S	u_N
0.228	0.893	68.1%	23.9%	7.8%	0.2%
Decentralized Equilibrium					
K_S/K_H	C/Y	u_Y	u_H	u_S	u_N
0.155	0.9	68%	26.8%	5.3%	0%

As expected, the social capital to human capital ratio is always sub-optimal and the consumption output ratio is always above the optimal level for that ratio, as well as allocation of human capital to pollution prevention is only done under the optimal choice. Moreover, the under-optimal ratio of social to human capital directly implies a under-optimal allocation to the social capital sector. Usually, in the World and Europe and Central Asia calibrations, the decentralized economy allocated more human capital to the final good than it was desired from the social planner point of view. However, this is inverted in the Low Income Countries calibration. Finally, allocation of human capital to school attendance is always above

the optimal level. The social planner increases the allocation of human capital to social capital accumulation and pollution prevention due to the two types of externalities explained above.

6.3 Sensitivity Analysis

As a first sensitivity analysis, we eliminate the externality due to the social capital market failure ($\sigma = 0$). We note that all the distortions described above still remain effective, although the distortion in the allocation of human capital to the social capital sector decreases. However, the social capital - environmental externality ($\kappa > 0$) is sufficiently strong to keep the investment in social capital below the optimal level. When setting $\kappa = 0$, the results change. We present results on the Low Income Countries, as it is the experiment where the distortions are bigger. However, the direction of distortions are equal in the three calibrations considered above.

Table 7 - Results on Distortions - Low Income Countries ($\sigma = \kappa = 0$)

Social Planner					
K_S/K_H	C/Y	u_Y	u_H	u_S	u_N
0.131	0.893	67.3%	27.71%	4.46%	0.5%
Decentralized Equilibrium					
K_S/K_H	C/Y	u_Y	u_H	u_S	u_N
0.132	0.9	67.8%	27.67%	4.49%	0%

Table 7 shows that without externalities caused by social capital, the allocation of social capital is above the optimal level and now, human capital allocation to school attendance is below the optimal level. Contrary to what happened in Table 6, this exercise shows overallocation of resources to the final good.

In models without social capital and natural capital (e.g. Reis and Sequeira [23]), no distortion used to affect human capital allocated to school attendance. In this model, there is the participation of social capital in human capital accumulation that also induces distortions in human capital allocated to school attendance. If we set $\alpha = 0$, we recover that result.

7 Conclusion

Achieving economic growth has always been a major concern of economic theory and policy. Several significant developments have appeared in the literature, especially in the last two decades. In particular, human capital has been included as an essential force for growth and the role of natural capital has been discussed thoroughly. More recently, attention has been drawn to the notion of social capital, seen as the level of trust, social norms, and social networking, as an additional explanatory factor for growth. Nonetheless, existing models typically include only two or, rarely, three types of capital, so that at best they can provide a partial view of the growth trends.

In this paper we propose an intertemporally efficient model containing the four categories of assets that can be defined, namely produced capital, natural capital, human capital, and social capital, in an attempt to emphasize the relationships between them in an inclusive model of sustainable growth.

We identify several types of externalities linked with social and natural capital. There are two types of externalities linked with social capital. First, there is a market failure for social capital, as social capital contributes to production but firms do not pay for it, so a social planner would produce more social capital than the amount produced in the market. Second, social capital increases environmental protection (or decreases pollution), an effect that is not internalized by the market. There are also two types of externalities linked with natural capital. First, in the decentralized market it does not exist the incentive to allocate human capital to environmental protection. Second, the effect of physical capital in pollution is not internalized in the decentralized economy. Contrary to what is usual to happen in models of endogenous growth without natural resources, this externality implies distortions in the output to capital ratio and in the consumption to output ratio, with an higher savings rate under the social planner framework. We discovered that the social planner always chooses to allocate more human capital to pollution prevention and also for social

capital accumulation. He usually allocates less human capital to the final good production and to school attendance.

Further work should be both empirical, using recently available data to ascertain the validity of the model's assumptions, and theoretical, in the search for intuition for general results with these or other functional forms. The use of specific values would also allow the study of the transition dynamics. Finally, work on inclusive sustainable growth must also be undertaken for models where uncertainty exists, in particularly when systems are complex and unstable. System resilience is turning out to be a key property in the search for sustainability and its consideration should be sought in discussions about economic growth in the real, natural world.

References

[1] Akbostanci, E, T. Serap, and G. Tunç (2009), "The Relationship Between Income and Environment in Turkey:Is There an Environmental Kuznets Curve?", *Energy Policy,* 37: 861-867.

[2] Bartolini, S. and L. Bonatti (2007), "Endogenous Growth, Decline in Social Capital and Expansion of Market Activities", *MPRA Working-Paper 3341.*

[3] Beugelsdijk, S., H. Groot, and A. van Schaik (2004), "Trust and Economic Growth: A Robustness Analysis", *Oxford Economic Papers*, 56, 118-134.

[4] Beugelsdijk, S. and S. Smulders (2004), "Social Capital and Economic Growth", Tilburg University, Department of Economics, *mimeo.*

[5] Bisin, B. and D. Guaitoli (2006), "Social Capital, Modernization and Growth", ASSET Conference 2006, *mimeo.*

[6] Brunnschweiler, C. and E. Bulte (2008) "The resource curse revisited and revised: a tale of paradoxes and red herrings", *Journal of Environmental Economics and Management*, 55, 248-264.

[7] Coleman, J. (1988), "Social Capital in the Creation of Human Capital", *American Journal of Sociology*, 94, S95-S120.

[8] Constantini, V. and S. Monni (2008), "Environment, Human Development and Economic Growth", *Ecological Economics*, 64, 867-880.

[9] Dasgupta, P. (2008), "Nature in Economics", *Environmental and Resource Economics*, 39, 1-7.

[10] Dinda, S. (2007), "Social Capital in the Creation of Human Capital and Economic Growth: a Productive Consumption Approach", *The Journal of Socio-Economics*, doi:10.1016/jsocec.2007.06.014.

[11] Economides, G. and A. Phillipopoulos (2008), "Growth Enhancing Policy is the Means to Sustain the Environment", *Review of Economic Dynamics*, 11, 207-219.

[12] England, R. (2000), "Natural Capital and the Theory of Economic Growth", *Ecological Economics*, 34, 425-431.

[13] Glaeser, E.L., D. Laibson, and B. Sacerdote (2002), "An Economic Approach to Social Capital", *Economic Journal*, 112, F437-F458.

[14] Guvenen, M. (2006), "Reconciling Conflicting Evidence on the Elasticity of Intertemporal Substitution: A Macroeconomic Perspective", *Journal of Monetary Economics*, 53(7), 1451-1472.

[15] Gylfason, T. and G. Zoega (2006) "Natural Resources and Economic Growth: The Role of Investment", The World Economy, 1091-1115.

[16] Jones, L., R. Manuelli, and H. Siu (2000), "Growth and Business Cycles", *NBER Working Paper No.7633*, National Bureau of Economic Research.

[17] Knack, S. and P. Keefer (1997), "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation", *Quarterly Journal of Economics*, 112, 1251-88.

[18] Kraev, E. (2002) "Stocks, Flows and Complementarity: Formalizing a Basic Insight of Ecological Economics", *Ecological Economics*, 43, 277-286.

[19] Maddison, A. (2003), *The World Economy: Historical Statistics*, OECD.

[20] Pearce, D., A. Markandya, and E. Barbier (1989), *Blueprint for a Green Economy*, Earthscan.

[21] Pezzey. J. and M. Toman (2005), "Sustainability and its Economic Interpretations", in *Scarcity and Growth Revisited - Natural Resources and the Environment in the New Millenium*, Editors R.D. Simpson, M.A. Toman, and R. Ayres, Resources for the Future, Washington DC.

[22] Putman, R. (1993), *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton, Princeton University Press.

[23] Reis, A. and T. Sequeira (2007), "Human Capital and Overinvestment in R&D", *Scandinavian Journal of Economics*, 109(3), 573-591.

[24] Roca, J., E. Padilla, M. Farré, and V. Galletto (2001), "Economic Growth and Atmospheric Pollution in Spain: Discussing the Environmental Kuznets Curve Hypothesis", *Ecological Economics,* 39, 85-99.

[25] Rodrigues, J., T. Domingos, P. Conceição, and J. Belbute (2005), "Constraints on Dematerialisation and Allocation of Natural Capital Along a Sustainable Growth Path", *Ecological Economics*, 54, 382-396.

[26] Rupasinga A. (2000), "Social Capital and Economic Growth: A County-Level Analysis", *Journal of Agricultural and Applied Economics*, 32(3), 565-72.

[27] Sequeira, T.N. and A. Ferreira-Lopes (2007) "An Endogenous Growth Model with Human and Social Capital Interactions", *mimeo.*

[28] Song, T., T. Zheng, and L. Tong (2008), "An Empirical Test of the Environmental Kuznets Curve in China: A Panel Cointegration Approach", *China Economic Review,* 19, 381-392.

[29] Smulders, S. (2005), "Endogenous Technological Change, Natural Resources, and Growth", in *Scarcity and Growth Revisited - Natural Resources and the Environment in the New Millenium*, Editors R.D. Simpson, M.A. Toman, and R. Ayres, Resources for the Future, Washington DC.

[30] Teachman J. D., K. Paasch, and K. Carver (1997), "Social Capital and the Generation of Human Capital", *Social Forces*, 75(4), 1343-1359.

[31] Temple, J. and P.A. Johnson (1998), "Social Capability and Economic Growth", *The Quarterly Journal of Economics*, 113(3), 965-990.

[32] Whiteley, P. (2000), "Economic Growth and Social Capital", *Political Studies*, 48(3), 443-466.

[33] World Bank (2001), *Global Development Network Growth - Macro Time Series*, New York University Development Research Institute.

[34] World Bank (2006), *Where is the Wealth of Nations? Measuring Capital for the 21st Century*, The World Bank, Washington DC.

A First Order Conditions for the Decentralized Equilibrium

The choice variables for the consumers are C, H_H, and H_S, so the first order conditions for the consumer problem yield:

$$\frac{\partial U}{\partial C} = \lambda_a \tag{40}$$

$$\lambda_H^{'} = \frac{\lambda_a W_H}{\xi} \tag{41}$$

$$\lambda_S^{'} = \frac{\lambda_a W_H}{\omega} \tag{42}$$

as well as:

$$\frac{\dot{\lambda}_a}{\lambda_a} = \rho + \delta_P - r \tag{43}$$

$$\dot{\lambda}_H^{'} = \rho\lambda_H^{'} - (\lambda_a W_H - \lambda_H^{'}\delta_H) \tag{44}$$

$$\dot{\lambda}_S^{'} = \rho\lambda_S^{'} - \left(\frac{\partial U}{\partial K_S} + \lambda_H^{'}\alpha + \lambda_S^{'}\Omega\right) \tag{45}$$

where λ_a is the co-state variable for the budget constraint, and $\lambda_H^{'}$ and $\lambda_S^{'}$ are co-state variables for the stocks of human and social capital, respectively.

B Growth Rates in the Decentralized Equilibrium

In the steady-state we can obtain the human capital growth rate of the decentralized equilibrium as follows. By using equation (44) and replacing it in $\frac{\dot{\lambda}_H^{'}}{\lambda_H^{'}} = \frac{\dot{\lambda}_a}{\lambda_a} + g_W$, which we get by (41), we find $\frac{\dot{\lambda}_a}{\lambda_a} = \rho + \delta_H - \xi - g_W$. After log-differentiating equation (27) we get $g_W = g_Y - g_{K_H}$. Substituting this last equation in the previous one and introducing both in $-\frac{1}{\tau}g_Y + \psi\left(1 - \frac{1}{\tau}\right)g_{K_H} = \frac{\dot{\lambda}_a}{\lambda a}$, which we calculated from (40), we get:

$$\left(1 - \frac{1}{\tau}\right)g_Y + g_{K_H}\left(\psi\left(1 - \frac{1}{\tau}\right) - 1\right) = \rho + \delta_H - \xi \tag{46}$$

By log-differentiating the production function (9) we get:

$$g_Y = g_{K_H}\left(\frac{\eta + \sigma}{1 - \beta}\right) \tag{47}$$

Substituting this last expression into (46) we get:

$$g_{K_H}^{DE} = \frac{\xi - \delta_H - \rho}{\left(\frac{\eta+\sigma}{1-\beta} + \psi\right)\left(\frac{1}{\tau} - 1\right) + 1} \tag{48}$$

By substituting equation (48) into (47) we find:

$$g_Y^{DE} = \frac{(\xi - \delta_H - \rho)\left(\frac{\eta+\sigma}{1-\beta}\right)}{\left(\frac{\eta+\sigma}{1-\beta} + \psi\right)\left(\frac{1}{\tau} - 1\right) + 1} \tag{49}$$

These growth rates are equal to the ones that we found in the social planner problem.

C Consumption-Output Ratios

Now we also demonstrate the relationship between the consumption to output ratio in the decentralized equilibrium and in the social planner solution. We use this expression in the calibration of the model.

From equation (1) we get:

$$\frac{\dot{K}_P}{K_P} = \frac{Y}{K_P} - \frac{C}{K_P} - \delta_P \tag{50}$$

Also, from this equation we get $g_{K_P} = g_Y = g_C$ because the growth rates have to be constant in steady state. Since we have shown that $(g_{K_P} = g_Y = g_C)^* = (g_{K_P} = g_Y = g_C)^{DE}$ the left hand-side of equation (50) is equal in the social planner and in the decentralized equilibrium and δ_P is a constant.

In the social planner problem by transforming equation (12) we obtain - $\frac{\dot{\lambda}_P}{\lambda_P} = -\frac{1}{\tau}g_C + \psi(1-\frac{1}{\tau})g_{K_S}$. Replacing this last equation in equation (17) we get:

$$\frac{\beta Y}{K_P}^* = \frac{\rho + \delta_P + \frac{1}{\tau}g_C - \psi(1-\frac{1}{\tau})g_{K_S}}{\left(1 - \frac{\eta\varphi}{\epsilon\beta}\frac{u_N^*}{u_Y^*}\right)} \tag{51}$$

Since we have shown that $g_{K_P}^* = g_{K_P}^{DE}$, hence $\left(\frac{Y}{K_P} - \frac{C}{K_P}\right)^* = \left(\frac{Y}{K_P} - \frac{C}{K_P}\right)^{DE}$. Using the fact that $\frac{C}{K_P}$ can be written has $\frac{C}{Y}\frac{Y}{K_P}$ and putting $\frac{Y}{K_P}$ in evidence, we obtain:

$$\left(\frac{Y}{K_P}\right)^* \left(1 - \frac{C}{Y}\right)^* = \left(\frac{Y}{K_P}\right)^{DE} \left(1 - \frac{C}{Y}\right)^{DE} \tag{52}$$

From the decentralized equilibrium, using (28) we note that $\left(\frac{\beta Y}{K_P}\right)^{DE} = r$ and that $r = Pmg_K$, thus:

$$\left(\frac{\beta Y}{K_P}\right)^{DE} = \rho + \delta_P + \frac{1}{\tau}g_C - \psi(1-\frac{1}{\tau})g_{K_S} \tag{53}$$

Hence

$$\frac{\beta Y}{K_P}^* = \frac{\left(\frac{\beta Y}{K_P}\right)^{DE}}{\left(1 - \frac{\eta\varphi}{\epsilon\beta}\frac{u_N^*}{u_Y^*}\right)} \tag{54}$$

Now substituting (54) into (52), we get:

$$\left(\frac{C}{Y}\right)^* = \left(\frac{C}{Y}\right)^{DE}\left[1 - \frac{\eta\varphi}{\epsilon\beta}\frac{u_N^*}{u_Y^*}\right] - \frac{\eta\varphi}{\epsilon\beta}\frac{u_N^*}{u_Y^*} \tag{55}$$

From this equation it becomes clear that $\left(\frac{C}{Y}\right)^* \neq \left(\frac{C}{Y}\right)^{DE}$, only because $u_N^* \neq 0$, and also that $\left(\frac{C}{Y}\right)^* < \left(\frac{C}{Y}\right)^{DE}$. We can use (30) and (33) to determine the value of $\left(\frac{C}{Y}\right)$ in the optimal solution dependent on known values:

$$\left(\frac{C}{Y}\right)^* = \frac{\left(\frac{C}{Y}\right)^{DE} - \beta(1 + \left(\frac{C}{Y}\right)^{DE}) + \frac{1+\left(\frac{C}{Y}\right)^{DE}\ (\Psi)}{g_Y+\delta_P}}{1 + \frac{1+\left(\frac{C}{Y}\right)^{DE}\ (\Psi)}{g_Y+\delta_P}}$$

where $\Psi = \rho + \delta_P + g_Y\left[\frac{1}{\tau} - \psi\left(1 - \frac{1}{\tau}\right)\left(\frac{1-\beta}{\sigma+\eta}\right)\right]$.

D Human Capital Shares

D.1 Social Planner

To obtain the share of human capital allocated to production in the social planner, we first use equation (14) which gives us: $\lambda_H = \frac{\lambda_S\omega}{\xi}$ and by substituting this expression into equation (13) we get:

$$\lambda_P = \frac{H_Y\lambda_S\omega}{\eta Y} \tag{56}$$

We know that $\frac{\partial U}{\partial C} = \lambda_P$ by equation (12) and we also know that $\frac{\partial U}{\partial K_S} = \psi\frac{\partial U}{\partial C}\frac{C}{K_S}$, hence $\frac{\partial U}{\partial K_S} = \psi\lambda_P\frac{C}{K_S}$. Substituing (56) in this last expression we get:

$$\frac{\partial U}{\partial K_S} = \psi\frac{H_Y\lambda_S\omega}{\eta Y}\frac{C}{K_S} \tag{57}$$

Substituting equation (14) into equation (16) we obtain:

$$\lambda_N = \frac{\lambda_S\omega K_P^{\varphi}}{m_0\epsilon H_N^{\epsilon-1}K_S^{\kappa}} \tag{58}$$

By equation (14) we get that $\frac{\dot{\lambda}_H}{\lambda_H} = \frac{\dot{\lambda}_S}{\lambda_S}$, which by equation (18) is equal to $\rho + \delta_H - \xi$. After substituting this last expression and also (57), and (58) into equation (19) and dividing the referred equation by λ_S we finally obtain (30).

The share of human capital allocated to pollution prevention in the social planner is calculated as described next. First, we substitute equation (16) into equation (13) to get:

$$\lambda_N = \frac{\lambda_P Y \eta K_P^{\varphi}}{H_Y m_0 \epsilon H_N^{\epsilon-1} K_S^{\kappa}} \tag{59}$$

By substituting (59) into (17) and dividing this last equation by λ_P we obtain:

$$\frac{\dot{\lambda}_P}{\lambda_P} = \rho + \delta_P - \frac{Y}{K_P}\left(\beta - \frac{\eta\varphi}{\epsilon}\frac{u_N}{u_Y}\right) \tag{60}$$

Secondly, we log-differentiate equation (1) to get:

$$\frac{Y}{K_P} = \frac{g_Y + \delta_P}{1 - \frac{C}{Y}} \tag{61}$$

Thirdly, log-differentiating equation (12) and using (22) gives us:

$$g_Y\left(-\frac{1}{\tau} + \psi\left(1 - \frac{1}{\tau}\right)\left(\frac{1-\beta}{\sigma+\eta}\right)\right) = \frac{\dot{\lambda}_P}{\lambda_P} \tag{62}$$

Finally, substituting (61) and (62) into (60) we get (33).

We get the share of human capital allocated to school time (32) from equation (2) and the share of human capital allocated to investing in social capital (31) from equation (4).

D.2 Decentralized Equilibrium

To obtain the share of human capital allocated to production in the decentralized equilibrium, we first use equation (42) which gives us:

$$\lambda_a = \frac{\lambda_S^{'}\omega}{W_H} \tag{63}$$

We know that $\frac{\partial U}{\partial C} = \lambda_a$ by (40) and we also know that $\frac{\partial U}{\partial K_S} = \psi \frac{\partial U}{\partial C} \frac{C}{K_S}$, hence $\frac{\partial U}{\partial K_S} = \psi \lambda_a \frac{C}{K_S}$. Substituing (63) in this last expression we get:

$$\frac{\partial U}{\partial K_S} = \psi \frac{\lambda_S^{'} \omega}{W_H} \frac{C}{K_S} \tag{64}$$

By the substitution of (41) into (42) and by log-differentiating it we get that $\frac{\dot{\lambda}_H^{'}}{\lambda_H^{'}} = \frac{\dot{\lambda}_S^{'}}{\lambda_S^{'}}$, which by equation (44) is equal to $\rho + \delta_H - \xi$. After substituting this last expression and also (64) into equation (45) and dividing this equation by $\lambda_S^{'}$ we get:

$$\xi - \delta_H - \frac{\omega \gamma}{\xi} - \Omega = \psi \frac{\omega}{W_H} \frac{C}{K_S} \tag{65}$$

We know by (27) that $W_H = \frac{\eta Y}{H_Y}$ and after substituting this expression into (65) we get the expression for the share of human capital allocated to production as in (35).

The share of human capital allocated to school time (37) and the share of human capital allocated to investing in social capital (36) are equal and obtained in the same way that in the social planner. In the decentralized equilibrium the share of human capital allocated to pollution prevention (38) is zero.

STRUCTURAL TRANSFORMATION AND THE ROLE OF FOREIGN DIRECT INVESTMENT IN PORTUGAL: A DESCRIPTIVE ANALYSIS FOR THE PERIOD 1990-2005

MIGUEL LEBRE DE FREITAS
E-mail: afreitas@egi.ua.pt
Universidade de Aveiro
RICARDO PAES MAMEDE
Instituto Superior de Ciências do Trabalho e da Empresa (ISCTE)

Abstract

In this paper we use a recent measure of the "income level of a country's exports" proposed by Hausmann et al. (2007) to characterize the structure of the Portuguese export basket, its recent evolution and the role of FDI in this process. We find that between 1990 and 2005 the improvement in the income content of the Portuguese export basket was achieved through a positive "structural transformation effect" that more than offset the negative effect of having a significant share of products exposed to an increasing competition from emerging economies. We find that the weight of exports with "high" and "very high" income content increased considerably in this period, with these two classes explaining more than half of the total export growth. Analysing the presence of FDI in the different export products, we find a higher than average share of foreign affiliated firms in products with "High" and "Very High" income content. These and other pieces of evidence suggest that FDI has played a relevant role both in the growth of Portuguese exports and in the increase of their income content.

JEL: C14, F14.
Key-Words: International trade, Foreign Direct Investment, The Portuguese Economy, Structural transformation.
Thematic Area: Growth Economics.

1. INTRODUCTION

In the current debate on the Portuguese economy, there is a view that the country's specialization pattern, traditionally dominated by low-skilled labour intensive products, is a major obstacle to convergence. According to this view, with the emergence of new trading partners in the international arena, the future performance of the Portuguese economy will depend critically on its ability to shift its specialization pattern towards goods with higher productivity content. In this paper, we investigate the extent to which the Portuguese economy has indeed become increasingly specialized in more sophisticated goods and whether such shift is more evident in sectors with a high presence of FDI.

The view that a country's economic performance is to a large extent linked to international trade has a long tradition in economic thinking, backing from Adam Smith and David Ricardo. At the theoretical level, many authors have emphasized the type and the characteristics of the sectors wherein a country specializes (Prebisch 1950, Singer, 1950, Kaldor, 1966, Thirlwall,

1979, Grossman and Helpman, 1991)[1]. Empirically, however, this idea has been difficult to test, because a measure of a country specialization pattern that reflects the quality of the goods being exported is not easy to define. In a recent contribution, Hausmann et al. (2007) propose a quantitative index that ranks traded goods in terms of their "implied income". This index (PRODY) is estimated as a weighted average of the per capita GDPs of the countries exporting a product, where the weights reflect the revealed comparative advantage of each country in that product. The authors then compute a measure of sophistication of a country export basket (EXPY) by calculating the export-weighted average PRODY for that country. The authors report a strong correlation between EXPY and per capita GDPs and also find that EXPY is a strong and robust predictor of subsequent economic growth, controlling for standard covariates[2].

In this paper, we compute a new vector of PRODY indexes, using 1995 and 2005 COMTRADE data for 1235 products and 81 countries. We then use this index to characterise the Portuguese export basket and to assess how well it has moved towards goods with higher income content. We document that in the period from 1995 to 2005 there has been indeed an upscale move of the Portuguese specialization pattern. Though using a different methodology, our evidence accords with the recent findings of Caldeira Cabral (2008) and Amador et al. (2007) who analysed the changing structure of Portuguese exports following the OECD classification of R&D intensities[3].

We then investigate the extent to which foreign direct investment (FDI) has played a key role in this change. Portuguese governments have made significant efforts to support FDI inflows, either through financial incentives (EU funds and tax benefits) or by providing complementary infrastructure. Despite the high year-on-year volatility, FDI net flows to Portugal have a clear upward trend, from 0,43% of GDP in the 1970s to 1,03% in the 1980s, 1,085% in the 1990s and 3,65% in the period 2000-2006 (UNCTAD, 2007). An obvious question is, thus, whether such an effort has helped or impaired the process of structural transformation.

The relationship between FDI and economic performance is a topic of controversy in the economic literature. Policymakers and academics often argue that FDI can be a source of benefits to host countries, through knowledge spillovers or by creating linkages from multinationals to domestic firms[4]. Accordingly, governments all over the world spend large amounts of resources to attract subsidiaries of multinational firms to their jurisdiction. Empirically, however, the literature has not being able to confirm the existence of positive

[1] A different question is whether the *identity* of the trade partners matters. The rationale is that a country importing goods primarily from technological leaders receives more technology than a country importing primarily from follower countries (Eaton an Kortum, 1996, 1999). The empirical evidence of this hypothesis remains, however, mixed (see Keller, 2004, for a survey).

[2] Other relevant contributions include Dalum et al. (1999) and Feenstra and Rose (2000). Dalum et al. (1999) focused on 11 manufacturing sectors in the OECD area for the period 1965-1988, and found that the characteristics of the specialization pattern are important to explain growth differentials. However, the impact seems to be gradually wearing off during the 1980s and their results are sensitive to alternative classifications of sectors into different technological categories that the authors consider. Feenstra and Rose (2000) develop a procedure to rank-order countries and commodities according to the "product-cycle" hypothesis, using disaggregated data on US imports. They find a strong relation between what they dubbed "advanced export structure" and high productivity levels and fast growth rates.

[3] Monthly updates using this methodology are available at the GEE homepage, in "Balança de Produtos Industriais Transformados por Grau de Intensidade Tecnológica" (http://www.gee.min-economia.pt/).

[4] Fosfuri et al. (2001) discuss the spillover effects related to the flow of skilled workers trained by multinationals to other firms in the host country. Rodriguez-Clare (1996) and Markusen and Venables (1999) examine the linkage effects between multinationals and domestic firms: the entry of multinationals may raise the demand for intermediate products which would not develop otherwise, triggering the supply of downstream industries and eventually creating the conditions for related upstream industries to develop (a similar reasoning in Trindade, 2005). Other theories pointing to inertia in the specialization patterns (which FDI may help disrupt) include learning-by-doing (Jovanovic and Nyarko, 1996, Hausmann and Klinger, 1997, and Hausmann and Rodrik, 2006) and information externalities (Hausmann and Rodrik, 2003).

externalities from FDI to host countries[5]. In the specific case of Portugal, there is anecdotic evidence of training spillovers and quality improvement effects on domestic suppliers (OECD, 2008, pp. 86-87). However, Flores et al. (2007) found no robust evidence of intra-sectoral spillover effects in the 1990s, as measured by the effect of FDI on domestic firm' labour productivity. Guimarães et al. (2000), analysing the role of agglomeration effects in the location decisions of establishments participated by foreign capital between 1982 and 1992, found a positive influence of industry-specific localization economies but no significant influence of foreign-specific agglomeration effects. This is suggestive of spillovers, but not necessarily related to the affiliation of capital.

This paper abstracts from spillovers and other indirect effects of FDI. Simply, we examine whether there is a tendency for foreign affiliated firms to operate in fast growing, non-traditional and income content export sectors, thus having a significant direct contribution to the process of structural transformation in Portugal.

The role of FDI in Portuguese exports was examined by Cabral (1996) and Magalhães and Africano (2007). Cabral (1996), using a panel of 1174 firms between 1986 and 1992 report that foreign firms tend to export more than domestic firms, everything else constant. Magalhães and Africano (2007), using data on bilateral trade and FDI flows for the period from 1995 to 2000, found a positive correlation between the presence of foreign affiliated capital in the Portuguese economy and the volume of exports to the affiliation country. Gonçalves and Guimarães (1997) do not examine export data, but report a significant difference between the production patterns of domestic and foreign firms operating in Portugal, thus concluding that FDI helps increase the diversification level of the economy[6]. Our paper adds to the previous analysis in that it examines the relationship between FDI and the structure of Portuguese exports, crossing information on export values at the product (SITC-4 rev 2) level and on the proportion of capital with foreign affiliation at the firm level. We then assess whether FDI has contributed to improve the specialization pattern of the Portuguese economy, using the PRODY index as a measure of income content.

The paper proceeds as follows. In Section 2, we analyse the relationship between RCA and PRODY at the product level from 1995 to 2005, in Portugal and in some other countries. In Section 3 we decompose the changes in the average income content of each country's exports into a "PRODY effect" and a "structural transformation effect". In Section 4 we investigate how the composition of the Portuguese export basket has evolved in terms of classes of PRODY. In Section 5 we evaluate the extent to which the sectors that most contributed to the Portuguese export growth have a large presence of FDI and whether the presence of FDI is more significant in products with higher income content. Section 6 concludes.

[5] Many authors remain sceptical about a positive relationship between FDI and economic performance of host countries (e.g., Rodrik, 2007, pp.119-120). Empirically, the evidence has been ambiguous (see Keller, 2004 for a survey). Still, some positive results were found. At the aggregate level, a commonly cited article is Borenztein et al. (1998), who found an important role for FDI in a cross-country growth regression, but only when the host economy has a minimum threshold stock of human capital. This suggests that FDI contributes to economic growth only when a sufficient absorptive capability of the advanced technologies is available in the host economy. Recent evidence with micro-data is also suggestive of important spillovers associated to FDI (Kugler, 2006). In general, empirical assessments face a basic problem of endogeneity: because multinationals are attracted to high-productivity industries and high-performance countries, in the absence of adequate instrumental variables, the direction of causality is difficult to assert (Aitken and Harrison, 1999).

[6] Other relevant studies addressing FDI in Portugal include Barbosa and Louri (2005), who investigate whether the profitability of foreign firms operating in Portugal differs from that of domestic firms; Barbosa et. al (2004), who examined the determinants of FDI; Mata and Portugal (2004), who compare the patterns of entry, survival and growth of domestic and foreign owned firms operating in Portugal; and Barbosa and Louri (2002), who analyse the determinants of ownership decisions.

2. INCOME CONTENT AND COMPARATIVE ADVANTAGES

In this paper, we use the Hausmann et al. (2007) PRODY index to assess the sophistication level of products. Formally, the index is defined, for each product, as the weighted average of per capita incomes of countries exporting that product, where the weights are proportional to the country's index of Revealed Comparative Advantage in that good (details in Appendix 1). Products with high values of PRODY are, by construction, those where high income countries play a major role with respect to the other trading partners. The implied assumption is that the presence of higher wages is stronger where comparative advantages are determined by factors other than labour cost, such as know how, technology, public infrastructures, research centres and so on.

Our calculations use international trade data at the product level (SITC-4 rev 2), from the UN-COMTRADE database, as extracted in September 2007 and per capita GDP levels (in PPP) by the International Monetary Fund, World Economic Outlook Database, April 2008. Both variables refer to 1995 and 2005, and countries for which there was no consistent data for those two years were excluded. This leaves us with 81 countries and data for 1235 products. Table 1 displays the estimated PRODY values for some products, the corresponding PRODY rank and the share in World exports, in 2005. As expected, agricultural commodities and raw materials appear at the bottom of the table.

Table 1: PRODY values for a sample of products

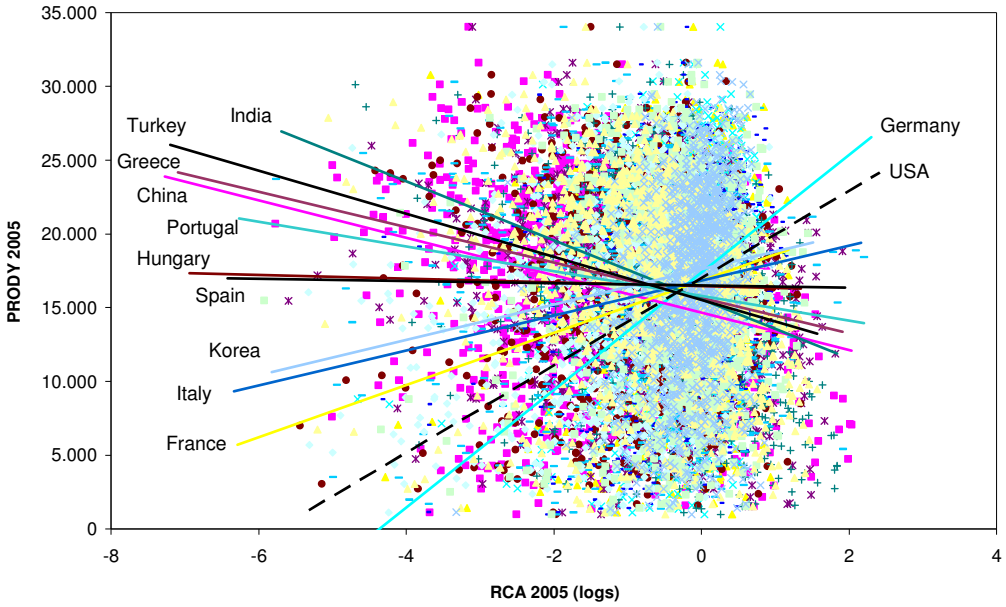
Code	Commodity	PRODY 05	Rank	Share of World exports (per cent)
2933	Heterocyclic compounds with nitrogen hetero-atom(s) only.	33.408	4	0,47
8411	Turbo-jets, turbo-propellers and other gas turbines.	27.010	82	0,71
3004	Medicaments (excluding goods of heading 30.02, 30.05 or 30.06)	26.024	108	2,13
8525	Transmission apparatus for radio-telephony, radio-broadcasting	24.156	196	1,89
8542	Electronic integrated circuits and microassemblies.	24.047	201	2,81
9018	Instruments and appliances used in medical, surgical, dental or veterinary ...	23.486	229	0,61
8473	Parts and accessories for use with machines of heading 84.69 to 84.72	23.244	240	1,89
8703	Motor cars and other motor vehicles principally designed for the transport ...	22.951	255	5,15
8471	Automatic data processing machines and units thereof	22.355	292	2,78
8802	Other aircraft (for example, helicopters, aeroplanes); spacecraft	21.886	330	0,88
8414	Air or vacuum pumps, air or other gas compressors and fans	21.457	344	0,43
8708	Parts and accessories of the motor vehicles of headings 87.01 to 87.05.	20.802	382	2,34
8536	Electrical apparatus for switching or protecting electrical circuits, or fo ...	20.455	401	0,59
8541	Diodes, transistors and similar semiconductor devices	18.685	512	0,47
8901	Cruise ships, excursion boats, ferry-boats, cargo ships, barges and similar ...	17.586	584	0,48
2701	Coal; briquettes, ovoids and similar solid fuels manufactured from coal.	17.237	610	0,44
8704	Motor vehicles for the transport of goods.	16.900	624	0,87
8528	Reception apparatus for television	16.114	664	0,58
7102	Diamonds, whether or not worked, but not mounted or set.	15.347	702	0,85
2709	Petroleum oils, crude	11.549	914	5,05
6204	Women's or girls' suits, ensembles, jackets, blazers, dresses, skirts	7.977	1069	0,46
2401	Unmanufactured tobacco; tobacco refuse.	2.407	1235	0,07
801	Coconuts, Brazil nuts and cashew nuts, fresh or dried	2.230	1236	0,02
1801	Cocoa beans, whole or broken, raw or roasted.	2.097	1238	0,03
5203	Cotton, carded or combed.	1.414	1242	0,00
2612	Uranium or thorium ores and concentrates.	1.211	1243	0,01
5304	Sisal and other textile fibres of the genus Agave, raw or processed but not ...	1.146	1244	0,00
905	Vanilla	1.075	1245	0,00

Sources: UN, COMTRADE database; IMF, World Economic Outlook Database

For illustrative purposes, Figure 1 assesses the linear relationship between our estimated 2005 PRODYs and the Balassa indexes of revealed comparative advantage (RCA) for 12 countries (China, France, Germany, Greece, Hungary, India, Italy, Korea, Portugal, Spain, Turkey and USA)[7]. Despite the high dispersion of the data, plotting a linear regression line helps in assessing the sign of the correlation between the two indexes. If significant, a negative correlation indicates a general tendency for a country to be specialized in goods with low income content. A positive correlation, in turn, indicates a general tendency for a country to be increasingly specialized in goods with higher income content.

[7] The Balassa RCA index is in Logs. Null coefficients of RCA became missing values..

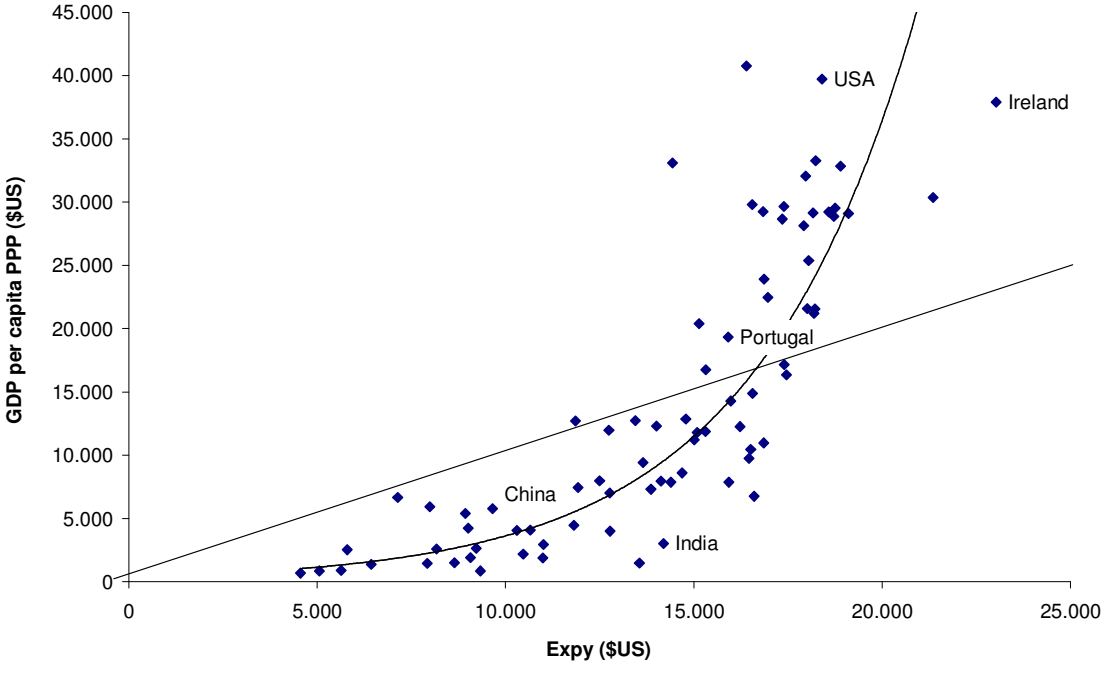
Figure 1: PRODY and Revealed Comparative Advantage in 2005 (China, France, Germany, Greece, Hungary, India, Italy, Korea, Portugal, Spain, Turkey, USA)



According to the figure, by 2005 India was the country in this sub-sample with a more negative correlation between comparative advantage and PRODY values, followed by Turkey, Greece, and China. The Portuguese specialization pattern was more favourable than in these countries, but less than those of Hungary and Spain. On the other hand, Korea, Italy, France, USA and Germany exhibited positive correlations between RCA and PRODY values, suggesting a tendency to be more specialized in "rich country goods".

Moving from a negative correlation towards a positive correlation involves the country becoming increasingly specialized in products with higher income content. This is what is meant by *structural transformation*.

Figure 2: EXPY and GDP per capita at PPP (2005, $US)



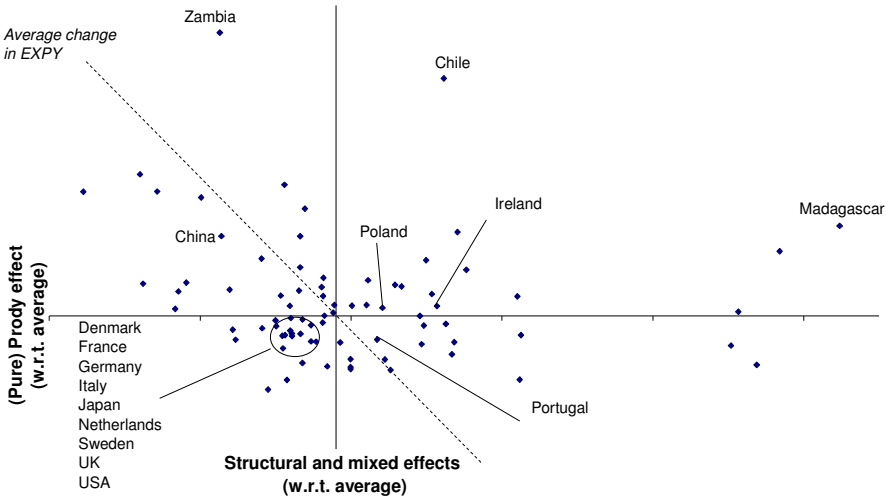
The data in Figure 1 is silent in respect to sector sizes (the RCA index actually measures sizes, but relative to the world average). To account for a country total export mass, Hausmann et al. (2007) proposed the EXPY index. This is the average PRODY in a country export basket, where the weights are the share of each product in a country exports (details in Appendix 1). Figure 2 mimics Figure 3 in Hausmann et al. (2007), relating EXPY values with GDP per capita for the countries in our sample. The figure confirms the positive relation between the two variables, with GDP per capita growing exponentially with EXPY. This supports the idea that rich countries export products that tend to be exported by rich countries, while poor countries export products that tend to be exported by other poor countries. Hausmann et al. (2007) also found that EXPY is a strong and robust predictor of subsequent economic growth, controlling for standard covariates[8]. These findings suggest that the type of goods in which a country specializes has important implications for subsequent economic performance. That being so, structural transformation, e.g., the change in the specialization pattern towards products with higher income content, should be part of the agenda for economic growth.

3. PRODY EFFECT VERSUS STRUCTURAL ADJUSTMENT EFFECT

PRODY indexes change over time, reflecting the changes in the world structure of trade and the changes in per capita GDP levels. Hence, EXPY indexes in two different points in time can either be computed at current PRODY or at base-year PRODY levels. Changes in EXPY at *current* PRODYs will, therefore, reflect changes in the country's structure of exports and changes in the implied value of exports.

Figure 3 describes how the changes in EXPY at current PRODYs break down into a "pure PRODY effect" (i.e., the change in EXPY that would have been observed if the PRODY values of the different products had changed the way they did, while the export structure remained the same) and other effects (this includes a "pure structural transformation effect" – i.e., the value of EXPY which would be observed had the PRODY values remained the same while the structure of exports evolve the way they did – and a mixed effect). The technical details and the figures for 81 countries are in Appendix 2.

Figure 3: Decomposing the changes in EXPY at current PRODYs between 1995 and 2005



[8] In their central case, the estimation results imply that a 10 percent increase in EXPY boosts growth by half a percentage point (p. 15 and Table 8, in the original). Because these results are not significantly affected by the presence of other variables, such as physical capital, human capital and institutional quality, the authors concluded that EXPY exerts an *independent* force on economic growth, leading them to conclude that "countries become what they export".

The horizontal and vertical axes in Figure 3 represent the sample average "pure PRODY effect" and the sample average "pure structural and mixed effects" (respectively) underlying the changes in EXPY values between 1995 and 2005. The dashed diagonal in the figure represents the average growth in EXPY across countries (weighted by GDP per capita in PPP in 2005). Dots to the right of this line represent countries whose EXPY value has increased above the average; dots to the left of the diagonal correspond to countries whose exports have experienced a decrease in income content in relative terms.

The figure reveals that the Portuguese EXPY level has increased slightly above the average, while the reverse happen to most OECD countries (other exceptions include Australia, Canada, Ireland and Poland). Portugal is located in the lower-right quarter of the graph, meaning that the change in the income content of its exports is accounted for by a positive structural transformation (plus mixed) effect, which was big enough to offset a negative PRODY effect. A negative PRODY effect means that, on average, the most important Portuguese exports in 1995 did not improve in terms of income content. In other words, had the Portuguese export basket remained stuck, its average income content would have grown less than the average. The reason is that a significant component of the Portuguese exports basket corresponds to traditional segments, where competition by emerging economies has been increasing. The positive structural transformation effect more than offset this negative effect, allowing the EXPY level in Portugal to grow above the average.

The Portuguese pattern contrasts with what was observed in other OECD countries: most developed countries have registered negative PRODY and structural and mixed effects[9]. In other words, not only their exports became on average less exclusive, but also the change in the structure of exports has led to a less 'rich country goods' export profile[10]. In contrast, Chile and Madagascar, for example, have improved significantly their EXPY values, due to both positive structural adjustment and value effects.

4. INCOME CONTENT, EXPORT SHARES AND EXPORT GROWTH IN PORTUGAL

Having established the relative importance of the structural transformation effect in the case of Portugal, we now focus on this component, abstracting from changes in EXPY caused by changes in PRODY values. Hence, the analysis proceeds at constant PRODYs[11]. In this section and in the following, we use trade data from the Portuguese National Institute of Statistics (INE), which includes data on confidential positions, thus being more accurate than the

[9] This evidence contrasts with Amador et al. (2007), who report a high persistence of the Portuguese specialization pattern, as compared to Spain and Ireland. However, their analysis does not take into account income contents. Weighting the exports shares with PRODY indexes, our analysis suggests that the structural adjustment effect was more significant in Portugal than in the cases of Ireland and Spain (see Appendix 2). Lebre de Freitas and Salvado (2008) discuss, on a comparative basis, how valuable the current productive experience is in preparing the country for further upscale moves.

[10] The analysis for Italy confirms Di Maio and Tamagni (2007). The authors found that the low performance of that country in the last two decades was mainly explained by the fact that Italy remained stuck in a number of products which PRODY values have declined, due to the entry of emerging economies in these markets. In the figure, Italy is on the lower-left corner, meaning lack of structural adjustment and specialization in products of declining value.

[11] Restricting attention to 2005 PRODY values, we are no longer constrained with the need to have a consistent sample of countries for the years 1995 and 2005. Therefore, from this point forward the PRODY values are computed using a larger sample of countries (93 instead of 81), allowing the PRODY index to reflect more accurately the world structure of international trade.

COMTRADE database.[12] The corresponding estimates of EXPY and export shares by classes of PRODY are displayed in Table 2. The table reveals that the average sophistication level of the Portuguese export basket (EXPY) has increased steadily over time, from 14.041 USD in 1990 to 16.603 USD in 2005. This is suggestive of future growth.

To get a sense on how this change came about, export volumes at constant PRODY values are split into 5 classes of PRODY. The 5 classes considered range from the 20% products with higher PRODY values to the 20% products with lower PRODY values (figures for 81 countries based on COMTRADE data are in Appendix 3).

Table 2 – The structure of Portuguese Exports by classes of PRODY

PRODY Class	1990		1995		2000		2005	
	Share on Exports	EXPY	Share on Exports	EXPY	Share on Exports	EXPY	Share on Exports	EXPY
Very High (top 20%)	6,2	1528	8,5	2118	9,4	2363	12,5	3097
High	21,6	4457	25,8	5392	32,8	6982	31,8	6727
Average	14,4	2390	14,2	2363	14,8	2460	16,3	2692
Low	32,1	3743	31,1	3673	27,0	3202	25,6	3049
Very low (20% lowest)	25,8	1923	20,4	1517	15,9	1195	13,9	1036
Total	**100**	**14041**	**100**	**15063**	**100**	**16202**	**100**	**16603**

Sources: own calculations, based on INE data.

Table 3: Structure of exports by classes of PRODY – Portugal

PRODY Class	1990		2005		Growth of exports 1990-2005	
	Exports (10^6 Euros)	Share on Exports	Exports (10^6 Euros)	Share on Exports	% Change	Contribution (percentage points)
Very High (top 20%)	718,1	6,2	3688,8	12,5	413,7	16,7
High	2508,6	21,6	9358,7	31,8	273,1	38,4
Average	1670,4	14,4	4792,1	16,3	186,9	17,5
Low	3737,2	32,1	7534,2	25,6	101,6	21,3
Very low (20% lowest)	3001,0	25,8	4082,1	13,9	36,0	6,1
Total	**11635**	**100**	**29456**	**100**	**153,2**	**100**

Sources: own calculations, based on INE data.

The table shows that there has been a steady increase in the share of products with "High" and "Very High" income content (from a total weight of 27.8% in 1990 to 44.3% in 2005), at the cost of the classes "Low" and "Very Low" (from 57.9% to 39.5%). This suggests that the increase in the average sophistication of the Portuguese export basket was achieved through a re-allocation of resources from products with low and very low implied productivity to products with higher implied productivity.

Table 3 examines the contributions of the different classes of PRODY to the growth rate of Portuguese exports between 1990 and 2005. According to these data, the growth rate of exports (at current prices) between 1990 and 2005 was of 153%. The classes growing above the average were those with "Very High" (413,7%), "High" (273,1%) and "Average" (186,9%) income content. In terms of contributions, these two classes, which represented little more than ¼ of the

[12] A major drawback of COMTRADE is the presence of a sizeable category of miscellaneous products, "9999-Commodities not specified according to kind", which accounted for 2,9% of the world trade in 2005. This category cannot be ignored while computing RCA indexes, but there is no point in computing its PRODY value. Because this category differs significantly over time and across countries, its presence complicates international and inter-temporal comparisons. In the case of Portugal, a major change in the statistical treatment of confidentiality has occurred in 2005, causing a large number of products previously classified elsewhere to be moved to the class 9999. As a result, the share of exports in this category jumped from nearly zero to 8.7%.

exports in the beginning of the period, accounted for 55% of total export growth. This confirms a trend towards a specialization pattern more based on "rich country goods".

5. FDI, EXPORT GROWTH AND STRUCTURAL TRANSFORMATION IN PORTUGAL

In this section, we assess the extent to which Foreign Direct Investment (FDI) had a role in the process of structural transformation of the Portuguese Economy, in the period from 1995 to 2005. For this purpose, we estimate the share of "foreign firms" in the Portuguese exports, by product category, using data collected by the Portuguese Ministry of Labour and Social Solidarity on the composition of firms' capital by nationality of owners. By "foreign firms", we mean those firms in which the proportion of capital owned by non-nationals is equal or greater than 50% (details in Appendix 4).[13]

We first assess the extent to which FDI firms have contributed to the growth of Portuguese exports. In Table 4, product categories are split into 5 groups of similar dimensions, according to their contribution to the growth of Portuguese exports in the period from 1995 to 2005. Here, we see that the top 20% products in terms of contribution to export growth concentrate 83% of the estimated exports by foreign firms in 2005. Table 4 also reveals that 12% of the estimated FDI exports are related to products which exports have declined between 1995 and 2005. Coincidently, this is the only group of products in which the share of FDI in total exports has diminished (from 32% in 1995 to 25% in 2005). This is suggestive of a strong impact of foreign firms on the variation of Portuguese exports, both positively and negatively[14] [15].

Table 4: The role o FDI in Portuguese exports by contribution to export growth

Contribution to export growth between 1995 and 2005:	number of product classes	share of exports (%)		contribution to export growth (%)	share of FDI in total exports (%)		share of exports by foreign firms (%)	
		1995	2005		1995	2005	1995	2005
very high (20% highest)	218	49	72	106	35	40	53	83
high	219	5	5	6	23	27	4	4
median	219	1	1	1	22	28	1	1
low	219	0	0	0	18	34	0	0
very low (lowest 20%)	219	42	17	-20	32	25	43	12
All products	1094	97	96	93	33	36	100	100

Sources: own calculations based on INE and GEP/MTSS, Quadros de Pessoal

Notes: the table does not include data on 140 product classes, for which there is no data available on the presence of FDI; the share of FDI in each group is calculated as the weighted average of the FDI shares in the exports in each product, with the weights given by the share of each product in the exports of the group.

We next investigate the role of FDI in the change of Portuguese specialization pattern, by organizing the export products according to their revealed comparative advantage (RCA) in

[13] Due to data limitations, in this section we restrict the analysis to 1.094 product categories (representing 96% of the Portuguese exports in 2005).

[14] Actually, the direction of causality cannot be disentangled on the basis of the available data: multinational companies also tend to be attracted by fast exporting sectors. In a formal investigation, Magalhães and Africano (2007) find a significant correlation between the stock of (inward) FDI inflows and exports, thus suggesting at least a causality running from FDI to export growth.

[15] It should be noted that these results are influenced by the bigger scale of foreign controlled firms with respect to the nationally controlled ones. To have an idea of the disproportion, in 2005 the average turnover of foreign-controlled firms in Portugal was about 24 times bigger than the average turnover of the remaining firms (source: Quadros de Pessoal database, GEP/MTSS). This figure considers all firms, independently of their involvement in international trade. If we were to consider only exporting firms, the contrast in the scales of foreign-dominated and other firms would surely be lower. Still, if we only consider firms with 50 employees or more, the average turnover of foreign-controlled firms in Portugal in 2005 was about 3.4 times higher than the average turnover of the remaining firms.

1995 and in 2005. In Table 5 we consider four types of products: the 'classics' (i.e., products in which Portugal had a revealed comparative advantage both 1995 and in 2005); the 'marginal' (products in which Portugal did not have a RCA in none of the years); the 'emerging' (products in which Portugal gained a RCA between 1995 and 2005); and finally the 'decaying' (products in which Portugal had a RCA in 1995 but not in 2005)[16].

Table 5: The role o FDI in Portuguese exports by evolution of RCA

Types of products	number of product classes	share of exports (%)		contribution to export growth (%)	share of FDI in total exports (%)		share of exports by foreign frims (%)	
		1995	2005		1995	2005	1995	2005
classics	175	67	54	35	26	26	54	41
rarities	682	12	15	19	33	46	12	20
emerging	110	10	24	45	64	52	21	36
decaying	51	8	2	-5	52	46	13	3
All products	1094	97	96	93	33	36	100	100

Sources: own calculations based on INE and GEP/MTSS, Quadros de Pessoal

Notes: the table does not include data on 140 product classes, for which there is no data available on the presence of FDI; the share of FDI in each group is calculated as the weighted average of the FDI shares in the exports in each product, with the weights given by the share of each product in the exports of the group.

According to Table 5, the 'emerging' products was the group that contributed the most to the increase in exports (45%), reflecting the role of non-traditional products to the expansion of the Portuguese export sector. As far as the role of FDI is concerned, we observe that the "emerging" group is also the one in which the share of FDI in total exports was larger, both in 1995 (64%) and in 2005 (52%). The last column in the right hand side of Table 5 examines the distribution of FDI-commanded exports. The table reveals that the group of "classics" is dominant in FDI exports, but with a loosing weight (41% in 2005, as compared to 54% in 1995). The non-traditional products ("emerging" plus "rarities"), in turn, are of increasing importance and, taken together, already accounted for 56% of the foreign-commanded exports in 2005 (33% in 1995).

We now investigate the role of FDI in the upscale move of the Portuguese specialization pattern. Table 6 analyses the presence of FDI in exports per class of income content. According to these estimates, the share of FDI in total exports increased from 33% in 1995 to 36% in 2005. In 2005, the classes of PRODY with higher presence of FDI were, respectively, the "High" and "Very High" (weights equal to 56% and 43%, respectively). Moreover, in that year, 63% of exports by foreign firms were accounted for by these two classes.

Table 6: The role of FDI in Portuguese exports by classes of PRODY[17]

Prody Class in 2005	number of product classes	share of exports (%)		contribution to export growth (%)	share of FDI in total exports (%)		share of exports by foreign frims (%)	
		1995	2005		1995	2005	1995	2005
very high (20% highest)	217	8	10	13	34	43	9	13
high	235	25	31	40	50	56	40	50
median	216	14	16	19	33	33	14	16
low	215	30	25	17	25	17	23	12
very low (lowest 20%)	211	20	13	4	23	24	14	9
All products	1094	97	96	93	33	36	100	100

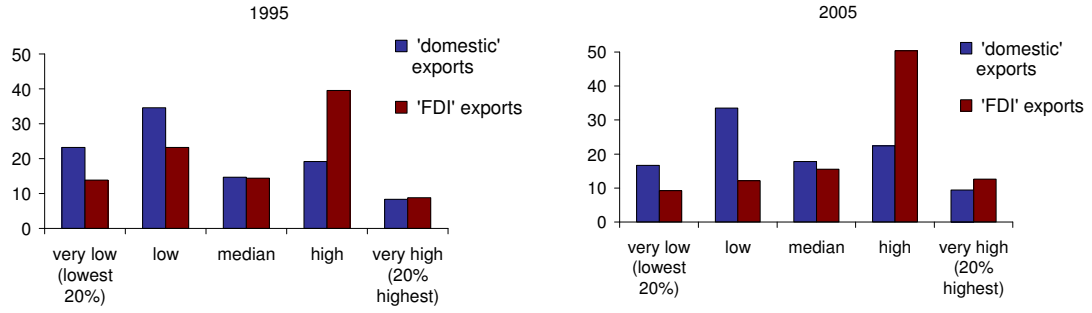
Sources: own calculations based on INE and GEP/MTSS, Quadros de Pessoal

Notes: the table does not include data on 140 product classes, for which there is no data available on the presence of FDI; the share of FDI in each group is calculated as the weighted average of the FDI shares in the exports in each product, with the weights given by the share of each product in the exports of the group.

[16] We partially borrow these expression from Boccardo et al. (2007).

[17] In this and in the following tables, the share of FDI in each group is calculated as the weighted average of the FDI shares in the exports in each product, with the weights given by the share of each product in the exports of the group (for further details see appendix 3).

Figure 4: Share of exports of different PRODY classes



Source: own calculations based on INE and GEP/MTSS

In Figure 4, we compare the distributions of exports by classes of PRODY, for domestic firms and foreign firms, in 1995 and in 2005. We observe that the distribution of FDI-led exports is more biased towards products with higher income content than the corresponding distribution of domestic firms (a simple Chi-square test rejects the hypothesis of equal distributions, at a 1 % significance level). Furthermore, while in the case of 'domestic' exports the shape of the distribution is approximately the same in 1995 and in 2005 (though with an increase in the weight of products with higher income content), in the case of FDI exports there is a visible change in the shape of the distribution (also confirmed by the Chi-square test). In particular, the distribution of foreign-commanded exports by class of PRODY has changed from a bi-modal to a one-modal one, with half of exports concentred in the class of "High" PRODY value.

Finally, we assess whether the increasing role of FDI in exports with high income content refers to traditional or to non-traditional sectors. In Table 7, we cross the information on exports by foreign affiliates per historical status (Table 5) with the information on exports by foreign affiliates per classes of PRODY (Table 6), for the year 2005. We observe that 31% of the FDI commanded exports correspond to "emerging" products with "High" income content and other 11% correspond to "rarities" with Very High income content".

Table 7: FDI exports by evolution of RCA and PRODY class

	Prody Class in 2005					
Type of products	**very low (lowest 20%)**	**low**	**median**	**high**	**very high (20% highest)**	**Total**
classics	6	11	9	14	1	41
rarities	0	0	2	6	11	20
emerging	2	1	3	31	0	36
decaying	0	0	2	0	0	3
All products	9	12	16	50	13	100

Sources: own calculations based on INE and GEP/MTSS, Quadros de Pessoal

Notes: the table does not include data on 140 product classes, for which there is no data available on the presence of FDI; the share of FDI in each group is calculated as the weighted average of the FDI shares in the exports in each product, with the weights given by the share of each product in the exports of the group.

Table 8 illustrates the results discussed in this section by providing information on the 20 product categories that have contributed the most for the growth in Portuguese exports between 1995 and 2005 (these were responsible for 60% of the total increase in exports during this period). In the table we see that FDI accounted for at least 2/3 of the exports in 2005 in 8 out of those 20 product categories. With two exceptions the share of FDI in this FDI-dominated products was already significant in 1995. Only 3 of these 8 cases consist in 'classic' exports, the others being non-traditional products. And in all but two of these products (namely, cigarrets and rubber tyres), the income content is either "High" or "Very High". This table also illustrates the relevance of the automotive and related industries in the processes discussed above: Motor cars and Parts and accessories of motor vehicles, both classified as products with high Prody values, are responsible for 19% of the growth observed in Portuguese exports.

Table 8: Top 20 products in terms of contribution to export growth

Code	Commodity	share of exports in 2005 (%)	contribution to export growth (%)	share of FDI in exports in 1995 (%)	share of FDI in exports in 2005 (%)	Prody value in 2005	RCA class
8.703	Motor cars and other motor vehicles principally designed for the transport ...	7	11	99	84	High	emerging
8.708	Parts and accessories of the motor vehicles of headings 87.01 to 87.05.	4	8	56	66	High	emerging
8.473	Parts and accessories for use with machines of heading 84.69 to 84.72	2	5	28	n.a.	Very High	emerging
2.710	Petroleum oils, other than crude	4	5	0	0	Low	classics
9.401	Seats (other than those of heading 94.02), whether or not convertible into ...	2	3	5	0	Median	classics
4.802	Uncoated paper and paperboard, of a kind used for writing	2	3	1	0	Very High	classics
8.527	Reception apparatus for radio-telephony, radio-telegraphy or radio-broadcas ...	3	3	93	98	High	classics
8.542	Electronic integrated circuits and microassemblies.	2	3	80	95	Very High	marginals
6.109	T-shirts, singlets and other vests, knitted or crocheted.	2	3	31	33	Very low	classics
4.011	New pneumatic tyres, of rubber.	1	3	75	93	Median	classics
7.601	Unwrought aluminium.	1	2	0	12	Median	emerging
2.402	Cigars, cheroots, cigarillos and cigarettes	1	2	4	85	Very low	emerging
3.004	Medicaments (excluding goods of heading 30.02, 30.05 or 30.06)	1	2	38	36	Very High	marginals
8.481	Taps, cocks, valves and similar appliances for pipes, boiler shells	1	1	14	78	High	emerging
7.214	Other bars and rods of iron or non-alloy steel, not further worked than for ...	1	1	0	0	Low	emerging
2.204	Wine of fresh grapes, including fortified wines	2	1	31	18	Low	classics
2.901	Acyclic hydrocarbons.	1	1	5	73	High	classics
4.504	Agglomerated cork (with or without a binding substance)	1	1	8	8	High	classics
8.480	Moulding boxes for metal foundry; mould bases; moulding patterns	1	1	4	6	High	classics
4.503	Articles of natural cork.	1	1	8	8	High	classics
Total of 20 products contributing most to export growth		**39**	**60**	**46**	**50**	**-**	**-**

6. CONCLUSIONS

In this paper, we document that the average income content of the Portuguese exports has grown above the world average in recent years. This evolution is related to a "structural transformation effect" (that is, a shift in the specialization pattern towards products with higher income content), rather than to changes in the implied income of traditional exports. On the contrary, given the increasing competition from emerging economies in the traditional segments, had the Portuguese export basket remained stuck, its average income content would have grown less than the average.

Analysing in greater detail the evolution in the Portuguese export structure, we find an increasing role of the classes of products with "High" and "Very High" income content, both in terms of growth and in terms of contribution to growth. Between 1990 and 2005, these two classes accounted for 55% of the total export growth. Though using a different methodology, our evidence accords with the recent findings of Caldeira Cabral (2008) and Amador et al. (2007) who analysed the changing structure of Portuguese exports following the OECD classification of R&D intensities.

As far as the role of FDI is concerned, we document three stylised facts. First, foreign affiliated firms have played a key role in the growth rate of Portuguese exports. In particular, we observe that the top 20% of products that most accounted for the growth in Portuguese exports concentrate 83% of the estimated exports by foreign firms in 2005. Second, we document that foreign affiliated firms have contributed significantly to the change in the Portuguese specialization pattern. In particular, we find that the share of FDI commanded exports in total exports is higher in the category of products in which Portugal recently achieved comparative advantage. Taken together, the non-traditional exports (e.g, those products in which Portugal had no revealed comparative advantage in 1995) accounted for 56% of the exports by foreign firms. Third, foreign affiliated firms have contributed to the upscale move of the Portuguese specialization pattern. For instance, we find that almost 2/3 of exports by foreign firms in Portugal in 2005 correspond to products with "High" and "Very High" income content. We also observe that the distribution of FDI-led exports is more biased towards products with higher income content than the corresponding distribution for domestic firms and that this bias has increased over time. Taken together, this evidence suggests that FDI has played a relevant role in the Portuguese export performance, both in terms of growth, diversification and upscale movement.

The evidence found in this paper is consistent with Cabral (1996) and Magalhães and Africano (2007), in that FDI has contributed to the expansion of Portuguese exports, and with Gonçalves and Guimarães (1997) in that foreign affiliated firms tend to exhibit a production pattern that differs significantly from that of domestic firms. The evidence in this paper does not, support, however, the IMF (2008, pp 97-103) claim that FDI did not contribute to boosting export performance or to upgrade Portuguese exports. The IMF conclusion is formulated observing that: (i) the sectors which experienced an increase in the shares of FDI since the mid-1990s were typically those with a lower growth of international demand, and (ii) rising FDI flows to "high-tech" sectors were offset by increasing "low-tech FDI". A drawback in the IMF analysis is that the authors used a high level of aggregation and examined FDI financial flows, rather than exports by foreign affiliated firms, broken down by income content, as we do in this paper.

It should be noted that a higher income content on exports does not necessarily correspond to a higher content of domestic value added on exports. As pointed out by Amaral (2006), some low tech traditional industries, like textiles and footwear, have a larger content of domestic valued added than high tech industries, like motor vehicles, which are more dependent on imports of intermediate inputs. Using the input-output methodology, Amaral (2006) estimates a decline in the domestic component of Portuguese exports, from 45% in 1995 to 44% in 2004, due to its changing structure. These estimates ignore, however, the indirect effects of innovation, including learning by doing, knowledge spillovers, information externalities and backward and forward linkages. As mentioned in Section 2, Hausmann et al. (2007) found that

the sophistication of a country export basket is a leading indicator of economic growth, after controlling for the standard co-variates. That being so, the results we reach in this paper that (i) Portugal has moved its specialization pattern towards "rich country goods"; and (ii) FDI has played a positive role in this process, seem to imply that FDI is indeed having a positive impact on the growth prospects of the Portuguese economy.

REFERENCES

AITKEN, B., HARRISON, A., 1999. "Do domestic firms benefit from direct foreign investmjent? Evidence from Venezuela". American Economic Review 89, 605-618.

AMADOR, J., CABRAL, S., MARIA, J., 2007. "International Trade Patterns over the Last Four Decades: How does Portugal Compare with other Cohesion Countries?", Banco de Portugal Working Paper Nº14/2007.

AMARAL, J. , 2006. "Evolução do comércio externo português de exportação: 1995-2004", GEE Papers Nº 1, Gabinete de Estratégia e Estudos, Ministério da Economia e Inovação, Lisbon.

BALASSA, B., 1965. "Trade Liberalization and Revealed Comparative Advantage". The Manchester School of Economic and Social Studies 33(2), pp. 99-123.

BARBOSA, N., GUIMARÃES, P., WOODWARD, D. , 2004. "Foreign entry in an open economy: the case of Portugal", Applied Economics 36 (5), 465-472.

BARBOSA, N., LOURI, H., 2002. "On the determinants of multinationals ownership preferences: evidence from Greece and Portugal". International Journal of Industrial Organization 20 (4), 493-515.

BARBOSA, N., LOURI, H., 2005. "Corporate performance: does ownership matter? A comparison of foreign- and domestic-owned firms in Greece and Portugal". Review of Industrial Organization 27 (1), 73-102.

BOCCARDO, J. , CHANDRA, V., LI, Y., OSORIO, I., 2007. "Why export sophistication matters for growth?". Mimeo.

BORENSZTEIN, E., DE GREGÓRIO, J., LEE, J-W, 1998. "How does foreign direct investment affect economic growth?". Journal of International Economics 45, 115–135.

CABRAL, S., 1996. Comparative export behaviour of foreign and domestic firms in Portugal. Banco de Portugal Working Paper Nº 8/96.

CALDEIRA CABRAL, M., 2008. "Export diversification and technological improvement: recent trends in the Portuguese Economy". GEE Papers Nº 6, Gabinete de Estratégia e Estudos, Ministério da Economia e Inovação, Lisbon.

DALUM, B., LAURSEN, K., VERPAGEN, B., 1999. "Does specialization matter for growth?" Industrial and Corporate Change, 8 (2), 267-288.

DI MAIO, M., TAMAGNI, F., 2007. "The Evolution of the World Trade and the Italian 'Anomaly': A New Look". Università degli Studi di Macerata, mimeo.

EATON, J., and KORTUN, S., 1999. "International patenting and technological diffusion: theory and measurement". International Economic Review 40, 537-570.

EATON, J., and KORTUN, S., 1996. "Trade in ideas: patenting and productivity in the OECD". Journal of International Economics, 40: 251-278.

FEENSTRA, R. C. and A. K. ROSE, 2000. "Putting things in order: patterns of trade dynamics and growth". The Review of Economics and Statistics 82(3), 369-382.

FLORES, R., FONTOURA, M., SANTOS, R., 2007. "Foreign Direct Investment Spillovers: Additional Lessons from a country study". The European Journal of Development Research 19 (3), pp. 372-390.

FOSFURI, A., MOTTA, M., RONDE, T., 2001. "Foreign direct investment and spillovers through workers' mobility". Journal of International Economics 53, pp. 205-22.

GONÇALVES, O., and GUIMARÃES, P., 1997. "O investimento directo estrangeiro na industria transformadora Portuguesa: uma abordagem sectorial e regional através do emprego para o período 1982-1992". Estudos de Economia 16-17 (3), 333-58.

GROSSMAN, G., HELPMAN, E., 1991. "Innovation and growth in the global economy". MIT Press, Cambridge.

GUIMARÃES, P., FIGUEIREDO, O., WOODWARD, D., 2000. "Agglomeration and the Location of Foreign Direct Investment in Portugal. Journal of Urban Economics 47, 115-135.

HAUSMANN, R. and B. KLINGER, 2007. "The structure of the product space and the evolution of the comparative advantage". Center for International Development, Harvard University.

HAUSMANN, R., HWANG, J. and D. RODRIK, 2007. "What you export matters". Journal of Economic Growth 12, 1-25.

HAUSMANN, R. and D. RODRIK, 2006. "Doomed to choose: industrial policy as predicament". Harvard University, mimeo.
HAUSMANN, R. and D. RODRIK, 2003. "Economic Development as Self-Discovery". Journal of Development Economics 72, 603-633.
INTERNATIONAL MONETARY FUND, 2007. World Economic Outlook Database, April.
INTERNATIONAL MONETARY FUND, 2008. "France, Greece, Italy, Portugal, and Spain – Competitiveness in the Southern Euro Area". IMF Country Report No. 08/145.
JOVANOVIC, B., NYARKO, Y. , 1996. "Learning by doing and the choice of technology". Econometrica 64 (6), 1299-1310.
KALDOR, N., 1966. "Causes of the Slow Rate of Growth of the United Kingdom". Cambridge University Press.
KELLER, W., 2004. "International Technological Diffusion", Journal of Economic Literature 42: 752-782.
KUGLER, M., 2006. "Spillovers from foreign direct investment: within or between industries?". Journal of Development Economics 80, pp 444-477.
LEBRE DE FREITAS, M., SUSANA SALVADO, 2008. "Industry similarities, comparative advantage and upscale opportunities for Portugal in the process of structural transformation: an empirical assessment based on 2005 trade data". Mimeo, paper presented at the 2008 meeting of the European Economic Association.
MAGALHÃES, M., and AFRICANO, A., 2007. "A Panel Analysis of the FDI impact on International Trade". NIPE WP 6, Universidade do Minho.
MARKUSEN, J., 2002. "Multinational firms and the theory of international trade". Cambridge MIT Press.
MARKUSEN, J., VENABLES, T., 1999, Foreign direct investment as a catalyst for industrial development. European Economic Review, 335-356.
MATA, J., PORTUGAL, P., 2004. "Patterns of entry, post-entry growth and survival: A comparison between domestic and foreign owned firms," Small Business Economics, 22 (3,4), 283-298.
MURPHY, K., SHLEIFER, A, VISHNY, R., 1989. Industrialization and the big push. Journal o Political Economy 97, 1003-1026.
OECD, 2008. "OECD Economic Surveys: Portugal". OECD, Paris.
PREBISCH, R., 1950. "The Economic Development of Latin America and its Principal problems". United Nations, Lake Success, New York.
SINGER, H., 1950. "The distribution of trade between investing and borrowing countries". American Economic Review 40, 473-485.
RODRIGUEZ-CLARE, A., 1996. "Multinationals, linkages, and economic development". American Economic Review 86, pp. 852-873.
RODRIK, D., 2007. "One Economics, Many Recipes: Globalization, Institutions, and Economic Growth". Princeton University Press, Princeton.
THIRLWALL, A. P., 1979. "The Balance of Payments constraint as an explanation of international growth differences". Banca Nazionale del Lavoro Quarterly Review, vol. 3, 245-253.
TRINDADE, V. 2005, The Big Push, Industrialization and International Trade: the role of exports, Journal of Development Economics 78 (1), 22-48.
UNCTAD, United Nations Conference on Trade and Development, 2007. World Investment Report 2007.

APPENDIX 1: DEFINITIONS OF PRODY AND EXPY

The PRODY index measures the "income content" of each product, as a weighted average of per capita incomes of the countries that export it. For each product i, the PRODY index is computed as:

$$PRODY_i = \sum_{c \in C} \sigma_{ci} Y_c \text{ , where } \sigma_{ic} = \frac{RCA_{ic}}{\sum_{d \in C} RCA_{id}}, \quad RCA_{ic} = \frac{X_{ic}/X_c}{X_i/X}, \quad C = \{1,2,....,M\},$$

where Y_C is real GDP per capita in the c-th country, M is the number of countries and the the weights σ_{ci} normalize the Balassa (1958) index of Revealed Comparative Advantage (RCA) of the c-country with respect to all the countries exporting in the same sector.

EXPY: measures the "sophistication level" of a country export basket, as an weighted average of the PRODYs of the products exported by that country. The income content of a country export basket, EXPY, is computed, for each country, as:

$EXPY_c = \sum_i s_i PRODY_i$, where $s_i = \frac{X_{ic}}{X_c}$ is the share of product i in the exports of country c.

APPENDIX 2 – DECOMPOSING THE GROWTH OF EXPY AT CURRENT PRODY

Let E_i^t be the value of EXPY of country *i* in year *t*, s_{ij}^t the share of product *j* in the total exports of country *i* in year *t*, and Pj^t the PRODY value of product *j* in year *t*. The change in EXPY from *t* to *t+n* can be decomposed as follows:

$$E_i^{t+n} - E_i^t = \sum_j s_{ij}^{t+n}.P_j^{t+n} - \sum_j s_{ij}^t.P_j^t$$

$$= \sum_j \left(s_{ij}^{t+n}.P_j^{t+n} - s_{ij}^t.P_j^t\right)$$

$$= \sum_j \left[\left(s_{ij}^{t+n} - s_j^t\right).P_j^{t+n} + s_{ij}^t.\left(P_j^{t+n} - P_j^t\right)\right]$$

$$= \sum_j \left(s_{ij}^{t+n} - s_{ij}^t\right).P_j^{t+n} + \sum_j s_{ij}^t.\left(P_j^{t+n} - P_j^t\right)$$

$$= \sum_j \left(s_{ij}^{t+n} - s_{ij}^t\right).P_j^t + \sum_j \left(s_{ij}^{t+n}.P_j^{t+n} + s_{ij}^t.P_j^t - s_{ij}^{t+n}.P_j^t - s_{ij}^t.P_j^{t+n}\right) + \sum_j s_{ij}^t.\left(P_j^{t+n} - P_j^t\right)$$

The first component of this expression is the pure "structural transformation effect" (it tells how the EXPY would have changed if the PRODY values of the different products did not change between 1995 and 2005), the last component gives the pure "PRODY effect" (it shows how the EXPY of a country would have changed if there had been no transformation in its export structure), and the component in the middle is the "mixed effect" (which takes into account the fact that the impact of changes in PRODY values on the country's EXPY are amplified when they refer to products which have gained weight in the country's export basket and vice-versa). The following table displays the results of this decomposition for 81 countries.

	EXPY 1995	EXPY 2005	EXPY growth		Pure Prody effect	Mixed effect	Pure structural effect
			rate	rank			
Argentina	9.909	12.964	31%	73	29%	2%	-1%
Australia	11.328	16.762	48%	33	42%	4%	2%
Austria	13.656	18.599	36%	67	35%	2%	-1%
Belize	5.731	7.150	25%	81	42%	-7%	-10%
Bolivia	6.825	11.038	62%	13	29%	-1%	34%
Brazil	10.231	15.063	47%	35	39%	1%	6%
Cameroon	6.681	11.054	65%	12	46%	-7%	27%
Canada	9.200	14.537	58%	19	42%	6%	9%
Chile	9.012	17.340	92%	6	70%	-20%	42%
China	6.875	9.736	42%	47	49%	-3%	-4%
Colombia	8.835	17.240	95%	4	31%	12%	52%
Costa Rica	10.981	13.794	26%	80	39%	-3%	-10%
Cote d'Ivoire	6.963	13.918	100%	3	39%	5%	56%
Croatia	10.800	15.478	43%	44	37%	-1%	7%
Cyprus	10.540	17.699	68%	11	35%	7%	25%
Czech Rep.	12.360	18.053	46%	38	38%	0%	7%
Denmark	13.468	18.578	38%	61	36%	2%	0%
Dominica	5.680	8.071	42%	46	32%	-13%	23%
Ecuador	7.418	12.810	73%	8	41%	7%	25%
Estonia	10.810	16.380	52%	29	39%	-2%	14%
Finland	14.324	19.569	37%	66	35%	-1%	2%
France	13.077	18.493	41%	48	37%	3%	2%
Germany	14.054	19.363	38%	62	36%	1%	1%
Greece	9.828	15.363	56%	23	37%	9%	11%
Guatemala	6.419	10.376	62%	14	41%	-5%	26%
Honduras	4.365	9.321	114%	2	47%	-2%	69%
Hong Kong SAR	11.293	17.337	54%	26	34%	3%	16%
Hungary	11.332	18.071	59%	16	37%	0%	22%
Iceland	13.440	18.952	41%	52	31%	6%	4%
India	9.322	14.455	55%	25	43%	4%	8%
Ireland	14.585	23.438	61%	15	39%	16%	6%
Israel	12.411	18.550	49%	31	43%	6%	0%
Italy	12.880	17.886	39%	59	36%	2%	1%
Japan	14.547	19.575	35%	70	34%	1%	0%
Jordan	8.314	11.962	44%	43	46%	-12%	10%
Kazakhstan	9.216	14.460	57%	21	56%	-9%	11%
Kiribati	4.527	5.854	29%	77	55%	-60%	35%
Kyrgyzstan	6.968	9.237	33%	72	80%	-62%	14%
Latvia	10.023	15.236	52%	28	49%	-5%	8%
Lithuania	10.177	15.041	48%	34	45%	-2%	5%
Madagascar	4.205	9.458	125%	1	50%	-5%	80%
Malawi	2.921	4.589	57%	20	38%	-5%	24%
Malaysia	12.387	17.095	38%	60	31%	0%	7%
Maldives	7.396	12.827	73%	7	49%	-13%	37%
Malta	13.293	18.710	41%	53	31%	5%	5%
Mauritius	7.582	11.988	58%	18	34%	1%	23%
Mexico	12.152	16.998	40%	54	35%	0%	5%
Morocco	6.791	10.775	59%	17	42%	-6%	22%
Mozambique	4.692	6.528	39%	58	55%	-86%	70%
Netherlands	13.044	17.928	37%	63	35%	1%	1%
New Zealand	11.848	17.120	44%	40	41%	0%	3%
Nicaragua	5.901	8.213	39%	57	57%	-31%	13%
Niger	3.985	5.159	29%	76	35%	-22%	17%
Norway	12.673	16.532	30%	75	36%	-3%	-3%
Oman	11.195	15.379	37%	64	37%	-4%	4%
Panama	6.111	10.357	69%	9	44%	-14%	39%
Paraguay	6.713	9.031	35%	71	36%	-10%	8%
Peru	6.233	8.984	44%	42	54%	-12%	2%
Poland	10.916	16.730	53%	27	39%	1%	13%
Portugal	11.058	16.394	48%	32	35%	5%	9%
Rep. of Korea	12.787	18.280	43%	45	34%	0%	9%
Rep. of Moldova	8.213	10.547	28%	78	41%	-15%	2%
Romania	10.241	14.465	41%	50	39%	-2%	4%
Saudi Arabia	10.863	15.360	41%	49	41%	-1%	2%
Singapore	13.903	18.792	35%	69	32%	3%	1%
Slovakia	11.472	17.148	49%	30	39%	2%	8%
Slovenia	12.629	18.561	47%	36	41%	3%	4%
Spain	12.507	17.475	40%	55	38%	1%	1%
Sweden	14.143	19.332	37%	65	37%	1%	-1%
Switzerland	15.117	21.842	44%	41	38%	6%	0%
TFYR of Macedonia	8.939	12.107	35%	68	42%	-8%	2%
Thailand	11.246	16.484	47%	37	32%	3%	11%
Togo	6.153	8.039	31%	74	42%	-40%	28%
Trinidad and Tobago	8.994	14.064	56%	22	52%	-9%	13%
Tunisia	8.683	12.668	46%	39	31%	4%	12%
Turkey	9.124	14.247	56%	24	33%	6%	17%
Uganda	4.493	8.732	94%	5	34%	-8%	68%
United Kingdom	13.689	19.312	41%	51	38%	2%	2%
Uruguay	10.645	13.523	27%	79	28%	2%	-3%
USA	13.700	19.078	39%	56	35%	2%	2%
Zambia	3.376	5.701	69%	10	76%	-87%	80%

APPENDIX 3 – EXPORT SHARES BY CLASS OF PRODY

	1990						1995						2000						2005					
	Very low	Low	Median	High	Very High	Total	Very low	Low	Median	High	Very High	Total	Very low	Low	Median	High	Very High	Total	Very low	Low	Median	High	Very High	Total
Argentina							39	25	20	12	4	100	35	27	20	13	4	100	36,8	26,1	22,7	10,7	3,8	100
Australia	26	20	33	14	7	100	24	18	33	15	11	100	20	22	33	13	13	100	17,5	18,6	42,4	11,7	9,8	100
Austria							6	16	20	34	24	100	5	14	18	37	24	100	4,3	17,1	18,6	35,3	24,7	100
Belize							69	27	2	3	1	100	62	36	1	1	0	100	67,8	29,4	1,3	0,7	0,7	100
Bolivia							64	23	11	2	0	100	61	11	14	12	1	100	42,3	18,1	37,9	1,1	0,5	100
Brazil	28	25	24	17	6	100	30	20	24	19	6	100	24	18	24	26	8	100	23,8	21,8	24,2	21,5	8,8	100
Cameroon							38	51	10	1	0	100	24	68	8	0	0	100	28,5	64,1	6,7	0,5	0,2	100
Canada	9	18	24	33	17	100	8	17	22	35	18	100	6	18	24	35	18	100	6,1	21,1	27,5	30,4	15,0	100
Chile	64	22	7	6	1	100	59	20	9	11	2	100	56	22	10	10	2	100	65,6	17,9	8,9	5,3	2,3	100
China							27	24	18	20	12	100	20	21	18	25	15	100	13,0	16,7	18,2	30,1	22,0	100
China Hong Kong SAR							15	17	18	25	24	100	13	15	17	27	29	100	9,7	11,0	13,4	25,0	40,9	100
Colombia							49	28	14	6	3	100	29	43	16	9	3	100	30,8	36,0	21,6	8,3	3,3	100
CostaRica							64	19	9	5	3	100	32	13	10	5	40	100	30,5	14,0	13,7	7,5	34,2	100
CotedIvoire																			54,8	37,8	3,2	3,4	0,8	100
Croatia							20	29	27	16	8	100	17	29	32	14	9	100	12,5	31,3	29,7	17,3	9,2	100
Cyprus	32	32	21	10	6	100	42	21	21	6	10	100	40	23	16	12	9	100	7,7	27,2	13,6	13,6	37,9	100
Czech Rep.							9	21	27	30	12	100	6	16	26	40	12	100	4,5	14,5	24,1	44,3	12,7	100
Denmark	10	17	27	20	26	100	9	16	27	21	27	100	9	19	21	21	30	100	7,7	19,6	19,7	22,5	30,6	100
Dominica							81	10	6	2	1	100	59	25	14	2	1	100	53,3	28,6	16,2	1,0	1,0	100
Ecuador							54	41	2	2	1	100	38	56	3	2	1	100	30,8	64,4	2,3	1,7	0,9	100
Estonia							21	29	21	15	14	100	13	26	20	11	30	100	8,9	28,6	20,1	17,1	25,2	100
Finland	6	15	14	24	41	100	4	14	15	25	42	100	3	13	12	21	51	100	3,7	13,5	11,5	22,1	49,2	100
France							7	17	22	35	19	100	6	15	20	35	25	100	5,4	15,9	19,2	36,0	23,6	100
Germany *	5	11	20	41	23	100	5	10	19	41	24	100	4	9	18	43	26	100	3,8	9,4	17,3	42,7	26,8	100
Greece	35	39	17	7	3	100	32	37	18	9	4	100	25	37	18	11	10	100	19,8	32,7	20,8	13,6	13,1	100
Guatemala							66	17	10	3	4	100	56	24	13	3	4	100	59,7	20,6	11,5	3,8	4,3	100
Honduras							86	9	3	1	0	100	72	18	7	3	0	100	61,4	23,4	8,2	4,8	2,2	100
Hungary							16	23	29	21	11	100	7	13	19	43	18	100	4,6	12,4	19,0	42,3	21,7	100
Iceland	46	37	12	4	1	100	49	32	12	4	2	100	43	29	21	3	3	100	38,7	27,2	20,8	6,7	6,6	100
India	48	13	27	7	5	100	44	15	26	9	6	100	37	20	25	10	8	100	27,2	26,7	26,2	12,1	7,8	100
Ireland							4	11	17	25	43	100	2	6	8	21	63	100	1,8	6,1	6,7	17,8	67,6	100
Israel							8	11	41	14	26	100	5	7	37	15	35	100	4,1	7,0	51,6	12,6	24,7	100
Italy							9	21	24	28	19	100	8	20	23	28	20	100	6,9	20,2	23,1	28,4	21,4	100
Japan	2	5	18	46	29	100	1	5	18	42	33	100	1	5	16	42	36	100	1,5	6,1	16,9	42,1	33,3	100
Jordan							42	24	11	11	11	100	29	22	25	15	10	100	43,1	24,1	16,6	7,2	9,0	100
Kazakhstan							27	38	23	7	4	100	18	67	12	2	1	100	14,9	74,5	8,9	1,4	0,3	100
Kiribati							79	21	0	0	0	100							81,2	11,5	6,6	0,2	0,5	100
Kyrgyzstan							41	31	10	15	3	100	62	25	6	6	2	100	64,3	22,6	7,3	4,8	1,0	100
Latvia							19	44	17	11	9	100	18	48	19	7	7	100	12,5	46,0	19,9	12,9	8,7	100
Lithuania							15	40	26	12	7	100	18	43	22	12	5	100	10,9	44,9	23,6	15,2	5,4	100
Madagascar	86	9	3	1	1	100	85	10	4	2	0	100	64	29	3	3	1	100	70,7	20,8	3,9	2,4	2,1	100
Malawi	95	3	1	1	0	100	93	3	2	1	0	100	93	2	3	1	0	100	92,6	3,1	2,2	1,6	0,5	100
Malaysia	23	30	14	12	21	100	17	14	17	22	30	100	9	12	17	23	39	100	11,0	14,4	18,1	25,2	31,3	100
Maldives							61	39	0	0	0	100	65	35	0	0	0	100	19,3	75,2	2,5	2,3	0,8	100
Malta							13	6	42	9	31	100	7	8	7	10	68	100	5,4	4,2	56,4	20,3	13,7	100
Mauritius							72	17	5	2	4	100	74	15	4	3	3	100	60,0	13,1	5,9	4,2	16,9	100
Mexico	12	48	10	24	5	100	11	25	17	36	11	100	8	22	17	38	15	100	6,4	27,0	18,7	34,2	13,7	100
Morocco							64	26	6	4	1	100							51,6	32,5	11,0	3,5	1,3	100
Mozambique							79	12	4	4	1	100	52	27	19	1	1	100	75,8	13,5	8,3	2,1	0,3	100
Netherlands							11	18	22	26	23	100	9	17	16	28	30	100	8,7	18,5	18,1	25,0	29,7	100
New Zealand	9	28	25	28	10	100	11	26	27	26	10	100	10	25	28	26	12	100	8,5	24,4	29,8	24,9	12,4	100
Nicaragua							62	20	4	7	7	100	75	17	6	1	1	100	64,8	26,3	6,4	1,5	1,0	100
Niger							90	3	2	5	0	100	78	6	5	11	1	100	88,3	2,9	2,6	5,7	0,5	100
Norway							6	54	23	8	10	100	3	64	20	5	7	100	2,8	58,7	26,4	4,8	7,3	100
Oman	2	93	1	3	0	100	6	82	3	8	1	100	4	85	3	7	1	100	1,9	73,2	23,2	1,3	0,4	100
Panama							73	14	9	1	4	100	57	29	10	1	2	100	56,3	35,3	7,4	0,6	0,5	100
Paraguay	78	19	3	0	0	100	81	14	3	1	1	100	80	14	4	1	1	100	69,0	21,5	7,5	1,0	1,0	100
Peru							73	21	4	1	1	100	70	23	4	2	1	100	71,8	21,3	4,2	2,1	0,6	100
Poland							20	24	37	14	7	100	13	20	35	24	7	100	8,6	19,9	34,1	29,3	8,0	100
Portugal	26	32	15	21	6	100	20	31	15	25	8	100	16	27	15	32	9	100	14,4	26,5	16,9	31,5	10,7	100
Rep.of Korea	17	25	20	21	17	100	9	17	21	27	24	100	7	16	19	28	30	100	3,1	13,5	18,4	33,3	31,6	100
Rep.of Moldova							27	51	11	7	4	100	34	48	7	8	3	100	34,5	48,3	9,9	5,4	1,9	100
Romania	8	43	25	21	4	100	23	38	23	11	4	100	27	35	19	11	7	100	20,2	38,7	20,6	16,8	3,7	100
SaudiArabia							1	89	7	3	0	100	1	93	4	3	0	100	0,6	88,0	8,0	2,9	0,4	100
Singapore	10	24	13	31	22	100	6	12	12	36	34	100	4	12	10	29	44	100	2,7	15,2	10,2	23,0	48,9	100
Slovakia							11	30	28	23	8	100	8	25	23	37	7	100	5,3	26,9	23,1	36,4	8,3	100
Slovenia							9	13	31	32	15	100	6	11	32	36	16	100	4,6	11,5	28,7	38,2	17,0	100
Spain	8	26	20	37	9	100	8	20	19	41	11	100	8	20	20	40	12	100	7,4	20,1	20,8	37,4	14,3	100
Sweden							3	13	14	33	37	100	2	12	14	28	43	100	2,9	14,9	14,8	29,4	38,1	100
Switzerland	4	8	18	25	44	100	3	8	17	26	46	100	2	7	17	23	50	100	2,7	8,8	12,3	19,4	56,8	100
TFYR of Macedonia							39	33	14	11	3	100	38	47	7	6	2	100	42,7	43,2	6,1	5,3	2,7	100
Thailand	45	16	14	11	14	100	31	15	17	21	15	100	21	14	20	20	24	100	17,2	14,9	22,1	30,1	15,7	100
Togo							68	25	3	3	1	100	82	10	4	3	0	100	65,4	26,9	6,7	0,6	0,5	100
Trinidad and Tobago							6	67	24	2	1	100	3	62	33	1	0	100	1,6	52,1	44,6	0,9	0,8	100
Tunisia							53	32	6	6	2	100	48	36	6	9	2	100	38,6	38,7	7,5	12,7	2,5	100
Turkey	48	33	11	6	2	100	41	35	14	8	2	100	34	30	18	14	3	100	24,4	29,8	24,9	17,8	3,1	100
Uganda							94	3	2	1	0	100	75	10	12	1	2	100	80,0	12,6	2,5	2,2	2,7	100
United Kingdom							5	17	17	33	29	100	3	16	16	30	34	100	3,3	17,7	15,9	27,8	35,2	100
Uruguay							38	27	22	11	2	100	33	32	20	11	4	100	22,3	42,4	24,3	7,2	3,7	100
USA							10	10	16	35	30	100	7	9	15	36	34	100	6,5	9,3	16,3	35,8	32,1	100
Zambia							91	5	3	1	0	100	85	9	2	2	1	100	90,2	5,7	3,4	0,4	0,3	100

* Former Fed Rep. of Germany in 1990

APPENDIX 4 – ESTIMATING THE ROLE OF FDI IN EXPORTS

Although we have data on exports at the product level (including confidential positions), we do not know how much of these exports are conducted by foreign-controlled firms. In order to estimate the share of FDI in the total exports of each product category, we used the «Quadros de Pessoal» database, which is compiled by the Portuguese Ministry of Labour and Social Solidarity. This database includes information on every firm with employed labour in Portugal, and contains a variable measuring the proportion of each firm's capital held by non-nationals.

We start with the concordance tables between the Combined Nomenclature of goods (at the 4 digit level of desegregation) and NACE (the Classification of Economic Activities in the European Community, at the 4 digit level of desegregation) for 1995 and 2005. There is a bi-univocal relation for 84% of the CN codes, but some of the product categories have more than on corresponding NACE code, as shown in the following table:

	CN codes	
Number of NACE codes for each CN code	**N.**	**%**
1	924	84
2	139	13
3 or more	24	3
Total	1094	100

Using the information in "Quadros de Pessoal", we computed the share of foreign-controlled firms (defined as those firms in which the proportion of capital owned by non-nationals is equal or greater than 50%) in the total sales turnover of each industry. Then, the share of FDI in the exports of each CN category was computed as the weighted average of FDI shares each industry turnover, with weights given by the turnover of that industry. Formally,

$$FX_i = \sum_j a_{ij} FT_j \ ,$$

where FX_i is the share of FDI in the exports of product i; FT_j is the proportion of foreign-affiliated firms' turnover in the total turnover of industry j; and α_{ij} is the weight of industry j in the total turnover of industries associated with the product i (according the concordance tables), i.e.,

$$a_{ij} = T_{ij} \Big/ \sum_j T_{ij} \ ,$$

where $T_{ij} = \begin{cases} \text{turnover of industry j} & \text{if j is associated with product i} \\ 0 & \text{otherwise.} \end{cases}$

UN NUEVO ORDEN ECONÓMICO: APRENDIENDO DE LA EXPERIENCIA

JOSE MARIA CANTOS
E-mail: JoseMaria.Cantos@uclm.es
RAQUÉL ALAMO CERRILLO
E-mail: Raquel.Alamo@uclm.es
Facultad de Ciencias Sociales de la
Universidad de Castilla-La Mancha

Resumen

La economía mundial está sufriendo una crisis comparable a la de 1929, por sus causas y por su dimensión. La mala praxis de la política económica ha conducido a una sobrevaloración de activos financieros y reales a partir de un continuo exceso de dinero en el circuito, en este caso, fruto de la descomposición del sistema de Bretton Woods. Más allá de las políticas a corto plazo para restablecer los precios de mercado y evitar la depresión de muchas economías nacionales, es preciso poner las bases de un nuevo orden económico internacional que considere el nuevo mapa económico del mundo. El planteamiento de la nueva estrategia debe realizarse partiendo del análisis económico y de la experiencia en política económica (aciertos y errores) en torno a 4 ejes: la política monetaria, las expectativas de los agentes, el comportamiento del sector público y la competitividad internacional.

Palabras clave: Recesión económica, política monetaria, keynesianismo, neutralidad del dinero, competitividad, expectativas, equivalencia ricardiana, nuevo orden económico
Área Temática: Economía del Crecimiento económico.

Abstract

The world economy is suffering a crisis comparable to that of 1929, for his causes and for his dimension. The bad praxis of the economic policy has driven to an overvaluation of financial and real assets from one it continued excess of money in the circuit, in this case, by the decomposition of the Bretton Woods' system. Beyond the political ones in the short term to restore the prices of market and to avoid the depression of many national economies, it is necessary to put the bases of a new economic international order that it considers to be the new world economic map. The exposition of the new strategy must be realized from the economic analysis and the experience in economic policy (wise decisions and bugs) concerning 4 axes: the monetary policy, the expectations of the economic agents, the behavior of the public sector and the international competitiveness.

Thematic Area: Growth Economics.

INTRODUCIÓN

El nuevo siglo guardaba celosamente una desagradable sorpresa a las principales economías del mundo, si bien y a diferencia de lo ocurrido en los últimos 80 años, la virulencia con que se ha producido ahora el cambio en los principales indicadores económicos ha constituido un fenómeno insólito por su rapidez y por su generalización a nivel mundial. A su lado, la todavía irresuelta crisis japonesa de los años noventa, así como la denominada crisis de las "puntocom" centrada en los países tecnológicamente más avanzados entre 2001-2002, no constituyen sino meras erupciones premonitorias de lo que ocurriría después.

Mucho se ha escrito en los últimos años sobre el riesgo de colapso del sistema económico en los Estados Unidos y en algunos países, donde el crecimiento se ha apalancado en el fuerte endeudamiento exterior merced a unos tipos de interés anómalamente bajos, en una superabundancia de crédito internacional, en unos crecimientos extraordinariamente elevados de los beneficios empresariales y, como consecuencia de todo, en una sobrevaloración de los activos. Mientras en 1980 los activos financieros ascendían a 1,2 veces el PIB mundial, en 2007 este coeficiente alcanzó 3,56 veces el PIB[1]. A todo este panorama provocado por las prácticas de crédito fácil, se ha unido en varios países avanzados una estrategia de crecimiento muy sesgada hacia la actividad inmobiliaria y de la construcción, al abrigo de las exigencias de la competitividad internacional, que ha terminado por infligir un enorme impulso a los precios de los activos de esta naturaleza.

A lo largo de 2008, muchas economías del mundo han entrado en recesión. Con unas tasas de desempleo que han crecido a un ritmo muy elevado, en un contexto donde las bolsas de valores han sufrido enormes depreciaciones en sus índices, donde el crédito se ha retraído hasta experimentar crecimientos próximos a cero y el precio de los activos reales se ha desplomado, el año 2009 se afronta con pesimismo ante la falta de respuesta clara a la situación planteada a nivel global[2].

El objetivo de este trabajo es abundar en la necesidad de plantear soluciones globales a partir de una nueva concepción de las relaciones económicas entre países y de la vuelta a los fundamentales del conocimiento teórico acumulado durante más de dos siglos de capitalismo. Por eso, tras partir de un breve diagnóstico de la crisis económica que sacude a muchas economías nacionales, en el apartado dos se profundiza sobre los conocimientos teóricos orientados a cuatro grandes cuestiones, la formación de las expectativas de los agentes económicos, el dinero y la política monetaria, el papel del sector público y la política fiscal y, finalmente, el comercio internacional y la moderna noción de competitividad. Para terminar, en el apartado tercero se trasladan unas conclusiones razonadas y ordenadas a partir de las cuatro cuestiones citadas, que bien podrían informar el nuevo orden económico que será necesario construir.

1. EL DIAGNÓSTICO DE UNA GRAN RECESIÓN

Como ha reconocido el Presidente del Comité Bancario de Basilea, en vísperas de la "cumbre" de Londres del G-20 (2/04/2009), la relajación de las autoridades monetarias frente al

[1] El Informe publicado por el McKinsey Global Institute en octubre de 2008, elabora el citado indicador al que denomina de "profundidad financiera". Los activos financieros (acciones y similares, deuda pública, deuda privada y depósitos) crecieron en 2007 un 18% nominal hasta alcanzar los 196 billones de dólares, aunque lo vienen haciendo a un ritmo anual del 9% desde 1980. Entre los países con un índice superior a 3,5 veces el PIB se encuentran España, Francia, Gran Bretaña, Islandia, Sudáfrica, Australia y México. Para EEUU, este coeficiente que ascendía a 5,09 veces el PIB en 2006, se ha reducido drásticamente a finales de 2007 como consecuencia de la pérdida de valor de los activos financieros.

[2] Las previsiones de marzo del Banco Mundial apuntan a una recesión de la producción global para 2009 de 2 puntos. Asimismo, la Organización Mundial de Comercio también prevé en marzo un retroceso del comercio del 9% para este año, algo desconocido desde la IIGM.

elevado ritmo de creación de dinero que han impulsado las grandes corporaciones financieras, se encontraría entre las causas del colapso actual, interpretación alineada con los planteamientos monetaristas de corte clásico que sobre-ponderan el papel de las instituciones financieras para explicar el comportamiento de la economía real. En efecto, se podría argumentar que la progresiva aceptación sin remilgos de un sistema fiduciario que renuncia en los años setenta al corsé del oro para abrazar prácticas hasta entonces no ortodoxas de la política monetaria, la concesión de carta de normalidad a los productos derivados financieros, unido al fácil recurso al endeudamiento exterior para financiar el déficit de la balanza comercial, han creado un castillo de naipes sobre unas bases inestables amparadas en unas expectativas de beneficios y de crecimiento que se retroalimentaban año tras año en muchos países[3]. Una situación descrita premonitoriamente por Galbraith (1954) con motivo de su análisis del Crash de 1929[4].

Frente a este enfoque instrumentalista del dinero, otros buscarían entre las características sociales e institucionales de un país, las principales causas de los desequilibrios sostenidos en el tiempo que afectan a los mercados y a la competitividad internacional. Así, las estrategias fiscales que inciden sobre el equilibrio ahorro-inversión y sobre los incentivos de los agentes económicos, las políticas regulatorias en los sectores clave que afectan a la eficiencia de los mercados, el impacto asimétrico del proceso de globalización económica sobre los mercados internos y sobre los tradicionales patrones que determinan la competitividad internacional de las economías, o el papel del sector público como creador autónomo de rigideces, estarían detrás de la diferente intensidad con la que la crisis está afectando a los países.

Al igual que ocurriera meses antes en EEUU, la entrada en recesión de la economía española se produce una vez constatado el crecimiento negativo del PIB durante los dos últimos trimestres de 2008. La secuencia ya es conocida: todo comienza con una disminución del crecimiento de la economía provocado, casi siempre, por la menor inversión y el empeoramiento del sector exterior. A continuación, la aparición del miedo a un futuro más incierto provoca que el consumo interno disminuya, los inventarios de las empresas crecen con rapidez, lo que contribuye al colapso de un sector industrial que intenta aliviarse del descenso de beneficios reduciendo su actividad y alumbrando desempleo. Las expectativas de rentas futuras de los consumidores caen presas del pánico y provocan la reducción drástica de la demanda de bienes duraderos y de aquéllos con elasticidad renta más elevada. Cuando las economías domésticas están muy endeudadas (como es el caso actual de varios países avanzados), uno de los sectores más afectados es el del crédito que sufre con especial intensidad el impacto de la morosidad y del desapalancamiento, frenando en seco la actividad del sistema financiero.

Junto a la característica de ser una crisis de madurez del sistema de economía de mercado centrada en los países más avanzados, podría decirse que son dos las circunstancias que agravan la recesión mundial actual y que la hacen diferente de otras recientes y la acercan a la de 1929. Primero, el enorme endeudamiento alcanzado por parte de varias de las principales economías del mundo (Estados Unidos, Gran Bretaña, España, Irlanda, Italia, entre otros). Segundo, la concurrencia en muchos países del listado anterior, de una importante sobrevaloración de los

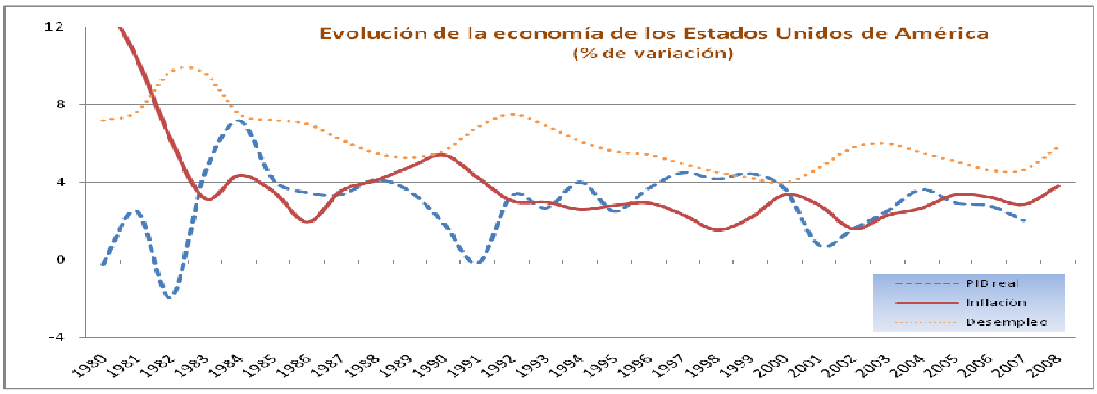
[3] "En ausencia de patrón oro, no hay forma de proteger los ahorros de la confiscación a través de la inflación. No hay depósito de valor seguro. Si lo hubiera, el gobierno tendría que ilegalizarlo, como se hizo en el caso del oro. (...) La política financiera del Estado del Bienestar requiere que no haya modo alguno de que los propietarios de riqueza se protejan. Este es el vil secreto de los estadistas del bienestar en sus diatribas contra el oro. El gasto deficitario es simplemente un plan para la confiscación de la riqueza. El oro (...) representa un protector de los derechos de propiedad. Si uno entiende esto, no tiene ninguna dificultad para entender la hostilidad política hacia el patrón oro" (Greenspan, 1966). Y, sin embargo, Presidente de la Reserva Federal de los Estados Unidos entre 1987-2006.

[4] El gran crash de 1929 comenzó a gestarse en la *burbuja* especulativa de la segunda mitad de la década de los veinte. El autor relata el gran crecimiento que experimentaron los precios de los activos financieros e inmobiliarios durante la década de 1920 donde, al igual que ahora, hubo algunos avisos previos (traslademos el escenario a la reciente crisis de las puntocom). Entonces, muchas advertencias de economistas y banqueros sobre la formación de *burbujas* insostenibles en los precios, fueron desatendidas y, aunque se hicieron intentos por pinchar la *burbuja* desde 1926, éstos resultaron infructuosos ante el optimismo que provocaban las nuevas invenciones del automóvil y la extensión de la electricidad a todo el territorio norteamericano. El autor se detuvo a analizar el papel que representó el extraordinario endeudamiento que se dio en los años previos al estallido de la burbuja: la proliferación en la bolsa de los investment trust que apalancaban su capital, invertían en acciones cotizadas e, incluso, en otros investment trust, proporcionando así ganancias potenciales que solo eran posibles en una situación de mercado con constantes subidas. Como ahora, casi todo el mundo se olvidó de que el apalancamiento funciona tanto al alza como a la baja.

activos inmobiliarios y de los financieros, alcanzada a partir de un extraordinariamente largo ciclo alcista en sus economías.

Centrándonos en la economía estadounidense, por ser donde se ha producido el estallido de la burbuja acumulada durante la fase ascendente del ciclo (14 años), el gráfico nº 1 muestra una elevada tasa de crecimiento del PIB anual para una economía madura (3,06% en promedio del período 1993-2007). Este ritmo de crecimiento ha sido compatible con una tasa de inflación anual más bien moderada (apenas el 2,64% en promedio del período) y con una tasa de desempleo del 5,2% en promedio anual, indicadores que reflejan una situación impensable a este lado del Atlántico.

Gráfico nº 1



Los rasgos genuinos de la economía norteamericana se vienen sosteniendo en un tradicional recurso al déficit exterior de la balanza de bienes y servicios (ver gráfico nº 2) que llegó a una cifra próxima a los 200.000 millones de dólares en 2006 y que impulsa un elevado crecimiento de la economía compatible con un reducido coste en términos de ajuste del valor de la moneda frente a otras divisas importantes, sobre todo, si se compara con el enorme volumen de endeudamiento que acumula el país (hoy alcanza el 300% del PIB). Como muestra el gráfico nº 3, desde el inicio de la fase alcista del ciclo hasta 2002, no se aprecia una tendencia clara de la relación de cambio con la libra y el yen (incluso con el joven euro).

Gráfico nº 2

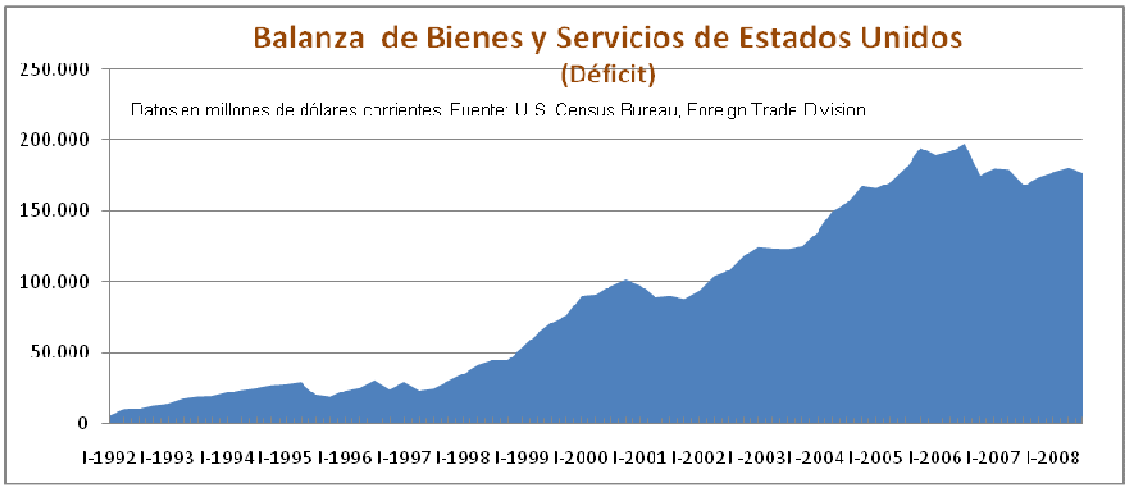
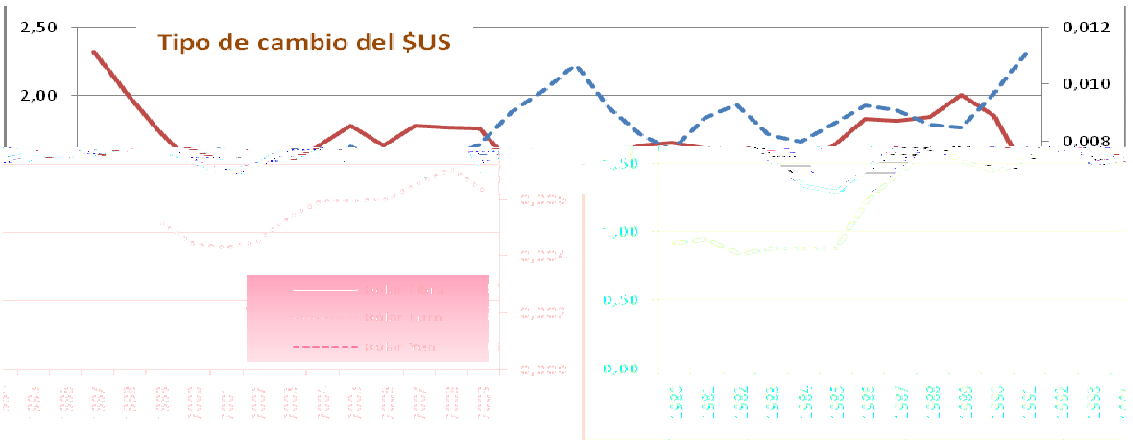


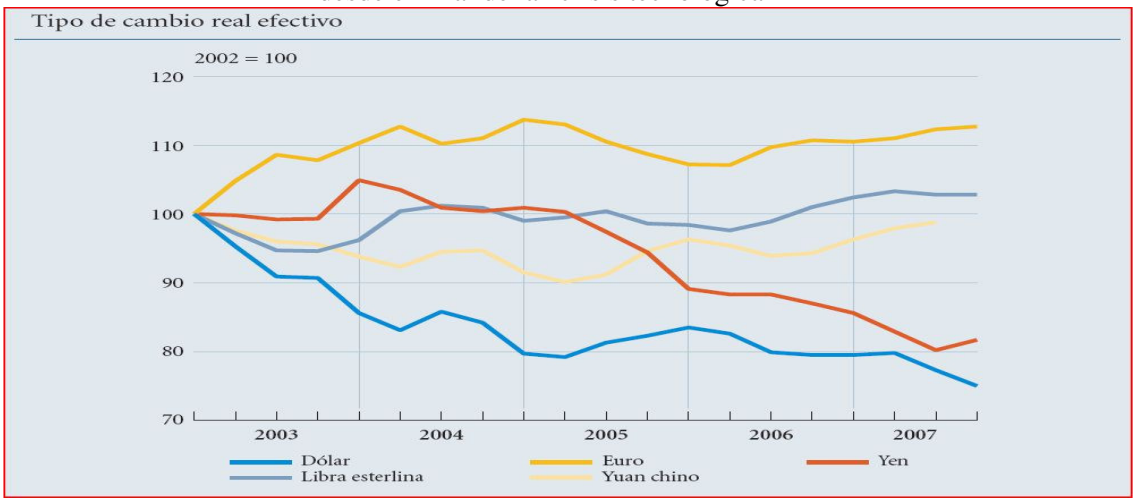
Gráfico nº 3



En cambio, como muestra con mayor claridad el gráfico nº 4, es a partir del fin de la crisis de las tecnológicas cuando el billete verde se deprecia frente al yuan chino, frente a la libra y, sobre todo, frente al euro.

Gráfico nº 4

Tipo de cambio real efectivo de las principales divisas
desde el final de la "crisis tecnológica"



El mantenimiento de unos tipos de interés muy bajos por parte de la Reserva Federal después de los acontecimientos del 11 de septiembre de 2001 y hasta finales de 2003, junto a su permisividad frente a prácticas crediticias del sector privado con desconsideración del riesgo sistémico, estrategia dentro de la que se encuadran las políticas que arrancan de la Administración Clinton (1999) para extender a la población humilde la concesión de hipotecas sobre vivienda, han hecho recaer en la política monetaria USA la principal responsabilidad por el *credit crunch* internacional, si bien, las mismas prácticas han sido replicadas, en mayor o menor medida, por otros países. Crisis financiera provocada por una laxa política de las autoridades monetarias, consentida y alentada por los gobiernos de aquéllos países que creían posible eludir el agravamiento que viene suponiendo el impacto de las economías emergentes sobre la competitividad internacional, aunque fuera a partir de fórmulas de crecimiento piramidal al modo del conocido como "esquema de Ponzi"[5].

[5] Como refleja el Informe Mackinsey, 2008, pág. 9, el crecimiento de los activos financieros ha sido mucho más intenso en el caso de los capitales, los depósitos y la deuda privados.

Hoy en día, este diagnóstico tiene una aceptación generalizada, con diferencias de matiz que no son importantes. Aunque hoy, lo urgente es proponer una estrategia que evite la vuelta al proteccionismo y que impida entrar en una situación similar a la de los años treinta del siglo pasado, este no es el objeto del presente trabajo, sino hacer una primera aproximación a las circunstancias que han conducido a una recesión mundial con síntomas de depresión y buscar soluciones a largo plazo. No deja de ser curioso que, mientras se acepta por consenso que en las prácticas financieras internacionales es preciso un *back to basic* o vuelta a las prácticas tradicionales del sistema bancario como institución de intermediación financiera, no se reconocen los errores cometidos en la apuesta estratégica por la economía real: los gobiernos siguen resistiéndose a aceptar que la globalización económica pone de manifiesto, con mayor nitidez y crudeza que nunca, cambios en la competitividad internacional de las economías y la urgencia de acometer medidas estructurales en aquéllos países que se ven perjudicados por la misma, incluido Estados Unidos.

La precipitación de los acontecimientos hace que apenas se hable de cómo hay que reorientar las políticas a largo plazo y del "sistema de valores" que debe presidir las relaciones económicas internacionales en el futuro, hasta el punto de que la ronda de reuniones al máximo nivel iniciada en Washington[6], apenas ocupa un ínfimo espacio entre las aportaciones de los economistas, muy centradas en el alocado magma de las medidas paliativas a corto plazo que vienen poniendo en marcha los gobiernos para alejar a sus economías del riesgo de la depresión. No obstante, a diferencia del panorama que inspiraba a lord Keynes en 1936, hoy hay conciencia sobre la necesidad de articular una salida global, a la vez que se cuenta con una rica experiencia de aciertos y errores en la política económica.

2. MEDIO SIGLO DE ORTODOXIA ECONÓMICA

Hasta la década de 1930, la teoría clásica, con los retoques dados por las escuelas liberales del XIX, constituía el enfoque mayoritario en economía que inspiraba de alguna manera las medidas de política económica en los países avanzados. En resumen, estos son sus principales postulados:

1. *Liberalismo político como informador de la teoría explicativa del comportamiento de los agentes económicos*. El «laissez faire» justifica el modelo de capitalismo puro, donde por ejemplo, los ciclos económicos son procesos naturales de ajuste que no requieren de ninguna acción por parte del gobierno (son una manera saludable de purgar los excesos cometidos). La "mano invisible" hace que las empresas produzcan lo que la gente desea, por lo que el gobierno no debe intervenir en los procesos de ajuste: la economía resuelve por sí sola sus problemas y los mercados funcionan en competencia perfecta.

2. *El mercado monetario se ajusta solo*. La renta que no se consume inmediatamente entrará en el mercado monetario como ahorro. Como el dinero inactivo no produce rendimiento, el ahorro volverá a la economía como inversión aumentando el capital productivo (el sistema bancario garantiza que el ahorro captado vuelva a la economía en forma de préstamo). El interés pagado por los prestatarios a los ahorradores asegura el equilibrio entre ahorro e inversión. El mercado monetario alcanza el equilibrio mediante un ajuste en el tipo de interés. Si el tipo de interés es alto, la gente se siente más inclinada a ahorrar o prestar, pero los prestatarios no demandarán grandes préstamos: el mercado tenderá al equilibrio de modo natural.

3. *Los precios y los salarios en la economía son flexibles.* En competencia perfecta, si existe un exceso de fuerza de trabajo (desempleo friccional) o de productos, el salario o el precio de éstos se ajustará para absorber dicho exceso. De este modo, si hay mucho paro, las empresas pueden emplear trabajadores por salarios más bajos y se reducirá el desempleo. Por tanto, no existe desempleo involuntario porque un ajuste en los salarios asegura que los

[6] Summit on Financial Markets and the World Economy (15-11-2008).

desempleados encontrarán otro empleo. Además, como los trabajadores necesitan comprar mercancías, aceptarán trabajos peor pagados.

4. *Teoría del comercio internacional.* El intercambio internacional es, en última instancia, un trueque disfrazado, y los metales preciosos se reparten por sí mismos entre los países que los necesitan, dirigiéndose siempre, de modo automático, a las naciones que poseen un poder adquisitivo en mercancías más elevado, sin que sea posible desvirtuar esta ley[7]. Para Ricardo, el comercio internacional se basa en la cooperación y no en la competencia entre países. Frente al Mercantilismo, la teoría de los costes comparativos predice que todo país saca provecho del libre cambio, aunque se realice de forma unilateral: el país que gana más es el más pobre, puesto que las ventajas del comercio internacional deben apreciarse sólo desde el punto de vista del consumidor. Posteriormente, J. S. Mill revisa esta proposición ricardiana.

5. *La producción queda determinada por el nivel de pleno empleo.* La oferta determina la demanda y el nivel de producción de equilibrio, siendo vertical la curva de oferta. Las variaciones en la demanda se traducen en variaciones en los precios.

6. *La política monetaria es ineficaz.* Las variaciones en la oferta monetaria sólo afectan al nivel de precios, sin que tengan ningún efecto sobre las variables reales (cantidad demandada, producción de equilibrio, salarios, etc., una vez depurado el efecto de los precios): el dinero es neutral.

7. *La Política fiscal también es ineficaz.* Las medidas de política fiscal tampoco influyen en las variables reales porque la economía siempre se encuentra en situación de pleno empleo y solo dan lugar a inflación[8]. No puede decirse, sin embargo, que Ricardo como el resto de los clásicos, crea en los efectos neutrales de la deuda pública.

Hasta principios del siglo XX han existido unos patrones de comportamiento por parte de los teóricos de la economía de mercado que respondían con bastante grado de fidelidad a esos siete postulados de la economía clásica. A su vez, estos postulados se asientan sobre los axiomas generales de que los consumidores tienen por objetivo maximizar la satisfacción que les produce el consumo de bienes y que los empresarios persiguen el máximo beneficio, lo que permite formular, a su vez, las leyes de la oferta y la demanda, leyes inmutables de la economía. La pregunta que cabe hacerse es en qué medida las aportaciones doctrinales posteriores refutan los postulados clásicos y en qué condiciones.

2.1. La ruptura con la tradición clásica

Sobre el primer postulado clásico (liberalismo), Keynes no cuestiona los fundamentos ético/políticos de la economía de mercado, sino la capacidad de este último por sí solo para conducir a la economía al pleno empleo. Ahora bien, aunque cabe reconocer varias alusiones en su obra al problema del crecimiento (Galindo, 2003), no puede decirse que los problemas del desarrollo y del largo plazo constituyeran su preocupación principal. La pérdida de fe en los automatismos reguladores de la economía abrió la puerta a la necesidad de la *política económica* o, más en rigor, a la *gestión macroeconómica*, es decir, a la intervención del Estado para alcanzar una situación de pleno empleo (Bustelo, 1999). Así, el paradigma keynesiano es

[7] Ricardo (1821;1985), pág 79 y ss.

[8] En cuanto a la forma de financiar el gasto público, a Ricardo se le atribuye el conocido principio de "equivalencia ricardiana": el gobierno puede financiar su gasto mediante los impuestos cobrados a los contribuyentes actuales o mediante la emisión de deuda pública. No obstante, si elige la segunda opción, tarde o temprano tendrá que pagar la deuda subiendo los impuestos por encima de lo que estos se ubicarían en el futuro si otra fuera la elección. La elección es entre pagar impuestos hoy o pagar impuestos mañana: si el gobierno decide financiar un gasto adicional a través de déficit, esto supone que los ciudadanos se darían cuenta de que tendrían que pagar impuestos mayores en el futuro y, por lo tanto, ahorrarán un dinero adicional para poder pagar los impuestos futuros. Este mayor ahorro por parte de los consumidores compensaría exactamente el gasto adicional del gobierno, de modo tal que la demanda agregada permanecerá inalterada. Sin embargo, Ricardo reconoce, posteriormente, que aunque este planteamiento es correcto en términos económicos, hay que rechazarlo como irreal. Al financiar el gasto de la guerra con préstamos, el individuo solamente tendrá que pagar 100 libras en un año y "considerará que hace bastante" con ahorrar esa mucho menor suma "para luego engañarse a sí mismo con la creencia de que es tan rico como antes". Así, "la financiación con préstamos es un sistema que tiende a hacernos menos ahorrativos, a cegarnos ante nuestra situación real" (Ricardo, 1817; pág. 247). Una revisión amplia sobre esta literatura se encuentra en Briotti, M.G. (2005).

una teoría *de y para* la crisis de la economía capitalista que busca dar respuesta a dos problemas centrales no resueltos por la ciencia convencional prevaleciente y que se encuentran en la base de estas crisis: la notable brecha entre la situación en que se encuentra la economía real y su nivel de producto potencial, y la concentración de la renta[9]. Esta preocupación le conduce a cuestionar profundamente el supuesto básico con que opera el universo conceptual neoclásico: el estado de equilibrio estacionario en el cual se encuentra de manera permanente la economía. La intersección de curvas de oferta y demanda, lugar en que los agentes que ofrecen o demandan bienes o factores maximizan su bienestar, representa una posibilidad dentro de muchísimas alternativas presentes en el plano cartesiano. Esta comprensión del fenómeno económico lo mueve a sostener que no existen mecanismos de mercado que actúen de manera eficaz frente a los desequilibrios que se observan con regularidad en la economía, de manera que se hace indispensable la presencia de un agente exógeno a las relaciones de mercado: el Estado. Agente que asume un doble rol: controla y administra la demanda agregada para procurar el pleno empleo y promueve el bienestar social para redistribuir el ingreso (argolla de enganche de los desarrollos posteriores del keynesianismo). El elemento que impulsa el crecimiento económico es el consumo, pero aquél no se encuentra vinculado a un *trade off* entre consumo y ahorro, sino al proceso efectivo de transformación del ahorro en inversión.

En relación al postulado número dos anterior (el mercado monetario se ajusta solo), Keynes sostiene que ahorradores e inversores tienen planes incompatibles, por lo que es posible que no se llegue a establecer un equilibrio en el mercado de valores. Ahorradores e inversores son grupos que no siempre se relacionan: están separados por intermediarios financieros (bancos). Durante una recesión, aunque el tipo de interés sea muy bajo, la inversión puede no igualar al ahorro, bien porque los prestatarios tengan pocas perspectivas de ventas, bien porque los bancos no se animen a prestar al considerarlo arriesgado, o bien porque los ahorradores prefieren esperar que haya intereses más altos. Esto causa la denominada "trampa de la liquidez", lo que implica que parte de los ahorros sean improductivos.

Keynes rebate el postulado número tres anterior, al afirmar que los precios y los salarios no son flexibles a la baja, como afirma la teoría clásica. Los salarios tienden un escaso (o nulo) recorrido a la baja porque los trabajadores no aceptan salarios que no les permitan vivir adecuadamente, lo cual se refuerza por las acciones de los sindicatos. Si los salarios se encuentran en ese mínimo, existirá desempleo. En el caso de los precios, las empresas que producen artículos muy caros prefieren reducir la producción y despedir a trabajadores que bajar los precios. Su poder monopolístico normalmente les permite actuar así (no hay competencia perfecta).

Por lo que atañe al cuarto postulado de los clásicos, poco se ha escrito sobre las propuestas keynesianas en relación a su concepción del comercio internacional y su papel en esta nueva economía, quizás porque este asunto constituyó uno de los capítulos más inestables de su pensamiento. Como afirma P. Schwartz (2001, pág. 35), Keynes adoptó a lo largo de su vida una actitud instrumental y cambiante respecto de la libertad de comercio. Su preocupación principal hasta 1930 fue evitar las deflaciones monetarias en Inglaterra causadas por la vuelta al patrón oro después de la IGM, dentro de un marco de pensamiento librecambista heredado de su maestro A. Marshall. A partir de entonces, abraza el proteccionismo debido a que el marco institucional reinante (salarios rígidos, poderosos sindicatos, subsidio de paro generoso…) impedía aumentar el empleo, y lo hace a partir de dos razonamientos: a) el arancel elevaría el nivel de precios nacional sin elevar los costes; b) el superávit comercial alcanzado equivaldría a una inversión en el exterior, aumentando el empleo interno y compensando la reducción inicial del poder de compra de los asalariados (derivada de la subida de precios). Claro que tampoco defendía abiertamente el abandono del librecambismo y, de hecho, si Inglaterra hubiera

[9] "Los principales inconvenientes de la sociedad económica en que vivimos son su incapacidad para procurar la ocupación plena y su arbitraria y desigual distribución de la riqueza y los ingresos" (Keynes, "Teoría General…", 1936).

abandonado el patrón oro en 1930 en vez de en 1932, Keynes probablemente no se habría separado de la doctrina librecambista (Harrod, 1963, pág. 431). Posteriormente, en la Teoría General solo hay un punto donde aborda los aspectos internacionales de la actividad económica, y es donde habla del impacto del grado de apertura de la economía sobre la velocidad de recuperación derivada de las medidas de reactivación, para concluir que cuanto más abierta estuviese la economía al comercio internacional, menor eficacia tendrían las medidas adoptadas, puesto que sus efectos positivos se desbordarían hacia los trabajadores de otros países[10]. Y es que el modelo y la política keynesiana no podrían funcionar en un marco de economías abiertas[11] [12]. Finalmente, Keynes acepta el sistema de Bretton Woods como patrón de un nuevo orden económico internacional más abierto, porque pensaba que los principales países iban a adoptar sus políticas y que la apertura de las economías quedaría limitada de facto[13].

En relación con el postulado del equilibrio de la economía, no existe un equilibrio único de la economía en situación de pleno empleo, puesto que la oferta agregada no es vertical (perfectamente inelástica) cuando existe una situación de bajo crecimiento económico o de estancamiento, sino que puede ser perfectamente elástica. Si el precio es muy bajo, muchas empresas prefieren reducir la producción en vez de vender con pérdidas (lo que ocurre cuando los costes fijos son muy elevados) y se puede dar equilibrio con desempleo (voluntario).

El postulado clásico de la ineficacia de la política monetaria queda reducido por Keynes al supuesto de una situación próxima al estancamiento económico, donde unos tipos de interés muy bajos provocan ineficacia de una política monetaria expansiva, al estar inmersa en una trampa de liquidez. Las malas expectativas empresariales contrarrestan el bajo coste financiero de la inversión.

Finalmente, Keynes discrepa sobre el postulado clásico de la ineficacia de la política fiscal. Cuando la economía se encuentra en bajo crecimiento o estancamiento, entonces un crecimiento de la demanda agregada (provocado, por ejemplo, por un aumento en el gasto público) no se traduce en un aumento exclusivo del nivel de precios (como en el enfoque clásico), sino que incrementa también la producción. En relación con la forma de financiar el gasto público, durante el primer tercio del S.XX la inmensa mayoría de la literatura económica había rechazado el viejo parecer mercantilista de que el pago de los intereses de la deuda no es más que una mera transferencia de la mano derecha a la izquierda[14]. Desde Keynes, se replantea completamente el papel del sector público y de muchos de los principios presupuestarios que habían regido hasta entonces la vida pública: la deuda no es más que un subproducto, un mal necesario, al servicio de las políticas de expansión de la demanda en búsqueda del pleno empleo. La deuda en circulación no supone problema alguno porque, en realidad, "nos la debemos a nosotros mismos", mientras la deuda privada "se debe a otros"[15]. Como afirman Buchanan y Wagner (1977, pág. 43), esta idea del déficit y la deuda como subproducto, parte de la concepción de que hay que sacrificar cualquier efecto derivado sobre el tamaño del sector público o los problemas que presente su financiación, para atender a lo verdaderamente importante —los efectos sobre la economía de una medida expansiva—, asignando un papel instrumental al déficit y a la deuda. Es la vuelta a la neutralidad del dinero.

Como resumen, puede decirse que la teoría keynesiana no invalida con carácter general ninguno de los enumerados principios de la economía clásica, pero sí que los declara

[10] Keynes, J. M. (1936): The General Theory of Employment, Interest and Money; X, iii

[11] Compartimos plenamente la opinión de Schwartz, 2001; Pág. 37.

[12] Según recoge Keynes en la Teoría General (pág. 339), en los años previos al abandono del patrón oro, Inglaterra trató de manejar el tipo de interés para mantener el equilibrio de la balanza de pagos, lo que resultó incompatible con la búsqueda del pleno empleo y condujo al abandono de la paridad con el oro. Poca memoria histórica tuvieron los responsables de la política económica en España entre 1988-1992 (en época del Ministro Solchaga), en lo que se conoció como "estrategia de crecimiento con deflación" y que acabó con tres devaluaciones consecutivas de la peseta en menos de un año. Y es que D. Hume predijo un final favorable de esta estrategia en el marco del patrón oro (léase, paridad fija), pero bajo el supuesto de flexibilidad de precios y salarios.

[13] Harrod (1963), pág. 609; citado por Schwartz (2001).

[14] Tan solo la escuela financiera italiana de finales del diecinueve, defiende la tesis de la equivalencia entre deuda e impuestos.

[15] Para tener una clarificadora visión de esta postura doctrinal véase Lerner, 1948. La creación de deuda no representa una reducción de riqueza nacional, al igual que tampoco reducen la renta de las generaciones futuras el pago de los intereses.

incompatibles con una situación de recesión económica. No obstante, lo más importante, es el corolario que se desprende de todo su planteamiento: se da entrada al Estado como agente económico activo.

2.2. Keynesianismo y socialdemocracia

Es a partir del Círculo de Viena (1922-1936) cuando todos los métodos y validaciones de las ciencias naturales fueron aplicados a la economía y, por ello, los principios económicos deben explicar el comportamiento de cualquier sociedad, al contrario de lo que ocurría con los principios de la escuela neoclásica que explicaban, justificaban y validaban únicamente el sistema capitalista. La experimentación es una característica de las ciencias nomotéticas, como señala Fernández Díaz, por lo que es a partir de entonces cuando también pasa a formar parte del acervo de la Ciencia Económica[16].

Atribuida seminalmente a Hicks (1937), Hansen (que la popularizó en EEUU) y Modigliani (que la formalizó e hizo dinámico el modelo), se da por aceptado que la gran síntesis neoclásica se refiere a las ideas y modelos de la tradicional escuela de pensamiento complementada con ideas y conceptos keynesianos. A pesar de estar inspirada en el aporte keynesiano, los supuestos, desarrollo y resultados son bastante diferentes dependiendo de los autores. La síntesis neoclásica tiene sus raíces en la revolución marginalista de finales del siglo XIX y ha sido la corriente económica principal hasta la década de 1970. Para Eichner (1983), la síntesis neoclásica se asienta en seis supuestos: las curvas de indiferencia, las curvas isocuantas, la pendiente positiva de la curva de oferta, el desarrollo teórico y las aplicaciones de la marginalidad, la IS-LM y la curva de Phillips.

El objetivo de este tipo de análisis es justificar el equilibrio de mercado. Se supone que existe un vector de precios que vacía todos los mercados y hace compatibles los planes de compra y venta de todos los agentes económicos. Cada agente está dotado de racionalidad, de manera que busca toda la información posible en la medida en que se lo permita su restricción presupuestaria. La sociedad es vista como un compendio de seres individuales y nunca como un organismo. Cada consumidor representativo tiene una "conducta optimizadora" que le impele a demandar la cesta de bienes necesaria para maximizar su función de utilidad; asimismo, las empresas tratan de obtener los mayores beneficios posibles al ofrecer cierto nivel de producción. Asumiendo que los mercados están en competencia perfecta, el comportamiento racional de cada agente económico conlleva que el agregado económico en su conjunto también tenga un comportamiento racional, estos es, los resultados agregados son consistentes, hipótesis que se encargó de rebatir Arrow (1951) desde un punto de vista teórico. Para la síntesis neoclásica el pleno empleo existe y, si no se observa, es porque también existen rigideces que impiden el ajuste en las variables.

La "síntesis" se refiere a la conciliación de las dos formas de entender cómo responde la economía ante estímulos exógenos a partir de la evidencia de que el pleno empleo no es el estado natural de las cosas, aunque lo que subyace en este planteamiento es la justificación teórica para una intervención sistemática del Estado en la economía: se trata de uno de los pilares fundamentales de la socialdemocracia. La herramienta básica del modelo se ensambla sobre un vector unidimensional donde se representan los planteamientos clásicos y los keynesianos como representativos de situaciones extremas en las que cabe modular la aplicación de un conjunto acotado de recetas. Unas pocas variables permiten identificar las condiciones en las que se encuentran el mercado de bienes (IS) y el mercado de dinero (LM) en cada momento, pudiendo elegir la combinación de medidas de política económica a aplicar. Nótese que la "síntesis" no cuestiona el fondo de ninguno de los siete postulados del clasicismo enunciados anteriormente (incluido el pleno empleo), sino la oportunidad de su vigencia en ciertos casos. El

[16] Fernández Díaz, A. et al. (1995); Pág. 35. De hecho, el tema de la clasificación de las ciencias es moderno y solo surge con fuerza a partir de la separación entre filosofía y las llamadas ciencias particulares [Fernández Díaz, A. y Rodríguez Sáiz, L. (1982); Pág. 20].

gran aporte diferencial que toma de Keynes, es la necesidad de que el Estado debe erigirse en un vigilante activo de la situación y en el garante de la restauración del equilibrio, en lo que después se dio en llamar un "déspota benevolente" que tanto criticara la Public Choice.

A partir de este principio informador del post-keynesianismo es como pueden entenderse los sucesivos desarrollos que han centrado el grueso de la literatura económica desde los años cincuenta, esto es, el análisis de los efectos de las medidas de política económica bajo determinados escenarios. En un marco teórico que ampara sin ambages la intervención sistemática del Estado en la actividad económica, al *cuanto* intervenir, daban respuesta las instituciones políticas representativas de las preferencias sociales de cada país (socialdemocracia v. sociedades liberales); y al *cómo,* se respondía desde las aportaciones positivas de las corrientes doctrinales dominantes en cada momento a partir de un catálogo de medidas no muy numeroso. No han sido muchos los aportes normativos sobre la economía en los últimos cincuenta años y menos los que han cuestionado algunas de las hipótesis básicas sobre el comportamiento de los agentes económicos ya asentadas a inicios del S. XX.

2.3. El pensamiento económico frente a las grandes cuestiones

En este apartado pretendemos hacer un resumen de cómo las corrientes del pensamiento económico han resuelto algunas de las cuestiones clave de la economía; cuestiones que se traen a colación aquí por la singular actualidad que representan en la gestación y posterior salida de la actual crisis económica. Hay que tener presente que las inercias creadas durante décadas por los diversos tratamientos doctrinales realizados sobre los problemas económicos, son las fuentes de donde han bebido los policy makers para hilvanar sus estrategias en la etapa de la globalización económica que arranca a principios de los años noventa, favorecidas por la caída de barreras políticas que impidieron su aparición décadas antes. Una lección fácil de extraer de estos últimos veinte años, es que la larga etapa del keynesianismo como corriente dominante es deudora del no menos largo período en que el mundo estuvo dividido en dos grandes bloques económicos.

2.3.1. El papel de las expectativas de los agentes económicos

A ningún teórico de la economía en estos últimos 200 años se le ha escapado la importancia que tienen las expectativas de los agentes económicos en la toma de decisiones. Los consumidores, las empresas, los inversores, los pensionistas y, en general, todos los colectivos sociales adoptan decisiones sobre consumo, ahorro e inversión basadas en la percepción que tienen de lo que está ocurriendo y de lo que va a ocurrir. Así, las expectativas no son otra cosa que las previsiones que los agentes realizan sobre la magnitud en el futuro de las variables económicas pero que afectan al momento presente, y lo hacen a partir de su experiencia pasada y de la actual.

Sin embargo, ante algo tan evidente, los economistas han tenido muchas dificultades para plasmar las expectativas en los modelos, dada la complejidad que conlleva conocer esta faceta del comportamiento humano. Hasta la amarga experiencia de la década de los veinte del siglo pasado, la Economía opinaba mayoritariamente que el punto de mira de los agentes económicos a la hora de formar sus expectativas se encontraba, sobre todo, en la capacidad del mercado guiada por la mano invisible. Los problemas sociales que generó la situación de estancamiento económico en Europa y EEUU, junto con el planteamiento keynesiano de un nuevo agente económico, el Estado, omnisciente y capaz de garantizar el equilibrio con pleno empleo (lo que no había sido capaz de hacer el sistema de precios libres), supuso un salto cualitativo importante en el mecanismo de formación de las expectativas. Además de poner el foco de atención en los tradicionales factores institucionales (guerras, alianzas gubernamentales, imperfecciones del mercado en cada país, sistema impositivo e impacto económico del régimen político, en general) que afectan de forma pasiva (no intencional), los agentes económicos deben considerar ahora, además, las decisiones de una política económica, activa e intencional, con el claro objetivo de intervenir en los mercados.

Con el paso del tiempo, para los economistas "el agente-tipo" cada vez va dejando menos al albur de la mano invisible la formación de sus expectativas, mientras que las decisiones de la política económica van centrando mayoritariamente la atención a la hora de formar esas expectativas, en detrimento de la lógica del mercado como fuente de opinión. Hasta que en los años setenta comienzan a adquirir notoriedad los propios "fallos del sector público" como causa de distorsión en las economías. Es entonces cuando la credibilidad de las medidas estatales sufre un gran deterioro (pérdida de omnisciencia del Estado), a la vez que se produce un elevado ritmo de apertura en las economías durante los años noventa que complica el cálculo de expectativas por parte de los agentes económicos. El *estatus quo* internacional que rige desde Bretton Woods se desmorona, con el consiguiente aumento en la complejidad del sistema económico y en la infravaloración de los riesgos y de la incertidumbre derivados de la creciente dificultad para el acceso a la información relevante. Todo este cúmulo de circunstancias ha afectado a la formación de las expectativas haciéndolas más volubles e, incluso, podrían contribuir a explicar por qué los ciclos son más acusados en los últimos veinte años, tanto en la parte alta (burbuja) como en la profundidad de la fase recesiva.

La teoría keynesiana es la primera en formalizar las expectativas, aunque de forma muy primaria. Y lo hizo al introducir esta variable como exógena al modelo a la hora de determinar la demanda de inversión, pero sin explicar cómo se generan aquéllas, porque obedecían a mecanismos que nada tienen que ver con el sistema económico[17]. Sin embargo, en el modelo keynesiano básico las expectativas siguen siendo desconocidas para el consumidor cuando toma decisiones sobre consumo y ahorro. Fue después, en los años sesenta, cuando los modelos keynesianos se ven obligados a introducir de alguna manera las expectativas con motivo de formalizar los primeros intentos de predicción económica: los agentes económicos utilizaban los datos de los últimos años para predecir la situación de los siguientes, lo que en el argot econométrico daba mucho juego al concepto de variable retardada. Surgen, entonces, los modelos con *expectativas adaptativas* (o adaptables), donde los agentes forman su opinión teniendo en cuenta su experiencia reciente[18]. Este argumentario fue rápidamente hecho suyo por los sindicatos de los países de la socialdemocracia para plantear los primeros problemas serios a la política económica en los años sesenta, dando lugar a la espiral salarios-precios que desembocó en una gran catarsis en los años setenta. Se suponía que los agentes ponderaban las tasas de inflación de los últimos años para estimar la de los siguientes; a la del último año se le daba una ponderación alta y a la de años anteriores, más baja. Era un enfoque que podía ser fácilmente expresado en forma matemática e incorporado a los modelos macroeconómicos al uso. Por ejemplo, en el caso de la inflación, los agentes económicos no eran capaces de utilizar *previsiones* en sus cálculos, sino que "se veían sorprendidos" permanentemente por la inflación del año anterior, único dato que utilizaban para calcular la inflación futura[19].

Para cubrir las insuficiencias formales que presentaba el modelo de expectativas adaptativas, a mediados de los años setenta se empieza a utilizar el modelo de las *expectativas racionales*, propuesto por economistas de la *Nueva Economía Clásica*. La hipótesis principal es que la gente aprende de sus propios errores, esto es, no incurre en errores sistemáticos a la hora de realizar sus predicciones; predicciones que puede que sigan siendo erróneas, pero esos errores ya no serán sistemáticos sino aleatorios[20]. Las consecuencias para la política económica que se derivan

[17] También Keynes introduce las expectativas cuando desarrolla formalmente la demanda especulativa de dinero. En efecto, además de los motivos transacción y precaución, la gente puede demandar dinero cuando espere un aumento en el tipo de interés futuro (*tipo de interés normal esperado*), lo que haría caer el precio de los activos financieros emitidos a largo plazo. En su modelo, las expectativas de evolución de los tipos de interés son rígidas a corto y a largo plazo.

[18] Se atribuye a Phillip Cagan (1956), la paternidad de esta hipótesis.

[19] En una situación inflacionista, los agentes esperan que la inflación del próximo año sea igual que la del anterior, por lo que la inflación esperada será siempre inferior a la real. Los errores que cometen los agentes son sistemáticos y, en este caso, crecientes.

[20] Aunque el modelo de expectativas racionales es propuesto inicialmente por Muth (1961), para criticar la hipótesis de las expectativas adaptativas, es Robert E. Lucas (1972) el que la desarrolla. Lucas es considerado el líder de la corriente de la Nueva Macroeconomía Clásica, heredera de la vieja Escuela de Chicago. El supuesto de la racionalidad de las expectativas se interpreta con frecuencia como una pretensión de que las previsiones de los agentes siempre deban ser acertadas, lo cual es una interpretación equivocada. Si las expectativas de los agentes son racionales, seguirán cometiendo errores, pero diferentes en cada ocasión, esto es, serán aleatorios. Una derivada de esta corriente de economistas es la que se desprende del trabajo de Robert Barro (1974), donde

de este modelo son sorprendentes para aquélla época. Dos casos: a) Se pone en tela de juicio, incluso, alguna de las predicciones del modelo de la ecuación cuantitativa del dinero, puesto que un aumento en la cantidad de dinero en circulación provocará un aumento de precios pero no así un aumento en la cantidad producida en el mercado de bienes; b) ante un desequilibrio o una amenaza para la economía, simplemente el efecto anuncio de una medida de política económica por parte del Gobierno hace que los agentes replanteen sus estrategias y se consiga el efecto esperado, aunque la medida anunciada no se ponga en práctica. Consecuencias que resultan del supuesto implícito en el modelo consistente en que los agentes económicos confían plenamente en la palabra de los responsables de la política económica, supuesto que, de ser refutado por la realidad (como todo parece indicar), hace que el propio modelo sea inconsistente, puesto que los agentes ya no aprenderían de sus errores.

Tanto la formalización del modelo de las expectativas adaptativas como el de las racionales, adolecen del mismo sesgo: parten del supuesto de que los agentes forman su opinión sobre el futuro a partir de las medidas que tomará el Ejecutivo de turno, pero descuidan el resto de variables que tienen que ver con el funcionamiento de los mercados y de las instituciones sociales. Dicho de otra forma, se desarrollan en un marco keynesiano de un Estado omnisciente y benevolente, donde el mercado y otras instituciones sociales pierden importancia como informadoras de comportamientos presentes y futuros de interés para los individuos y las empresas. Un sesgo que bien podría tener efectos procíclicos bajo hipótesis propias de la teoría de los ciclos electorales o de la de los grupos de presión. La actual forma de incluir las expectativas en los modelos todavía deja mucho que desear.

2.3.2. El papel del dinero

Uno de los debates más intensos en toda la historia de la literatura económica ha sido el que gira sobre el dinero y su papel en el sistema económico. Como medio de cambio, superó las ineficiencias del trueque. Como unidad de cuenta, su mérito ha servido para fijar la relación de cambio con otros bienes o activos financieros. Como instrumento de conservación del valor, permite los intercambios futuros, para lo que se requiere que el bien utilizado como tal sea duradero y estable. La elección de la unidad monetaria debe cumplir, finalmente, con el requisito de su aceptación general por la sociedad y la facilidad de su uso en los intercambios. Ante la dificultad de encontrar un bien cuyo valor intrínseco reuniera estas características, las sociedades han optado por la creación del denominado dinero fiduciario, esto es, monedas y billetes no respaldados (al menos de forma directa) por metales preciosos (oro o plata) ni por el valor intrínseco de algún otro bien; simplemente por una declaración de respaldo de un banco central basada en la confianza y que se acepta por los agentes económicos.

Sin embargo, la comodidad que ofrece el dinero fiduciario y su importante contribución al desarrollo económico de las sociedades ha generado un desapego respecto de su significado seminal, lo que ha provocado a lo largo de la Historia reciente nuevos problemas, algunos de ellos muy graves, que se resumen en la falta de sintonía de su valor facial con los bienes y servicios que lo deben respaldar[21]. La importancia de controlar el proceso de creación de dinero fiduciario es algo que no ha escapado a los economistas desde los autores del mercantilismo, escuela que centra su objeto en el dinero-mercancía junto al interés por la balanza comercial (metales preciosos). Para los mercantilistas, indudablemente, el dinero es riqueza porque es poder de compra y las monedas no deberían salir del país, a la vez que habría que gravar con sobretasa a las monedas extranjeras, pagar las importaciones con mercancías nacionales y repatriar las ganancias obtenidas en el extranjero. Además, estos autores formulan por primera vez dos de los principales efectos del dinero sobre la actividad económica de un país. Por un

abraza el viejo y controvertido teorema de la equivalencia ricardiana para afirmar la neutralidad de la deuda pública y, por tanto, demostrar que no tiene efectos sobre la producción real, haciendo inútiles las medidas de política fiscal.

[21] La importancia del dinero la resume Galbraith (1975) con un curioso aserto: "la mayoría de las cosas de la vida (los automóviles, las amantes, el cáncer…) solo son importantes para aquéllos que las tienen. En cambio, el dinero es tan importante para los que lo tienen como para los que carecen de él".

lado, Bodin es de los primeros en relacionar las ingentes cantidades de oro y plata que entran en Europa en el S. XVI, con el crecimiento sostenido de los precios[22]. Este crecimiento es una consecuencia del aumento de la oferta de dinero-mercancía en relación a su demanda, puesto que el nivel de precios es el inverso del valor del dinero-mercancía. Bien podría señalarse esta aportación como un primer broche en la polémica sobre la neutralidad del dinero en la economía. En segundo lugar, los mercantilistas observan que el aumento sostenido de dinero-mercancía provoca una disminución de los tipos de interés.

Desde mediados del S. XVII hasta 1776, la teoría monetaria se articulaba en torno a dos corrientes de pensamiento. Una afirmaba que el dinero estimula el comercio (John Law, Jacob Vanderlint y George Berkely), destacando el efecto del dinero sobre la producción y el empleo (no neutralidad) e ignorando en gran medida la posible relación entre dinero y precios. La otra corriente coincide con la teoría cuantitativa del dinero, centrada en la relación dinero-precios (John Locke, Richard Cantillon y David Hume)[23]. Un buen ejemplo del pensamiento monetario de principios del S. XIX es el Bullion Report de 1810, el primer informe oficial contra la política monetaria discrecional[24].

Salvo en el hecho de que la riqueza procede de la agricultura y no del superávit comercial con el resto del mundo, la escuela de la Fisiocracia no hace planteamientos muy diferentes sobre el dinero. En el caso de la economía clásica, aunque ya con una formulación más refinada, la política monetaria es ineficaz para influir sobre las variables reales de la economía y solo afectará al nivel general de precios, lo que equivale a decir que el dinero es neutral[25]. Para Smith, la cantidad de dinero que circula en una economía no forma parte de su renta ni representa riqueza, salvo que se trate de dinero con valor intrínseco; asimismo, lo importante no es la cantidad sino el poder de compra: "el valor de todo el papel moneda nunca puede exceder al valor de todo el oro y la plata que hubieran sido requeridos en ausencia de él" (La riqueza…). Son muchos los pasajes donde Smith manifiesta la necesidad de cautela para que las emisiones del Banco de Inglaterra no sobrepasen el stock de oro y plata, alertando sobre la necesaria cautela sobre los préstamos entre bancos, porque son "de muy corto plazo de forma tal que no se estimule mucho su uso", afirmaciones que hoy gozan de palpitante actualidad.

Por su parte, Ricardo asumió la teoría cuantitativa del dinero y defendió una vuelta a la convertibilidad. Sostuvo que, tanto las disminuciones como los aumentos del nivel de precios, se regulan por los cambios en la cantidad de dinero, aunque nunca se refirió al dinero como un depósito de valor ni a la demanda de dinero. El dinero es un mero regulador del valor que no crea vínculo alguno entre el presente y el futuro. Su visión del crédito fue también excesivamente restrictiva: no consideró los cheques como instrumento de circulación (como hizo Cantillon), sino como medios de economizar el uso del dinero. Dado que no consideraba los cheques como instrumentos monetarios, no podían afectar a los precios, circunstancias todas estas que hicieron de su teoría del dinero algo muy restrictivo.

[22] Probablemente, el primer debate que ha trascendido sobre la relación dinero-inflación se produce entre Bodín y Malestroit, puesto que para este último no puede hablarse de una subida de los precios a largo plazo; en todo caso, bastaría con aplicar la ortodoxia monetaria consistente en mantener el contenido metálico de las monedas (Bodin, 1568). Ahora bien, la primera formulación teórica del valor-escasez, se atribuye a Martín de Azpilcueta (Escuela de Salamanca, 1556), adelantándose en unos doce años a Bodin, para constatar que en los países en que los metales preciosos son escasos, el precio de los bienes es inferior.

[23] Al verse el S. XVIII salpicado de guerras, fueron muchas las presiones para aumentar la oferta monetaria. Apenas se había instaurado el papel de curso forzoso en Inglaterra a finales de ese siglo cuando todos empezaron a ponderar los modos y maneras de volver lo más rápidamente posible al dinero metálico, como enseñara Adam Smith.

[24] A partir de 1808, el aumento de la emisión de billetes empezó preocupar, al punto que Francis Horner, miembro del Parlamento Británico, propuso en la Cámara de los Comunes que se nombrase un comité para investigar el elevado precio de los metales preciosos. En junio, vio la luz un informe redactado por Horner, William Huskisson y Henry Thornton, debatido meses después, pero rechazadas sus conclusiones. Sostenía que la excesiva cantidad de billetes de banco influía en al valor del papel moneda y habría provocado un aumento en el precio de los metales preciosos, a la vez que generaba desconfianza del público en el papel moneda.

[25] "(…) Puesto que el dinero representaba meramente una técnica conveniente de cambio, ya para el cálculo, ya como intermediario, era indiferente para las relaciones productivas esenciales y por tanto no podía afectar el sistema de las proporciones de cambio. Un aumento o disminución de la cantidad de dinero no podía afectar la relación existente entre los precios, puesto que tendía a afectarlos a todos por igual" (*La riqueza de las naciones*, A. Smith).

Por el contrario, John Stuart Mill también aceptaba la teoría cuantitativa, pero le añadió algunas modificaciones para corregir la visión restrictiva de Ricardo. Mill era consciente (como Cantillon y Hume) de que las rígidas conclusiones de la teoría cuantitativa se basaban en el supuesto de una distribución equitativa del nuevo dinero en relación con las tenencias iniciales, por lo que cualquier otra distribución alteraría la estricta proporcionalidad entre dinero y precios. Además, él creía que la teoría cuantitativa estricta solo se sostenía para el dinero metálico, añadiendo que en situación de pleno empleo un aumento del crédito bancario podría disminuir el tipo de interés. No puede olvidarse también la aportación de Henry Thornton al pensamiento monetario clásico, donde establece una relación entre tipo de interés natural y tipo de interés bancario, a la vez que enuncia el postulado del ahorro forzoso[26].

La versión moderna (enfoque macroeconómico) de la teoría cuantitativa del dinero, aparece con Fisher en 1911, superando las formulaciones más imprecisas que dieran, sucesivamente, Azpilcueta, Bodin y Hume: la cantidad de dinero en una economía nacional se puede aproximar por la oferta de dinero que realiza el banco central (considerada como dada). Para una velocidad de circulación del dinero constante y puesto que, en pleno empleo, puede considerarse fijo el número de transacciones, el nivel de precios quedaría determinado por dichas tres variables. Asimismo, en una situación de equilibrio donde se igualan la oferta y la demanda de dinero, la ecuación cuantitativa se puede plantear como una función de demanda de dinero: $M/p = 1/v*Y$ [27]. Aunque Fisher coincide con Knut Wicksell en que la política monetaria se encuentra en la base de los ciclos económicos y de las crisis, ambos mantuvieron serias discrepancias en torno al papel de la política económica, trasladándose la polémica cincuenta años después a los monetaristas y keynesianos.

Por su parte, Cambridge desarrolla otra formulación de esta teoría en las personas de Marshall, primero y de Pigou, después, a partir de un enfoque microeconómico que intenta determinar los factores que explican por qué los individuos mantienen dinero en efectivo (el principal motivo de demanda es el de poder hacer transacciones). Si la velocidad de circulación del dinero depende de las preferencias individuales y el dinero cumple, además, el papel de ser reserva de valor, la demanda dependerá del nivel de riqueza y del tipo de interés; en la medida en que Pigou asume que el tipo de interés es constante y la riqueza se aproxima por la renta nominal, entonces la demanda de dinero individual será proporcional a la renta nominal ($M=k*p*Y$), conclusión similar a la de Fisher, aunque con enfoque distinto.

Lavington (1921) hace un análisis de la demanda de dinero desde el enfoque marginalista, aunque su trabajo lo reformula Fisher en 1930, siendo Hicks en 1935 y Patinkin en 1956, los que desarrollan en detalle el enfoque marginalista de la demanda de dinero. El dinero entra a formar parte de los argumentos de la función de utilidad de los agentes, de forma que el "efecto-riqueza" o de "saldos reales" afecta al gasto agregado de la economía: la propensión a gastar de los agentes económicos dependerá, no solo del nivel de ingreso, sino también del valor real de su riqueza, de la cual las tenencias monetarias son un componente importante. Una variación en el nivel de precios afectará el valor real de las tenencias monetarias y otros componentes de la

[26] El tipo del rendimiento sobre el capital invertido (determinado por la frugalidad y la productividad) regula el tipo de interés bancario sobre los préstamos. Si el tipo bancario está por debajo del primero, la competencia por los préstamos a las empresas elevara el tipo bancario y, al contrario ocurrirá si el tipo bancario está por encima del primero, obligando a los bancos a disminuir los tipos a fin de conseguir préstamos. Por lo tanto, la cuestión de determinar la cantidad óptima de los préstamos bancarios depende de una comparación entre el tipo del rendimiento sobre el capital y el tipo de interés sobre los préstamos bancarios. En cuanto al postulado del ahorro forzoso, para Thornton un aumento de la cantidad de dinero produce un aumento de capital, así como un aumento de precios, con tal de que una parte del nuevo dinero fuera a los empresarios. Si éstos convierten este nuevo dinero en capital, entonces los efectos de producción acompañarían a los precios más altos asociados con el aumento del dinero, de ahí que el dinero no fuera estrictamente neutral, como sostenía Hume. Además, Thornton sugirió la posibilidad de que un aumento de los billetes en condiciones de desempleo general llevaría a un aumento de la producción y del empleo, más que un aumento de los precios. Thorton creía en la neutralidad del dinero solo como una proposición a largo plazo y, aun así, bajo ciertas condiciones (vid *Thorton,* 1802), reimpreso y comentado por Hayek en 1938.

[27] Fisher (1911). En realidad, la ecuación cuantitativa en todas sus versiones (también denominada ecuación de cambio o ecuación de Fisher) no es sino una tautología, puesto que como se aprecia mejor en su versión básica (PT = MV), indica que el valor monetario de las transacciones debe ser idéntico al valor de todos los bienes y servicios vendidos.

riqueza, pudiendo modificar la capacidad de gasto de los agentes económicos y, por lo tanto, el nivel de la demanda agregada[28].

Antes de escribir la Teoría General, Keynes ya había dedicado tiempo a las cuestiones monetarias. En 1923 escribió su Tratado sobre la Reforma Monetaria (duramente criticado por Hayek), abogando porque los países mantuvieran la estabilidad de precios interna y el uso de tipos de cambio flexibles para, posteriormente, en 1930, exponer su teoría sobre el ciclo del crédito, de origen wickselliana. En realidad, para el planteamiento keynesiano, sólo después de alcanzar la economía el pleno empleo, los aumentos en la cantidad de dinero se traducirían en variaciones similares en los precios (Keynes, Teoría General, 263), concediendo en este caso validez a la teoría cuantitativa del dinero. De aquí que se desprenda la posibilidad de un intercambio (*trade-off*) entre desempleo e inflación, pues a medida que se incrementa la actividad económica y el empleo, los precios tienden a aumentar y viceversa. Es, precisamente, esta idea la que desarrolla A. W. Phillips (1958) cuando compara el desempleo con el ritmo de aumento de los salarios, que fueron considerados por él como un indicador aproximado de la inflación[29].

Finalmente, como reacción al keynesianismo surge la conocida como "nueva teoría cuantitativa" de la mano de Milton Friedman, en la década de 1950, quien justifica la demanda de dinero por la utilidad derivada de mantener saldos reales en la medida en que sirven para realizar transacciones. Así, de acuerdo con la teoría del comportamiento del consumidor, la demanda de dinero dependerá del coste de oportunidad de mantener dinero (el tipo de interés de otros activos) y de la riqueza del individuo que se aproxima por la renta permanente ($M/P=g(r,\pi)*Yp$). Para Friedman (1953), $g(r,\pi)$ es una función estable a largo plazo y podría aproximarse a la velocidad de circulación del dinero[30]. El "monetarismo" en el que se inserta la versión de Friedman pretende poner énfasis en el preeminente papel que debe reservarse a la política monetaria y surge como reacción al papel secundario a que la relega el keynesianismo a favor de la política fiscal, donde los precios dependían más de los niveles de empleo y de la tasa de remuneración a los factores: mientras exista desempleo en la economía, los aumentos en la cantidad de dinero -debido a su efecto expansivo sobre la demanda agregada- se traducirían en incrementos en la actividad económica y en los precios.

Mención aparte requiere la escuela austríaca de economía, centrada en la demanda y en la utilidad a partir del enfoque del individualismo metodológico. Enfrentada al historicismo alemán, es conocido su rechazo al método de inducción, afirmando que las ideas colectivas o agregadas no pueden tener un fundamento a menos que se basen sobre los componentes individuales. La economía debe construirse sobre las ideas del interés propio y la maximización de la utilidad, fenómenos que se producen en toda clase de regímenes. El fundamento de la teoría austriaca es que el capital, el dinero y las monedas están sujetos a las leyes de la oferta y la demanda, como cualquier otro bien (L. von Mises). Por lo tanto, su precio refleja una realidad de mercado y transmite información[31]. Los austríacos ven en las monedas respaldadas por activos reales (tierras, metales preciosos, petróleo) la solución a la manipulación monetaria causante de los ciclos y ponen como ejemplo el éxito del patrón oro durante más de 200 años.

L. von Mises refuta la teoría cuantitativa, afirmando con rotundidad la no neutralidad del dinero: *"el cambio de poder adquisitivo no modifica simultáneamente los precios de las*

[28] vid. Lavington (1921), Hicks (1935) y Patinkin (1956).
[29] Para Friedman, a largo plazo la curva de Phillips es vertical y no existe *trade-off* alguno entre inflación y desempleo.
[30] Esta proposición viene desarrollada por Friedman desde 1953; véase también Fiedman (1957) y Friedman (1959).
[31] Llama la atención la enorme actualidad con la que puede describirse la crisis actual a partir de los argumentos de la escuela austríaca creada por Karl Menger, sobre todo, a partir de las aportaciones de L. von Mises, F. von Hayek y M. Rothbard. El precio del dinero transmite las preferencias temporales de los ahorradores, la oferta de capital disponible y la urgencia de los inversores por disponer de él. En la medida en que los bancos centrales y su monopolio de creación monetaria manipulan el precio del dinero (tipos de interés), transmiten a los inversores señales erróneas. Los tipos de interés demasiados bajos estimulan la inversión por encima de lo justificado por el capital disponible. A medida que los inversores y las empresas, cargados de liquidez, pujan al alza los recursos productivos disponibles en la economía, suben los precios nominales hasta igualar la cantidad de dinero. Estas subidas descubren la mentira, revelan la ilusión monetaria y demuestran que dichas inversiones no eran rentables. En este momento, sólo una nueva inyección monetaria puede evitar que explote la burbuja, prolongando la especulación y empeorando la mala asignación de los recursos, agravando las consecuencias de lo inevitable.

distintas mercancías y servicios" (p.94)[32]. Como afirma Rothbard (1974, pág. 19), *"todo incremento de la oferta dineraria tiende, desde luego, a reducir el valor de la unidad monetaria; ahora bien, en cuánto efectivamente llegue a descender, si es que, en definitiva, se reduce el valor del dinero, depende de lo que, al tiempo, esté sucediendo con la demanda, con la utilidad marginal de valuta, con el afán de los agentes por disponer de medios de pago"*. La política monetaria en ningún caso es neutral y sus efectos llegan a la economía productiva antes o después.

En resumen, desde J. S. Mill y salvo para la Nueva Macroeconomía Clásica, se reconoce en mayor o menor medida el impacto del dinero sobre la economía real, con efectos sobre los precios que dependerán del momento del ciclo y de otros factores estructurales. En la medida en que el dinero no sea neutral, la economía sufrirá los efectos nocivos de las inyecciones injustificadas de liquidez o se beneficiará del impacto expansivo de la política monetaria, según los posicionamientos de cada cual conocidos ya desde el Siglo XVII. No queda mucho más por demostrar en este terreno, donde el consenso va siendo cada vez mayor.

2.3.3. Deuda versus impuestos y la ilusión fiscal

Desde A. Wagner, se tiene asumido que el tamaño del gasto público en la economía guarda cierta relación positiva con la renta o la producción de un país. Además, para este autor la causalidad está clara: es la presión del gasto público la que impulsa mayores ingresos en el presupuesto[33]. No obstante, no son pocos los problemas que aparecen a la hora de contrastar empíricamente esta hipótesis, a juzgar por la polémica doctrinal que ha existido. Lo que sí cierto, es que nada se desprende de la obra de Wagner en el sentido de que quepa una prescripción normativa o de que se aviste el riesgo de la aparición de desequilibrios presupuestarios de forma sistemática, sobre todo porque entre lo que denominó principios *lógicos* de la imposición, se recoge el de la suficiencia de los impuestos para atender el gasto[34].

Salvo el caso de la escuela financiera italiana de finales de siglo XIX, desde los mercantilistas hasta bien entrado el siglo XX nadie defendía los efectos neutrales (económicos e intergeneracionales) de la deuda pública. La concepción keynesiana de que la deuda en circulación no supone problema alguno porque, en realidad, "nos la debemos a nosotros mismos, mientras la deuda privada se debe a otros", se encuentra nítidamente expuesta y respaldada por Lerner[35]. Aunque durante los primeros años de keynesianismo, el déficit y la deuda no requerían mayor atención, los autores de la síntesis neoclásica ya se vieron en la obligación (sobre todo, a partir de los años sesenta) de formalizar los efectos esperados de la política fiscal sobre el déficit público. Y La respuesta se hizo depender del punto de la curva *LM* en que se encuentre la economía, asumiendo que la pendiente de la curva *IS* decrecerá conforme la economía se encuentre en una situación más próxima al pleno empleo: una reducción en la tasa impositiva puede dar lugar, por tanto, a un aumento o a una reducción de la recaudación de impuestos[36].

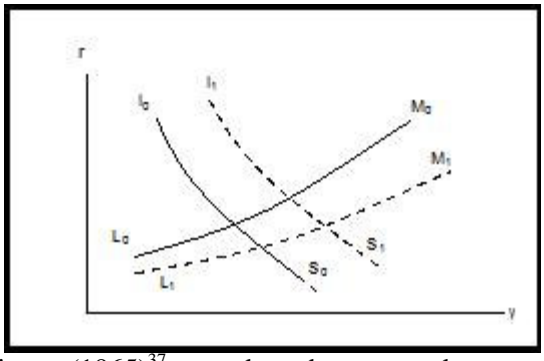
[32] "A la economía política matemática le di el golpe mortal demostrando que la cantidad de dinero y el poder adquisitivo de la unidad monetaria no son inversamente proporcionales"; Mises (1912, p.93).

[33] Véase García Villarejo y Salinas Sánchez (1994), pág. 91-92. El motivo de esta ley *perpetua* y *universal*, en palabras del propio Wagner, se encuentra en " ... la presión del progreso social y los cambios resultantes de las esferas relativas del sector público y privado de la economía, especialmente en la economía pública nacionalizada".

[34] Fuentes Quintana, E. (1990), pág. 78. Como se afirma en Musgrave (1987), Wagner se mostró más cauteloso que Dietzel a la hora de defender este círculo virtuoso de la deuda en la exposición de la hipótesis de crecimiento del gasto público [Wagner (1883), pág. 184].

[35] Para tener una clarificadora visión de esta postura doctrinal, véase Lerner (1948). La creación de deuda no representa una reducción de riqueza nacional, al igual que el pago de los intereses tampoco reduce la renta de las generaciones futuras.

[36] La condición para que dé lugar a un aumento en la recaudación es que la propensión marginal a consumir [c'], más la propensión marginal neta a invertir, incluido el efecto del cambio endógeno de la tasa de interés a largo plazo de LM $[\delta i/\delta y - (\delta i/\delta r)(k'/l')]$, sea mayor que uno. Así, cuando el proceso se da en el momento inicial dentro de un punto de la curva *LM* donde su pendiente (k'/l') es muy reducida, es muy probable que la recaudación impositiva neta resultante sea superior respecto a la del momento previo a la medida fiscal, lo que indica que este tipo de medidas es más efectivo cuando la economía se encuentre poco activa.



Siguiendo a Smith y Tiegen (1965)[37], cuando se hacen eco de un memorándum del Council of Economics Advisers de EEUU, la situación del momento se caracterizaba por elevados tipos de interés y fuerte desempleo de factores productivos, situándose en una zona de la curva LM bastante empinada con una propensión marginal a invertir ($\delta i/\delta y$) muy reducida, por lo que era probable que una rebaja de impuestos aumentara el déficit. El efecto de dicha rebaja sería aumentar el ingreso disponible, por un lado y, por otro, aumentar los activos financieros netos del sector privado. Este proceso acaba por aumentar el ahorro neto (los activos financieros del sector privado), ceteris paribus, en la misma cuantía que se incrementan los pasivos del sector público.

El efecto estimulante de la rebaja impositiva (y del aumento del déficit) supondría un desplazamiento de la curva *IS* desde I_0S_0 hasta I_1S_1 de la figura adjunta, lo que, a su vez, se traducirá en un efecto final sobre la renta, mayor o menor, en función de si la curva *LM* se desplaza desde L_0M_0 hasta L_1M_1, lo que dependerá de la forma en que se financie el déficit. Si la financiación se realiza sin aumentar las reservas del sistema bancario, la oferta monetaria no se desplazará y la curva *LM* permanecerá inmóvil. En caso contrario, la curva *LM* se desplazará hacia la derecha hasta L_1M_1, potenciando el efecto expansionista del déficit.

Sin embargo, también hay un procedimiento de financiación del déficit consistente en vender bonos al banco central, lo que es equivalente a una monetización de la deuda, aumentando las reservas del sistema y desplazando la curva *LM* hacia L_1M_1. Llegados a este punto, no deja de sorprender cómo la mayoría de los economistas keynesianos han optado por preferir aceptar el desafío de la carga de la deuda y convivir con ella, frente a la alternativa de la financiación del déficit mediante creación de dinero[38]. En efecto, en una situación de demanda insuficiente y, en los supuestos del modelo keynesiano simple, el coste de oportunidad de un mayor gasto gubernamental es nulo, puesto que no hay riesgo de inflación. Recurrir a la financiación con deuda en lugar de emitir dinero daría lugar a una situación de *second best*, haciendo innecesario el pago de retribución alguna a los compradores de títulos de deuda. Ante esta inconsistencia del argumento keynesiano, se tuvo que recurrir a la interpretación de que el endeudamiento público era diferente del privado: "...en el contexto macroeconómico, la sociedad no podría nunca ser deudora de manera comparable a la de una persona, familia, compañía o administración local o, incluso, a la de un gobierno central que se hubiera endeudado mediante crédito exterior" (Buchanan y Wagner, 1977 pág. 63). Este argumento

[37] Smith, W L; Tiegen, R L (1965). Hay tres casos en los que la financiación del déficit por parte del Tesoro no aumentará la oferta de dinero: *i)* Vendiendo bonos al público contra sus saldos a la vista en bancos comerciales; *ii)* Vendiendo bonos al público no bancario directamente, sin afectar las reservas bancarias; *iii)* Vendiendo bonos a los bancos comerciales, por lo que éstos tendrán que reducir los préstamos al sector privado. Estos tres procedimientos constriñen los mercados monetario y crediticio pero, mientras que con el primero aumenta el ingreso, el producto y la demanda de dinero, subiendo los tipos de interés a lo largo de L_0M_0, en el segundo caso el exceso de bonos en el mercado reduce sus precios subiendo los tipos de interés en el mercado de bonos y, después, en el mercado monetario, a lo largo de L_0M_0. En el tercer caso, la venta por el sistema bancario de títulos privados eleva los tipos de interés de los bonos y, finalmente, los tipos en el mercado monetario. En estos tres casos, la demanda de inversión se retrae al subir los tipos de interés a lo largo de L_0M_0, contrarrestando en parte el desplazamiento inicial de la curva *IS*.

[38] Buchanan y Wagner (1977;1983), pp. 60 y ss.

implica la inexistencia de transferencia de coste alguno para los contribuyentes futuros, sean los mismos del momento del endeudamiento o no[39].

El influjo de la *"Teoría General"* en la literatura económica en torno al papel de la deuda pública y sus efectos llega hasta los años cincuenta sin apenas contestación, pero entonces se producen las primeras reacciones de calado sobre la interpretación ortodoxa. Así, Buchanan (1958) en sus *Public Principles of Public Debt*, el propio Musgrave (1959), Bowen, Davis y Kopf (1960) y Modigliani (1961), son algunos de los primeros economistas que arremetieron contra la ortodoxia keynesiana en materia de deuda pública. En realidad, si la deuda era inocua para las generaciones futuras, el gobierno podría incurrir en déficits públicos sin límite alguno[40].

El nudo central de la discusión gira en torno a cómo demostrar que la carga de la deuda se traslada a las generaciones futuras. Así, Buchanan entiende que los adquirentes de la deuda lo hacen de forma *voluntaria* contra su consumo futuro a cambio de un interés, por lo que no cabe hablar de sacrificio para la generación actual en la medida en que no soporta carga alguna. Esta interpretación tiene un marcado carácter financiero, desconociendo los problemas que se derivan de la carga real trasladada en forma de cambios en el ahorro y en la inversión por parte de la generación presente. Posteriormente, Buchanan y Wagner (1977) intentan aportar un argumento que rubrican como *síntesis postkeynesiana* de la discusión sobre la equivalencia fiscal, partiendo de la preocupación por conocer el por qué de la elección fiscal de esta vía de financiación[41]. Negando el argumento ricardiano, la sustitución de impuestos por deuda da lugar a una reducción del precio *percibido* de los bienes y servicios públicos por los ciudadanos. Como consecuencia de ello, se produce un *efecto renta*, en términos de Hicks, dando lugar a un cierto incremento en el gasto privado. También se produce un *efecto sustitución* como consecuencia de la alteración de los precios relativos que da lugar a una percepción de abaratamiento de los bienes públicos y al consiguiente aumento de su demanda, exigiendo niveles deseados superiores de presupuesto. Esta inercia es capitalizada políticamente, venciéndose la resistencia ciudadana al establecimiento de mayores impuestos[42].

Un sorprendente planteamiento metodológico en la polémica equivalencia deuda-impuestos es el dado por Barro (1974), donde pretende contestar a tres cuestiones: *i)* si los agentes privados perciben o no la deuda pública como riqueza neta; *ii)* si la deuda y los impuestos son formas equivalentes de financiar el gasto público y, *iii)* si la carga de la deuda se traslada o no a las generaciones futuras. Una vez desarrollada su argumentación basada en la hipótesis de expectativas racionales y en la optimización dinámica con restricciones intertemporales, sólo contesta afirmativamente a la segunda cuestión, lo cual no es extraño si se tiene en cuenta que

[39] Como resumen bien González-Páramo y Contreras (1987), pág. 349, si el uso de recursos públicos significa una reducción en la utilización de los privados —aunque no ocurrirá así en caso de subempleo de factores productivos—, esto es independiente del mecanismo de financiación, puesto que en la financiación con deuda, el pago de intereses por parte de los contribuyentes futuros constituye una transferencia monetaria a los tenedores de los títulos, mientras que la carga en bienes reales la soportará la generación presente.

[40] El argumento keynesiano sobre la no traslación de la carga de la deuda lo expresan muy bien Buchanan y Wagner (1977;1983), pp. 231-232: "Ya que, de acuerdo con esta 'teoría', la emisión de deuda no impone ninguna carga sobre el contribuyente futuro, apenas habría que prestar ninguna atención a la cuestión concerniente a la posible capitalización neta de las futuras cargas fiscales y su probable influencia sobre el comportamiento del gasto en los períodos durante los cuales se emite la deuda. El pago de intereses a los tenedores de títulos de la deuda puede hacerse a partir de los impuestos y estos impuestos pueden, por supuesto, capitalizarse y tratarse como cargas corrientes. Pero, como quiera que tales impuestos se cobran precisamente para pagar el interés a los tenedores nacionales de estos títulos, el pago de estos intereses también puede capitalizarse en valores de activos corrientes, equilibrando las cargas que corresponden al pago del impuesto. En conjunto, sus efectos se compensan y dejan el efecto del impacto inicial del gasto de la operación de financiación por medio de deuda sin que se vea afectado por el cómputo del valor actual". Sin embargo, como reflejan a continuación los autores, este argumento desconoce el impacto de la puesta en circulación de la deuda sobre la disminución de los activos privados.

[41] Buchanan y Wagner (1977;1983), pág. 232 y ss.

[42] En la anterior interpretación, es necesario efectuar algunas precisiones adicionales. En primer lugar, para la Public Choice no es tan importante el efecto renta como el efecto sustitución, despreciado por los keynesianos. Y es que, dentro del primero, podría tener cabida el argumento ricardiano de que el ciudadano contribuyente actual puede anticipar los impuestos futuros que están implícitos en el endeudamiento, si bien, no hay necesidad de llevar el argumento hasta sus últimas consecuencias, rechazando la hipótesis de la racionalidad extrema. Por otro lado, los keynesianos habían mostrado interés por el efecto renta porque el objetivo de la actuación política era incidir sobre la demanda agregada estimulando el gasto privado. En tercer lugar, el argumento de la Public Choice incorpora claramente la hipótesis de la existencia de cierta ilusión fiscal. Dos apoyos empíricos a esta proposición se encuentran en Oates (1975) y en Green (1973).

las tres se encuentran relacionadas. La resolución de estas cuestiones tiene su importancia para plantear, posteriormente, sus propuestas de política presupuestaria: si los bonos gubernamentales no son percibidos por el sector privado como riqueza neta, entonces los efectos esperados de una política fiscal expansionista sobre la demanda agregada se encuentran claramente cuestionados. Cuando el efecto marginal de la deuda sobre la riqueza neta sea próximo a cero, puede no ser cierto el argumento de la no neutralidad en los cambios en el stock de dinero, a la vez que un cambio en el stock de deuda pública puede no tener efectos sobre la formación de capital, mientras que habrá efectos fiscales neutrales sobre la demanda agregada, los tipos de interés y la formación de capital, como consecuencia de los cambios en las cantidades relativas de impuestos y deuda para un nivel determinado de gasto.

Aunque Drazen y, parcialmente, Williamson (1986) y Mundell, entre otros, admiten la equivalencia ricardiana, en general, ha habido más detractores que partidarios de esta idea, entendiéndose mayoritariamente que la deuda facilita el gasto y que no cabe distinguir diferencias en el plano normativo entre el deudor público y el privado. En suma, no parece que la hipótesis de Barro se cumpla, dada la dificultad de preexistencia de los restrictivos supuestos en que se fundamenta. La investigación empírica rechaza la equivalencia ricardiana en su forma pura, aunque algunos estudios han encontrado efectos ricardianos en el comportamiento del ahorro[43]. No obstante, merecen destacarse las consecuencias de la hipótesis de Barro para la política económica y que se traducen en el conocido "modelo de suavización fiscal", enmarcado en el esquema teórico del dictador benevolente para combatir los efectos de las fluctuaciones cíclicas provocadas, no por circunstancias de la demanda (modelo keynesiano), sino por guerras o recesiones profundas. El propio Barro (1986), testó el modelo de suavización fiscal para EEUU y Reino Unido, obteniendo evidencia sobre lo acertado de sus predicciones[44].

En contra de la equivalencia ricardiana y, sobre todo, en contra de las políticas de *mala administración* keynesianas, se posiciona la Public Choice. Como pionero de este posicionamiento, Buchanan desarrolla un enfoque positivo para explicar las malas prácticas públicas, y lo hace rescatando el viejo argumento de la ilusión fiscal que desarrollara A. Puviani (1903): una teoría positiva, por cuanto intenta explicar cómo actúan y se comportan las clases gobernante y gobernada a partir de la consideración sicológica que tiene el contribuyente respecto de las instituciones fiscales. Se trataba de resolver lo que él llamo la paradoja fiscal: ¿cómo puede explicarse que las clases sociales más pobres paguen tributos muchas veces abusivos sin que perciban los beneficios correspondientes al sacrificio que soportan? La respuesta no estaría en la fuerza de la clase dominante, sino que tiene que explicarse por las ventajas que se derivan de la actividad financiera. Puviani se acoge a la teoría hedonista, pero no fundamentada en términos de utilidad marginal, sino en la comparación de costes y beneficios *subjetivos* producidos por la actividad financiera —donde juega un papel importante la *ilusión*— y en la teoría de los *males de la coacción y bienes de la corrupción*[45].

Entre las muchas influencias que ha tenido la teoría de la ilusión financiera de Puviani, hay una especialmente relevante: la teoría de la elección pública. Buchanan ha sido el que más ha contribuido a actualizar los postulados de Puviani. En contra de lo que afirmaba Fasiani —la teoría de la ilusión financiera sólo sería aplicable en un estado *monopolista*—, para Buchanan (1960) la teoría de la ilusión financiera sí que tiene sentido dentro de un enfoque democrático

[43] Véase a este respecto Briotti (2005).

[44] El modelo de suavización fiscal predice que la relación Deuda/PNB aumenta en los períodos bélicos y decrece en los períodos de paz, puesto que el gobierno fija los impuestos de forma que en el primer período exista déficit y en el segundo, superávit.

[45] Hay que tener en cuenta que la ilusión que afecta a la masa ciudadana para inducirles a aceptar una actividad financiera, no supone que ese comportamiento sea, necesariamente, irracional. La irracionalidad conduce a elecciones inconsistentes, sin que quepa hacer predicciones sobre el comportamiento de los individuos, cosa que sí es posible hacer con los individuos afectados por un fenómeno de ilusión. Por otro lado, el poder de la clase dominante a través de la ilusión financiera no es el mismo para con todos los ciudadanos, porque existen algunos —normalmente de renta alta y media— que carecen de ilusión financiera, para los que se aplica la teoría de *los bienes de la corrupción,* consistente en dar favores públicos por parte de las clases dominantes a las clases dominadas —exenciones, bonificaciones fiscales, favores en concesiones, etc.—. Finalmente, el recurso a los *males de la coacción* permite la adhesión de toda la clase dominada [Para una explicación más detallada sobre la teoría de la ilusión financiera de Puviani, véase Fuentes Quintana (1990), pp. 195-203.

del proceso financiero. No obstante, en la medida en que Buchanan centra su teoría en cómo las instituciones fiscales condicionan la determinación y realización de las preferencias de los individuos, afectando a las elecciones sociales, su enfoque es diferente del de Puviani porque, en éste último, es la clase dominante la que impone a la mayoría sus propias elecciones a través del fenómeno de la ilusión. En efecto, para Buchanan no existe motivación para las ilusiones, que se dan tal y como se presentan —como instituciones que pueden condicionar la elección social—. La idea intuitiva en que se sustenta la ilusión fiscal en Buchanan es que los contribuyentes no entienden la limitación presupuestaria intertemporal del gobierno, por eso, insisten estos autores en la crítica al keynesianismo por haber atacado el principio de equilibrio presupuestario de forma sistemática. El keynesianismo es, en sí, propulsor de la ilusión fiscal del contribuyente ocultando las consecuencias de sus políticas en el medio y largo plazo[46].

Como resumen, puede decirse que hasta Keynes, pocos economistas han defendido la neutralidad de la financiación del presupuesto público con deuda o con impuestos, ni siquiera Ricardo. Es con motivo del triunfo de las ideas keynesianas cuando se da cabida a esta equivalencia, más por motivos estratégicos (lo importante son los efectos económicos a corto plazo del gasto público) que dogmáticos, si bien la "síntesis" posterior intenta establecer las condiciones óptimas del recurso a la deuda, aunque con inconsistencias importantes. Al margen de la propuesta de Barro y de unos pocos partidarios, en la década de los sesenta se inicia un posicionamiento mayoritario de los economistas en contra de la neutralidad de la deuda, donde la teoría de la ilusión fiscal ha liderado la oposición a los planteamientos keynesianos en este terreno.

2.3.4. La ecuación de la competitividad internacional

Nunca ha resultado fácil para la literatura económica definir el término competitividad de una nación, aunque sus orígenes se remonten a la época mercantilista con su propuesta de utilizar el saldo favorable del comercio como instrumento de riqueza y progreso de las naciones (no todos los países ganan, juego de suma cero). También los clásicos consideraron el comercio como motor del crecimiento, pero partían de la idea de que el comercio ofrece oportunidades a todos los países para mejorar su riqueza (teoría de la ventaja absoluta de Smith y teoría de la ventaja comparativa de Ricardo, donde no tiene por qué haber perdedores: juego de "suma-suma")[47]. Mientras que la ventaja absoluta presupone ventaja competitiva dentro del mismo sector productivo (de dos países), la ventaja comparativa se fundamenta en la ventaja competitiva entre sectores de un mismo país. El problema de la teoría de Ricardo es que resulta incompleta, al menos por dos razones: primero, porque predice un excesivo grado de especialización, mientras que, en la práctica, los países producen muchos productos, algunos de los cuales compiten incluso con los importados; segundo, porque, al definir el comercio internacional por diferencias en niveles de productividad entre países, no explica el por qué de tales diferencias[48].

Son los trabajos de Heckscher (1919) y Ohlin (1924), los que intentan superar estas dos insuficiencias al considerar que las diferencias en productividad surgen de las distintas dotaciones de factores productivos (teoría de las proporciones de los factores)[49]. Cada país exporta el bien que utiliza intensamente el factor productivo más abundante, por lo que la ventaja comparativa de un país dependerá de los factores en los que esté bien dotado y de menor

[46] Una de las mejores exposiciones del ideario de la Public Choice se encuentra en Buchanan y Wagner (1977).

[47] En el planteamiento de Smith, un país solo puede encontrar ventaja absoluta en la producción de un bien. En cambio, para Ricardo podría existir comercio internacional incluso si un país no fuera superior a otro en la producción de un bien (ventaja comparativa). Un país se especializará en el bien donde tiene la mayor ventaja absoluta y el otro país lo hará donde tiene la menor ventaja absoluta: lo importante son los costes relativos [el desarrollo de esta teoría se encuentra en Ricardo (1817)].

[48] Krugman (2000). En el modelo de Ricardo subyacen tres hipótesis implícitas: los salarios se determinan dentro del mercado laboral de capa país; la economía se encuentra en pleno empleo en los dos países; a medio y largo plazo, la balanza comercial está equilibrada en ambos países

[49] En la versión básica del teorema se utiliza como hipótesis que la estructura de la demanda es idéntica en los dos países que se comparan e independiente del nivel de renta.

coste: la ventaja comparativa depende del coste y éste, de la abundancia del factor. El teorema de Heckscher-Ohlin ha sido ampliado por tres aportes posteriores. Primero, Samuelson (1948) concluye que el libre comercio terminará por igualar los factores de producción entre países, eso sí, en ausencia de costes de transporte y de barreras comerciales, y con una tecnología idéntica, pero donde no se precisa la existencia de inversiones extranjeras que conlleven la transferencia de factores de producción (Teorema de igualdad del precio del factor, Samuelson (1948)). El segundo aporte es el conocido como Stolper-Samuelson (1954), e implica que a medio y largo plazo el comercio internacional impone la necesidad de acometer políticas de redistribución de la renta, vía sistema fiscal[50]. Finalmente, el Teorema de Rybczynski (1955) predice que, con precios constantes, un aumento en la dotación de un factor incrementará en mayor proporción el output de un bien intensivo en ese factor y reducirá el output del otro, lo que implica como corolario que un país puede modificar sus dotaciones relativas de factores alterando sus decisiones de inversión (para Smith y Ricardo, las dotaciones de factores son fijas).

En 1953, la conocida como "paradoja de Leontief" sintetiza la más conocida aportación al teorema de Heckscher-Ohlin, advirtiendo sobre los límites de dicho teorema al demostrar que la teoría de las proporciones de los factores no tiene suficiente capacidad explicativa[51]. A pesar de los empeños de varios economistas y del propio Samuelson en trabajos posteriores, no pudo resolverse tal paradoja y el teorema de Heckscher-Ohlin quedó en entredicho ante su falta de capacidad explicativa de una buena parte del comercio internacional. Trabajos interesantes en esta dirección son los dados por Vernon (1966), bajo la hipótesis del ciclo del producto, según la cual dentro de las cuatro fases por las que atraviesa la fabricación de un producto (introducción, crecimiento madurez y declive), en la primera, la producción de bienes de exportación es intensiva en trabajo, mientras que en las últimas es intensiva en capital[52]. Linder (1961), por su parte, desarrolla la conocida como "teoría de la similitud del país" para explicar el comercio internacional: cada país producirá, de modo principal, para su mercado local y exportará parte de su output a otros países de similares características productivas.

La última aportación a lo que se entiende por teoría económica tradicional del comercio, corresponde a Krugman y Lancaster, de forma separada, con el enfoque basado en las economías de escala[53]: existen dos tipos de comercio, el comercio interindustrial y el comercio intraindustrial. El primero refleja la ventaja comparativa, ya que los países que son relativamente similares y, en consecuencia, tienen algunas diferencias comparativas, podrían involucrarse en el comercio interindustrial. Para el caso de dos países con idénticas dotaciones de factores, el modelo de Heckscher-Ohlin predeciría la no existencia de comercio. Con la presencia de economías de escala, sin embargo, el comercio genera importantes beneficios debido a la especialización de cada país. Podría concluirse que el comercio entre países con diferÿÿÿÿs dotaciones de factores es, en gran medida, interindustrial, pero cuando las dotaciones son similares, el comercio es, en gran parte, intraindustrial (y podría explicar el comercio entre países desarrollados que fabrican bienes similares, p.e. coches).

Sin embargo, probablemente el cambio de acento en el comercio internacional como motor del crecimiento que se produce a partir de los años noventa hacia un énfasis sobre la competitividad internacional de las economías como elemento indispensable para la pura supervivencia, todavía se encuentra en fase de asimilación por las sociedades y aún no se ha incorporado plenamente al recetario de los responsables de la política económica. No por

[50] Puesto que el libre comercio beneficia al factor abundante y perjudica al escaso (por ejemplo, el trabajo), éste tiene que competir a nivel internacional, a pesar de que sea bastante inmóvil entre países, puesto que el trabajo está incluido en los bienes que se intercambian y su precio se ve afectado. Se precisan políticas redistributivas para evitar que un grupo de personas pierda de forma permanente.

[51] Leontief (1956). Siendo EEUU el mayor país en abundancia de capital, descubrió que los bienes que competían con las importaciones requerían un 30% más de capital que el incorporado en los bienes que exportaba; hallazgo que da nombre a la citada paradoja.

[52] La predicción del modelo de Vernon fue bastante aceptable durante las dos primeras décadas posteriores a la IIGM en los Estados Unidos, pero después decayó: el propio autor lo justificó por la gran diversificación geográfica de las empresas capaces de introducir nuevos productos y por la reducción de las diferencias en los mercados nacionales de los países avanzados.

[53] Véase Krugman (1979) y Lancaster (1979).

casualidad, ha sido en fechas recientes cuando los economistas se han referido a la competitividad internacional como la capacidad de producir bienes y servicios para superar los test en los mercados internacionales, a la vez que mejoran las condiciones de vida de los ciudadanos de los países participantes[54]. A mayor abundamiento, hoy se entiende que las naciones no plantean el problema como una competencia entre ellas, sino que consideran que las relaciones económicas con el exterior deben constituir una preocupación interna, por encima de todo (Baldwin, 1995).

La moderna teoría económica del comercio surge a partir del modelo de Porter (1990a), para quien el comercio seguía siendo un factor esencial en el éxito competitivo de un país, pero aceptando la existencia de otros factores muy importantes[55]. Su modelo intenta explicar cuestiones para las que la teoría tradicional carece de respuestas: ¿por qué la innovación es un proceso continuo y no puntual en muchas industrias? ¿Qué nación puede obtener economías de escala y en qué industrias? ¿Por qué ciertas empresas son capaces de obtener ventajas en determinadas industrias y en otras no? Porter observa que una gran parte del comercio mundial tiene lugar entre naciones avanzadas con dotaciones similares de factores, así como un importante volumen de comercio entre sucursales nacionales de empresas multinacionales (algo extraño para la teoría convencional). El "modelo del diamante" de la ventaja nacional de Porter refleja la situación de una economía cada vez más globalizada, con un fuerte grado de concentración empresarial, con un mayor número de multinacionales, una variada demanda, una oferta cada vez más diferenciada y una competencia crecientemente intensificada en torno a territorios. Las ventajas de una nación respecto a otra no se basan en las dotaciones de los factores y en los costes comparativos, sino en las elecciones estratégicas y en la capacidad de las industrias para innovar y mejorar (la prosperidad nacional se crea, no se hereda, [Porter, 1990b]). Este planteamiento supone el abandono de la idea de *nación naturalmente competitiva,* a favor de otra idea consistente en analizar los determinantes de la productividad de los recursos utilizados, donde el gobierno tiene el papel de estimular los vértices del diamante (apoyar la creación de factores especializados, responsabilizarse de la educación primaria y secundaria, la investigación, la sanidad y las infraestructuras básicas).

Aunque, como afirma el propio Porter (1998), él no pensaba que su propuesta sería la dominante en la literatura económica durante tanto tiempo, lo cierto es que lo sigue siendo quince años después. En su modelo, el entorno nacional continúa influyendo en el potencial de la ventaja competitiva de la empresa, donde la disponibilidad de recursos sigue teniendo mucha importancia en la competitividad del mercado local, remarcando la importancia de las ideas de Smith y Ricardo. Por eso, han surgido críticas al hecho de que el modelo no incorpore apropiadamente la actividad de las multinacionales, lo que podría superarse planteando dos diamantes, uno interno y otro para las relaciones internacionales con los principales socios comerciales. Otra crítica es la que argumenta que el modelo no funciona cuando se utiliza para medir la competitividad de los países pequeños o poco desarrollados. A partir de estas críticas, han surgido otros modelos con la intención de completar el de Porter. Este es el caso del de Moon, Rugman y Verbeke (1995) que ofrece el concepto de "doble diamante generalizado", donde se incluyen las empresas del mercado local y de las extranjeras en propiedad[56]. Por otro lado, Cho (1994) redefine el modelo de Porter con nueve factores, añadiendo que un determinante de la competitividad de una nación es la posición competitiva relativa entre países similares en desarrollo económico[57].

[54] En este sentido se han pronunciado Krugman (1990), Zysman y Tyson (1983), Cohen y Zysman (1987), Porter (1987), Tyson (1992) y OECD (1997).

[55] El modelo de Porter analiza la ventaja competitiva nacional identificando cuatro grupos de variables (condiciones de los factores, condiciones de la demanda, sectores afines y auxiliares y estructura y rivalidad de las empresas), más dos factores: el azar y el gobierno.

[56] Definen la competitividad internacional como la capacidad de las empresas para mantener el valor añadido a largo plazo, a pesar de la competitividad internacional.

[57] Cho considera que en la etapa de gran desarrollo económico de una nación, el papel de los empresarios se reduce a favor del mayor papel de los directivos.

Como señala Thomas L. Friedman (2000), "El mundo se ha convertido en un lugar cada vez más entremezclado y hoy, independientemente de ser una empresa o un país, las amenazas y oportunidades se derivan cada vez más de con quién están conectados; un sistema de globalización caracterizado por una sola palabra: la Web. En un sentido más amplio, hemos pasado de un sistema construido alrededor de la división y los muros (eg., *Muro de Berlin*) a un sistema que se construye cada día a través de la integración y las páginas Web"; en definitiva, la globalización se apoya en la integración y, en esta nueva forma de entender las relaciones humanas en torno a la tecnología, la educación adquiere todavía mayor importancia. En palabras de Levinson (2000), la educación se ha convertido en el requisito para acceder a la economía basada en el conocimiento, mientras que la tecnología constituye el instrumento a través del cual se aporta educación a la sociedad. Hoy, al comercio internacional se le considera el motor que impulsa el crecimiento, mientras que la competitividad internacional sería el combustible que alimenta dicho motor.

En la actualidad, todo el bloque de conceptos sobre el comercio internacional y la competitividad que arrancan del trabajo de Porter y que posteriormente han sido moldeados por varios autores, constituye un cuerpo de conocimiento que sitúa en el epicentro el concepto de productividad. A esta transformación también ha contribuido el reciente Nobel de Economía de 2008, Paul Krugman[58], si bien su incidencia en el debate se ha dispersado entre aportaciones para mejorar el enfoque clásico de Heckscher-Ohlin (como ya se ha visto) y otras que se podrían enmarcar ya dentro del nuevo enfoque de la competitividad internacional pero que, a diferencia del modelo de Porter, no han tenido influencia significativa en la predicción del comercio internacional. En el nuevo enfoque de Porter, la competitividad de un país está sujeta a cambios, tanto en el nivel micro como en el nivel macro, pudiendo ser considerada como un fenómeno esencialmente del nivel micro, pero influido por determinados parámetros del nivel macro. Este argumento se encuentra presente en las dos principales aplicaciones empíricas que actualmente se utilizan para evaluar la competitividad internacional y que incluyen a un buen número de países. Se trata de los informes anuales que elaboran, hoy por separado, el World Economic Forum (WEF), con su *Global Competitiveness Report,* y el International Institute for Management and Development (IMD), en su *World Competitiveness Yearbook*. Aunque ambas organizaciones tienen su sede en Suiza, desde 1996 publican sus informes por separado[59]. Mientras que para el WEF la clave del éxito radica en crear las condiciones microeconómicas adecuadas que permitan alcanzar un crecimiento económico rápido y sostenible, para el IMD se encuentra en ofrecer un entorno en el que las empresas puedan competir con éxito[60].

Ambas metodologías utilizan similares factores de competitividad para elaborar sus índices por países. Mientras el IMD utiliza 4 factores y 20 subfactores, el WEF utiliza 11 factores, varios de los cuales coinciden con los subfactores del otro estudio. En el WEF, los fundamentos microeconómicos de desarrollo de las naciones surgen de las estrategias de las empresas y sus prácticas operativas, así como de los *inputs* empresariales, las infraestructuras, las instituciones y las políticas que constituyen el entorno en el que compiten las empresas de una nación. Si no existen mejoras a nivel micro, las políticas macro no tendrán éxito, razón por la cual los modelos macroeconómicos habrían fracasado a la hora de explicar fluctuaciones económicas a corto plazo (Sachs y Warner, 2000). El modelo del WEF se basa en la idea de que el estándar de vida viene determinado por la productividad de la economía de una nación, medida por el valor de los bienes y servicios producidos por cada unidad de factor (trabajo, capital y recursos naturales), por lo que es fundamental crear las condiciones para que tenga lugar un crecimiento rápido y sostenido en términos de productividad[61].

[58] Véase Krugman (1990), Krugman (1991), Krugman (1994) y Krugman (1999).
[59] http://www.weforum.org/ y http://www.imd.ch/, respectivamente.
[60] Estas son las ideas vehiculares que señalan A. Warner y S. Garelli, de forma reiterada, con motivo de la presentación de los sucesivos informes anuales de WEF y de IMD, respectivamente. Desde que en 1996 ambas instituciones decidieron emitir por separado sus informes, el WEF se asoció con Porter (Harvard University) y, en consecuencia, asumió sus postulados.
[61] El resultado son dos índices: uno, denominado de "Competitividad Actual" (identifica los factores que permiten una alta competitividad de una nación medida por el nivel del PIB pc); otro, sobre "Crecimiento de la Competitividad" (evalúa los factores

El informe del IMD parte de un modelo que presupone que cada país gestiona su entorno a partir de los cuatro factores fundamentales de la competitividad y bajo un modus operandi determinado por razones históricas y por valores de la sociedad, no siempre explícitas. Los cuatro ejes sobre los que gira la competitividad son: atracción versus agresividad, proximidad versus globalidad, activos versus procesos y toma de riesgos individual versus cohesión social. A partir de estos cuatro ejes, Garelli (2000) establece combinaciones que permiten explicar los comportamientos de áreas geográficas respecto a la competitividad. Por ejemplo, el modelo de "globalidad y toma de riesgo" es propio del mundo anglosajón, frente al modelo de "proximidad y cohesión social" que caracteriza a la mayoría de los países europeos, donde importan más las consecuencias sociales de la competitividad, circunstancia que le permite proponer un factor determinante de la competitividad: el sistema de valores de un país. Este factor explicaría cómo los países atravesarían cuatro fases sucesivas, desde la perspectiva colectiva hacia la perspectiva individual, en un proceso que no es reversible pero que sí se puede gestionar: "trabajo duro" (muchas horas de trabajo, p.e. Corea y muchos países asiáticos), "riqueza" (preferencia por el incremento de la riqueza personal, p.e. Singapur), participación social" (preferencia por cuestiones sociales, p.e. EEUU u Europa años 60) y "autorrealización" (preferencia por desarrollar el logro personal, p.e. EEUU y Europa hoy). La explotación de resultados que realiza el IMD incluye cuatro índices: el *ranking* de la competitividad global, el *ranking* de los factores determinantes de la competitividad, la estructura de competitividad del país y los perfiles de la competitividad de cada nación.

Como resumen de este apartado, puede constatarse que, desde las aportaciones de Smith y Ricardo, la preocupación de la literatura económica por la competitividad de las naciones en el escenario internacional no ha perdido de vista la productividad como condicionante principal, si bien, la progresiva apertura de las economías nacionales ha trasladado el debate hacia los condicionantes internos de la cada una de ellas, esto es, el problema es visto como algo interno y no como algo sobrevenido desde el entorno internacional. Carece de sentido plantear modelos de crecimiento para un mercado nacional, incluso en sectores no expuestos directamente a la competencia internacional.

3. EL FUTURO ORDEN ECONÓMICO INTERNACIONAL

De la misma manera que entendemos los avances de la ciencia a partir de la refutación de ideas anteriores y de los nuevos descubrimientos, las normas de la convivencia en los regímenes democráticos cambian como consecuencia de la experiencia, de los errores advertidos en el pasado y de la evolución de las variables ambientales. Si bien es cierto que si tomamos como indicador el número de naciones soberanas, la afirmación anterior no superaría el contraste empírico (los regímenes democráticos son claramente minoritarios en este mundo), no es menos cierto que la importancia relativa de las naciones se mide por su poderío económico en el concierto mundial; y en el grupo de las naciones pioneras sí que prima la democracia como sistema de organización social[62].

Hoy bastaría con comparar el mapa de naciones con presencia significativa en el comercio internacional, en relación con el mapa existente en 1939, para advertir cuánto se ha ensanchado el mundo de las naciones económicamente relevantes. Y para advertir, también, que el gran crecimiento experimentado por el comercio internacional ha sido compatible con el aumento de la riqueza de la mayoría de ellas, refutando la teoría mercantilista en este aspecto. Por contra, las

que contribuyen al crecimiento futuro de la competitividad medido por la tasa de variación del PIB pc). Junto a estos índices, el WEF elabora dos índices auxiliares: el Índice de Competitividad Medioambiental y el Índice de Perfil de cada país (ventajas y desventajas competitivas del país respecto a cada factor de competitividad).

[62] Es muy discutible cuál sería el mejor indicador de poderío económico de una nación, donde podrían considerarse variables como el producto nacional, la población total, el tamaño del territorio, la productividad, la disponibilidad de recursos estratégicos a nivel mundial, etc., pero hay que tener en cuenta que algunas de estas variables están correlacionadas. No obstante, cualquier indicador de producto por habitante, ponderado por el tamaño de la población, podría ser el más representativo.

naciones económicamente avanzadas (entonces y ahora) han caído con reincidencia en situaciones propias de la crisis generalizada de 1930, merced a una indisciplina monetaria que, como entonces, ha forzado el crecimiento económico más allá de sus posibilidades y ha impulsado una revalorización de los activos muy por encima del subyacente, lo que ha sido posible gracias al comportamiento asimétrico de los agentes económicos frente al riesgo ante situaciones de alza y de recesión. Siguiendo el patrón de los avances de la ciencia y después de 30 años de vacío regulatorio en el orden económico internacional, parece más que razonable reflexionar sobre una serie de aspectos a partir de los que debería llegarse a un consenso amplio; un nuevo orden económico que dé respuesta a cuestiones centradas en los cuatro ejes del apartado anterior y ponga fin al símil económico del mito de Sísifo que ya nos ha llevado a dos grandes crisis generalizadas:

- *Dinero y política monetaria*. Como ya se ha visto, el temor a que una mala praxis gubernamental basada en un aumento inconsistente de la oferta monetaria ocasione trastornos a la economía, ya quedó perfectamente identificado desde los autores del mercantilismo. La filtración íntegra del exceso de dinero hacia los precios que postulaba Ricardo, era una consecuencia del modelo de pleno empleo permanente en el que se enmarca su visión de la economía, así como propia de una época en la que el capitalismo era poco más que incipiente, de manera que las consecuencias nocivas se circunscribían a una economía nacional con un impacto visible sobre el precio de los bienes de consumo. Sin duda, han sido algunos economistas de la Escuela Austríaca los que primero advirtieron de los serios riesgos que un exceso de dinero tendría sobre el capital y sobre la génesis y amplificación de los ciclos económicos (von Mises). La utilización de los tipos de interés por parte de la política monetaria al objeto de elevar la inversión, supone una distorsión de las señales del mercado que se transmite a los inversores, haciendo ficticiamente rentables algunos proyectos de forma que, una vez trasladado el exceso de liquidez a los precios, la falta de rentabilidad de dichos proyectos solo puede ser suplida por nuevas inyecciones monetarias con el fin de que no estalle la burbuja, agravando las consecuencias de lo inevitable. En la medida en que el aumento de la oferta monetaria no afecte a todos los sectores por igual, como entienden los austríacos y el propio Mill (cuando revisa el planteamiento de Ricardo), el dinero no será neutral en la economía.

  De hecho, la práctica de sujetar las divisas al patrón oro durante una buena parte del S. XX, supone el mejor reconocimiento de esta circunstancia. A sensu contrario, el abandono de la paridad oro del dólar en 1969 ha alentado (cuando no provocado) al menos tres episodios de estallido de la burbuja de precios de los activos en los países avanzados: principios de los noventa, la crisis de las "puntocom" y la actual crisis, cuya virulencia y generalización recuerda a la que dio origen a la Gran Depresión. En una economía tan globalizada como la actual y que cuenta con varias naciones y áreas económicamente relevantes, empieza a constituir un anacronismo la existencia de dos divisas de reserva internacional controladas por sus respectivas autoridades monetarias. El riesgo de que la estrategia utilizada por los EEUU durante décadas consistente en imprimir moneda (real o virtualmente) y exportarla al resto del mundo vía déficit comercial, se extienda al área de los países del euro, no es sino cuestión de tiempo, lo que podría provocar una nueva crisis de proporciones inimaginables. La lógica debería imponer la creación de una divisa global, vinculada a una cesta de monedas relevantes, con las características incorporadas de ser una reserva de valor, con capacidad para denominar en ella las emisiones de deuda pública y determinadas operaciones privadas, convertible automáticamente a las monedas de referencia. Por otro lado y en paralelo, se necesita disponer de un código internacional de conducta en materia de política monetaria que afecte, tanto las prácticas autónomas de las autoridades monetarias, como a las regulatorias dirigidas a las entidades financieras con capacidad para crear dinero.

- *Expectativas y política económica*. Como Se ha visto anteriormente, una de las consecuencias que ha traído la presencia activa del sector público en la economía, derivada, a su vez, del aval doctrinal del keynesianismo durante décadas, ha sido el drástico cambio operado en la formación de las expectativas de los agentes económicos. La creciente presencia del sector

público en la economía después de la IIGM también llevó aparejados cambios profundos en la formación de las expectativas de los agentes económicos privados, pasando a ser dependientes de las medidas de política económica y de las decisiones de multitud de organismos públicos y de empresas públicas, alejadas de las que hubiera tomado el mercado en su ausencia. Ejemplo extremo de lo anterior es uno de los corolarios del modelo de expectativas raciones, al admitir que simplemente el "efecto anuncio" de las autoridades puede dar lugar a la fijación de expectativas. Incomprensiblemente, sin embargo, venimos observando durante décadas cómo las propias autoridades económicas y monetarias rehúyen responsabilidad alguna en las cada vez más agudas crisis que se vienen produciendo en las economías avanzadas. El menoscabo que viene sufriendo la relación esfuerzo-recompensa como fuente del beneficio empresarial y de la utilidad percibida por el consumidor, ha sido suplido por el patrón de conducta que representan las decisiones discrecionales del Estado, hasta el punto de que el principio de seguridad jurídica se ha convertido en un activo sujeto a una elevada prima de riesgo. La inestabilidad de la actuación pública ante los ojos de los agentes económicos aumenta la complejidad de los mercados distorsionando la percepción del riesgo.

Probablemente sea este el mejor momento en los últimos 30 años para reconocer y corregir los excesos cometidos por la mala praxis del Estado en la economía. Es preciso reformular el modelo de intervención del Estado en la economía, en dos direcciones: por un lado, en lo que se refiere a los objetivos y a la forma de acometer las políticas regulatorias de la actividad privada, dotándolas de mayor transparencia y estabilidad; y, por otro, revisando el "trade-off" entre neutralidad y redistribución de la renta, con una mayor consideración de los incentivos. Es preciso volver a un modelo de comportamiento en el que los agentes económicos privados fijen principalmente sus expectativas a partir del binomio esfuerzo-recompensa, y donde el Estado se dedique a fijar sus objetivos de forma transparente y, sobre todo, estable. Se hace necesario un código internacional de conducta de las actuaciones del sector público con trascendencia económica que incluya indicadores de seguimiento.

- *Disciplina presupuestaria e ilusión fiscal*. Hoy está claro que entre las causas de la crisis financiera y económica internacional no se encuentra la indisciplina presupuestaria del sector público, lo que sí se esgrimió con fuerza en la crisis de los setenta y ochenta en muchos países. Ni siquiera podría encontrarse una relación causal entre el excesivo tamaño del sector público y la sobrevaloración de los activos que hoy se da en muchas economías avanzadas, razón por la cual esta crisis se parece mucho más a la de 1929 que a otras posteriores. La teoría de la ilusión fiscal como principal argumento explicativo del tamaño del sector público y de su desequilibrio, junto con las teorías positivas fundamentadas en los ciclos electorales, en la burocracia o en el ciclo de la deuda, no parecen aportar mucho sobre las causas del problema actual.

Sin embargo, están siendo muy controvertidas las medidas de corto plazo de política fiscal que están sometiendo a las cuentas públicas a fuertes desequilibrios con el fin de rescatar al sector financiero de su situación general de insolvencia. Aunque no es objetivo de este trabajo entrar en consideraciones de corto plazo, es preciso señalar que una cifra superior a 3 billones de euros que costaría sólo el rescate del sector financiero mundial con cargo a los contribuyentes, plantea un problema adicional a mayor plazo a los ya existentes: la socialización de las pérdidas privadas en un contexto de deflación de activos puede derivar, bien en una prolongación *sine die* de la crisis, puesto que bajo el argumento de no equivalencia ricardiana, la deuda acabará creando distorsiones importantes (subida de impuestos, déficits públicos continuados, deterioro de expectativas, crowding out, proteccionismo comercial, … que estrecharán el sector privado rebajando las tasas de crecimiento), bien en una situación de hiperinflación, en la medida en que todo este proceso se acompañe de fuertes inyecciones de dinero y de unos tipos de interés forzados a la baja para provocar la reflación de las economías. De hecho, es probable que concurran las dos circunstancias (estanflación) y que se ponga de

manifiesto, de nuevo, la dificultad de encontrar un ajuste voluntario del desequilibrio existente[63].

- *El nuevo escenario de la competitividad internacional*. La frontera física entre países ha formado parte hasta fechas muy recientes de los supuestos subyacentes a los modelos económicos, tanto en el plano normativo como en el positivo. Las implicaciones de este proceder han hecho que se supongan como dados una serie de condicionantes de la situación interna de cada país, de forma que segmentando el mercado entre bienes comerciables y no comerciables, la política económica podía centrarse en los factores de coste que afectan a los primeros, en cuya solución jugaban un importante papel las barreras arancelarias y administrativas, los contingentes, los argumentos de seguridad nacional, las estrategias de dumping fiscal y, en último extremo, las devaluaciones competitivas. Los modelos de comercio exterior, aún con dificultades, eran capaces de explicar una buena parte del comportamiento internacional de las economías, hasta que se produjo una aceleración importante en su grado de apertura desde finales de los ochenta que impulsó un doble cambio: el corrimiento desde la competencia en precio hacia la competencia en innovación por diferenciación de productos; y la deslocalización geográfica de la producción buscando la ventaja competitiva en costes. Las consecuencias de esta nueva estrategia no se han hecho esperar: la aparición de nuevos países muy competitivos con fuerte superávit comercial; el agravamiento del déficit comercial de aquéllos países donde el binomio costes-innovación no ha podido mantener la posición competitiva; las devaluaciones explícitas o consentidas de las divisas de los países perdedores; y la reasignación de recursos hacia la educación, la investigación y la innovación.

Los efectos del cambio en el ranking de competitividad de los países son traumáticos para muchos de ellos. Los países que han buscado salirse de la restricción de la competitividad forzando el crecimiento a partir de la producción de bienes no comerciables (EEUU, España, Reino Unido, Irlanda, en el caso de la construcción residencial), con la actual crisis han constatado que este sector no es totalmente impermeable a las reglas de la competitividad internacional, puesto que una parte de los recursos necesarios son obtenidos del resto del mundo, compitiendo con el sector de bienes comerciables y con el propio sector público. En aquéllos países donde el déficit comercial ha podido ser compensado por la balanza favorable de transferencias procedente de los retornos de inversiones exteriores (actividad multinacional), el saldo neto del binomio costes-innovación ha podido mantener el indicador de competitividad, aunque ahora estén sufriendo los efectos de la burbuja de los activos. Por el contrario, en aquéllos países donde su fuerte déficit comercial esté financiado en gran medida por operaciones financieras exteriores con escasa vinculación productiva, al problema de sobrevaloración de activos se añadirá al de una drástica reducción del ritmo de crecimiento económico al acusar más las restricciones de liquidez. Esto explica la diferente posición de España frente a EEUU en estos momentos descontando, incluso, la ventaja que representa el dólar como moneda de reserva internacional y la posibilidad de recurrir a la devaluación. Si a estas circunstancias sumamos la diferente contribución de los factores internos (excluidos los costes) a la competitividad internacional para los distintos países, puede entenderse la diferente posición que ambos países ocupan en el ranking de competitividad que proponen tanto el WEF como el IMD durante los últimos años. En definitiva, la moderna teoría del comercio internacional nos dice que países con elevada renta relativa (desventaja en costes) sólo pueden mantener o mejorar su competitividad a largo plazo con esfuerzos adicionales en innovación, con el fin de mantener la productividad y los retornos internacionales de su innovación. Puede decirse que, en un escenario de globalización económica, las prácticas que buscan el crecimiento económico al margen de la competitividad internacional, a largo plazo están

[63] Para Estados Unidos, la burbuja de activos que se creó en la década de 1920 y que derivó en el Crash de 1929, tuvo como respuesta el proteccionismo comercial, el fuerte y sostenido desequilibrio en las cuentas públicas y políticas monetarias inflacionistas, en un proceso de ajuste del que formó parte la IIGM como solución más efectiva que el propio *New Deal* a la hora de deflacionar los activos.

abocadas al fracaso. Por la misma razón, una vuelta al proteccionismo generalizado a nivel mundial es poco probable en época de paz.

Como siempre es bueno recurrir a la experiencia para forjar un mejor análisis de la actualidad, hoy tiene más vigencia que nunca la acertada visión de Galbraith sobre la Gran Depresión que hiciera en 1954, aunque la centrara más en las causas del Crash de 1929 que en las ineficaces medidas adoptadas por las autoridades norteamericanas hasta el estallido de la IIGM. Aunque no es objeto de interés de este trabajo abundar en las medidas de salvamento del sistema financiero adoptadas por los gobiernos de las principales economías del mundo durante las últimos meses, incluido el reciente plan del Presidente Obama para utilizar hasta un billón de dólares (conjuntamente con el sector privado) en la adquisición de activos depreciados en poder de los bancos, resultan preocupantes los efectos a largo plazo que el enorme desequilibrio de las cuentas públicas y la superabundancia monetaria tendrán sobre el ahorro, los incentivos, la inflación, las expectativas y, en definitiva, sobre el ritmo de crecimiento de las economías en esos países. Aunque en las políticas de prevención de salud es costumbre vacunar a personas sanas (inocular pequeñas dosis de agente maligno para crear anticuerpos), a ningún médico se le ocurriría curar al enfermo de sida inyectándole más VIH.

BIBLIOGRAFÍA

ARROW, K. J. (1951): "Social Choice and Individual Values"; tésis doctoral, Columbia Univ.

AZPILCUETA, M. (1556): "Tratado de las rentas de los beneficios eclesiásticos". Véase análisis de su obra en Muñoz de Juana, R. (1998): "Moral y economía en la obra de Martín de Azpilcueta"; Eunsa: Ediciones Universidad de Navarra, S.A. Barañáin.

BALDWIN, R.E. (1995): "The problem with competitiveness"; Proceedings of EFTA´s, 35th Anniversary Workshop; EFTA, Geneva.

BARRO, R. (1974): "Are Government Bonds Net Wealth?"; Journal of Political Economy, vol. 82, nº 6, pp. 1095-1117.

BARRO, R (1986): "U.S. Deficits since World War I"; Scandinavian Journal of Economics, vol. 88, nº 1, pp. 193-222.

BODIN, J. (1568): "Paradoxes de M. de Malestroit touchant le fait des monnaies et l'enrichissement de toutes choses".

BRIOTTI, M. G. (2005): "Economic Reactions to Public Finance Consolidation: a Survey of the Literature", European Central Bank, Occasional Paper, No. 38, Oct. 2005.

BUCHANAN, J; WAGNER, R (1977): "Democracy in deficit. The political legacy of Lord Keynes", Academic Press, New York (v.c. Déficit del sector público y democracia, RIALP Ed., Madrid, 1983).

BUSTELO, P. (1999): **"**Teorías Contemporáneas del Desarrollo Económico", Ed. Síntesis, Madrid.

CAGAN, PH. (1956): "The Monetary Dynamics of Hyperinflation", in Friedman, Milton (ed.), Studies in the Quantity Theory of Money, 1956, Chicago: University of Chicago Press, ISBN 0226264068.

CHO, D.S. (1994): "A dynamic approach to international competitiveness: The case of Korea"; Journal of Far Eastern Business; 1(1): 17-36.

COHEN, S. Y ZYSMAN, J. (1987): "Manufacturing matters"; New York: Basic Books.

EICHNER, A. S. (1983): "Why Economics Is Not Yet a Science"; Armonk, New York: M. E. Sharpe.

FERNÁNDEZ DÍAZ, A. ET AL. (1995): "Curso de Política Económica". MacGraw-Hill. Madrid.

FERNÁNDEZ DÍAZ, A. Y RODRÍGUEZ SÁIZ, L. (1982): "Introducción y Metodología de la Política Económica"; Ed. ICE. Madrid.

FISHER, I. (1911): "The Purchasing Power of Money"; New York; Macmillan.

FRIEDMAN, M. (1953): "Essays in Positive Economics", University of Chicago Press, reprinted 1970, pp. 3-43; [v.c. (1967): "Ensayos sobre economía positiva"; Ed. Gredos].

FRIEDMAN, M. (1957): "A Theory of the Consumption Function"; [v.c. (1985): "Una teoría de la función de consumo"; Alianza Editorial].

FRIEDMAN, M. (1959): "A Program for Monetary Stability"; [v.c. (1962): "Un Programa de estabilidad monetaria y reforma bancaria"; Ed. Deusto (Barcelona)].

FRIEDMAN, T. L. (2000): "The lexus and the olive tree"; First Anchor Books Edition.

FUENTES QUINTANA, E. (1990): "Hacienda Pública", Imp. Rufino García Blanco. Madrid.

GALBRAITH, J. K. (1954;2008): The Great Crash 1929; Penguin Books (reprint). V.e.: El Crash de 1929; Ariel, 2008

GALBRAITH, J. K. (1975;1983): El dinero; Ed. Orbis, Barcelona.

GALINDO, MIGUEL A. (2003): "Keynes y el Nacimiento de la Macroeconomía", Ed. Síntesis, Madrid, 238 pp.
GARCÍA VILLAREJO, A Y SALINAS SÁNCHEZ, J. (1994): "Hacienda Pública, general y de España"; Ed. Tecnos; Madrid.
GARELLI, S. (2000): "Competititiveness of Nations: The Fundamentals"; The World Competitiveness Yearbook. IMD, Switzerland.
GONZÁLEZ-PÁRAMO, J. Y CONTRERAS, C. (1987): "Traslación intergeneracional de la carga de la deuda pública", Papeles de Economía española, nº 33, pp. 347-361.
GREEN, K. V. (1973): "An Empirical Test of the Wagner Debt Illusion Hypothesis", en International Institute of Public finance: Issues in Urban Public Finance, Saarbrucken, pp. 208-225.
GREENSPAN, A. (1966): "Gold and Economic Freedom"; The Objectivist published; reprinted in "Capitalism: The Unknown Ideal", Rand Published (1957).
HARROD, R. F. (1963): "The Life of the John Maynard Keynes"; McMillan, London.
HECKSCHER, E.F. 1949 (1919): "The Effects of Foreign Trade on the Distribution of Income". En Howard. S. Ellis & Lloyd A. Metzler, eds. Readings in the theory of international trade. Philadelphia: The Blakiston Co. pp. 272-300. La traducción completa se encuentra en Harry Flam and M. June Flanders, eds., Heckscher-Ohlin Trade Theory, Cambridge, MA: MIT Press, 1991, pp. 39-69).
HICKS, J. R. (1935): "A Suggestion for Simplifying the Theory of Money"; Economica.
KEYNES, J. M. (1936): "The General Theory of Employment, Interest and Money"; Macmillan Cambridge University Press, for Royal Economic Society.
KRUGMAN, P. (1979): "Increasing Returns, Monopolistic Competition and international Trade"; Journal of International Economics, 9: 469-479.
KRUGMAN, P. (1990): "Rethinking International Trade". MIT Presss, Cambridge, MA.
KRUGMAN, P. (1991): "Geography and Trade"; Leuven University Press & The MIT Press.
KRUGMAN, P. (1994): "Competitiveness. A dangerous obsession"; Foreign Affairs, March/April:1-17.
KRUGMAN, P. (1999): "Development, Geography, and Economic Theory"; The MIT Press, Cambridge, MA.
KRUGMAN, P. (2000): "Ricardo´s Difficult Idea"; http://web.mit.edu/krugman/www/ricardo.htm
LANCASTER, K. J. (1979): "Variety, Equity and Efficiency"; New York: Columbia University Press.
LAVINGTON, F. (1921): "The English Capital Market"; Methuen & co. Ltd., London.
LEONTIEF, W. (1956): "Factor Proportions and the Structure of American Trade: Further Theoretical and Empirical Analysis"; Review of Economics and Statistics, 38:386-407; reprinted in: W. Leontief, 1966, Input-Output Economics, Chap. 6.
LERNER, A. (1948): "Burden of the National Debt", in "Income, Employment and Public Policy", Essays in honour of Alvin Hansen, W.W. Norton, New York.
LEVINSON, M. (2000): "Education for Technology or Technology for Education: The Dilemma of the New Economy"; The Global Competitiveness Report 2000. World Economic Forum.
LINDER, S.B. (1961): "An Essay on Trade and Transformation"; New York. Wiley.
LUCAS, R. E. (1972): "Expectations and the Neutrality of Money"; JPE.
MCKINSEY GLOBAL INSTITUTE (2008): "Mapping Global Capital Markets: Fifth Annual Report"
MISES, L. V. (1912): "The Theory of Money and Credit"; Foundation for Economic Education, New York, 1971; translation from 2ª Germany Ed., 1924.
MOON, H.C., RUGMAN, A.M. Y VERBEKE, A. (1995): "The generalized double diamond approach to international competitiveness"; En Alan M. Rugman, editor, Research in Global Strategic Management: A Research Annual, 5: 97-114.
MUSGRAVE, R. (1987): "La carga de la deuda: una breve historia", Papeles de Economía Española, nº 33, pp. 362-364.
MUTH, J. F. (1961): "Rational Expectations and the Theory of Price Movements"; Econometrica, 1961, 29.3:315-335.
OATES, W.E. (1975): "Automatic Increases in Tax Revenues. The Effect on the Size of Public Budget", en W E Oates (Ed.): Financing the New Federalism, Baltimore: John Hopkins Press; pp. 139-160.
OECD (1997): "Industrial Comeptitiveness". Paris.
OHLIN, B. (1924;1933): Interregional and International Trade, Harvard University Press, Cambridge.
PATINKIN, D. (1956): "Money, Interest, and Prices: an Integration of Monetary and Value Theory"; University of Chicago Press.
PHILLIPS, A. W. (1958); "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1961-1957"; Economica, November pp. 283 – 299.
PORTER, M. (1987): "From Competitive Advantage to Corporate Strategy"; Harvard, Business Review, May-June: 43-59.
PORTER, M. (1990a): "The Competitive Advantage of Nations"; The Free Press.

PORTER, M. (1990b): "¿Dónde Radica la Ventaja Competitiva de las Naciones?"; Harvard Deusto Business Review. Especial 100; Lo Mejor en Gestión de HDBR. IV Trimestre.

PORTER, M. (1998): "International Competitiveness of Spanish Companies: Microeconomic Foundations and the Role of Cluster"; Conferencia en Madrid organizada por el Institute for International Research.

PUVIANI, A. (1903): "Teoria della Ilusione Finanziaria"; Palermo, Laterza (v.c. "La teoría de la ilusión financiera", A. Rodriguez Bereijo, IEF, Madrid, 1972).

RICARDO, D. (1817): "On the Principles of Political Economy and Taxation"; [v.c. "Princios de la economía política y la tributación" (1973); Ed. Ayuso; Madrid].

RICARDO, D. (1821): "On the Principles of Political Economy and Taxation". Selección publicada por editorial Orbis, Barcelona, 1985, con el título Principios de economía política y tributación (selección), vid., pp. 79 y ss.

ROTHBARD, M. N. (1974;1985): "Lo esencial de Ludwig von Mises"; Unión Editorial, Madrid.

RYBCZYNSKI, T. (1955): "Factor Endowments and Relative Commodity Prices", Econometrica NS, vol. XXII (1955), 336-41.

SACHS, J. Y WARNER, A. (2000): "Globalization and International Competitiveness: Some Broad Lessons of the Past decade"; The Global Competitiveness Report 2000; World Economic Forum; Oxford University Press: 18-27.

SAMUELSON, P. (1948): "International Trade and the Equalization of Factor Prices", Economic Journal 58, 163-184; reprinted in: P.A. Samuelson, 1966b, Chap. 67.

SCHUARTZ, P. (2001): "El comercio internacional en la historia del pensamiento económico", IUDEM, Documento de Trabajo 2001-3.

SMITH, W L; TIEGEN, R L (Ed.) (1965): "Readings in Money, National Income and Stabilization Policy", Homewood, I11, R. D. Irwin (in Council of Economics Advisers: Financing a Federal Deficit, and Fiscal Policy in Perspective.

STOLPER, W. F. Y SAMUELSON, P. (1941): "Protection and Real Wages"; Review of Economics Studies 9: 50-73. Reprinted in: J. Bhagwati (ed.), 1969: 245-268; and in: P.A. Samuelson, 1972: 831-846.

THORTON, H. (1802): "Enquiry into the Nature and Effects of the Paper Credit of Great Britain".

TYSON, L. (1992): "Who´s Bashing Whom: Trade Industries"; Washington: Institute for International Economics.

VERNON, R. (1966): "International Investment and International Trade in the Product Cycle"; Quarterly Journal of Economics, 80: 190-207.

ZYSMAN J. Y TYSON, L. (1983): "American Industry in International Competition, Government Policies and Corporate Strategies"; Ithaca, N.Y., London: Cornell, University Press.

FOREIGN AID TO PAKISTAN: IMPACT ON GDP WITH LEAST SQUARE ESTIMATION

ASMA SALMAN
E-mail: asmaiqbal@yahoo.com
HUI XIAO FENA
E-mail: xfhui@hit.edu.cn
School of Management
Harbin Institute of Technology

Abstract
Foreign direct investment flows have increased dramatically in recent decades. Developing countries have long tried to attract a reasonable amount of FDI, and easing foreign exchanges regimes, and removing restrictions have become increasingly popular.
Foreign direct investment (FDI) is a potent weapon of economic development, especially in the current global context. It enables a capital-poor country like Pakistan to build up physical capital, create employment opportunities, develop productive capacity, enhance skills of local labor through transfer of technology and managerial know-how, and help integrate the domestic economy with the global economy There are many forms of foreign resources like Foreign Direct Investment, Grants & Loans, Foreign Portfolio Investment, Export Credit, Official Aid, Technical Assistance, Emergency Relief, Project and non project aid. This paper reviews impact of foreign aid on development. By performing Least Squares we analyses the trend and impact of foreign aid on the economic development of Pakistan during a three decade period. From the analysis it is clear that the foreign aid and inflow has a positive impact on the GDP.

Keywords: GDP, Foreign Aid, Least Squares, FDI.
Thematic Area: Growth Economics.

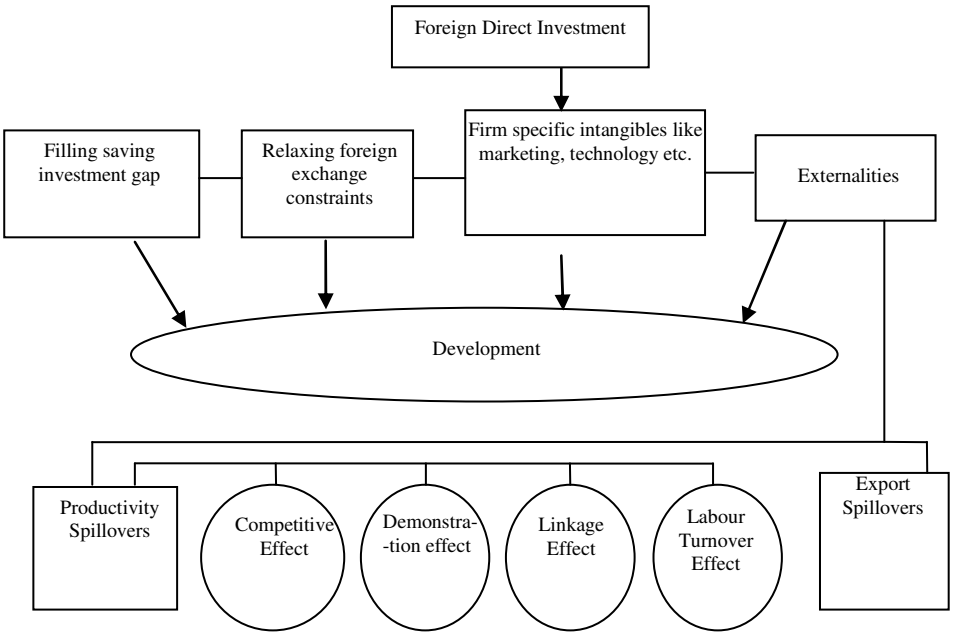
1. INTRODUCTION

The role of foreign direct investment (FDI) has been widely recognized as a growth-enhancing factor in developing countries. There are a variety of channels through which FDI can promote economic growth, in the host country. The most important being is technology transfer and spillovers. Literature on economic growth has established the importance of technological progress in economic development. FDI often leads to technology transfer to affiliates of multinational firms in the host countries. Spillover can occur through the interaction of multinational firms with domestic suppliers, customers and worker mobility. Therefore, FDI can have impact on income [GAO (2004)]. Most of the developing countries rely primarily on FDI as a source of external finance because FDI stimulates economic growth more than other sources of capital inflows. Particularly, FDI is supposed to be less volatile to offer financial resources, transfer of modern technology, market access and managerial know-how.

Financial resources are largely used to expand productive capacity by increasing fixed investment in the host countries, while transfer of technology and managerial know-how improves productive capacity. Furthermore, FDI brings various networks such as sales and procurement networks to the host countries, which can be used to expand their business opportunities. FDI also increases competitive pressures to the local firms that result in an improvement in technical and allocative efficiency in the host country.

It can be argued that economic growth depends on technological progress and FDI can play a crucial role because it facilitates technology diffusion. Zhang (2001) has noted that FDI is likely to be an engine of host country's economic growth, because (i) inward FDI may enhance capital formation and employment generation, (ii) FDI may promote manufacturing exports, (iii) FDI may bring resources into host country such as, management know-how, access of skilled labour to international production networks, and established brand names, and (iv) FDI may result in technology transfers and spillover effects. In the light of important contributions that FDI delivers to both home and host countries, it is useful to analyse its impacts on growth of the domestic economy. Figure-1 has identified the following linkages between FDI and development.

Fig.1 Linkages between FDI and Development; further details at Pradhan (2003)



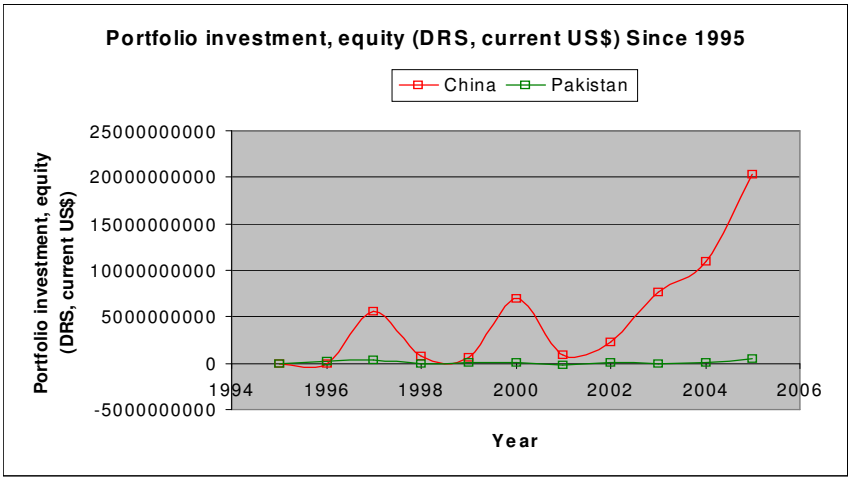
Pakistan has experienced GDP growth in the 6-8% range in 2004-07, spurred by gains in the industrial and service sectors. Poverty levels have decreased by 10% since 2001, and Islamabad has steadily raised development spending in recent years, including a 52% real increase in the budget allocation for development in financial year 2007. Since independence Pakistan has had to depend on foreign assistance in its development efforts and to balance its international debt payments. In 1960 the World Bank organized the Aid-to-Pakistan Consortium to facilitate coordination among the major providers of international assistance. The consortium held 92% of Pakistan's outstanding disbursed debt at the end of June 1991. The consortium's members include the United States, Canada, Japan, Britain, Germany, France, and international organizations such as the World Bank and the Asian Development Bank (ADB). The World Bank accounted for 26% of the outstanding debt, and the ADB, which was the largest lender in the early 1990s, accounted for 15%. Most non-consortium funding comes from Saudi Arabia and other oil-producing Middle Eastern countries. Most aid is in the form of loans, although the proportion of grants increased from around 12% in the late 1970s to around 25% in the 1980s, mainly because of food aid and other funds directed toward Afghan refugees. With the decline in this aid after 1988, the proportion of grants decreased to 16% in 1992. The United States has been a major provider of aid since independence and was the largest donor in the 1980s. The

unsustainable economic growth has been blamed mainly on the high inflation rate, a mounting fiscal deficit, increasing foreign debt and debt servicing, weak foreign demand for Pakistani products, low level of physical and human capital, unfavorable weather, political instability, and, among other factors, a deteriorating law and order situation in the country. Pakistan is a rapidly developing country and a major emerging market with an economic growth rate of 7 % per annum for four consecutive years up to 2007. Despite being a very poor country in 1947, Pakistan's economic growth rate was better than the global average during the subsequent four decades, but imprudent policies led to a slowdown in the late 1990s. The 2005 estimate of foreign debt was close to US$40 billion. Pakistan's gross domestic product, as measured by purchasing power parity (PPP), is estimated to be US$ 475.4 billion while its per capita income (PCI) stands at $2,942. The poverty rate in Pakistan is estimated to be between 23% and 28%. Pakistan's GDP growth rates have seen a steady increase over the last 5 years. However, inflationary pressures and a low savings rate, among other economic factors, could make it difficult to sustain a high growth rate.

1.2 Overview of the Financial and Trade Policies in Pakistan

Economic growth of developing countries is heavily based on the financial sector's credit allocation. Overall financial development is necessary for economic growth at the macro-level [Andersen and Tarp (2003); Khan and Senhadji (2000); Levine (2002)]. A more advanced intermediation enables firms to raise and manage large amount of funds more effectively, resulting in a rapid economic development. Particularly, the development of financial sector is important for developing countries because bank-based system has greater impact on growth at the early stage than does a market-oriented system [Fase and Abma (2003); Tadesse (2002); Iimi (2004)]. The graph gives us a clear idea about the portfolio investments in china and Pakistan, our interest however is Pakistan. Data is available from 1995.

Graph 1: Portfolio Investment



Source: World Bank

2. LITERATURE REVIEW

Pakistan has been relying on the foreign aid to support its development programs since independence. And the aid still has a larger proportion in the foreign capital inflows to Pakistan. The role of foreign economic assistance in economic development and growth remains continuous in economic literature. Some studies proved its positive impact on the economic development, while some studies highlighted its negative effects. Chenery and Strout [3]

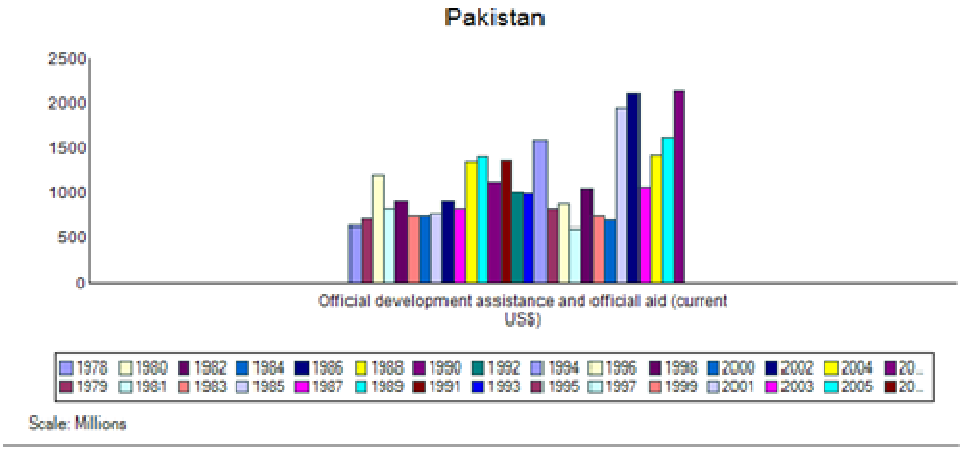
concluded, on the basis of empirical evidence from LDCs, that foreign capital has a positive effect on the economic growth. Burnside and Dollar [2] in their study found that aid has a positive impact on economic growth in developing countries with a good fiscal policy, while it has negative effect on economic growth in the presence of poor policy. Leff analyzed its negative impact on growth. They argued that foreign aid could adversely affect the economic growth by substituting the domestic saving. Mwafaq Dandan Al-Khaldi [7] concluded that the foreign aid has positive effect in the economic development. Robert and Nicholas found that the relationship among foreign aid, government policy and economic growth is tenuous and depends heavily on the particular set of countries included in their study. In case of Pakistan, the foreign aid is a major form of the foreign capital inflow and has a significant role for the country's economic development. In Pakistan, several economists have tried to find out the role of foreign aid in economic development of Pakistan. As, Shabbir & Mahmood (1992) and Khan & Rahim [9] (1993) concluded that the aid has accelerated the rate of growth of GDP. Aslam [1] (1987) examined that the public FCI did not affect the domestic investment significantly, while the private FCI covered the domestic saving-investment gap. Khan (1993) concluded that foreign aid has played an extremely important role in influencing the pace of development, especially investments and imports have to a large extent depended upon the amount of foreign aid. However, this dependence on foreign aid, on the other hand, has led to the emergence of rising debt burden.

In short, we can conclude on the review of the above literature foreign capital has stimulated the economic growth on one hand and has substituted the domestic savings on the other hand. And it caused a severe debt serving problems in Pakistan. It is open to question how far the LDCs could achieve self-sustaining economic growth with foreign aid. Could foreign aid reduce poverty, inequality and unemployment? It helps traditionalism to be upgraded and they think it helps modernization and adaptation. But critics say, concessional resources from aid donors delay self-reliance of the recipient countries, thwart domestic saving, prop up authoritarian and repressive regimes, and distort domestic development. They also think foreign fosters dependence syndrome, which proves difficult to break. Such criticism led to evaluation of the impact of foreign aid. It raised questions whether or not aid works. The Development Advisory Committee (DAC) of the Organization for Economic Cooperation & Development (OECD) states that aid has helped build key development-promoting institutions like agricultural universities as well as technical and enterprise management training institutes in many countries. A number of econometric studies estimated that about 23% of foreign-capital inflows offset the declines in domestic savings.

3. TRENDS OF FOREIGN AID IN PAKISTAN

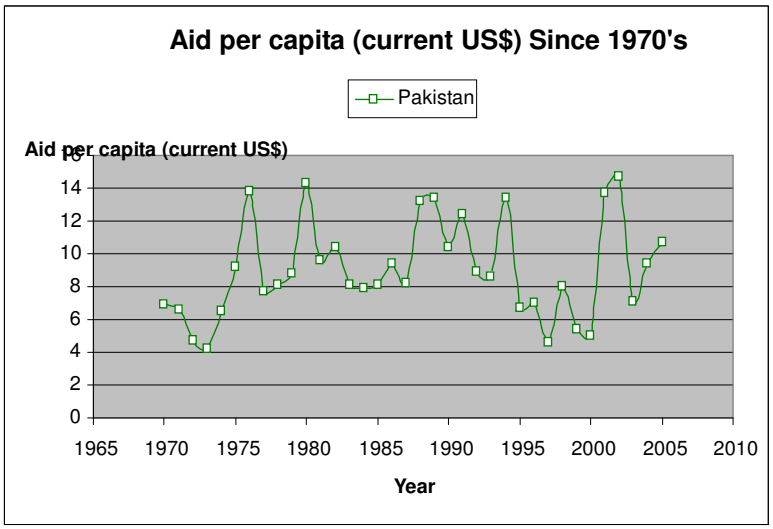
The trends and the patterns have shown that the FDI, portfolio investment and borrowing through private sources have also increased sharply. But Pakistan is still unable to attract such FCIs. Pakistan lacks physical, financial & human capital as well as political & macroeconomic stability. So, it has to rely on foreign aid or foreign debt. The role of foreign aid varies from one country to another. Like other developing countries, FCI in Pakistan is being widely considered as an important vehicle for economic growth. Pakistan has introduced a wide range of incentives, congenial for local and foreign investors and has increasingly tended to turn to FCI as source of capital, technology, managerial skills and market access needed for sustained economic development.

Table 1. Pakistan: official development assistance (FA) and official aid (Current US$ Millions)



Due to low income the taxable capacity remains lower, i.e. government earning also remain low. In such situations the underdeveloped countries have to face saving investment deficit as well as the deficit in the balance of payment. The two- gap model suggests that developing countries have to rely on the foreign capital inflow to fill these two gaps: The import-export gap and the saving investment gap. Foreign aid can involve a transfer of financial resources or commodities (e.g., food or military equipment) or technical advice and training. The resources can take the form of grants or concessional credits (e.g., export credits). Regardless the fact that all the under developed countries need foreign capital inflows for their development, the amount and the form of the country size and the economic circumstances of the country are the major determinants of the volume and the form of the foreign capital inflows. For instance the least developed countries of Africa have been relying on the foreign aid, while the developing countries of the East- Asia are largest beneficiary of the foreign direct Investment. The composition of aid is a big factor as well.

Graph 2. Capita aid to Pakistan (1970-2007)



Source: World Bank

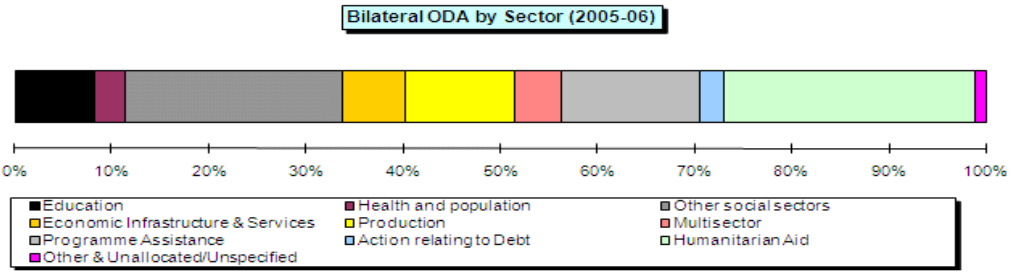
The Figures (1&2) depict an altering behavior of foreign aid over the time (1978-2007) in terms of both FA as well as Per Capita Aid. From 1978 ~1988, there is a rising trend. There is another period of decline in aid from 1990 to 1998, while there is swift increase in the aid after 2000 in terms of both FA and Per Capita aid. But the over all trend of foreign aid in Pakistan (during 1978-2007) is increasing. In most of LDCs, (including Pakistan) the population growth rate was about 3%. Thus the growth rate of 6% was assumed as a target for the rapid economic development. According to the Harrod-Domar [8] Growth Model, to achieve the target of 6% growth rate the savings (investments) must be equal to the 18-20%. The Two-Gap Model suggests that the Poor countries have to rely on the foreign capital inflows (FCI) to fill the two Gaps: Import-Export Gap and the Savings Investment Gap. The most common type of foreign aid is official development assistance (FA), which is assistance given to promote development and to combat poverty.

Pakistan

Receipts	2004	2005	2006
Net ODA (USD million)	1 424	1 626	2 147
Bilateral share (gross ODA)	23%	50%	50%
Net ODA / GNI	1.5%	1.5%	1.7%
Net Private flows (USD million)	160	883	1 519

For reference	2004	2005	2006
Population (million)	152.1	155.8	159.0
GNI per capita (Atlas USD)	600	690	770

Top Ten Donors of gross ODA (2005-06 average)	(USD m)
1 IDA	744
2 United States	410
3 AsDF	268
4 Japan	187
5 United Kingdom	159
6 Turkey	91
7 Germany	59
8 Norway	51
9 EC	50
10 Canada	47



Sources: OECD, World Bank.

The primary source of FA - which for some countries represents only a small portion of their assistance, is bilateral grants from one country to another, though some of the aid is in the form of loans, and sometimes the aid is channeled through international organizations and nongovernmental organizations (NGOs). For example, the International Monetary Fund (IMF), the World Bank, and the United Nations Children's Fund (UNICEF) have provided significant amounts of aid to countries and to NGOs involved in assistance activities. So, LDCs (including Pakistan) have to rely on the Foreign aid, rather than portfolio investments. We can see an increase in the net private inflows and an increase in the GNI as well, during 2004-2006. The top 10 donors of gross FA to Pakistan is shown in the upper-right corner of the Table No. 1. It shows IDA as the major donor, providing 744 million US$ to Pakistan in 2005- 2006 while USA being the second highest. After analyzing the list it becomes clear that Pakistan is highly dependent upon the USA, IDA, AsDF and Japan, as they are the top 4 donors to Pakistan in 2005-06.

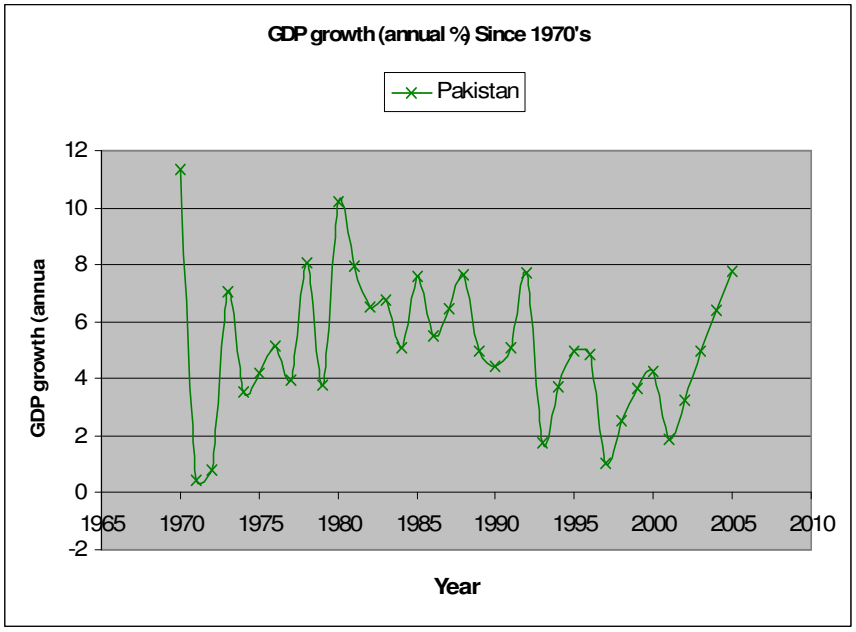
The lower portion of Table No. 2 depicts the sector-wise allocation of the FA received by the Pakistan in 2005-06. It is clear that the major share of the bilateral FA is goes to the 'humanitarian aid' and 'program assistance', since Pakistan faced a major natural disaster in the form of earthquake in 2005, whereas the share of the education, infrastructure & services, health

& population and the other social sectors remained low. We can see a significant change in the ratio of aid to action related debt, which was most dominating in the years 2002-2003. Hence, the above analysis depicts that Pakistan's economy is heavily dependent upon the IDA, USA, Japan and AsDF for the foreign aid and the major portion of the foreign economic assistance goes back to the repayment of the loans and the debt servicing.

4. MATERIALS AND METHODS

This study attempts to analyze the effectiveness of the foreign aid in economic development of Pakistan during the period 1978-2007 using for this purpose different statistical techniques. The study is based on various sources of data bases like a) World Development Reports, b) World Development Indicators, c) Government of Pakistan, Ministry of Finance, and d) Publications and Central Bank of Pakistan Reports OECD. The present paper aims to study; a) The composition of foreign aid in Pakistan since 1978; and b) The impact of foreign aid in Pakistan.

Graph 3. GDP growth in Pakistan since 1970's



Source: World Bank

Many types of studies were under taken by different authors in order to understand the impact of foreign capital inflows on the economic development of a country and many methods and variables were used for this purpose. Some of these studies focused on the study of impact of aid on the domestic saving, investment and capital formation, while, some other researchers paid much attention to study the role of aid on the debt burden, GDP growth etc. Some other studies focused upon the impact of aid on different sectors of the economy.

It is difficult to analyze the effect of foreign aid on all the sectors and variables as described above in a single study. There for, I narrow down my analysis to the impact of foreign FA and debt on the GDP. The impact of FA on the GDP growth rate was examined in many studies, in order to analyze the impact of official development assistance in Pakistan the data on the FA and the GDP given in following Table 2 is used. The regression model is fitted to the data. The quadratic model is given as:

$GDP = \alpha + \beta 1\ FA + \beta 2\ FA2$

The results of the estimated model are given as:
GDP = -5471.2245 + 85.4 FA - 0.019 FA2

Coefficient Data: Overall Significance of model
(F – Statistic) = 28.776544 *

The calculated values for the coefficients are as follows:

R squared = 0.685532 = 68%
Standard Error = 15792.44321
Correlation Coefficient (r) = 0.811254 = 81 %

The result depicts the positive effect of aid on the GDP in Pakistan (during 1978-2007) as the β_1 (the coefficient of regression) has a positive sign. But the β_2 has negative sign showing that the GDP increases as the FA increases but at the decreasing rate. All the coefficients are significant. The value of R-Squared is 58%, that is, it explains that the 58% variations in GDP explained with the help of the FA. And there is 76% correlation among the FA and GDP in Pakistan. This positive correlation between GDP and FA is also depicted in Figures (1&3). The burden of excessive dependence on the foreign aid has been caused by the: firstly, the shift in the composition of foreign aid from grants to hard loans has, over the time, thereby reducing the amount of net aid available for financing the imports and investments. Secondly, the terms and conditions attached to tie credits have imposed both economic as well as political costs on the country (Khan [9], 1993). The increased debt burden in Pakistan can be depicted by the Figure 4. It shows that the amount of external debt rises over the period of the 1978-2007 in Pakistan. It is clear that the overall debt burden also rises as the flow of foreign capital also increased over the same period.

Table 3: Foreign Aid and GDP Growth in Pakistan since 1975

Years	FA (Current US$ in Millions)	GDP (Current US$ in Millions)
1975	664.1	11340.0
1976	1013.9	13338.5
1977	588.8	15126.1
1978	635.4	17820.1
1979	711.1	19708.0
1980	1183.1	23689.7
1981	823.1	28100.6
1982	915.5	30726.0
1983	727.9	28691.9
1984	729.4	31151.8
1985	769.6	31144.9
1986	916.5	31899.1
1987	820.2	33351.5
1988	1355.4	38472.7

Table 3 (cont.)

Years	FA (Current US$ in Millions)	GDP (Current US$ in Millions)
1989	1446.3	40171.0
1990	1129.3	40010.4
1991	1371.1	45452.0
1992	1014.1	18635.2
1993	1005.4	51478.4
1994	1605.6	51894.8
1995	823.8	60636.1
1996	884.1	63320.2
1997	596.3	62433.3
1998	1053.0	62192.0
1999	732.9	62973.9
2000	702.8	73321.0
2001	1948.2	71496.2
2002	2138.2	71485.1
2003	1065.7	82350.0
2004	1421.0	96114.9
2005	1626.4	109502.1
2006	2147.2	126867.4
2007	..	143597.5

5. SUMMARY OF FDI INFLOWS IN PAKISTAN

The higher level of saving and investment is necessary to increase the rate of capital formation. However, in developing countries the level of domestic savings falls below the desired level because of low per capita income. In the case of Pakistan, domestic savings account for less than 20 percent of the GDP. This gap between domestic savings and desired level of investment can be filled by the transfer of resources from outside. FDI is one of the most important sources. To increase the level of foreign capital inflows, liberalisation of trade and investment regime by relaxing controls and offering financial and trade incentives like tax concessions and tariff reductions should be needed. Furthermore, host country should pursue active liberalisation policies to overcome trade deficit, and encourage investment in export-led sectors. To ensure that FDI stimulates domestic economic activity, the host country should make it mandatory for the foreign investor to use a certain amount of locally made inputs in the production of final goods.

The domestic policies opted by the host countries have an important influence on the decisions of foreign investment. To attract FDI, the host country should adopt concrete and investor friendly policies, strong infrastructure are the pre-conditions to restore the confidence of foreign investors.

After following somewhat restrictive economic policies, the government of Pakistan initiated market-based reforms in the 1990s. These reforms included gradual liberalisation of

trade and investment regime by providing various trade and fiscal incentives to foreign investors through tax concessions, credit facilities, tariff reduction and easing foreign exchange controls [Khan (1997)]. In the early 1990s, the government undertook a number of policy and regulatory measures14 to improve the business environment in order to attract foreign investment [Anwar (2002)]. In order to encourage FDI, restrictions on capital inflows and outflows were gradually lifted. Foreign investors were allowed to hold 100 percent of the equity of industrial project on repatriable basis without any prior approval. Furthermore, investment shares issued to non-residents could be exported, and remittance of dividends and disinvestments proceeds was permissible without any prior permission of State Bank of Pakistan (SBP). In 1994, restrictions on some capital transactions were partially relaxed, and foreign borrowing and certain outward investments were allowed to some extent. Full convertibility of the Pak-rupee was established on current international transactions. The establishment of an interbank foreign exchange market also marked an important step towards decentralising the management of foreign exchange and allowing market forces to play a greater role in exchange rate determination.

Pakistan's foreign investment regime mainly consists of three components. (I) Regulatory, (ii) economic, and (iii) socio-political. Regarding privatisation and deregulation, Pakistan has opted very liberal regulatory regime. The regulatory framework for foreign investment consists of three laws facilitating and protecting foreign investors; (i) Foreign Private Investment (Promotion and Protection) Act 1976, (ii) Furtherance and Protection of Economic Reforms Act 1992, and (iii) Foreign Currency Accounts (Protection) Ordinance 2001. In addition Bilateral Agreements include: investment protection with 43 countries and avoidance of double taxation with 51 countries. To protect the intellectual property rights (IPRs), Pakistan has also updated IPR laws to bring them in compliance with international requirements particularly, those mandatory under the Agreement on Trade Related Intellectual Property Rights (TRIPS) of the WTO. The salient features of the Pakistan's regulatory regime are: These measures include:

(1) Removal of the requirement of government approval of foreign investment,

(2) Permission of foreign equity participation of up to 100 percent,

(3) Permission to negotiate the terms and conditions of payment of royalty and technical fees suited to foreign investors for transferring technology,

(4) Liberalizing of foreign exchange regime,

(5) Permission of remittances of principal and dividends from FDI and portfolio investment including an extensive set of fiscal incentives and allowances to foreign investors,

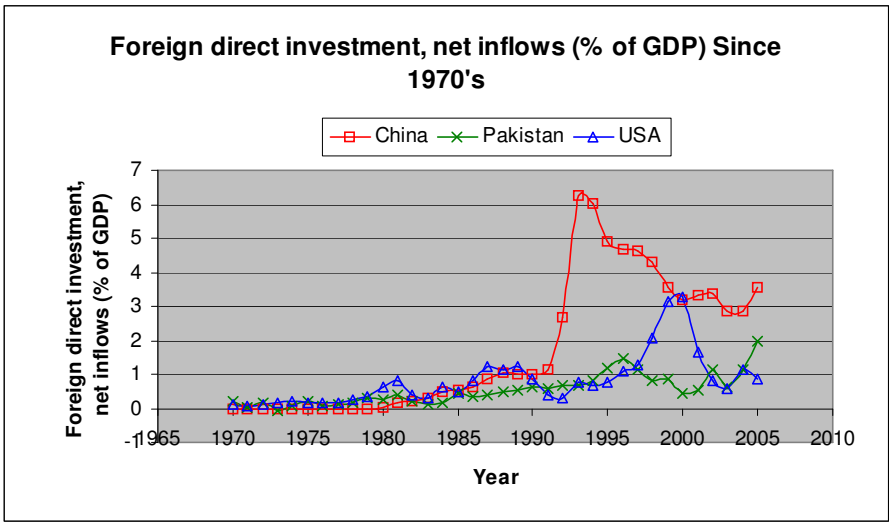
(6) Convertibility of Pak-rupee from July 1994,

(7) Liberalisation of import policy, and

(8) Opening up the sectors of agriculture, telecommunications, energy and insurance to FDI in 1997.

Further details can be seen at Zaidi (2004)

Graph 4: FDI's Net Inflows



Source: World Bank

Freedom to bring, hold and take out foreign currency from Pakistan in any form. Privatisation of an enterprise is fully protected. Neither it can be renationalized, nor can the government take over any foreign enterprise. Original FDI as well as profits earned can be repatriated to the country of origin. Equal treatment is provided to the foreign investor and local investor in terms of import and export of goods. FDI is not subject to taxes in addition to those levied on domestic investment. Foreign currency accounts are fully protected and they cannot be frozen. All the economic sectors16 including services sector are open to FDI, foreign equity up to 100 percent is allowed in all sectors. However, foreign equity up to 80 percent is allowed in agriculture sector. There is no lower limit on the size of FDI in manufacturing sector. However, in services, infrastructure and social sectors the minimum amount of foreign equity investment is $0.3 million. No government sanction is required to set up any industry, in terms of field of activity, location, and size, except arms and ammunitions, high explosives, radioactive substances, security printing, currency and mint and alcoholic beverages. No double taxation on income earned by foreign investors. Pakistan has also rationalised its tariff regime. Custom duty on import of most of the primary raw material is not more than 5 percent, while on the imported machinery is between 0 and 10 percent. Copyright law has been amended while laws regarding patents; industrial designs and trademarks have been re-enacted. There is no requirement for obtaining no objection certificate (NOC) from provincial governments for locating the project anywhere in the country except in areas that are notified as negative areas. But due to the inconsistency of government policies, the level of FDI remained low as compared to other developing countries. Pakistan has received comparatively higher amount of FDI over the past two decades due to its market-oriented investment policies and enabling environment for investment. FDI inflows to Pakistan can be explained in terms of its size and percentage of gross domestic product (GDP). Due to inconsistent policies, the flow of FDI was insignificant until 1991. However, the flow of FDI steadily increased in the post-liberalisation period. Actual inflows of FDI to Pakistan have increased from $119.6 million in the 1975-79 to $3299.8 millions Except for some sectors of strategic importance in the 1995-99.

6. CONCLUSION

The current study shows both positive as well as the negative effects of foreign aid on the economic development. Foreign aid, on positive side, has helped in boosting the GDP Growth through structural transformation of the economy, laid foundations of the industrial and agricultural sectors, provided technical assistance, policy advice and modern technology, assisted in overcoming the budget deficits. As the regression analysis of the GDP and the FA confirms the positive affect of Official Development Assistance (FA) on the GDP. Thus, the overall impact of the aid on the economic development is positive. The policies are also important in the effectiveness of foreign aid, as the aid has a more positive impact on growth in with good fiscal, monetary, and trade policies. What is the impact of foreign financial assistance on developing countries' economic development? They are facing increasing debt burden. They cannot invest in the needed sectors. On the other hand, increasing underdevelopment due to governance problems develops dependence on assistance from the developed countries. Problems associated with foreign aid occur partially because of failures on the part of aid recipients and also because the donor countries value strategic considerations more than the development concerns. Donor countries impose conditional ties that do not seem to work because the steady flow of aid is meant to provide income sources to many interest groups in the donor countries. A dominant concern remains their income, and not necessarily the well-being of the aid recipients. Poor countries need aid that is delivered in a predictable fashion, without too many strings attached to minimize transaction costs and maximize the value for money. All too often, they get aid that is unpredictable, uncoordinated and tied to purchases from the donor countries. This implies that to take the advantage of positive interaction between FDI and growth, one should liberate the economy particularly, stimulate financial sector development in the economy.

Our findings suggest that FDI plays an important role in contributing to economic growth. However, domestic financial sector development is crucial for positive effects to realise that has not been shown before. We also provide evidence that the link between FDI and growth is causal, where FDI promotes growth through financial sector development. Furthermore, the results suggest that better domestic financial conditions not only attract foreign companies but also allow host economy to maximise the benefits of foreign investments.

BIBLIOGRAPHY

ASLAM, N. “The Impact of Foreign Capital Inflow on Savings and Investment: The Case of Pakistan”. Pakistan Development Review. Islamabad: *Pakistan Institute of Development Economics*. (1987). 26 (4), 787- 89.

BURNSIDE and DOLLAR. “Aid Policies and Growth. Macroeconomics and Growth Division, Policy Research Department, World Bank”. Available at [http://econ.worldbank.org]

CHENERY, H.B. and A.M. STROUT “Foreign Assistance and Economic Development”. American Economic Review. *New York: American Economic Association.* (1966). 56 (4), 179-773.

DOMAR, E. DCAPITAL EXPANSION, “Rate of Growth, and Employment”. *Econometrica,* (1946) 14, 137–147

GAO, TING (2004) FDI, Openness and Income. *The Journal of International Trade and Economic Development*. 13: 3, 305–323

HANSEN and TARP “Aid Growth and Regressions”. *Centre for Research in Economic Development and International Trade* (CREDIT), University of Nottingham. CREDIT Research (2000). Paper No.00/7

KHAN, S. M., and A. S. SENHADJI, “Financial Development and Economic Growth: An Overview”. *International Monetary Fund*, Washington, D.C. IMF Working (2000) Paper 00/209

KHAN, N.Z. and E. RAHIM (1993). Foreign Aid, Domestic Savings and Economic Growth (Pakistan: 1960-1988). *Pakistan Development Review.* Islamabad: Pakistan Institute of Development Economics. 32 (4), 1157-67.

LEVINE, R., and S. ZERVOS Stock Markets, Banks and Economic Growth. *American Economic Review,* (1998). 88, 537–558

MWAFAQ DANDAN AL-KHALDI "Impact of Foreign Aid on Economic", Development in Jordan (1990-2005), *Journal of Social Sciences* (2008). 08-114-29

PRADHAN, JAYAPRAKASH Foreign Direct Investment and Economic Growth in Developing Countries: Further Evidence from Panel Data. *Asian Economic Review*. (2003). 45: 2, 197–217.

SOLOW, R. M,. "A Contribution to the Theory of Economic Growth. Quarterly" *Journal of Economics* (1956) 70, 65–94.

WORLD BANK, World Development Indicators: Online Database. Washington DC: (2005). *The World Bank*. at http://devdata.worldbank.org/query

ZAIDI, HUSSAIN H. (2004) Snags in the Inflow of FDI. *Dawn*, the International Edition, 09 August. Also at http://DAWN.com

ZANG, KEVIN HONGLIN (2001) How Does Foreign Investment Affect Economic Growth in China? *Economics of Transition* 9: 3, 679–693.

ÁREA VI

ECONOMÍA DE LA REGULACIÓN Y DE LA CONCURRENCIA

ECONOMIA DA REGULACÃO E DA CONCORRÊNCIA

O PROBLEMA DE FALÊNCIA: UMA ABORDAGEM EXPERIMENTAL ACERCA DO COMPORTAMENTO DOS AGENTES ECONÓMICOS USANDO REDES BAYESIANAS

FILIPE COSTA DE SOUZA
E-mail: filipecostadesouza@hotmail.com
FRANCISCO S. RAMOS
E-mail: fsr@ufpe.br
Universidade Federal de Pernambuco/PIMES

Resumo

Frequentemente, os estudos sobre leis de falência focam apenas os aspectos relativos à adequabilidade das leis ao ambiente econômico. Este trabalho utiliza o método experimental para analisar como aspectos socioeconômicos e culturais influenciam o comportamento individual em situações típicas do problema de falência, impactando assim, nos resultados esperados para lei. Neste sentido, adota-se Redes Bayesianas para determinar a relação de causalidade entre as variáveis analisadas. Com isso, torna-se possível inferir sobre comportamentos egoístas, altruístas e fraudadores, e se suas causas potenciais estão ligadas a aspectos como religião, inadimplência familiar, titularidade ou dependência de planos de saúde, entre outros. Assim, torna-se possível concluir que uma lei ótima de falência além de se adequar ao ambiente econômico, também deve considerar o perfil socioeconômico e cultural dos agentes.

Palavras chave: Leis de Falência, Economia experimental, Redes Bayesianas.
Área Temática: Economia da Regulação e da Concorrência.

Abstract

Frequently, the studies about bankruptcy laws focus only on the aspects related to the adequability of the laws to the economic environment. This work evaluates, by experimental method, how social, cultural and economics aspects of the individuals can influence on their behavior in typical situations of the bankruptcy problem, impacting on the expected result of the law. In this sense, was adopted Bayesians Networks to determinate the causality relationship among the analyzed variables. So, was possible to detect egoists altruist and fraudster behavior, and that their potential causes are related to aspects like religion, familiar nonpayment, health insurance and others. Thus, we conclude that an optimal bankruptcy law must adequate to economic environment, and also considers socioeconomics and cultural aspects of the agents.

Key words: Bankruptcy laws, Experimental economics, Bayesians Networks.
Thematic Area: Regulation and Competence Economics.

INTRODUÇÃO

A contínua necessidade de aprimorar os ambientes para negócios tem levado as nações a reformarem suas leis de falência. Essas reformas são, muitas vezes, inspiradas em aspectos de legislações estrangeiras, com vieses bastante distintos na defesa do direito dos credores. Com o Brasil não foi diferente. O antigo sistema falimentar brasileiro data de 1945, não sendo compatível com a dinâmica da economia nacional. Assim, em 1993, iniciaram-se as discussões sobre a nova lei, a qual foi concluída em 2005 (Lei 11.101, de 09 de fevereiro) e entrou em vigor no dia 09 de junho do mesmo ano.

Porém, como garantir que aspectos inspirados em leis estrangeiras se aplicarão no contexto local, alcançando os objetivos traçados pela lei? Com essa preocupação este trabalho se propõe a investigar experimentalmente como aspectos socioculturais e econômicos podem influenciar o comportamento dos indivíduos em situações típicas do problema de falência como o uso de penalidades para inibir fraudes e o conflito entre os credores na divisão dos ativos.

O método experimental foi escolhido por permitir a exposição dos agentes econômicos a situações hipotéticas controladas, permitindo uma melhor compreensão das causas do comportamento individual. Um questionário foi elaborado e aplicado e, para a avaliação dos resultados, foi utilizada uma análise estatística dos dados coletados. Além disso, para modelar qualitativa e quantitativamente as relações de causa e efeito entre as variáveis envolvidas no processo de tomada de decisão dos agentes econômicos, adotou-se Redes Bayesianas – RBs. Tais redes foram escolhidas por permitirem uma fácil interpretação das relações causais entre as variáveis do problema, via uma representação gráfica, além de quantificarem estas relações por meio de probabilidades condicionais.

A seção 2 a seguir será define o problema de falência. Na seção 3 serão introduzidas as Redes Bayesianas. A seção 4 apresenta o experimento e seu desenvolvimento. Na seção 5 serão apresentados os principais resultados do estudo e na seção 6 será analisado o comportamento dos indivíduos, na busca de causas potenciais que possam explicá-los. A seção 6 apresenta as conclusões.

1. O PROBLEMA DE FALÊNCIA

Na literatura econômica, o problema de falência é tradicionalmente abordado como a alocação de recursos divisíveis de um devedor falido entre seus *n* credores, de modo que a soma dos montantes reclamados supera o valor disponível. Ademais, o montante total a ser dividido, ou seja, o patrimônio do devedor é conhecido por todos, isto é, os credores fazem a mesma avaliação monetária do bem ou bens (GARCÍA-JURADO, GONZALES-DÍAZ & VILLAR, 2006). Aumann & Maschler (1985) definem formalmente a questão: o problema de falência é dado por um par *(E; d)*, com $d = (d_1,..., d_n)$, $0 \le d_1 \le ... \le d_n$ e $0 \le E \le \sum_{i=1}^{n} d_i$, onde E representa o patrimônio a ser repartido, e d_i representa o valor do crédito devido ao agente $i \in N = \{1, \ldots, n\}$, sendo N o conjunto dos credores.

A solução do problema, ou melhor, a alocação final dos recursos entre os *n* credores, é representada por uma n-upla $x = (x_1,..., x_n)$ de números reais. Na tentativa de se obter essa solução, utiliza-se uma função *f* que, dados os parâmetros *E* e *d,* consegue chegar a uma alocação x: $f(E; d) = x$ (AUMANN & MASCHLER, 1985). Tais funções são conhecidas como regras de falência (*bankruptcy rules*). Ademais, a solução obtida deve respeitar três restrições:

- Eficiência de Pareto da alocação final dos recursos, ou seja, $\sum_{i=1}^{n} x_i = E$.
- Os credores não podem receber mais do que o valor de seus créditos, isto é, $x_i \le d_i$.
- Os credores não são obrigados a pagar as reivindicações dos demais, ou seja, $0 \le x_i$.

O leitor pode perceber que inúmeras formas de divisão são capazes de satisfazer essas restrições, e assim chegar a uma alocação dos ativos entre os credores. Com base nessa idéia, algumas regras de alocação ou regras de falência são propostas na literatura. Herrero, Moreno-Ternero & Ponti (2003), e García-Jurado, Gonzales-Díaz & Villar (2006) expõem as algumas delas: prêmios iguais, perdas iguais e prêmios proporcionais.

- **Prêmios iguais** (*equal-awards*): o patrimônio é dividido igualmente entre os credores, assim, cada participante recebe $x_i = \min\{d_i, \lambda\}$, onde λ resolve $\sum_{i=1}^{n} \min\{d_i, \lambda\} = E$.

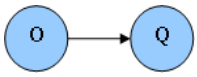
- **Perdas iguais** (*equal-losses*): todos os credores têm a mesma perda monetária, assim, cada credor recebe $x_i = \max\{0, d_i - \lambda\}$, onde λ resolve $\sum_{i=1}^{n} \max\{0, d_i - \lambda\} = E$.
- **Prêmios Proporcionais**: cada participante recebe um montante do patrimônio proporcional ao valor de seu crédito, ou seja, o credor i recebe $x_i = \lambda d_i$, onde $\lambda = E / \sum_{i=1}^{n} d_i$.

Logo, como base na exposição do problema de falência, a literatura deriva duas novas questões. A primeira se refere à escolha de mecanismos que maximizem ou preservem o valor de *E*. Essa preocupação surge, pois, gerentes ou donos de firmas em falência podem tentar fraudar a empresa para obter ganhos pessoais, depreciando assim, o patrimônio a ser dividido entre os demais credores (problema de risco moral). A segunda questão por sua vez, se refere ao conflito existente, e as diferentes possibilidades existentes para a divisão dos ativos. Assim, leis devem ser estabelecidas indicando as diretrizes a serem seguidas.

2. REDES BAYESIANAS

Segundo Pearl (1988), Redes Bayesianas são modelos probabilísticos gráficos que representam um conjunto de variáveis aleatórias e o relacionamento existente entre elas. Conceitualmente, uma Rede Bayesiana é um grafo acíclico direcionado representado pelo par (V, R), onde V é um conjunto não vazio de nós (referente às variáveis aleatórias), e R é o conjunto de pares ordenados de elementos distintos de V (representando as relações de causa e efeito entre os membros de V). Os elementos de V são definidos de forma a não permitirem ciclos, isto é, seus arcos são unidirecionais; logo, partindo-se de qualquer um dos elementos de V é impossível retornar ao mesmo. A Figura 3.1 apresenta a idéia básica de uma Rede Bayesiana: dados um grafo acíclico direcionado, G=(V, R), onde V={O, Q} e o par (O, Q) ∈ R, tem-se que O é pai, causa potencial, ou ancestral de Q, e Q é filho, efeito potencial, ou descendente de O.

Figura 3.1: Estrutura gráfica de uma Rede Bayesiana.

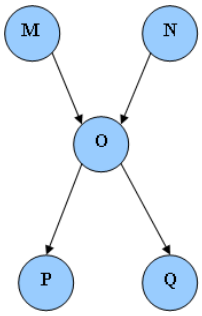


Assim, para se construir uma Rede Bayesiana, é preciso identificar inicialmente as possíveis variáveis intrínsecas ao problema e os prováveis relacionamentos de dependência entre elas. Segundo Santana *et al* (2007), essa escolha pode ser baseada na experiência do profissional (o especialista) ou por testes estatísticos e evidências na literatura. Deste modo, só após a definição dos relacionamentos é que se deve partir para o cálculo das probabilidades. Por exemplo, na Figura 1, poder-se-ia considerar ***O*** como sendo a variável "Gênero" e Q representando a variável "Ter um comportamento fraudador".

Conforme citado anteriormente, a dependência entre as variáveis pode ser interpretada como uma relação de causa e efeito, onde as relações probabilísticas são representadas por probabilidades condicionais. Desta forma, assumindo variáveis discretas, diz-se que '***O***' exerce influência sobre '***Q***' e que as probabilidades condicionais P(Q| O) e P(Q|¬O) – sendo '¬Q' o evento complementar de Q – informam o poder da causalidade entre O e Q. Essas duas probabilidades aliadas a P(O) compõem o conjunto de parâmetros de quantificação da Rede Bayesiana. Além disso, comparando as probabilidades marginais com as probabilidades condicionais, pode-se detectar de forma qualitativa a influência - positiva ou negativa - de uma variável sobre a outra.

Conforme Korb & Nicholson (2004) o uso das probabilidades condicionais permite três tipos de raciocínio: uma análise diagnóstica, uma análise preditiva ou prognóstica, e uma análise intercausal. A Figura 3.2 ilustra os possíveis raciocínios utilizando-se Redes Bayesianas.

Figura 3.2: Possíveis raciocínios via Redes Bayesianas.



Na análise diagnóstica, o raciocínio segue a direção oposta a da rede, ou seja, partindo-se de novas informações sobre o efeito, atualizam-se as crenças sobre as possíveis causas. Retomando o exemplo, tem-se que, dada uma nova informação sobre o comportamento do individuo (por exemplo, o indivíduo cometeu uma fraude - efeito potencial), é possível inferir sobre a probabilidade de este ser do sexo masculino (causa potencial). Assim, a análise diagnóstica é dada por P(causa| efeito) e, no exemplo, P(O=é do sexo masculino| Q=fraudou a firma).

Por outro lado, a análise prognóstica ou preditiva segue a direção da Rede Bayesiana, isto é, a partir de novas informações sobre a causa potencial, se atualizam as crenças sobre o efeito. O cálculo das probabilidades é feito da seguinte forma: P(efeito|causa). Assim, para um dado valor de 'O', ou seja, diante do fato que se tem uma pessoa do sexo masculino, pode-se realizar um prognóstico sobre se esta pessoa é propensa ou não a cometer fraudes através da probabilidade condicional P(Q= "ser fraudador"| O="é do sexo masculino").

Por fim, por meio da análise intercausal, é possível avaliar como duas ou mais variáveis afetam um efeito comum, como por exemplo, a P(O| M, N), ou ainda P(Q| M, O). Assim, diz-se que a análise intercausal é representada por uma estrutura em forma de 'V' na Rede Bayesiana (*v-structure*).

3. O EXPERIMENTO

Para a definição da unidade amostral partiu-se de duas considerações iniciais: a primeira foi a familiaridade dos participantes com o ambiente empresarial e com situações hipotéticas. Em segundo lugar, se procurou indivíduos que pudessem responder o questionário em lugares fixos, viabilizando o experimento. Assim, foram aplicados 174 questionários com estudantes[1] de cursos de pós-graduação e graduação em Administração, Contabilidade, Economia, Engenharia de Produção e Ciência da Computação de uma Universidade Federal.

Como instrumento de pesquisa foi utilizado um questionário, dividido em três partes básicas: dezesseis questões sobre aspectos socioeconômicos e culturais dos participantes; duas situações sobre penalidades no combate a fraude, e duas situações sobre a barganha entre os credores.

As perguntas socioeconômicas apresentadas no início do questionário tinham por objetivo permitir um maior conhecimento da amostra e, a partir dessas características, possibilitar a realização de inferências sobre o comportamento dos participantes nos casos posteriormente analisados ao longo do experimento. Para isso, após a revisão da literatura e avaliação de outros questionários já utilizados em experimentos ou em pesquisas de campo, foram selecionadas 16 questões que procuravam verificar aspectos como: gênero, nível de escolaridade dos familiares (dado que a amostra tinha um nível de escolaridade dos indivíduos quase homogêneo), nível de

[1] O trabalho com estudantes universitários é uma prática comum em experimentos econômicos, por permitir uma redução nos custos da pesquisa, tornado-a viável. Trabalhos como de Bianchi (1998), Castro & Ribeiro (2000) são apenas alguns exemplos dessa prática.

renda familiar (renda *per capita*), nível de inadimplência e religião, entre outros. A seguir será apresentada a modelagem dos casos referentes a fraude e a barganha entre os credores.

3.1. Uso de penalidades para inibir o risco moral

Foram definidos dois casos com o objetivo de avaliar se o uso de penalidades legais inibiria os proprietários ou gerentes de empresas em falência a realizarem ações que depreciariam o patrimônio da firma, na tentativa de garantir benefícios pessoais. No primeiro caso, chamado de **Caso 1**, os participantes eram inicialmente inquiridos se fraudariam uma firma em processo falimentar a fim de obter ganhos pessoais, caso não existisse possibilidade de serem penalizados pela fraude. Em seguida, foi questionado se eles permaneceriam com a postura de fraudar caso pudessem ser presos por até 6 anos na eventualidade que a fraude fosse descoberta. A Figura 4.1 mostra os questionamentos realizados.

Figura 4.1: os Casos 1 e 2.

Caso 1: Você é gerente de uma empresa que está falindo, sendo essa sua única fonte de renda. É possível fraudar alguns documentos, de modo a transferir recursos da firma para a sua conta pessoal. Além disso, o governo não é capaz de descobrir a fraude e, desse modo, você nunca será penalizado. Nessas circunstâncias você cometeria a fraude?

☐ Sim ☐ Não

Caso 2: Agora, imagine-se novamente no Caso 1. Porém, caso o governo descubra a fraude, você pode ser preso por até 6 anos. Nessas circunstâncias você cometeria a fraude?

☐ Não.
☐ Sim, independente do valor.
☐ Sim, dependendo do valor ganho.

3.2. A barganha entre os credores

Com relação aos casos referentes à barganha, os estudantes se depararam com dois problemas. Eles deveriam escolher como se daria a divisão dos ativos da firma falida entre três credores, sendo ele um dos credores. O questionamento é exposto na Figura 4.2.

Figura 4.2: A barganha entre os credores.

Você é gerente de um banco. Uma empresa que acaba de falir deve dinheiro a três instituições financeiras, inclusive a sua. Assim, vocês devem decidir simultaneamente (sem direito a negociação) como será realizada a divisão da firma. Cada cenário abaixo apresenta diferentes valores para a firma, desse modo, em cada quadro você deve escolher a melhor opção seguindo as seguintes regras.

Os quadros do questionamento consistiam de diferentes cenários, onde os participantes deveriam escolher dentre quatro alternativa, qual seria a mais apropriada para a partilha dos bens entres os credores, como mostra a Figura 4.3. As alternativas utilizadas foram as regras de falência definidas na seção 2 (prêmios iguais, perdas iguais, e divisão proporcional), além de uma nova alternativa, chamada aqui de '*venda da firma*'. Tal opção tinha como propósito

representar um ponto de desacordo entre os participantes, exprimindo, com isso, a diferença de senioridade (prioridade de pagamento) existente entre eles. Deste modo, as demais opções (prêmios iguais, perdas iguais, e divisão proporcional) tinham como objetivo representar a forma de pagamento dos credores em caso de reorganização da firma. As denominações X, Y e Z foram utilizadas para não influenciar os participantes durante o experimento, e representam a divisão proporcional, ganhos iguais e perdas iguais respectivamente. O leitor pode perceber que, em todos os cenários, a soma dos ganhos dos participantes nas opções X, Y ou Z é de 30.000, superando assim, a soma dos valores da opção de venda da firma.

Os valores escolhidos para os créditos foram inspirados nos valores propostos por Herrero, Moreno-Ternero & Ponti (2003), pois cada credor teria uma forma preferida de fazer a divisão dos ativos diferente da dos demais (outros valores também foram testados, mas não apresentavam de forma tão direta e simples tais características). No início do processo, cada participante recebia um tipo, que poderia ser A, B ou C, imutável durante todo o experimento. A amostra foi composta por 57 participantes do tipo A, 57 participantes do tipo B e 60 participantes do tipo C (totalizando 174 participantes).

Figura 4.3: A barganha entre os credores.

Cenário 1:		( )	( )	( )	( )
Tipo	**Seu crédito**	**Venda da Firma**	**X**	**Y**	**Z**
A	49.000	10.000	14.700	12.500	16.500
B	5.000	0	1.500	5.000	0
C	46.000	0	13.800	12.500	13.500
Cenário 2:		( )	( )	( )	( )
Tipo	**Seu crédito**	**Venda da Firma**	**X**	**Y**	**Z**
A	5.000	5.000	1.500	5.000	0
B	46.000	2.400	13.800	12.500	13.500
C	49.000	2.600	14.700	12.500	16.500

Cenário 3:		( )	( )	( )	( )
Tipo	**Seu crédito**	**Venda da Firma**	**X**	**Y**	**Z**
A	46.000	20.000	13.800	12.500	13.500
B	49.000	0	14.700	12.500	16.500
C	5.000	0	1.500	5.000	0

Assim, baseados nos seus tipos, os participante deveriam responder duas vezes que opção iriam propor para a divisão da firma em cada cenário, sem direito a negociação com os demais, sendo cada escolha regida por uma regra diferente:

Regra 1:

- A opção vencedora será aquela que tiver TODOS os votos.
- Caso não se chegue a um consenso, a opção vencedora será a "Venda da Firma".

Regra 2:

- A opção vencedora será aquela que obtiver a MAIORIA dos votos.
- Caso nenhuma opção obtenha dois ou mais votos, a opção vencedora será a "Venda da Firma".

A Regra 1 foi inspirada nas legislações pró-credor, permitindo que o credor tipo A, que representa um credor com garantia real, tivesse maiores facilidades na venda da firma. Os credores tipo B e C representariam credores sem garantias reais. Já a Regra 2 foi inspirada nas

legislações pró-credor com viés mais moderado, onde o conselho dos credores decidiria o futuro da firma[2].

Além disso, com base na regra 1, dois comportamentos individuais foram definidos: o comportamento egoísta e o comportamento altruísta. O indivíduo egoísta foi definido como o jogador tipo A que no cenário 2 escolhe a 'venda da firma' deixando valores menores para os outros participantes, quando deveria escolher a opção 'divisão igual' (Y), a qual permitiria ganhos maiores para os demais jogadores e preservaria o valor de seu ganho. Assim, a atitude egoísta levaria à venda da firma quando esta poderia ser reorganizada. Por sua vez, o indivíduo altruísta foi definido como o jogador tipo A que, no cenário 3, não escolheu a venda da firma, reduzindo seus ganhos certos, mas permitindo ganhos maiores ao demais participantes. A atitude altruísta poderia levar à reorganização da firma quando esta deveria ser liquidada.

Tanto o comportamento egoísta quanto o altruísta levam em conta os ganhos dos demais participantes na escolha individual. Além disso, o egoísmo e o altruísmo são duas variáveis distintas, não devendo ser vistas como atitudes excludentes. Um participante pode demonstrar os dois comportamentos ao mesmo tempo. Maiores explicações serão apresentadas nos resultados do experimento na seção 6.

4. O PERFIL DA AMOSTRA

Analisando-se algumas mediadas de tendência central e variabilidade é possível perceber, com mostra a Tabela 5.1 que a idade média dos participantes era de aproximadamente 26 anos, o número médio de filhos era de 0,28 (mais de 85% da amostra não tinha filhos), e o número médio de moradores por residência era de 3,74 indivíduos. Segundo o IBGE[3], o tamanho médio das famílias em Pernambuco é de 3,62 indivíduos, valor este bem próximo do encontrado na pesquisa.

Tabela 5.1: Idade, número de filhos e número de moradores por residência.

	Média	Mediana	Moda	Desvio padrão	Max	Min
Idade	26,1	24	23	7,43	63	19
Nº de filhos	0,28	0	0	0,79	0	5
Nº de moradores	3,74	4	4	1,3	1	7

Fonte: elaborada pelos autores

Analisando o nível de escolaridade das mães dos participantes, tem-se que a moda foi o nível superior completo, com 60 observações, ou seja, aproximadamente 34,68% da amostra. Porém, em mais de 50% dos casos, as mães não têm nível superior completo (um participante não respondeu o nível de escolaridade da mãe). Com relação ao nível de escolaridade do pai, assim como no da mãe, a moda foi o nível superior completo com 46 observações, ou seja, aproximadamente 26,74%. Contudo, aproximadamente 58% dos pais têm escolaridade inferior ao nível superior completo (dois participantes não responderam o nível de escolaridade do pai). Os dados completos podem ser visto na Tabela 5.2.

Com relação à renda familiar dos participantes, a moda foi as famílias dos ganharem mais de 6000 reais, com 52 observações, ou seja, aproximadamente 31% das observações. A distribuição completa da renda familiar dos participantes pode ser vista na Tabela 5.3 (cinco pessoas não responderam esta pergunta).

[2] A regra 1 foi inspirada na legislação de falência da Inglaterra, pois permite que o credor A, que representaria um credor com garantia real, pudesse vender a firma a qualquer instante, independente da vontade dos demais participantes, indicando o viés pró-credor da lei inglesa. Além disso, tal regra torna a reorganização, representada pelas opções X, Y e Z, mais difícil de ser alcançada. A regra 2, por sua vez, foi inspirada na legislação da Alemanha, pois o sistema de insolvência germânico permite uma grande participação dos credores no processo decisório. O voto de cada credor é proporcional ao seu crédito, sendo a opção vencedora a que tiver mais de 50% da aprovação dos créditos votantes. Como neste experimento para que a soma dos créditos ultrapassasse 50% do total dos créditos exigiria dois ou mais votos, a regra foi substituída pela regra da maioria.

[3] Fonte: IBGE, Diretoria de Pesquisas, Coordenação de Índices de Preços de Orçamentos Familiares 2002-2003.

Tabela 5.2: Tabela de freqüência da escolaridade da mãe e pai dos participantes.

Grau de escolaridade	Mãe				Pai			
Categorias	**Qtd.**	**ΣQtd.**	**%**	**Σ%**	**Qtd.**	**ΣQtd.**	**%**	**Σ%**
Não alfabetizado	0	0	0	0	0	0	0	0
Fundamental incompleto	10	10	5,78	5,78	16	16	9,30	9,30
Fundamental completo	10	20	5,78	11,56	13	29	7,56	16,86
Médio incompleto	10	30	5,78	17,34	9	38	5,23	22,09
Médio completo	49	79	28,33	45,67	44	82	25,57	47,66
Superior incompleto	10	89	5,78	51,45	18	100	10,47	58,13
Superior completo	60	149	34,68	86,13	46	146	26,74	84,87
Especialização Superior	21	170	12,14	98,27	13	159	7,58	92,45
Mestrado	3	173	1,73	100,00	7	166	4,07	96,52
Doutorado	0	173	0	100,00	6	172	3,48	100,00

Fonte: elaborada pelos autores

Tabela 5.3: Tabela de freqüência do nível de renda familiar.

Categorias	Qtd.	ΣQtd.	%	Σ%
Até 400	0	0	0	0
De 400 a 600	3	3	1,77	1,77
De 600 a 1000	8	11	4,73	6,50
De 1000 a 1200	6	17	3,55	10,05
De 1200 a 1600	13	30	7,69	17,74
De 1600 a 2000	16	46	9,47	27,21
De 2000 a 3000	24	70	14,20	41,41
De 3000 a 4000	21	91	12,43	53,84
De 4000 a 6000	26	117	15,38	69,22
Mais de 6000	52	169	30,78	100,00

Fonte: elaborada pelos autores

Quando questionados se a família teve alguma dívida nos últimos 12 meses, os participantes apresentaram as seguintes respostas como mostra a Tabela 5.4. O tipo de dívida mais comum se enquadrou na categoria de despesas com água, eletricidade e telefone, com aproximadamente 29% de inadimplentes. Além disso, quando analisado o número de categorias em que a família dos participantes tem dívidas, constatou-se que, aproximadamente, 56% dos participantes não tiveram inadimplências nos últimos 12 meses, e apenas 4% tiveram dívidas não pagas em todas as categorias, como mostra a Tabela 5.5.

Tabela 5.4: Relação de inadimplência por categoria.

Tipo da dívida	Pagou (%)	Inadimplente (%)	Total (%)
Aluguel ou prestação da casa, apartamento, condomínio, etc.	89,66	10,34	100
Despesas com água, eletricidade, telefone.	70,69	29,31	100
Pagamento de prestações de bens ou serviços adquiridos.	82,78	17,22	100

Fonte: elaborada pelos autores

Tabela.5.5: Tabela de freqüência do número de categorias com dívidas.

Categorias	Qtd.	Σ Qtd.	%	Σ%
Sem dívidas	97	97	55,75	55,75
Dívida em apenas 1 categoria	62	159	35,63	91,35
Dívida em 2 categorias	8	167	4,60	95,98
Dívida em todas as categorias	7	174	4,02	100,00

Fonte: elaborada pelos autores

Finalizando a análise das características socioeconômicas da amostra, avaliou-se a religião dos participantes. A Tabela 5.6 mostra que 56,4% dos participantes se denominam católicos, 20,9% sem religião, 11% são evangélicos, 9,3% espíritas, e 2,3% se enquadram na categoria outra. Como a amostra é composta apenas por estudantes universitários, também comparou-se os dados do experimento com os do Censo Demográfico de 2000, para pessoas com mais de 15 anos de escolaridade.

Tabela 5.6: Relação entre composição da amostra por religião e Censo 2000.

Religião	Experimento (%)	Censo 2000* (%)
Católico	56,4	74
Sem religião	21	6,7
Evangélico	11	9,3
Espírita	9,3	6,7
Outra	2,3	3,3
Total	100	100

**Fonte: IBGE, Censo Demográfico 2000.*

Pela comparação, pode-se perceber que a maior diferença existente está na religião Católica e nos indivíduos que se classificaram sem religião. Esse fato pode ser explicado, pois diferente das outras religiões, a Católica tem um número mais elevado de indivíduos não praticantes, o que pode levar tais indivíduos a se avaliarem sem religião, mesmo sendo católicos de batismo.

5. RESULTADOS

5.1. Penalidades inibem fraudes?

Avaliando inicialmente o Caso 1, tem-se que 36 indivíduos (aproximadamente 20% da amostra) afirmaram que, nas circunstâncias descritas, fraudariam a firma e 138 afirmaram o contrário (que não fraudariam). Esses indivíduos que afirmaram cometer a fraude foram denominados neste trabalho de 'fraudadores'. É importante ressaltar que os mesmos não apresentaram penalidades morais que os inibissem de cometer o ato ilícito. Entretanto, quando existe a possibilidade de ser penalizado, apenas 5,7% da amostra (10 participantes) permanecem com a postura de fraudar a firma. Desses, nove afirmaram fraudar dependendo do valor a ser ganho e apenas um afirmou que cometeria a fraude, independentemente do valor a ser ganho.

Em seguida, foi verificado se a possibilidade de punição por até 6 anos inibiria as fraudes. Para isso, foi utilizado o teste Chi Quadrado de independência (HEALY, 1945), cujas hipóteses são:

H_0: o ato de fraudar independe da penalidade, isto é, as variáveis são independentes.

H_1: o ato de fraudar depende da penalidade, ou seja, as variáveis são dependentes.

Com o p-valor=0,00 e um nível de significância de 5%, tem-se que é possível rejeitar H_0, isto é, as variáveis são dependentes: portanto, a penalidade auxilia a reduzir as fraudes.

5.2. A alocação de recursos pela barganha

O cenário 1 pela regra 1

No cenário 1 existem dois equilíbrios de Nash em estratégia pura[4], que ocorrem quando todos os participantes escolhem a estratégia 'divisão proporcional' (X), ou quando todos escolhem a estratégia 'prêmios iguais' (Y). Assim, no cenário 1, um total de 143 participantes escolheu uma das possíveis estratégias que levariam a um equilíbrio, ou seja, aproximadamente 83% da amostra. Neste caso, a moda foi a opção 'prêmios iguais' com 76 votos, fato que

[4] O percentual de participantes que escolheu um dos possíveis equilíbrios dos jogos em questão serão destacadas com um '*'.

poderia indicar que, em circunstância semelhante à proposta, a divisão igual dos ganhos entre os participantes poderia ser um ponto focal. A Tabela 6.1 mostra as escolhas dos participantes, no cenário 1 pela regra 1.

Tabela 6.1: Escolhas dos participantes no cenário 1 pela regra 1.

Alternativa	Tipo A (%)	Tipo B (%)	Tipo C (%)	Todos (%)
Venda da Firma	15,79	12,28	3,33	10,34
Proporcional (X)	29,83*	24,56*	60,00*	38,51
Ganhos iguais (Y)	36,84*	61,40*	33,33*	43,68
Perdas iguais (Z)	17,54	1,76	3,33	7,47
Total	100	100	100	100

O cenário 2 pela regra 1

No cenário 2 o único equilíbrio de Nash ocorre quando todos os participantes escolhem a estratégia 'prêmios iguais'. Entretanto, alguns fatos curiosos foram constatados. Neste cenário, a moda foi a 'divisão proporcional', com 65 votos, desviando-se do equilíbrio teórico do jogo. Esse desvio foi provocado, em grande parte, pelos votos dos participantes dos tipos B e C. Como explicação para este fato, pode-se supor a descrença na cooperação do participante tipo A, ou ainda, o fato de o crédito do participante tipo A ser de apenas 5.000, o que levaria à escolha da estratégia 'divisão proporcional' pelos demais participantes como uma divisão mais justa. Além disso, nesse cenário pode-se destacar um comportamento específico dos participantes do tipo A: o egoísmo. Foi constatado que dos 57 participantes tipo A, 22 escolheram a alternativa "venda da firma", ou seja, cerca de 39%. A Tabela 6.2 mostra a escolha de cada tipo de participante no cenário 2 pela regra 1.

Tabela 6.2: Escolhas dos participantes no cenário 2 pela regra 1.

Alternativa	Tipo A (%)	Tipo B (%)	Tipo C (%)	Todos (%)
Venda da Firma	38,60	10,53	3,33	17,24
Proporcional (X)	17,54	59,65	35,00	37,36
Ganhos iguais (Y)	42,11*	22,81*	30,00*	31,61
Perdas iguais (Z)	1,75	7,01	31,67	13,79
Total	100	100	100	100

O cenário 3 pela regra 1

Partindo para a análise do cenário 3, o equilíbrio de Nash do jogo ocorre quando o indivíduo tipo A escolhe a opção de venda da firma, independente da escolha dos demais participantes. Para os participantes B, a moda foi a opção 'divisão proporcional', e para o jogador C foi a opção 'perdas iguais', representando as estratégias que lhes dariam maiores ganhos. Para o jogador A, a moda foi a opção venda da firma, como já era esperado, com 54% dos votos. Porém, pode-se perceber que quase metade dos participantes do tipo A se desviaram da solução teórica. Assim, tais indivíduos foram classificados como 'altruístas', pois preferiram uma alternativa que reduz seus ganhos, mas que permite uma divisão mais igualitária. A Tabela 6.3 mostra a escolha de cada tipo no cenário 3 pela regra 1.

Tabela 6.3: Escolhas dos participantes no cenário 3 pela regra 1 .

Alternativa	Tipo A (%)	Tipo B (%)	Tipo C (%)	Todos (%)
Venda da Firma	54,38*	3,51*	5,00*	20,69
Proporcional (X)	26,32	42,11*	25,00*	31,04
Ganhos iguais (Y)	15,79	21,05*	68,33*	35,63
Perdas iguais (Z)	3,51	33,33*	1,67*	12,64
Total	100	100	100	100

Durante o problema alocação de recursos, foi possível perceber que aproximadamente 39% dos participantes, se comportaram de modo egoísta; e aproximadamente 46% se comportaram de modo altruísta. Então, foi verificado se há dependência entre o comportamento egoísta e altruísta. Do mesmo modo que na seção 4, foi utilizado o teste Chi Quadrado:

H_0: as variáveis são independentes.

H_1: as variáveis são dependentes.

Com o p-valor=0,00 e um nível de significância de 5%, tem-se que é possível rejeitar H_0, isto é, as variáveis são dependentes; além disso, a correlação entre as variáveis é de -0,44, indicando que indivíduos altruístas tendem a ser não egoístas.

Outros trabalhos experimentais como Berg, Dickmaut & McCabe (1995), Straub & Murnighan (1995), Castro & Ribeiro (2000), Haley & Fessler (2005) e Bettinger & Slonim (2006) também apontam evidências de altruísmo, sendo essa generosidade fortemente associada à idéia de reciprocidade em jogos repetidos.

O cenário 1 pela regra 2

Nesse cenário, os indivíduos tipo A poderiam alcançar o equilíbrio escolhendo qualquer estratégia, exceto a venda da firma, e o mesmo ocorreria para o indivíduo tipo C. Por sua vez, o participante tipo B poderia escolher as opções 'proporcional' ou 'prêmios iguais'. A Tabela 6.4 mostra a escolha de cada jogador no cenário 1 pela regra 2.

Pela análise dos resultados, é possível perceber que a maior freqüência de jogadas dos participantes tipo B e C reside nas opções que os levaria ao melhor resultado (divisão igual e proporcional, respectivamente). Porém para o indivíduo tipo A, o qual perdeu o poder de barganha com a mudança da regra, a opção escolhida foi a 'divisão proporcional', indicando uma preferência por ganhos mais seguros e menores, por meio da cooperação, do que arriscar ganhos maiores.

Tabela 6.4: Escolhas dos participantes no cenário 1 pela regra 2

Alternativa	Tipo A (%)	Tipo B (%)	Tipo C (%)	Todos (%)
Venda da Firma	7,01	8,77	0	5,17
Proporcional (X)	42,11*	42,11*	70,00*	51,72
Ganhos iguais (Y)	24,56*	45,61*	18,33*	29,31
Perdas iguais (Z)	26,32*	3,51	11,67*	13,80
Total	100	100	100	100

O cenário 2 pela regra 2

Neste cenário, os indivíduos tipo A poderiam escolher tanto a opção de venda da firma, quanto a opção 'divisão igual'. Os demais jogadores poderiam escolher qualquer estratégia, exceto a opção de venda da firma. A Tabela 6.5 mostra a escolha de cada tipo no cenário 2 pela regra 2.

Tabela 6.5: Escolhas dos participantes no cenário 2 pela regra 2.

Alternativa	Tipo A (%)	Tipo B (%)	Tipo C (%)	Todos (%)
Venda da Firma	14,04*	3,51	5,00	7,47
Proporcional (X)	31,58	57,89*	40,00*	43,10
Ganhos iguais (Y)	50,88*	19,30*	21,67*	30,46
Perdas iguais (Z)	3,50	19,30*	33,33*	18,97
Total	100	100	100	100

Na Tabela 6.5 é possível perceber que apenas 14% dos participantes escolheram a opção de venda da firma, valor bem inferior aos 39% de indivíduos egoístas pela regra 1. Mesmo considerando a regra 2, tal opção representa uma atitude racional. Em contrapartida, as opções

'divisão proporcional' e 'divisão igual' tiveram suas freqüências ampliadas, reforçando uma preferência pela cooperação após a perda do poder de barganha.

O cenário 3 pela regra 2

Nesse cenário, o jogador tipo A deveria escolher a venda da firma, e os participantes B e C poderiam optar pelas estratégias 'divisão proporcional' ou 'divisão igual'. A Tabela 6.6 mostra a escolha de cada tipo no cenário 3 pela regra 2.

Tabela 6.6: Escolhas dos participantes no cenário 3 pela regra 1.

Alternativa	Tipo A (%)	Tipo B (%)	Tipo C (%)	Todos (%)
Venda da Firma	22,81*	3,51	3,33	9,77
Proporcional (X)	45,61	35,09*	31,67*	37,36
Ganhos iguais (Y)	17,54	29,82*	63,33*	37,36
Perdas iguais (Z)	14,04	31,58	1,67	15,51
Total	100	100	100	100

Avaliando a Tabela 6.6, é possível perceber que apenas 22,8% dos participantes do tipo A escolheram a venda da firma, porém a moda foi a opção 'divisão proporcional'. Esse comportamento evidencia novamente que os indivíduos preferem garantir a segunda melhor opção (*second best*), a arriscarem receber valores pequenos.

De um modo geral, realizando uma avaliação de todos os cenários pela regra 2, a moda dos votos na coluna 'todos' foi a divisão proporcional, fato que ajuda a corroborar a idéia de que tal estratégia de divisão é uma forma natural de partilha em um processo de falência (HERRERO, MORENO-TERNERO & PONTI, 2003, GARCÍA-JURADO, GONZALES-DÍAZ &VILLAR, 2006).

6. AVALIANDO O COMPORTAMENTO INDIVIDUAL VIA REDES BAYESIANAS

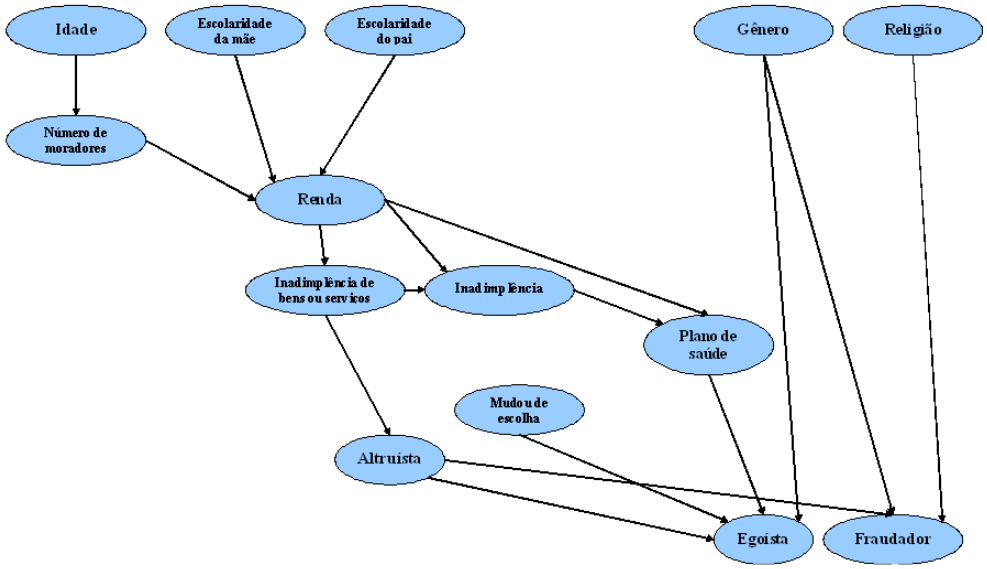
Após o cálculo dos equilíbrios e constatação de algumas características dos participantes, foi realizada a definição de um perfil para os agentes de comportamento egoísta, altruísta e que fraudariam a firma no Caso 1, procurando possíveis explicações (causas) para esses comportamentos.

A Figura 7.1 exibe a Rede Bayesiana correspondente aos dados. Este modelo permite realizar inferências diagnósticas e prognósticas sobre as características potenciais de indivíduos egoístas, altruístas e fraudadores no Caso 1 (fraudador), além, é claro, de permitir uma análise intercausal, dando uma idéia integrada do problema. Vale ressaltar que, devido a limitações no tamanho da amostra, a análise intercausal só será realizada de modo qualitativo, pelo gráfico da Rede Bayesiana.

Para escolher a topologia da Rede Bayesiana foram realizados inicialmente testes estatísticos para definir a relação de dependência entre as demais variáveis do questionário e os comportamentos definidos (altruísta, egoísta e fraudador). Em seguida, se procurou razões lógicas (como o aspecto temporal) e evidências presentes na literatura para definir a causalidade entre as variáveis onde a dependência foi detectada[5].

[5] * indica que a relação de dependência é significante a 1%, ** indica que a relação de dependência é significante a 5%, *** indica que a relação de dependência é significante a 10%. Quando o nível de significância não for especificado, está implícito que ele é de 5%.

Figura 7.1: Estrutura da Rede Bayesiana do experimento.



O indivíduo Fraudador: Inicialmente tem-se que a probabilidade de um indivíduo ser fraudador é de aproximadamente 0,20. Realizando uma análise diagnóstica, é possível perceber que o indivíduo fraudador tende a ser: homens*, não evangélicos** e não altruístas***. Portanto:

- P(Fraudador| Homem)=0,28, por sua vez, P(Fraudador| Mulher)=0,068.
- P(Fraudador| Evangélico)=0, enquanto, P(Fraudador| Não evangélico)=0,23.
- P(Fraudador| Altruísta)=0,038, enquanto, P(Fraudador| Não Altruísta)=0,23.

O indivíduo Egoísta: Inicialmente tem-se que probabilidade de um indivíduo ser egoísta é de aproximadamente 0,39. Realizando uma análise diagnóstica é possível perceber que o indivíduo egoísta tende a ser: Mulher***, não tem plano de saúde**, não são altruístas*, são mais cooperativos quando perdem o poder de barganha[6]* e a família teve alguma inadimplência nos últimos 12 meses***. Ou seja:

- P(Egoísta| Mulher)=0,56, enquanto, P(Egoísta| Homem)=0,31.
- P(Egoísta| tem plano de saúde)=0,32, e P(Egoísta| Não tem plano de saúde)=0,83.
- P(μ=Egoísta| Altruísta)=0,15, e P(Egoísta| Não altruísta)=0,58
- P(Egoísta| Mudou de escolha)=0,53, e P(Egoísta| Não mudou de escolha)=0,17.
- P(Egoísta| Não Pagou todas as contas)=0,52, e P(Egoísta| Pagou todas as contas)=0,28

O indivíduo Altruísta: A probabilidade de um indivíduo ser altruísta é de aproximadamente 0,46. Realizando uma análise diagnóstica é possível perceber que o indivíduo altruísta tende a ser: "A família não teve dívidas de bem e serviços nos últimos 12 meses**; realizando uma análise prognóstica, eles tendem a ser: não egoístas* e não fraudadores***. Portanto:

- P(Altruísta| Pagou os bens e serviços)=0,53, e P(Altruísta| Não pagou os bens e serviços)=0,17
- P(Altruísta| Não egoísta)=0,62, enquanto, P(Altruísta| Egoísta)=0,18.
- P(Altruísta| Fraudador)=0,12, enquanto, P(Altruísta| Não fraudador)=0,51.

[6] Indica que os indivíduos egoístas mudaram de opção no cenário 2 quando passaram da regra 1 para a regra 2, onde a sua escolha depende também da escolha dos demais jogadores.

7. CONCLUSÃO

Como primeira constatação, tem-se o já esperado resultado de que penalidades desestimulam a prática de atividades ilícitas. Ademais, o uso de Redes Bayesianas se mostrou adequado ao objetivo proposto por permitir a identificação de comportamentos que se desviam dos pressupostos teóricos, como o altruísmo e o egoísmo, bem como, o cálculo das probabilidades de um indivíduo com determinadas características socioculturais e econômicas de se enquadrar nesses comportamentos.

O comportamento altruísta e egoísta, como definidos, poderia levar à venda da firma em situações em que ela deveria ser reorganizada, ou ao caso contrário, a reorganização da firma quando todos os fatores indicam a sua venda. Esses comportamentos (egoísta, altruísta e fraudador) apresentam como causas potenciais aspectos como religião, inadimplência familiar, titularidade ou dependência de planos de saúde (aspectos relacionados à renda), entre outros, fornecendo uma visão integrada das variáveis. Contudo, é evidente que os cenários apresentados são muito limitados e não têm o objetivo de simular fielmente uma situação de falência real, apenas tentam isolar os efeitos que influenciam na tomada de decisão e, conseqüentemente, mesmo em situações mais complexas deveriam ser levadas em consideração.

Tais constatações ressaltam a importância da análise desses variáveis no processo de escolha individual em atividades ligadas ao problema de falência e, portanto, também devem ser consideradas pelos formuladores de leis durante a elaboração ou reforma das mesmas. Assim, espera-se que uma lei de falência ótima e as teorias que a suportam, além de se adequarem ao ambiente econômico, também levem em consideração o perfil sociocultural e econômico da população, de modo garantir de forma mais apropriada que os objetivos traçados para a lei sejam alcançados.

Trabalhos futuros podem ser direcionados no sentido de avaliar a possibilidade de criação de lei de falência para pessoa física no Brasil e seus principais impactos econômicos, bem como avaliar experimentalmente como os agentes econômicos reagem a combinações entre benefícios e penalidades, como incentivos a comportamento não fraudatório.

REFERÊNCIAS

AUMANN. R. J.; MASCHLE. M, (1985). Game Theoretic analysis of a bankruptcy problem from Talmud. *Journal of Economic Theory*. 36, 195-213.

BETTINGER, E.; SLONIM, R. (2006). Using experimental economics to measure the effects of a natural educational experiment on altruism. *Journal of Public Economics*. 90, 1165-1148.

BERG, J.; DICKMAUT, J.; MCCABE, K. (1995) Trust, Reciprocity, and Social History. *Game and Economic behavior*. 10, 122-142.

BIANCHI. A. M. (1998). Are Brazilian economists different?. *Revista Brasileira de Economia*. 52, 3, 427-439.

CASTRO, J. D; RIBEIRO, E. P. (2000). Um teste empírico para a teoria dos jogos: o modelo da racionalidade egoísta. In: XXVIII Encontro Nacional de Economia, 2000, Campinas, SP. Anais do XXVIII Encontro Nacional de Economia. Campinas, SP: ANPEC.

GARCÍA-JURADO, I.; GONZALES-DÍAZ, J.; VILLAR, A. (2006). A non-cooperative approach to bankruptcy problems. *Spanish Economic Review*. 8, 189-197.

HALEY, KEVIN; FESSLER, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Beha*vior. 26, 245-256.

HERRERO, C.; MORENO-TERNERO, J. D; PONTI, G. (2003). An experiment on bankruptcy, *working paper*.

KORB. H. B.; NICHOLSON. A. E. (2004). Bayesian Artificial Intelligence. New York: Chapman & Hall.

PEARL, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Califórnia: Morgan Kaufmann.

SANTANA A. L. *et al*. (2007). Strategies for improving the modeling and interpretability of Bayesian networks. *Data & Knowledge Engineering*. 63, 91–107.

STRAUB, P.G; MURNIGHAN. J.K. (1995) An experimental investigation of ultimatum games: information, fairness, expectation, and lowest acceptable offers. *Journal of Economic Behavior and Organization.* 27, 345-364.

DESCONTOS DE QUANTIDADE COM EFEITOS DE FIDELIZAÇÃO: UMA PROPOSTA DE ANÁLISE JUS CONCORRENCIAL

ANA OLIVEIRA-BROCHADO
abrochado@concorrência.pt
Autoridade da Concorrência

Resumo

A análise dos descontos com efeitos de fidelização recebeu uma importância crescente após as decisões do Tribunal de Primeira Instância relativamente aos casos Michelin II e British Airways. Os descontos com efeitos de fidelização incluem vários tipos de incentivos atribuídos por uma empresa dominante aos seus consumidores ou distribuidores quando um determinado nível de compras é alcançado. De acordo com a jurisprudência comunitária, os descontos com efeitos de fidelização oferecidos por empresas dominantes são incompatíveis luz do artigo 82 do Tratado CE, na medida em que tendem a restringir a liberdade do comprador na escolha do fornecedor, e por isso limitam o acesso dos concorrentes ao mercado e tendem a encerrar o mercado. No entanto, dependendo das circunstâncias, os descontos podem ter também efeitos pró-concorrenciais. Logo, este trabalho pretende apresentar os princípios económicos chave que devem ser utilizados para aferir os efeitos dos desconto e discutir o contexto e que estes descontos, quanto utilizados por empresas dominantes, geram efeitos de exclusão.

Palavras chave: posição dominante; concorrência; descontos de fidelização, Europa.
Área Temática: Economia da Regulação e da Concorrência.

Abstract

The analyses of loyalty rebates by dominant firms has received an increasing attention since two high-profile decisions of the Court of First Instance in 2003 on cases Michelin II e British Airways. Loyalty discounts include several different schemes of incentives granted by a dominant undertaking to customers and distributors if a certain level of purchases is met. According to established case law, fidelity discounts offered by dominant undertakings are incompatible with article 82 EC, because tend to restrict the buyer's freedom to choose his source of supply, thus hindering competitors' access to the market and tend to foreclose the market. Nevertheless, depending on the circumstances, loyalty rebates may also have pro-competitive effects. Thus, the aim of this paper is to outline key economic principles that should be employed to assess the effects of rebates and to discuss the circumstances under which discount schemes by dominant firms can have exclusionary effects.

Key words: dominant position, competition, loyalty rebates, EU.
Thematic Area: Regulation and Competence Economics.

1 INTRODUÇÃO

Os descontos condicionais são atribuídos pelas empresas aos seus clientes (consumidores finais ou distribuidores) como recompensa de um comportamento específico, normalmente associado às quantidades adquiridas. Efectivamente, os descontos condicionais mais frequentes são concedidos para premiar os clientes quando as suas compras realizadas durante um período de tempo específico ultrapassam um determinado escalão, podendo ser concedidos sobre a totalidade das compras (descontos retroactivos) ou apenas sobre as aquisições efectuadas acima do escalão (descontos progressivos). Os descontos condicionais de natureza retroactiva podem

gerar efeitos de fidelização quando o limiar para a obtenção do desconto representa uma parte substancial das necessidades do comprador no período de referência. Por este motivo, os descontos retroactivos, quando oferecidos por empresas dominantes, podem produzir efeitos efectivos ou potenciais de encerramento semelhantes aos das obrigações de compra exclusiva (Comissão Europeia, 2009).

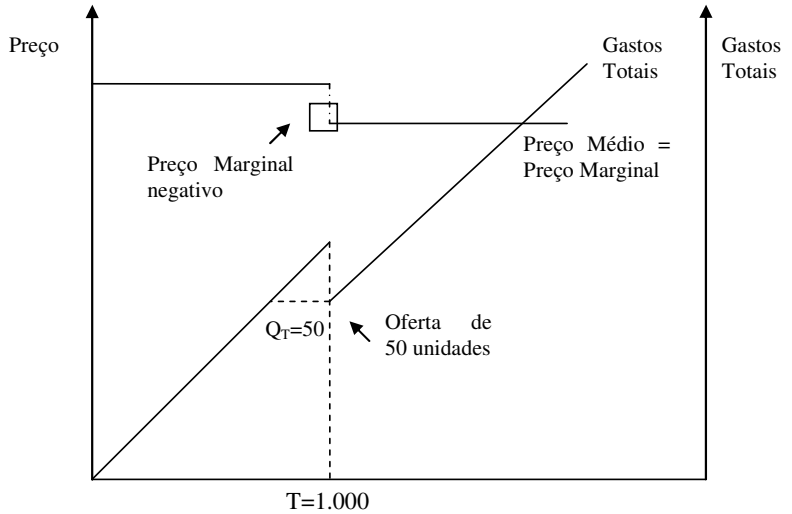
A análise dos efeitos dos descontos de fidelização por empresas em posição dominante registou um crescimento exponencial na literatura após os Acórdãos dos Tribunais europeus relativos aos casos *Michelin II* e *British Airways.* Neste contexto, o presente trabalho propõe-se descrever as características dos descontos retroactivos (ponto 2) sintetizar as principais conclusões desta literatura (ponto 3), apresentar a abordagem de análise da Comissão constante das Orientações relativas à aplicação do artigo 82.º do Tratado CE (ponto 4), com o propósito de desenvolvimento de uma proposta metodológica para este tipo de prática (ponto 5).

2 CARACTERÍSTICAS DOS DESCONTOS RETROACTIVOS

Dado que os descontos com efeitos de fidelização são descontos condicionais definidos em função das compras dos consumidores, importa verificar como distinguir entre os descontos de quantidade e descontos de quantidade com efeitos de fidelização, com efeitos restritivos da concorrência. De acordo com Greenlee e Reitman (2005), pelos menos três características identificam os descontos com efeitos de fidelização, a referir: *(i)* os consumidores apenas se qualificam para o desconto se adquirirem uma parte substancial das suas compras à empresa dominante; *(ii)* o desconto normalmente é aplicado a todas as compras, em alternativa às compras incrementais acima de um determinado limite, *i.e.*, possui uma natureza retroactiva; e *(iii)* o esquema de desconto frequentemente conduz a uma discriminação em termos de preços, na medida em que diferentes consumidores que adquirem a mesma quantidade de um determinado vendedor podem pagar preços distintos.

Nos descontos aplicados de forma retroactiva, *i.e.* a todas as compras desde que um dado nível de compras seja atingido, o preço marginal é negativo quando a quantidade adquirida aumenta uma unidade para atingir o limite a partir do qual o desconto é aplicável. De outro modo, observa-se uma descontinuidade no montante a pagar pelas compras quando o limite é atingido. Para exemplificar as características dos descontos com efeitos de fidelização, considere-se o seguinte exemplo ilustrativo (Frederico, 2005). Uma empresa em posição dominante atribui um desconto de 5% sobre o preço de venda unitário de € 1 se o cliente adquirir pelo menos 1.000 unidades. Se o cliente adquirir 999 unidades paga € 999, mas se adquirir mais uma unidade (1.000) paga apenas € 950, *i.e.*, o mesmo montante como se apenas tivesse adquirido 950 unidades. Esta característica dos sistemas de descontos retroactivos também pode ser interpretada pela oferta de algumas unidades (50) na proximidade do limite do desconto. O preço marginal em torno de 1.000 é, deste modo, negativo (Gráfico *infra*).

Gráfico 1: Características dos desconto retroactivos



3 ANÁLISE ECONÓMICA DOS DESCONTOS COM EFEITOS DE FIDELIZAÇÃO

Na jurisprudência europeia existência de um efeito potencialmente indutor da fidelização é suficiente para a prática ser proibida, dado que uma responsabilidade especial é requerida as empresas em posição dominante, de não dificultarem a entrada /crescimento dos rivais (OFT, 2005). De seguida são sintetizados alguns elementos considerados relevantes na análise dos descontos retroactivo, a referir: *(i)* efeito de sucção; *(ii)* efeitos horizontais e verticais; *(iii)* o papel do tempo; *(iv)* a incerteza.

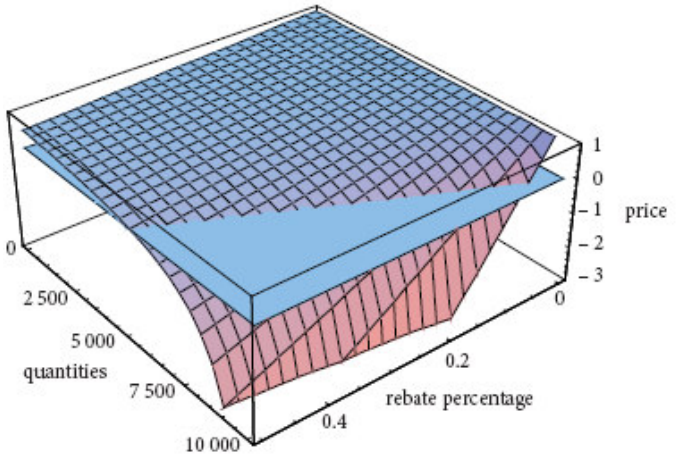
3.1 O efeito de sucção nos descontos retroactivos

Considere-se o sistema de descontos com as seguintes características: a empresa cobra um preço p por qualquer quantidade adquirida até $x^T - 1$ e existe apenas um escalão, em que é aplicado o desconto $\alpha \in (0,1]$, para quantidades adquiridas iguais ou superiores a $x^T \in Z_{++}$, no período T. Do exposto resulta que o preço das $x^T - 1$ unidades é *p,* o preço da última unidade até se atingir x^T é $p - x^T \alpha p$ e o preço das unidades a partir de $x^T + 1$ é $p^T = p(1-\alpha)$. Seja x o montante adquirido à empresa dominante num cenário de concorrência pelo mérito, e seja $x*$ a restante quantidade até ao limite, pelo que $x = x^T - x* \forall\, x* \leq x^T$, com $x, x^T \in Z_{++}$, o preço efectivo $p_c(x*)$ que o concorrente deveria oferecer para tornar indiferente a aquisição de $x*$ à empresa dominante (Maier-Rigaud, 2006) é dado por (1):

$$p_c(x*) = \frac{p(x* - \alpha x^T)}{x*} \qquad (1)$$

A expressão (1) pode ser interpretada como o incentivo que o comprador tem para não mudar as suas aquisições para um concorrente da empresa dominante ('*switching costs'*). O grau do efeito de indução na aquisição de mais quantidades à empresa dominante (relativamente ao volume que seria adquirido na ausência do desconto), *i.e.*, o efeito de fidelização é superior quanto maior foi o limite em termos de quantidade a partir do qual o desconto é aplicado e a percentagem do desconto (Kamann e Bergmann, 2005). O carácter retroactivo do desconto, que permite ao cliente beneficiar de uma redução do preço em todas as aquisições efectuadas durante um dado período de tempo cria o chamado efeito de sucção '*suction effect*'. Deste modo a expressão (1) permite avaliar o efeito de sucção do sistema de descontos (Maier-Rigaud, 2006). Depois de conhecida a quantidade $x*$, o preço que um concorrente necessitaria oferecer pode ser lido directamente a partir da Figura 1.

Figura 1: Efeito de sucção nos descontos retroactivos



Fonte: Maier-Rigaud e Vaigauskaite (2006: 21)

O efeito de sucção é em princípio mais forte na última unidade de produto adquirida antes do limiar para a aplicação do desconto ser excedido. No entanto, o que é relevante para a avaliação do efeito indutor da fidelização não é esta última unidade, mas o efeito de encerramento do desconto em montantes comercialmente viáveis oferecidos pelos concorrentes da empresa em posição dominante (rivais activos no mercado ou concorrentes potenciais).

O efeito de sucção revela que o desconto retroactivo desconto permite à empresa em posição dominante usar a parte inelástica ('*non contestable*') da procura de cada comprador, *i.e.*, o montante que seria, de qualquer forma adquirido (x), como alavancagem para diminuir a preço na parte elástica ('*contestable*') da procura, *i.e.*, o montante relativamente ao qual o comprador poderia encontrar ou até mesmo preferir substitutos[1] (Tom et al., 2000; Maier-Rigaud, 2006; Heimler, 2005; OCDE, 2002). Quanto maior o montante que constitui o limiar e quanto maior a percentagem de desconto, maior o efeito de sucção próximo do limite.

3.2 Efeitos horizontais e verticais dos descontos

A análise da jurisprudência sobre descontos com efeitos de fidelização aplicados por empresas em posição dominante revela duas situações em que estes podem constituir uma exploração abusiva[2] (Kamann e Bergmann, 2005). Um efeito negativo possível dos sistemas de descontos é o encerramento do mercado a concorrentes actuais e concorrentes potenciais, na medida em que os descontos impedem ou limitam a escolha dos clientes entre várias alternativas de oferta dado que o sistema de descontos mantém ou fortalece a posição dominante dificultando a manutenção ou o crescimento da concorrência residual ou potencial (efeito horizontal). Outro efeito negativo possível é a discriminação de preços entre compradores no mercado a jusante (efeito vertical)[3], *i.e.*, a imposição aos parceiros comerciais de diferentes condições para serviços idênticos.

Para além dos efeitos potenciais de encerramento, um desconto com efeito de fidelização, ao criar incentivos para relações de exclusividade ou quase exclusividade, geram una redução da concorrência intra-marcas e prática de preços superiores (Lin, 1990; Faella, 2008; Klemperer, 1987).

3.3 A variável tempo nos descontos

Existe jurisprudência relativamente ao papel do período de referência dos descontos em descontos do tipo retroactivo. Efectivamente, no caso *Michelin* I o Tribunal considerou que "(…) *any system under which discounts are granted according to the quantities sold during a relatively long reference period das the inherent effect, at the end of that period, of increasing pressure on the buyer to reach the purchase figure needed to obtain the discount or to avoid suffering the expected loss for the entire period*"[4]. No caso *Michelin* II o tribunal entendeu que

---

[1] *Vide*, Acordão do Tribunal de Primeira Instância de 30 de Setembro de 2003, relativo ao caso *Manufacture française des pneumatiques Michelin contra Comissão das Comunidades Europeias* (*Michelin II*), processo T-203/01 (*§*162 e §163) e Acordão do Tribunal de Primeira Instância de 17 de Dezembro 2003, relativo ao caso *British Airways plc contra Comissão das Comunidades Europeias* (*British Airways*), Processo T-219/99 (§ 277 e §278).

[2] *Vide*, Acordão do Tribunal de Justiça de 13 de Fevereiro de 1979, relativo ao caso *Hoffmann-La Roche & Co. AG contra Comissão das Comunidades Europeias* (*Hoffmann-La Roche*), Processo 85/76 (§ 90), Acordão do Tribunal de Justiça de 9 de Novembro de 1983, relativo ao caso *N.V. Nederlandsche Banden-Industrie-Michelin contra Comissão das Comunidades Europeias* (*Michelin I*), Processo 322/81 (§ 73), Acordão do Tribunal de Primeira Instância de 7 de Outubro de 1999, relativo ao caso *Irish Sugar plc contra Comissão das Comunidades Europeias* (*Irish Sugar*), Processo T-228/97 (§114), Acordão *Michelin II*, já citado (§ 60) e Acórdão *British Airways,* já citado *(*§ 270).

[3] Num grande número de casos da Comissão e Decisões dos Tribunais Europeus consideraram, para além do efeito de encerramento, a discriminação que resulta do sistema de descontos (Acordão *Hoffmann-La Roche*, já citado, *Irish Sugar*, já citado e *British Airways*, já citado).

[4] *Vide*, Acordão *Michelin I*, já citado (§ 81).

"(…) *it cannot be denied that the loyalty-inducing nature of a system of discounts calculated on total turnover achieved increases in proportion to the length of the reference period*"[5].

Há algum consenso na literatura de que o período de referência dos descontos deve ser contemplado na análise dos seus efeitos. Em particular, é frequentemente referido que quanto maior o período de referência maiores os custos de mudança e como resultado maior o efeito de encerramento (Kallaugher e Sher, 2004; Gyselen, 2004). De acordo com Kallaugher e Sher (2004) "*the first step would be to consider whether the system does generate substantial switching costs. We would suggest that a system with reference periods of three months or less is unlikely to create such costs, as long as the customer has a credible ability to switch to a rival supplier at the end of the period*." O'Donoghue (2003: 23) argumenta que "*the longer the reference period is, the greater the customer's marginal incentives to buy from the dominant firm will be towards the end of the period*." No entanto, este autor identifica dois cenários em que o impacto do período de referência pode ser diferente: *(i)* se a proporção das necessidades dos compradores afectos pelo esquema de descontos com efeitos de fidelização é reduzida, o período de referência dos descontos é considerado irrelevante; *(ii)* por outro lado, se grande parte das necessidades dos compradores são afectadas pelos esquemas de descontos com efeitos de fidelização, o período de referência é um entre vários elementos que deve ser incorporado na análise.

Um posicionamento distinto foi recentemente apresentado por Maier-Rigaud (2005). Este autor analisa o efeito o efeito do período de referência dos descontos na aferição dos custos de mudança '*switching costs*' associados aos descontos. Um desconto retroactivo permite à empresa usar a parte elástica da procura como alavancagem para a redução da parte elástica da procura, e deste modo aumentar os custos de mudança para os compradores. Segundo o autor, enquanto que elementos como *(i)* a percentagem de desconto e o *(ii)* limite a partir do qual o desconto é aplicado influenciam de forma directa o grau em que os descontos com carácter de fidelização geram um encerramento potencial, e por isso devem ser considerados pelas autoridades competentes na avaliação dos efeitos de encerramento do mercado decorrentes destes descontos[6], o tempo é um factor que exerce apenas uma influencia indirecta. O autor argumenta que "*the length of the reference period in rebate schemes is irrelevant for the economic assessment of switching costs and a potential foreclosure effect*" Maier-Rigaud (2005: 273) (…)"*of course elapsed time will often be correlated with quantity, so that the suction effect may be most pronounced towards the end of the period. However, such a correlation in only due to the fact that total quantity is non-decreasing in time."* Maier-Rigaud (2005: 274).

3.4 O papel da incerteza

A incerteza por parte dos clientes da empresa em posição dominante relativamente aos limites a partir dos quais o desconto será aplicado em períodos sucessivos e ao seu valor é um aspecto relevante na avaliação do efeito de lealdade e de encerramento do sistema de descontos (Heimler, 2005). De facto, a incerteza pode induzir a um aumento no efeito fidelizador e consequentemente de encerramento dos descontos, dado que origina estratégias para a minimização do risco por parte dos clientes[7] (Kamann e Bergmann, 2005).

4 A PROPOSTA DE ANÁLISE DA COMISSÃO

Na análise dos efeitos dos descontos retroactivos é importante ter presente que os concorrentes da empresa dominante não podem competir pelas aquisições totais dos compradores. Deste modo, a questão relevante é a de saber se o sistema de descontos dificulta a oferta de montantes comercialmente viáveis aos consumidores individuais por parte dos concorrentes da empresa dominante. Deste modo, a Comissão Europeia (2005) sugere o cálculo

[5] *Vide*, Acordão *Michelin II*, (§ 88).

[6] Os custos de mudança são uma função crescente da percentagem de desconto e do montante anteriormente adquirido pela empresa dominante.

[7] *Vide*, Acordão *Hoffman la Roche* (§ 98 a 100), Acordão *Michelin I* (§ 83 e 84) e Acordão *Michelin II*, já citado (§ 111 e 141).

do preço efectivo que o concorrente deveria pagar para vender uma quantidade relativa à sua quota comercialmente viável, compensando o cliente da perda do desconto. Quanto mais baixo o preço efectivo relativamente aos custos totais médios da empresa dominante, maior o efeito de fidelização gerado pelo desconto. Da comparação entre o preço efectivo e o custo total médio da empresa dominante podem surgir as seguintes situações: *(i)* o preço efectivo é inferior ao custo total médio da empresa dominante $(p_c < ATC)$ - neste caso pode ser difícil ou mesmo impossível para concorrentes eficientes competir com a empresa dominante para esta parte da procura; *(ii)* o preço efectivo é superior ao custo total médio da empresa dominante $(p_c > ATC)$ - pode ser possível aos concorrentes eficientes competir com a empresa dominante no longo prazo igualar a oferta da empresa dominante, mas pode continuar a funcionar como um desincentivo à expansão e à entrada de novos concorrentes. A principal razão para se considerar o custo total médio como um custo *benchmarking* abaixo do qual o sistema de descontos conduz a um preço exclusionário efectivo é que a alavancagem entre a parte da procura 'contestável' e não 'contestável' permitir ao sistema de descontos operar sem um prejuízo de lucros e operar no longo prazo.

A Comissão Europeia (2005) propõe a metodologia bietápica para a avaliação dos efeitos dos descontos que de seguida se descreve. Numa primeira etapa deverá ser calculada a percentagem das aquisições dos consumidores que em média uma nova empresa deve capturar, no mínimo, para que o preço efectivo seja pelo menos tão elevado como o custo total médio da empresa dominante. Esta percentagem é designada por quota de mercado necessária ou '*required share*' (RQS). No caso da quota de mercado necessária diferir significativamente para grupos de consumidores por existirem diferentes descontos, não será calculada a quota média para todos os consumidores, mas a quota média por grupo de consumidores com um desconto semelhante. Deve ser avaliada a importância destes grupos de consumidores para a entrada e a expansão dos concorrentes potenciais e actuais, respectivamente. Considerando o caso simplificado de um desconto uniforme, a *required share* (RQS) é dada por (3):

$$RQS = \frac{\alpha p}{p - ATC} \qquad (3)$$

Em que α representa a percentagem de desconto, p o preço sem o desconto e ATC o custo total médio de produção do produto pela empresa dominante. A '*required share'* pode, assim, ser definida como a percentagem das aquisições que uma nova empresa deve capturar para que o preço efectivo que resulte do desconto iguale o custo total médio da empresa dominante.

Se não for claro, a partir do calculo da '*required share*', que o sistema de descontos gera um efeito de encerramento, pode ser aferida a quota comercialmente viável '*commercially viable share'* que um concorrente eficiente ou novo concorrente espera oferecer e comparar este valor com o da '*required share'*. A '*commercially viable share'* é a percentagem das aquisições de um consumidor que uma empresa nova eficiente pode esperar capturar e determina em que parte da procura os consumidores o efeito dos descontos deve ser aferido. Quando a '*required share'* excede a '*commercially viable share*' o sistema de descontos terá em princípio um efeito de encerramento que reduz a concorrência, dado que o preço efectivo que resulta do sistema de descontos sob esta '*commercially viable share'* será inferior ao custo total médio da empresa dominante. Da comparação entre a CVS e a

RQS três situações podem surgir:

CVS=RQS	=>	Pc=ATC
CVS<RQS	=>	Pc<ATC
CVS>RQS	=>	Pc>ATC

Se as percentagens dos descontos diferem entre consumidores, a '*required share'* pode exceder a '*commercially viable share*' apenas para alguns consumidores. Neste caso deve ser considerado o rácio e a importância destes dois grupos de consumires em consideração na avaliação. O mesmo se aplica se, no caso de descontos com escalões, a '*required share'* se situar acima da '*commercialy viable share'* para alguns descontos e para outros não.

5 EFEITOS PRO-CONCORRENCIAIS DOS DESCONTOS

Quando a análise económica revela que os sistemas de descontos são abusivos, na medida em que geram efeitos de encerramento e/ou são discriminatórios, deve ser de seguida equacionado se o sistema de descontos não é objectivamente justificado (Loewenthal, 2005; Rousseva, 2006).

De facto, vários estudos têm discutido a natureza dos efeitos pro-concorrenciais dos descontos de fidelização, a referir:

(i) Redução do preço no mercado a jusante, que beneficia os consumidores (Kallaugher e Sher, 2004);

(ii) Implementação de discriminação de segundo grau, que pode aumentar as vendas e gerar eficiências (Kolay e Shaffer, 2004; Geradin e Petit, 2006);

(iii) Fornecimento da recompensa adequada aos esforços de promoção dos produtos da empresa dominante no mercado a jusante (Spector, 2005; OFT, 2005) e uma melhor coordenação da cadeia de distribuição (Gavil, 2004);

(iv) Recuperação de investimentos específicos efectuados pelo produtor para satisfazer as necessidades de clientes particulares (Temple, 2002) ou para melhorar a capacidade de oferta dos distribuidores quando os concorrentes utilizam os mesmos canais de distribuição (Spector, 2005);

(v) Resposta ao poder de compra dos compradores (OFT, 2005).

Como uma regra geral, uma empresa não abusa da sua posição dominante à luz do artigo 82.º do Tratado CE se conseguir fornecer uma justificação objectiva para a sua conduta. Efectivamente, a empresa em posição dominante deverá demonstrar o preenchimento das seguintes condições (Comissão Europeia, 2009): *(i)* as eficiências obtidas são um resultado da conduta em análise; *(ii)* a conduta em análise é indispensável para a ocorrência das eficiências; *(iii)* as eficiências beneficiam os consumidores; *(vi)* a concorrência relativamente a uma parte substancial do mercado não é afectada.

6 DISCUSSÃO

Desde o acórdão Hoffman La Roche, os Tribunais Comunitários consideraram que os sistemas de descontos oferecidos por empresas em posição dominante são potencialmente abusivos à luz do artigo 82.º do Tratado se induzirem em efeitos de fidelização, na medida em que podem encerrar o mercado, tal como os contratos de exclusividade.

Casos subsequentes confirmaram a necessidade de apreciação de todas as circunstâncias, nomeadamente os critérios e as modalidades da concessão de descontos e apurar se esses descontos se destinam, através de uma vantagem que não assenta em qualquer prestação económica que a justifique, a suprimir ou restringir a possibilidade de o comprador escolher as suas fontes de abastecimento, impedir o acesso ao mercado dos concorrentes, aplicar a parceiros comerciais condições desiguais para prestações equivalentes ou reforçar a sua posição dominante através de uma concorrência falseada[8] (Waelbroeck, 2005).

No entanto, a aplicação do artigo 82.º a práticas de preços é uma matéria complexa. As empresas dominantes têm uma responsabilidade acrescida de não afectar negativamente a estrutura da concorrência, já enfraquecida. No entanto, estas não estão proibidas de concorrer pelos preços ou adoptar outros comportamentos normais. Em condições particulares, os descontos de fidelização podem ter um efeito exclusionário. No entanto, em muitos mercados, representam uma prática comum e são susceptíveis de gerar ganhos de eficiência.

Com base na revisão da literatura *supra*, propõe-se o estudo do desconto de escalões de volume considerando os seguintes aspectos:

[8] *Vide,* Acórdão *Michelin II*, já citado (§ 62).

0.	**A empresa possui uma posição dominante no mercado relevante?**
I.	**O desconto de escalões de volume tem efeitos discriminatórios?**
II.	**O desconto de escalões de volume tem efeitos de fidelização?**
II.1.	Estudo de características gerais dos descontos que potenciam o seu efeito fidelização
II.1.1.	O desconto é aplicado de forma retroactiva
II.1.2.	Natureza progressiva do desconto, com grandes diferenças entre escalões
II.1.3.	Incerteza relativamente ao valor dos descontos em períodos sucessivos e aos limites a partir do qual são calculados
II.1.4.	Período de tempo alargado de aplicação
II.1.5.	Efeito de sucção
II.1.6.	Ajustamento das quantidades em períodos sucessivos
II.1.7.	A procura dos clientes assume valores próximos do limite do desconto
II.2.	Aplicação da metodologia proposta pela Comissão para a análise dos efeitos de fidelização dos descontos de escalões de volume
II.2.1.	Cálculo do preço efectivo
II.2.2.	Cálculo da quota necessária
III.	**O desconto de escalões de volume tem efeitos pró-concorrenciais?**

REFERÊNCIAS

COMISSÃO EUROPEIA (2005). DG Competition discussion paper on the application of Article 82 of the Treaty to exclusionary abuses. Comissão Europeia.

COMISSÃO EUROPEIA (2009). Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings. Comissão Europeia.

FAELLA, G.2008 (2008), "The antitrust assessment of loyalty discounts and rebates". *Journal of Competition Law and Economics*, Vol. 4, n.º 2, pp. 375-410.

FREDERICO, G. (2005). "When are rebates exclusionary?" *European Competition Law Review*, Vol. 9, pp. 477-480.

GAVIL, A. (2004). "Exclusionary Distribution Strategies by Dominant Firms: Striking a Better Balance", *Antitrust Law Journal*, Vol. 72, n.º 1, pp. 3-82.

GERADIN, D. E PETIT, N. (2006). "Price discrimination under EC Competition Law: another antitrust doctrine in search of limiting principles?" *Journal of Competition Law and Economics*, Vol. 2, n.º 3, pp. 479-531.

GREENLEE, P. E REITMAN, D. (2005) "Distinguishing competitive and exclusionary uses of loyalty discounts." *The Antitrust Bulletin*, Vol. 50, n.º 3, pp. 441-463.

GYSELEN, L. (2003). "Rebates: Competition on the Merits or Exclusionary Practice?", in Claus-Dieter Ehlermann and Isabela Atanasiu (eds.) *European Competition Law Annual 2003: What Is an Abuse of a Dominant Position?*. Oxford/Portland Oregon: Hart Publishing.

HEIMLER, A. (2005). "Below-cost pricing and loyalty-inducing discounts: are they restrictive and, if so, when?" *Competition Policy International*, Vol. 1, nº. 2, pp. 149-172.

KALLAUGHER, J. E SHER, B. (2004). "Rebates Revisited: Anti-competitive effects and exclusionary abuse under article 82." *European Competition Law Review*, Vol. 25, n.º5, pp. 263-285.

KAMAN, K. E BERGMANN, E. (2005). "The granting of rebates by market dominant undertakings under article 82 of the EC Treaty." *European Competition Law Review*, Vol. 83, pp. 103-115.

KAMANN, H.G. E BERGMANN, E. (2005). "The granting of rebates by market dominant undertakings under article 82 of the EC treaty". European Competition Law Review, Vol. 2, pp. 83-89.

KLEMPERER, P. (1987). "The Competitiveness of Markets with Switching Costs". *Rand Journal of Economics*, Vol. 18, n.º 1, pp. 138-150.

KOLAY, S, SHAFFER, G. E ORDOVER, J. A. (2004). "All-Units discounts in retail contracts." *Journal of Economics and Management Strategy*, Vol. 25, n.º 5, pp. 429-459.

LIN, Y. J. (1990). "The Dampening of Competition Effect of Exclusive Dealing". *Journal of Industrial Economics*, Vol. 39, n.º 2, pp. 209-223.

LOEWENTHAL, P.J. (2005). "The defense of objective justification in the application of article 82 EC". *World Competition*, Vol. 28, n.º 4, pp. 455-477.

MAIER-RIGAUD, F. P. (2005). "Switching costs in retroactive rebates: what's time got to do with it?." *European Competition Law Review*, Vol. 25, n.º 5, pp. 272-276.

MAIER-RIGAUD, F. P. E VAIGAUSKAITE, D. (2006). "Prokent/Tomra, a textbook case?", *Competition Newsletter*, Vol. 2, pp. 19-24.

MAIER-RIGAUD, F. P: (2006). "Article 82 rebates: four common fallacies." *European Competition Journal*, Vol. 2, pp. 85-100.

O' DONOGHUE, R. (2003). "Over-Regulating lower prices: time for a rethink on pricing abuses under article 82 EC". *2003 EU Competition Law and Policy Workshop Proceedings*, 44.pag.

OCDE (2002). Loyalty and Fidelity Discounts and Rebates, proceedings of the Roundtable on loyalty or fidelity discounts and rebates. OCDE.

OFT (2005). Selective price cuts and fidelity rebates. Economic Discussion Paper. Office of Fair Trading.

ROUSSEVA, E. (2006). "The concept of 'objective justification' of an abuse of dominant position: can it help to modernize the analysis under article 82 ET?" *Competition Law Review*, Vol. 2, n.º 2, pp. 27-72.

SPECTOR, DAVID (2005) "Loyalty Rebates: An Assessment of Competition Concerns and a Proposed Structured Rule of Reason". *Competition Policy International*, Vol. 1, n.º 2, pp. 89-114.

TOM, W., BALTO, N. AVERITT. (2000). "Anticompetitive aspects of market-share discounts and other incentives to exclusive dealing". *Antitrust Law Journal*, Vol. 67, pp. 615-627.

WAELBROECK, D. (2005). "Michelin II: a per se rule against rebates by dominant companies?" *Journal of Competition Law and Economics*, Vol. 1, n.º 1, pp. 149-171.

ESTUDO DA EFICIÊNCIA DOS AEROPORTOS MUNDIAIS

ANA OLIVEIRA-BROCHADO
E-mail: abrochado@concorrência.pt
Autoridade da Concorrência

Resumo

Recentemente estudos de eficiência dos aeroportos têm recebido uma atenção considerável. Este trabalho aplicada a análise envolutória de dados para aferir o desempenho de 107 aeroportos. As medidas de eficiência são baseadas em quatro medidas de input incluindo os custos operacionais, número de trabalhadores, número de pistas e área do terminal, e quatro medidas de output incluindo as receitas comerciais, o tráfego de passageiros, o número de movimento e o volume de carga transportado. Apesar da eficiência global dos aeroportos ser aceitável (ET=0,76, ETP=0,84, EE=0,90), há espaço de melhoria para alguns aeroportos. As medidas de desempenho foram usadas numa segunda etapa numa regressão Tobit em que variáveis ambientais, estruturais e de gestão foram incluídas. Concluiu-se que os aeroportos com uma maior percentagem de receitas aeronáuticas, uma menor percentagem de tráfego internacional, uma maior percentagem de passageiros de ligação, dominadas por uma companhia principal e com um maior rácio de passageiros/movimentos possuem em termos esperados maior eficiência.

Palavras chave: Eficiência, DEA, Tobit, aeroportos, mundo.
Área Temática: Economia da Regulação e da Concorrência.

Abstract

Recently, considerable attention has been focused on airport operational efficiency. This paper applies data envelopment analysis to assess the performance of 107 airports. Efficiency measures are based on four resource input measures including airport soft costs, number of employees, number of runways and terminal area, and four output measures including commercial revenues, passenger flow, number of movements and total cargo transportation. Although the overall performance of airports is acceptable, (ET=0,76, ETP=0,84, EE=0,90), there is room for improvement for some airports. The performance measures are also used in a second stage Tobit regression in which environmental, structural and managerial variables are included. We conclude that airports with a higher share of aeronautical revenues, lower percentage of international traffic, larger percentage of connecting traffic, dominated by a major carrier and with larger ratio passenger/movement are expected to have higher efficiency.

Key words: Efficiency, DEA, Tobit, airports, world.
Thematic Area: Regulation and Competence Economics.

1. INTRODUÇÃO

Os estudos de eficiência no sector aeroportuário são importantes para vários *stakeholders* do aeroporto (Graham, 2005). De facto, os estudos de desempenho operacional permitem aos reguladores o conhecimento da eficiência ao longo do tempo para a definição das taxas aeroportuárias, auxiliam os governos na tomada de decisões sobre novas infra-estruturas e/ou expansão das existentes, como resposta ao crescimento do número de passageiros (ou do número de passageiros esperado) de um dado aeroporto, fornecem informação aos investidores em aeroportos privados, que preferem realizar aplicações em aeroportos mais eficientes, com um maior lucro esperado e permitem aos operadores dos aeroportos a realização de actividades

de *benchmark* do seu desempenho com os aeroportos de referência (Pels et al., 2001). As companhias aéreas preferem desenvolver as suas operações em aeroportos mais eficientes e os governos locais pretendem que os aeroportos sejam eficientes, por forma a competirem por passageiros e por negócios (Sarkis, 2000).

Consequentemente, tem-se assistido ao crescimento do número de estudos divulgados na literatura sobre características operacionais e de desempenho no sector aeroportuário. Adler e Berechman (2001), Francis et al. (2002), Lin e Hong (2006), Tseng et al. (2008) avaliaram o desempenho operacional dos grandes aeroportos mundiais, Chang e Wang (2002), Park (2003) e Lin e Chen (2004) estudaram o desempenho dos aeroportos localizados na região Ásia Pacífico, Humphreys e Francis (2002) e Pels et al. (2003) avaliaram o desempenho dos aeroportos europeus e Gillen e Lall (1997), Sarkis (2000), Bazargan e Vasigh (2003) e Sarkis e Talluri (2004) estudaram o desempenho dos aeroportos americanos. Os estudos do desempenho operacional têm sido também efectuados considerando os aeroportos de um dado país: Yu (2000, 2005), Yang e Chang (2000) e Hsu (2003), em Taiwain, Fernandes e Pacheco (2002) e Pacheco e Fernandes (2003) no Brasil, Hooper e Hensher (1997) e Abbott e Wu (2002), na Austrália, Martin e Roman (2001) em Espanha e Yoshida e Hiroyoshi (2004), no Japão.

É importante ter presente que um aeroporto pode ser considerado ineficiente por vários motivos (Pels et al, 2003). Desde logo, surge a problemática das indivisibilidades. A construção de uma nova pista pode criar, em muitos casos, um excesso de capacidade no curto prazo. Adicionalmente, pode não ser possível ajustar a nova capacidade à procura adicional, na medida em que o comprimento de uma pista depende de requisitos técnicos e de segurança (peso e velocidade de aterragem ou descolagem das aeronaves que usam a pista). A eficiência dos aeroportos pode igualmente ser influenciada por regulamentações do governo (por exemplo, o limite de horas de operação ou o controle de ruído) e por circunstâncias físicas desfavoráveis (nevoeiro ou vento) nas quais o aeroporto opera, fora do controle dos operadores. A forma de gestão/propriedade dos aeroportos e a forma de determinação das taxas aeroportuárias por parte do reguladores, que podem não ter o incentivo de atingir o óptimo social podem também exercer influência sobre a eficiência dos aeroportos (Domney et al., 2005).

O presente estudo pretende: *(i)* determinar os níveis de eficiência de um conjunto de aeroportos localizados na Europa, Ásia Pacífico e América do Norte através da Análise Envolutória de Dados (DEA); e *(ii)* identificar potenciais factores explicativos das ineficiências identificadas através de modelos de regressão Tobit. De seguida são apresentadas as opções metodológicas do presente trabalho, nomeadamente a amostra considerada, os métodos de análise dos dados utilizados e a descrição das variáveis utilizadas para descrever as operações do aeroporto. Posteriormemte, são apresentados e analisados os resultados e finaliza-se com uma conclusão.

2. METODOLOGIA

2.1. Amostra

No presente estudo são considerados 107 observações para o ano de 2006: 23 aeroportos localizados na Europa, 63 na América do Norte e 21 na região Ásia Pacífico. Os dados foram obtidos a partir do Airport Benchmarking Report (2007).

2.2. Métodos

Neste trabalho adopta-se a abordagem bi-etápica proposta por Gillen e Lall (1997): o DEA é utilizado para obter um indicador da eficiência dos aeroportos, que é considerado como a variável dependente numa regressão Tobit, para a identificação dos factores que influenciam a eficiência.

O conceito de eficiência associa-se à comparação entre os valores dos inputs e outputs observados relativamente a uma determinada actividade e os seus valores óptimos. Esta

eficiência pode ser avaliada pelo ajustamento de uma fronteira de produção eficiente. A fronteira fornece o output máximo (*i.e.* óptimo), dados os inputs (orientação para o output) ou o input mínimo dados os outputs (orientação para os inputs). Qualquer output abaixo do óptimo ou qualquer input acima do óptimo correspondem a operações ineficientes.

O DEA baseia-se em vários problemas de programação linear para derivar a fronteira de produção. Assumindo que o objectivo de cada aeroporto é maximizar os seus inputs, mantendo o nível de output constante, a eficiência técnica pura (ETP) do aeroporto *o* (*o=1,…,n*) pode ser calculada como a solução do seguinte problema de programação linear (Banker et al., 1984):

$$Min\,\theta_o - \varepsilon\left(\sum_{i=1}^{m} s_i^- \sum_{r=1}^{s} s_r^+\right)$$

$$s.a.:$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta_o x_{io},\ i = 1,...,m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro},\ r = 1,...,s$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\theta_o, \lambda_j, s_i^-, s_r^+ \geq 0, \forall\, i \text{ e } r$$

$$\varepsilon \text{ é uma constante não Arquimediana (na ordem de } 10^{-5} \text{ ou } 10^{-6})$$

em que *n* é o número de aeroportos e *m* e *s* são o número de inputs e de outputs, respectivamente. As variáveis x_{ij} e y_{rj} representam os montantes do input *i* consumido e o montante do output *r* produzido pelo aeroporto *j*, respectivamente. Se $ETP = 1$ e todas as folgas de inputs e de outputs, s^- e s^+ forem iguais a zero, então o aeroporto é tecnicamente eficiente. Se $ETP < 1$, o aeroporto é tecnicamente ineficiente. Este modelo, que acomoda rendimentos variáveis à escala é conhecido por modelo BBC. A ETP é obtida pela diferença entre o rácio de outputs e inputs de um dado aeroporto e o rácio do aeroporto com a melhor prática, que são atribuídas a práticas de gestão e não à escala (Abbott e Wu, 2002). Se for retirada a equação $\sum_{j=1}^{n} \lambda_j = 1$, obtém-se uma tecnologia de rendimentos constantes à escala (Charlesÿÿt al., 1978), e a eficiência técnica de um aeroporto *o* é definida por $ET = \lambda_o$. Este modelo é designado na literatura por CCR (iniciais dos autores). A eficiência de escala (*EE*), i.e., a proporção de inputs que pode se reduzida depois da ETP ser eliminada se os ajustamentos à escala forem possíveis é obtida através do quociente $EE = ET / ETP$. Se um aeroporto possuir $EE = 1$, opera a rendimentos constantes à escala. Se $EE < 1$ o aeroporto é ineficiente em termos de escala e pode poupar inputs através de um ajustamento da escala operacional. Se a escala deve ser expandida ou diminuída depende da escala de operação actual. Adicionalmente, se $\sum_{j=1}^{n} \lambda_j < 1$, o aeroporto trabalha com rendimentos crescentes à escala e se $\sum_{j=1}^{n} \lambda_j > 1$, com rendimentos decrescentes à escala.

O DEA fornece uma medida de eficiência (abordagem de Farrell) e não uma explicação da eficiência (abordagem de Leibenstein) (Button e Weyman-Jones, 1994). No entanto, é possível usar os indicadores de eficiência produzidos pelo DEA como uma variável dependente numa análise de regressão Tobit (Tobin, 1958). Deste modo, a segunda etapa da metodologia proposta por Gillen e Lall (1997) procura explicar a eficiência (variável dependente) através de um conjunto de factores exógenos pelo modelo Tobit definido por:

$y_o^* = \beta' x_o + \varepsilon_o$, em que $y_o = y_o^*$ se $y_o > 0$ e $y_o = 0$, caso contrário; $\varepsilon \sim N(0,\sigma^2)$, em que y_o^* é a variável latente para o aeroporto *o* (*o=1,…,n*) e y_o as pontuações de eficiência.

2.3. Inputs, Outputs e Determinantes da Eficiência

Os aeroportos prestam serviços a dois grupos de consumidores: as companhias aéreas e os passageiros, existindo uma procura complementar entre estes dois grupos. De facto, os aeroportos fornecem às companhias aéreas as infra-estruturas necessárias para a aterragem e descolagem de aeronaves e movimentação de passageiros e carga (tais como as pistas e os terminais). Por estes serviços aeronáuticos os aeroportos cobram taxas aeroportuárias, que estão sujeitas a regulação na maioria dos países (Marques e Brochado, 2008) e que constituem a fonte directa de rendimento dos aeroportos. Os aeroportos prestam aos passageiros serviços não aeronáuticos, tais como o parqueamento e o retalho (alimentar ou não). Destes serviços os aeroportos obtêm receitas comerciais, tais como rendas de concessões e receitas de parqueamento, que constituem uma fonte indirecta de receita dos aeroportos.

Assim, com o objectivo de melhor aproximar a análise às operações dos aeroportos e tendo por base a revisão da literatura efectuada, são considerados os seguintes outputs relativos às actividades aeronáuticas do aeroporto: número de passageiros, volume de carga e número de movimentos. Quanto às actividades não aeronáuticas, considerou-se um índice de receitas comerciais (deflacionado pela Paridade do Poder de Compra).

Do lado dos inputs consideram-se quatro categorias variáveis: número de funcionários que trabalham directamente para o operador do aeroporto, índice de gastos operacionais não relacionados com o trabalho e o capital (deflacionado pela Paridade do Poder de Compra), número de pistas e área do terminal. As duas últimas variáveis propõe-se medir o capital do aeroporto em termos físicos. A inclusão de um indicador para os gastos operacionais é importante, na medida em que vários aeroportos efectuam a subcontratação do factor trabalho.

Os níveis de eficiência dos aeroportos são afectados por um grande número de factores, incluindo factores que não são controlados pelos operadores. Neste trabalho pretende-se analisar o impacto nos níveis de eficiência obtidos de um conjunto de variáveis, a referir: *(i)* peso dos passageiros internacionais; *(ii)* peso dos passageiros em trânsito; *(iii)* peso da empresa dominante no aeroporto; *(iv)* rácio passageiros/movimento; *(v)* peso das receitas comerciais. A tabela que se segue apresenta as estatísticas descritivas das variáveis consideradas.

Tabela 1. Estatísticas descritivas dos Inputs, Outputs e outras variáveis

	Variável	Média	Desvio-padrão	Mínimo	Máximo
Outputs	N.º passageiros (10^3)	19.007	15.842	997	83.189
	N.º movimentos	251.691	182.180	7.784	992.427
	Volume de carga (ton)	442.212	653.855	4.351	3.544.575
	Receitas não aeronauticas	1,00	0,84	0,02	5,12
Inputs	N.º de funcionários	887	1.178	115	8.872
	Gastos operacionais	1,00	2,07	0,01	13,77
	N.º de pistas	3	1,30	1	7
	Dimensão do terminal (m^2)	176.520	157.701	7.500	634.100
Outras variáveis	Peso dos passageiros internacionais	0,33	0,35	0,00	1,00
	Peso dos passageiros em trânsito	0,24	0,20	0,00	0,77
	Peso da empresa dominante	0,46	0,18	0,19	0,87
	Rácio passageiros/movimento	76,50	33,85	22,00	168,00
	Peso das receitas comerciais	0,49	0,12	0,18	0,76

Os passageiros internacionais requerem mais serviços e recursos do que o tráfego doméstico, pelo que se espera que exerçam um efeito negativo na eficiência (Gillen e Lall, 1997; Oum e Yu, 2004). No entanto, os passageiros internacionais tendem a gerar mais receitas de concessão do que os passageiros domésticos numa base per capita, pelo que o efeito do peso

do tráfego internacional na eficiência dos aeroportos dependerá do balanceamento entre estes dois factores.

Relativamente aos passageiros em trânsito, espera-se que exerçam um efeito positivo nos níveis de eficiência, na medida em que os passageiros de ligação, frequentes nos grandes *hubs*, normalmente não requerem o mesmo número de serviços que os passageiros com voos directos.

Espera-se que a dominância de uma empresa num aeroporto permita uma melhor coordenação entre o aeroporto e a empresa dominante e consequentemente gere uma maior eficiência do que a coexistência de várias empresas concorrentes num aeroporto. Outra explicação possível é o facto das companhias aéreas escolherem os aeroportos mais eficientes como *hubs* (Gillen e Lall, 1997; Sarkis, 2000).

Relativamente ao rácio passageiros/movimento, trata-se de um indicador do nível de utilização do aeroporto. Intuitivamente, valores superiores para este rácio devem corresponder a aeroportos mais eficientes. No entanto, a interpretação deve ser tomada de forma cuidada, na medida em que uma utilização excessiva pode conduzir a outputs indesejáveis, tais como o congestionamento e os atrasos.

O peso das receitas comerciais é utilizado como um indicador do grau de diversificação das actividades do aeroporto. De acordo com Oum et al. (2003), os aeroportos que se concentram nas actividades aeronáuticas tendem a ter menores eficiências, na medida em que uma estratégia de diversificação através do aumento das receitas de actividades não aeronáuticas permite explorar a complementaridade da procura entre entre serviços aeronáuticos por parte das companhias aéreas e de serviços comerciais por parte dos passageiros.

3. APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Nesta secção analisam-se as eficiências técnicas e de escala médias e identificam-se os aeroportos que pertencem à fronteira. Adicionalmente, estuda-se o impacto na eficiência de um conjunto de características da operação do aeroporto.

3.1. Níveis de eficiência médios

A Tabela que se segue apresenta os resultados médios dos modelos CCR (ET) e BBC (ETP) para a amostra global e por região (América do Norte, Ásia Pacífico e Europa), bem como a eficiência de escala. A média da eficiência técnica (ET) foi 0,76 em 2006. De outro modo, observa-se uma ineficiência de 0,24 que indica a possibilidade de redução dos inputs, em média, de 24%, mantendo os mesmos níveis de output. A eficiência técnica pura (ETP) registou valores de 0,84, e a eficiência de escala de 0,90. Estes resultados revelam que, em termos médios, as ineficiências dos aeroportos considerados se devem principalmente a ineficiências puras, relativamente a ineficiências de escala. Tal sugere que os operadores devem numa primeira etapa remover eficiências técnicas e só depois eficiências de escala. Conclui-se ainda que, em média, os recursos e as capacidades organizacionais estão a ser bem aproveitadas pela maioria dos aeroportos.

Os valores médios da eficiência técnica são, de acordo com o teste Kruskal-Wallis, estatisticamente diferentes nas três regiões consideradas $\left(\chi^2 = 6{,}975; gl = 2; p = 0{,}002\right)$. De facto, observam-se valores médios da ET superiores para aeroportos localizados na região Ásia Pacífico (0,83), relativamente à América do Norte (0,77) e Europa (0,67). Os valores médios produzidos pelo modelo BBC são também distintos $\left(\chi^2 = 12{,}415; gl = 2; p = 0{,}031\right)$, de 0,95 na região Ásia Pacífico, 0,84 na América do Norte e 0,75 na Europa. Estes resultados são consistentes com os obtidos por Tseng et al. (2008), para uma amostra de 20 grandes aeroportos mundiais no período de 2001 a 2005.

As diferenças nos valores médios fornecidos pelos modelos CCR e BBC apontam para a existência de rendimentos à escala distintos nos dados, *i.e.*, a relação entre os inputs e os outputs depende da magnitude dos respectivos valores.

Tabela 2. Valores médios da eficiência técnica, eficiência técnica pura e eficiência de escala

		América do Norte	Ásia Pacífico	Europa	Global
ET	Média	0,77	0,83	0,67	0,76
	Desvio-padrão	0,19	0,16	0,23	0,20
	Mínimo	0,27	0,53	0,27	0,27
	Máximo	1,00	1,00	1,00	1,00
ETP	Média	0,84	0,95	0,75	0,84
	Desvio-padrão	0,15	0,10	0,22	0,17
	Mínimo	0,44	0,65	0,36	0,36
	Máximo	1,00	1,00	1,00	1,00
EE	Média	0,92	0,88	0,90	0,90
	Desvio-padrão	0,12	0,15	0,15	0,13
	Mínimo	0,39	0,53	0,40	0,39
	Máximo	1,00	1,00	1,00	1,00

De acordo com o modelo de rendimentos constantes à escala em 2006 os aeroportos que pertenciam à fronteira eram os seguintes: MAS, EDI, GVA, LGW, LHR (Europa), ATL, DTW, EWR, JFK, LAS, MDW, MEM, OAK, RDU, SAN, SDF, SNA e YUL (América do Norte) e AKL, BKK, BNE, SEL, SYD, TPE (Ásia Pacífico). Observa-se que 25 aeroportos operavam com rendimentos decrescentes à escala, 58 com rendimentos crescentes e 24 constantes à escala no mesmo ano. Cerca de 14 aeroportos registaram níveis de eficiência inferiores a 0,50 e 25 inferiores a 0,60, de acordo com o modelo CCR, e 7 e 25 de acordo com o modelo BBC, respectivamente.

3.2. Determinantes da eficiência

Os resultados da aplicação do modelo Tobit considerando como variável dependente as ET revelaram que as variáveis explicativas consideradas são estatisticamente na explicação da eficiência. Tal como esperado, a estratégia e diversificação, medida pelo peso das receitas não aeronáuticas, gera um efeito positivo na eficiência $\left(\hat{\beta} = 0{,}57,\, p < 0{,}01\right)$. Deste modo conclui-se que uma gestão proactiva de diversificação a base de receitas do aeroporto expandindo as suas actividades comerciais e outras oportunidades de negócio não aeronáuticas, gera um efeito positivo nos níveis de eficiência. Trata-se de um resultado interessante, na medida em que esta variável é controlada pelo operador aeroportuário. Relativamente às restantes variáveis conclui-se que os passageiros internacionais geram um efeito negativo na eficiência $\left(\hat{\beta} = -0{,}03,\, p < 0{,}01\right)$ e as restantes variáveis – peso dos passageiros em trânsito $\left(\hat{\beta} = 0{,}01,\, p < 0{,}01\right)$, peso da empresa de maior dimensão $\left(\hat{\beta} = 0{,}06,\, p < 0{,}1\right)$e rácio passageiros movimento $\left(\hat{\beta} = 0{,}12,\, p < 0{,}01\right)$ exercem um efeito positivo.

4. CONCLUSÃO

Os estudos de avaliação da eficiência no sector aeroportuário fornecem informação importante para vários *stakeholders* do aeroporto. Neste trabalho efectuou-se um estudo da eficiência de 107 aeroportos mundiais utilizando a abordagem bietápica de Gillen e Lall (1997). Numa primeira fase, foram ajustados modelos DEA orientados para a minimização dos inputs considerando rendimentos constantes (modelo CCR) e rendimentos variáveis (BBC) à escala. Usando os dois modelos, foi possível verificar se existem rendimentos variáveis à escala na amostra considerada (diferenças dos valores médios de eficiência). Em geral concluiu-se que o desempenho global dos aeroportos é aceitável e que a fronteira é constituída por aeroportos de

todo o mundo (América do Norte, Europa e Ásia Pacífico). Não obstante, vários aeroportos em todo o mundo continuam ineficientes.

Numa segunda etapa procurou-se determinar a influencia de um conjunto de variáveis na eficiência, tende-se concluído que o peso dos passageiros internacionais gera um efeito negativo na eficiência e o peso dos passageiros em transito, pelo da empresa dominante, rácio movimentos/passageiro e peso das receitas comerciais exerce um efeito positivo. Deste modo conclui-se que há factores fora do controle do aeroporto no curto e médio prazo que afectam a sua eficiência. Dado que se concluiu que há características dos aeroportos que afectam a sua eficiência, pode ser útil em estudos posterior reagrupar as eficiências.

Em termos futuros pretende-se considerar uma amostra de dados em painel por forma a analisar a eficiência numa perspectiva dinâmica, e incluir outputs indesejáveis, como o congestionamento, ruído e atrasos o modelo de eficiência e estudar o efeito do poder de mercado do aeroporto na sua eficiência.

REFERÊNCIAS

ABBOTT, M. E WU, S. (2002). "Total factor productivity and efficiency of Australian Airports", *The Australian Economics Review*, Vol. 35, n.º 3, pp. 244-260.

ADLER, N. E BERECHMAN, J. (2001). "Measuring airport quality from the airlines viewpoint: an application of data envelopment analysis", *Transport Policy*, Vol. 8, n.º 3, pp. 171-81.

BANKER, R., CHARNES, A. E COOPER, W. (1984). "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis", *Management Science*, Vol. 30, n.º 9, pp. 1078-1092.

BAZARGAN, M. E VASIGH, B. (2003). "Size versus efficiency: a case study of US commercial airports", *Journal of Air Transport Management*, Vol. 9, n.º 3, pp. 187-193.

BUTTON, K.J. E T.G. WEYMAN-JONES (1994). "X-inefficiency and technical efficiency", *Public Choice,* Vol. 80, n.º 1-2, pp. 83-104.

CHANG, Y-H. E WANG, D-C. (2002), "A study on the aviation performance evaluation – using the international airports in Asian Pacific as an example", *Civil Aviation Quarterly*, Vol. 4, n.º 1, pp. 89-107.

CHARNES, A., COOPER, W. E RHODES, E. (1978). "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research*, Vol. 2, n.º 6, pp. 429-444.

DOMNEY, M,D., HEATHER, I. M., CHEN, E. (2005). "Natural monopoly privatization under different regulatory regimes. A comparison of New Zealand and Australian airports", *International Journal of Public Sector Management*, Vol. 18, n.º 3, pp. 274-292.

FERNANDES, E. E PACHECO, R. R. (2002). "Efficient use of airport capacity", *Transportation Research* Part A, Vol. 36, n.º 3, pp. 225-238.

FRANCIS, G., HUMPHREYS, I. E FRY, J. (2002). "The benchmarking of airport performance". *Journal of air management*, Vol. 8, n.º 4, pp. 239-247.

GILLEN, D. E LALL, A. (1997). "Developing measures of airport productivity and performance: an application of data envelopment analysis". *Transportation Research* Part E: Logistics and Transportation Review, Vol. 33, n.º 4, pp. 261-273.

GRAHAM, A. (2005). "Airport benchmarking: a review of the current situation", *Benchmarking: An international Journal*, Vol. 12, n.º 2, pp. 99-111.

HOOPER, P. G. E HENSHER, D. A. (1997). "Measuring factor productivity of airports: an índex number approach", *Transportation Research* Part E: Logistics and Transportation Review, Vol. 33, n.º 4, pp. 249-259.

HSU, R-M (2003). "An evaluation of the operational performances of Asian-Pacific regional operations center – an empirical study on the cargo zone in Taiyuan airport", *Journal of the Border Police*, Vol. 2, n.º 1, pp. 217-237.

HUMPHREYS, I. E FRANCIS, G. (2002). "Policy issues and planning of UK regional airports". *Journal of Air Transport Management*, Vol. 10, n.º 4, pp. 249-258.

LIN, L. C. E HONG, C. H. (2006). "Operational performance evaluation in international major airports: an application of data envelopment analysis", *Journal of Air Transport Management*, Vol. 12, n.º 6, pp. 342-51.

LIN, L-C. E CHEN, Y-J. (2004). "The application of DEA of operational performance efficiency in the international airports in Asian-Pacific", *Journal of the Chinese Institute of Transportation*, Vol. 16, n.º 2, pp. 155-144.

MARQUES, R. C. E BROCHADO, A. O. (2008). "Airport regulation in Europe: Is there need for a European Observatory?", *Transport Policy*, Vol. 15, n.º 3, pp. 163-172.

MARTIN, J. C. E ROMAN, C. (2001). "An application of DEA to measure the efficiency of Spanish airports prior to privatization", *Journal of air transport management*, Vol. 7, n.º 3, pp. 149-157.

OUM, T. H., YU, C. E FU, X. (2003). "A comparative analysis of productivity performance of the world's major airports: a summary report of the ATRS global airport benchmarking reserahc report – 2002". *Journal of transport management*, Vol. 9, n.º 5, pp. 285-297.

PACHECO, R. R. E FERNANDES, E. (2003). "Managerial efficiency of Brazilian Airports", *Transportation Research Part A*, Vol. 37, n.º 8, pp. 667-80.

PARK, Y. (2003). "An analysis of competitive strength of Asian major airports". *Journal of Air Transport Management*, Vol. 9, n.º 6, pp. 353-60.

PELS, E., NIKAMP, P. E RIETVELD, P. (2003). "Inefficiencies and scale economies of European airport operations". *Transportation Research* Part E, Vol. 39, n.º 5, pp. 341-361.

SARKIS, J. (2000). "An analysis of the operational efficiency of major airports in the United States", *Journal of Operations Management*, Vol. 18, n.º 3, pp. 335-351.

SARKIS, J. E TALLURI, S. (2004). "Performance based clustering for benchmarking of US airports", *Transportation Research*, Part A, Vol. 38, n.º 5, pp. 329-346.

TOBIN, J. (1958). Estimation of relationships for limited dependent variables, *Econometrica*, Vol. 26, n.º 1, pp. 24-36.

TSENG, K-J., HO, J-F. E LIU, Y-J. (2008). "A study on the performance evaluation of major international airports in the world". *Journal of modeling in management*, Vol. 3, n.º 1, pp. 71-81.

TSENG, K-J., HO, JOW-FREI E LIU, YUAN-JING (2008). "A study on the performance evaluation of major international airports in the World". *Journal of Modeling in Management*, Vol. 3, n.º 1, pp. 71-81.

YANG, J-R. E CHANG, Y-J. (2000). "An analysis of the service quality in Taiwan's airports: from the airliners' perspective", *Civil Aviation Journal Quartely*, Vol. 2, n.º 1, pp. 1-28.

YOSHIDA, Y. E HIROYOSHI, F. (2004). "Japanese-airport benchmarking with the DEA and endogenous-weight TFP methods: testing the criticism of overinvestment in Japanese regional airports", *Transportation Research* Part E, Vol. 40, n.º 6, pp. 533-546.

YU, M-M. (2000). "An analysis of the changes in the key factors and productivity of Taiyuan airport", *Civil Aviation Journal Quarterly*, Vol. 2, n.º 1, pp. 97-123.

YU, M-M. (2005). "The influence of unplanned output on the efficiency and output loss of Taiwain's airports", *Journal of Management*, Vol. 22, n.º 2, pp. 241-259.

LA ANTIGÜEDAD Y LAS DIFERENCIAS DE ESFUERZO ENTRE TRABAJADORES DE DISTINTAS ZONAS GEOGRÁFICAS: UN ESTUDIO DE LOS ACCIDENTES DE TRABAJO

ALFONSO MORAL DE BLAS
E-mail: amoral@eco.uva.es
ÁNGEL MARTÍN ROMÁN
E-mail: angellm@eco.uva.es
Facultad de Ciencias Sociales Jurídicas y de la Comunicación de Segovia
Universidad de Valladolid
JUAN CARLOS RODRIGUEZ CABALLERO
E-mail: jcrc@eco.uva.es
Facultad de Ciencias Económicas y Empresariales de Valladolid
Universidad de Valladolid

Resumen

El objetivo de este trabajo consiste en analizar las diferencias en el reporte de los accidentes conocidos dentro de la literatura como de difícil diagnóstico entre la población nacional e inmigrante. El fin de utilizar este tipo de accidentes es que en ellos puede parecer más verosímil que se produzcan episodios de absentismo laboral encubierto. Así se podría obtener una medida del esfuerzo relativo ejercido por los trabajadores inmigrantes en relación con los nacionales. Desde un punto de vista metodológico se realiza una estimación probit para obtener las principales conclusiones. Partiendo de esta estimación se descompone en dos partes la diferencia observada en la incidencia de este tipo de accidentes, una debida a las distintas características de cada grupo y otra provocada por la diferente repercusión de esas características. Con este fin se aplica una generalización de la descomposición de Oaxaca-Blinder para modelos no lineales.

Palabras Clave: Inmigración, riesgo moral, absentismo, descomposición no lineal.
Área Temática: Economía de la Regulación y de la Competencia.

Abstract

The objective of this article is to investigate the differences in the report of hard-to-diagnose injuries between Spanish and immigrant workers. We use this kind of accidents because is more common to find hidden labour absenteeism episodes in them. So it could be obtained a measure of the relative effort exercised by the immigrant workers in comparison with the national ones. From a methodological point of view we perform a probit estimation to obtain the main conclusions. Starting from these results we decompose the observed difference in the incidence of this type of accidents in two parts, one part justified by the characteristics of each group, and other part which is due to the different yield of those characteristics. Because of that we apply a generalization of the Oaxaca-Blinder decomposition for non linear models.

Thematic Area: Regulation and Competence Economics.

1. INTRODUCCIÓN

La inmigración se ha manifestado, sobre todo durante los últimos 10 años, como uno de los factores más determinante a la hora de explicar la historia reciente del mercado de trabajo español. Este resultado se sustenta en la importancia que ha adquirido la participación de la inmigración en las variables laborales básicas. Actualmente representan en torno al 15% de la población activa y ocupada y casi el 25% de la población desempleada.

El análisis de los efectos de las migraciones sobre el mercado de trabajo ha generado abundante literatura económica en los años recientes. Los trabajos se han centrado tanto en estudiar el impacto sobre las oportunidades de empleo y salarios de los trabajadores nativos como en el análisis de la asimilación de los inmigrantes en el mercado de trabajo del país de destino. Excelentes revisiones de esta literatura a nivel internacional pueden encontrarse en los trabajos de Borjas (1999), Friedberg y Hunt, (1995), Lalonde y Topel (1997) y Card (2005).

En España los estudios sobre las consecuencias de la inmigración en el mercado de trabajo han proliferado en consonancia con el fuerte crecimiento que han experimentado en los últimos años los flujos de inmigrantes. Una rigurosa relación de estos estudios aparece en el trabajo de Simón et al. (2007). Aquí encontramos trabajos que ofrecen una panorámica de la situación de la inmigración en el mercado de trabajo español (Izquierdo y Jimeno, 2005; Consejo Económico y Social, 2004; Garrido y Toharia, 2003 y Carrasco et al., 2003) y trabajos que se centran en cuestiones más concretas como la ocupación, el desempleo y las diferencias salariales. Así Dolado et al. (1997) encuentra un efecto positivo de la regularización de principios de los 90 sobre la tasa de paro y el salario. Amuedo-Dorantes y De la Rica (2005) constatan la existencia de una emigración localizada hacia aquellas zonas con mayores tasas de empleo. Carrasco et al. (2008) encuentran un efecto muy leve de la inmigración sobre el empleo de los nativos, efecto que viene determinado en parte por el nivel de asimilación de los inmigrantes.

Existe además, para el caso español, otro grupo de trabajos que analiza de forma específica la asimilación de los inmigrantes, esto es, los resultados laborales de los inmigrantes en relación con los de los trabajadores nativos. Así, Iglesias y Llorente (2006) observan, en base a la comparación en el tiempo de una batería de indicadores laborales asociados a diferentes grupos de trabajadores, que no existe un proceso de asimilación claro de las pautas laborales de los inmigrantes hacia los parámetros que definen el comportamiento laboral de la población nacional. Amuedo-Dorantes y De la Rica (2006) confirman que los inmigrantes están segregados en ocupaciones de bajos salarios a la vez que dan cuenta de la mayor movilidad ocupacional de los inmigrantes a medida que aumenta el tiempo de residencia en nuestro país. Sanromá et al. (2006) analizan la asimilación salarial de los inmigrantes en base a la estrecha relación existente entre sobreeducación y diferencia salarial; y observan en el colectivo de inmigrantes una mayor incidencia de la sobreeducación responsable de una parte importante de la diferencia salarial entre nativos e inmigrantes. Para terminar, el propio trabajo de Simón et al. (2007) investiga directamente las diferencias salariales y encuentra que estas se explican en gran medida por la diferencia en las características y por la localización de los inmigrantes en ocupaciones y establecimientos diferentes.

Un aspecto que no ha sido muy tratado en la literatura es la relación entre la mano de obra inmigrante y los problemas de seguridad y salud laboral. Algunos trabajos empíricos para Estados Unidos encuentran un mayor riesgo de accidente laboral en los trabajadores de color y del sur. (Worrall y Butler, 1983; Butler y Worrall, 1983; Bartel y Thomas, 1985; y Graham y Shakow, 1990, Richardson et al., 2004). Por su parte Richardson y Loh (2004) encuentran para el período 1996-2001 un aumento del riesgo de accidente laboral mortal entre los trabajadores inmigrantes en Estados Unidos (especialmente entre los procedentes de America Latina) contraponiéndose al descenso del riesgo observado en los trabajadores autóctonos. De otro lado, Bauer et al (1999), en un trabajo realizado para Alemania, encuentra que la regulación alemana sobre seguridad laboral se ha mostrado efectiva en la reducción de accidentes leves, a la vez que no encuentra diferencias en la procedencia de los accidentados. Finalmente, el trabajo de Ahonen y Benavides (2008) para datos españoles encuentra un menor riesgo, tanto de sufrir accidentes mortales como no mortales, en los trabajadores inmigrantes[1].

[1] En una versión previa de este trabajo los autores encontraron que el riesgo de sufrir accidentes de trabajo era mayor para los trabajadores inmigrantes que para los nativos. Esta diferencia se producía como consecuencia de una

El resto del trabajo se organiza de la siguiente manera. El apartado dos se dedica al estudio pormenorizado de los datos empleados en el análisis empírico. En el apartado tres se detalla la metodología econométrica utilizada. Los resultados obtenidos se analizan detalladamente en el punto cuatro. Finalmente, en el punto cinco se resumen las principales conclusiones obtenidas.

2. DATOS.

El objetivo de este trabajo es analizar los factores determinantes del porcentaje de accidentes de difícil diagnóstico que sufren diferentes grupos poblacionales. Por lo tanto los datos estadísticos que más información proporcionan en relación a este fenómeno provienen de la Estadística de Accidentes de Trabajo (EAT), que elabora anualmente el Ministerio de Trabajo y Asuntos Sociales a partir de los partes de accidente comunicados por las empresas mediante el sistema informático Delt@[2]. Estos datos administrativos incluyen el total de accidentes que se han producido en un mismo año y que han supuesto al menos un día de baja para el accidentado y contienen información sobre las características personales del trabajador accidentado, sobre el tipo y grado de la lesión y sobre las características de la empresa en la que estaba empleado el trabajador en el momento del accidente.

Desde 2003 esta base aporta información sobre la nacionalidad del trabajador accidentado, sin embargo es desde 2004 cuando recibe un tratamiento informático más adecuado que corrige la mala consignación del código de nacionalidad que sufría más del 15% de los accidentes. Para este trabajo se utilizarán los datos referidos al año 2005.

Teniendo en cuenta los fundamentos del modelo que proponemos, se va a prestar especial atención a aquellos accidentes que han producido una lesión en la espalda, esguinces y torceduras (en la literatura denominadas *lesiones de difícil diagnóstico*), por que son los más susceptibles de verse relacionados con un comportamiento discrecional del trabajador. La delimitación de este tipo de lesiones a partir de la estadística de accidentes de trabajo puede realizarse en base a dos variables o una combinación de las mismas (la parte del cuerpo lesionada y el tipo de lesión). En nuestro caso se ha atendido a la clasificación de la descripción de la lesión, considerando las "Dislocaciones, esguinces o distensiones" (códigos 30, 31, 32 y 39) como lesiones de “difícil diagnóstico”. Estas lesiones son las más numerosas, constituyendo más del 40% del total de lesiones por accidente laboral. Otra variable que recibirá especial atención es la referida a la nacionalidad del trabajador accidentado. La estadística de accidentes de trabajo consigna esta información a través de la variable país. Para este trabajo las nacionalidades se van a agrupar en diferentes zonas geográficas, así tendremos por un lado España, y por otro Äfrica, Asia, América Latina, Europa del Este y Países Desarrollados[3]

El cuadro 1 muestra una primera intuición de los resultados previsibles. Se aprecia que más del 90% de los accidentes afectan a trabajadores de nacionalidad española y que son también los trabajadores españoles los que reportan un mayor porcentaje de accidentes de difícil diagnóstico, concretamente el 44%. En cuanto a los accidentados inmigrantes, en torno al 40% procede de América Latina, cerca de un 30% son africanos y el resto procede, sobre todo, de Europa del Este y de Países desarrollados y en menor medida de Asia. Por lo que se refiere al porcentaje de accidentes de difícil diagnóstico que estos trabajadores declaran, este varía entre el 38% de los latinos y los procedentes de países desarrollados y el 30% de los accidentados que proceden de Asia.

Se comprueba, por lo tanto, la existencia de unas diferencias que pueden indicar la necesidad de un tratamiento individualizado. Sin embargo estas diferencias pueden estar ocultando otro tipo de información, como puede ser el hecho de que los trabajadores nacionales e inmigrantes no realizan las mismas actividades o no ocupan los mismos puestos. Para solucionar este problema en el cuadro 1 se procede a desagregar los accidentes por 10 ramas de

asignación errónea de la nacionalidad a los casos en los que ésta, o bien no constaba, o bien no seguía la clasificación oficial, en la base de datos utilizada. Véase Ahonen y Benavides (2006).

[2] En algunos estudios se ha utilizado el módulo de accidentes y enfermedades laborales contenido en la Encuesta de Población Activa (EPA) pero esta encuesta sólo se realizó en el año 1999.

[3] Dentro del grupo de países desarrollados se incluirán todos los pertenecientes a Europa Occidental junto con Austria, Oceanía y América del Norte (Estados Unidos y Canadá). Por su parte en América Latina se incluyen todos los países de América del Sur y Central.

actividad y por 10 códigos de ocupación a la vez que se analizan las diferentes nacionalidades. Se puede comprobar como, con independencia de la rama de actividad considerada o del código de ocupación, y siempre que el número de datos es suficiente, se sigue manteniendo que el porcentaje de accidentes de difícil diagnóstico que reportan los trabajadores nacionales es superior al de los inmigrantes.

Cuadro1. Proporción de accidentes de difícil diagnóstico y número total de accidentes según nacionalidad del accidentado y tipo de trabajo.

		Desarr.	Europa E.	A. Latina	Asia	África	España	Total
C.N.A.E.	Agricultura y pesca	0,41 350	0,36 1066	0,34 1398	0,26 81	0,32 2210	0,42 31378	0,41 36483
	Energía y agua	0,27 70	0,23 146	0,32 94	0,75 4	0,31 51	0,40 9945	0,39 10310
	Industria extractiva	0,42 172	0,36 337	0,39 793	0,33 117	0,37 1027	0,44 36557	0,43 39003
	Industria manufact.	0,31 577	0,27 1108	0,31 2223	0,35 187	0,25 2130	0,38 105352	0,38 111577
	Otras manufactur.	0,36 429	0,31 1048	0,37 2464	0,29 296	0,30 2303	0,41 82705	0,40 89245
	Construcción	0,33 3294	0,34 4079	0,35 12260	0,34 583	0,33 10817	0,40 227426	0,40 258459
	Comercio y reparaciones	0,39 2150	0,37 1574	0,38 8777	0,25 852	0,31 4189	0,45 174535	0,44 192077
	Tte. Y comunicación	0,42 564	0,40 553	0,42 1615	0,48 61	0,41 600	0,51 55862	0,50 59255
	Instituciones fras	0,46 695	0,36 681	0,42 4076	0,28 355	0,34 2545	0,50 80738	0,49 89090
	Otros servicios	0,50 600	0,47 295	0,53 2367	0,48 82	0,39 998	0,54 122164	0,53 126506
C.N.O.	Fuerzas Armadas	0,55 31	0,40 35	0,36 138	0,63 8	0,33 126	0,46 4753	0,46 5091
	Dirección	0,47 30	0,56 9	0,41 39	1,00 1	0,21 19	0,49 3409	0,49 3507
	Tcos. Cientificos	0,46 114	0,37 27	0,52 100	0,42 12	0,41 49	0,55 15877	0,55 16179
	Tcos. Apoyo	0,53 340	0,59 83	0,55 369	0,64 28	0,56 90	0,57 26321	0,57 27231
	Empleados admvos.	0,49 330	0,50 72	0,43 760	0,21 19	0,41 187	0,55 39450	0,55 40818
	Trabajadores servicios	0,38 1516	0,39 812	0,40 6010	0,23 561	0,30 2063	0,48 127257	0,48 138219
	Cualificados agr. y pesca	0,34 158	0,34 353	0,33 425	0,26 27	0,33 616	0,42 18279	0,41 19858
	Artesanos y cualificados	0,34 3198	0,32 4336	0,35 11333	0,30 707	0,31 9489	0,40 307104	0,39 336167
	Operadores maquinaria	0,38 917	0,36 1189	0,38 2985	0,31 179	0,31 1746	0,43 133907	0,43 140923
	No cualificados	0,39 2267	0,34 3971	0,39 13908	0,34 1076	0,34 12485	0,45 250305	0,44 284012
Total		0,38 8901	0,34 10887	0,38 36067	0,30 2618	0,32 26870	0,44 926662	0,44 1012005

Fuente: Elaboración propia a partir de datos procedentes de la Estadística de accidentes de trabajo de 2005.

En dicho cuadro se observa que los porcentajes son mayores en las actividades más relacionadas con el sector terciario de la economía y con las ocupaciones que exigen mayores niveles de cualificación. Estos resultados ponen de manifiesto que las diferencias observadas entre nacionalidades no se explican por la actividad desarrollada o el puesto ocupado y que se revela necesario un estudio más pormenorizado.

Para terminar el análisis de los datos es necesario aclarar que, antes de realizar el análisis empírico posterior, se ha procedido a realizar una depuración de los datos con el fin de que se pudieran incluir todas las variables explicativas. En primer lugar, y con el fin de poder incluir la variable que mide los días de baja, se eliminan todos aquellos accidentes sin fecha de baja y a los que la estadística les consigna una duración estándar que no es real. En segundo lugar se prescinde de todos los accidentes cuya indemnización supera el máximo legal, así se eliminan aquellos datos anormalmente altos que pueden deberse a fallos de registro y falsear la variable indemnización. Para terminar se eliminan los accidentes relacionados con contratos sin código conocido y con trabajadores de más de 100 años. Una vez realizados todos estos ajustes se dispone de 861530 accidentes referidos al año 2005 que son los que luego se emplearán en el análisis empírico.

3.- METODOLOGÍA

Un primer aspecto relevante desde el punto de vista metodológico lo constituye la definición de la variable dependiente. Como se ha explicado previamente, la base de datos esta constituida por microdatos relativos a trabajadores accidentados, por lo tanto no se dispone del dato correspondiente a los trabajadores sujetos a riesgo para construir un indicador de frecuencia como el denominado índice de incidencia elaborado para datos agregados. Por otro lado, y como se ha apuntado en la literatura (Smith, 1990; Butler et al. 1996; Card y McCall, 1996; Bolduc et al. 2001), los problemas de riesgo moral que se estudian en este artículo son más frecuentes en aquellos accidentes asociados a lesiones de difícil diagnóstico (básicamente esguinces, torceduras y lumbalgias).

Con estos antecedentes, la estrategia empírica seguida en esta investigación consiste en construir una variable dependiente que mida la probabilidad de tener un accidente con una lesión de difícil diagnóstico condicionada a haber sufrido un accidente. Así, se obtiene un índice de frecuencia que puede ser utilizado para medir si existen diferencias injustificadas en la frecuencia del reporte de accidentes de difícil diagnóstico entre las distintas nacionalidades. Nótese que, en este punto, seguimos el mismo procedimiento para construir nuestra variable dependiente que Campolietti y Hyatt (2006).

Disponemos por lo tanto de una variable dependiente que tomará el valor 1 si el accidente es de difícil diagnóstico y cero en el caso contrario. A partir de aquí la primera parte del trabajo se dedica a realizar un análisis probit donde se regresará esta variable contra diferentes grupos de variables explicativas. Estos grupos siempre incluirán todas las nacionalices estudiadas, y se irán modificando con la progresiva introducción de otras variables independientes. De esta manera se podrán extraer conclusiones sobre la evolución seguida por los coeficientes de cada una de las nacionalidades cuando se recogen diferentes características de los accidentados.

En segundo lugar se realiza una generalización de la descomposición de Oaxaca Blinder para modelos no lineales. Con esta descomposición se consigue determinar que parte de la diferencia observada entre el porcentaje de accidentes de difícil dialgnóstico que manifiesta cada una de las nacionalidades se justifica por que los grupos son diferentes, y que parte es consecuencia de una respuesta diferente a características similares.

Descomposición no lineal.

Oaxaca (1973) y Blinder (1973) generalizaron un método para estudiar las diferencias entre grupos poblacionales que ha aportado multitud de resultados sobre todo en temas de discriminación salarial. Sin embargo, la descomposición de Oaxaca-Blinder presenta dos

limitaciones fundamentales. De un lado, se refiere solo al anásis de modelos lineales, y de otro, plantea problemas de identificación si se intenta calcular la aportación de variables individuales en presencia de grupos de dummies (Oaxaca y Ransom, 1999).

La literatura ya incluye bastante evidencia sobre trabajos que descomponen modelos no lineales. Even y Macpherson (1990), Fairlie (1999), Hernanz y Toharia (2006), Motellon y Lopez-Bazo (2005) o Yun (2005) realizan descomposiciones para modelos probit, Nielsen (1998) hace una aproximación para modelos logit y Fairlie (2003) realiza otra aplicación donde se estudian tanto modelos logit como probit. Yun (2004) consigue desarrollar una generalización de la metodología de *Oaxaca-Blinder* que permite realizar la descomposición para cualquier tipo de relación funcional y que consigue calcular la aportación de cada variable sin incurrir en problemas de identificación.

De acuerdo con Yun, si disponemos de una variable dependiente *Y,* que es función de una combinación lineal de variables independientes *XB,* a través de una función φ no lineal ($Y = \varphi(XB)$). Se puede descomponer la diferencia en la media de la variable dependiente entre dos grupos poblacionales 1 y 2 de acuerdo a la siguiente expresión:

$$\overline{Y}_1 - \overline{Y}_2 = \left[\overline{\varphi(X_1 B_1)} - \overline{\varphi(X_2 B_1)}\right] + \left[\overline{\varphi(X_2 B_1)} - \overline{\varphi(X_2 B_2)}\right] \quad (1)$$

El primer sumando de la parte derecha de la expresión (1) respondería a la diferencia justificada e indicaría, para unos mismos coeficientes, la parte explicada por las distintas características de cada grupo. El segundo sumando representaría la diferencia injustificada, o lo que es lo mismo, para unas mismas características, la parte debida al distinto rendimiento de estas.

4. RESULTADOS

El análisis empírico del trabajo comienza con la realización de un test de igualdad de proporciones para determinar si es coherente plantearnos un comportamiento diferencial entre los trabajadores nacionales y los inmigrantes. En el cuadro 2 se muestran los resultados de este test y se comprueba la validez de este planteamiento. El estadístico Z manifiesta la relevancia de las diferencias observadas para un nivel de significación del 1%, alcanzando un máximo cuando se comparan los accidentados españoles con los inmigrantes sin distinguir la procedencia.

En ese mismo cuadro también se observa que la procedencia importa. Así se puede observar que el porcentaje de accidentes de difícil diagnóstico es un 15% más alto en España que en el grupo de países que se han denominado como desarrollados, pero este porcentaje sube hasta el 50% si se comparan con los accidentados de procedencia asiática.

Cuadro 2. Test de igualdad de proporciones entre el porcentaje de accidentes de difícil diagnóstico de los accidentados españoles y los inmigrantes.

	Media	Diferencia	Estadístico Z
Total inmigrantes	0,36	-0,09	-43,27
Paises Desarrollados	0,38	-0,06	-9,99
Europa del Este	0,34	-0,10	-17,98
Países Latinoamericanos	0,38	-0,06	-21,06
Asia	0,30	-0,15	-13,51
África	0,32	-0,12	-33,36

Fuente: Elaboración propia a partir de datos procedentes de la Estadística de accidentes de trabajo de 2005.

Una vez comprobada la heterogeneidad en el comportamiento de los accidentados, la segunda parte del trabajo intenta determinar si ese comportamiento diferencial se mantiene cuando se controla por distintos grupos de variables explicativas. Con esta finalidad se estimará, mediante un probit, cual es la influencia de cada una de las nacionalidades sobre la probabilidad de reportar una lesión de difícil diagnóstico.

El análisis se realizará de una forma progresiva, en primer lugar se estimará el modelo I donde las únicas variables explicativas son las dummies que indican la nacionalidad del accidentado. Posteriormente se analizarán modelos cada vez más completos donde se irán introduciendo diferentes grupos de variables. Así en el modelo II se incluyen variables personales como el sexo, la edad, la cualificación[4] o la experiencia del individuo. En el modelo 3, además de las anteriores se añaden características del accidente sobre como, cuando y donde se produjo. El modelo IV incluye las variables referidas a la empresa indicando cual es su rama de actividad, su tamaño, si se trata de una empresa de trabajo temporal o si efectuó algún tipo de prevención de riesgos. Y por último el modelo V añade las variables referidas a la lesión como su gravedad, la parte del cuerpo lesionada, o la forma en que esta tuvo lugar.

Cuadro 3. Efecto sobre la probabilidad de sufrir un accidente de difícil diagnóstico de las variables que recogen la procedencia del accidentado según la complejidad del modelo.

		Modelo I	Modelo II	Modelo III	Modelo IV	Modelo V
Países Desarrollados	**Coef**	-0,06	-0,05	-0,05	-0,04	-0,03
	Est. Z	-10,02	-7,68	-7,78	-7,23	-4,52
Europa del Este	**Coef**	-0,10	-0,07	-0,07	-0,07	-0,05
	Est. Z	-18,05	-13,23	-11,74	-12,60	-7,95
Países Latinoamericanos	**Coef**	-0,06	-0,05	-0,05	-0,06	-0,04
	Est. Z	-21,12	-16,87	-16,95	-17,70	-11,28
Asia	**Coef**	-0,15	-0,12	-0,11	-0,11	-0,06
	Est. Z	-13,56	-11,39	-9,68	-9,98	-4,81
África	**Coef**	-0,12	-0,09	-0,08	-0,09	-0,05
	Est. Z	-33,50	-24,82	-22,21	-23,56	-13,37
Características personales			*	*	*	*
Características del accidente				*	*	*
Características de la empresa					*	*
Característica de la lesión						*
Log Verosimilitud		-589117,6	-582773,5	-561548,9	-561878,6	-440626,7

Fuente: Elaboración propia a partir de datos procedentes de la Estadística de accidentes de trabajo de 2005.

[4] Como Proxy de la cualificación se utilizan variables que indican la ocupación y el régimen de cotización a la Seguridad Social

Los resultados de estas estimaciones aparecen recogidos en el cuadro 3. Se puede apreciar como en el modelo I se trasladan las diferencias que ya se observaban en el análisis descriptivo. La probabilidad de sufrir un accidente de difícil diagnóstico es 6 puntos porcentuales menor si el accidentado inmigrante procede de un país desarrollado o latinoamericano, 10 si procede de Europa del Este, 12 si es africano y 15 si procede de un país asiático.

Por otro lado, parece lógico suponer que habrá determinadas características que pueden influir en estos porcentajes y que pueden explicar parte de las diferencias. El modelo II reduce en cierta medida esas diferencias que ahora se sitúan entre los 5 puntos porcentuales para los inmigrantes de países desarrollados y latinoamericanos, y los 12 puntos que se mantienen con los que proceden de Asia. Este resultado indica que existen diferencias en edad, sexo y cualificación que explican parte de la diferencia entre inmigrantes y nacionales.

Los modelos II y III no aportan mucho más a la estimación, por lo tanto, no existen grandes diferencias entre las circunstancias que rodean al accidente de un trabajador inmigrante y al de un trabajador nacional. Y tampoco se aprecia que las características de las empresas en que trabajan los accidentados nacionales o extranjeros afecten demasiado al porcentaje de accidentes de difícil diagnóstico que estos declaran.

Finalmente el modelo IV muestra otra fuerte reducción en la probabilidad objeto de estudio. Cuando se introducen variables que indican aspectos del diagnostico o de la lesión propiamente dicha las diferencias oscilan entre los 3 puntos porcentuales de los accidentados procedentes de países desarrollados y los 6 de los procedentes de Asia. Este resultado indica que las lesiones que sufren los trabajadores nacionales si son sensiblemente diferentes a las que sufren los inmigrantes. Este resultado puede estar en gran medida relacionado con cuestiones relativas al esfuerzo y al riesgo moral que se analiza en este trabajo. Es decir, si un trabajador pretende reportar un accidente que ha tenido lugar fuera del puesto de trabajo, o que ni siquiera a tenido lugar, este seguramente sea un accidente leve, normalmente tendrá una determinada caracterización (contusión o sobreesfuerzo) y posiblemente se sitúe cierta parte del cuerpo (extremidades, cuello o espalda).

La última parte del estudio econométrico se dedica a realizar una descomposición de la diferencia de porcentaje de accidentes de difícil diagnóstico que reportan los accidentados nacionales frente a los diferentes grupos de inmigrantes. Para ello se han vuelto a aplicar estimaciones tipo probit, ahora para cada grupo poblacional de forma independiente, aunque hemos prescindido de las variables referidas al diagnóstico porque, como se ha explicado previamente, pueden estar explicando parte del comportamiento discrecional de los agentes.

Cuadro 4. Descomposición de la diferencia en el porcentaje de accidentes de difícil diagnóstico que reportan los trabajadores nacionales y los inmigrantes según procedencia.

	Diferencia	Componente justificado	Componente injustificado	% Justificado	% injustificado
Total inmigrantes	0,09	0,02	0,06	28%	72%
Países Desarrollados	0,06	0,02	0,04	33%	67%
Europa del Este	0,10	0,03	0,06	35%	65%
Países Latinoamericanos	0,06	0,01	0,05	18%	82%
Asia	0,15	0,05	0,10	31%	69%
África	0,12	0,04	0,08	33%	67%

Fuente: Elaboración propia a partir de datos procedentes de la Estadística de accidentes de trabajo de 2005

Los resultados de la descomposición no lineal aparecen en el cuadro 4 y se recogen en cinco columnas. La primera se refiere a la diferencia en el porcentaje de accidentes de difícil diagnóstico que presentan los trabajadores nacionales frente al conjunto de inmigrantes y frente a cada una de las procedencias. Las otras cuatro columnas se refieren a la descomposición no lineal, las dos primeras en términos absolutos y las dos siguientes en porcentajes.

El componente justificado indica, como ya se ha explicado, cuales son las diferencias que se explican porque cada grupo de trabajadores tiene unas características personales diferentes, sufre accidentes distintos o no pertenecen al mismo tipo de empresa. Por su parte el componente injustificado refleja la distinta repercusión de características similares, es decir, que trabajadores de iguales características que trabajan en sitios parecidos y con las mismas condiciones repercuten diferente porcentaje de accidentes de difícil diagnóstico según sea su procedencia.

Los resultados indican que solo una pequeña parte de la diferencia observada se explica por la heterogeneidad de los grupos de trabajadores. Dicho de otra manera, si las variables introducidas afectaran a los trabajadores inmigrantes como a los nacionales, el porcentaje de accidentes de difícil diagnóstico que reportarían los inmigrantes sería dos puntos mayor. Si se distingue por la procedencia de los inmigrantes, los trabajadores asiáticos aumentarían 5 puntos el reporte de este tipo de accidentes, los africanos 4, los europeos del este 3, los de países desarrollados 2 y los latinoamericanos 1.

La mayor parte de la diferencia es consecuencia de que características similares repercuten de forma diferente en los trabajadores dependiendo de cual sea la procedencia. Es este componente injustificado el que se puede identificar con posibles situaciones de riesgo moral o con diferencias relativas en el esfuerzo de los trabajadores. Parece lógico suponer que los trabajadores que más se esfuerzan son menos propensos a declarar accidentes que se han producido fuera del puesto de trabajo o que ni siquiera han tenido lugar. Por lo tanto, el hecho de que la mayor parte de las diferencias no se explique por la heterogeneidad de los grupos, pone de manifiesto que los trabajadores inmigrantes presentan mayores niveles de esfuerzo relativo que los nacionales, o también que presentan menos situaciones de absentismo encubierto. Si ampliamos el análisis a las distintas procedencias, se puede afirmar que los trabajadores que más se esfuerzan y menos situaciones de riesgo moral manifiestan serían los asiáticos y africanos

Efectos de la antigüedad en el esfuerzo de los trabajadores inmigrantes.

Anteriormente se ha podido comprobar la existencia de diferencias en el esfuerzo asociadas a la procedencia de los trabajadores. Sin embargo, parece lógico suponer que según aumenta el tiempo de residencia en nuestro país[5], ese comportamiento cambie y que las diferencias de esfuerzo de unos y otros trabajadores se reduzcan. Para analizar este supuesto se utiliza nuevamente la descomposición no lineal de la diferencia en el porcentaje de accidentes de difícil diagnóstico que reportan trabajadores nacionales e inmigrantes, y se estudia el componente injustificado en función de la antigüedad de los trabajadores.

En el cuadro 5 se recoge la diferencia observada en el porcentaje de accidentes de difícil diagnóstico, y el componente injustificado de esa diferencia según aumenta la antigüedad en el puesto de trabajo que están desempeñando en el momento de sufrir el accidente.

[5] La proxi utilizada para medir el tiempo de residencia en nuestro país es la antigüedad en el puesto de trabajo.

Cuadro 5. Evolución de las diferencias en el porcentaje de accidentes de difícil diagnóstico entre trabajadores nacionales e inmigrantes según aumenta la antigüedad en el puesto.

Antigüedad	Diferencia Observada	Diferencia no justificada	
		Términos absolutos	Términos porcentuales
Menos de 1 año	0,080	0,067	84%
De 1 a 2 años	0,072	0,064	89%
De 2 a 5 años	0,081	0,063	77%
De 5 a 10 años	0,081	0,059	73%
Más de 10 años	0,096	0,054	56%

Fuente: Elaboración propia a partir de datos de la EAT.

Los resultados observados ponen de manifiesto que, aunque las diferencias observadas aumentan con la antigüedad de los trabajadores, el componente injustificado se reduce tanto en términos absolutos como relativos. Para los trabajadores con menos antigüedad, de los 8 puntos porcentuales más de accidentes de difícil diagnóstico que reportan los trabajadores españoles, el 84 % (6,7 puntos) permanece inexplicado y puede asociarse a un mayor nivel de esfuerzo de los trabajadores inmigrantes. Sin embargo, para los trabajadores con antigüedad superior a los 10 años, aunque la diferencia real casi alcanza los 10 puntos porcentuales, ya solo un 56% (5,4 puntos) permanece inexplicada, y por lo tanto las diferencias de esfuerzo parecen reducirse.

Merece la pena comentar que las diferencias injustificadas ahora tienen de forma generalizada un valor mayor que la que se observaba cuando se analizaba toda la muestra. Eso es debido a que se ha eliminado de la regresión probit la experiencia como variable explicativa porque los resultados se muestran para una ordenación de trabajadores determinada por su antigüedad.

5. CONCLUSIONES

El propósito del presente trabajo era tratar el problema de la inmigración desde una vertiente un poco distinta, relacionada con los problemas de seguridad y salud que se producen en el mercado de trabajo. Con esta perspectiva, se pretende aportar una medida del esfuerzo relativo de los trabajadores en función de su procedencia, y para ello se miden las posibles situaciones de absentismo encubierto que se manifiestan a través de la declaración de cierto tipo de accidentes.

El simple análisis descriptivo ya pone de manifiesto la existencia de diferencias sensibles entre el porcentaje de accidentes de difícil diagnóstico que reportan los trabajadores accidentados nacionales y los inmigrantes. Diferencias que se mantienen y amplían cuando se analizan diferentes procedencias para esos procesos migratorios. Este resultado, ya de por si significativo, se magnifica si tenemos en cuenta que parte de esas diferencias se mantienen cuando se incluyen en el modelo factores explicativos como las características personales del trabajador, aspectos relativos al puesto ocupado, detalles sobre la empresa en que trabaja, la descripción del tipo de accidente que ha sufrido o incluso de lesión que ha soportado.

El análisis empírico desarrollado pone de manifiesto que, solo una pequeña parte de la diferencia en el porcentaje de accidentes de difícil diagnóstico que reporta cada grupo, se debe a que los trabajadores accidentados inmigrantes son diferentes a los nacionales, realizan labores distintas o sufren otro tipo de accidentes. Por lo tanto, existe otra parte de esa diferencia que indica que en circunstancias similares, los trabajadores nacionales declararían un mayor porcentaje de ese tipo de accidentes que los inmigrantes. Es esta parte no justificada por las

distintas características de cada grupo de trabajadores la que parece indicar que existe un mayor nivel de esfuerzo relativo en los trabajadores inmigrantes, en especial asiáticos y africanos, y que se traduce en que estos declaran un menor porcentaje de accidentes relacionados con posibles situaciones de riesgo moral. Finalmente, es de destacar que las diferencias en el esfuerzo relativo a favor de los trabajadores inmigrantes parecen reducirse a medida que aumenta la antigüedad en el puesto de trabajo como indicador del tiempo de permanencia en nuestro país.

BIBLIOGRAFÍA

AHONEN, E. Q. Y BENAVIDES, F. G. (2008): "Riesgo de lesion por accidente laboral en trabajadores extranjeros", Gaceta Sanitaria, 22:1, 44-47.

AHONEN, E. Q. Y BENAVIDES, F. G. (2006): "Risk of fatal and non-fatal occupational injury in foreign workers in Spain", Journal of Epidemiology and Community Health, 60, 424-426.

AMUEDO-DORANTES, C.; DE LA RICA, S. (2005): "Immigrants' Responsiveness to Labor Market Conditions and Its Implications on Regional Disparities: Evidence from Spain", IZA DP 1557.

AMUEDO-DORANTES, C.; DE LA RICA, S (2006): "Labor Market Assimilation of Recents Immigrants in Spain", IZA DP 2104.

BARTEL, ANN P., AND LACY GLENN THOMAS. (1985.) "Direct and Indirect Effects of Regulation: New Look at OSHA's Impact." Journal of Law and Economics, Vol. 28, pp. 1-25.

BAUER, T. K.; MILLION, A.; ROTTE, R.; ZIMMERMANN, K.F. (1999): "Immigration Labor and Workplace Safety", IZA DP nº 16.

BLINDER, A. S. (1973): 'Wage discrimination: reduced form and structural estimates', Journal of Human Resources, 8, 436-455.

BOLDUC, B., FORTIN, B., LABRECQUE, F. Y LANOIE P. (2001): 'Workers' compensation, moral hazard, and the composition of workplace injuries', Journal of Human Resources, 37(3), 623-652.

BUTLER, R. J., DURBIN, D. L. Y HELVACIAN, N. M. (1996): 'Increasing claims for soft tissue injuries in workers' compensation: cost shifting and moral hazard', Journal of Risk and Uncertainty, 13, 73-87.

BUTLER, RICHARD J., AND JOHN D. WORRALL. 1983. "Workers' Compensation: Benefit and Injury Claims Rates in the Seventies." Review of Economics and Statistics, Vol. 65, pp. 580-589.

BORJAS G. J. (1999): "The Economic Analysis of Immigration", en Ashenfelter O. y Card D. (eds.). Handbook of Labor Economics, vol. 3A, ed. North-Holland.

CAMPOLIETTI, M. AND HYATT, D. E. (2006) Further evidence on the "Monday effect" in workers' compensation, Industrial and Labor Relations Review, 59(3), 438-50.

CARD, D. (2005), "Is New Immigration Really so Bad?", The Economic Journal, vol. 115, 300–323.

CARD, D. Y MCCALL, B. P. (1996): 'Is workers' compensation covering uninsured medical cost? Evidence from the "monday effect"', Industrial and Labor Relations Review, 49(4), 690-706.

CARRASCO, R.; SERRANO, C.; MALO, M.A. (2003): "Mercado de trabajo e inmigración", en A. Izquierdo (ed.) Inmigración, mercado de trabajo y protección social en España, ed. Consejo Económico y Social.

CARRASCO, R.; JIMENO, J.F.; ORTEGA, A.C. (2008): "The Effect of Immigration on the Employment Opportunities of Native-Born Workers: Some Evidence for Spain", journal of population economics 3/08.

COLLADO, D.; ITURBE, I.; VALERA, G. (2004): "Quantifying the Impact of Immigration on the Spanish Welfare State", International Tax and Public Finance, 11(3), 335-353.

CONSEJO ECONÓMICO Y SOCIAL (2004): La inmigración y el mercado de trabajo en España, Informe 2/2004.

DOLADO, J.J.; JIMENO, J.F.; DUCE, R. (1997): "Los efectos de la inmigración sobre la demanda relativa de trabajo cualificado vs. poco cualificado: evidencia para España", Cuadernos Económicos del ICE, n.º 63, 11-30.

EVEN, W. E. Y MACPHERSON, D. A. (1990): 'Plant size and the decline of unionism'. Economic letters 32, 393-398

FAIRLEI, R. W. (1999): 'The absence of the African-American owned business: an analysisis of the dynamics of self-employment'. Journal of Labor Economics, 17, 80-108.

FAIRLEI, R. W. (2003): 'An extension of the Oaxaca-Blinder decomposition technique to Logit and Probit models'. Working Paper, Economic Growth Center, Yale University.

FRIEDBERG, R. M.; HUNT, F. (1995): "The Impact of Immigrants on Host Country Wages, Employment and Growth", Journal of Economic Perspectives, 9(2), págs. 23–44.

GARRIDO, L.; TOHARIA, L. (2004): "La situación laboral de los españoles y los extranjeros según la Encuesta de Población Activa", Economistas, nº 99, 74-86.

GRAHAM, JULIE, AND DON M. SHAKOW. 1990. "Labor Market Segmentation and Job-Related Risks.", American Journal of Economics and Sociology, Vol. 49, pp. 307-324.

HERNANZ, V. AND TOHARIA, L. (2006) Do temporary contracts increase work accidents? A microeconometric comparison between Italy and Spain, Labour, 20(3), 475-504.

IGLESIAS, C.; LLORENTE, R. (2006): "¿Integración o segmentación laboral de los inmigrantes en el mercado de trabajo? Un análisis de cohortes", SERVILAB DT-07/2006.

IZQUIERDO, M.; JIMENO, J.F. (2005): "Inmigración: desarrollos recientes y consecuencias económicas", Boletín Económico del Banco de España, febrero, 41-49.

LALONDE, R.J.; TOPEL, R.H. (1997): "The economic impact of international migration and the economic performance of migrants", en Rosenzweig and Stark (eds) Handbook in population and family economics, ed. Elsevier Science.

MOTELLÓN, E. Y LÓPEZ-BAZO, E. (2005): "Discriminación por género en el acceso a la contratación indefinida". VIII Encuentro de Economía Aplicada. Murcia.

NIELSEN, H. S. (1998). "Discrimination an detailed decomposition in a logit model". Economic letters 61, 115-120

OAXACA, R. (1973): "Male-female wage differentials in urban labour markets", International Economic Review, 14, 693-709.

OAXACA, R. AND RANSOM, M. (1999): "Identification in detailed wage decompositions," Review of Economics and Estatistics, 81:1, 154-157

RICHARDSON, D.B., LOOMIS, D., BENA, J. AND BAILER, J. (2004): "Fatal Occupational Injury Rates in Southern and Non-Southern, By Race and Hispanic Ethnicity", American Journal of Public Health, 94:10, 1756-61.

RICHARDSON, S. AND LOH, K. (2004): "Foreign-born workers: trends in fatal occupational injuries, 1996-2001.", Monthly Labor Review, June 2004, 42-53.

SANROMÁ, E.; RAMOS, R.; SIMÓN, H. (2006): "Inmigración reciente en España: sobreeducación y asimilación en el mercado de trabajo", IX Encuentro de Economía Aplicada, Jaen.

SIMÓN, H.; SANROMÁ, E.; RAMOS, R. (2007): "Segregación laboral y estructura salarial de nativos e inmigrantes. Un análisis con datos emparejados empresa-trabajador", VII Jornadas de Economía Laboral.

SMITH, R. S. (1990): 'Mostly on modays: is workers' compensation covering off-the-job injuries?', in: Philip S. Borba and David Appel (eds.), Benefits, Costs, and Cycles in Workers' Compensation. (Boston: Kluwer Academic Publishers),

WORRALL, JOHN D., AND RICHARD J. BUTLER. (1983). "Health Conditions and Job Hazards: Union and Nonunion Jobs. Journal of Labor Research, Vol. 4, pp. 338-347.

YUN, M. (2004): 'Decomposing differences in the first moment', Economics Letters, 82, 275-280

YUN, M. (2005): 'Normalized equation and decomposition analysis: computation and inference'. IZA Discussion Paper, 1822, Tulane University.

ÁREA VIII

ECONOMÍA MONETARIA Y FINANCIERA

ECONOMIA MONETÁRIA E FINANCEIRA

CALENDAR EFFECTS IN STOCK MARKETS: CRITIQUE OF PREVIOUS METHODOLOGIES AND RECENT EVIDENCE IN EUROPEAN COUNTRIES

MARIA ROSA BORGES
E-mail: mrborges@iseg.utl.pt
ISEG - Instituto Superior de Economia e Gestão
UECE - Research Unit on Complexity and Economics
Universidade Técnica de Lisboa (Portugal)

Resumo

Este estudo analisa os efeitos "dia da semana" e "mês do ano" nos índices de mercado de dezassete países europeus, no período 1994-2007. Discutem-se os modelos e testes empíricos utilizados em estudos anteriores, e propõe-se uma especificação mais simples, utilizável para a detecção de todos os tipos de efeitos de calendário. Reconhecendo que os rendimentos não são normalmente distribuídos, e os problemas de autocorrelação e heteroescedasticidade dos resíduos, são utilizadas metodologias de estimação robustas, incluindo *bootstrapping* e modelos GARCH. Apesar de os rendimentos tenderem a ser menores nos meses de Agosto e Setembro, não encontramos evidência consistente da existência de efeitos de calendário em todos os países, sendo a evidência mais favorável apenas específica de alguns países. Utilizando regressões com janelas flutuantes, verifica-se que os efeitos de calendário mais fortes não são estáveis durante o período em análise, lançando dúvidas adicionais sobre a sua significância económica.

Palavras-Chave: Efeito "dia da semana"; Efeito "mês do ano", Eficiência do mercado, Mercados Europeus.
Área Temática: Economia Monetária e Financeira.

Abstract

This paper examines day of the week and month of the year effects in seventeen European stock market indexes in the period 1994-2007. We discuss the shortcomings of model specifications and tests used in previous work, and propose a simpler specification, usable for detecting all types of calendar effects. Recognizing that returns are non-normally distributed, autocorrelated and that the residuals of linear regressions are variant over time, we use statically robust estimation methodologies, including bootstrapping and GARCH modeling. Although returns tend to be lower in the months of August and September, we do not find strong evidence of across-the-board calendar effects, as the most favorable evidence is only country-specific. Additionally, using rolling windows regressions, we find that the stronger country-specific calendar effects are not stable over the whole sample period, casting additional doubt on the economic significance of calendar effects.

Key words: Day-of-the-week effect; Month effect, Market efficiency, European stock markets.
Thematic Area: Monetary and Financial Economics.

1. INTRODUCTION

Several empirical studies have studied the phenomena of calendar effects in stock markets, where returns tend to show higher (or lower) than average returns is specific calendar periods. The calendar effects that have attracted more interested, fueled by favorable evidence, are: (i) the weekend effect, where Monday returns tend to be lower than on other days of the week, and sometimes Friday returns are higher; and (ii) the January effect, revealed in the fact that daily returns tend to be higher in this month, than in other months of the year. Other calendar effects that have been studied include day of the month effects, where higher returns tend to be

concentrated in specific periods of the month, and holiday effects, where we observe the behavior of returns after holidays (no trading days).

The study of calendar effects is relevant, in financial economics, because some types of calendar effects are inconsistent with the efficient market hypothesis. If the flow of information is continuous, and prices reflect all information, we would expect to find that Monday returns are around three times higher than other weekday returns, because there are three calendar days between the market closing of Friday, and the market closing of Monday. But even if we admit that the flow of information is negligible on weekends, Monday returns should at least be as high as other weekday returns. However, none of these two hypotheses is confirmed in the US market, nor in several other markets. Monday returns are in fact lower than other weekday returns. On the other hand, month effects are not necessarily inconsistent with market efficiency, because it is possible that the flow of information to the markets is specially concentrated in one, or some, of the months of the year. In any case, there is no strong evidence that January higher returns are caused by a relatively higher flux of good news, and so calendar effects remain at odds with both the hypothesis of: (i) market efficiency and (ii) rational behavior of investors. The study of calendar effects is also relevant for financial managers, financial counselors, market professionals and investors in general, and all those interested in developing profitable trading strategies.

This paper looks exclusively at day of the week effects and month of the year effects, in European stock markets. It makes several contributions to the literature on calendar effects in stock market returns. First, we discuss the shortcomings of previously used models for the detection of calendar effects, and we propose a simpler model specification that overcomes those shortcomings. Second, we recognize non-normality and autocorrelation in stock market returns, and time-dependent variance of the residuals of linear regressions, and apply appropriate statistical methodologies to tackle these problems, including the bootstrap approach and the GARCH model, adding statistical robustness to our results. Third, we examine the time-stability of the most significant calendar effects in the period under study. Fourth, we use observations from a set of seventeen countries of the same economic region, allowing us to conclude if calendar effects are across-the-board effects in that region or only country-specific effects. This is important to know, because some possible explanations for calendar effects, like psychological traits of investors, would imply across-the-board effects, while other explanations, like those related to fiscal motivations or market structure, allow for country-specific calendar effects. Five, we use data from recent years, from 1994 to 2007, on West and Central European stock markets, thus adding and updating international evidence on calendar effects.

The remainder of the paper is organized as follows. In section 2, we present some of the more relevant previous studies and results on day of the week effects and month of the year effects. In section 3, we present the data, including several descriptive statistics. In section 4, we discuss alternative model specifications and their shortcomings, and the different statistical methodologies we use for estimating the calendar effects. Section 5 contains the results of the model estimations and also includes an examination of the time-stability of the detected calendar effects. In Section 6, we present the conclusions and suggestions for further research.

2. LITERATURE REVIEW

Many researchers have studied the phenomenon of seasonalities in price movements in stock markets, related to specific calendar periods. These regularities are known as calendar effects. The most commonly studied calendar effects, which we also cover in the present study are: (i) the day of the week effect, and (ii) the month of the year effect. There are several studies which focus on other types of calendar effects, like the behavior of daily returns after holidays, or the behavior of returns in the first trading days of each month, but those are beyond the scope of this paper. In this section, we present a short review of previous works on day of the week and month of the year effects, and of the results they find.

2.1. The day of the week effect

Cross (1973) is among the group of authors that first studies a day of the week effect, namely, the weekend effect. He observes several US market indexes, without performing statistical tests, and finds that stocks have a negative return over the weekends. French (1980), Keim and Stambaugh (1984), Rogalski (1984), and Smirlock and Sarks (1986) examine the Standard & Poor's and the Dow Jones Index and conclude that Monday returns are on average negative. However, Rogalski (1984), using OLS regressions, F-tests and t-tests, observes that the Monday effect is negative but not statistically significant. In the nineties, Chang et al. (1993) and Kamara (1997) confirm the validity of the weekend effect.

Using the same approach as Rogalski (1984), other authors, including Jaffe and Westerfield (1985a, 1985b), Condoyanni et al. (1987) and Chang, et al. (1993), study non-US markets, including Japan, Singapore, Australia, Canada, UK, and other European countries, and find that Monday returns are on average negative and statistically significant. Other studies find a day of the week effect in different days. Brooks and Persand (2001) observe significant negative returns on Tuesdays in Thailand and Malaysia, and a significant Wednesday effect in Taiwan. Jaffe and Westerfield (1985a) and Dubois and Louvet (1996) confirm that daily returns in some Pacific countries tend to be negative on Tuesdays.

This apparent consensus is challenged by a set of more recent studies. Sullivan, Timmermann and White (2001) use a non traditional approach (a bootstrap procedure) and conclude that calendar effects no longer remain statistically significant. Rubinstein (2001), Mabberly and Waggoner (2000), Schwert (2001), Steeley (2001), Kohers et al. (2004) and Hui (2005) undertake international studies and show that this market anomaly is recently becoming weaker, particularly in developed markets. More recently, Chukwuogor-Ndu (2006) analyze the day of the week effect in stock market returns in fifteen European countries and finds corroborative evidence in only seven of those markets. He also finds significant negative returns on Tuesdays, in some of these countries. Basher and Sadorsky (2006), using different models for detecting the day of the week effect, conclude that a majority of the twenty one emerging stock markets they examine do not have such an effect, but some countries do exhibit strong day of the week effects, even after considering for conditional market risk. Overall, there is mixed evidence on day of the week effects, as more recent studies, using more advanced statistical procedures, have cast some doubt on the favorable evidence from the initial studies.

2.2. The month of the year effect

A month of the year effect exists if returns tend to be higher or lower in a specific month, when compared with the other months of the year. The most commonly reported month effect is the tendency for returns to be higher in January, although other month effects have also been reported. The first studies, by Rozeff and Kinney (1976), Dyl (1977) and Brown et al. (1983) analyze the US stock market and observe significant higher returns in January than in the other months of the year. Also, Gultekin and Gultekin (1983) study seventeen countries using both non-parametric and parametric tests, and conclude that January returns are significantly higher when compared with the other months, in thirteen of those countries.

Keim (1983) links the January effect to a small-firm effect, and a set of international studies find that small firms achieve larger rates of returns than larger firms, and that this is particularly evident in January (Aggarwal, Rao and Hiraki, 1990). Reinganum (1983) also finds that the January effect is largely due to the behavior of prices of small firms, and related to a tax-loss selling hypothesis as proposed by Brown, et al. (1983), who argues that selling pressure at the end of the tax year depresses price that rebound back in January. A study of the UK market by Menyah (1999) finds an April effect for small firms, besides a January effect for larger firms.

Ho (1990) examines twelve stock markets, including Australia, Japan, Korea, New Zealand, Singapore, Thailand, UK and US, and finds evidence corroborative of the January effect as he observes that average returns on January are higher than other months at a 95% level of confidence. More recently, Haugen and Jorion (1996), Tonchev and Kim (2004) and Rosenberg (2004) reach empirical findings similar to prior studies. In balance, the evidence of a January effect is mostly confirmatory, although the reasons why it exists are still under discussion.

3. DATA

We collect from Reuters daily data on seventeen Western and Central European stock market indexes, for the period beginning in January, 1994 through to December, 2007. The countries and respective stock market indexes are, in alphabetical order: Austria (ATX), Denmark (OMXC20), Finland (OMXHPI), France (CAC40), Germany (DAX), Greece (ASE), Hungary (BUX), Iceland (OMXIPI), Ireland (ISEQ), Italy (MIBTEL), Netherlands (AEX), Norway (OSEAX), Poland (WIG), Portugal (PSI20), Spain (IBEX), Switzerland (SMI) and United Kingdom (FTSE).

For all indexes, daily returns are computed as:

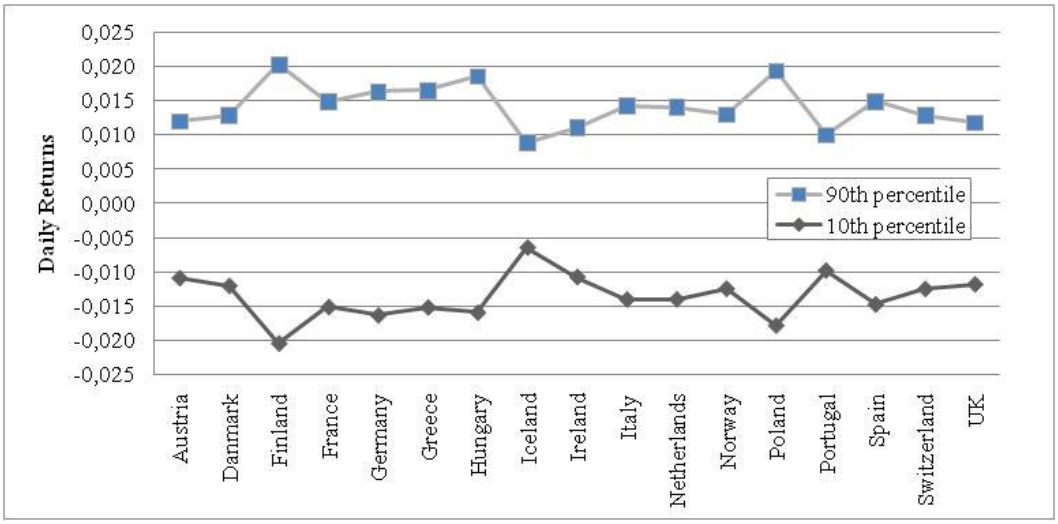
$$r_t = \ln\left(P_t / P_{t-1}\right) \quad (1)$$

Where r_t is the daily return of the stock market index and P_t is the stock index at date t. When the stock market is closed on a weekday, we do not compute the daily return both for that day and for the following weekday, this resulting in two missing observations. Thus all daily returns are computed with a lag of one calendar day, except Mondays, which have a lag of three calendar days. This results in an average of 3430 observations per country, with a maximum of 3595 observations for the United Kingdom and a minimum of 3279 observations for Poland. The descriptive statistics for daily returns of all seventeen stock market indexes are presented in Table 1.

INSERT TABLE 1

From Table 1, we find that the mean daily returns range between 0.016% in United Kingdom and 0.089% in Iceland. In that fourteen year period, the maximum daily returns have been registered in Finland (+14.6%), Hungary (13.6%) and the Netherlands (+ 9.5%), while the minimum daily returns happened in Hungary (-18.0%), Finland (-17.4%) and Poland (-11.3%). The standard error of the mean is lower in Iceland (0.0001313), Portugal (0.0001691) and the United Kingdom (0.0001751), suggestive of lower return volatility, and higher in Finland (0.0003283), Poland (0.0003026) and Hungary (0.0002907), a signal of relatively higher volatility. The 90th percentile daily return ranges from 0.87% in Iceland to 2.02% in Finland. The 10th percentile daily return ranges between -2.05% in Finland and -0.65% in Iceland. Figure 1 confirms, visually, that Finland, Hungary and Poland had wider ranges between the 10th and the 90th percentile, and the narrower ranges were in Iceland, Portugal and Ireland.

Figure 1. 10th and 90th Percentiles of Daily Returns by Country (1994-2007)



In all countries, the distribution of returns is negatively skewed, which means that the left tail (negative returns) concentrates more extreme observations than the right tail. Kurtosis ranges from 5.3 in Denmark to 16.3 in Hungary. In all cases, kurtosis is above 3, which is the expected value for a normal distribution. Thus, all daily return distributions are leptokurtic, meaning that relative to normal distributions, they have both higher peaks and fatter tails (a higher probability of extreme values). The non-normality of the daily returns distributions is also confirmed by Jarque-Bera, Shapiro-Wilk and Shapiro-Francia tests, at least at a 1% significance level, for all countries.

4. METHODOLOGIES

The approach we use in this paper is to analyze daily returns of stock market indexes, comparing the daily returns on specific calendar periods, such as the day of the week and the month of the year, with the daily returns of the remaining days, outside the period under scrutiny. Calendar effects can be studied either using observations of returns of individual stocks of a specific country, or by examining the behaviour of a stock market index (as in French, 1980, Keim and Stambaugh, 1984, Rogalski, 1984, Chang et al, 1983, Basher and Sodorsky, 2006). Officer (1975) claims that calendar effects are more easily detected in market indexes or large stock portfolios than in individual stock prices.

4.1. Discussion of the model specification

When using stock market indexes, a common approach in the literature consists in estimating the following formula. We present the case for the study of the day of the week effect, coding Monday as 1, Tuesday as 2, Wednesday as 3, Thursday as 4 and Friday as 5:

$$r_t = \alpha + \beta_2^* D_{2t} + \beta_3^* D_{3t} + \beta_4^* D_{4t} + \beta_5^* D_{5t} + \varepsilon_t \qquad (2)$$

Where r_t is the daily return of the stock market index, D_{it} are dummy variables which take on the value of 1 if the corresponding return for day t is a Tuesday, Wednesday, Thursday or Friday, respectively and 0 otherwise. Because the dummy for Monday is missing, the constant α captures the mean return on Mondays; β^*_i are coefficients which represent the mean excess daily returns on the remaining days of the week, relative to Mondays; finally, ε_t is the error term.

In this specification, the t-tests of the β^*_i coefficients inform us if they are statistical significant, i.e., if the excess daily returns on Tuesdays, Wednesdays, Thursdays and Fridays, either positive or negative, are significantly different from Mondays' mean return. If we hold an *a priori* belief that an effect exists on one of the specific days, say Monday, this is the best specification. However, if we have no previous expectation on which of the days a calendar effect might exist (or not), the above specification is no longer appropriate. For example, if we want to investigate whether a Thursday effect exists, in the above specification, the coefficient β^*_4 would inform us if Thursdays' returns are statistically different from Mondays' returns, but the model tells us nothing whether Thursdays are different from Tuesdays', Wednesdays' and Fridays' returns. This shortcoming can be overcome by estimating five different models, one for each day of the week, in each case omitting the dummy variable for the day of the week under scrutiny.

But if our purpose is to test a Monday effect, it might be more appropriate to test the mean daily return of that day against the mean daily return of the pool of all non-Monday days, instead of testing Mondays separately against each of the other weekdays. We believe it makes more sense to recognize a Monday effect, if we find out that the mean daily return of that day differs significantly from the mean daily return of non-Mondays, rather than in the case where we find out that Mondays differ from Wednesdays and Thursdays, but not from Tuesdays and Fridays. If, for example, only the coefficient β^*_4 in equation (4) is found to be significant, did we find a Monday effect, or a Wednesday effect?

An alternative specification is to include the dummy variables for all weekdays (all five of them) while excluding the intercept, in order to avoid the dummy variable trap,

$$r_t = \beta_1 D_{1t} + \beta_2 D_{2t} + +\beta_3 D_{3t} + \beta_4 D_{4t} + \beta_5 D_{5t} + \varepsilon_t \quad (3)$$

In this case, the β_i capture the mean daily return for each of the days of the week, but the *t*-tests for those coefficients only inform us if they are significantly different from zero. If the time period under study is sufficiently long, it is to be expected that mean daily return is positive, whilst a very small number[1]. Therefore, the significance of the *t*-tests is biased in favor of accepting positive excess returns, and against accepting negative excess returns. This specific bias can be corrected if we construct our data set with excess daily returns, instead of daily returns.

However, there would still remain a bias, if the excess returns are constructed by deducting the mean daily returns for the all sample. For example, the excess return on Mondays would not be relative to non-Mondays, but rather relative to all days of the week including Mondays. A simple example illustrates this. Suppose we have the same number observations for each of the days of the week, and that the mean returns for Mondays through Fridays are: 0.001, 0.02, 0.02, 0.02 and 0.02, respectively. The overall mean return is 0.0162, while the mean return for Non-Mondays is 0.02 and the mean return including Mondays but excluding one other day of the week is 0.01525. By deducting the overall mean return, we have the following excess returns: -0.0152 for Mondays and 0.0038 for all other weekdays. By deducting the other-day means returns, we obtain the following excess returns, instead: -0.019 for Mondays and 0.00475 for all other weekdays. By using the overall mean daily return to calculate excess daily returns, we would underestimate the absolute value of the excess returns relative to other days, thus biasing the analysis against the detection of existing calendar effects.

We claim that a simpler approach, which overcomes all these shortcomings, is to estimate five equations separately, each aiming to detect a specific day of the week effect:

$$r_t = \alpha + \beta_i D_{it} + \varepsilon_t \quad (4)$$

With this specification, if we include only the dummy variable for Mondays, α captures the mean daily return of non-Mondays, and β_1 is the excess return of Mondays, relative to non-Mondays. The *t*-test of β_1 tells us if this effect is significant. The same arguments apply to β_2, β_3, β_4 and β_5, for detecting other days of the week effects. Note that an OLS regression of this equation is formally identical to performing a two-group mean comparison test between the mean daily return of a specific weekday and the mean daily return of all other weekdays.

All the previous discussion can be transposed to month effect analysis, where the only difference is that we need twelve different dummies, M_i (i=1 to 12), each taking on the value of 1 if the corresponding return for day t is of January, February, through December, respectively and 0 otherwise. With this approach, month effect analysis is more burdensome, because we need to estimate twelve separate equations, one for each month:

$$r_t = \alpha + \beta_i M_{it} + \varepsilon_t \quad (5)$$

This model specification is so general, that any specific calendar effect can be studied this way, like for example, the trading days after holidays, the trading days between Christmas and New Year, the first five trading days of each month, the first one hundred days after a new President elected, and so on. We just need to construct the dummy variable to take the value of 1 in the relevant days.

[1] In this paper, we study calendar effects in eighteen European countries, between 1994 and 2007. From table 1, we can see that the mean daily returns for that fourteen-year period ranges between a maximum of 0.0008868 (in Sweden and Iceland) and a minimum of 0.0001556 in the United Kingdom.

4.2. Estimation procedures

The first studies of calendar effects (French, 1980, Gibbons and Hess, 1981, Jaffe and Westerfield, 1985) employ the linear regression model (OLS) which assumes that the data are normally distributed, serial uncorrelated and with constant variance (Wooldridge, 2003). Connolly (1989, 1991) points out several specific problems that may arise when using this approach: (i) the stock market index returns are likely to be autocorrelated (ii) the residuals are possibly non-normal; (iii) and the variance of the residuals may not be constant.

It is a well documented fact that financial market returns suffer time-dependent changes in volatility (Fama, 1965, Lau et al., 1990, Kim and Kon, 1994). Engle (1982) proposes the use of autoregressive conditional heteroskedasticity (ARCH) models in order to correct the variability in the variance of the residuals. These models assume that the variance of the residuals $\left(\sigma_t^2\right)$ are not constant over time an that the error term can not be modeled $\varepsilon_t \approx iid\left(0,\sigma_t^2\right)$, as assumed in OLS regressions. The generalized version of these models (GARCH) is developed by Bollerslev (1986), where the variance of the residuals is expressed as the sum of a moving-average polynomial of order *q* on past residuals (the ARCH term) plus an autoregressive polynomial of order *p*, on past variances (the GARCH term):

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \lambda_i \sigma_{t-i}^2 \qquad (6)$$

The simplest form is GARCH (1,1), which is estimated by maximum likelihood, and includes only one lag both in the ARCH term (last period's volatility) and in the GARCH term (last period's variance). In more recent studies, different versions of the GARCH model have been used by several authors in the study of calendar effects (Choudry, 2000 and Chen et al., 2001). Choudry (2000) applies the GARCH model to a research on a day of the week effect in seven East Asian countries. By analyzing the estimated coefficients of the dummy variables and coefficients, he finds significant effects in three of those countries, and also in ARCH and GARCH terms.

In this paper, we aim to detect calendar effects in the following cases, with no *a priori* restriction on which periods those effect might be revealed: (i) month effects and (ii) day-of-the-week effects. We first address the problem of heteroskedasticity by regressing the models in Stata 10 software, with the option of robust standard errors switched on. The non-normality of the data is tackled by applying the non-parametric bootstrap approach, with 1000 replications for each model regression, and then using the standard errors and confidence intervals resulting from the distribution of the estimated coefficients. The use of the bootstrap approach in the study of calendar effects has been applied before (Sullivan, Timmermann and White, 2001). We perform a test of ARCH effects on our data, and confirm that it is present in the data for all seventeen countries. Therefore we re-estimate all our models with the GARCH(1,1) approach.

5. RESULTS

Considering the discussion in the previous section, we use the following procedures to detect calendar effects in the daily returns of all seventeen stock market indexes, in the period 1994 to 2007. All statistical tests and estimations are computed in Stata 10 software

5.1. OLS regressions

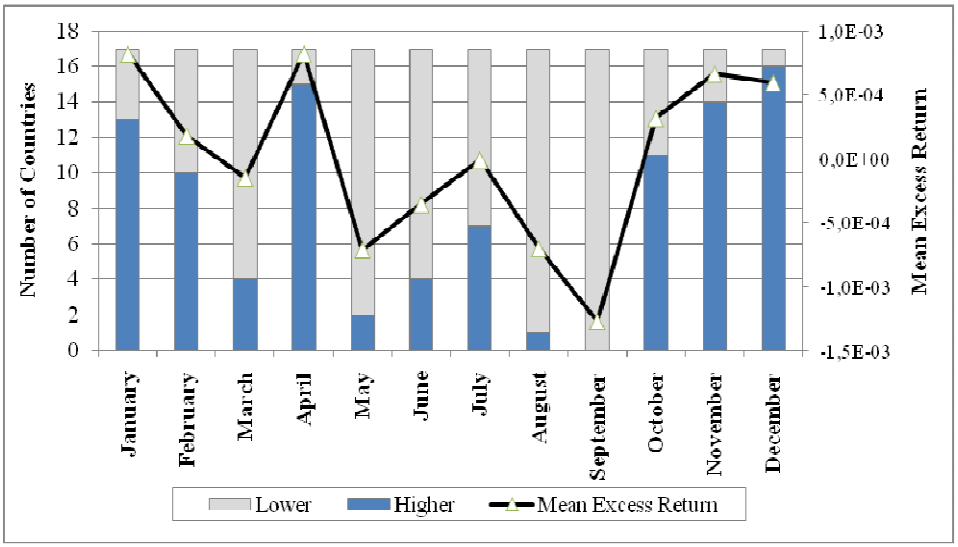
We start by computing individual OLS regressions for each of the seventeen countries, using model (4) for day of the week effects and (5) for month effects. Therefore, we perform a total of 85 regressions for day of the week effects and 204 regressions for month effects. Given the non-normality of the data, all OLS regressions were computed with robust standard errors. The results for the β_i coefficients are presented in Table 2.

INSERT TABLE 2

One of the important critiques to previous studies of calendar effects, is that it may well be exclusively a result of data mining (Sullivan, Timmermann and White, 2001), based on the idea that if we squeeze a particular sample or time series hard enough, all sorts of regularities may start to appear. To control for data-snooping, Cooper, McConnell and Octchinnikov (2006) propose a randomized-bootstrap procedure, and Schwert (2003) suggests the use of data from other countries.

Our first results, in Table 2, are not immune to the critique of Sullivan, Timmermann and White (2001). In day of the week effects, the number of significant coefficients is 2 at the 1% significance level, 6 at the 5% and 5 at the 10%. Given that we compute 85 regressions, the number of significant coefficients that we might expect to find, in random data, would be around 1 (at 1%), around 4 (at 5%) and around 8 (at 10%). So, our overall results for day of the week effects are not very different from those we might expect to obtain, in a randomly constructed sample. For month effects, the number of significant coefficients is 6 (at 1%), 10 (at 5%) and 18 (at 10%). As we have 205 regressions, again, our global results are similar to the number of significant coefficients we might expect to find in random data at 5% (10) and 10% (20). However, at the 1%, we expect to find around 2 significant coefficients in random data, but we have 6. Also, as becomes apparent in Figures 2 and 3 below, there is some concentration of the significant coefficients in specific months and days of the week, and this also justifies further investigation.

Figure 2. Number of Countries with Higher and Lower Daily Returns by Month



What are the detected month effects? First, January returns tend to be higher than in other months, but are only significant, at 5%, in four (Hungary, Iceland, Poland and Portugal) of the seventeen countries. Second, for most countries, daily returns tend also to be higher in April (but not significant at 5% level) and in the last three months of the year, October, November and December (but only significant in one or two countries). Third, all countries show lower returns than average in August (except Iceland, where it is one of the stronger months) and September. The stronger across-the-board month effect in European countries is clearly September, with significant negative excess returns for two countries at 1%, four countries at 5%, and another four countries at 10%.

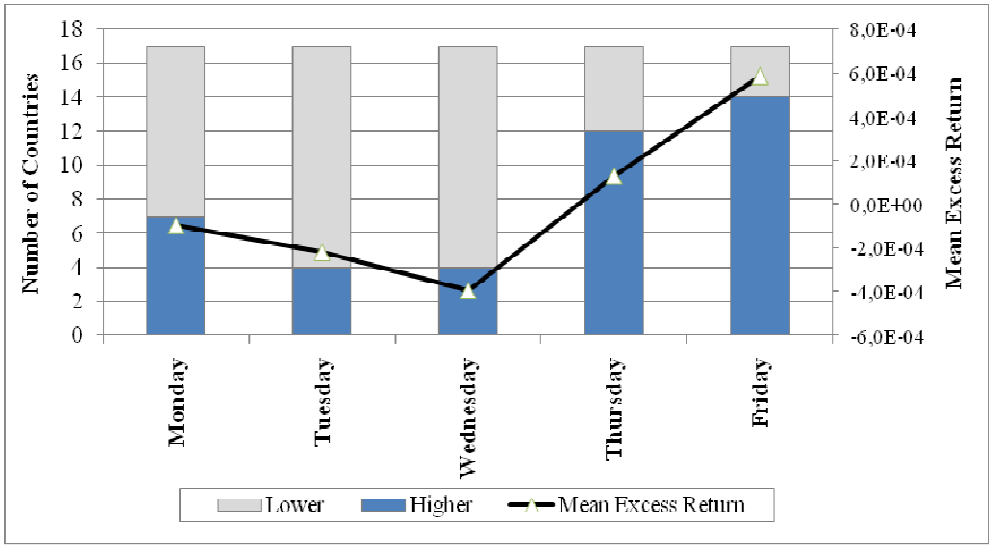
The fact that all seventeen countries show negative excess returns in September needs to be addressed. If no month effects existed, the probability that in any given month excess returns are negative is 0.5, for any country. If we assume independence between the seventeen stock

markets, the probability that all countries have negative excess returns in the same month is 0.5^{12} = 0.0244%. As there are 12 months, the probability of getting this result in our study would be small, i.e., 0.0244% x 12 = 0.293%. However, we know that most of these stock markets are strongly correlated with each other, and so the independence assumption does not hold. Therefore, contemporaneous movements in all stock markets are expected to happen, and are not necessarily evidence of an investor behavior based month effect.

Taken together, the lower daily returns on August and September justify further investigation of the reasons behind that behavior. Although that is behind the scope of this paper, we propose that the part of the answer might possibly be related to changing behavior of both personal and institutional investors (postponing investment decisions?) related to the enjoyment of summer holidays. As we collect no proof of this, readers should consider this only as suggestion for further investigation.

Figure 3 illustrates the results for day of the week effects.

Figure 3. Number of Countries with Higher and Lower Daily Returns by Weekday



Overall, the individual coefficients for daily excess returns are not significant, but Figure 3 shows that mean excess returns tend to be negative and decreasing in the first three days of the week, in most countries, while excess returns are positive in Thursdays and Fridays, in most countries. The last day of the week seems to be, in average, the strongest day of the week. However, only five countries (Finland, Greece, Iceland, Ireland and Norway) have excess mean returns on Fridays which are significantly different from other weekday returns, at the 5% level.

5.2. Bootstrap Approach and GARCH Model

To test the robustness of the results presented in section 5.1, bootstrapping can be applied to the OLS regressions, as Sullivan, Timmermann and White (2001) propose. These authors warn against the dangers of data mining in the study of calendar effects, claiming that most of the obtained results are only "chimeras" and the product of data mining. We use the bootstrap command in Stata 10, with 1000 replications, to all the OLS regressions. This methodology executes the OLS regressions 1000 times, bootstrapping the statistics of the β_i, by re-sampling observations (with replacement) from the data. Because this is a non-parametric approach, it is not affected by the non-normality of the data.

Additionally, because we know that there are ARCH effects in our sample daily returns, we re-estimate the models for month and weekday effects using the ARCH command in Stata 10, allowing for a GARCH(1,1) process, by means of maximum likelihood. In all estimations, both the ARCH term and the GARCH are significant at

the 1% level, confirming that periods of high and low volatility in the residuals are grouped.

As additional evidence, we compute a Kruskal-Wallis equality-of-populations rank test, by dividing daily returns in different groups (months or days of the week), and determining if the null hypothesis that all groups come from the same population. This is similar to a one-way analysis of variance with the data replaced by their ranks. Because this a non-parametric test, it does not depend on the data being normally distributed. The null hypothesis (no effects) is rejected in weekday effects for Greece, Iceland and Poland, and in month effects for Austria, Iceland and Portugal.

Except for the Kruskal-Wallis test, we do not show the obtained results for the bootstrap approach and GARCH (1,1) directly, to avoid burdensome tables. We choose to report exclusively, in Table 3, which day of the week effects remain significant, at a level of 5%, for each of the statistical methodologies applied.

INSERT TABLE 3

In day of the week effects, the following days/countries are significant in all statistical methodologies: (i) positive Fridays in Greece, Iceland, Ireland and Norway; (ii) positive Tuesdays in Germany, and (iii) negative Mondays in Iceland. The GARCH model additionally uncovers: (i) negative Tuesdays in Poland and Greece; and, (ii) negative Mondays in Greece. Overall, the two countries who reveal a stronger day of the week effect are clearly Iceland and Greece, consistently with the weekend effect extensively documented in the literature. Nevertheless, our overall results are very clear in demonstrating that there is no across-the-board weekend effect in European stock markets, as in most countries it is non-existing (including Austria, Denmark, France, Hungary, Italy, Poland, Portugal, Spain, Switzerland and United Kingdom).

Table 4 reports on month effects detected by all the statistical methodologies applied.

INSERT TABLE 4

As in day of the week effects, we find no overall effect covering the full spectrum of countries under study, as there are no month effects in Denmark, Finland, Ireland and UK, and the effects initially detected in some countries do not resist to more robust statistical methodologies, such as France, Hungary, Italy, Netherlands, Norway, Poland, Spain and Switzerland. On the other hand, the stronger month effects include: (i) Iceland has positive excess returns in August, and negative in October; (ii) Austria has positive excess returns in February and negative in September; (iii) Portugal has positive excess returns in January, and negative in May; and, (iv) Greece has negative excess returns in June.

In our sample, Iceland is clearly the country with stronger calendar effects, revealed both on days of the week and months.

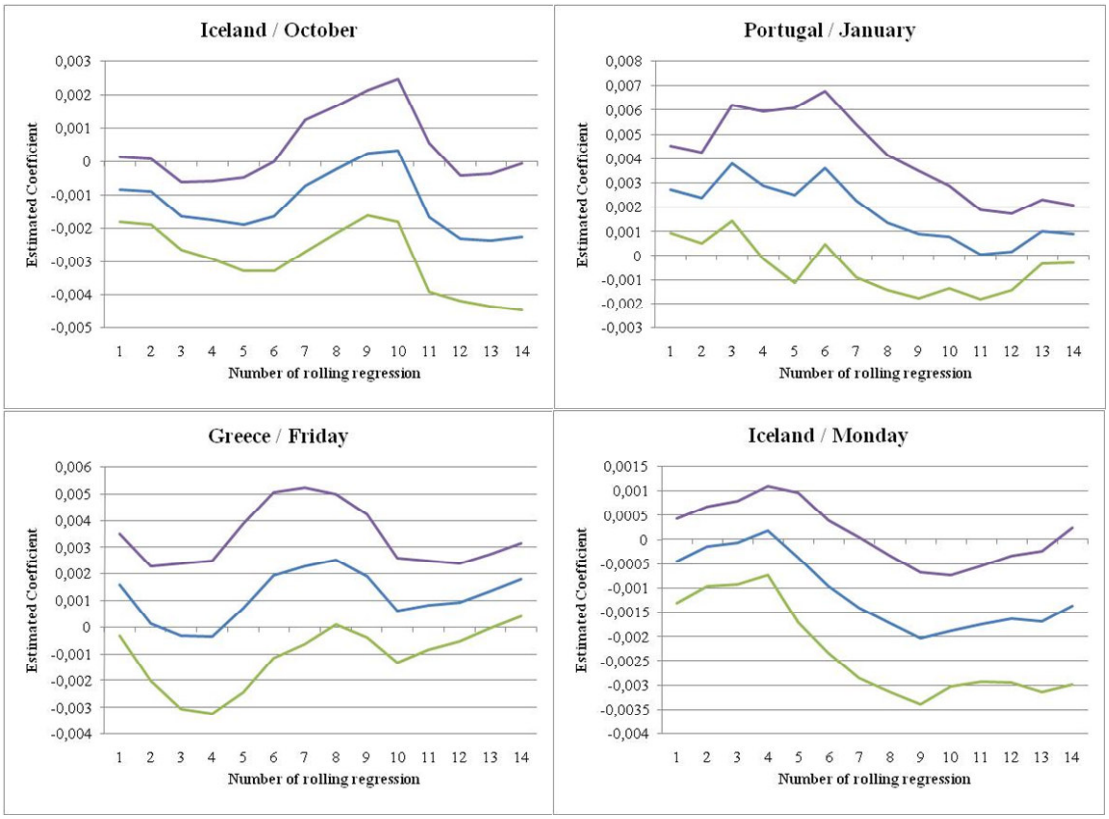
5.3. Time-stability of day of the week and month effects

Our sample covers a period of fourteen years, between 1994 and 2007. As an additional robustness check, we investigate if the calendar effects detected are stable, over the whole period under analysis. It may be the case that the global result is affected by short-run phenomena, in only a few of the years under study. The purpose of this section is to shed some light on this.

To test the time stability of the coefficients, we compute rolling window OLS regressions on the most significant coefficients detected in day of the week and monthly effects, both for positive and negative excess returns. We choose four cases: (i) for positive day of the week: Greece/Friday; (ii) for negative day of the week: Iceland/Monday; (iii) for positive month: Portugal/ January; (iv) and for negative month: Iceland/October. In the rolling window OLS regressions, we use a window size of 1000 observations (roughly equivalent to four or five years of observations), and a

step size of 200 observations. This means that the first regression uses observations [1;1000], the second regression uses observations [201;1200] and so on. Given the number of observations available, we compute 14 rolling regressions for each coefficient. In Figure 4, we show the evolution of the β_i coefficients for these four strong calendar effects, and also the upper and lower bounds on its 95% confidence interval.

Figure 4. Coefficients for Calendar Effects in Rolling-Window Regressions



In all four cases, the coefficient that captures the calendar effect fluctuates significantly, as the windows of observations evolve. In the case of Iceland/October, the coefficient becomes positive both in the 9th and 10th regression, decreasing again sharply after that. In more than half the regressions, the upper bound on the 95% confidence interval is a positive return. In the case of Portugal, the January effect seems to be due mainly to the observations in the first years in the sample, as the effect wears out in the last eight windows of observations. In the case of Greece, the Friday effect changes radically from window to window, with periods of higher returns shortly followed by periods of lower returns. It is only in the last window that the lower bound of the confidence interval is clearly positive. Finally, the Monday effect in Iceland seems to be only a recent phenomenon, as it did not exist in the first windows of observations.

Taken together with the results of the previous sections, this evidence of high instability of the calendar effects coefficients casts further doubt on the significance of the month and day of the week effects in European stock markets.

6. CONCLUSIONS

There is an extensive body of research documenting day of the week and month of the year effects, particularly in US markets, although international evidence is constantly growing, but with mixed results. Some studies reveal that calendar effects that were strong in the 1970s, 1980s and 1990s, have become weaker in more recent years, both in developing and developed

markets. Is it the case that markets are becoming more efficient and calendar effects are being arbitraged away, or is it the case that more recent and more powerful statistical methodologies no longer detect those effects, casting doubt on previous studies?

There is more than one reason why the findings of previous studies may need to be re-assessed. First, as we discuss in section 4 of this paper, model specifications may have been inadequate for the detection of calendar effects, in particular the use of *t*-tests on models with multiple dummy variables. Second, given the non-normality and other problems in the data, the use of linear regressions may have lead to the incorrect rejection of the null hypothesis of no calendar effects. Third, a large number of studies cover only one country or only a few countries, causing some authors to assign a disproportionate importance to a specific detected effect, which may well be a spurious result delivered by intensive data mining. Studies covering a large number of countries help authors to maintain a skeptical point-of-view on country-specific effects, and these types of studies are still in minority. Fourth, markets develop over time, and widespread knowledge of some types of calendar effects may have lead to their exploitation by arbitrageurs, thus eroding such effects. So, investigators need to look at new and more recent data, frequently enough, and to keep checking the time-stability of previous results.

In our study, we apply robust statistical methodologies and consider only the calendar effects that are significant under all alternative methodologies. Our main findings are the following.

First, we find no strong convincing evidence of an across-the-board calendar effect in West and Central European countries. In particular, there are no statistically significant across-the-board January effects or weekend effects. European countries seem to be mostly immune to day of the week effects, even though the group of seventeen countries, taken together, does tend to show higher daily returns on Thursdays and Fridays, and lower in Mondays in Tuesdays. If any, the only across-the-board calendar effect that warrants further investigation is the general tendency for lower returns in the holiday months of August and September. All the calendar effects are basically country-specific.

Second, the number of significant coefficients we detect is very similar to the number we would expect to find in random data. Even though there is some concentration on specific months / days of the week, our results are not immune to the critique that the calendar effects we detect are exclusively a result from intensive data mining. This skeptical view is reinforced by the fact that the statistically stronger calendar effects are not stable over time. In fact, when we use different sub-samples of the data, the stronger calendar effects that we detect in the whole sample, are not detected in several of those sub-samples. So, some of the apparently stronger calendar effects may well not be the result of economic motives, market microstructure, or behavioral traits, but may rather be only lucky snapshots of capricious movements in the stock market indexes.

Finally, we suggest as avenues for further research, the following. First, some preliminary results we obtain on day of the month effects, which we do not report, signal that this may be the type of calendar effect more relevant in European countries, justifying specific research. Second, the use of data on firms instead of indexes, allows the study of calendar effects by firm characteristics. Third, we need more studies using broader sets of countries, to determine if calendar effects are across-the-board or only country-specific. Fourth, a closer look at the low August / September returns in Europe, and the study of the reasons behind that effect, if it is confirmed. Fifth, there are several alternative variants of the GARCH model, like TGARCH and IGARCH; which one fits the data better? Sixth, we need to improve on the microeconomics of calendar effects. We should strive to find the true economic (or behavioral) rationale behind calendar effects. We need to a better understanding on why calendar effects are expected to exist, and under what circumstances.

REFERENCES

AGGARWAL, R., R. RAO and T. HIRAKI, "Regularities in Tokyo Stock Exchange Security Returns: P/E, Size and Seasonal Influences", *Journal of Financial Research*, 13, 1990, pp. 49-263.

BASHER, S. and P. SADORSKY "Day of the Week Effects in Emerging Stock Markets", *Applied Economic Letters*, 13, 2006, pp. 621-628.

BOLLERSLEV, T. "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31, 1986, pp.307-327.

BROOKS, C. and G. PERSAND "Seasonality in Southeast Asian Stock Markets: Some New Evidence on Day of the Week Effects, *Applied Economic Letters*, 8, 2001, pp. 155-158.

BROWN, P., D. KEIM, A. KLEIDON and T. MARSH "New Evidence on the Nature of Size-related Anomalies in Stock Prices", *Journal of Financial Economics*, 12, 1983, pp.33-56.

CHANG, E., J. PINEGAR and R. RAVICHANDRAN "International Evidence on the Robustness of the Day of the Week Effect", *Journal of Financial and Quantitative Analysis*, 28, 1993, pp 497-513.

CHEN, G., C. KWOK and O. RUI, "The Day of the Week Regularity in the Stock Markets of China", *Journal of Multinational Financial Management*, 11, 2001, pp. 139-163.

CHOUDHRY, T. "Day of the Week Effect in Emerging Asian Stock Markets: Evidence from the GARGH Model", *Journal of Financial Economics*, 10, 2000, pp. 235-242.

CHUKWUOGOR-NDU, C. "Stock Market Returns Analysis, Day of the Week Effect, Volatility of Returns: Evidence from European Financial Markets 1997-2004", *International Research Journal of Finance and Economics*, 1, 2006, pp. 112-124

CONDOYANNI, L., J. O'HANLON, and C. WARD "Day of the Week Effect on the Stock Returns: International Evidence", *Journal of Business Finance and Accounting*, 14, 1987, pp. 159-174.

CONNOLLY, R. "An Examination of the Robustness of the Weekend Effect", *Journal of Financial and Quantitative Analysis*, 24, 1989, pp. 133-169.

Connolly, R. "A Posterior Odds Analysis of the Weekend Effect", *Journal of Econometrics*, 49, 1991, pp. 51-104.

COOPER, M., J. McCONNELL and A. OCTCHINNIKOV "The Other January Effect", *Journal of Financial Economics*, 82, 2006, pp. 315-341.

CROSS, F. "The Behavior of Stock Prices on Fridays and Mondays", *Financial Analysts Journal*, 29, 1973, pp. 67-69.

DYL, E. "Capital Gains Taxation and Year-end Stock Market Behavior, *Journal of Finance*, 32, 1977, pp. 165-175.

DUBOIS, M. and P. LOUVET "The Day of the Week Effect: The International Evidence", *Journal of Banking and Finance*, 20, 1996, pp. 1463-1485.

ENGLE, R. "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation", *Econometrica*, 50, 1982, pp. 987-1008.

FAMA, E. "The Behavior of Stock Prices", *Journal of Business*, 38, 1965, pp. 34-105.

FRENCH, K. "Stock Returns and the Weekend Effect", *Journal of Financial Economics*, 8, 1980. pp. 55-80.

GIBBONS, M. and P. HESS "Day of the Week Effects and Asset Returns", *Journal of Business*, 54, 1981, pp. 579-596.

GULTEKIN, M. and B. GULTEKIN "Stock Market Seasonality: International Evidence", *Journal of Financial Economics*, 12, 1983, pp. 469-481.

HAUGEN, R. and P. JORION "The January Effect: Still There After all These Years", *Financial Analysts Journal*, 52, 1996, pp.27-31.

HO, Y., "Stock Return Seasonalities in Asia Pacific Markets", *Journal of International Financial Management and Accounting*, 2, 1990, pp. 47-77.

HUI, T. "Day of the Week Effects in US and Asia-Pacific Stock Markets During the Asian Financial Crisis: a Non-parametric Approach", *The International Journal of Management Science*, 33, 2005, pp. 277-282.

JAFFE, J. and R. WESTERFIELD "The Weekend Effect in Stock Returns: the International Evidence", *Journal of Finance*, 41, 1985a, pp. 433-454.

JAFFE, J. and R. WESTERFIELD "Patterns in Japanese Common Stock Returns: The International Evidence", *Journal of Financial and Quantitative Analysis*, 20, 1985b, pp. 243-260.

KAMARA, A. "New Evidence on the Monday Seasonal in Stock Returns", *Journal of Business*, 70, 1997, pp. 63-84.

KEIM, D. "Size Related Anomalies and Stock Return Seasonality: Further Empirical Evidence", *Journal of Financial Economics*, 12, 1983, pp. 12-32.

KEIM, D. and R. STAMBAUGH "A Further Investigation of the Weekend Effect in Stock Returns. *Journal of Finance*, 39, 1984, pp. 819-835.

KIM, D. and S. KON "Alternative Models for the Conditional Heterocedasticity of Stock Returns", *Journal of Business*, 67, 1994, pp. 563-598.

KOHERS, G., N. KOHERS, V. PANDEY and T. KOHERS "The Disappearing Day of the Week Effect in the World's Largest Equity Markets", *Applied Economic Letters*, 11, 2004, pp. 167-171.

LAU, A., H. LAU and J. WINGENDER "The Distribution of Stock Returns: New Evidence Against the Stable Model", *Journal of Business and Economic Statistics*, 8, 1990, pp. 217-223.

MABERLY, E. and D. WAGGONER, "Closing the Question on the Continuation of the Turn of the Month Effects: Evidence from the S&P 500 Index Future Contracts", Federal Reserve Bank of Atlanta, 2000.

MENYAH, K. "New Evidence on the Impact of Size and Taxation on the Seasonality of UK Equity Returns", *Review of Financial Economics*, 8, 1999, pp.11-25.

OFFICER, R. "Seasonality in Australian Capital Markets: Market Efficiency and Empirical Issues", *Journal of Financial Economics*, 2, 1975, pp. 29-52.

REINGANUM, M. "The Anomalous Stock Market Behavior of Small Firms in January: Empirical Tests for Tax-Loss Effects", *Journal of Financial Economics*, 12, 1983, pp. 89-104.

ROGALSKI, R "New Findings Regarding Day of the Week over Trading and Non-trading Periods: a Note, *Journal of Finance*, 39, 1984, pp. 1903-1614.

ROSENBERG, M "The Monthly Effect in Stock Returns and Conditional Heteroskedasticity", *American Economist*, 48, 2004, pp. 67-73.

ROZEFF, M. and W. KINNEY, "Capital Market Seasonality: The Case of Stock returns", *Journal of Financial Economics*, 3, 1976, pp. 379-402.

RUBINSTEIN, M. "Rational Markets: Yes or No? The Affirmative Case", *Financial Analysts Journal*, 57, 2001, pp. 15-29.

SCHWERT, G. "Anomalies and Market Efficiency", in G. Constantinides et al., *Handbook of the Economics of Finance*, 2001, North Holland, Amsterdam.

SMIRLOCK, M. and L. STARKS "Day of the Week and Intraday Effects in Stock Returns", *Journal of Financial Economics*, 17, 1986, pp. 197-210.

STEELY, J. "A Note on Information Seasonality and the Disappearance of the weekend Effect in UK Stock Market", *Journal of Banking and Finance*, 25, 2001, pp. 1941-1956.

SULLIVAN, R., A. TIMMERMANN, and H. WHITE, "Dangers of Data mining: The Case of Calendar Effects in Stock Returns", *Journal of Econometrics*, 105, 2001, pp. 249-286.

TONCHEV, D. and T. KIM "Calendar Effects in Eastern European Financial Markets: Evidence from the Czech Republic, Slovakia and Slovenia", *Applied Financial Economics*, 14, 2004, pp. 1035-1043.

WOOLDRIGE, J. *Introductory Econometrics: A Modern Approach*, South Western College Publishing, USA, 2003.

Table 1. Descriptive Statistics for Daily Returns (1994-2007)

Country	Index	Observa-tions	Mean	Standard Deviation	Standard Error of Mean	Maximum	Minimum	10th Percentile	90th Percentile	Skewness	Kurtosis
Austria	**ATX**	3318	0.0003136	0.0104380	0.0001812	0.05135	-0.08427	-0.01098	0.01185	-0.7388	7.8385
Denmark	**OMXC20**	3418	0.0003746	0.0108016	0.0001848	0.04970	-0.06258	-0.01211	0.01278	-0.3467	5.3195
Finland	**OMXHPI**	3404	0.0004392	0.0191563	0.0003283	0.14563	-0.17425	-0.02056	0.02024	-0.5260	10.797
France	**CAC40**	3462	0.0001716	0.0132970	0.0002260	0.07002	-0.07678	-0.01522	0.01475	-0.1227	5.7002
Germany	**DAX**	3465	0.0002944	0.0144468	0.0002454	0.07270	-0.09431	-0.01644	0.01623	-0.2084	6.1468
Greece	**ASE**	3595	0.0004321	0.0151520	0.0002527	0.07661	-0.09615	-0.01527	0.01640	-0.0804	7.5173
Hungary	**BUX**	3391	0.0008001	0.0169282	0.0002907	0.13616	-0.18033	-0.01604	0.01853	-0.8530	16.274
Iceland	**OMXIPI**	3353	0.0008868	0.0076047	0.0001313	0.06970	-0.07053	-0.00650	0.00872	-0.3603	11.543
Ireland	**ISEQ**	3421	0.0003270	0.0100765	0.0001723	0.05835	-0.06124	-0.01082	0.01100	-0.4229	6.8666
Italy	**MIBTEL**	3462	0.0001890	0.0123341	0.0002096	0.06832	-0.10648	-0.01411	0.01416	-0.4008	7.2839
Netherlands	**AEX**	3498	0.0002164	0.0134391	0.0002272	0.09517	-0.07531	-0.01409	0.01395	-0.1519	7.6900
Norway	**OSEAX**	3428	0.0004033	0.0114155	0.0001950	0.08016	-0.06352	-0.01254	0.01286	-0.4416	7.1434
Poland	**WIG**	3279	0.0002076	0.0173266	0.0003026	0.07893	-0.11344	-0.01793	0.01931	-0.5573	8.4188
Portugal	**PSI20**	3389	0.0002643	0.0098439	0.0001691	0.06941	-0.09590	-0.00986	0.00986	-0.7989	11.346
Spain	**IBEX**	3398	0.0003309	0.0128992	0.0002213	0.06323	-0.07339	-0.01483	0.01483	-0.2756	5.8470
Switzerland	**SMI**	3435	0.0002231	0.0115262	0.0001967	0.07462	-0.07331	-0.01256	0.01273	-0.2068	7.3045
UK	**FTSE**	3595	0.0001556	0.0104980	0.0001751	0.05903	-0.05885	-0.01189	0.01168	-0.2053	6.0160

Table 2 (a). Differences in Mean Returns: Month Effects and Day of the Week Effects (1994-2007)

	Austria	Denmark	Finland	France	Germany	Greece	Hungary	Iceland	Ireland
Month Effects									
January (m1)	0.0008370	0.0006718	-0.0000672	0.0004758	0.0004642	0.0012756	0.0030919***	0.0014591***	0.0005312
February (m2)	0.0012117*	0.0000033	-0.0005953	-0.0003238	-0.0001027	0.0000724	-0.0006621	0.0006754	0.0001269
March (m3)	-0.0001552	-0.0004547	-0.0001447	0.0005065	0.0000163	-0.0003734	-0.0005116	-0.0001510	-0.0001220
April (m4)	0.0011506*	-0.0000662	0.0019942	0.0013475	0.0008637	0.0012275	0.0007798	0.0000061	0.0003006
May (m5)	-0.0002729	0.0000892	-0.0016483	-0.0009067	-0.0005842	-0.0003769	-0.0014368	-0.0005309	-0.000579
June (m6)	-0.0002558	-0.0002900	-0.0003972	-0.0002228	0.0003801	-0.0018381**	0.0001130	-0.0002897	-0.0003666
July (m7)	-0.0002316	0.0004002	-0.0006373	-0.0005493	0.0000795	0.0011410	0.0006777	-0.0000392	-0.0005222
August (m8)	-0.0009997	-0.0000412	-0.001093	-0.0012453	-0.0015085*	-0.0008802	-0.0010083	0.0012296***	-0.0000880
September (m9)	-0.0014473**	-0.0011481*	-0.0003406	-0.0019372**	-0.0023202***	-0.0002694	-0.0017642*	-0.0003333	-0.0010632*
October (m10)	-0.0009800*	0.0000347	0.0017162	0.0012442	0.0006597	-0.0009277	-0.0001182	-0.0010512**	0.0004554
November (m11)	0.0004969	0.0004069	0.0015542	0.0011597	0.0018291**	0.0003771	-0.0007940	-0.0010134**	0.0006287
December (m12)	0.0010656*	0.0004432	-0.0004157	0.0005599	0.0004149	0.0006386	0.0019474*	0.0001833	0.0008309
Weekday Effects									
Monday (d1)	0.0003267	0.0002133	0.0003584	-0.0001953	0.0001200	-0.0011393*	0.0010105	-0.0010800***	-0.0007305*
Tuesday (d2)	-0.0000216	-0.0001378	-0.0013173	0.0001918	0.0012919**	-0.0009483	-0.0002630	-0.0002404	-0.0000191
Wednesday (d3)	-0.0005173	0.0000549	-0.0010811	-0.0003940	-0.0002125	0.0002494	-0.0006958	-0.0003118	-0.0002824
Thursday (d4)	0.0007971*	0.0001466	0.0004348	0.0000647	-0.0006418	0.0002865	-0.0009669	0.0004697	0.0001105
Friday (d5)	-0.0005772	-0.0002778	0.0016806**	0.0003375	-0.0005415	0.0015450**	0.0009558	0.0011560***	0.0008772**

Notes: *Denotes significance at the 0.1 level. **Denotes significance at the 0.05 level. ***Denotes significance at the 0.01 level.

Table 2 (b). Differences in Mean Returns: Month Effects and Day of the Week Effects (1994-2007)

	Italy	Netherlands	Norway	Poland	Portugal	Spain	Switzerland	UK
Month Effect								
January (m1)	0.0012869*	-0.0003875	0.0006695	0.0023601**	0.0018092***	0.0002735	-0.0004143	-0.0004209
February (m2)	-0.0001054	0.0002576	-0.0000159	0.0010675	0.0008848*	0.0008105	-0.0002606	0.0000201
March (m3)	0.0005654	-0.0004310	-0.0000091	-0.0006989	-0.0000436	-0.0007650	0.0003083	-0.0000293
April (m4)	0.0009929	0.0013887	0.0010869	0.0009733	-0.0006660	0.0009011	0.0009134	0.0007768
May (m5)	-0.0013138*	-0.0002559	-0.0002246	-0.0020804*	-0.0011030*	-0.0003470	0.0000091	-0.0005743
June (m6)	-0.0008792	-0.0001735	0.0000732	-0.0000233	-0.0008001	-0.0006307	0.0000303	-0.0005211
July (m7)	-0.0001825	-0.0000570	0.0003118	0.0007452	-0.0002300	-0.0007125	-0.0004382	0.0000865
August (m8)	-0.0009048	-0.0007284	-0.0009551	-0.0002820	-0.0009143	-0.0011146	-0.0012106*	-0.0002095
September (m9)	-0.0012298*	-0.0020503**	-0.0018248***	-0.0016862	-0.0012040**	-0.0009877	-0.0010732	-0.0009836
October (m10)	-0.0001614	0.0008677	0.0004716	-0.0002076	0.0008932	0.0008898	0.0007582	0.0008577
November (m11)	0.0014259*	0.0011520	0.0002235	-0.0002220	0.0007718	0.0016059**	0.0012480*	0.0004334
December (m12)	0.0007903	0.0006395	0.0003714	0.0008053	0.0006780	0.0003548	0.0002129	0.0005978
Weekday Effect								
Monday (d1)	-0.0005973	0.0005714	-0.0001770	0.0008422	-0.0003908	-0.0006186	-0.0000761	-0.0000452
Tuesday (d2)	0.0000217	-0.0001557	-0.0004651	-0.0016038**	-0.0000050	0.0002262	-0.0001401	-0.0001057
Wednesday (d3)	-0.0003295	-0.0004713	-0.0007308	-0.0010887	0.0001411	-0.0005240	0.0000865	-0.0005845
Thursday (d4)	0.0002900	-0.0003432	0.0004058	0.0012438*	-0.0001887	0.0000621	-0.0001172	0.0001690
Friday (d5)	0.0006044	0.0004175	0.0009807**	0.0007086	0.0004365	0.0008570*	0.0002453	0.0005648

Notes: *Denotes significance at the 0.1 level. **Denotes significance at the 0.05 level. ***Denotes significance at the 0.01 level.

Table 3. Weekday Effects by Country (At Significance Level: 5%)

Country	Index	Higher Returns than Other Weekdays			Lower Returns than Other Weekdays			Kruskal – Wallis Test
		Regression	Bootstrap /Regression	GARCH Model	Regression	Bootstrap /Regression	GARCH Model	
Austria	**ATX**	-	-	-	-	-	-	0.3485
Denmark	**OMXC20**	-	-	-	-	-	-	0.9661
Finland	**OMXHPI**	Friday*	Friday*	-	-	-	-	0.1232
France	**CAC40**	-	-	-	-	-	-	0.9454
Germany	**DAX**	Tuesday*	Tuesday*	Tuesday*	-	-	-	0.1189
Greece	**ASE**	Friday*	Friday**	Friday**	-	-	Mon**, Tue**	0.0005**
Hungary	**BUX**	-	-	-	-	-	-	0.2635
Iceland	**OMXIPI**	Friday**	Friday**	Friday**	Monday*	Monday**	Monday**	0.0001**
Ireland	**ISEQ**	Friday*	Friday*	Friday*	-	-	-	0.1305
Italy	**MIBTEL**	-	-	-	-	-	-	0.3035
Netherlands	**AEX**	-	-	-	-	-	-	0.2455
Norway	**OSEAX**	Friday*	Friday*	Friday*	-	-	-	0.0507
Poland	**WIG**	-	-	-	Tuesday*	-	Tuesday*	0.0302*
Portugal	**PSI20**	-	-	-	-	-	-	0.8178
Spain	**IBEX**	-	-	-	-	-	-	0.2686
Switzerland	**SMI**	-	-	-	-	-	-	0.9568
UK	**FTSE**	-	-	-	-	-	-	0.2400

Notes: *Denotes significance at the 0.05 level. **Denotes significance at the 0.01 level.

Table 4. Month Effects by Country (At Significance Level: 5%)

Country	Index	Higher Returns than Other Months			Lower Returns than Other Months			Kruskal – Wallis Test
		Regression	Bootstrap /Regression	GARCH Model	Regression	Bootstrap /Regression	GARCH Model	
Austria	**ATX**	-	Feb*, Apr*	Feb*	Sep*	Sep*	Sep**	0.0383*
Denmark	**OMXC20**	-	-	-	-	-	-	0.9791
Finland	**OMXHPI**	-	-	-	-	-	-	0.7046
France	**CAC40**	-	-	-	Sep*	-	-	0.4088
Germany	**DAX**	Nov*	Nov*	-	Sep**	Sep*	-	0.5714
Greece	**ASE**	-	-	-	Jun*	Jun*	Jun*	0.2311
Hungary	**BUX**	Jan**	Jan*	-	-	-	-	0.0744
Iceland	**OMXIPI**	Jan**, Aug**	Jan**, Aug**	Aug**	Oct*, Nov*	Oct*, Nov*	Oct*	0.0012**
Ireland	**ISEQ**	-	-	-	-	-	-	0.6429
Italy	**MIBTEL**	-	Nov*	-	-	-	-	0.0847
Netherlands	**AEX**	-	-	-	Sep**	Sep*	-	0.5963
Norway	**OSEAX**	-	-	-	Sep**	Sep*	-	0.5902
Poland	**WIG**	Jan*	Jan*	-	-	-	-	0.1285
Portugal	**PSI20**	Jan**	Jan**	Jan**	Sep*	May*	May*	0.0180*
Spain	**IBEX**	Nov*	Nov*	-	-	-	-	0.3301
Switzerland	**SMI**	-	Nov*	-	-	-	-	0.6630
UK	**FTSE**	-	-	-	-	-	-	0.7609

Notes: *Denotes significance at the 0.05 level. **Denotes significance at the 0.01 level.

EXCHANGE RATE AND INTEREST RATE VOLATILITY IN A TARGET ZONE: THE PORTUGUESE CASE

ANTÓNIO PORTUGAL DUARTE
E-mail: portugal@fe.uc.pt
JOÃO SOUSA ANDRADE
E-mail: jasa@fe.uc.pt
ADELAIDE DUARTE[1]
E-mail: maduarte@fe.uc.pt
Faculdade de Economia
Universidade de Coimbra

Abstract

This work examines the participation of the Portuguese economy in the ERM of the EMS based on the main predictions of the target zone literature. The exchange rate distribution reveals that the majority of the observations lie close to the central parity, thus rejecting one of the key predictions of the Krugman (1991) model. Using a M-GARCH model however we confirm that there is a trade-off between exchange rate volatility and interest rates differential volatility. These results express the increased credibility of the monetary policy, due manly to the modernisation of the banking and financial system and to the progress made in terms of the disinflation process under a target zone. In accordance to these results we can say that the participation of the Portuguese escudo in a target zone was crucial to create the conditions of stability, credibility and confidence necessary for the adoption of a single currency.

Key Words: Credibility, M-GARCH, volatility and target zones.
Thematic Area: Monetary and Financial Economics.

1 – INTRODUCTION

The target zones are one of the exchange rates regimes that have been more often used throughout history, from the classical gold standard until the current exchange rate arrangement in the ERM II of the European Union. However, the traditional debate concerning the choice of appropriate exchange rate regime has neglected this fact, focusing its attention on the choice between fixed or flexible exchange rates[2].

The functioning of the ERM of the EMS remains one of the best examples of a target zone regime. From March 13th, 1979 the majority of the currencies fluctuated within a band of ±2,25% built around a central parity reference. The exchange rates could float more or less freely within an exchange rate band, committing the respective central banks to intervene whenever the target zone was threatened, which happened whenever the exchange rate reached one of the pre-established limits.

After some initial theoretical research on the functioning of exchange rate target zones, with major contributions from McKinnon (1982, 1984), Williamson (1985), Williamson and Miller (1987) and Dumas (1989), the literature on the subject has known a revival with the introduction by Krugman (1991) of the first model of a nominal exchange rate target zone in continuum time

[1] The authors thank Pedro Bação and Marta Simões for their helpful comments. The usual disclaimer applies.
[2] See Frankel (1999), Reinhart (2000), Fischer (2001), Calvo and Reinhart (2002), Bordo (2003), Reinhart and Rogoff (2004) and Levy-Yeyati and Sturzenegger (2005).

with rational expectations and flexible prices, using the structure of regulated Brownian motion[3].

The basic target zone model has quite interesting implications for the behaviour of the exchange rate and the interest rates differential. The model predicts that the statistical distribution of the exchange rate must be U-shaped or bimodal, with a greater number of observations lying close to the edges of the band. Additionally, it predicts that the exchange rate is much less variable the closer it is to the edges of the band and that there is a negative relationship between the exchange rate and the interest rate differential. Therefore, we should find evidence of a trade-off between the volatility of the exchange rate and the volatility of the interest rates differential. These predictions can be tested against real world data in order to confirm the validity of the model and the credibility of the bands.

However, although the former model leads to interesting predictions regarding exchange and interest rates differential dynamics these were rejected by empirical analyses of data. Examples can be found in Diebold and Nason (1990), Meese and Rose (1991), Flood, Rose and Mathieson (1991), Svensson (1991a, b), Frankel and Phillips (1992) and Lindberg and Soderlind (1994). These empirical results gave birth to the development of a set of critical extensions that include Svensson (1991c), Froot and Obstfeld (1991), Bertola and Caballero (1992) and Bertola and Svensson (1993), just to name a few.

This paper analyses the functioning of the Portuguese exchange rate target zone under the participation of the Portuguese escudo in the ERM of the EMS. Our main aim is to study the behaviour and volatility of exchange and interest rates based on the predictions of the first generation of target zones models. Our contribution to the literature stems from the fact that we are analysing a currency from the periphery of the system, the Portuguese escudo, while most previous work has focused on the Nordic countries and on fluctuation bands of the ERM of the EMS considered to be more stable and credible.

The paper is structured as follows. Section 2 examines the functioning of the Portuguese exchange rate target zone during the participation of the Portuguese escudo in the ERM of the EMS. Section 3 analyses the distribution and volatility of exchange and interest rates under the Portuguese exchange rate target zone. Section 4 explores the existence of a trade-off between exchange rate volatility and interest rates differential volatility under a M-GARCH model. Finally, Section 5 draws some conclusions.

2 – THE PORTUGUESE EXCHANGE RATE TARGET ZONE

On April 6 1992, the framework of the Portuguese monetary and foreign exchange policy changed when the Portuguese escudo joined the ERM of the EMS. This was made possible by the better convergence of the Portuguese inflation rate with the EU's average level and, in particular, with the values recorded by Germany. The central parity was fixed at 178,735 and 86,9393 escudos for the ECU and for the Deutschmark, respectively, and the Portuguese escudo was allowed to fluctuate within a band of ±6%.

As there was a formal commitment to keep the Portuguese escudo within the band from this date, the credibility of the disinflation policy increased, facilitating the pursuit of the price stability goal. This foreign exchange policy course was maintained until the end of 1998, in spite of the disturbances that affected the EMS. Table 1 summarizes these events, allowing us to identify the main features of the Portuguese exchange rate target zone while the Deutschmark was the reference currency.

[3] Details of this literature can be found in Svensson (1992a), Magnier (1992), Bertola (1994), Garber and Svensson (1995), Pansard (1997) and Kempa and Nelles (1999). For a literature survey see the introduction in Krugman and Miller (1992).

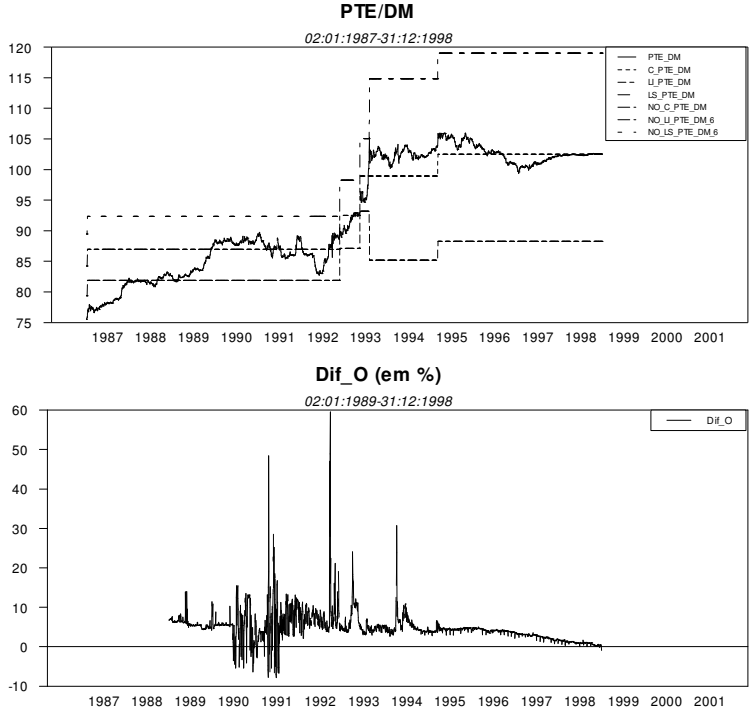
Table 1: Bands for the Portuguese Target Zone (PTE/DM)

Period / Date	Band	PTE/DM (a)		
		Lower Edge (b)	Central Parity	Upper Edge (b)
6 April 1992 (c)	±6%	81,9	86,9393	92,336
23 November 1992 (d)	±6%	87,108	92,488	98,232
13 May 1993 (e)	±6%	93,197	98,9177	105,042
2 August 1993 (f)	±15%	85,179	98,9177	114,811
6 March 1995 (g)	±15%	88,277	102,505	119,033

Source: Banco de Portugal.
Note: (a) Portuguese escudos necessary to buy one Deutschmark.
(b) The lower (upper) edge represents the maximum appreciation (depreciation) permitted to the Portuguese escudo against the Deutschmark.
(c) Membership of the Portuguese escudo of the ERM of the EMS with a bandwidth of ±6%.
(d) Realignment in the EMS with a devaluation of 6% of the Spanish's peseta and the Portuguese escudo.
(e) Realignment in the EMS with a devaluation of 8% of the Spanish's peseta and 6,5% of the Portuguese escudo.
(f) Enlargement of the bands of the ERM to ±15%, with the exception of the guilder/deutschmark exchange rate that maintained the ±2,25% band.
(g) Realignment in the EMS with a devaluation of 7% of the Spanish's peseta and 3,5% of the Portuguese escudo.

In Figure 1 we depict the behaviour of the Portuguese escudo exchange rate against the Deutschmark and the evolution of the overnight interest rates differential between Portugal and Germany as part of the PTE/DM band[4]. Besides the target zone period, we simulated, from January 2 1987 to April 5 1992, an unofficial band of ±6%, with an unofficial central parity (No_C_PTE_DM) and unofficial intervention edges (No_LI_PTE_DM_6 and No_LS_PTE_DM_6) equal to that adopted on joining[5].

Figure 1: PTE/DM Exchange Rate and Overnight Interest Rate Differential



[4] See Appendix I for a description of the data used.
[5] See Appendix II for a complete list of variables and their description.

Figure 1 shows that the PTE/DM exchange rate had been relatively stable since the beginning of the 1990s. This was the result of the pegging of the Portuguese escudo to the Deutschmark, which allowed the Portuguese currency to benefit from the credibility, stability and discipline associated with the tacit acceptance of the anti-inflationary stance of the Bundesbank's monetary policy.

Concentrating on the target zone period, it is possible to confirm that after joining the ERM, the Portuguese escudo registered a significant nominal appreciation and an almost immediate decline to a value near the lower edge of its band. Furthermore, everything seems to indicate that the realignments were anticipated, given the high interest rates differential before the realignments. After the widening of the bands, the exchange rate again stabilises, but now within the implicit band of ±6%[6].

The effectiveness of the nominal stabilisation policy of the escudo allowed the Portuguese economy to substantially reduce inflation deviation against Germany. From 1992 onwards there was a gradual disinflation process in Portugal, expressed on a continuous decline in the inflation rate which coincided with the participation of the Portuguese escudo in the exchange rate target zone.

Nevertheless, despite the fact that Portugal has made great strides in its European integration process, it should also be mentioned that disinflation brought about a drop in competitiveness due to the real appreciation of the escudo. This affected output growth and, consequently, employment, particularly in the sectors more open to international competition[7]. In comparison to other European countries, the costs of the disinflation process in Portugal seem relatively modest, at least at first[8]. However, the more aggressive effects of the exchange rate appreciation policy are felt in the long run, even if everything seems to indicate that in the short run the costs of the disinflation are relatively low. The findings of Rebelo and Végh (1995), European Commission (2004) and Abreu (2006) pointed in this direction. The main issue seems to lie in the fact that Portugal has joined the ERM of the EMS with a currency too appreciated.

The issues recently raised by the Bruegel policy brief seem to support these conclusions[9]. The first seven years of participation in the euro zone show that some countries, such as Ireland, have been successful, but others, such as Portugal, have faced some adjustment costs. In the case of Ireland, the appreciation of the real exchange rate was quickly adjusted by substantial improvements in terms of productivity and by upward movements in the value of tradable goods. In Portugal, however, the slow growth of productivity and the composition of exports, based on traditional sectors with low added-value, made the country more vulnerable to international competition from countries with low production costs, especially China and new European Union member states.

It was nevertheless possible to begin a cycle of exchange rate stability. Thus, improved competitiveness must be based on structural adjustments. Competitiveness requires productivity gains, which will depend to a great extent on the modernisation of the tradable goods sector and also to a depreciation of the real exchange rate and a tighter budgetary discipline, which may require an extended period of wage moderation. Careful monitoring of credit expansion is also essential to prevent an overshooting of domestic prices, especially because Portugal has to adjust to lower interest rates that it hasn't historically been accustomed to.

3 – DISTRIBUTION AND VOLATILITY UNDER THE PORTUGUESE EXCHANGE RATE TARGET ZONE

According to the basic target zone model proposed by Krugman (1991), the behaviour of the exchange rate within the band depends on an aggregate fundamental and its expected rate of change, as can be described by the following equation:

[6] The use of interest rates with longer maturity times does not change these results. Similar conclusions were also obtained for the PTE/ECU exchange rate. Due to space limitations, we are only presenting the most important results. The other analyses are available from the corresponding author on request (email: portugal@fe.uc.pt).
[7] See Duarte (2001, 2002) and Macedo (2006).
[8] For suggestive evidence see Barbosa and Machado (1996).
[9] See Ahearne and Pisani-Ferry (2006).

$$s(t) = f(t) + \alpha E_t\left[ds(t)\right] / dt, \qquad \forall t \text{ and } \alpha > 0, \tag{1}$$

where *s(t)* is the log of the nominal exchange rate at time *t*, *f(t)* is the fundamental at time *t*, α is the absolute value of the semi-elasticity of the exchange rate with respect to its expected rate of change and E_t is the expectations operator conditional on the available information at time *t* according to the rational expectations hypothesis.

The fundamental is the sum of two components,

$$f(t) = m(t) + v(t), \tag{2}$$

the domestic money supply, *m(t)*, and a term representing a composite money demand shock, usually referred to in the literature on target zones as "velocity", *v(t)*. The model assumes that "velocity" is an exogenous stochastic process, whereas the money supply is a stochastic process controlled by the monetary authorities. The question is then how the presence of a credible floating band may affect the behaviour of the exchange rate.

In the absence of any intervention, a situation common in a free floating regime, it is assumed that the money supply *m(t)* is kept constant. As a consequence, the fundamental is simply equal to "velocity", *f(t) = v(t)*. It is thus assumed that "velocity" follow a Brownian motion with drift μ and instantaneous standard deviation σ:

$$dv(t) = \mu\, dt + \sigma\, dz(t), \ \mu \text{ and } \sigma \text{ positive parameters and } v(0) > 0, \tag{3}$$

where *z(t)* is a Wiener process with $E_t\left[dz(t)\right] = 0$ and $E_t\left[\left(dz(t)\right)^2\right] = dt$, that is, *f(t)* is the equivalent of a continuous random walk[10].

This assumption implies that the exchange rate under a free floating regime is also a Brownian motion. Therefore, changes in the fundamental will translate into equal changes in the exchange rate, *ds(t)=df(t)*.

In a target zone, it is assumed that the intervention rule is based on a specific floating band for the fundamental, $f_L \le f(t) \le f^U$, and that, if necessary, the fundamental will be regulated to remain within the band. This implies that the fundamental follows a regulated Brownian motion with constant drift and instantaneous standard deviation[11]:

$$df(t) = \mu\, dt + \sigma\, dz(t) + dL(t) - dU(t), \tag{4}$$

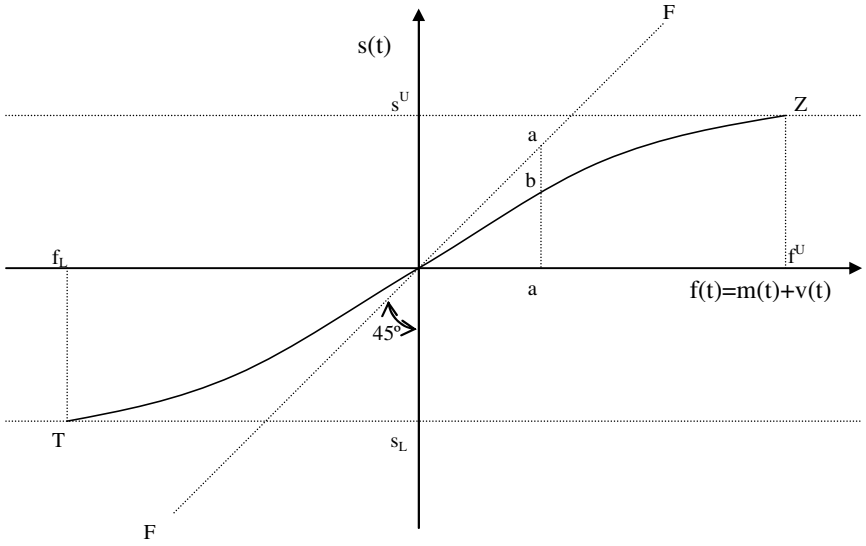
where *L(t)* and *U(t)* are the lower and upper regulators, defined as continuous and increasing functions of *t*, so that *dL(t)* represents increases in the money supply, positive only if $f(t)=f_L$ and *dU(t)* represents decreases in the money supply, positive only if $f(t)=f^U$. It is thus necessary to assume that the probability distribution of the interventions is determined by the current level of the fundamental, $s(t) = s\left(f(t)\right)$, so that the exchange rate function in a target zone will be flat at the edges of the fundamental band and tangent to the edges of the exchange rate band:

$$s'(f_L) = s'(f^U) = 0. \tag{5}$$

Under these circumstances, the exchange rate function establishes a non-linear relationship between the exchange rate and its fundamental, as illustrated in Figure 2.

[10] See Merton (1992), Campbell, Lo and Mackinlay (1997), Maddala and Kim (1999) and Demange and Rocher (2005).
[11] Harrison (1985) and Karatzas and Shreve (1997) provide a formal presentation of these processes.

Figure 2: Exchange Rate in a Perfectly Credible Target Zone



The straight-line FF represents the equilibrium exchange rate in the free floating case. A shock in *v(t)* leads to a proportional change in *f(t)* and *s(t)*. According to equation (5), the exchange rate target zone function is tangent to the edges of its floating band, where $s_L=s(f_L)$ and $s^U=s(f^U)$, represented by the curves TZ, non-linear, and S-shape, respectively.

The behaviour of the exchange rate in a target zone with perfect credibility leads to two main results. First, the slope of the curve TZ is always less than one. This feature is called "the honeymoon effect", a reference by Krugman (1987: 19) to a "target zone honeymoon". The exchange rate function thus appears less sensitive to changes in the fundamental than the corresponding free floating exchange rate. Moreover, the part of the adjustment supported by the exchange rate in a target zone is not constant, but decreases as the exchange rate moves away from the central parity. The "honeymoon effect" thus implies that a perfectly credible target zone is inherently stabilizing[12].

Second, the curve TZ becomes flatter, reaching a zero slope at the edges of the band. At the edges of the target zone, the exchange rate function is tangential to the horizontal dashed lines that represent the edges of the exchange rate band. This result, represented by equation (5), is known as "smooth pasting" conditions, which correspond to the continuity conditions for the solution of the basic model. Given these conditions, the model predicts that the statistical distribution of the exchange rate must be U-shaped or bimodal.

We analysed the asymptotic distribution of the exchange rate within the band based on the work of Harrison (1985), Svensson (1991a) and Bertola and Caballero (1992). Since the exchange rate function, *s(f)*, is strictly increasing and invertible only on the interior of the band, the asymptotic density function of the exchange rate, $\varphi^s(s)$, is given by:

$$\varphi^s(s) = \frac{\varphi^f\left(f(s)\right)}{\left|s'\left(f(s)\right)\right|}, \tag{6}$$

for $s_L < s < s^U$, where *f(s)* denotes the inverse of *s(f)*. Since the fundamental is uniformly distributed, the exchange rate will change slowly near the edges of the band. The exchange rate will thus tend to fix itself in the region where it changes more slowly and $\varphi^s(s)$ should present a peak near s_L and s^U.

[12] Returning to Figure 2, if we consider that a positive random shock in *v(t)* increases the fundamental from the origin to point *a*, under a free floating regime the exchange rate increases by the same amount. However, in a target zone, agents recognize that there is a high probability of a future contraction in the money supply. Thus, agents expect the future appreciation of the exchange rate. This results in an equilibrium exchange rate that is less than *a*, at point *b*.

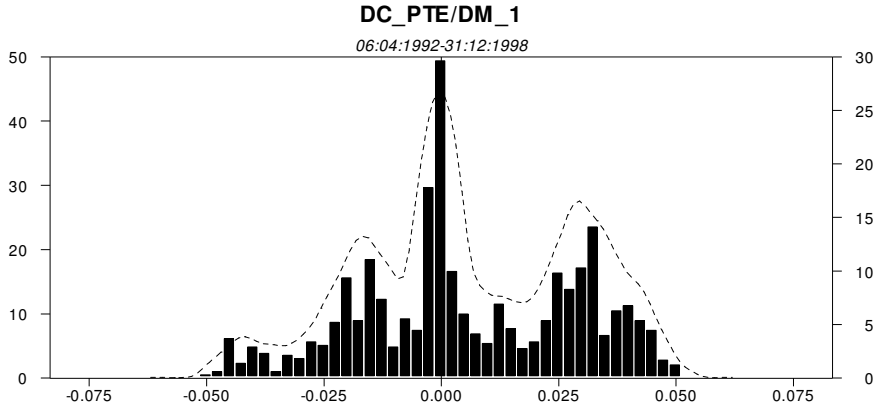
The model also implies that the exchange rate will be much less variable near the edges of the band. According to the model, the instantaneous variability of *s(t)* is directly proportional to the slope of the curve TZ:

$$V(t)\left[ds(t)\right] = \left[s'\left(f(t)\right)\sigma\right]^2 dt\,, \qquad (7)$$

where *V(t)[ds(t)]* is the conditional variance of the changes in the exchange rate. The variability of the exchange rate thus reaches a maximum at the centre of the exchange rate band and decreases as the exchange rate gets closer to the edges of the band.

Figure 3 shows the frequency distribution and the Kernel estimation of the PTE/DM exchange rate function for exchange rate regime 1 (Whole period in the ERM of the EMS[13]).

Figure 3: Frequency Distribution and Kernel Estimation of the Probability Density Function of the PTE/DM Exchange Rate (Regime 1)



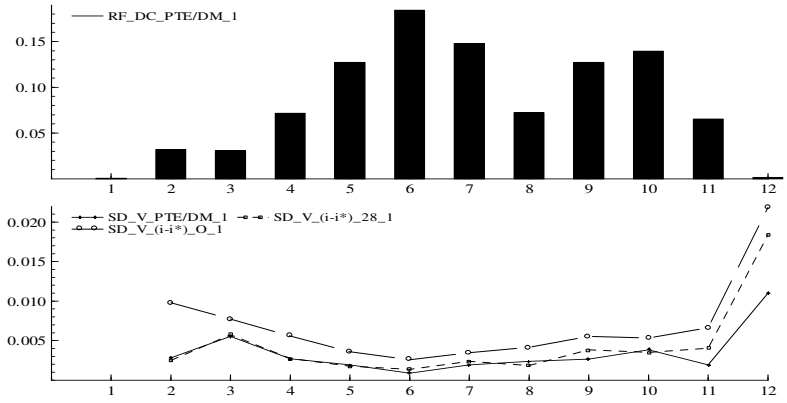
We used the Epanechnikov function[14] for the analysis of the density distribution of the exchange rate. As we can see, most of the observations of the PTE/DM exchange rate often lie close to the central parity[15]. Despite the nominal exchange rate stabilization process, enabled by the participation of the Portuguese escudo in an exchange rate target zone regime, the U-shaped or bimodal density of the exchange rate is rejected by the statistical analysis of data[16].

We also evaluated the distribution of the exchange rate series in specific intervals of exchange variation of identical width. From this analysis, we proceeded to study the exchange rate and the interest rates volatility in each sub-interval used in the histograms. For this, we considered twelve sub-intervals in the ±6% fluctuation band.

As can be seen from Figure 4 (top), most of the observations lie close to the central parity for the whole target zone period, thus rejecting once again the U-shaped or bimodal density of the exchange rate[17].

[13] See Appendix I for a description of the remaining periods used in the empirical analysis.
[14] RATS 6.2 uses the Normal distribution as Kernel, represented by the dotted line, while the histogram is represented by the vertical bars. See Pagan and Ullah (1999).
[15] The left axis concerns the histogram, as the right axis corresponds to the Kernel estimation.
[16] For the remaining exchange rate regimes, see Figure A.1 in Appendix IV.
[17] At the bottom of Figure 4 we represent the exchange rate and interest rates differential volatility in each sub-interval considered in the histogram. We return to this subject later. For the remaining exchange rate regimes, see Figure A.2 in Appendix IV.

Figure 4: Exchange Rate and Interest Rate Distribution and Volatility (PTE/DM, Regime 1)



The analysis of the histograms also allowed us to conclude that only in the period immediately after the Portuguese escudo joined the ERM (exchange rate regimes 2 and 8) were there a larger number of observations lying close to the limits of the band, although this is not very pronounced near the upper edge. The autonomy in monetary policy making as well as in exchange rate policies, possibly due to the maintenance of capital movement restrictions, explains the stability of the exchange rate near the edges of the band.

For the remaining exchange rate regimes, the U-shape in the distribution of the PTE/DM exchange rate is clearly rejected by the analysis of the histograms[18]. Table 2 summarizes the fundamental nature of these results.

Table 2: Relative Frequency of PTE/DM Exchange Rate Observations in Twelve Sub-Intervals of the Band

Regime	Lower Edge (a) (%)	Centre of the Band (b) (%)	Upper Edge (c) (%)	Distribution
1	13.50	53.17	33.33	∩
2	48.18	33.58	18.24	\ : ∪
3	46.67	53.33	0	—
4	9.44	6.78	83.78	/
5	6.42	77.74	15.84	∩
6	53.31	37.28	9.41	\
7	4.58	57.03	38.39	∩
8	50.33	31.79	17.88	\ : ∪
9	9.56	55.45	34.99	∩
10´	*49.61*	*48.31*	*2.08*	—
11	*0*	*94.71*	*5.29*	∩

Note 1: (a) Sum of the relative frequencies of the sub-intervals 1 to 4 (%).
(b) Sum of the relative frequencies of the sub-intervals 5 to 8 (%).
(c) Sum of the relative frequencies of the sub-intervals 9 to 12 (%).

Note 2: ∪= bi-modal or U-shaped distribution (larger number of observations near the edges of the band); ∩= uni-modal or unshaped distribution (larger number of observations near the centre of the band); \= uni-modal distribution in the form of "U cut to the right" (larger number of observations near the lower edge of the band); /= uni-modal distribution in the form of "U cut to the left" (larger number of observations near the upper edge of the band); — = Undetermined distribution.

Suggested reading: In regime 1, 13,5% of the PTE/DM exchange rate observations are located near the lower edge of the band, 53,17% concentrated in the centre of the band and 33,33% located close to the upper edge of the band. Thus, the PTE/DM exchange rate presents a uni-modal or unshaped distribution, indicating a larger number of observations in the sub-intervals related to the centre of the band.

[18] The results are not very different from those in Flood, Rose and Mathieson (1991), Bertola and Caballero (1992), Magnier (1992) and Lindberg and Soderlind (1994) for the case of the fluctuation bands of the countries of the centre of the EMS and for the Swedish unilateral target zone.

This lack of conformity with theoretical predictions, instead of indicating a lack of credibility of the bands, may be due to the existence of intra-marginal interventions, which were frequent during the period under analysis. Since the intra-marginal interventions limit the amount of time spent by the exchange rates near the edges of the band, their occurrence take the exchange rates close to the central parity, thus contributing decisively to the rejection of the first prediction of the model.

Table 3 presents the annual variation of the official foreign exchange reserves held by the Portuguese monetary authorities from 1990 to 1997. Its evolution allows the frequency of intra-marginal interventions to be evaluated and also helps to understand how the Portuguese authorities managed to stabilize the behaviour of the Portuguese escudo within the band, even during the periods of disturbance in the ERM.

Table 3: Annual Variation of the Official Foreign Exchange Reserves

Year	Variation of Foreign Exchange Reserves (10^6 PTE)		
	Foreign Currencies	Officials ECU	Total (a)
1990	412304	28350	440654 (b)
1991	809500	-4597	804903 (c)
1992	84218	-44350	39868
1993	-113106	89212	-23894
1994	-304108	-23914	-328022
1995	-72200	-53400	-125600
1996	79200	28800	108000
1997	332700	54900	387600
1992	*-932*	*-528*	*-1460 (d)*

Source: Adapted from the Reports of the Banco de Portugal.

Notes: (a) Positions at the end of period. Assets of the monetary authorities, excluding gold and other assets.

(b) Variation evaluated initially in U.S. dollars. Subsequently it was used as conversion factor the PTE/USD exchange rates on December 29 1989 and December 31 1990.

(c) Variation evaluated initially in U.S. dollars. Subsequently it was used as conversion factor the PTE/USD exchange rate on December 31 1990.

(d) Variation evaluated in U.S. dollars.

As can be seen, during the periods of greater instability in the EMS, when the central parity of the Portuguese escudo was realigned, there was a fairly significant reduction in the foreign exchange reserves held by the Portuguese monetary authorities[19]. A few months after the beginning of the participation of the Portuguese escudo in the exchange rate mechanism, the Portuguese central bank began to lose quite substantial amounts of foreign exchange reserves. The realignment on November 23 1992 only temporarily relieved the pressure on the Portuguese currency, with the foreign exchange reserves again declining in March and April 1993. The realignment of the following month introduced some stability to the foreign exchange market. In an attempt to defend the original central parity with which the Portuguese escudo joined the exchange rate target zone of the EMS, and afterwards the parity adopted in November 1992, the Banco de Portugal lost in little more than six months around 30% of the foreign exchange reserves it held in late August 1992[20].

In Table 4 the results of the foreign exchange intervention of the Banco de Portugal during those months are presented.

[19] In 1992 the total foreign exchange reserves, expressed in Portuguese escudos, increased by 39868 million escudos, but when evaluated in USD, they fell by 1460 million dollars. The variation in escudos can only be justified by the fact that, in terms of end of period positions there was a depreciation of the Portuguese currency, which increased from 134,184 escudos per dollar, to 146,758 escudos per dollar.

[20] See Loureiro (1998).

Table 4: Monthly Change in the Foreign Exchange Reserves of the Banco de Portugal

Period / Date	Change in the Foreign Exchange Reserves (10^6 PTE)
September 1992	-513393
October	-117789
November	-82416
December	127574
January 1993	124123
February	10399
March	-436538
April	-113102

Source: Adapted from Loureiro (1998).
Note 1: The evolution observed in the foreign exchange reserves reflects only the foreign currency reserves of the Banco de Portugal, and does not include gold or foreign exchange reserves held by the Treasure.
Note 2: Positive (negative) changes mean purchase (sale) of foreign currency by the Banco de Portugal.
Note 3: The changes in the foreign exchange reserves were evaluated initially in Deutschmarks. Subsequently, the PTE/DM monthly average exchange rate was used as conversion factor. It is assumed that the change in foreign exchange reserves results from the intervention in a single market, the Deutschmark market. This arises from the way in which the system operates, characterized by the fact that the Bundesbank is responsible for operating on the U.S. dollar foreign exchange market, while at the same time the other central banks have to operate through the ERM, to keep their currencies aligned with the Deutschmark.

As can be observed, in the periods preceding the realignments, the foreign exchange reserves suffered a marked fall in the attempt to defend the central parities. The foreign currency sales helped to cope with the speculative attacks on the Portuguese currency. We can also see that the foreign exchange reserves were restored in the periods immediately following the realignment. The change of central parity relieved the pressure on existing foreign exchange reserves, avoiding their total depletion. By using these two ways it was possible to keep the exchange rate in the target zone regime and, more importantly, close to the central parity. This analysis makes it is possible to explain the results lack of conformity with the predictions of the literature.

On the other hand, the fact that the Portuguese escudo floated within a wide band of ±6% implied a lower probability that the edges of the target zone would be reached. If a narrow band of ±2,25% had been adopted, the exchange rates would certainly have remained near the boundaries of the target zone more often.

However, the proximity of the exchange rates to the edges of the band should not be regarded by itself as a measure of lack of credibility, as it may simply indicate that the target zone is being tested by the agents operating in the market. It is also important to analyse the volatility of the exchange rate as it gets closer to the edges of the band. According to the literature, evidence of lower volatility near the edges of the band is an indicator of credibility. In most exchange rate regimes, the PTE/DM exchange rate volatility tends to increase as the exchange rate approaches the edges of the band, particularly the upper edge, contrary to theoretical predictions[21]. However, this lack of conformity with the predictions of the basic model should not be interpreted as indicating the lack of credibility of the Portuguese escudo floating band. It is important to remember that the escudo very rarely approached the edges of the band, where the existence of less volatility in the exchange rate should be more obvious.

The exchange rate volatility analysis also confirms that the realignments are preceded by periods of above average volatility, largely as a result of speculative attacks and fear of a sudden depreciation in the currency value. The periods immediately after the realignments, however, are characterized by below-average exchange rate volatility, due to the fact that the weak competitive position of the currency has been restored[22].

[21] For a complementary analysis see Table A.1 in Appendix III.
[22] See Figure A.3 in Appendix IV.

4 – TRADE-OFF BETWEEN EXCHANGE RATE VOLATILITY AND INTEREST RATES DIFFERENTIAL VOLATILITY UNDER THE PORTUGUESE EXCHANGE RATE TARGET ZONE

The basic model has also implications for the behaviour of the interest rates differential that emphasize the special features of the adjustment process in a target zone following a monetary shock. Svensson (1991c, 1992a, b) shows that in narrow target zones, with and without devaluation risk, the foreign exchange risk premium is likely to be very small. We thus assume that the foreign exchange risk premium is zero and that uncovered interest rate parity continually holds, that is:

$$i(t) - i^*(t) = E_t\left[ds(t)\right]/\,dt\,. \tag{8}$$

Letting $\delta(t)$ denote the interest rates differential, we can combine (8) with equation (1) in order to represent the interest rates differential:

$$\delta(t) = \frac{s(t) - f(t)}{\alpha}\,. \tag{9}$$

In the case of a free floating regime it follows that the interest rates differential is constant and equal to the drift of the fundamental.

Since under a target zone $s'\left(f(t)\right) < 1$, it follows that the interest rates differential will be decreasing in the fundamental, $\delta'\left(f(t)\right) < 0$. Furthermore, recalling that the exchange rate is an increasing function of the fundamental, the basic model implies a negative deterministic relationship between the exchange rate and the interest rates differential.

We can find here the main difference between a free floating regime and a target zone. In a free floating regime, the exchange rate is responsible for the whole adjustment process. In a target zone, the exchange rate and interest rates "share" that responsibility between them. The volatility spillover from the exchange rate to the interest rates can be analysed through the instantaneous variability of the interest rates differential. The variability of $\delta(t)$ is computed using Ito's lemma:

$$d\delta(t) = \frac{\delta\left(f(t)\right)}{\alpha}\,dt + \delta'\left(f(t)\right)\sigma dz(t)\,, \tag{10}$$

where $\left[\delta'\left(f(t)\right)\sigma\right]^2 dt$ represents the conditional variance of the changes in the interest rates differential. It follows that the interest rates differential's volatility is minimum at the centre of the band and increases near the edges of the band, replacing here the exchange rate as the main variable in the adjustment process[23]. There is thus a trade-off between the instantaneous variability of the exchange rate and the instantaneous variability of the interest rates differential[24].

We examined the correlation between the estimated values of the conditional variances of the exchange rate and interest rates differentials under a M-GARCH (Multivariate Generalized-AutoRegressive Conditional Heteroscedasticity) model[25], in a further attempt to validate the theoretical predictions.

The autoregressive conditional heteroscedasticity (ARCH) was proposed by Engle (1982) in order to explain the evolution of the residuals of financial time series models. A very general form of an ARCH model is given by:

[23] Returning to Figure 4 (bottom), we observe that the interest rates differential volatility seems to behave more in accordance with the implications of the theory.

[24] See Svensson (1991a, 1994).

[25] See Hamilton (1994), Campbell, Lo and MacKinlay (1997), Johnston and DiNardo (1997) and Tsay (2002).

$$y_t = x_t \beta + u_t \tag{11}$$

$$h_t = \text{var}(u_t) = \sigma^2(u_t), \tag{12}$$

in which there is an ARCH (q) process if:

$$h_t = c_0 + a_1 u_{t-1}^2 + a_2 u_{t-2}^2 + ... + a_q u_{t-q}^2, \tag{13}$$

where, by assumption, in the context of our analysis, the dependent variable is the daily depreciation (appreciation) rate of the Portuguese escudo against the Deutschmark.

Instead, Bollerslev (1986) proposed a GARCH (p, q) model where the variance term depends on the lagged residuals (squared), as well as on the lagged variances. This leads to a greater persistence of volatility but with a relatively smaller number of parameters. The GARCH (p, q) model for the variance is given by:

$$\begin{aligned} h_t &= c_0 + a_1 u_{t-1}^2 + a_2 u_{t-2}^2 + ... + a_q u_{t-q}^2 + b_1 h_{t-1} + b_2 h_{t-2} + ... + b_p h_{t-p} = \\ &= c_0 + \sum_{i=1}^{q} a_i u_{t-i}^2 + \sum_{j=1}^{p} b_j h_{t-j} \ . \end{aligned} \tag{14}$$

We used a two equations GARCH model (or multivariate GARCH), M-GARCH (p, q), with *p=1*, *q=1*:

$$Y_t = X_t \beta + \mu_t, \tag{15}$$

where $\boldsymbol{Y_t}$ is a vector with the exchange rate variation and the interest rates differential, and $\boldsymbol{X}_t$ a matrix with the deterministic variables. The exchange rate is always modelled with the constant. In the case of the interest rates differential, in addition to the constant, the model includes a trend and an auto-regressive value (the interest rates differential lagged one day)[26].

In principle we have four variants of the M-GARCH (1, 1) model.

Variant 1 is the separate estimation of the variance for each of the variables in the model:

$$\mathbf{H}_{ij}(t) = c_{ij} + a_{ij} u_i(t\text{-}1) u_j(t\text{-}1) + b_{ij} \mathbf{H}_{ij}(t\text{-}1), \tag{16}$$

where $\boldsymbol{H_{ij}(t)}$ represents the conditional matrix of the variances/covariances of the residuals.

Variant 2 takes into account a more important dynamic between the variances. This model requires a positive semi-defined matrix[27]:

$$\mathbf{H}(t) = \mathbf{C'C} + \mathbf{A'u}(t\text{-}1)\mathbf{u'}(t\text{-}1)\mathbf{A} + \mathbf{B'H}(t\text{-}1)\mathbf{B}. \tag{17}$$

Variant 3 is based on the assumption that the covariance between the conditional variances of the residuals is null, which implies that $a_{ij}=0$ and $b_{ij}=0$, with $i \neq j$[28].

Finally, the fourth variant of the M-GARCH (1, 1) model is based on the assumption of constant correlation between the conditional variances, which allows to overcome the over-parameterisation problem[29].

In accordance to the purpose of our study, the option was naturally for the use of the first two variants of the M-GARCH (1, 1) model, since we want to find the relationship between the conditional variances.

[26] See Engle, Lilien and Robins (1987) and Engle and Kroner (1993).
[27] See Baba et al. (1990) and Engle and Kroner (1993).
[28] See Bollerslev et al. (1988).
[29] See Bollerslev (1990).

Table 5 presents the results of the correlation analysis between the estimated values of the conditional variances of the PTE/DM exchange rate and of the overnight interest rate differential, according to variant 1, for the whole period[30].

Table 5: Correlation between CVr_TxV_PTE/DM and CVr_(i-i*)_O
(Variant 1; Whole Period)

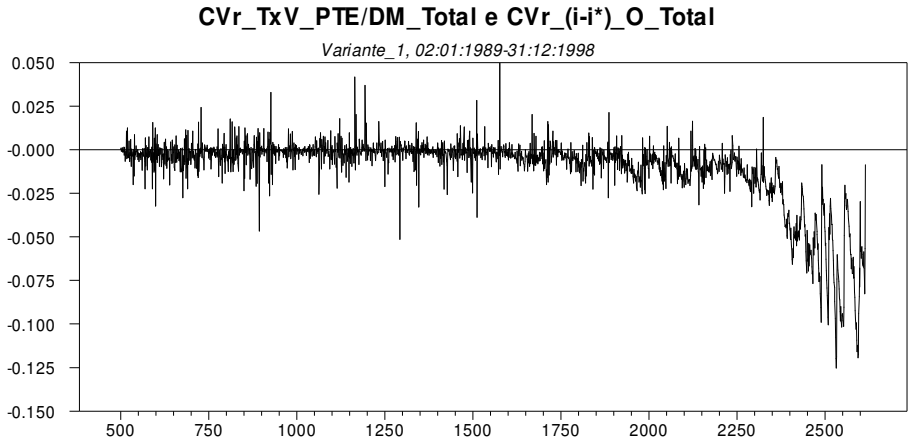
Variable	Coefficient	Standard Error	T statistic	Significance
1- Constant	1,9783e-06	1,1429e-05	0,17310	0,863
2- Constant	6,2772e-03	4,7366e-04	13,25257	0,000
3- Trend	-1,9301e-06	1,3781e-07	-14,00585	0,000
4- (i-i*)_O{1}	0,9567	4,7862e-03	199,88852	0,000
5- C (1, 1)	-5,7192e-11	0,0000	0,00000	0,000
6- C (2, 1)	-2,2643e-08	7,6961e-09	-2,94210	0,003
7- C (2, 2)	9,5160e-08	1,7838e-08	5,33458	0,000
8- A (1, 1)	0,0556	5,5260e-03	10,06433	0,000
9- A (2, 1)	9,2753e-03	0,0206	0,44953	0,653
10- A (2,2)	0,2894	0,0280	10,31747	0,000
11- B (1, 1)	0,9528	3,6375e-03	261,93571	0,000
12- B (2, 1)	-0,2363	0,5336	-0,44279	0,658
13- B (2, 2)	0,8041	0,0129	62,17371	0,000

Note: GARCH Model – BFGS Estimation; Observations used 2117; Convergence in 632 iterations; Final criterion was 0,0000005 < 0,0000100; Log Likelihood 19183,473.

The first constant relates to the PTE/DM exchange rate. The second constant and trend concern the overnight interest rate differential, followed by the values of the lagged coefficient of the interest rates differential. The constant in the variance equation is designated by *C*. The "ARCH" parameters (lagged squared residuals), in increasing lag order, are designated by *A*. The "GARCH" parameters (lagged variance), in increasing lag order, are designated by *B*.

The results of this analysis allow us to conclude that in at least one period there is a negative correlation between the estimated values of the conditional variances of the Portuguese escudo rate of change against the Deutschmark and the overnight interest rate differential. It is also worth mentioning that there is a trade-off between the PTE/DM exchange rate volatility and the overnight interest rate differential volatility. Figure 5 illustrates this.

Figure 5: Correlation between the Conditional Variances of the Exchange Rate and Overnight Interest Rate Differential (Variant 1; Whole Period)



[30] See Table A.2 in Appendix III for the interest rates with a maturity of 28 to 32 days.

The use of both variant 2 and the interest rate differential with a maturity of 28 to 32 days, led to similar results[31]. Possible explanations for this situation are the modernisation of the banking and financial system, and the progress made in terms of disinflation policy, which allowed the interest rate to serve as an alternative variable to the exchange rate in the adjustment process following a monetary shock, thereby facilitating the pursuit of the main objective of price stability. In this context, the participation of the Portuguese escudo in a target zone was crucial to create the conditions of stability, credibility, and confidence, necessary for the adoption of the single currency.

5 – CONCLUSION

In this study we have analysed the participation of the Portuguese escudo in the ERM of the EMS according to the literature on target zones. Although the results do not support the theoretical predictions, this does not mean that the participation in a target zone did not exercise a stabilizing effect on the exchange rate.

The analysis of the distribution of the exchange rate has confirmed that most of the observations tended to lie near the central parity, thus rejecting the U-shaped distribution implied by the theory. For most of the exchange rate regimes analysed, volatility tended to increase as the exchange rate approached the edges of the band, contrary to the predictions of the basic model. Interest rates differential volatility, on the other hand, seemed to behave in line with theoretical predictions.

The analysis of the correlation coefficients between the estimated values of the conditional variances of the exchange rate and the interest rates differential under an M-GARCH model allowed the identification of a negative correlation between the two variables. This result reveals the presence of a trade-off between the exchange rate volatility and the interest rates differential volatility. This is particularly interesting given the presence of a negative trend in the interest rate differential which, instead of representing inefficient actions in the foreign exchange market, reveals the high degree of macroeconomic stability achieved in the Portuguese economy. The downward path eventually reflects the increased credibility of the conduct of monetary policy, allowing the continued pursuit of exchange rate stability, in the context of the ultimate objective of price stability.

The integration process of the Portuguese economy should therefore be used as an example by other small open economies in the sense that they may benefit from participating in one of the dominant monetary areas, otherwise they will be more exposed to speculative attacks, especially in the case of real appreciation of their currencies. But from eight years of EMU, winner members have to conduct solid real convergence policies. Overappreciation currencies can be fatal for new members.

REFERENCES

ABREU, ORLANDO (2006), "Portugal's Boom and Bust: Lessons for Euro Newcomers", *ECFIN Country Focus*, Vol. 3, No. 16, December.

AHEARNE, ALAN AND JEAN PISANI-FERRY (2006), "The Euro: Only for the Agile", *Bruegelpolicybrief*, No. 1, February.

BABA, YOSHIHISA, ROBERT ENGLE, DENNIS KRAFT AND KENNETH KRONER (1990), "Multivariate Simultaneous Generalized ARCH", *Cambridge University Press*, Vol. 11, No.1, pp. 122-50.

BANCO DE PORTUGAL, "Relatório do Conselho de Administração do Banco de Portugal", years of 1990, 1992, 1993, 1995 and 1997.

BARBOSA, ANTÓNIO AND JOSÉ MACHADO (1996), "O Custo da Desinflação Portuguesa: Uma Nota", *Boletim Económico*, Banco de Portugal, September, pp. 47-50.

BERTOLA, GIUSEPPE (1994), "Continuous-Time Models of Exchange Rates and Intervention", in Frederick van der Ploeg (ed.), The Handbook of International Macroeconomics, *Blackwell Economics Handbooks*, July, pp. 251-98.

[31] See Table A.3 in Appendix III. The remaining results, and their graphical analysis, may be obtained from the corresponding author on request (e-mail: portugal@fe.uc.pt).

BERTOLA, GIUSEPPE AND LARS SVENSSON (1993), "Stochastic Devaluation Risk and the Empirical Fit of Target-Zone Models", *Review of Economic Studies*, Vol. 60, pp. 689-712.

BERTOLA, GIUSEPPE AND RICARDO CABALLERO (1992), "Target Zones and Realignments", *American Economic Review*, Vol. 82, No. 3, June, pp. 520-36.

BOLLERSLEV, TIM (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, Vol. 31, pp. 307-27.

BOLLERSLEV, TIM (1990), "Modeling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model", *Review of Economics and Statistics*, Vol. 72, pp. 498-505.

BOLLERSLEV, TIM, ROBERT ENGLE AND JEFFREY WOOLDRIDGE (1988), "A Capital Asset Pricing Model With Time Varying Covariances", *Journal of Political Economy*, Vol. 96, pp. 116-31.

BORDO, MICHAEL (2003), "Exchange Rate Regime Choice in Historical Perspective", Working Paper 9654, *National Bureau of Economic Research*, April.

CALVO, GUILLERMO AND CARMEN REINHART (2002), "Fear of Floating", *Quarterly Journal of Economics,* Vol. 127, No. 2, pp. 379-408.

CAMPBELL, JOHN, ANDREW LO AND ARCHIE MACKINLAY (1997), The Econometrics of Financial Markets, *Princeton University Press*, Princeton.

DEMANGE, GABRIELLE AND JEAN-CHARLES ROCHER (2005), Méthodes mathématiques de la finance, *Economica*, 3rd Edition, September.

DIEBOLD, FRANCIS AND JAMES NASON (1990), "Nonparametric Exchange Rate Prediction?", *Journal of International Economics*, Vol. 28, No. 3-4, May, pp. 315-32.

DUARTE, ANTÓNIO PORTUGAL (2001), "Paridade do Poder de Compra e das Taxas de Juro: Um Estudo Aplicado a Três Países da UEM", *Estudos do GEMF*, No. 4, Faculdade de Economia da Universidade de Coimbra.

DUARTE, ANTÓNIO PORTUGAL (2002), "O Processo Desinflacionista Português: Análise de Alguns Custos e Benefícios", *Estudos do GEMF*, No. 1, Faculdade de Economia da Universidade de Coimbra.

DUMAS, BERNARD (1989), "Pricing Physical Assets Internationally: A Non-Linear Heteroscedastic Process for Equilibrium Real Exchange Rates", Working Paper No. 25, *University of Pennsylvania*.

ENGLE, ROBERT (1982), "Autoregressive Conditional Heteroscedasticity With Estimates of the Variance of United Kingdom Inflation", *Econometrica*, Vol. 50, pp. 987-1008.

ENGLE, ROBERT AND KENNETH KRONER (1993), "Multivariate Simultaneous Generalized ARCH", *Econometric Theory*, Vol. 11, pp. 122-50.

ENGLE, ROBERT, DAVID LILIEN AND RUSSELL ROBINS (1987), "Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model", *Econometrica*, Vol. 55, pp. 391-407.

EUROPEAN COMMISSION (2004), "The Portuguese Economy after the Boom", *Occasional Paper*, No. 8, European Economy.

FISCHER, STANLEY (2001), "Exchange Rate Regimes: Is the Bipolar View Correct?", *Journal of Economic Perspectives*, Vol. 15, No.2, pp. 3-24.

FLOOD, ROBERT, ANDREW ROSE AND DONALD MATHIESON (1991), "An Empirical Exploration of Exchange Rate Target Zones", IMF Working Paper No. 91/15, *International Monetary Fund*, February.

FRANKEL, JEFFREY (1999), "No Single Currency Regime is Right for All Countries or at All Times", *Essays in International Finance*, Vol. 215, Princeton University.

FRANKEL, JEFFREY AND STEVEN PHILLIPS (1992), "The European Monetary System: Credible at Last?", *Oxford Economic Papers*, Vol. 44, No. 4, pp. 791-816.

FROOT, KENNETH AND MAURICE OBSTFELD (1991), "Exchange Rate Dynamics Under Stochastic Regime Shifts: A Unified Approach", *Journal of International Economics*, Vol. 31, November, pp. 203-29.

GARBER, PETER AND LARS SVENSSON (1995), "The Operation and Collapse of Fixed Exchange Rate Regimes", in Gene M. Grossman and Kenneth Rogoff (eds.), Handbook of International Economics, Vol. III, *North Holland*.

HAMILTON, JAMES (1994), Time Series Analysis, *Princeton University Press*, Princeton.

HARRISON, MICHAEL (1985), Brownian Motion and Stochastic Flow Systems, *John Wileys & Sons*, New York.

HENDRY, DAVID AND JURGEN DOORNIK (2001), Empirical Econometric Modelling Using PcGive 10, Vol.I and II, *Timberlake Consultants*.

JOHNSTON, JACK AND JOHN DINARDO (1997), Econometric Methods, 4th Edition, *McGraw Hill*, New York.

KARATZAS, IOANNIS AND STEVEN SHREVE (1997), Brownian Motion and Stochastic Calculus, *Springer-Verlag*, 2nd Edition, June.

KEMPA, BERND AND MICHAEL NELLES (1999), "The Theory of Exchange Rate Target Zones", *Journal of Economic Surveys,* Vol. 13, No. 2, April, pp. 173-210.

KRUGMAN, PAUL (1987), "Trigger Strategies and Price Dynamics in Equity and Foreign Exchange Markets", Working Paper 2459, *National Bureau of Economic Research*, December.

KRUGMAN, PAUL (1991), "Target Zones and Exchange Rate Dynamics", *The Quarterly Journal of Economics*, Vol. CVI, No. 3, August, pp. 669- 82.

KRUGMAN, PAUL AND MARCUS MILLER (1992), Exchange Rate Targets and Currency Bands, Centre for Economic Policy Research and National Bureau of Economic Research, *Cambridge University Press*.

LEVY-YEYATI, EDUARDO AND FEDERICO STURZENEGGER (2005), "Classifying Exchange Rate Regimes: Deeds versus Words", *European Economic Review*, Vol. 49, No. 6, pp. 1603-35.

LINDBERG, HANS AND PAUL SODERLIND (1994), "Testing the Basic Target Zone Model on Swedish Data: 1982-1990", *European Economic Review*, Vol. 38, pp. 1441-69.

LOUREIRO, JOÃO (1998), Euro-Análise Macroeconómica, *Vida Económica*, Porto.

LUTKEPOHL, HELMUT AND MARCUS KRATZIG (2004), "Applied Time Series Econometrics", in Helmut Lutkepohl and Markus Kratzig (eds.), *Cambridge University Press*, Cambridge.

MACEDO, JORGE BRAGA DE (2006), "Competitividade Portuguesa na Economia Global", *Working Paper*, No. 493, Faculdade de Economia da Universidade Nova de Lisboa.

MADDALA, G. S. AND IN-MOO KIM (1999), Unit Roots, Cointegration and Structural Change, *Cambridge University Press*, Cambridge, January.

MAGNIER, ANTOINE (1992), "Théorie des zones cibles et fonctionnement du SME", *Economie et Prévision*, No. 104 1992-3.

MCKINNON, ROBERT (1982), "Currency Substitution and Instability in the World Dollar Standard", *American Economic Review*, Vol. 72, No. 3.

MCKINNON, ROBERT (1984), An International Standard for Monetary Stabilisation, *Institute for International Economics*, Washington.

MEESE, RICHARD AND ANDREW ROSE (1991), "An Empirical Assessment of Non-Linearities in Models of Exchange Rate Determination", *Review of Economic Studies*, Vol. 58, No. 3, May, pp. 603-19.

MERTON, ROBERT (1992), Continuous-Time Finance, *Blackwell*, June.

PAGAN, ADRIAN AND AMAN ULLAH (1999), Nonparametric Econometrics, *Cambridge University Press*, Cambridge.

PANSARD, FABRICE (1997), "La détermination du taux de change en zones cibles: une synthèse des théories", Document de Travail No. 1997-12/EI, *Économie Internationale*, November.

REBELO, SÉRGIO AND CARLOS VÉGH (1995), "Real Effects of Exchange Rate-Based Stabilization: An Analysis of Competing Theories", Working Paper 5197, *National Bureau of Economic Research*, July.

REINHART, CARMEN (2000), "The Mirage of Floating Exchange Rates", *American Economic Review Papers and Proceedings*, Vol. 90, No.2, pp. 65-70.

REINHART, CARMEN AND KENNETH ROGOFF (2004), "The Modern History of Exchange Rate Arrangements: A Reinterpretation", *The Quarterly Journal of Economics*, Vol. 119, No. 1, February, pp. 1-48.

SVENSSON, LARS (1991a), "Target Zones and Interest Rate Variability", *Journal of International Economics*, Vol. 31, No. 1/2, August, pp. 27-54.

SVENSSON, LARS (1991b), "The Simplest Test of Target Zone Credibility", *IMF Staff Paper*, Vol. 38, September, pp. 655-65.

SVENSSON, LARS (1991c), "The Term Structure of Interest Rate Differentials in a Target Zone: Theory and Swedish Data", *Journal of Monetary Economics*, Vol. 28, No. 1, August, pp. 87-116.

SVENSSON, LARS (1992a), "An Interpretation of Recent Research on Exchange Rate Target Zones", *Journal of Economic Perspectives*, Vol. 6, No. 4, pp. 119-44.

SVENSSON, LARS (1992b), "The Foreign Exchange Risk Premium in a Target Zone With Devaluation Risk", *Journal of International Economics*, Vol. 33, No. 1/2, August, pp. 21-40.

SVENSSON, LARS (1994), "Why Exchange Rate Bands? Monetary Independence in Spite of Exchange Rates", *Journal of Monetary Economics*, Vol. 33, No. 1, February, pp. 157-99.

TSAY, RULY (2002), Analysis of Financial Time Series, *Wiley*, New York.

WILLIAMSON, JOHN (1985), The Exchange Rate System, *Institute for International Economics*, Washington.

WILLIAMSON, JOHN AND MARCUS MILLER (1987), Target and Indicators: A Blueprint for the International Coordination of Economic Policy, *Institute for International Economics*, Washington.

APPENDIX I – THE DATA

We used time series data with daily frequency in an attempt to cover the period from January 2, 1987 to December 31, 1998, which gives a total of 3130 potential observations. The observations corresponding to holidays and weekends were left out of the sample. In addition to the period when a target zone was officially functioning, between April 6 1992 and December 31 1998, we extended the analysis to a period in which Portugal adopted a crawling peg and a managed floating system, when the Portuguese escudo was pegged to the Deutschmark.

The exchange rate and interest rates data was taken from the Banco de Portugal (Long Series: Monetary and Financial Statistics of the Banco de Portugal) and the Bundesbank (Bundesbank Time Series Database). We used nominal exchange rates of the Portuguese escudo against the Deutschmark (PTE/DM). This exchange rate is computed by the Banco de Portugal as the average daily currency price. Interest rate data is available from January 2, 1989 and is measured as the average of the daily transactions held in the Money Market. Since we used daily data, we chose overnight interest rates and interest rates with a maturity of 28 to 32 days. Based on the examination of the evolution of the interest rates over time, we left out the outliers, defined as observations with values higher than 25% of the arithmetic average of the previous thirty observations relative to the observation under analysis. This methodology made the analysis of the volatility of the series feasible.

Unless otherwise stated, all the series have been transformed into natural logarithms. In the case of the interest rates, we used the natural logarithm of 1 plus the interest rate (%), divided by 100. In order to maximize the number of available observations, we were forced to extrapolate missing values whenever there were breaks in the series. The breaks occurred mainly in the Portuguese Money Market interest rates series with a maturity of 28 to 32 days. To compute the missing values, we used an extrapolation method based on an AR1 process with trend. The empirical analysis was applied to different sub-periods, also known as exchange rate regimes. We considered eleven exchange rate regimes:

Regime	Dates (Sub-periods)	Description	Potential Number of Observations
1	06:04:1992-31:12:1998	Whole Period in the ERM of the EMS	1759
2	06:04:1992-22:11:1992	Membership of the ERM - 1st Realignment	165
3	23:11:1992-12:05:1993	1st Realignment - 2nd Realignment	123
4	13:05:1993-05:03:1995	2nd Realignment - 3rd Realignment	472
5	06:03:1995-31:12:1998	3rd Realignment- EMU membership	999
6	06:04:1992-01:08:1993	Narrow band Period	345
7	02:08:1993-31:12:1998	Wide band Period	1414
8	06:04:1992-15:12:1992	Restrictions on Capital Mobility	182
9	16:12:1992-31:12:1998	Free Capital Mobility	1577
10	*02:01:1987-30:09:1990*	*Portuguese Escudo Crawling Peg*	*976*
10´	*02:01:1989-30:09:1990*	*Portuguese Escudo Crawling Peg*	*455*
11	*01:10:1990-05:04:1992*	*Pegging of the PTE to the DM*	*395*

Note: Exchange rate regime 10´presents a smaller number of observations since the interest rates database only contains 455 observations. When the analysis focused on the exchange rate we used the sample corresponding to exchange rate regime 10. When the analysis concerned both the exchange rate and the interest rates we used sample 10´.

The five observations immediately before and after the realignments and the date corresponding to the enlargement of the bands were excluded from the sample, in all the analyses. This was to avoid biases in the analysis. Most results were obtained using RATS 6.2, PcGive 10, and Jmulti 4.1[32].

[32] See www.estima.com, for RATS 6.2, Hendry and Doornik (2001), for PcGive 10, and Lutkepohl and Kratzig (2004), for Jmulti 4.1 (www.jmulti.de).

APPENDIX II – VARIABLES USED IN THE EMPIRICAL ANALYSIS

PTE/DM = PTE_DM: Nominal exchange rate of the Portuguese escudo against the Deutschmark

USD: U.S. Dollar

C_PTE_DM: Official central parity of the Portuguese escudo against the Deutschmark

LI_PTE_DM: Official lower edge for PTE/DM

LS_PTE_DM: Official upper edge for PTE/DM

No_C_PTE_DM: Unofficial central parity of the Portuguese escudo against the Deutschmark

No_LI_PTE_DM_6: Unofficial lower edge for PTE/DM and an unofficial exchange rate band of ±6%

No_LS_PTE_DM_6: Unofficial upper edge for PTE/DM and an unofficial exchange rate band of ±6%

Dif_O: Differential between the Portuguese overnight interest rate and the German overnight interest rate (%)

Dif_28: Differential between the Portuguese interest rate with a maturity of 28 to 32 days and the German interest rate with a maturity of 28 to 32 days (%)

$(i-i^*)$_O: Differential between the Portuguese overnight interest rate and the German overnight interest rate (log)

$(i-i^*)$_28: Differential between the Portuguese interest rate with a maturity of 28 to 32 days and the German interest rate with a maturity of 28 to 32 days (log)

RF_DC_PTE/DM: Relative frequency of DC_PTE/DM (%)

SD_V_PTE/DM: Standard deviation of the PTE/DM variation (log)

SD_V_$(i-i^*)$_O: Standard deviation of the variation of $(i-i^*)$_O

SD_V_$(i-i^*)$_28: Standard deviation of the variation of $(i-i^*)$_28

CVr_TxV_PTE/DM: Conditional variance of the rate of change of the Portuguese escudo against the Deutschmark

CVr_$(i-i^*)$_O: Conditional variance of $(i-i^*)$_O

CVr_$(i-i^*)$_28: Conditional variance of $(i-i^*)$_28

APPENDIX III – TABLES

Table A.1: Exchange Rate and Interest Rate Volatility in Twelve Sub-Intervals of the Band (PTE/DM)

Regime	Exchange Rate (a)				Interest Rate Differential Overnight (b)				Interest Rate Differential Maturity of 28 to 32 days (c)				Volatility		
	Minimum		Maximum		Minimum		Maximum		Minimum		Maximum		Exchange Rate (a)	Interest Rate Overnight (b)	Interest Rate 28 to 32 days (c)
	Standard Deviation (%)	Sub-Intervals	Standard Deviation (%)	Sub-Intervals	Standard Deviation (%)	Sub-Intervals	Standard Deviation (%)	Sub-Intervals	Standard Deviation (%)	Sub-Intervals	Standard Deviation (%)	Sub-Intervals			
1	0,08625	I_6	1,09932	I_{12}	0,25811	I_6	2,18513	I_{12}	0,13531	I_6	1,83876	I_{12}	∪	∪	∪
2	0,18888	I_2	0,97383	I_8	0,60932	I_8	1,23089	I_9	0,17470	I_3	0,76075	I_9	∩	∪	—
3	0,21328	I_7	0,33621	I_5	0,72791	I_6	0,90201	I_3	0,18404	I_5	0,75829	I_3	—	—	∪
4	0,19052	I_{11}	1,09932	I_{12}	0,18562	I_2	2,18513	I_{12}	0,16953	I_5	1,83876	I_{12}	—	⌋	⌋
5	0,04594	I_6	0,19801	I_{10}	0,12508	I_5	0,55742	I_{10}	0,08756	I_6	0,31187	I_{10}	∪	⌋	∪
6	0,27716	I_6	0,98555	I_{10}	0,64246	I_8	1,20590	I_9	0,15438	I_8	0,74500	I_9	—	∪	—
7	0,04594	I_6	0,26696	I_{10}	0,12508	I_5	0,66388	I_{11}	0,08756	I_6	0,40474	I_{11}	∪	⌋	⌋
8	0,18888	I_2	0,98555	I_{10}	0,67213	I_7	1,20590	I_9	0,14730	I_8	0,74500	I_9	—	∪	—
9	0,08625	I_6	1,09932	I_{12}	0,18562	I_2	2,18513	I_{12}	0,13531	I_6	1,83876	I_{12}	∪	⌋	⌋
10´	*0,11178*	*I_3*	*0,19476*	*I_4*	*0,18979*	*I_3*	*1,86544*	*I_7*	*0,13282*	*I_9*	*0,74216*	*I_7*	—	∩	∩
11	*0,15781*	*I_6*	*1,03457*	*I_{10}*	*0,78696*	*I_{10}*	*3,24793*	*I_7*	*0,26865*	*I_6*	*0,90948*	*I_8*	⌋	⌉	—

Note 1: (a) Standard deviation of the variations of the PTE/DM exchange rate in each sub-interval considered in the histogram.
(b) Standard deviation of the variations of the overnight interest rate differential in each sub-interval considered in the histogram.
(c) Standard deviation of the variations of the interest rate differential with a maturity of 28 to 32 days in each sub-interval considered in the histogram.

Note 2: I_j with j= 1, 2, 3,..., 11= Sub-intervals of the exchange rate band;
∪= Minimum volatility in the centre of the band (sub-intervals I_5, I_6, I_7, I_8) and increasing near the edges of the band (sub-intervals I_1, I_2, I_3, I_4, I_9, I_{10}, I_{11}, I_{12});
∩= Maximum volatility in the centre of the band (sub-intervals I_5, I_6, I_7, I_8) and decreasing near the edges of the band (sub-intervals I_1, I_2, I_3, I_4, I_9, I_{10}, I_{11}, I_{12});
⌊= Increasing volatility near the lower edge of the band (sub-intervals I_1, I_2, I_3, I_4);
⌈= Decreasing volatility near the lower edge of the band (sub-intervals I_1, I_2, I_3, I_4);
⌋= Increasing volatility near the upper edge of the band (sub-intervals I_9, I_{10}, I_{11}, I_{12});
⌉= Decreasing volatility near the upper edge of the band (sub-intervals I_9, I_{10}, I_{11}, I_{12});
— = Volatility with undefined behaviour.

Suggested reading: For instance, in regime 2 which covers the period from the date of joining the ERM and the first realignment of the Portuguese escudo against the Deutschmark, the exchange rate volatility is greatest in the centre of the band and decreases as the exchange rate gets closer to the edges of the band. The overnight interest rate differential volatility is smallest in the centre of the band and increasing as the exchange rate gets closer to the edges of the band, and there is therefore a trade-off between the two volatilities. The results relating to the volatility of the interest rate differential with a maturity of 28 to 32 days are not conclusive. However we can see that there is less volatility as the term structure of the interest rates increases.

Table A.2: Correlation between CVr_TxV_PTE/DM and CVr_(i-i*)_28
(Variant 1; Whole Period)

Variable	Coefficient	Standard Error	T statistic	Significance
1- Constant	2,0383e-06	1,1434e-05	0,17826	0,859
2- Constant	3,3537e-03	4,9228e-04	6,81266	0,000
3- Trend	-1,0155e-06	1,4691e-07	-6,91256	0,000
4- (i-i*)_28{1}	0,9707	4,3473e-03	223,28392	0,000
5- C (1, 1)	-5,8879e-11	1,1600e-10	-0,50756	0,612
6- C (2, 1)	-4,7243e-08	2,7780e-08	-1,70059	0,089
7- C (2, 2)	3,8869e-08	9,8079e-09	3,96301	0,000
8- A (1, 1)	0,0554	4,0697e-04	136,15167	0,000
9- A (2, 1)	-2,1731e-03	1,2178e-03	-1,78441	0,074
10- A (2,2)	0,1518	0,0144	10,53108	0,000
11- B (1, 1)	0,9530	1,0848e-03	878,46365	0,000
12- B (2, 1)	-0,9873	8,4243e-04	-1172,00597	0,000
13- B (2, 2)	0,8768	8,8324e-03	99,26874	0,000

Note: All Observations
GARCH Model – BFGS Estimation
Convergence in 119 iterations. Final Criterion was 0,00000072 < 0,0000100
Observations used 2117
Log Likelihood 20279,110

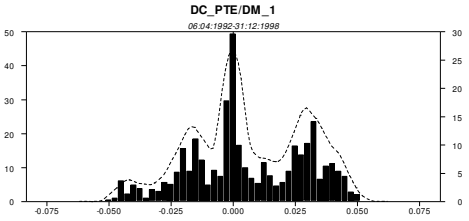
Table A.3: Correlation between CVr_TxV_PTE/DM and CVr_(i-i*)_O
(Variant 2; Whole Period)

Variable	Coefficient	Standard Error	T statistic	Significance
1- Constant	6,1536e-06	1,2324e-05	0,49932	0,618
2- Constant	5,4842e-03	4,8619e-04	11,28016	0,000
3- Trend	-1,6746e-06	1,4168e-07	-11,81932	0,000
4- (i-i*)_O{1}	0,9604	4,8173e-03	199,36618	0,000
5- C (1, 1)	9,7106e-06	4,8309e-06	2,01009	0,044
6- C (2, 1)	-2,5792e-04	3,5794e-05	-7,20576	0,000
7- C (2, 2)	6,7666e-09	1,2683e-04	5,33520e-05	0,999
8- A (1, 1)	-0,1972	0,0100	-19,68569	0,000
9- A (1, 2)	-0,2280	0,0344	-6,62200	0,000
10- A (2,1)	0,0135	3,1292e-03	4,31864	0,001
11- A (2, 2)	-0,5225	0,0262	-19,95680	0,000
12- B (1, 1)	0,9821	1,4145e-03	694,32568	0,000
13- B (1, 2)	-0,0319	5,7684e-03	-5,52804	0,000
14- B (2, 1)	4,4883e-03	8,3214e-04	5,39371	0,000
15- B (2, 2)	0,9026	6,7621e-03	133,47746	0,000

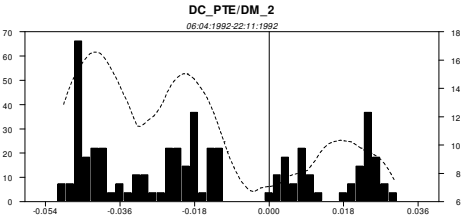
Note: M-GARCH Model, BEKK – BFGS Estimation; Observations used 2117; Convergence in 94 iterations; Final Criterion was 0,0000044 < 0,0000100; Log Likelihood 19141,904.

APPENDIX IV – FIGURES

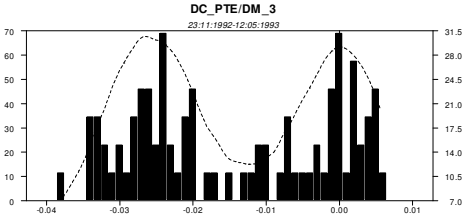
Figure A.1: Frequency Distribution and Kernel Estimation of the Probability Density Function of the PTE/DM Exchange Rate (Regimes 1 to 11)



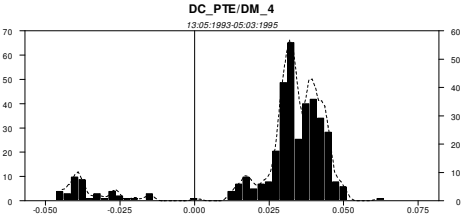
Regime 1



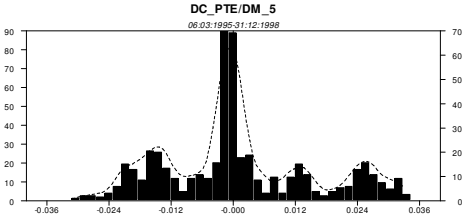
Regime 2



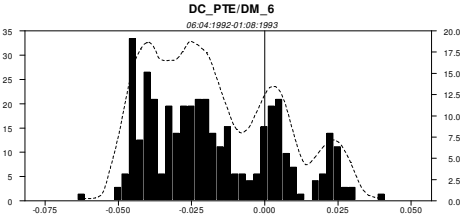
Regime 3



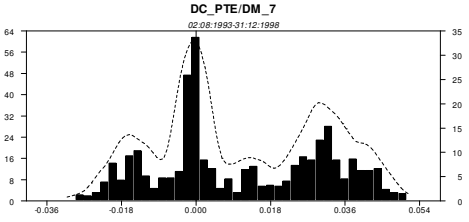
Regime 4



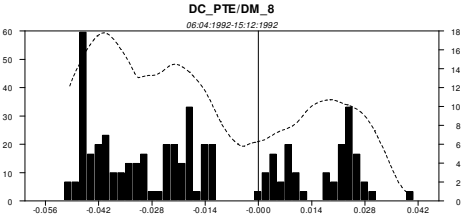
Regime 5



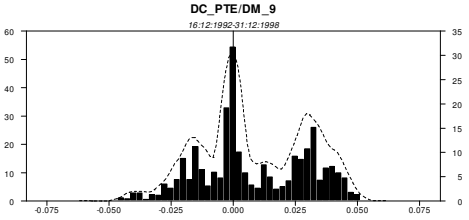
Regime 6



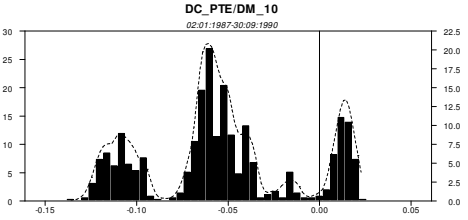
Regime 7



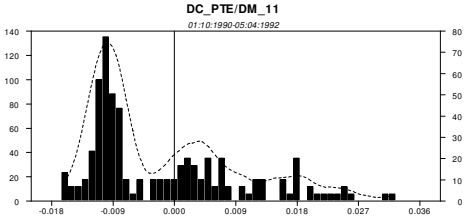
Regime 8



Regime 9



Regime 10



Regime 11

Figure A.2: Exchange Rate and Interest Rate Distribution and Volatility (PTE/DM, Regimes 1 to 11)

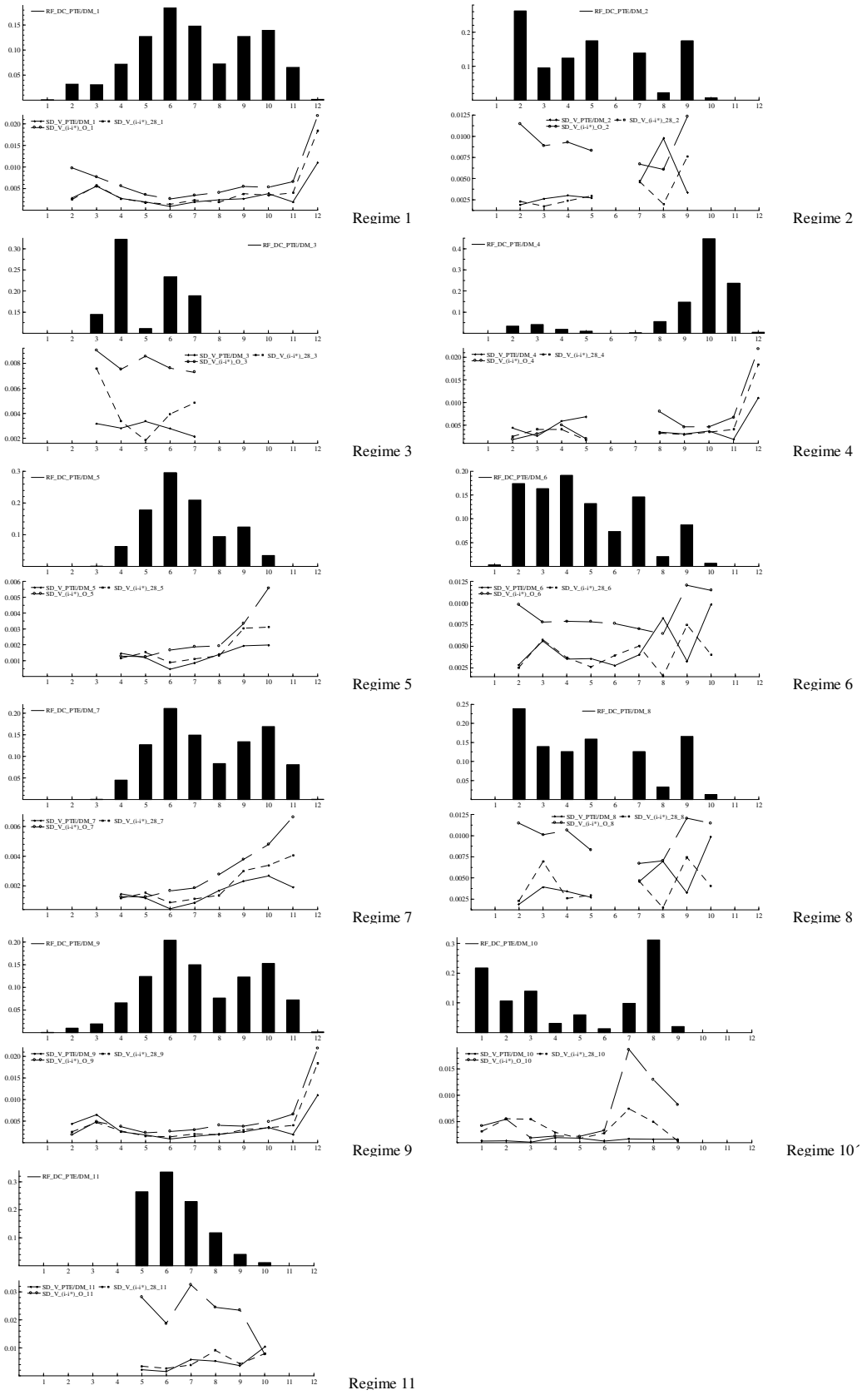
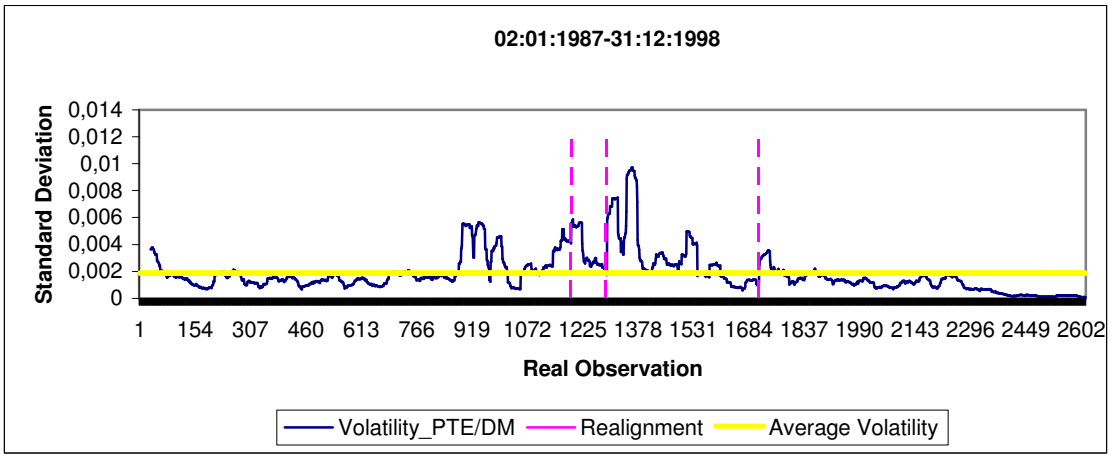


Figure A.3: PTE/DM Exchange Rate Volatility



Note: The volatility consists of the standard deviation of the variation of the PTE/DM exchange rate from the 30 most recent observations.

CRISIS ECONÓMICA Y FINANCIERA MUNDIAL: FUNDAMENTOS, MECANISMOS DE PROPAGACIÓN Y ANÁLISIS DE LAS INEFICIENCIAS CREADAS POR LA AUSENCIA DE COORDINACIÓN

CARMEN DOMINGUEZ DOMECH
E-mail: cdominguez@iberdrola.es
JUAN MANUEL MENÉNDEZ BLANCO
E-mail: jmmenend@ccee.ucm.es
Dpto. Economía Aplicada III
Universidad Complutense de Madrid

Resumen

Tras un periodo de expansión económica en la década de los noventa del pasado siglo y primeros años del actual en la economía europea y mundial sin precedentes tras la Segunda Guerra Mundial, la coincidencia de una crisis financiera globalizada con un periodo de recesión en la mayoría de las economías más desarrolladas obliga a revisar los fundamentos de lo que algunos autores han denominado "la crisis perfecta", analizar los mecanismos de transmisión de las fluctuaciones financieras sobre la economía real y evaluar las alternativas de política económica y financiera adoptadas por los Gobiernos. Adicionalmente se consideran las ineficiencias generadas por las distintas actuaciones de los países que si bien, por un lado, están salvando de la quiebra a una parte del sistema financiero dificultan, por otro, la corrección y/o eliminación de las fluctuaciones financieras.

Palabras clave: crisis económica y financiera, ineficiencias, planes de estimulo a la economía, política monetaria y financiera.
Área Temática: Economia Monetária e Financeira.

Abstract

After the economic expansión of the decade of 1990 of the last century and first years of the current century, with no precedents since the Second World War, in the current situation of financial crisis and a recesión period in most of the developed economies it is time to check the basis of what some analysts call "the perfect crisis". It is needed to check the transmisión mecanisms of the financial flectuations over the economy and evaluate the alternatives in political and financial policy selected by the Governments. Also it is needed to analize the unefficiencies generated by the diferents measures taken by the countries, measures that by one side are avoiding the bankrupcy of part of the financial system but on the other hand is difficulting the corection and or the elimination of the economic fluctuations.

Thematic Area: Monetary and Financial Economics.

1. INTRODUCCIÓN

Tras un periodo de expansión económica en la década de los noventa del pasado siglo y primeros años del actual sin precedentes en la economía europea y mundial tras la Segunda Guerra Mundial, la coincidencia de una crisis financiera globalizada -cuya magnitud y profundidad aún no puede valorarse- con un periodo de recesión en la mayoría de las economías más desarrolladas –cuyo final es difícil de pronosticar- obliga a revisar los fundamentos de lo

que algunos autores han denominado "la crisis perfecta", analizar los mecanismos de transmisión de las fluctuaciones financieras sobre la economía real y evaluar las alternativas de política económica y financiera adoptadas por los Gobiernos, así como proponer otras acordes.

Adicionalmente, no es ajeno a todo ello el debate ideológico sobre el carácter y naturaleza de esta crisis y las políticas económicas a aplicar; controversia ya clásica entre quienes defienden un diagnóstico keynesiano y un mayor activismo e intervención pública y quienes confían en el mercado, sin excluir políticas de regulación específicas en el ámbito económico y financiero.

Nuestro interés, desde el punto de vista investigador, se centra en dos aspectos concretos: el primero, relacionado con el impacto de la crisis financiera en la economía real, tratado en los dos siguientes epígrafes y, que en gran medida, se manifiesta en una corrección importante a la baja de las previsiones macroeconómicas de crecimiento y de generación de empleo elaboradas por los expertos del eurosistema para la zona del euro en diciembre de 2008 (BCE, 2008); proyecciones, por otro lado, menos satisfactorias que las realizadas para fuera de la zona euro.

Tras analizar las interacciones entre los dos componentes de la "crisis perfecta", revisar las intervenciones de la Reserva Federal de EE UU y del BCE – en ocasiones distintas dado el disímil contexto fundacional y operativo en el que ambos bancos centrales aplican la política monetaria- se exponen en el siguiente epígrafe – y se reflexiona críticamente sobre ellos- los distintos planes propuestos y aplicados de estimulo a la economía y las medidas conducentes a favorecer la solvencia y liquidez de los mercados e instituciones financieras.

La magnitud de la crisis financiera ha exigido de los Gobiernos la adopción de medidas mayoritariamente coyunturales destinadas a evitar la quiebra de las instituciones financieras, lo que genera ineficiencias de varios tipos: de origen (en función del país o área económica), de elección (dentro de cada país, son los Gobiernos los que están decidiendo a qué entidades salvar y a cuales dejar quebrar), normativas (derivadas, por un lado, de los distintos marcos legales de entre países, incluso de la eurozona; y, por otro, las que se producirán en el futuro como consecuencia de las distintas medidas que están adoptando los Gobiernos de cada país), y competitivas (el apoyo recibido por las entidades puede entenderse como que no se está favoreciendo la eficiencia y competitividad de las entidades, dado que los errores de gestión acaban siendo difuminados por las ayudas gubernamentales).

Este conjunto de ineficiencias- aplicadas al contexto actual- no ha tenido un tratamiento significativo en la literatura académica y científica por lo que hemos considerado novedoso su estudio dado que son, a nuestro juicio, una de las causas principales de la crisis financiera y de las distintas actuaciones por parte de los gobiernos. Este es el segundo aspecto de interés de esta investigación, que se aborda con detalle en el epígrafe quinto.

2. ¿FIN DE UN CICLO ECONÓMICO?: ALGUNAS REFLEXIONES SOBRE EL ORIGEN, CAUSAS E IMPLICACIONES DE LA CRISIS ECONÓMICA Y FINANCIERA

En la cuarta fase del paradigma tecno-económico [1] asociado a la nueva economía del conocimiento, en general, y las TIC, en particular y tras un periodo de crecimiento de la economía europea y mundial sin precedentes desde la Segunda Guerra Mundial, desde el verano del 2007 comienza a gestarse una crisis económica y financiera, también sin precedentes[2]; crisis que ha renovado a nivel académico el interés por el fenómeno de las burbujas y nos ha recordado que los ciclos –aún con distintas manifestaciones, duración y amplitud que en el pasado- siguen existiendo en un contexto en el que la economía europea parecía haber

[1] Siguiendo a Carlota Pérez (2001), en cada revolución tecnológica se distinguen 5 fases: 1ª) Irrupción tecnológica; 2ª) Auge financiero; 3ª) Punto de inflexión, reflexión, y de reconsideración del proceso; 4ª) Edad de oro; y 5ª) Etapa de madurez. Considerando esta clasificación y este orden, tras la crisis financiera asociada a las empresas *punto com* –fechada en el 2001- comenzaría una segunda etapa de crecimiento y expansión económica.

[2] La mayoría de las economías avanzadas están en recesión, escenario desconocido desde el final de la Segunda Guerra Mundial.

encontrado en los esquemas de política monetaria y fiscal aplicados[3] una senda de crecimiento estable de las variables macroeconómicas.

Esta crisis no se corresponde, sin embargo- con las burbujas asociadas a las revoluciones tecnológicas- la burbuja financiera asociada a algunas empresas del sector tecnológico y, más en concreto, a las *punto com* ya sucedió en el 2001- si no que tiene orígenes, causas y fundamentos distintos: una política de bajos tipos de interés en la economía norteamericana durante un periodo sostenido de tiempo provoca una expansión del volumen de crédito para mantener los márgenes de la banca; primero, a un perfil de cliente de mayor riesgo y, posteriormente, a través de la titulización y el apalancamiento -y en un entorno de globalización- se produce la transferencia de ese riesgo a nivel mundial.

La expansión inicial del crédito se basa en unos activos, mayoritariamente inmobiliarios, cuyo precio disminuye aceleradamente al pinchar la burbuja, de forma que ahora tenemos unos créditos basados en activos en deflación, y posteriormente a través de la titulización y globalización, se ha creado más dinero sin soporte real.

A corto plazo es justificable, por la gravedad de la crisis, la urgencia de las medidas adoptadas por la UE y Norteamérica pero, en gran medida, hay que considerar que únicamente se está produciendo una transferencia de riesgo a los Estados y que se trata de medidas coyunturales ya que a medio plazo será necesario establecer un nuevo marco normativo común para el sistema financiero que limite y regule los riesgos, así como que promueva que las entidades de crédito eficientes vuelvan a asumir los riesgos inherentes a su actividad, pasado ya el periodo de intervención estatal inicial.

Las proyecciones macroeconómicas elaboradas por los expertos del eurosistema para la zona del euro en diciembre de 2008 (BCE, 2008)[4] reflejan la desaceleración económica mundial, las condiciones de financiación más restrictivas junto con las tensiones en los mercados financieros, la aceleración del ajuste en el mercado de la vivienda y la intensificación de los efectos de contagio provenientes de EE UU. Todo ello conduce a estimaciones de crecimiento medio anual del PIB real que oscilan en un intervalo del 0,8%-1,2% para el 2008, -1,0%-0,0% para el 2009, y entre el 0,5%-1,5% para el 2010[5].

Tras estas previsiones está un débil comportamiento de la demanda interna del PIB: aún con políticas de estimulo de la inversión pública, la inversión total se verá afectada negativamente por las restrictivas condiciones de financiación y por la debilidad de la actividad económica mundial. Por otro lado, el consumo privado se espera que crezca a tasas muy moderadas, fundamentalmente, por la pérdida de renta real disponible que supondrá el creciente desempleo y, en menor medida, por el aumento de la tasa de ahorro debido al motivo precaución.

Mención especial requiere el impacto de la crisis en el mercado de trabajo, en el que se prevé que para el periodo 2009-2010 la tasa de crecimiento de la población activa se situará en niveles muy inferiores a los observados desde el 2002, debido principalmente al efecto "desmotivación" causado por una coyuntura en la que la tasa de desempleo aumentará significativamente.

En este contexto macroeconómico, el riesgo de que se agudice la interacción negativa entre el sistema financiero y la economía en general es una amenaza crítica: la combinación de la acumulación de pérdidas, la caída de los precios de los activos y una desaceleración más acusada de la economía ha generado graves dudas acerca de la viabilidad de una franja cada vez más amplia del sistema financiero (Informe sobre la Estabilidad Financiera Mundial, octubre de

[3] En concreto nos estamos refiriendo al diseño de política monetaria que surge de la introducción del euro como moneda única en un conjunto de países y la independencia del BCE en el logro de su objetivo principal – la estabilidad de precios- y a las reglas fiscales que introducen el Tratado de Maastricht y el Pacto de Estabilidad y Crecimiento.

[4] En relación con las estimaciones publicadas en septiembre del 2008 sobre el crecimiento del PIB real los intervalos proyectados para el 2008 y el 2009 se han ajustado a la baja de forma significativa, reflejando el deterioro de las perspectivas económicas a corto plazo y los efectos de las perturbaciones financieras. Las previsiones de otras instituciones (FMI, Comisión Europea, Encuesta a expertos en previsión económica, *Consensus Economics Forecats*, OCDE) son compatibles con los intervalos de las proyecciones de los expertos del Eurosistema, aún existiendo diferencias metodológicas, temporales, y de datos entre las instituciones que no hacen estrictamente comparables las proyecciones.

[5] Las proyecciones de las tasas de crecimiento anual del PIB real son más satisfactorias fuera de la zona euro: 3,9% de promedio en el 2008, 2,4% de promedio en el 2009 y el 3,6% en el 2010.

2008). Por un lado, el proceso de desapalancamiento se ha acelerado y se ha vuelto más desordenado y, por otro, las intervenciones dirigidas a aliviar las tensiones de liquidez y resolver la situación de las instituciones financieras en dificultades no han conseguido sino muy parcialmente devolver la confianza al mercado.

Aunque las instituciones financieras han venido captando capital para reforzar sus balances, el proceso de desapalancamiento del sector bancario exigirá ventas de activos, un crecimiento más lento de los nuevos activos y cambios radicales en los modelos operativos de los bancos, ya que muchas de las fuentes habituales de ingresos han desaparecido. Adicionalmente, el progresivo proceso de reducción del riesgo asumible, implicará necesariamente un periodo de menor rentabilidad.

La continua incertidumbre en torno a la valoración de activos- antes considerados de bajo riesgo- ha complicado la tarea de determinar los niveles de capitalización adecuados. Por ello, la mayoría de los participantes en el mercado, agencias calificadoras de riesgo y reguladores coinciden en el que el nivel de reservas de capital tendrá que ser más elevado de lo que se había previsto; además debería basarse en un análisis prospectivo del riesgo más que en una aplicación mecánica de los coeficientes legales[6].

EE UU sigue siendo el epicentro de la crisis financiera[7]: se ha continuado debilitando el mercado de la vivienda, contribuyendo la desaceleración económica a un mayor deterioro de los préstamos, al aumento de la morosidad en los pagos de hipotecas de primera clase, en el sector inmobiliario comercial y en los préstamos empresariales y de consumo.

Cuando pensábamos que el ciclo económico estaba cada vez más vinculado a la evolución del ciclo tecnológico, dado el impacto e influencia que la revolución tecnológica asociada a las TIC había tenido sobre el gran ciclo expansivo de la economía estadounidense y europea desde la década de los 90 hasta muy recientemente, constatamos como la profundidad del ciclo económico depresivo actual está, fundamentalmente, relacionado con la crisis del sector financiero y el estallido de la burbuja inmobiliaria, fenómenos ambos interrelacionados.

En relación a los distintos planes de estimulo de la economía que se están aplicando y las medidas de apoyo al sector financiero – que se expondrán y revisarán en los epígrafes siguientes- subyace un debate ideológico sobre el carácter y naturaleza de esta crisis y las políticas económicas a aplicar; controversia ya clásica entre quienes defienden un diagnóstico keynesiano y un mayor activismo e intervención pública y quienes confían en el mercado, sin excluir políticas de regulación específicas en el ámbito financiero.

Existe un alto consenso entre los economistas en la poca utilidad –y las contraindicaciones- de utilizar la expansión del gasto público como instrumento para superar la crisis económica[8], aunque uno de sus efectos se manifieste en la debilidad de la demanda agregada, en especial del consumo y de la inversión privada.

Por otro lado, hay dos aspectos de esta crisis sobre los que es interesante reflexionar para evitar tentaciones intervencionistas erróneas:

En primer lugar, los activos en que se sustentaba la riqueza real (inmuebles, acciones y capital humano) están perdiendo valor desconociéndose cuanto más se depreciarán y cuando tocarán suelo.

Esto implica que a medida que la recesión avance y/o las expectativas de salir de ella se retrasen, cada vez más deudores se convertirán en insolventes y las entidades financieras tendrán dificultades para lograr los recursos necesarios para mantener sus ratios de liquidez y solvencia, dañados por la posición económico-financiera de las empresas y familias. El riesgo para bancos y cajas de ahorro de perder el dinero prestado y, en consecuencia, su resistencia a

[6] Un tema de discusión interesante sería si se considera necesario reformular los coeficientes de capital y de solvencia y, en concreto, adaptarlos al ciclo.

[7] Las cifras de pérdidas declaradas en relación con los préstamos y los activos titulizados siguen en aumento, superando los 1,4 billones de dólares, cifra muy por encima de la estimada en el *Global Financial Stability Report* de abril de 2008 (Informe de Estabilidad Financiera Mundial, 2008).

[8] Solo dos premios Nobel de Economía, Paul Krugman y Joseph Stiglitz, defienden ese tipo de medidas.

renovar la financiación y a conceder nuevos créditos se ve agravada por la ausencia y/o la depreciación de los activos que respaldan los préstamos en caso de impago[9].

En segundo lugar, la combinación del enorme endeudamiento privado existente –de hecho la crisis tiene en el exceso de creación de crédito uno de sus principales orígenes- con la aplicación de políticas fiscales expansivas -que originarán déficit y endeudamiento públicos de tal magnitud que podrían incumplir los criterios de Maastricht y el Pacto de Estabilidad y Crecimiento- provocarían mayores necesidades de financiación y a un mayor coste. Los mercados percibirán a estos países como de alto riesgo y el crédito será más caro y escaso, por lo que la escalada alcista del endeudamiento del Estado va a empeorar la situación financiera y crediticia de las familias y de las empresas por otra nueva vía.

Así pues además de que la crisis financiera y una posible depresión no invalidan toda la macroeconomía aprendida desde 1936 –como afirma Robert J. Barro- la política fiscal expansiva no sirve para solucionar la raíz del problema y sí puede generar desequilibrios adicionales. Como señala Rodrigo Rato y de Figaredo (2009): "las medidas puestas en práctica para salir de las anteriores crisis pusieron la semilla de la siguiente. De ahí la necesidad de que los policy makers no sólo atajen la actual debilidad económica sino que se planteen la estrategia del retiro a tiempo de los estímulos que resulten excesivos".

Finalmente, y éste es uno de los datos más importantes para el futuro del crecimiento económico, la tasa de retorno de las inversiones del sector no financiero ha sido en el 2008 en EE UU del 10%, siendo la tasa de retorno media de los últimos 50 años del 7%. Como remarca Xavier Sala-i-Martín (2008), el crecimiento económico de un país no viene precedido de aumentos de la bolsa, pero si de elevadas tasas de retorno en el sector financiero: "para entendernos: mientras Wall Street ha hecho sus locuras financieras, Sillicon Valley ha seguido innovando y eso, a la larga, es lo que determina el crecimiento de la economía".

3. SOLVENCIA, LIQUIDEZ Y PRESIONES EN LOS MERCADOS INTERBANCARIOS: INTERVENCIONES Y POLÍTICA MONETARIA DEL BANCO CENTRAL EUROPEO

Desde el comienzo de la crisis financiera y lo que en aquel momento aún era una desaceleración económica, distintas opiniones de responsables políticos y de instituciones privadas y públicas, así como del mundo académico reclamaban al Banco Central Europeo descensos en el precio oficial del dinero[10]. Tras ello solían darse dos tipos de argumentaciones: la primera se relacionaba con el excesivo coste financiero que suponía el pago de préstamos hipotecarios y, en general, el elevado coste de financiación que perjudicaba a la actividad empresarial; la segunda hacía referencia al caso estadounidense, en el que la Reserva Federal ya había realizado movimientos a la baja en el tipo de interés oficial, ampliando el diferencial respecto al de la eurozona.

Como es frecuente comparar las distintas actuaciones que a veces llevan a cabo la Reserva Federal de EE UU y el BCE ante la evolución del entorno económico y financiero es necesario revisar el contexto fundacional y operativo en el que ambos bancos centrales aplican la política monetaria.

Una primera diferencia radica en los distintos objetivos establecidos en los estatutos de ambas instituciones: mientras que el BCE tiene como objetivo primordial el mantenimiento de la estabilidad de precios a medio plazo, entendiendo que es una premisa básica para el desarrollo, el objetivo de la política monetaria señalada a la Reserva Federal en su acta

[9] Por ejemplo, para el caso español aproximadamente dos tercios de la riqueza de las familias tiene origen inmobiliario. Las acciones de las empresas, y la bolsa en general, han experimentado una corrección bajista cuyo final nadie pronostica con certidumbre y, por otro lado, el desempleo creciente supone que una parte de la población activa va a tener restringido adicionalmente su acceso al crédito.

[10] En algunas declaraciones de los citados se culpaba incluso a la política monetaria aplicada por el BCE de no colaborar en la solución de la crisis ni apoyar la recuperación económica.

fundacional de 1913 es la promoción del pleno empleo, la estabilidad de precios y la moderación de los tipos de interés a largo plazo, sin priorización alguna.

Un segundo grupo de diferencias se relaciona con los tipos de intervenciones y mecanismos de intervención con que cuenta cada entidad, así como las opiniones de sus dirigentes.

Una tercera diferencia se refiere al marco operativo de ambas instituciones: el BCE ha definido un procedimiento muy concreto consistente en fijar un objetivo cuantitativo de inflación (IPC interanual cercano al 2% a medio plazo) y dos pilares para valorar si existen riesgos para su cumplimiento: un análisis económico y otro monetario (evolución de los depósitos, créditos, liquidez del sistema, etc.); por su parte, la Reserva Federal actúa de forma discrecional al no establecer objetivos cuantitativos ni de inflación ni de crecimiento.

El diseño de ambas instituciones explica por qué en ocasiones reaccionan de forma tan diferente. Por ejemplo, cuando a finales del año 2000 aparecieron los primeros síntomas de la crisis económica que se avecinaba, el tipo de referencia de la Reserva Federal estaba en el 6,5% y el del BCE en el 4,75%. A finales de 2001, la Reserva Federal había bajado su tipo de referencia hasta el 1,75% mientras que el BCE solo lo hizo hasta el 3,25%.

Pese a que el crecimiento del PIB europeo se redujo dos tercios a lo largo del año, en mayo de 2001 el IPC armonizado alcanzó su mayor nivel desde la creación del BCE en 1999: el 3,1%. El BCE, como señalan sus estatutos, prefirió asegurar el descenso de la inflación y el control de la estabilidad de precios aún a costa de crecimiento económico. Por el contrario, en EE UU el IPC alcanzó el 3,6% pero eso no impidió a la Reserva Federal luchar contra la desaceleración económica.

Sin embargo, un fenómeno común a ambos lados del Atlántico son las presiones existentes en el mercado interbancario: la combinación de riesgos de liquidez y solvencia desembocó en un periodo de elevados diferenciales entre las tipos de interés a corto plazo y volúmenes de transacciones sustancialmente reducidos, así como una tensión inusitadamente prolongada en los mercados de financiación.

En relación con el primer aspecto mencionado en el párrafo anterior, los procedimientos de fijación de tipos a corto plazo –como el tipo interbancario de oferta de Londres (LIBOR) y el Euribor- deben perfeccionarse pues por un lado el diferencial LIBOR –OIS[11] en dólares de EE UU aumentó considerablemente incorporando los temores en torno a suspensiones de pagos y, por otro, las presiones de liquidez en dólares de EE UU se desbordaron hacia otras monedas (libra esterlina y euro, fundamentalmente), afectando a los diferenciales Euribor-OIS y LIBOR-OIS en libras esterlinas.

Es interesante entender de qué manera influyó la crisis en los tipos de interés como canal de transmisión de la política monetaria en el contexto de tres tendencias a largo plazo: una aceleración de la actividad de las instituciones cuasi-bancarias, un uso más extendido de los mercados de financiación mayoristas y una preferencia mayor por los instrumentos a corto plazo en lugar de una sólida y estable base de depósitos como medio de financiación. Aunque, en general, estas tendencias infundieron estabilidad a la transmisión mediante los tipos de interés, a lo largo del último año la fluida relación tradicional entre el tipo de referencia y el tipo activo se alteró drásticamente.

Con el fin de corregir estos inconvenientes, sería recomendable:

1. Mejorar la infraestructura de los mercados de financiación: estos tipos de referencia mencionados inspirarían más confianza si el cálculo del LIBOR estuviera basado en una muestra más amplia de bancos y cotizaciones que incluyera fuentes no bancarias de financiación (lo mismo seria válido en relación al cálculo del Euribor).

2. Lograr que las autoridades estén más atentas a los riesgos de crédito y de liquidez: el principal motivo de la amplitud de los diferenciales interbancarios fueron los riesgos de dificultades bancarias (tanto desde el punto de vista del crédito como de la liquidez).

3. Otorgar a los mercados de dinero un respaldo indirecto ilimitado: es necesario diseñar con cuidado los mecanismos de crédito de los bancos centrales que buscan reestablecer el funcionamiento de los mercados interbancarios para transmitir la política monetaria. Sería

[11] OIS es la tasa de los *swaps* con un índice a un día.

deseable que los participantes en el mercado empezarán a operar entre sí permitiendo a los bancos centrales una retirada ordenada.

4. Fomentar la cooperación y comunicación entre los bancos centrales, pues el último episodio de problemas de liquidez se contrarrestó gracias a la cooperación entre los principales bancos centrales, que atendieron las necesidades de financiación en moneda extranjera.

En este sentido, la respuesta del BCE a los problemas financieros en la eurozona se concretó en el compromiso de proporcionar toda la liquidez requerida durante todo el tiempo que fuera necesaria; esto se materializó en la implantación de un sistema de suministro de la misma sin límite de cantidad y a plazos que cubren hasta los 6 meses, en la ampliación de las divisas en las que se conceden los préstamos y en una importante extensión del conjunto de activos admitidos como garantía de los préstamos concedidos.

El cambio en la orientación de la política monetaria del BCE no se produjo hasta bien entrado el verano de 2008 cuando la desaceleración del crecimiento disminuyó las tensiones inflacionistas. Los efectos de agudización de la crisis financiera sobre la demanda y el retroceso de los precios de las materias primas relajaron sustancialmente la presión existente sobre los precios, costes y salarios. Todo ello permitió al Consejo de Gobierno del BCE reducir el tipo de interés oficial, en dos movimientos separados por tan solo 21 días laborables, en 1 punto porcentual hasta el 3,25%[12]. Posteriores reducciones han conducido a que en la actualidad –febrero de 2009- el tipo de interés vigente se haya situado en el 2%[13].

Un ejemplo de cómo ha afectado la política monetaria y financiera al sector real de la economía a lo largo de este periodo de crisis se refleja adecuadamente en el caso español: el ajuste en el sector inmobiliario ya había comenzado de una forma gradual en el 2007, corrigiendo los excesos que se habían originado en un contexto global de relajación financiera, de infravaloración del riesgo y bajo unas condiciones monetarias y financieras particularmente holgadas para España. El dinamismo del sector residencial se había ido moderando a medida que el cambio en la orientación de la política monetaria iniciado a finales de 2005 se fue trasladando al coste de financiación de las familias y de las empresas y que las expectativas de revalorización del precio de la vivienda se fueron recortando.

Esta pauta gradual se vio alterada por la irrupción e intensificación de la crisis financiera internacional hacia el verano de 2008: la demanda de viviendas se redujo de forma drástica y sus precios, tras varios años de aumentos cada vez más moderados, empezaron a descender en términos reales. Este escenario provocó además un recorte significativo en la iniciación de viviendas y llevó a numerosas empresas inmobiliarias a tener que reestructurarse, llegando algunas de ellas a entrar en situación concursal[14]. Las implicaciones de todo ello para el empleo total de la economía están siendo bien conocidas pues por primera vez en los últimos catorce años se está produciendo una destrucción de empleo y éste está actuando como el principal canal de propagación del ajuste del sector inmobiliario hacia el resto de la economía.

4. MEDIDAS DE ESTIMULO DE LA ECONOMÍA Y PLANES NACIONALES DE APOYO AL SISTEMA FINANCIERO APLICADOS EN LA UE Y EE.UU.: LA NECESIDAD DE UNA ACCIÓN COORDINADA

Se enumeran, a modo de síntesis de lo ya expuesto, los principales acontecimientos que han propiciado a lo largo del 2008 la crisis del sistema financiero internacional.

1. En el año 2008 ha existido una altísima volatilidad en todos los mercados, tanto de *equity* como de crédito, así como en los mercados de divisas y *commodities*.

[12] El primero de estos recortes tuvo lugar en el marco de un movimiento coordinado con otros grandes bancos centrales.

[13] Otro dato interesante es que el Euribor (2,22% de media mensual con fecha 10 de febrero) se ha ido acercando al precio oficial del dinero, reflejando el descenso de las tensiones en los mercados interbancarios.

[14] Además, las tasas de crecimiento de la financiación a las empresas y las familias españolas son muy inferiores a las existentes previamente y se acercan progresivamente al avance nominal del PIB (Fernández Ordóñez, 2008).

2. Se han producido elevadas pérdidas en los mercados bursátiles a nivel mundial, con un fuerte aumento de las primas de riesgo.

3. La aversión al riesgo ha llegado a límites extremos produciéndose una liquidación masiva e indiscriminada de las posiciones que, junto con las pérdidas acumuladas, provoca que los inversores se muestren reacios a comprar, aún con unas valoraciones atractivas.

4. En el sector financiero se ha iniciado una crisis de liquidez, que se ha trasladado al resto de sectores, con la consecuencia de una restricción generalizada del crédito.

5. A los problemas del acceso a la financiación se une el grave deterioro de los costes de la deuda, con los *spreads* de crédito en máximos históricos.

6. En lo relativo a las *commodities* y el petróleo, tras alcanzar máximos históricos en julio de 2008 los precios se han desplomado.

7. En cuanto a las divisas, el dólar, con mínimos históricos frente al euro en julio de 2008 se aprecia rápidamente en el cuarto trimestre al convertirse en moneda refugio.

8. El rápido deterioro de las condiciones tanto macroeconómicas como microeconómicas ha provocado una reacción de los Gobiernos a nivel mundial, con el anuncio de múltiples medidas para paliar los efectos de la recesión: inyecciones de liquidez, recapitalizaciones de bancos, medidas de apoyo específicas para determinados sectores, garantías gubernamentales en emisiones de deuda, compra de activos tóxicos, rebajas fiscales, etc.

9. A esto se han unido las medidas tomadas por los principales Bancos Centrales, con sucesivas rebajas de los tipos de interés oficiales, que han continuado en el comienzo de 2009 y se espera que sigan produciéndose al dejar de ser la inflación un problema debido a la crisis de consumo y demanda en la economía real, y ante la amenaza de la deflación.

La magnitud de la crisis ha obligado a los Gobiernos de los distintos países a actuar rápidamente para evitar la quiebra del sistema financiero y responder al deterioro acelerado de la economía real. Las medidas adoptadas han consistido principalmente en:

1. Inyecciones de liquidez
2. Recapitalizaciones de bancos
3. Garantías gubernamentales en emisiones de deuda
4. Medidas de apoyo específicas para determinados sectores
5. Compra de activos tóxicos
6. Reducciones impositivas
7. Inversión en infraestructuras

La falta de un diagnostico en profundidad de la situación ha llevado a que las medidas se hayan quedado insuficientes sistemáticamente, asistiendo continuamente a nuevos planes de los Gobiernos destinados a salvar el sistema financiero[15].

A continuación se analizan cada uno de estos tipos de medidas y los efectos que inicialmente están ocasionando.

1. Inyecciones de liquidez

Un primer conjunto de medidas ha estado encaminado a inyectar liquidez en determinadas entidades financieras, mientras otras se han dejado quebrar. Las entidades financieras están utilizando las inyecciones de liquidez para recapitalizarse y cubrir los vencimientos de deuda inmediatos que tenían sin que llegue dinero a la economía real. Esto está generando muchas quejas de los Gobiernos que, como el de España, piden a la banca que traspase dinero a la economía real.

Sin embargo, lo que determina la magnitud de la crisis actual es precisamente la enorme masa de crédito respaldada, en primer lugar, por unos activos sobrevalorados y, en segundo lugar, por titulizaciones y demás formas de monetización de estos primeros créditos ligados a

[15] Así lo admitió en EE UU Henry Paulson, Secretario del Tesoro estadounidense:"A pesar de nuestro gran plan de rescate, más instituciones financieras van a ir a la quiebra".

activos sobrevalorados. Por lo tanto, es razonable afirmar que para contener la crisis financiera sería necesario un recorte sustancial de la financiación ajena, tanto pública como privada.

De esta forma, las inyecciones de liquidez si bien se han justificado por la necesidad de evitar la quiebra de determinadas instituciones financieras y con ello las consecuencias que podría tener sobre el sistema financiero mundial, se trata de una medida paliativa que no aborda el problema en profundidad y que, incluso, podría ser perjudicial a largo plazo, aunque a corto plazo hayan sido necesarias para evitar un mal mayor: la quiebra del sistema financiero mundial y el estrangulamiento de la economía real.

2. Recapitalizaciones de bancos

Al intensificarse las inyecciones de liquidez efectuadas por los Gobiernos en determinadas instituciones financieras, la segunda medida que se ha ido aplicando ha sido la de entrar a formar parte del capital de determinados bancos a los que se ha ayudado. De esta forma se pretende un mayor control en las entidades que han precisado ayuda estatal; control directo razonable dado que son precisamente sus gestores los que las han llevado a la situación en la que actualmente se encuentran.

La entrada de representantes del Gobierno en los Consejos de las entidades trata de controlar las decisiones adoptadas y es una medida menos drástica que la intervención completa de la entidad. Un argumento adicional es que de esta forma se pueden poner limitaciones a los mecanismos de retribución de los gestores de las entidades, una de las causas de la crisis que vive el sistema[16].

Sería razonable, por lo tanto, establecer sistemas retributivos en función de objetivos a más largo plazo de las entidades (aunque la menor asunción de riesgos derivase en menores rentabilidades). Ya en el 2008 el FMI animó a los bancos a vincular la remuneración de los directivos a objetivos a largo plazo.

Otra característica de esta medida es su temporalidad. En el Reino Unido, por ejemplo, Gordon Brown dejó claro que esta ayuda de fondos no supondrá que el Gobierno será un inversor permanente en los bancos. "El Gobierno no será un inversor permanente. Con el tiempo, tenemos intención de eliminar todas estas inversiones de una forma ordenada".

3. Garantías gubernamentales en las emisiones de deuda

Ante el deterioro de la calidad crediticia de los bancos, y la mencionada subida de los *spreads* de crédito, una de las medidas propuestas por los gobiernos para ayudar a la banca ha sido ofrecer garantía estatal a las emisiones bancarias, de forma que los bancos puedan financiarse a un coste menor. Esta es una medida conocida y aplicada en el pasado, que provoca el conocido efecto de expulsión o *crowding out* afectando negativamente a otros sectores productivos.

Ante la escasez de recursos en el sistema y la aversión al riesgo que se produce en un momento de crisis financiera, los inversores prefieren entregar sus fondos a las emisiones con mayores garantías; en este caso, las bancarias con garantía estatal o bien a las propias emisiones estatales. Esto genera que empresas de otros sectores productivos tengan que pagar elevados *spreads* para colocar su deuda y poder financiarse.

La medida ayuda al sector bancario y, por tanto, puede ser eficaz en la crisis financiera pero en un momento como el actual -en que coexiste con una crisis de la economía real- provoca

[16] Sistemas de incentivos basados en la evolución de la acción a corto plazo han inducido a los gestores de las entidades a tomar decisiones que pudieran beneficiar, también a corto plazo, el precio de la acción: se han avanzado al momento presente beneficios futuros, creando "dinero sobre dinero" sin una base real que lo respalde, y asumiendo apalancamientos máximos con el fin de generar los máximos beneficios posibles a corto plazo. Las entidades recapitalizadas están viendo cómo desde los Gobiernos se incita a los directivos a renunciar a sus *bonus* y las autoridades quieren establecer retribuciones máximas.

efectos perniciosos para los sectores productivos creadores de empleo que pueden verse expulsados de los canales de financiación.

4. Medidas de apoyo específicas para determinados sectores

Cuando la crisis además de afectar al sistema financiero afecta a la economía real -en lo que algunos autores han denominado "la crisis perfecta"- exige también articular medidas de apoyo para determinados sectores productivos más generadores de empleo.

Como acabamos de señalar, medidas como garantías gubernamentales a las emisiones de deuda bancaria pueden generar el efecto expulsión de otros sectores productivos, que pueden verse muy limitados en su acceso a los recursos financieros al no poder competir con la garantía estatal. Por otra parte, el hecho de que haya una crisis de la economía real hace que sea necesario tomar medidas de apoyo a los sectores productivos más creadores de empleo: industria, construcción, automóvil, etc.

Actualmente muchos de estos sectores están solicitando ayudas gubernamentales para poder sobrevivir y evitar despidos masivos. El aspecto clave, en este sentido, es que hay una crisis de demanda generalizada en la economía que hace necesarios ajustes para adaptar la producción a la demanda real[17].

Por otra parte, la urgencia que genera una crisis como la actual, hace que determinados Gobiernos hayan aplicado medidas de apoyo rápidas para determinados sectores, como a la industria del automóvil en el caso de Alemania, sin que se haya realizado previamente un análisis exhaustivo del modelo productivo que cada país debe tener en futuro. Lo importante es determinar el peso que estos sectores deberían tener en el futuro en los modelos productivos de cada país considerando el ajuste que se esta produciendo en la demanda para evitar que estas medidas acaben siendo perjudiciales a largo plazo[18].

5. Compra de activos tóxicos

Otra de las medidas que se han propuesto es la compra de activos- denominados tóxicos- a las entidades financieras para eliminar de su balance los riesgos que ponen en cuestión su solvencia. Esta compra sería asumida por entidades gubernamentales y se caracteriza por tener el objetivo de eliminar de raíz los activos que están en la base de la crisis.

El punto de partida se fecha en la época de bajos tipos de interés vigentes en los mercados durante años, con unas entidades de crédito a las que su negocio se les hacía más pequeño por margen y debían aumentarlo por volumen. La única manera de ampliar negocio era otorgar préstamos a personas de un perfil de riesgo mayor, o a los que otras entidades rechazaran para poder cobrarles un *spread* de interés mayor, o bien aumentar el número de operaciones a través de la titulización, creando dinero basado en dinero sin respaldo de ningún bien.

En este contexto, surgen las hipotecas *subprime* y se intensifican los procesos de titulización de estas hipotecas en las llamadas *mortgage back securities*. Con la globalización, esos activos tóxicos son vendidos por entidades de todo el mundo lo que acaba contaminando sus balances. La propia regulación de control de la solvencia de las entidades financieras -Basilea II- actuaba como estímulo de la titulización.

La compra de estos activos tóxicos trata de eliminarlos de los balances de los bancos, con la contrapartida –negativa- de que los Gobiernos se ven obligados a endeudarse para comprarlos y que un exceso de deuda de los Gobiernos con el fin de comprar estos activos a las entidades

[17] Estos ajustes requieren, a nuestro juicio, una mayor flexibilidad en el mercado de trabajo; de hecho, en economías como la estadounidense, estos ajustes se hacen de forma rápida en contraposición a los mercados de trabajo más rígidos existentes en la UE y, en particular, en España.

[18] Vuelven a plantearse, como es habitual, los conflictos entre los problemas y soluciones de corto plazo y los de una dimensión temporal mayor.

puede poner en riesgo la propia calificación crediticia de los estados: el riesgo no se elimina, sólo se traspasa[19].

Si bien es cierto que la credibilidad de las agencias de *rating* ha quedado en entredicho con la crisis financiera, sus calificaciones siguen siendo las que marcan el *spread* a pagar en las emisiones de deuda, continúan marcando las reglas del juego, y Gobiernos – como el de España- y otras administraciones públicas ya han sufrido rebajas en su calificación crediticia.

6. Reducciones impositivas

Una crisis de dimensión de la actual exige aplicar medidas en todos los ámbitos y, por lo tanto, también fiscales. El problema que se plantea es que casi todas las medidas que se han detallado hasta ahora requieren de una amplia inyección de fondos estatales y los Gobiernos van a entrar en un periodo de déficits presupuestarios. Las rebajas fiscales profundizarían aún más esta tendencia.

No obstante, como ya se ha comentado la gravedad de esta denominada "crisis perfecta" está en el hecho de que coinciden una crisis financiera y una crisis de la economía real: en concreto, asistimos a un proceso de deflación de activos y acusada caída de la demanda. Las reducciones impositivas, siempre que aumenten la renta disponible para el consumo, estimularán la demanda.

En este sentido, un aspecto objeto de reflexión es determinar cuales son los niveles "razonables" de demanda pues entre las razones de la crisis actual está el hecho de que se ha incentivado el consumo inmediato en el corto plazo, avanzando al momento presente decisiones de inversión de toda una vida financiadas a través de un endeudamiento excesivo; esto ha llevado a crecimientos desmesurados de la demanda que han hecho estallar los riesgos del sistema.

El hecho actual es que estamos viviendo un proceso de desapalancamiento de la población, que quiere reducir sus niveles de endeudamiento y prefiere destinar al ahorro los aumentos de la renta disponible. Ello en si puede ser positivo y necesario pero si las rebajas fiscales no se traducen en aumentos de demanda que, a su vez, generen mayor actividad económica y recaudación de impuestos por vía indirecta, las medidas fiscales pueden acabar ahondando más el problema del déficit que están asumiendo los Gobiernos.

7. Inversión en infraestructuras

Este tipo de medidas están directamente relacionadas con la parte de esta "crisis perfecta" que está relacionada con la economía real: el Gobierno invierte para estimular la economía en el contexto de la tradición keynesiana. No obstante, hay que recordar que el propio Keynes considera que las medidas de estímulo de inversión en infraestructuras debían tener un carácter temporal.

Un aspecto destacable de este tipo de medidas es que la inversión en infraestructuras tiene efectos también a largo plazo en la economía, mientras que otras de las medidas relatadas son más limitadas en sus efectos temporales.

Se revisan, a continuación, los distintos planes que se están aplicando en EE UU y los países de la zona euro.

Estados Unidos

En EE UU, el Plan de rescate propuesto por Bush quedó paralizado por el cambio de Presidente, siendo el primer país que se ha propuesto tomarse más tiempo para realizar un diagnóstico de la situación antes de tomar un nuevo paquete de medidas. A cierre del tercer

[19] En este sentido, hay que preguntarse y analizar cuál es el nivel de riesgo que pueden asumir los Gobiernos sin entrar ellos mismos en riesgo.

trimestre de 2008, la deuda norteamericana suponía un 358% de su PIB, por encima del 300% que condujo a EE UU a la Gran Depresión.

Recientemente, el nuevo Presidente Obama ha conseguido sacar adelante su paquete de medidas de apoyo a la economía. En febrero de 2009 el Senado y la Cámara de los Representantes acordaron que el plan definitivo de la Administración Obama contra la crisis esté dotado con 789.000 millones de dólares, unos 40.000 millones menos de lo previsto en un principio.

El plan prevé la creación de 3,5 millones de empleos y destina más de una tercera parte de los fondos a reducciones de impuestos para la clase media. Se dedicará el 35% del paquete (unos 275.000 millones de dólares) a asumir las reducciones fiscales. Otros 150.000 millones serán destinados a inversiones en infraestructuras. Además, el paquete crea un fondo de estabilización para ayudar a los Estados con problemas para cumplir sus presupuestos, entre los que se encuentran algunos de los mayores, como California, Florida y Nueva York.

Se observa que, tras las medidas iniciales adoptadas por el Presidente saliente destinadas a responder a la crisis financiera, el Plan de Obama se centra más en medidas que tratan de estimular la economía real mediante fuertes inversiones en infraestructuras -que consumirán la mayor parte del presupuesto- y recortes de impuestos a los ciudadanos y las empresas.

Es reseñable y merece reflexión el hecho de que se estén abordando políticas de estabilización también para los Estados con problemas pues como se ha comentado debe evaluarse el riesgo máximo que pueden asumir los Estados en un momento como el actual.

Entre otros aspectos, se amplía de 26 a 46 semanas el período de subsidios para desempleados, aunque éste se extiende a 59 en estados con altas tasas de desempleo. También prevé recortes tributarios escalonados de entre 400 dólares para individuos y hasta 800 dólares para familias hasta 2010, dependiendo de sus salarios. Además, autoriza fondos para los Gobiernos estatales, la mayoría de los cuales afrontan graves presiones presupuestarias[20].

Quienes compren casa por primera vez, entre enero y agosto de este año, podrán recibir un crédito tributario de 8.000 dólares, 500 más de lo que estipula la ley actual. Los universitarios recibirían un crédito de hasta 2.500 dólares para su matrícula y otros gastos educativos como la compra de libros y ordenadores.

Es destacable también que, finalmente, parece que el plan ha desterrado una vuelta al proteccionismo más fuerte. Ante presiones del sector exportador y de los socios comerciales de EEUU el plan retiró la controvertida cláusula "*Buy American*" que, salvo algunas excepciones, exige el uso de hierro, acero y productos manufacturados estadounidense en obras de infraestructura financiadas por el plan. Ésta se aplicará, según el texto, "de forma consecuente con las obligaciones de EEUU bajo acuerdos internacionales por lo que las obras públicas podrán contar con productos extranjeros.

La mayoría de los expertos vaticina que el paquete de medidas tomará tiempo para surtir efecto y aunque no frenará la caída de empleos de inmediato sí comenzará a estimular su crecimiento para fines de 2009.

Zona euro

En Europa las medidas tomadas por los distintos países han abarcado desde nacionalizaciones parciales a la extensión de los seguros para los depósitos bancarios. Cada país ha realizado su análisis por separado impulsando medidas diferentes.

Desde el inicio de las turbulencias financieras, las autoridades económicas europeas trataron de establecer un marco de respuesta conjunta y coordinada con otras iniciativas internacionales, particularmente con las recomendaciones del *Financial Stability Forum* (FSF). El plan de trabajo quedó recogido en la hoja de ruta que el Consejo ECOFIN aprobó en octubre de 2007, cuyos objetivos son mejorar el funcionamiento de los mercados (incluyendo la actuación de las

[20] Por ejemplo, California, con un déficit de cerca de 41.000 millones de dólares, recibirá 26.000 millones, y buena parte de esos fondos se destinará a programas de educación y servicios médicos en los próximos 18 meses.

agencias de calificación), incrementar la transparencia, mejorar la valoración de activos y reforzar la supervisión, la gestión de riesgos y la cooperación entre supervisores.

Sin embargo, el extremo agravamiento de la crisis financiera global desde mediados de septiembre requirió una actuación decidida de los Gobiernos que iban anunciando medidas no coordinadas, país por país. No obstante, la cumbre de los países europeos del G 8 -el 4 de octubre- y la posterior declaración conjunta de los Jefes de Estado y de Gobierno de la UE recogieron inicialmente el compromiso político de actuar coordinada y decididamente para garantizar la estabilidad del sistema financiero. Posteriormente -el 7 de octubre- los ministros de economía y finanzas de la UE acordaron tomar todas las acciones necesarias para restablecer la confianza y el funcionamiento adecuado de los mercados. Los ministros decidieron elevar la cobertura mínima del fondo de garantía de depósitos hasta 50.000 euros y establecieron unos principios comunes sobre los que deberían pivotar las iniciativas nacionales: protección de los intereses de los contribuyentes, responsabilidad de accionistas y directivos, y respeto de las reglas del mercado único y de la competencia.

Tras este compromiso, el 12 de octubre los Jefes de Estado y de Gobierno de la zona del euro lanzaron un plan de acción concertado. Mientras que el BCE apoyaba con sus medidas las condiciones de liquidez a muy corto plazo, el plan de los Gobiernos pretende facilitar la financiación a medio plazo y restaurar las ratios de capital de las entidades para que sigan realizando su labor de intermediación en la economía. De forma inmediata, los países miembros dieron a conocer algunas características de los planes nacionales que, en algunos casos, venían a complementar medidas anunciadas o aprobadas con anterioridad.

En primer lugar, para que los mercados de crédito vuelvan a funcionar con normalidad y los bancos se puedan financiar a medio plazo, los Gobiernos de la UEM contemplan sistemas temporales de garantía sobre las nuevas emisiones bancarias de deuda senior (con plazo hasta cinco años). Además, algunos países, como Francia y Austria, ofrecen préstamos garantizados a su sistema bancario, mientras que otros -como España e Italia- consideran la compra o canje de activos de alta calidad que permitan la refinanciación de las entidades.

Por otro lado, el plan también facilita que las entidades restauren sus ratios de capital. Para ello, varios Gobiernos han anunciado inyecciones de capital a entidades solventes y, en caso necesario, inyecciones de emergencia para evitar el posible riesgo sistémico. Todas estas actuaciones deberán cumplir unos requisitos, como tener un carácter temporal y realizarse en condiciones de mercado, entre otros.

Finalmente, el plan de acción concertado en la UEM también incluye una mejora de los mecanismos de cooperación entre países y una flexibilización de los criterios contables. En relación con este último aspecto, la Comisión Europea ya ha emitido un reglamento para que la depreciación de activos no siga introduciendo presiones adicionales sobre las entidades europeas.

En definitiva, se trata de medidas dirigidas a restaurar la confianza y mitigar las presiones de liquidez en el corto plazo y de financiación en el medio plazo, contribuyendo a que el sistema bancario pueda seguir canalizando de forma eficiente los recursos de los ahorradores hacia los individuos y empresas con necesidades de financiación y, por tanto, reduciendo el impacto de las tensiones financieras sobre el crecimiento económico.

En cualquier caso, a medio plazo es preciso poner en marcha las reformas identificadas en la hoja de ruta que aprobó el Consejo ECOFIN en octubre de 2007. Además, la naturaleza global de esta crisis y la creciente interrelación transfronteriza de mercados y entidades requerirán una estrecha cooperación internacional de los gobiernos, los reguladores, autoridades supervisoras y los bancos centrales.

5. INEFICIENCIAS Y OPORTUNIDADES DE ARBITRAJE CREADAS POR LA FALTA DE COORDINACIÓN

Como se ha detallado en el punto anterior, la magnitud de la crisis financiera ha exigido de los Gobiernos la adopción de medidas mayoritariamente coyunturales destinadas a evitar la quiebra de las instituciones financieras, lo que genera ineficiencias en varios sentidos.

1. Ineficiencias de origen

En una economía globalizada se están implantando diferentes medidas para ayudar a las entidades financieras en Norteamérica y dentro de Europa en los distintos países. De esta forma, cada entidad financiera, dependiendo de su país de origen, está recibiendo unas medidas de apoyo u otras y diferentes magnitudes relativas de ayuda financiera[21].

El problema de las ineficiencias de origen es que en una economía globalizada los efectos de las quiebras de determinados bancos afectan a otros muchos bancos de distintos países: los Gobiernos de cada país están decidiendo salvar principalmente a aquellos bancos cuyos problemas repercutirían más en su propio país, con menor atención a los posibles efectos fuera de sus fronteras. Así, el Gobierno francés acudió con rapidez al rescate de Societé Generale, el Gobierno británico al de bancos Royal Bank of Scotland (RBS), Lloyds TSB y HBOS y los Gobiernos de Bélgica, Holanda y Luxemburgo anunciaron una inyección para rescatar a Fortis.

Dado que dependiendo de su origen se salvan unos bancos u otros y que en una economía globalizada el negocio bancario es global, dirigentes de algunos países han llegado a pedir al resto que apliquen medidas similares a las puestas en práctica en sus países para poder devolver la confianza al sistema financiero. En el Reino Unido, Gordon Brown pidió al resto de los gobiernos que ayuden a recapitalizar a sus bancos en dificultades.

2. Ineficiencias de elección

Dentro de cada país, son los Gobiernos los que están decidiendo qué entidades salvar, dejándose quebrar a algunas de ellas (Lehman Brothers, por ejemplo), mientras otras son objeto de ayudas financieras muy amplias. El apremio con que han debido efectuarse las inyecciones de liquidez ha impedido que se hubiera realizado previamente un análisis en profundidad de la viabilidad de cada entidad para determinar sobre una base objetiva qué entidades pueden ser salvadas y cuales no tendrán otra opción que fusionarse con otras o irremediablemente ir a la quiebra.

Cada país, en función de la importancia que se considera podría tener para su sistema la quiebra de una u otra entidad, decide o no apoyarla. También han influido otros factores, como el juicio que los dirigentes políticos han hecho de la dirección de cada una de estas entidades. En EE UU llama la atención como se dejó quebrar a Lehman Brothers o Washington Mutual, mientras se salvaban otras entidades como Fannie Mae o Freddy Mac y se impulsaban fusiones bancarias con respaldo gubernamental entre otras entidades. EE UU sí acudió al rescate de su primer banco, Citigroup.

3. Ineficiencias normativas

Las ineficiencias normativas pueden ser de dos tipos. En primer lugar, las que se producirán en el futuro como consecuencia de las distintas medidas que están adoptando los Gobiernos de cada país ante la crisis financiera. En lo relativo a la normativa bancaria, la ausencia de un regulador global implica que cada país impulsa medidas reguladoras de forma aislada, lo que genera marcos regulatorios del sistema financiero más o menos favorables dependiendo del país: el caso de las distintas garantías dadas a los depósitos bancarios por cada país de Europa es un claro ejemplo de ello.

[21] En Francia, Sarkozy insistió en que el Estado francés no dejará quebrar ningún banco y que, "en caso de urgencia", podrá aportar directamente su garantía a la refinanciación de los bancos con problemas. En ese caso, como ocurrió con el franco-belga Dexia, el Estado asumirá el control y cambiará la dirección del mismo. En EE UU, se ha realizado una selección de los bancos a los que salvar y los que dejar quebrar.

En segundo lugar están las ineficiencias normativas que se derivan del distinto marco legal existente en cada país: la magnitud de los problemas de solvencia y liquidez de cada banco se valora en su país de origen en función de su propia normativa nacional.

Por ello, están surgiendo propuestas hacia sistemas universales de supervisión que puedan evitar en el futuro estas ineficiencias normativas[22].

4. Ineficiencias competitivas

La necesidad de tomar medidas rápidas de apoyo a las instituciones financieras hace que, en cierta medida, se esté apoyando a los gestores de las entidades que han entrado en crisis y que en el mercado están recibiendo un severo castigo. Los críticos de las medidas gubernamentales temen que éstas pudieran poner en riesgo el dinero de los contribuyentes y alentar a las instituciones financieras a tomar riesgos desmedidos con la convicción de que el gobierno solucionará sus problemas.

El apoyo recibido por las entidades puede entenderse como que los Gobiernos actuarán siempre como benefactores últimos; de esta forma no se está favoreciendo la eficiencia y competitividad de las entidades, dado que los errores de gestión acaban siendo difuminados por las ayudas gubernamentales. En determinados casos no se deja actuar al mercado y se evita la quiebra de entidades que no son competitivas[23].

En este sentido, Juergen B. Donges se pregunta cómo se evitará la aparición del llamado "riesgo moral" que crean las instituciones estatales cuando actúan como "prestamista de última instancia" y además se comprometen a impedir que grandes bancos quiebren, lo que podría inducir en el futuro a las entidades financieras nuevamente a asumir mayores riesgos en sus operaciones que los que pueden gestionar con eficacia.

6. REFLEXIONES FINALES

Siguiendo a Juergen B. Donges, un intervencionismo estatal desmesurado en la economía tal vez pueda dar seguridad a corto plazo pero no resuelve los problemas existentes, sino que sólo los traslada al futuro. Sin embargo, una apuesta, por el empresario emprendedor y por políticas económicas que creen condiciones favorables a la iniciativa privada orientarían un futuro próspero para las economías y la sociedad. Precisamente es en tiempos de crisis cuando hay que comprometerse con las reformas estructurales.

A pesar de un cierto sentimiento anti-mercado reciente, Europa disfrutará de más crecimiento, riqueza y empleo en el futuro si se mantiene cerca del capitalismo reformado. Tanto la competencia como proceso de descubrimiento (Hayek), como la evolución del mercado a través de los procesos de destrucción creativa (Schumpeter) son elementos indispensables para una economía dinámica. Una de las lecciones de esta crisis económica y financiera es recordarnos la necesidad de invertir en actividades de investigación, desarrollo e innovación (I+D+i) con el fin de abrir nuevos mercados tanto en el propio país como en el extranjero, con productos de calidad, de contenido tecnológico alto y con buenos atributos medioambientales.

No hay que olvidar que actualmente estamos asistiendo a un proceso de transferencia de los riesgos asumidos por el sector financiero en pasados años a los Estado pero no se están tomando aún medidas para mitigar y controlar los riesgos que han llevado a la situación actual. El riesgo se está traspasando a los Estados, y es necesario preguntarse el nivel máximo que pueden asumir y la conveniencia de retirar a tiempo estímulos que resulten excesivos.

---

[22] En el Reino Unido el primer ministro Brown ha instado a los Gobiernos del mundo a adoptar reglas trasnacionales sobre transparencia, así como las más estrictas normas de conducta para garantizar que no vuelven a asumirse riesgos totalmente irresponsables. Según Brown, debería también ponerse en marcha un nuevo sistema de supervisión para hacer frente a los enormes flujos de capitales entre las distintas instituciones financieras. El Foro de Estabilidad Financiera y un Fondo Monetario Internacional "reformado" deberían también jugar un papel "*no sólo en la resolución de la crisis sino también en su prevención*", agrega Brown.

[23] En este sentido, un debate de actualidad en España es el relativo al funcionamiento de las Cajas de Ahorro, entidades de crédito peculiares por sus objetivos, sistema de gobierno, ausencia de ánimo de lucro, y que no tienen en la competitividad la base de su modelo de negocio.

En un futuro cercano, el sector financiero deberá asumir de nuevo los riesgos propios de su actividad, pero con un mayor control y regulaciones globales que permitan acotar las actuales ineficiencias mencionadas en este artículo; ineficiencias que están presentes desde el origen de la crisis financiera y las medidas aplicadas para solventarla y cuyo estudio, evaluación cuantitativa y corrección abre nuevas líneas de investigación.

7. BIBLIOGRAFÍA

ABADÍA, LEOPOLDO, (Enero 2009). "La crisis Ninja y otros misterios de la economía actual". http://leopoldoabadia.blogspot.com/search/label/%2B%20ANEXO%201%20Crisis%20NINJA

BANCO DE ESPAÑA. (2008a). Informe Anual 2007. http://www.bde.es/informes/be/infanu/infanu.htm

BANCO DE ESPAÑA. (2008b). Informe de Estabilidad Financiera 11/2008. http://www.bde.es/informes/be/estfin/completo/IEF200811.pdf

BANCO DE ESPAÑA, Boletín Económico, (Octubre 2008). http://www.bde.es/informes/be/boleco/2008/be0810.pdf

BANK FOR INTERNATIONAL SETTLEMENTS, (Noviembre 2008). "Supervisory guidance for assessing banks' financial instrument fair value practices".. http://www.bis.org/publ/bcbs145.htm

BCE. (2008a). "The Implementation of Monetary Policy in the Euro Area", General Documentation on Eurosystem Monetary Policy Instruments and Procedures. http://www.ecb.int/pub/pdf/other/gendoc2008en.pdf?0dd0816ab94907f0c24ac1c534ae0f06

BCE. (2008b). Recommendations of the Governing Council of the European Central Bank on government guarantees for bank debt, 20 October 2008. http://www.ecb.int/pub/pdf/other/recommendations_on_guaranteesen.pdf

BCE. (2008c). Recommendations of the Governing Council of the European Central Bank on the pricing of recapitalisations, 20 November 2008. http://www.ecb.int/pub/pdf/other/recommendations_on_pricing_for_recapitalisationsen.pdf

BCE. (2008d). Proyecciones macroeconómicas elaboradas por los expertos del eurosistema para la zona del euro, 4/12/2008. http://www.ecb.int/pub/pdf/other/eurosystemstaffprojections200812es.pdf

BLANFEIN, LLOYD. (2009). "Do not destroy the essential catalyst of risk". Financial Times, 8 de febrero de 2009. http://www.ft.com/cms/s/0/0a0f1132-f600-11dd-a9ed-0000779fd2ac.html

DE RATO, RODRIGO. (2009). "La globalización, los flujos financieros y la crisis: lecciones de la experiencia", Discurso de investidura como Doctor Honoris Causa por la Universidad Rey Juan Carlos de Madrid, Madrid.

DONGES, JUERGEN B, (2008-2009). "No hay un conflicto entre controlar la inflación y estimular el crecimiento económico". Revista Vida Académica de San Telmo. Programa de Continuidad www.santelmo.org/revista/n31/VA_continuidad.pdf

FERNÁNDEZ, MIGUEL. (2008a). Comparecencia del Gobernador ante la Comisión de Presupuestos del Congreso, Madrid, 7 de octubre de 2008. http://www.bde.es/prensa/intervenpub/gobernador/mfo071008.pdf

FERNÁNDEZ, MIGUEL. (2008b). Comparecencia del Gobernador del Banco de España ante la Comisión de Presupuestos del Senado, Madrid, 19 de noviembre de 2008. http://www.bde.es/prensa/intervenpub/gobernador/mfo191108.pdf

FINANCIAL STABILITY FORUM, (Junio 2008). "Update on the Implementation of the FSF's Recommendations. Report by the FSF Chairman to the G8 Finance Ministers". http://www.fsforum.org/publications/r_0806.pdf

FINANCIAL STABILITY FORUM, (Septiembre 2008), " Ongoing and Recent Work Relevant to Sound Financial Systems".. http://www.fsforum.org/publications/on_0809.htm

FINANCIAL STABILITY FORUM, (Septiembre 2008). "Report of the Financial Stability Forum on Enhancing Market and Institutional Resilience Follow-up on Implementation". http://www.fsforum.org/press/pr_081009f.pdf

FMI. (2008). "Global Financial Stability Report. Financial Stress and Deleveraging: Macrofinancial Implications and Policy", World Economic and Financial Surveys, octubre de 2008. http://www.imf.org/external/pubs/ft/gfsr/2008/02/pdf/text.pdf

PÉREZ RAMÍREZ, JORGE (2007). "Los derivados de crédito". Publicado por el Banco de España. http://www.bde.es/informes/be/estfin/numero3/ef0303.pdf

ROLDÁN, JOSÉ M..(2008). "El papel del modelo de "originar para distribuir" en la crisis financiera de 2007", Revista de Estabilidad Financiera, Banco de España, nº 15, noviembre de 2008. http://www.bde.es/informes/be/estfin/numero15/ief0115.pdf

ROUBINI, NOURIEL. (2009). "Nouriel Roubini on Prospects for 2009". Financial Times, 3 de febrero de 2009. http://www.ft.com/cms/s/0/89829f7a-f1d1-11dd-9678-0000779fd2ac.html

TILMAN, LEO M. (2008). "Financial Darwinism: Create Value or Self-Destruct in a World of Risk". ISBN 978-0-470-38546-3.http://www.financialdarwinism.com/?cat=3

UBIDE, ÁNGEL. (2008). "Anatomy of a modern credit crisis", Revista de Estabilidad Financiera, nº 14, Banco de España, mayo de 2008. http://www.bde.es/informes/be/estfin/numero14/IEF200814-4.pdf

WORLD ECONOMIC FORUM (2009), "World Economic Forum Annual Meeting 2009 Shaping the Post-Crisis World Davos-Klosters, Switzerland 28 January – 1 February" http://www.weforum.org/en/media/publications/index.htm

THE EX-DIVIDEND DAY ANOMALY IN THE SPANISH STOCK MARKET.

JOSEP GARCÍA BLANDÓN[*]
E-mail: josep.garcia@iqs.edu
MÒNICA MARTÍNEZ BLASCO
LUCINIO GONZÁLEZ SABATÉ
Facultat d'Economia, IQS
Universitat Ramon Llull

Resumen

Este trabajo investiga el comportamiento de las rentabilidades de las acciones y los volúmenes de negociación alrededor de las fechas ex-dividendo, en el mercado bursátil español, utilizando la metodología de eventos. Existe un amplio consenso en la literatura sobre el hecho de que los precios de las acciones caen en menos cuantía que el dividendo pagado en la fecha ex-dividendo. Este comportamiento que indicaría una preferencia por las ganancias de capital respecto a los dividendos, se explica en términos de ventajas fiscales. Contrariamente al consenso existente no observamos rentabilidades anormales en la fecha ex-dividendo. Este hallazgo es consistente con el hecho de que en la actualidad los dividendos y las ganancias de capital en España tributan al mismo tipo fiscal. Se observa sin embargo, volúmenes negociación anormalmente elevados

Palabras Clave: Fecha ex-dividendo, rendimientos anormales, volúmenes anormales, estudio de eventos.
Área Temática: Economía Monetaria y Financiera.

Abstract

This paper investigates the behavior of stock returns and trading volumes around ex-dividend dates, in the Spanish stock market, following the event studies methodology. There is a big consensus in the literature about the fact that stock prices fall by less than the dividend paid on ex-dividend days. This behavior indicates a preference for capital gains over dividends, generally explained in terms of tax advantages. Contrary to the existing consensus we do not observe significant abnormal returns on ex-dividend days. This finding is consistent with the fact that nowadays dividends and capital gains in Spain are taxed at the same rate. In addition, abnormally high trading volumes are observed around ex-dividend dates, especially for high yield stocks.

Key-words: Ex-dividend days, abnormal returns, abnormal volumes and event studies.
Thematic Area: Monetary and Financial Economics.

1. INTRODUCTION

In the academic field of corporate finance, few topics if any have received more attention than dividend policy. The understanding of the factors that determine the firm's dividend policy has challenged financial economists for various decades. However, despite the huge number of investigations in the field, dividend policy has been, and still is, a puzzling question. Lintner (1956) undertakes a pioneer empirical investigation about the determinants of dividend policy by interviewing a set of firm's managers. The author concludes that firms establish long term

[*] Via Augusta, 390, 08017, Barcelona, Spain. Ph.: +34 932 672 000. Fax: + 34 932 056 266. E-mail: josep.garcia@iqs.edu.

benchmark for the dividend distribution rate. Since managers are usually reluctant to reduce dividend payments, when benefits increase, dividends increase at a lower rate, trying to converge to the target payout. One year before, Campbell and Beranek (1955) investigate the effects of dividend payments in stock prices of a small sample of companies quoted in the New York Stock Exchange, reporting that prices dropped 90% of the dividend paid on ex-dividend dates. The authors, however, do not offer any explanation for this fact. Nevertheless, the crucial moment for the theoretical modelling of dividend policy arrives with the classic paper by Miller and Modigliani (1961) establishing the irrelevance of dividend policy. The authors propose a one-period model with perfect capital markets, no frictions such as transactions costs and taxes and under certainty conditions. Assuming that dividend policy will affect neither the firm's investment program nor its capital structure, higher dividends would require the issue of new stock without affecting the value of the firm. Thus, a dollar paid in dividends should cause a dollar drop in the price of the stock. However, when taxes are considered, Elton and Gruber (1970) support a tax effect hypothesis: since dividends are usually taxed at a higher rate than capital gains, the drop of the stock price should be lower than the dividend paid to make investors indifferent between selling the stock cum-dividend and holding the stock, obtain the dividend and selling the stock ex-dividend. Therefore, the higher the difference between dividends and capital gains tax rates, the lower the drop of the price will be. In fact, they argue that stockholders marginal tax rates could be inferred from the ex-dividend day stock price behaviour. Elton and Gruber investigate stock price adjustment on ex-dividend dates for the period: April 1966-Mars 1967, observing that stock prices drop on average the 77.7% of the dividend paid. By sorting the sample into deciles according with the firm's dividend yield, the authors find firms with higher yields showed higher adjustment in prices on ex-dividend dates. They interpret this result as a proof of a tax-clientele situation. Since investors in shares with high dividend yield should have relatively low marginal tax rates, the dividend value after taxes should not be far to the dividend paid. On the contrary, investors with high marginal tax rates would likely invest on shares with low dividend yields. Since them, many studies have been performed in different stock markets (with different tax regimes for dividends and capital gains) following Elton and Gruber methodology, most of them supporting the tax-effect hypothesis. Among these papers we can mention Poterba and Summers (1984), Robin (1991), and Graham et al. (2003) for the US, Athanassakos (1996) for Canada, Hayashi and Jagannathan (1990) for Japan, and Poterba and Summers (1985) for the United Kingdom. A straightforward way for testing the tax-effect hypothesis is by investigating stock price behaviour around ex-dividend days in countries or periods where neither dividends nor capital gains were taxed. In this line, Frank and Jagannathan (1998) and Yahyaee et al. (2007) observe the ex-dividend day anomaly in Hong Kong and Oman respectively, questioning the tax-clientele hypothesis.

Although the tax-clientele hypothesis constitutes the most widely accepted explanation for the ex-dividend price behaviour, financial economists have explored other possible explanations, alone or in most cases interacting with taxes. In this line, Kalay (1982) constitutes one of the most important contributions, by arguing that stockholders' marginal tax rates could not be inferred from the ex-dividend-day stock price behavior and by offering an alternative explanation for the fact that stock prices did not fully adjust the dividend paid. The author, without disregarding the importance of taxes, poses an alternative explanation known as the "short-term trading hypothesis". Without transaction costs, investors with the same tax rate on dividends and capital gains can buy the stock cum-dividend and sell it ex-dividend if the ex-dividend stock price drop is less than the dividend payment. If the received dividend and the tax savings from the capital loss exceed transaction costs, such a strategy will be profitable. Accordingly, without transaction costs, prices drop should equal the dividend paid. The short-term trading hypothesis is also consistent with the two main results observed by Elton and Gruber (1970): stock price drops less than the dividend payment and positive correlation between price drops and dividend yields. Focusing on volumes instead that on returns, Lakonishok and Vermaelen (1986) report higher trading volumes before and after ex-dividend days. They also observe higher volume increases for those stocks with higher yields. Furthermore, stocks experienced abnormal price increases before ex-dividend days and

abnormal price decreases afterwards. While abnormal price increases positively depend on dividend yield, abnormal trading volumes negatively depend on it.

Other papers suggests that factors different from taxes or transaction costs cause prices to drop less than the amount of dividend paid on ex-dividend days. These papers focus on microstructure arguments. One of them, Bali and Hite (1998), states that price discreteness rather than taxes or transaction costs is the responsible of this anomaly. In their model, since variation in stock prices were limited to multiples of a tick, the ex-day price drop would be the dividend paid rounded to the next smaller tick, thus lower than the amount of the dividend. This explanation is known as the price-discreteness hypothesis. Another paper focusing on microstructure arguments is Frank and Jagannathan (1998). They point out that because the collection and reinvestment of dividends is annoying for small investors they would not buy shares before ex-dividend days but rather afterwards. Since this is not the case for market makers, most transactions will occur at the ask price before the ex-dividend date and at the bid price later. Accordingly, the observed drop in stock prices by less than the dividend paid will be caused by the bid-ask bounce. The authors investigate the issue in the Hong Kong Stock Market, were neither dividends nor capital gains were taxed, observing that stock prices adjust approximately the half of the dividend paid on ex-dividend days. Jakob and Ma (2004) investigate both, the price discreteness hypothesis and the bid-ask bounce, concluding that when price discreteness was eliminated ex-dividend prices drop even less, unlike the price discreteness hypothesis predict. Their results, however, support the bid-ask bounce as an explanation of the ex-dividend price anomaly. In the same line, Yahyaee et al. (2007) investigate the issue in the Oman stock market, where taxes on dividends or capital gains did not exist and stock prices had been decimalized. They find prices drop by less than the dividend paid, suggesting the bid-ask bounce as the most suitable explanation.

More that thirty years later than their first investigation, Elton et al. (2002) investigate the ex-dividend anomaly on a sample containing municipal bond funds (where dividends are tax-advantaged) for which the ex-dividend price drop should be greater than the dividend if taxes mattered, as well as a set of securities (taxable bonds and domestic common stock funds) for which the drop should be, in general, less than the dividend. Their results support the tax effect hypothesis and are inconsistent with market microstructure arguments.

Three decades ago, Black (1976) stated: ‘‘The harder we look at the dividend picture, the more it seems like a puzzle, with pieces that just don’t fit together’’ (p. 5). Later on, Allen and Michaely (1995) conclude that much more empirical and theoretical research on the subject was required before a consensus could be reached. Elton et al. (2002) argue that since their article on taxes and ex-dividend behavior published in 1970, over 100 articles have appeared in leading journals of financial economics examining the issue. Nowadays the number of articles is clearly much higher. Nevertheless, the question is still a puzzle. In a recent review paper on dividend policy, Bhattacharyya (2007) states: “Despite decades of study, we have yet to completely understand the factors that influence dividend policy and the manner in which these factors interact” (p.1). In the same line, the classical Brealey, Myers and Allen (2006) textbook list the dividend issue as one of the ten important unsolved problems in finance, and explicitly mention the necessity of more empirical research.

This paper contributes to the existing literature on the ex-dividend day anomaly in various ways. First, we investigate the Spanish stock market, where there is only a previous investigation on the issue, covering the period 1980-1992. Since financial markets microstructure is one of the most accepted explanations of the ex-dividend day anomaly, empirical papers investigating the issue in markets other than the US should be especially welcomed. Not only taxes, but also the characteristics of the Spanish Stock Exchange have changed dramatically in the last 25 years, making it necessary, in our opinion, to update the dataset. The fact that after the last tax reform, dividends and capital gains are taxed at the same rate, while in the 1980-1992 period, capital gains had a better tax treatment compared with dividends, made in our opinion our investigation particularly interesting. Secondly, our method is different to the one proposed first by Elton and Gruber (1970) and later used by most empirical papers. Unlike them, we follow the Brown and Warner (1985) methodology for event studies. The usefulness of event studies, according to Kothary and Warner (2005), arises from the fact that the magnitude of abnormal performance provides a measure of the (unanticipated) impact of this type of event, in our case dividend payments, on stocks returns. Last but not least, a final contribution of the paper is that we have investigated not only price movements but also

trading volumes around ex-dividend days. The joint consideration of returns and volumes provides better possibilities to explain the results.

The remaining of the paper is as follows: next section discusses the methodology we propose to investigate the ex-dividend day anomaly in the Spanish stock market. Section 3 and 4 present the dataset and a summary of the Spanish tax system for dividends and capital gains. Then, results are analyzed in the context of previous investigations. The final section reports the main conclusions of the investigation.

2. METHODOLOGY

Elton and Gruber (1970) were the first to propose a plausible explanation for the observed anomalous behavior of stock prices around ex-dividend days. Since them, many authors have followed their approach to investigate the issue in other stock markets and/or time periods. Their methodology can be expressed by equation (1), where P_c is the cum-dividend price, P_c the ex-dividend price and P_0 the opening price of the last day before the ex-dividend day. Finally, t_g and t_d are the marginal tax rates for capital gains and dividends, respectively. Under equation (1), two strategies consisting on, on one hand, sell the stock cum-dividend and on the other obtain the dividend and sell the stock ex-dividend will provide the same net return.

$$P_c - (P_c - P_0) \times t_g = P_c - (P_e - P_0) \times t_g + D \times (1 - t_d) \qquad (1)$$

Equation (1) can be expressed as,

$$\frac{(P_c - P_e)}{D} = \frac{(1 - t_d)}{(1 - t_g)} \qquad (2)$$

From equation (2) we conclude that price drop should reflect marginal tax rates of marginal investors. In addition, marginal tax rates could be inferred by observing stock prices drops in terms of dividends.

Most empirical papers investigating ex-dividend price behavior have applied the Elton and Gruber (1970) framework, some of them, as Kalay (1982), incorporating other factors as transaction costs. Other papers, as Kato and Loewenstein (1995), Naranjo et al. (2000) and Yahyaee et al. (2007) propose a regression framework, where abnormal returns are regressed against a number of factors, including dividend yield, abnormal trading volume, transaction costs, etc.

In this paper, we propose the classical Brown and Warner (1985) event studies methodology, being the payment of a cash dividend the studied event. According to previous research, ex-dividend price behaviour is expected to depend on the stock dividend yield. Thus, we have selected two sets of stocks, with high and low dividend yield, to carry out our investigation. Abnormal returns have been computed as the difference between actual and normal returns, while normal returns are defined as expected returns without conditioning on the occurrence of the event.

Thus, abnormal return for stock i on day t will be expressed as,

$$AR_{it} = R_{it} - E(R_{it} | x_t) \qquad (3)$$

Where AR_{it} is the abnormal return of stock i on day t, R_{it} is its actual return calculated as $\ln\left(\frac{P_t + D_t}{P_{t-1}}\right)$, where P_t and D_t are the closing price and the dividend paid on day t respectively, and $E(R_{it} | x_t)$ its expected return for day t. Finally, X_t is the conditioning information set for the expected return on day t. There are two predominant models to compute expected or normal returns: the constant mean return model and the market model. While the first assumes normal return is the mean return over a given period, the market model assumes normal return is given by a linear relationship between the stock return and the market return. In

this paper, expected returns have been estimated throw three different models: the already mentioned constant mean (model 1) and market model (model 3), and a variation of the constant mean model, where average returns have been calculated taking into account the day of the week (model 2). The last model has been introduced because some papers (eg. Garcia (2007)) have provided evidence of a day of the week effect in the Spanish Stock Exchange.

For model 1, normal daily return for each event window has been computed through the average daily return of the estimation period: [-90, -20]. For model 2, a similar approach has been followed but five average returns have been calculated, one for each weekday. Then, the abnormal return for day t has been computed through the difference between its actual return and the average return that corresponds to that day. For example if day t is Monday, we use the Monday average return and so on. Regarding the market model, daily market returns have been computed through the *Indice General de la Bolsa de Madrid* (IGBM) and normal returns have been estimated through ordinary least squares. The use of the three proposed models will provide our research with more robust results. The selected event period, that is, the time period where the effect of dividend payments on returns will be investigated, starts on day -10 and ends on day +20, around the ex-dividend date (day 0). Normal returns have been estimated using the same estimation period [-90, -20].

After estimating daily average abnormal returns for each firm, the average abnormal return on day t for each sample $\left(\overline{AR_t}\right)$ was calculated:

$$\overline{AR}_t = \frac{1}{N}\sum_{i=1}^{N} AR_{it} \qquad (4)$$

Cumulative average abnormal return $\left(\overline{CAR_t}\right)$ has been obtained by adding the average daily abnormal return for different intervals through the event window [-10, +20].

$$\overline{CAR} = \sum_{t=a}^{b} \overline{AR}_t \qquad (5)$$

Our first hypothesis is that average abnormal return will be zero on the ex-dividend day for the two samples of stocks.

Lakonishok and Vermaelen (1986) suggest that the exam of stock trading volumes around ex-dividend dates can provide important information about the causes of the ex-dividend day anomaly. They argue that if short-term traders play a major role on ex-dividend price behavior, a net increase in trading volumes should be observed around ex-dividend days. Therefore, we have investigated stock trading volumes around ex-dividend days with a similar framework as the one we have used to investigate abnormal returns. The abnormal trading volume (AV) of a given stock has been estimated as the difference between its actual and normal trading volumes, all in euro values. Following, among others Lakonishok and Vermaelen (1986), Athanassakos (1996) and Koski and Scruggs (1998), we have estimated normal trading volumes through the mean-adjusted model. In addition, to be consistent with our approach to estimate normal stock returns, we have also estimated normal trading volumes through a day or the week mean-adjusted model. Thus, we define abnormal trading volume, for stock i on day t, AV_{it} as:

$$AV_{it} = \frac{V_{it}}{\left(\sum_{t=-20}^{t=-50} V_{jt} + \sum_{t=30}^{t=60} V_{jt}\right) * \frac{1}{62}} \qquad (6)$$

Where, V_{it} is the traded volume in euros of stock i on day t. As we did with returns, once abnormal daily volumes have been computed for each firm i, the average abnormal volume on day t $\left(\overline{AV_t}\right)$ is calculated for the two samples as:

$$\overline{AV}_t = \frac{1}{N}\sum_{i=1}^{N} AV_{it} \qquad (7)$$

Our second hypothesis is that the average abnormal trading volume will be zero on the ex-dividend day.

Our two hypotheses have been tested through the Brown and Warner (1985) methodology. As usually, the standard deviation of each sample returns has been used to assess the

significance of the event-window average abnormal return, through the z statistic. In addition, given the small size of our stock samples, Corrado (1989) non-parametric rank test has been performed.

3. DATA AND SAMPLE SELECTION

As we discussed in the previous section, we investigate the behavior of stock prices around ex-dividend dates taken into account the stock dividend yield. Accordingly, we have selected two sets of stocks. The first one is formed by high dividend yield stocks, and the second by dividend paying stocks with low dividend yield. In order to minimize potential problems associated to the homogeneity of the samples, all selected stocks belong to the IBEX-35 index. On the one hand, the set of high yield stocks is formed by those stocks belonging to the IBEX-Top Dividend index, with a dividend yield higher than four percent in each year of the period 2006-2008. On the other hand, the set of low yield stocks includes those paying dividend stocks that do not belong to the IBEX-Top Dividend index and with a dividend yield below the two percent in each year of the period 2006-2008. We have removed from the samples those companies that pay dividends more than twice a year to avoid the overlapping of estimation periods and windows events. As a result of applying the former conditions, the group of high yield stocks is formed by: ACS, Endesa, Enagás, FCC, Mapfre, Repsol, Telefónica and Telecinco, while the low yield group includes: Abengoa, Cintra, Gamesa, Gas Natural and Red Eléctrica Española. During this period we have forty-one dividend payments: eleven for the low yield group and thirty for the high yield sample.

Our period of investigation is from June 2006 to June 2008. Although there are some examples of investigations using periods of similar length, among them the pioneer paper by Elton and Gruber (1970), the election of a relatively short-length period of investigation responds to two main reasons, both involving methodological issues. The first one has to do with the event study approach we follow. As Kothary and Warner (2005: pg.17) stated, horizon length has a big impact on event study test properties. "First, short-horizon event study methods are generally well-specified, but long-horizon methods are sometimes very poorly specified (….). Second, short-horizon methods are quite powerful if (but only if) the abnormal performance is concentrated in the event window. In contrast to the short-horizon tests, long-horizon event studies (even when they are well-specified) generally have low power to detect abnormal performance". The second reason specifically involves the tax clientele hypothesis. Since differential taxation between dividends and capital gains constitutes the most widely accepted explanation of the ex-dividend day anomaly, the period of investigation should be characterized by certain stability in the taxation of dividends and capital gains.

4. SPANISH TAXATION ON DIVIDENDS AND CAPITAL GAINS

The last reform of the Spanish tax system for dividends and capital gains was conducted through the *Ley 35/2006, de 28 de noviembre*. One of the purposes of the reform was to minimize tax induced distortions on tax payers' investment decisions. Thus, investments returns after taxes should respect the pre-taxes scheme. Under the new tax system, dividends and capital gains are taxed at a fixed 18% rate. Before the reform, capital gains generated in more than one year were taxed at a fixed 15% rate, while capital gains generated in one year or less were taxed at the investor marginal tax rate. On the other hand, dividends were taxed at the investor marginal tax rate, although there was a deduction for double taxation consisting in the 40% of the received dividend multiplied by 1.4. After the reform, this deduction has been substituted by a lineal exemption of the first 1.500 € of dividends. Thus, high income taxpayers enjoy comparatively lower double taxation benefits.

Although it is arguably if the tax reform has achieved its neutrality goal, we can easily agree that the Spanish tax system is now more neutral than it was before the reform. Therefore, the usually mentioned tax advantage of capital gains over dividends is now very limited, beyond the

obvious fact that is the investor who decides when capital gains will be taxed, because capital gains are taxed when the asset is sold, while this is not the case with dividends.

On the other hand, corporations enjoy deductions for double taxation of dividends that are generally established in the 50% of the dividends received but which can reach the 100% under certain conditions. In the last situation, the double taxation of dividends is completely eliminated.

5. RESULTS

Tables 1 and 2 present average abnormal returns and cumulative average abnormal returns for low and high dividend yield stocks respectively. Both returns have been computed through the three estimation methods discussed in section 2. Results obtained with the three models are highly consistent for the low yield sample but not so much so for the high yield sample. As expected, the market model (model 3) provides more precise results, since the variance of the average abnormal return is lower compared with the other two models. This is not surprising since stocks in the sample, all included in the IBEX-35 index, are expected to show an important market component, whose variability is undertaken and therefore eliminated by the market model. This result supports Cable and Holland (1999) findings' that the market model performs better than other models proposed in the literature. Accordingly, we will focus our discussion on the results obtained with the market model.

Table 1. Average abnormal returns for low yield stocks

	Model 1		Model 2		Model 3			
t	$\overline{AR_t}$	Z_t	$\overline{AR_t}$	Z_t	$\overline{AR_t}$	Z_t	Corrado	$\overline{CAR_t}$
-10	0.43	0.66	0.62	1.02	0.18	0.34	0.24	0.18
-9	0.64	0.99	0.70	1.14	0.75	1.43	1.43	0.93
-8	-0.12	-0.18	-0.31	-0.51	-0.49	-0.92	-0.65	0.45
-7	0.19	0.29	0.28	0.46	0.35	0.66	1.00	0.79
-6	-1.13	-1.74	-1.20	-1.97 *	-0.95	-1.81	-0.99	-0.16
-5	-0.34	-0.53	-0.15	-0.25	-0.54	-1.02	0.08	-0.69
-4	0.01	0.02	0.03	0.04	0.22	0.42	0.29	-0.47
-3	0.72	1.11	0.53	0.87	0.15	0.28	-0.02	-0.32
-2	-0.86	-1.33	-0.85	-1.40	-1.29	-2.45 *	-2.91 **	-1.61
-1	1.68	2.59 *	1.52	2.50 *	1.58	3.00 **	2.94 **	-0.04
0	**-0.14**	**-0.21**	**-0.01**	**-0.01**	**-0.21**	**-0.40**	**-0.42**	**-0.25**
1	-0.25	-0.39	-0.30	-0.49	0.28	0.53	0.78	0.03
2	-0.43	-0.67	-0.56	-0.92	-0.65	-1.24	-0.52	-0.62
3	-0.24	-0.37	-0.17	-0.29	-0.14	-0.27	-0.89	-0.76
4	0.29	0.44	0.24	0.40	0.26	0.50	-0.25	-0.50
5	-0.09	-0.14	0.04	0.07	0.53	1.00	-0.05	0.03
6	-0.30	-0.46	-0.34	-0.56	-0.07	-0.13	0.07	-0.04
7	0.30	0.47	0.18	0.29	0.66	1.26	1.81	0.62
8	-0.04	-0.07	0.02	0.04	0.03	0.06	0.27	0.65
9	-1.03	-1.58	-1.07	-1.76	0.07	0.13	0.66	0.72
10	-0.38	-0.59	-0.25	-0.41	-0.29	-0.55	-1.01	0.43
11	0.13	0.20	0.08	0.14	-0.34	-0.64	-0.38	0.09
12	1.16	1.78	1.03	1.69	0.50	0.95	0.54	0.59
13	-0.97	-1.49	-0.90	-1.48	-0.18	-0.35	0.03	0.41
14	0.74	1.15	0.70	1.15	0.07	0.13	0.40	0.48
15	0.23	0.36	0.36	0.60	0.22	0.42	0.38	0.70
16	0.37	0.56	0.32	0.53	-0.22	-0.42	-1.13	0.48
17	0.86	1.32	0.69	1.14	0.49	0.93	1.06	0.97
18	0.32	0.50	0.41	0.68	-0.07	-0.13	0.19	0.90
19	0.65	0.99	0.56	0.92	0.65	1.24	0.81	1.55
20	-0.11	-0.16	0.06	0.09	0.32	0.61	0.10	1.87

*Significant at a 0.05 level. **Significant at a 0.01 level.

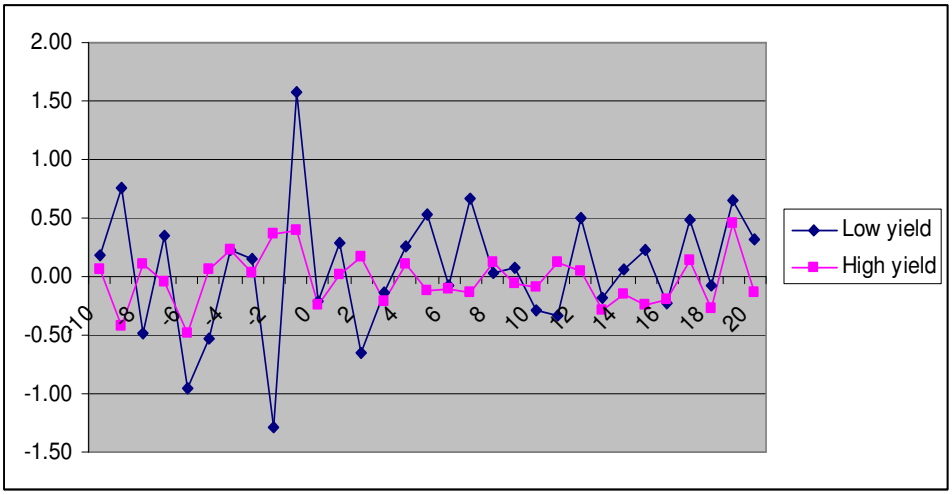
Table 1 shows non-significant abnormal returns for low dividend yield stocks on the ex-dividend day $(t=0)$. Therefore, our null hypothesis that mean abnormal return on the ex-dividend day was zero can not be rejected for the low dividend yield sample in any of the three models. This result is somehow surprising since it contradicts most of empirical evidence worldwide showing that stock prices fall less than the dividend paid. In our case, ex-dividend days are not different than ordinary days in terms of abnormal returns. In addition, although non-significant, the ex-dividend day abnormal return is negative while according to previous research should be positive. Interestingly, a positive and significant abnormal return is also observed on the day prior to ex-dividend date $(t=-1)$. As before, the same result is observed independently of the model used to estimate abnormal returns. Low dividend yield stocks do not show any significant abnormal return after ex-dividend days.

Table 2. Average abnormal returns for high yield stocks

	Model 1		Model 2		Model 3			
t	$\overline{AR_t}$	Z_t	$\overline{AR_t}$	Z_t	$\overline{AR_t}$	Z_t	Corrado	$\overline{CAR_t}$
-10	-0.02	-0.08	-0.14	-0.48	0.06	0.34	0.95	0.06
-9	-0.52	-1.72	-0.53	-1.79	-0.43	-2.25 *	-2.43 **	-0.36
-8	0.03	0.11	0.20	0.66	0.11	0.57	0.48	-0.25
-7	-0.14	-0.48	-0.18	-0.59	-0.05	-0.27	0.24	-0.30
-6	-0.57	-1.89	-0.59	-1.98 *	-0.48	-2.53 **	-1.16	-0.78
-5	-0.08	-0.26	0.00	0.01	0.07	0.36	0.21	-0.71
-4	0.22	0.74	0.17	0.58	0.22	1.17	1.18	-0.49
-3	-0.24	-0.80	-0.22	-0.74	0.03	0.14	0.30	-0.47
-2	0.49	1.64	0.48	1.60	0.36	1.90	1.36	-0.11
-1	0.43	1.44	0.40	1.34	0.39	2.07 *	1.56	0.28
0	**-0.57**	**-1.91**	**-0.67**	**-2.25 ***	**-0.24**	**-1.29**	**-0.62**	**0.04**
1	-0.11	-0.35	-0.19	-0.62	0.01	0.07	0.69	0.05
2	-0.15	-0.51	-0.02	-0.07	0.17	0.89	1.55	0.22
3	-0.51	-1.68	-0.43	-1.45	-0.21	-1.12	-0.37	0.01
4	-0.79	-2.62 **	-0.80	-2.69 **	0.10	0.54	-0.14	0.11
5	-0.31	-1.02	-0.30	-1.00	-0.12	-0.63	0.34	0.00
6	-0.68	-2.27 *	-0.63	-2.13 *	-0.10	-0.54	-1.19	-0.11
7	0.28	0.92	0.42	1.42	-0.13	-0.70	-0.11	-0.24
8	-0.20	-0.68	-0.19	-0.62	0.12	0.64	-0.17	-0.12
9	-0.39	-1.29	-0.45	-1.51	-0.05	-0.28	-0.76	-0.17
10	0.10	0.33	0.07	0.22	-0.09	-0.47	-0.87	-0.26
11	0.24	0.81	0.22	0.73	0.12	0.61	0.73	-0.15
12	0.22	0.73	0.30	1.02	0.05	0.24	-0.01	-0.10
13	-0.14	-0.46	-0.15	-0.51	-0.29	-1.52	-1.46	-0.39
14	-0.37	-1.23	-0.22	-0.72	-0.15	-0.78	-0.05	-0.54
15	-0.71	-2.37 *	-0.77	-2.58 **	-0.24	-1.28	-0.23	-0.78
16	0.12	0.40	0.12	0.40	-0.19	-1.01	-1.42	-0.97
17	0.22	0.72	0.07	0.23	0.14	0.72	0.67	-0.83
18	-0.50	-1.65	-0.37	-1.25	-0.27	-1.42	-1.14	-1.10
19	0.30	0.98	0.24	0.80	0.45	2.36 *	2.17 **	-0.66
20	0.09	0.31	-0.01	-0.04	-0.14	-0.73	-1.03	-0.79

*Significant at a 0.05 level. **Significant at a 0.01 level.

Graph 1. Average abnormal returns for high and low yield stocks.



As it can be seen in table 2, high dividend yield stocks show negative average abnormal returns on the ex-dividend day, statistically significant at a 0.05 level respectively for models 2, while non-significant for models 1 and 3. However, according with the market model we can not reject the null hypothesis that mean abnormal return on ex-dividend days is zero. As for the low yield sample, we observe positive and significant abnormal returns on day $t-1$.

As discussed in section 2, given our relatively low number of events, especially for the low yield sample, we have performed Corrado's non-parametric rank test, for detecting significant abnormal returns. The test has been performed for the abnormal returns obtained with the market model. The result of the test supports non-significant abnormal returns on ex-dividend dates, both for low and high yield stocks, and the observed positive and significant abnormal return for the day $t-1$ for the low yield sample.

Jointly considered, tables 1 and 2 show negative although non-significant average returns on ex-dividend dates for the two samples, while positive and statistically significant returns the day before ex-dividend dates. These results do not show a marked different return behavior according to the stock yield. Graph 1 shows average abnormal returns, as reported from model 3, for low and high yield stocks. As it can be seen, high yield stocks show higher volatility in abnormal returns around ex-dividend dates compared with low yield stocks. This fact can be explained by the higher number of events for high compared to low yield stocks.

We do not observe a marked behavior for cumulative abnormal returns, either for low or high yield stocks. Nevertheless, we had made no a priori hypothesis about the behavior this variable.

We have found only two comparable papers for the Spanish stock market: Ruiz and Espitia (1996) and Espitia and Ruiz (1997). Both of them investigate the same set of companies during the same time period: 1980-92, but with different methodologies. The authors observe that the ex-dividend price fall, in average terms, was significantly lower than the dividend amount. Thus, unlike our findings, average returns on ex-dividend days were positive. In our opinion, the most suitable explanation of these contradictory findings could rely on the different time period investigated. Capital gains have traditionally benefited of a better tax treatment by the Spanish tax system compared with dividends. As an example, until 1996 capital gains generated in more than five years, were tax-exempted, while for those generated in shorter periods there existed tax deductions. Therefore, the period covered by Ruiz and Espitia (1996) and Espitia and Ruiz (1997) was characterized by a tax advantage of capital gains compared to dividends. As we discussed in section 4, the situation is quite different nowadays.

As exposed at the beginning of the paper, Elton and Gruber (1970), and many authors since them, have viewed the ex-day stock return as an indicator of how market participants value lower-taxed capital gains relatively to higher-taxed dividends. Since one dollar dividend was less worthy than one dollar of capital gain, the ex-dividend price drop should be lower than the dividend paid, and thus ex-dividend days should show positive returns. However, as Naranjo et

al. (2000) state, this interpretation implicitly assumes that there is one investor tax clientele that determines the ex-day return. However, the ex-day return may more properly be viewed as being determined by the interaction of investors with different tax-induced valuations on dividends and capital gains. Since market participants show different tax profiles, the authors propose a tax heterogeneity explanation for ex-dividend day returns. This explanation is especially interesting for high dividend yield stocks, that should be particularly affected by tax-based trading. Naranjo et al (2000) report negative returns on ex-dividend days for high dividend yield stocks since 1975, after the introduction of negotiated commission rates. They explain this result in terms of the dividend capturing behavior of corporations more heavily taxed on dividends than on capital gains. Under low transaction costs these companies would buy high yield stocks before ex-dividend days and sell them afterwards. Our results, that abnormal returns for high yield stocks are negative and statistically significant at a 0.05 level for model 2 and almost at a 0.05 level for model 1, but non-significant for low yield stocks support this interpretation, particularly when trading volumes are also considered.

Table3. Average abnormal volumes for low yield stocks

	Model 1			Model 2		
t	$\overline{AV_t}$	Z_t	Corrado	$\overline{AV_t}$	Z_t	Corrado
-10	2.05	6.44 **	1.49	2.06	6.65 **	1.11
-9	1.14	0.86	1.66	1.13	0.83	1.43
-8	1.04	0.27	0.12	1.12	0.78	0.44
-7	0.87	-0.80	-0.60	0.89	-0.71	-0.39
-6	1.73	4.47 **	-1.66	1.64	3.99 **	-1.70
-5	1.17	1.02	0.82	1.15	0.95	0.51
-4	0.96	-0.23	-0.12	0.92	-0.47	-0.23
-3	0.88	-0.76	-0.67	0.93	-0.44	-0.12
-2	1.05	0.29	0.80	1.06	0.37	0.71
-1	0.87	-0.82	-0.63	0.93	-0.46	-0.47
0	**1.07**	**0.44**	**-0.49**	**1.03**	**0.16**	**-0.83**
1	1.19	1.15	0.84	1.14	0.89	0.52
2	1.07	0.45	0.15	1.16	1.00	0.78
3	0.85	-0.91	-1.62	0.84	-0.98	-1.37
4	1.06	0.38	-0.15	1.09	0.55	-0.34
5	0.82	-1.12	-0.73	0.80	-1.28	-1.19
6	0.99	-0.03	0.39	0.96	-0.23	0.20
7	1.26	1.58	0.37	1.35	2.23 *	0.84
8	0.83	-1.06	-0.87	0.83	-1.09	-0.82
9	0.85	-0.92	-0.60	0.88	-0.77	-0.56
10	1.18	1.13	0.45	1.15	0.92	0.32
11	1.08	0.47	0.98	1.05	0.30	0.82
12	1.26	1.58	1.03	1.36	2.25 *	1.39
13	0.99	-0.04	-0.30	0.98	-0.10	-0.33
14	1.05	0.32	0.22	1.06	0.38	0.42
15	1.18	1.12	1.18	1.15	0.96	0.90
16	1.07	0.44	0.34	1.03	0.19	0.15
17	1.40	2.47 *	1.36	1.49	3.11 **	1.76
18	1.14	0.85	0.92	1.14	0.89	0.87
19	0.94	-0.39	0.15	0.95	-0.30	-0.05
20	1.15	0.94	1.74	1.13	0.81	1.51

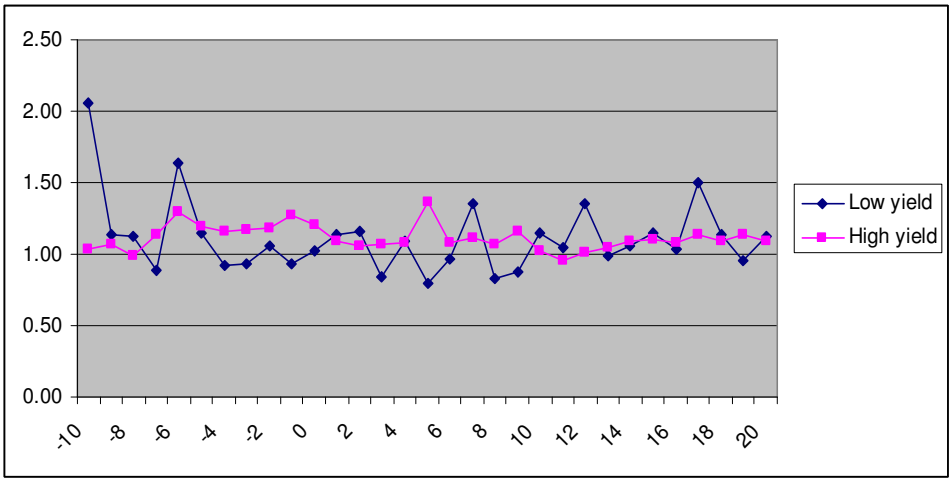
*Significant at a 0.05 level.**Significant at a 0.01 level.

Table 4. Average abnormal volumes for high yield stocks

			Model 1			Model 2
t	$\overline{AV_t}$	Z_t	Corrado	$\overline{AV_t}$	Z_t	Corrado
-10	1.00	-0.03	0.08	1.03	0.42	0.29
-9	1.12	1.25	1.18	1.07	0.85	0.41
-8	0.96	-0.40	-0.26	0.99	-0.07	0.20
-7	1.09	0.94	1.50	1.14	1.76	2.06 *
-6	1.40	4.28 **	0.71	1.30	3.68 **	0.49
-5	1.14	1.47	1.54	1.19	2.39 *	1.77
-4	1.17	1.80	1.87	1.16	2.00 *	1.31
-3	1.14	1.53	1.74	1.16	2.04 *	2.15
-2	1.17	1.78	1.45	1.18	2.27 *	1.74
-1	1.31	3.34 **	2.60 **	1.27	3.39 **	2.31 *
0	**1.18**	**1.89**	**1.48**	**1.20**	**2.50 ***	**1.40**
1	1.08	0.91	1.09	1.09	1.10	1.16
2	1.02	0.20	-0.36	1.06	0.71	-0.12
3	1.07	0.79	1.09	1.07	0.87	0.72
4	1.10	1.06	0.06	1.08	1.03	0.03
5	1.34	3.67 **	1.20	1.36	4.43 **	1.03
6	1.07	0.73	-0.23	1.08	0.95	-0.09
7	1.07	0.79	0.77	1.11	1.41	0.85
8	1.06	0.66	0.33	1.06	0.80	0.55
9	1.19	2.08 *	0.40	1.16	1.94	0.23
10	1.01	0.10	-0.39	1.03	0.35	-0.26
11	0.97	-0.31	-0.21	0.95	-0.60	-0.85
12	0.97	-0.35	-0.26	1.02	0.19	0.29
13	1.04	0.46	0.16	1.05	0.62	0.09
14	1.10	1.09	1.48	1.10	1.19	1.76
15	1.07	0.74	0.95	1.10	1.23	1.27
16	1.10	1.03	0.70	1.08	1.01	0.29
17	1.11	1.17	0.53	1.14	1.72	0.91
18	1.09	0.92	-0.07	1.09	1.14	-0.21
19	1.18	1.93	1.00	1.14	1.72	0.70
20	1.06	0.68	-0.64	1.09	1.07	-0.77

*Significant at a 0.05 level. **Significant at a 0.01 level.

Graph 2. Average abnormal volumes for high and low yield stocks.



The simultaneous investigation of returns and trading volumes strongly improves the ability to explain the ex-dividend day anomaly. The behavior of trading volumes around ex-dividends

days, showed by tables 3 and 4, supports the tax-heterogeneity hypothesis. First of all, because as suggested by Naranjo et al. (2000), under this hypothesis we should expect higher trading volumes around ex-dividend dates. For high yield stocks, table 4 shows significant abnormal trading volumes on four days of the period with model 1, and eight days with model 2, in all cases, average abnormal volumes are above one, indicating abnormally high trading volumes. Nevertheless, for low yield stocks, significant abnormal volumes are observed only in three days with model 1 and five days with model 2. The fact that abnormal trading volumes around ex-dividend days are higher for high than for low yield stocks is consistent with high dividend yield stocks being the most suitable candidates for dividend capturing behavior. Graph 2 shows abnormal trading volumes obtained from model 2 for low and high yield stocks. As it can be seen, while for high yield stocks abnormal volumes are above 1 for all days but two in the analyzed period, being particularly high the days before ex-dividend dates, for high yield stocks, days with high trading volumes are usually followed by days with low volumes.

As posed by Lakonishok and Vermeulen (1998), while no abnormal trading volumes should be observed under the basic Elton and Gruber (1970) framework, similar positive and negative abnormal volumes should occur under Kalay (1984) model. The reason is that investors with high marginal tax rates would prefer to sell before ex-dividend days and the opposite will prevail for those investors with low tax rates. In the case that all the categories of traders will exactly match each other, no abnormal volume should be observed. Our results support Lakonishok and Vermeulen (1998) model predicting: 1) Trading volume will increase abnormally before and after the ex-dividend day and 2) Because potential trading profits are relatively higher for high yield stocks, the excess trading activity will be more pronounced for high yield stocks.

6. CONCLUSIONS

The so-called dividend puzzle constitutes one of the most important lines of research in finance. However, in spite of the huge number of papers that have investigated the issue during the last decades, the puzzle is far from being solved.

This paper investigates the labelled ex-dividend day anomaly, characterized by significantly abnormal returns on this date. We have also investigated the behavior of trading volumes around ex-dividends days to better understand the causes of the ex-dividend day anomaly. Empirical evidence internationally available almost unanimously agrees about the existence of positive abnormal returns on ex-dividend days, indicating that one euro of capital gains is more valuable that one euro of dividends. Contrary to most of the existing evidence, we report negative although non-significant abnormal returns on ex-dividend days and positive and significant abnormal returns for the day before dividend payments. These results have been observed for high and low dividend yield stocks. We also observe higher trading volumes around ex-dividend dates, especially for high yield stocks.

Our results are difficult to conciliate with the two most accepted explanations for the ex-dividend anomaly, the traditional tax-clientele hypothesis, as posed by Elton and Gruber (1970) and the short-selling hypothesis, both predicting stock prices should drop less than the amount of dividend paid. On the contrary, they suggest a dividend capturing behavior by certain investors that would prefer dividends over capital gains, probably due to tax reasons. Deductions to avoid double taxation of dividends, especially important for corporate investors, might play a key role in encouraging this dividend capturing behavior. Nevertheless, more research is needed, for example, by examining the order size around ex-dividend days, to determine the validity of this hypothesis as an explanation of the ex-dividend day anomaly.

REFERENCES

ALLEN, F. AND R. MICHAELY. 1995. Dividend policy. In Jarrow, R.A., V. Maksimovic and W.T. Ziemba (Eds). *Finance*. Elsevier.

ATHANASSAKOS, G. 1996. Tax-induced trading volume around ex-dividend days under different tax regimes: the Canadian experience 1970–1984. *Journal of Business Finance and Accounting* 23: 557–584.

BALI, R. AND G.L. HITE. 1998. Ex-dividend day stock price behavior: Discreteness or tax-induced clienteles?. *Journal of Financial Economics* 47: 127-159.

BHATTACHARYYA, N. 2007. Dividend policy: a review. *Managerial Finance* 33(1): 4-13.

BLACK, F. 1976. The dividend puzzle. *Journal of Portfolio Management* 2: 5-8.

BREALEY, R.A., S.C. MYERS AND F. ALLEN. 2006, *Fundamentals of financial management*. McGraw-Hill.

BROWN, S. AND J. WARNER. 1980. Measuring security price performance. *Journal of Financial Economics* 8: 205-258.

CABLE, J. AND K. HOLLAND. 1999. Regression vs. non-regression models of normal returns: Implications for event studies. *Economics Letters* 64:81-85.

CAMPBELL, J. AND W. BERANEK. 1955. Stock price behaviour on ex-dividend dates. *Journal of Finance* 10: 425-429.

CORRADO, C. J. 1989. A nonparametric test for abnormal security-price performance in event studies. *Journal of Financial Economics* 23: 385–95.

ELTON, E.J. AND M. GRUBER. 1970. Marginal shareholder tax rates and the clientele effect. *Review of Economics and Statistics* 52: 68-74.

ELTON, E. J., M.J. GRUBER AND C.R. BLAKE. 2002. Marginal stockholder tax effects and ex-dividend day behavior thirty-two years later. *Working Paper*.

ESPITIA, M. AND F.J. RUIZ. 1997. Ex-dividend day stock price falls on the Spanish stock market. *Applied Financial Economics* 7: 481-492.

FRANK, M. AND R. JAGANNATHAN. 1998. Why do stock prices drop by less than the value of the dividend? Evidence from a country without taxes. *Journal of Financial Economics* 47: 161-188.

GARCÍA, J. 2007. Return autocorrelation anomalies in two European stock markets. Revista de Análisis Económico 22 (1): 59-70.

GRAHAM, J.R., R. MICHAELY AND M. ROBERTS. 2003. Do price discreteness and transactions costs affect stock returns? Comparing ex-dividend pricing before and after decimalization. *Journal of Finance* 58: 2613-2637.

HAYASHI, F., AND R. JAGANNATHAN. 1990. Ex-day behavior of Japanese stock prices: New Insights from new methodology. *Journal of the Japanese and International Economies* 4: 401-427.

JAKOB, K., AND T. MA. 2004. Tick size, NYSE rule 118, and ex-dividend day stock price behaviour. *Journal of Financial Economics* 72: 605-625.

KALAY, A. 1982. The ex-dividend day behavior of stock prices: A re-examination of the clientele effect. *Journal of Finance* 37: 1059-1070.

KATO, K. AND U. LOEWENSTEIN. 1995. The ex-dividend day behavior of stock prices: The case of Japan. *Review of Financial Studies* 8: 8I7-847.

KOSKI, J. AND J. SCRUGGS. 1998. Who trades around ex-dividend day? Evidence from NYSE audit file. *Financial Management* 27: 58-72.

KOTHARI, S.P., AND J.B. WARNER. 2005. Econometrics of event studies. In Eckbo, B. (Ed.) *Handbook of Corporate Finance: Empirical Corporate Finance*. Elsevier/North-Holland.

LAKONISHOK, J. AND T. VERMAELEN. 1986. Tax induced trading around ex-dividend days. *Journal of Financial Economics* 16: 287-320.

LINTNER, J. 1956. Distribution of incomes of corporations among dividends, retained earnings and taxes. *American Economic Review* 46: 97-113.

MILLER, M.H. AND F. MODIGLIANI. 1961. Dividend policy, growth and the valuation 20 of shares. *Journal of Business* 34: 411-33.

NARANJO, A., M. NIMALENDRAN AND M. RYNGAERT. 2000. Time variation of ex-dividend day stock returns and corporate dividend capture: A re-examination. *Journal of Finance* 50: 2357-2372.

POTERBA, J.M., AND L. SUMMERS. 1984. New evidence that taxes affect the valuation of dividends. *Journal of Finance* 39: 1397-1415.

POTERBA, J.M. AND L. SUMMERS. 1985. The Economic Effects of Dividend Taxation. In Altman, E. and M. Subramanyam (Eds.). *Recent Advances in Corporate Finance.* Richard Irwin.

ROBIN, A. 1991. The Impact of The 1986 Tax Reform Act on Ex-Dividend Day Returns. *Financial Management* 20: 60-70.

RUIZ, F.J., AND M. ESPITIA. 1996. La formación de precios de las acciones alrededor del pago de dividendos en el mercado de capitales español. *Revista Española de Financiación y Contabilidad* 25 (86): 179-198.

YAHYAEE, K., T. PHAM, AND T. WALTER. 2007. Ex-dividend behaviour in the absence of taxes and price discreteness. Paper presented at the 2007 EFMA conference, Vienna, Austria.

THE WORLD MACROECONOMIC DILEMMA 2008-2012: WORLD MONETARY STABILIZATION PLAN.

GONZALO PEREZ-SEOANE
e-mail: Gps@macroeconomicmodel.com
World-Markets Research and Consulting
Global Hedge Fund

Abstract

What global macroeconomic consequences will emerge from the economic and monetary policies adopted by the governments in the face of the current crisis? The present document contributes some novel theoretical and empirical ideas different from the general Keynesian and Neoclassical spirit. The considerations of this paper stem from the Natural Macroeconomic Model (NMM); this is a positive general equilibrium model of global scope.

Codigo JEL: E, F, G
Thematic Area: Monetary and Financial Economics.

1. 2008 WORLD ECONOMIC CRISIS. WATER EVEN IN THE FOUNDATIONS of THE SYSTEM.

The world economic system has commenced a complex process of destruction of wealth. The present crisis has overwhelmed the most prestigious Economic Institutes and Central Banks. However this should not be seen as a criticism since a process of destruction of wealth once initiated is, in truth, simply unpredictable. There is no model that can gauge the feeling of uncertainty and widespread panic. The Central Banks have acted so far in a sensitive and diligent manner with regard to the situation. But given the high degree of globalization of the world economy, the present crisis could well become the most voracious of all. The future is bleak and the social cost is already beginning to be high. Naturally, a great deal of uncertainty is hanging over the Financial System. Unfortunately, this reasonable doubt about the fragility of the Economic and Financial System is surpassing, to a greater or lesser extent, the effectiveness of the measures being taken, the foreseeable stages and the expected time frame of the crisis itself. Uncertainty and lack of confidence do not make the best foundation on which to build the economy.

But this world economic crisis has not only affected the real economy but has also dealt a blow to the doctrinal spirit which inspires it. We witness almost daily "*revolutionary*" aspects relating to the behavior of the world economy. The widespread adjustment of the financial markets and macroeconomic variables shows, in a very pragmatic way, the reality, the vast and intense world economic interrelationship. All the indications are that *the world economy is interdependent and interrelated*. Moreover, even a child can see that in the financial markets, the values of assets adjust over time in relative terms from one side of the planet to the other (with the exception of the price of money). This pragmatic undeniable evidence leads us, surprisingly, to a very simple conclusion: *the world is round and the economy as well*. The prices of the economy as a whole, despite being geographically very far apart at times, communicate with each other; they maintain a continuous dialogue which is no more than the real monetary substratum of what we call "economic globalization".

Nevertheless, all this evidence so perceptible at present in the fields of macroeconomics and finance would seem to be interpreted as being contrary to the basic tenets of Neoclassical economic thought which is the intellectual support of the current system. Thus, it is clear that a notorious silence extends over the intellectual background of the crisis. However, in spite of the silence, the waters of the current storm have reached the very foundations of the System. It may be stated that the current crisis has not only overwhelmed traditional Western economic policy but has also represented a severe setback for the economic thought underlying it. In short, a great new challenge for the future in which, once again, we will undoubtedly encounter the controversial work of Friedman: "*The Methodology of Positive Economics (1953)"*. Here, as it is well known, he suggests that "*economic theories should be evaluated according to empirical criteria*". Or, in other words, one day the empirical evidence will determine and substantiate true economic theory.

2. THE ECONOMIC YEAR 2009

The world economic outlook for 2009 is not optimistic. The information in the World Bank report "Global Economic Prospects" indicates that the crisis will get much worse in 2009 with an important number of high-income countries experiencing recession. The volume of world trade will fall drastically and inflation will remain low in the more-developed countries. The scenario described by the World Bank presents a situation of (near) deflation in the more-developed economies.

With respect to the developing countries –Asia, Latin America, Eastern Europe-, the analysis of 2009 data shows a general fall in GDP of these economies but in a situation of higher inflation. The IMF inflation forecasts indicate that the developing economies will continue with high inflation throughout 2009. Thus, the developing and peripheral economies could be entering into a situation of stagflation.

World economic growth will fall to 0.5% in 2009 according to the IMF.

Given the scenario described coupled with a risk of possible deflation, the Governments of the more-developed economies are conducting a dual and pragmatic -"Keynes-Fisher"- approach to the problem:

1.- *approval of substantial public investment packages (and tax incentives)*
In an effort to check the economic decline. These measures will have an important effect on the fiscal deficit and an immediate increase in the government debt of each nation.

2.- *future increase in the monetary mass*
The quantitative loosening of monetary policy links in with the reflections of Friedman and Schwartz in "*A Monetary History of the United States"* (1963) on the root causes of the Great Depression, reflections which are shared by the U. S. Federal Reserve.
One of the conclusions of Friedman and Schwarz was that in the short term, variations (decreasing) in the rate of growth of money can have devastating consequences either on prices or on real production. As a result, as is well known in the 1929 Depression, the recession became a disaster.

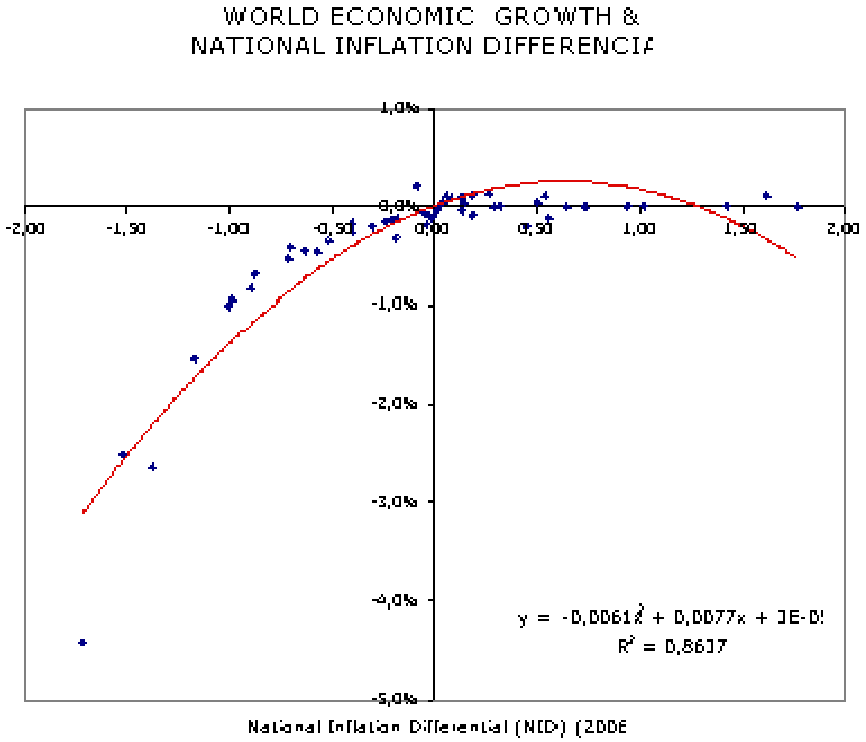
While both policies are the most suitable for the current situation, a lesson which was learnt at a high cost in the Great Depression and later years, nevertheless there would seem to exist at least one transcendent aspect which has not been taken into account from the macroeconomic point of view: *the degree of present globalization is not the same as 1929.* Thus, in a round world, *what world consequences will result from the above policies?.*

3. THE INFLATION DIFFERENTIAL AND THE ASYMMETRICAL IMPACT ON WORLD GROWTH

It is well known that the above policies generate inflation. But, is there a relationship between economic growth and inflation?. Will the increase in the level of prices in the developed countries have an effect on the inflation and state of economic growth in the developing nations?.

The empirical evidence provided by the Natural Macroeconomic Model (NMM) indicates that *the economic and monetary policies mentioned above could indeed give rise to serious economic alterations at the world level*.

The effects and the relationship between inflation (monetary policy) and economic growth have been intensely studied by many economists. The NMM (see dispersion graph below) shows the *existence of a universal relationship between the inflation differential and the economic growth of each nation.*



1	EMU (19 Countries)	2	Argentina	21	Hong Kong	40	Romania
	Austria	3	Australia	22	Hungary	41	Russia
	Belgium-Lux	4	Bahrain	23	Iceland	42	Singapore
	Chipre	5	Bolivia	24	India	43	QATAR
	Finland	6	Brazil	25	Indonesia	44	OMAN
	France	7	Bulgary	26	Israel	45	South Africa
	Germany	8	Canada	27	Japan	46	Sweden
	Greece	9	Chile	28	Korea	47	Switzerland
	Irland	10	China Main	29	Latvia	48	Thailand
	Italy	11	Kuwait	30	Lithuania	49	U.Kingdom
	Malta	12	Colombia	31	Malaysia	50	Uruguay
	Netherlands	13	Costa Rica	32	Saudi Arabia	51	United States
	Portugal	14	Croatia	33	Mexico	52	Ucrania
	Slovakia	15	UA Emirates	34	New Zealand	53	Venezuela
	Slovenia	16	Czech Rep	35	Norway	54	Turkey
	Spain	17	Denmark	36	Paraguay	55	Belarus
	Vatican	18	Ecuador	37	Peru	56	Iran
	Monaco	19	Egypt	38	Philippines		
	Andorra	20	Estonia	39	Poland		
	San Marino						

Considerations:

1.- Friedman Rule, the Tobin Effect, Taylor Principle, and the Phillips Curve

To the extent that a nation has a low inflation differential relative (positive) to the other nations with which it trades and competes, the GDP will grow more than its competitors. On the other hand, if a nation has a high inflation differential in relative terms (negative), the GDP will grow less than its competitors. *Thus, the main cause of the short term relative competitiveness of a nation, in terms of economic growth, is the inflation differential.*

The Friedman Rule, the Tobin Effect, the Taylor Principle are well known although there exist many other related studies, Phelps, Lucas, Mundell … The final question is the same in all of them: is there an ideal monetary policy in terms of interest rates, inflation and economic growth?.

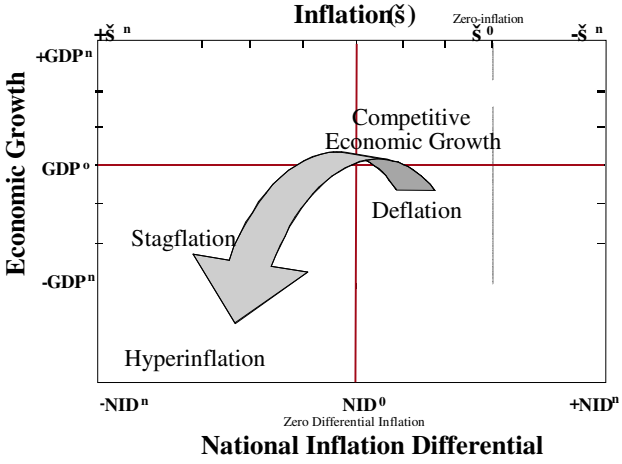
According to the NMM, the most favorable monetary policy of a nation is that in which the inflation differential is "zero". But, the best inflation differential determines, also, an *"optimal real interest rate"* –equilibrium- (Taylor Principle) for each country; and an"*optimal nominal interest rate*" according to the Fisher effect.

The NMM, just like Friedman, views the monetary mass as being the key variable to give direction to monetary policy. The growth rate of the monetary mass is one of the appropriate tools to use to attain the optimal rate of the inflation differential. Here, contrary to the opinion of Friedman, the growth rate of the monetary mass would *not be constant but variable*. In other words, the inflation differential of a nation is not determined by domestic monetary policy. The inflation differential is the result of the monetary policy of a country in relative terms to the actions of all the other monetary policies. Thus the dynamic nature of the complete monetary process means that the rate of creation of money, while it is a domestic action, must necessarily be dynamic if the objective is to reach a "*monetary optimum*".

The inflation differential, in view of the above, is *a new strategic monetary variable* since it helps to adjust the domestic monetary policy in its quest for greater relative competitiveness in terms of growth. From a different point of view, it could be said that the inflation differential highlights the changes in the relative world purchasing power within the domestic monetary policy.

Is it absolutely necessary to always look for the monetary optimal?. Possibly in the short term different monetary strategies may be followed, it difficult to say … but in the end the important thing is to ensure a stable economic growth since this affects the level and growth of employment (Okun's law). *Stable economic growth occurs from being close to the monetary optimal.*

How would the Phillips curve fit in with the statement on the inflation differential?. *The Phillips curve is absolutely true, but only on occasions.* The situation may arise where a country has increasing inflation and that it is close to its "zero" inflation differential, so the unemployment rate will decrease. What is important is not inflation in itself but rather the inflation differential and its grade of convergence to its own "optimal value". Each country has a distinct optimal value.The graph below shows the path of economic growth of each nation for different levels of inflation differential: competitive economic growth, deflation, stagflation or hyperinflation.



2.- Reformulation of the Fisher equation of exchange

Taking into consideration the empirical demonstration just presented, the Fisher equation of exchange can be reformulated from a global point of view or *in terms of purchasing power parity*. The reformulation would mean the acceptance of something quite simple: *the level of production of each nation is affected by the relative changes between the domestic and foreign purchasing powers.* Friedman proposed that monetary policy should have control over money supply but what really matters is the supply of money expressed in terms of purchasing power (not textual) "*The Quantity Theory of Money, A Restatement*" (1956). Thus, the Fisher exchange equation would be as follows:

$$M^{NID} V = P Q^{NID}$$

NID = National Inflaci ón Diferencial

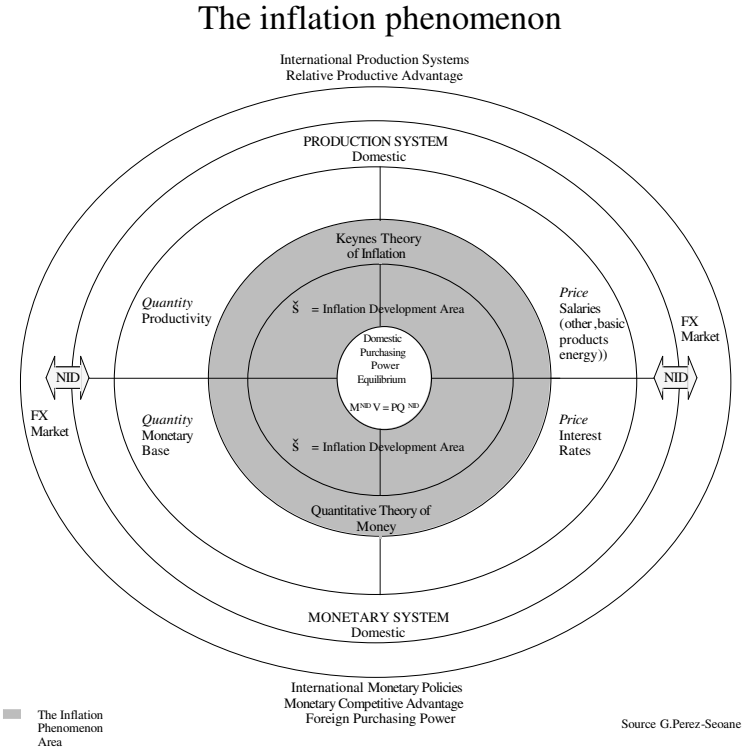
3.- The price gaps between different production systems are corrected in a natural way through variations in the exchange rates.

The (free) exchange rate market corrects the price deviations among all existing economies in a natural way.

Thus, the devaluation policies which aim to force a more rapid international competitiveness of economies via prices are often dangerous. Each time devaluation is used as an instrument of monetary and/or economic policy, the value of financial assets is affected in relative terms and the value of the domestic production system is altered in relative terms of other production systems. The consequences and posterior scenarios may be different since they ultimately depend on:

a) the degree of reaction caused on the various factors which act on the financial markets (distrust), -stock market, bond market, alteration of the flow of capital, slowing down or not of direct foreign investment and domestic investment -;
b) and the reaction of the system of production itself. In general, the joy in the exporting system of production (sole beneficiary of the devaluation) will soon come to an end, since the greater pressure of the external demand will tend to correct the relative imbalance in prices by increasing their value. Thus, over time, the devaluation can give rise to a further increase in the domestic price level which will feed back into the existing inflation, reducing the competitiveness temporarily gained and further slowing the economic growth towards a state of stagflation of hyperinflation. It is very attractive and simple to devaluate from a political point of view, but dangerous in macroeconomic terms and even more so in the current convulsed situation. Prices together with the law of supply and demand always relate, act and react in relative terms.

Inflation is the most complex macroeconomic variable. Inflation can be produced domestically; or it can be imported from external sources without consent and cost (loss of competitiveness). Regardless of its ultimate origin, (outside the scope of this paper), nevertheless, it is important to reiterate that the exchange market acts in a silent and permanent way restoring equilibrium between the financial system and the production system. This is reflected in the diagram below.



4. Uncertain "Nash equilibrium" in macroeconomics

In macroeconomics "*the Nash equilibrium"* is uncertain. Although (international) macroeconomics may be defined as a non-cooperative game (game of strategy), the empirical evidence discovered through the NMM suggests the non existence of the "Nash equilibrium". The world economic system permanently tends to general equilibrium, although the characteristic of this general equilibrium is that it always produces real winners and real losers.

Game theory is a wonderful tool for the analysis of human behavior and for situations of conflict among humans. But game theory is not only uncertain in the explanation of the macro-economy (and its global functioning) but is also uncertain in the explanation of just one price (free). The economy is prices, and to understand how prices come into being, relate with each other and adjust over time is to understand the economy. We are well aware that the formation and adjustment of prices are based on the law of supply and demand (Marshall), but, alongside this, prices relate with each other in relative terms. The ultimate reason for this is complex. In any case, prices (free) do not move nor adjust in a "*behavioral*" way but rather in a "*rational*" way. Thus, for example, *the prices of all financial assets at the world level,* along with the transitory variations in supply and demand, *move, relate and adjust in terms of rational expectations*. But these rational expectations, in the end amount to a single axiom; *maximum profit with minimal risk*: Thus, the whole *world Financial System functions, relates and adjusts in a permanent manner in terms of "real relative gains"* (to the delight of Von Mises and Hayek); the functioning of the system over time is empirically demonstrable. The well-known "*dilemma of the prisoner*" does not exist. *All financial assets are "prisoners of the same dilemma": higher returns with lower risk.*

5.- "Competitiveness and Productivity", key macroeconomic distinction

The above claim that the inflation differential is the actual mechanism that determines the level of competitiveness of a nation in the short term might seem surprising but it is really quite the opposite. To understand the simplicity and validity of the above, it is interesting to briefly consider the distinction between the terms "*competitiveness and productivity*".

The empirical evidence of the NMM shows that, in the medium to long term, the causes of growth of any nation are: a) *productivity, (since we are talking about a global setting: relative productivity), (Solow); b) the quantitative and qualitative growth in the human capabilities and skills (population growth – human capital), (Marshall, Kuznets, Schultz, Simon, Becker, Sen…).*

What really happens? Productivity should be one of the two most important causes which determines economic growth; however inflation has introduced perturbations in the logical economic growth of the nations giving rise to a situation where this monetary effect takes priority, to a certain extent, over the other causes of growth. Thus, the inflation differential directs all the variations in the GDP of the nations. If we accept that the nations with lower productivity are more vulnerable and the nations with higher productivity are less vulnerable; then to the extent that a nation is more productive and maintains low inflation differentials, its economic growth will accelerate at the direct expense of the interconnected nations which are worse positioned in both variables.

The phenomenon just described has been perceived and described in various ways for some time by many prestigious economists. Thus, for example, the following consideration by Samuelson[1] is of interest "when we compare economic results, we must distinguish a nation's competitiveness from its productivity. Competitiveness refers to the extent to which a nation's goods can compete in the marketplace; this depends primarily upon the relative prices of domestic and foreign products. *Competitiveness* is, however, quite distinct from a nation's *productivity*, which is measured by the output per unit of input, for example, the output per hour worked. Competitiveness in the U.S. fell sharply during the 1980s. But the cause was not a deterioration in productivity growth. Rather, deteriorating competitiveness during the 1980s arose because the appreciation of the dollar raised American prices relative to those of its trading partners. In fact, there was no major change in the overall trend in productivity growth during the 1980s, and if anything, productivity probably grew more rapidly during that period than in the prior decade. *Make sure you understand this fundamental point about competitiveness: nations are not inherently uncompetitive. Rather, they become uncompetitive when their prices move out of line with those of their trading partners because of an overvalued exchange rate*".

[1] *Samuelson Paul y Nordhaus William, "Economía",* p.754, *McGraw Hill, 1996*

Would this have anything to do with the Paradox of Leontief (1953)?. Could the inflation differential be an important cause of the increase or decrease in the exports of each country?.

4. MACROECONOMIC SITUATION 2008- 2012

With reference to the current situation, if inflation increases too much in the more-developed economies as a consequence of the monetary expansion and the excessive public deficit, it is possible that:

a.- *Financial Market*

The fall in production against a backdrop of rising inflation in the emerging economies (stagflation), will, no doubt, lead to devaluations. Such devaluations will have different effects, but given the current climate of instability and insecurity, perhaps, they will give rise to notorious exchange collapses and massive capital outflows. Where will this capital end up?. Will the Central Bank reserves be wiped out in countries like Iceland and others?. Will some countries default on their foreign public debt?. Will the IMF be able to cover the temporary insolvencies of a large number of countries?. Who will finance the IMF?. Many economies could enter into states of hyperinflation in view of the above.

Unfortunately the world of devaluations can become very complicated. The devaluations of some may encourage others. Will we witness an especially massive "jumble sale" of economies in their effort to be more competitive in relative terms as quickly as possible?. Latin America is already heading in that direction.

In this situation, we have the possibility of the emergence of protectionist processes, processes which are harmful to the freedom that should guide the action of the world economy: nationalizations, excessive protections of foreign trade.

In view of the above, it might be concluded that the future situation of the emerging

countries could become extremely complicated while at the same time the actions taken by the developed countries would worsen this situation. The world economic recovery would be compromised; the volatility of the exchange markets and the stock exchange would be even more affected as would the credibility of the actions taken and ultimately the expectations of recovery of the entire system.

b.- *World Trade*

When there are significant increases in the levels of inflation, the currencies in countries of the Western world tend to depreciate at the world level (US$, Euro, Pound, Yen…). As these economies are the main importers of the world (excepting China), the less advantageous relative position of the exchange rate will slow down the imports and ultimately world trade will not make the desired recovery, this will have a direct effect on the exports of emerging economies. Alongside this, it must be taken into account that in the current situation of a severe contraction of the world economy, exports will inevitably decrease in all countries, with particular impact in those countries where this variable (net exports) represents an important part of the GDP.

c.- *Commodities*

The loss of buying power of the more important currencies will lead to a *future* rise in the price of commodities (especially energy and gold). The relationship between the buying power of the dollar and oil prices is well known. Similarly, it is well known that an increase in the price of commodities (especially oil) directly affects inflation at the world level with the elasticity of the emerging economies much higher than that of the developed economies. The marginal gain from the rise in prices of the commodities (energy) will only benefit a small number of countries, the producers and exporters of oil and related industries. The remaining countries face a more delicate economic scenario with possible endogenous increases in inflation.

Could additional measures be taken as a precaution against a possible severe reprimand of the emerging economies?.
The answer may be affirmative.

5. WORLD MONETARY STABILIZATION PLAN

Perhaps it would be desirable to stabilize, temporarily, the entire world monetary system.

1.- Provisional system of "fixed parities" at the world level

An important part of the existing stress and imbalances occurs or inevitably affects the exchange rate system. The exchange rate system, while it is the best of the possible systems, at the present time, can only introduce more instability into the financial markets and the existing relationship among the different production systems. Therefore, it would be desirable to temporarily establish a system of fixed parities at the world level (all existing currencies) (with narrow bands or not … it is irrelevant). A period of 24-36 months would be sufficient time for calm to return to the markets and for the world economy to commence a new cycle of growth.

2.- World Monetary mass

Control of the world monetary mass during the stated period in all countries. This means the non-increase of the monetary mass and obviously in no case a decrease. The non-increase of the monetary mass would be a necessary evil so that the system of parities is fair and credible from an economic standpoint. In a global system of fixed parities, any increase in the monetary mass of one country increases the buying power of this country with respect to the remaining countries.

3.- Preferential interest rates in countries in deflation (developed)

The interest rates in developed countries should be nominally zero, negative in real terms.

4.- Preferential interest rates in countries in stagflation (emerging)

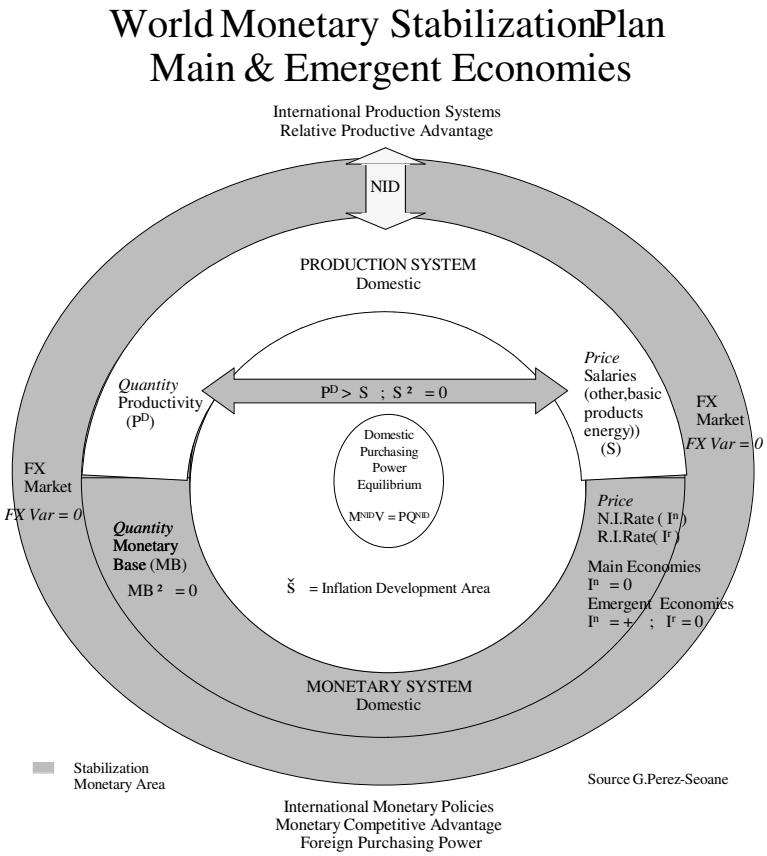
The interest rates in emerging economies and economies in development should be nominally positive, zero in real terms. The non-increase of the monetary mass will encourage the stabilization of price levels and indeed their decrease. The non-negative (zero) real interest rates in these economies will avoid massive capital outflows

5.- Productivity versus nominal wages

According to the NMM, the causes of the generation of inflation are varied. From the point of view of this model, some quantitative hypotheses would be correct as would some Keynesian hypotheses. In a hypothetical environment of high monetary stability, inflation, from the point of view of the NMM, would only have a small "*vent*" through which to surface, the inflation which arises from the mismatch between the increase in wages and the amount produced (productivity) (cost inflation – Keynes). Thus, it is sufficient that the variation in the nominal salaries in all the most developed economies should remain at zero. In the emerging countries, this would also be desirable even if there is a certain loss in purchasing power (there is a loss at the domestic level but there is a gain in the exterior). In this hypothetical scenario of almost complete monetary stability, the *competitiveness* of a nation and its *relative productivity,* even though they are distinct economic terms and effects, would tend to converge. The greater relative productivity will relaunch the economies of some countries more rapidly than others. Any deterioration in the relative productivity will reduce the potential of economic growth.

6.- Change in the existing expectations

An important part of the short-term evolution of each economy is linked to the expectations which are shared by the different economic and social agents. The current deterioration in the economic expectations at the world level is particularly striking. Moreover, there does not seem to be anything inherent in the present situation that could rescue the world economy from its state of prostration and apathy. In view of this, perhaps we need some new brilliant political action, an action which would change the tendency of the economic expectations thus avoiding future risks and giving credibility to the stability of the world economic and financial systems.



6. THE MACROECONOMIC DILEMMA 2008-2012

It is true that for the more developed economies, many of them beginning to undergo a process of deflation, to ask them to forego the increase in the monetary mass as one of the best medicines to promote a more rapid recovery, is a painful decision. This measure would lead to a slower recovery.

However, it must be taken into account that with a totally stabilized monetary system, the "*serious and unpredictable*" deterioration of the developing and peripheral economies could be avoided. The responsibility of these latter economies in the crisis is impossible to justify, moreover the deterioration of these economies would have general consequences. In terms of population, these economies represent more than two thirds of the planet.

This is the real current macroeconomic dilemma:

> *should the first economies of the world act in a sovereign way in the face of the world Economic System as a whole, using all available medicines to alleviate their domestic crisis; although, these same medicines could cause undesirable side effects in countless third countries?.*

But the real world is more complex. Is it possible, from a political view point, to achieve a system of fixed parities at the world level all at once?. Will all the countries be fully responsible for their obligations with respect to monetary policy, or the temporary non-action on the monetary base?. Will all this have influence at the geopolitical world level?.

7. A POSSIBLE HOPE FOR THE LONG-TERM: "EARLY WARNING SYSTEM FOR FINANCIAL CRISES"

During 2008 some politicians declared the necessity of constructing an "*early warning system for financial crises*" (Mr. G.Brown, U.K. Prime Minister). Is such a system possible? The truth is that nobody has every explicitly specified such an idea.

Prior to the consideration of the real possibility of an "*early warning system*", it is necessary to tackle a controversial question *what is the origin of the financial crises*?. The answer to this question could help to situate the primary objective which should cater for a possible "early warning system for financial crises".

To answer the question it is of interest to distinguish between "*formal cause*" and "*efficient cause*" of the financial crises.

Formal origin

The formal origin of all the financial crises is always the same: "financial bubble". Just like the flu virus, the financial crises, whenever they reappear, always portray some type of mutation whether it is in their active origin or whether it is due to other monetary, macroeconomic and/or environmental circumstances. In the current crisis, together with the U.S. real estate bubble and the widespread sub-prime assets, the transitory raw material and food bubble.

Efficient origin

Financial crises are always linked to a process of distrust. This coupled with a lack of credibility on behalf of the main actors in the market of the values of assets, or even worse, a lack of credibility in the functioning of the System itself, its institutions and instruments which is the case at present. Thus, financial crises are no more than downward adjustments, abrupt and widespread, in the monetary value which assets have accumulated over time. These adjustments impact, to a greater or lesser extent, on the real economy through well known side effects: paralysis of investment, credit restriction (illiquidity), fall in consumption, increase in unemployment …ultimately an enormous loss of *monetary* wealth.

Thus, independently of the formal origin of each crisis, *the efficient cause is, also, always the same, "the lack of some objective and explicitly shared reference on the accumulated value of assets over time (economic equilibrium value)"*. It is true that many mathematical and statistical techniques exist for the valuation of assets and moreover these techniques are universally shared by the economic and financial community. But, to date there is no link (empirical) between the accumulation of value of assets over time and the economic development of society. This link, were it to exist, could be called the "*economic equilibrium value*". Economic history shows, time after time, that after a period of intense accumulation of monetary value in assets (whatever the asset), the market vacillates about the veracity of the values assigned. Ultimately, this market uncertainty about the assigned accumulated value to the assets will end up converging in time with some formal cause (at present subprime crisis and credit crisis) thus instigating a process of destruction of monetary value which always advances relentlessly, spurred on by an irrational feeling of panic. The fall in value of some assets (market), forces other assets to adjust in terms of relative value. The greater the level of globalization the more the effects transcend from the domestic sphere to the world stage.

Despite what numerous economists say, there is no cyclicity of any type in financial crises, rather there is an unpunctual and disordered process of distrust which transmits its distress to the rest of the economy. Thus, *financial crises are no more than flu's self-generated periodically by the Financial System itself, these flu's, strange though it may seem, purge the not completely objective monetary valuation accumulated by the assets over time*. Once a process of destruction of value has commenced and moreover this process has been transmitted to the economic system, as Keynes would say, "there is nothing inherent in this situation that is capable of rescuing it". Investment disappears into thin air and aggregate demand plummets, leaving State aid as one of the few options to stimulate and correct the situation.

Is the existence of an early warning system possible?

Yes, it is possible.

At the world level, the financial markets function based on the law of markets (Marshall's Law). But, alongside this, the financial markets tend to permanently adjust in terms of "*real relative profitability"*. This curious financial and macroeconomic phenomenon means that, from a long term perspective, the financial markets function in a harmonic, rational, interdependent and empirically demonstrable manner at the world level. The disclosure and acceptance, theoretical and pragmatic, of this singular macroeconomic phenomenon of natural origin or at least with no conscious human cooperation, would make the existence of *economic equilibrium values* possible for distinct markets of key assets (world stock market, world bond market, world exchange rate market, short term public debt market, world property market, …). *The economic equilibrium values would allow us to link the monetary value attained by the assets with the very foundations of the world economic system, its own development over time. All this would help to avoid new processes of destruction of wealth thus cooperating in a greater future worldwide financial stability.* The invisible hand of Smith would be visible. To the satisfaction of all, *the "invisible hand" in the financial market arises simply from the free, healthy and rational arbitration between millions of agents who make up this market, always with the ultimate objective of obtaining the greatest profitability with the least possible risk*. In other words, the invisible hand emerges from the rationality of the monetary action of the human being, from the individual and/or collective common sense. But in order that the "*invisible hand*" continues its action of rebalancing the markets over time, *"price freedom"* should remain the most valued banner of the world economy.

8. CONCLUSION

The macroeconomic policies adopted by the world's largest economies to curb the current financial crisis are fully appropriate and necessary from the domestic macroeconomic perspective. The monetary authorities and Governments have shown a commendable coordination, a coordination without precedent in the history of economics.

But the world is round and the economy also. Thus, curiously, the policies adopted may generate very serious macroeconomic side-effects, especially in the emerging economies and other economies in development... Were these effects to occur, they could ultimately end up in reverse ("boomerang effect") and impact on the most developed economies.

The world economy has need of a novel and swift action. An action which avoids future risks (emerging economies). An action which ensures the full stability of the world financial and economic system.

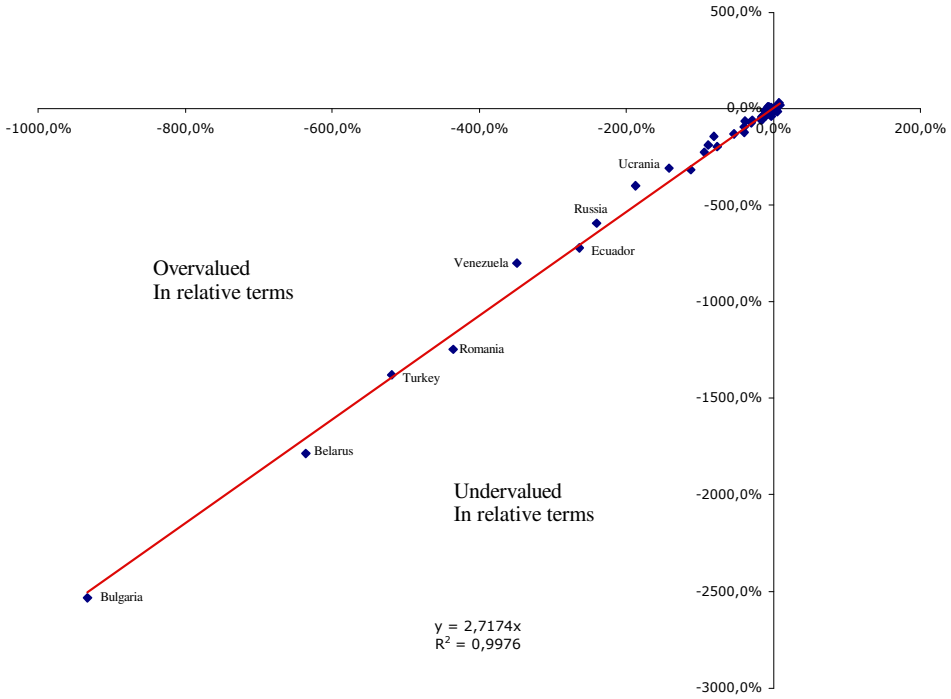
Perhaps, the world monetary system should be stabilized in its totality. Once stabilized, it will be easier to peacefully solve the bank-credit-illiquidity crisis and the world economy will foreseeably begin a tremendous recovery.

The year 2008 will take an outstanding place in the history of economics. A change has begun.

2009 will consolidate this change.

14 February, 2009

The World Exchange Rate Market Function
Empirical Demonstration December 2007

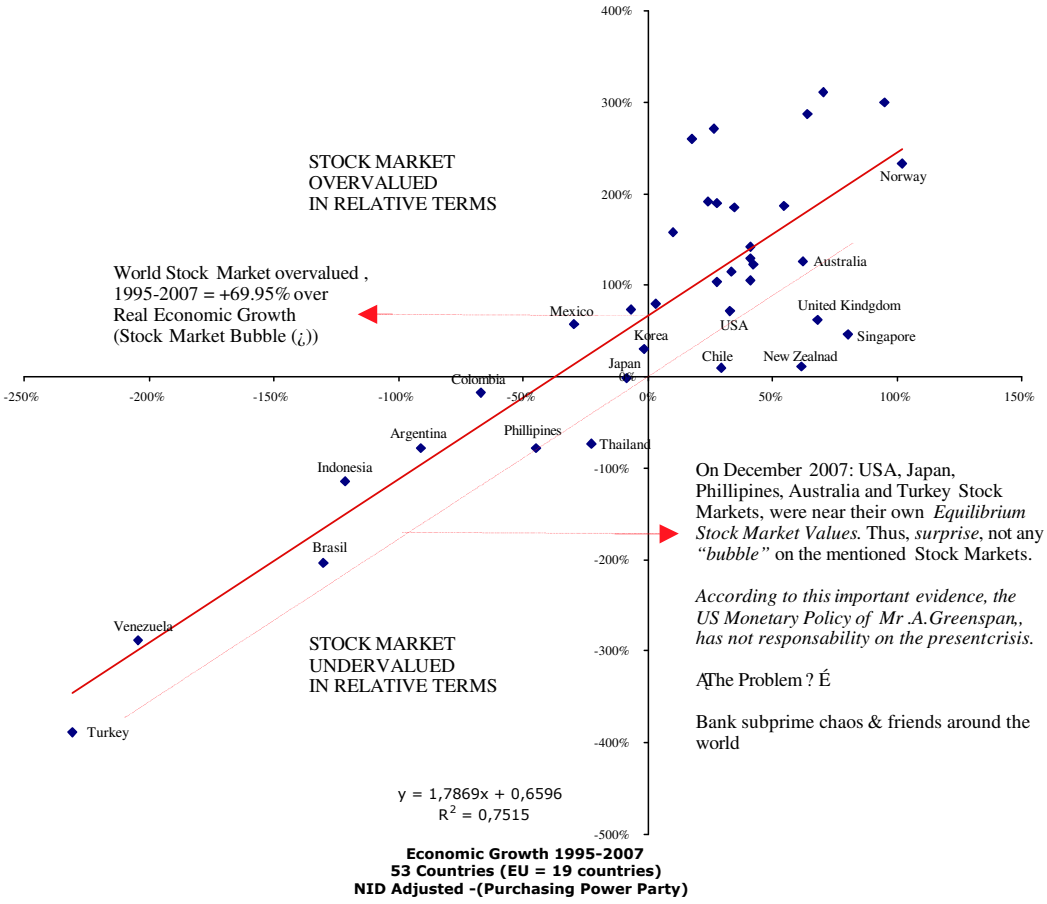


National Inflation Differential 1995-2007

CURRENCIES SDRs (IMF) / 56 Foreign Exchange - 75 Countries

1	EMU (19 Countries)	2	Argentina	21	Hong Kong	40	Romania
	Austria	3	Australia	22	Hungary	41	Russia
	Belgium-Lux	4	Bahrain	23	Iceland	42	Singapore
	Chipre	5	Bolivia	24	India	43	QATAR
	Finland	6	Brazil	25	Indonesia	44	OMAN
	France	7	Bulgary	26	Israel	45	South Africa
	Germany	8	Canada	27	Japan	46	Sweden
	Greece	9	Chile	28	Korea	47	Switzerland
	Irland	10	China.Main	29	Latvia	48	Thailand
	Italy	11	Kuwait	30	Lithuania	49	U.Kingdom
	Malta	12	Colombia	31	Malaysia	50	Uruguay
	Netherlands	13	Costa Rica	32	Saudi Arabia	51	United States
	Portugal	14	Croatia	33	Mexico	52	Ucrania
	Slovakia	15	UA Emirates	34	New Zealand	53	Venezuela
	Slovenia	16	Czech Rep	35	Norway	54	Turkey
	Spain	17	Denmark	36	Paraguay	55	Belarus
	Vatican	18	Ecuador	37	Peru	56	Iran
	Monaco	19	Egypt	38	Philippines		
	Andorra	20	Estonia	39	Poland		
	San Marino						

The World Stock Market Function
Empirical Demonstration December 2007



On December 2007: USA, Japan, Phillipines, Australia and Turkey Stock Markets, were near their own *Equilibrium Stock Market Values.* Thus, *surprise*, not any *"bubble"* on the mentioned Stock Markets.

According to this important evidence, the US Monetary Policy of Mr .A.Greenspan,, has not responsability on the presentcrisis.

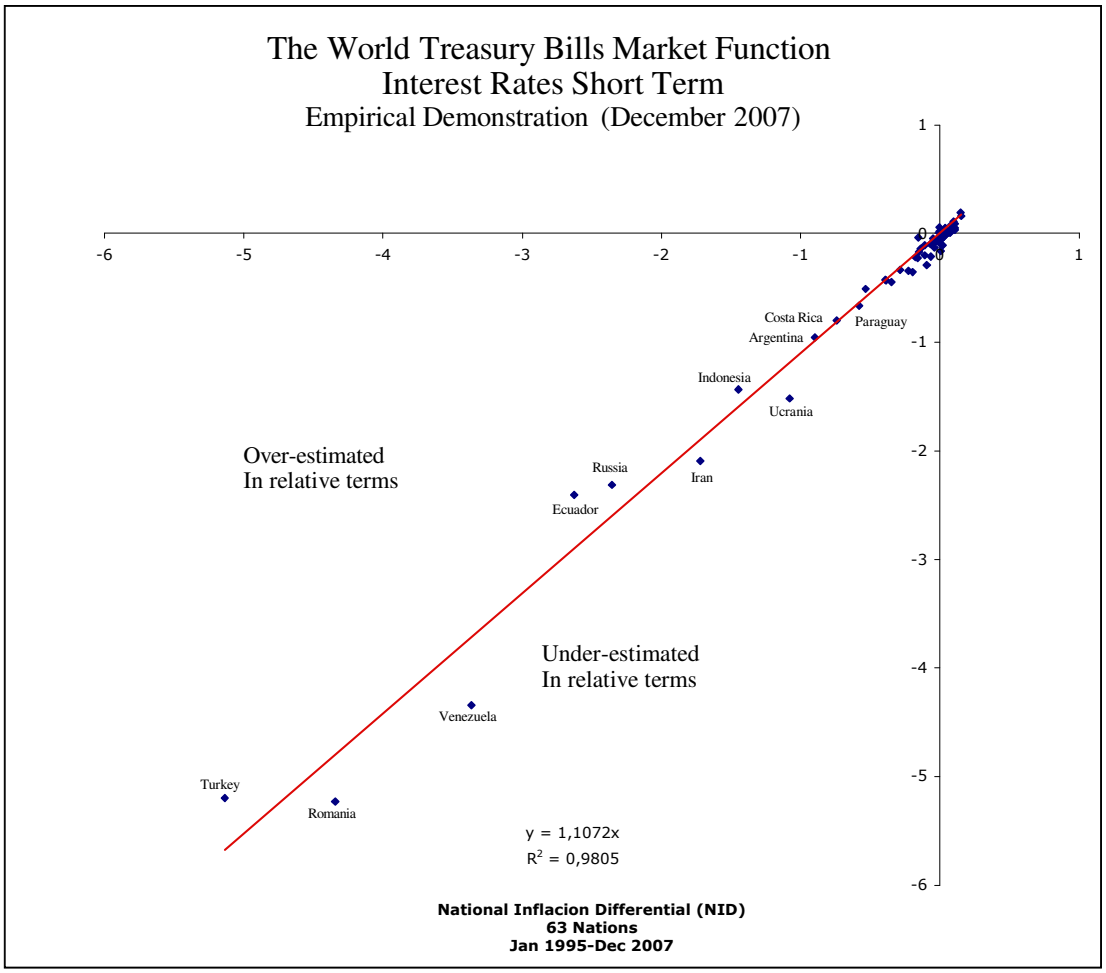
AThe Problem ? É

Bank subprime chaos & friends around the world

STOCK MARKET INDICES

1	Austria	13	Brazil	25	New Zealand
2	Belgium-lux	14	Canada	26	Norway
3	France	15	Chile	27	Peru
4	Germany	16	Colombia	28	Philippines
5	Greece	17	Denmark	29	Singapore
6	Italy	18	Hong Kong	30	Sweden
7	Netherlands	19	India	31	Switzerland
8	Portugal	20	Indonesia	32	Thailand
9	Spain	21	Israel	33	U.Kingdom
10	EMU	22	Japan	34	United States
11	Argentina	23	Korea	35	Venezuela
12	Australia	24	Mexico	36	Turkey

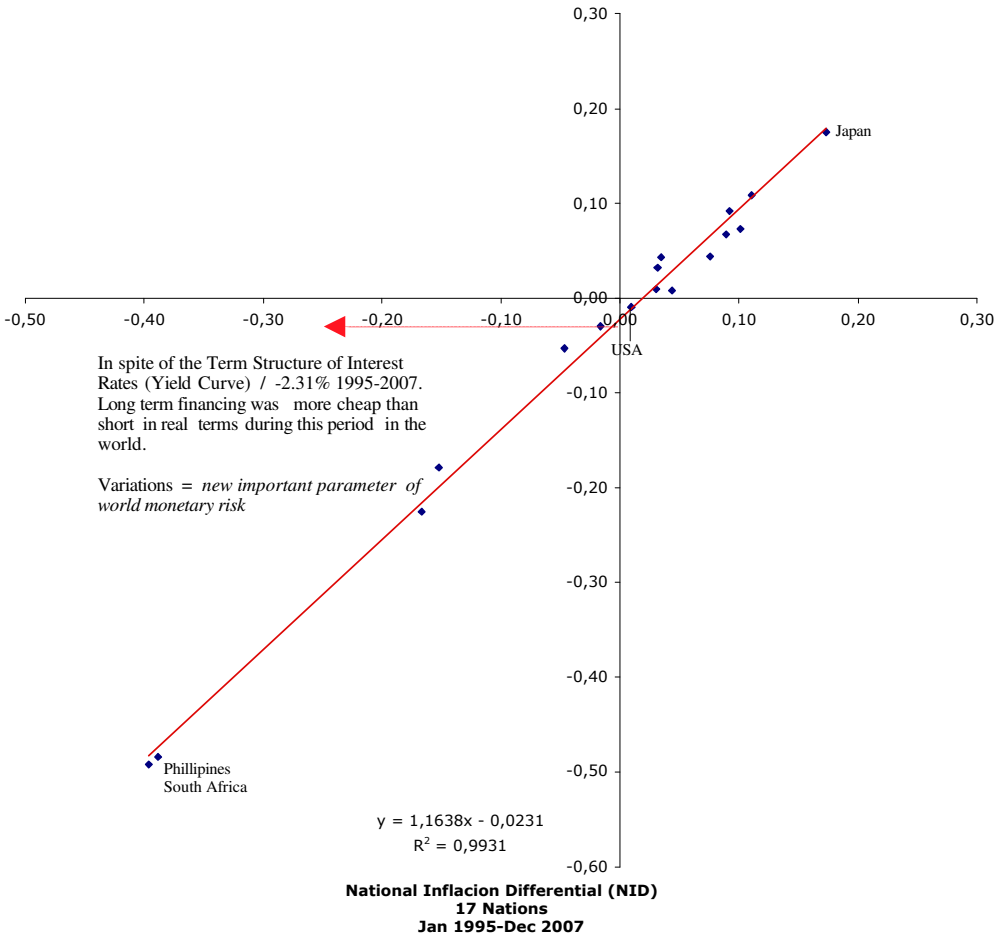
Source: IMF / Morgan Stanley



TREASURY BILLS COUNTRIES

1 Austria	22 Colombia	43 Mexico
2 Belgium-lux	23 Costa Rica	44 New Zealand
3 Finland	24 Croatia	45 Norway
4 France	25 United Arab Emirate	46 Paraguay
5 Germany	26 Czech Rep	47 Peru
6 Greece	27 Denmark	48 Philippines
7 Ireland	28 Ecuador	49 Poland
8 Italy	29 Egypt	50 Romania
9 Netherlands	30 Estonia	51 Russia
10 Portugal	31 Hong Kong	52 Singapore
11 Spain	32 Hungary	53 OMAN
12 EMU	33 Iceland	54 South Africa
13 Argentina	34 India	55 Sweden
14 Australia	35 Indonesia	56 Switzerland
15 Bahrain	36 Israel	57 Thailand
16 Bolivia	37 Japan	58 U.Kingdom
17 Brazil	38 Korea	59 United States
18 Canada	39 Latvia	60 Ucrania
19 Chile	40 Lithuania	61 Venezuela
20 China.Main	41 Malaysia	62 Turkey
21 Kuwait	42 Saudi Arabia	63 Iran

The World Government Bond Market Function
Long Term Interest Rates
Empirical Demonstration (December 2007)



BOND COUNTRIES

1	Austria	11	Philippines
2	Belgium-lux	12	South Africa
3	Germany	13	Sweden
4	Australia	14	Switzerland
5	Canada	15	Thailand
6	Denmark	16	U.Kingdom
7	Japan	17	United States
8	Korea		
9	New Zealand		
10	Norway		

BIBLIOGRAPHY

BARRO, ROBERT, "Macroeconómía", McGraw Hill, 1996.
BECKER, GARY, "The Essence of Becker", R.Febrero y P.S.Schawartz, 1995
BREALEY RICHARD - MYERS STEWARD, "Principles of Corporate Finance", Mc Graw Hill, 1984.
CAMPBELL R.MCCONNELL y STANLEY L.BRUE, "Economía", McGraw Hill, 1997.
CANTILLON, RICHARD, "Ensayo sobre la Naturaleza del Comercio en general", Fondo de Cultura Económica, 1950
DORNBUSCH RUDIGER - FISCHER STANLEY, "Macroeconomía", McGraw Hill, 1994.
Ekulund Robert - Hebert Robert, "Historia de la Teoría Económica y su Método", McGraw Hill, 1997.
FERNÁNDEZ DÍAZ ANDRÉS, PAREJO GAMIR JOSÉ ALBERTO, RODRIGUEZ SAIZ LUIS, "Curso de Política Económica", Editorial AC, 1993.
FISHER, IRVING, "La ilusión de la moneda estable", Ediciones Oriente Madrid, 1930.
FISHER, IRVING, "The Purchasing Power of Money", Augustus M.Kelley, New York, 1963
FISCHER STANLEY, DORNBUSCH RUDIGER, SCHMALENSEE RICHARD, "Economía", McGraw Hill, 1989.
FOREMAN-PECK JAMES, "Historia Económica Mundial", Prentice Hall International, 1995.
FRIEDMAN MILTON, "Paradojas del Dinero", Ediciones Grijalbo, 1992.
HEYMAN, TIMOTHY, "Inversión contra inflación", Editorial Milenio, 1987.
GRANT, ROBERT, "Dirección Estratégica", MacGrawhill, 1996
KONDRATIEFF, NICOLAI, "Los ciclos económicos largos", Akal, 1979.
KRUGMAN PAUL - OBSTFELD MAURICE, "Economía Internacional", McGraw Hill, 1997.
MARSHALL, ALFRED, "Principios de Economía", Editorial Aguilar, 1963
MARX, KARL, "El capital", Siglo Veintiuno Editores, 1987.
MUÑOZ CIDAD, CÁNDIDO, "Estructura Económica Internacional", Civitas 1996.
PORTER MICHAEL E., "The Competitive Advantage of Nations", Macmillan Press, 1990.
SAMUELSON PAUL y NORDHAUS WILLIAM, "Economía", McGraw Hill, 1996.
SEN, AMARTYA, "Bienestar, Justicia y Mercado", Ediciones Paidós, 1997
SIMON, JULIAN, "The Ultimate Resource" Priceton University Press, 1981.
SMITH ADAM, "La Riqueza de las Naciones", Alianza Editorial, 1996.

THE PIGS, DOES THE GROUP EXIST? AN EMPIRICAL MACROECONOMIC ANALYSIS BASED ON THE OKUN LAW[1]

JOÃO SOUSA ANDRADE
E-mail: jasa@fe.uc.pt
Faculdade de Economia
Universidade de Coimbra

Abstract

Will the current crisis accelerate the PIGS collapse? We approach the subject by comparing the responses of the unemployment rate to an output shock on those economies (Portugal, Italy, Greece and Spain) with those of a benchmark economy – the USA. Our methodological strategy relies on one of the pillars of empirical macroeconomics the Okun Law (OL) which we incorporated in a VAR model. We addressed two drawbacks usually present in OL, the interdependency problem and the non-stationarity problem. We have included in our models the participation rate as a way to overcome the former problem and for the later one we have analysed the time series properties of the variables used on our models. We propose stable VAR models for each of the economies involved and also a fixed-effects panel-VAR for the PIGS. The time for the absorption of shocks and the disequilibrium levels are much more favourable to USA, but we conclude also that in terms of unemployment we re not allowed to consider the PIGS as a homogenous group.

Key words: Okun Law, C-I, VAR, Participation rate, Stability and Impulses.
JEL codes: C32, C51, J21 and E24.
Thematic Area: Monetary and Financial Economics.

INTRODUCTION

The acronym PIGS is used to define a group of countries within the Eurozone, that share similar macroeconomic imbalances. They have large current account deficits and high unemployment as compared to the average of the European Monetary Union. These countries suffered a loss of competitiveness, registered high values of public deficits and in the last years the unit labour costs gap with Germany has not been reduced.

According to (Evans-Pritchard 2009), quoting a BNP Paribas report and senior executive Hans Redeker's declaration "the 'PIGS' quartet was now facing "collapse"", not only direct foreigner investment but also portfolio investment had abruptly declined in UEM as a result of the bad economic performance of PIGS.

The inspection of such a disastrous situation, as mentioned above, will be accomplished through the study of the results of an output shock in the PIGS area, measured by the levels of unemployment. We represent the current crisis by an output shock and compare the "bad" performances of those countries and the area with a benchmark economy, not an European one but the United Sates. As (Knotek 2007) has suggested in a pre-crisis context we can answer to our question with the application of the Okun Law (OL) with a special emphasis to the stability

[1] I thank Adelaide Duarte for the comments and suggestions. But only the author is responsible for the omissions or inaccuracies, which may still exist.

problem of the estimated relation. As the statistical supply of labour is not independent of the conjuncture we included in our study the interdependence created by the participation rate.

Besides the introduction and conclusion the paper is organized in two parts. In the first one, "The Okun Law: presentation and meaning", we present and comment the OL, which is central to our methodological strategy, and we discuss the interdependencies that crucial for OL but which are usually excluded from its formulation. In the second part, "An application to the PIGS area", we build a model for each one of the PIGS, for all the group, and for the USA economy.

I. THE OKUN LAW: PRESENTATION AND MEANING

1. The Okun Law

From equation (1.1) that expresses the relation between actual output (Y_t) and the unemployment rate gap (U_t-U_t*), (Okun 1962) as suggested to measure the potential output.

$$Y_t^P = Y_t\left[1+0,032\cdot\left(U_t - U_t*\right)\right] \qquad (1.1)$$

The value of U_t* in (1.1) has been taken by Okun as constant and equal to 4%[2]. This value can be associated to the NAIRU concept, the natural rate of unemployment, or simply to a "Benchmark Unemployment Rate" ((Lovell 2004), p.351). The original central idea was to obtain a measure of potential output[3]. But very early the OL was known by the relation 3/1, which represents the idea that a reduction of the unemployment rate of 1% requires a growth rate of 3%, and the idea was very appealing to economists and politicians[4]. Based on this interpretation the door was opened to be known, later on, as the OL: a stable relationship between the unemployment rate and the output growth rate. Based on the effects of employment on production, *mutatis mutandis*, it was proposed a link between a reduction on unemployment and an increase on output[5]. OL is no more than an empirical regularity and it was classified by (Blinder 1997) as one of the features of the nuclear model of the economy[6]. As was described on a well-known handbook, this robust empirical generalization[7] "is one of the most reliable generalizations that macroeconomists have found" ((Hall and Taylor 1988), p. 136).

[2] In Okun, A. (1962). "Potential GNP: its measurement and significance." American Statistical Association, Proceedings of the Business and Economic Statistics Section, pp.98-103, reprint in The Political Economy of Prosperity, W. Norton, New York, 1970, pp.132-45. words: "that four percent unemployment is a reasonable target under existing labor market conditions", p.98 cited in Attfield, C. and B. Silverstone (1998). "Okun's law, cointegration and gap variables." Journal of Macroeconomics **20**(3): 625-37.. When Okun has developed his relation the Council of Economic Advisers of Kennedy had proposed a target of 4% for unemployment (Lovell, M. C. (2004). Economics with Calculus. River Edge, World Scientific Publishing Company.). It's also interesting to remember that William Vickrey in his presidential address to the AEA, in 1993, has made the reference to a value of 1.5%.

[3] Okun, A. (1962). "Potential GNP: its measurement and significance." American Statistical Association, Proceedings of the Business and Economic Statistics Section, pp.98-103, reprint in The Political Economy of Prosperity, W. Norton, New York, 1970, pp.132-45., pp. 136-7 and Kahn, G. (1996). "New Estimates of the U.S. Economy's Potential Growth Rate." Contemporary Economic Policy **14**(4): 1-16..

[4] Gordon, R. (1984). "Unemployment and Potential Output in the 1980s." Brookings Papers on Economic Activity **2**: 537-64. has proposed a relation of 2/1, that he use in his well known manual, Gordon, R. (1990). Macroeconomics. Glenview, Scott, Foresman/Little, Brown Higher Education. .

[5] As had insisted Prachowny, M. F. J. (1993). "Okun's Law: Theoretical Foundations and Revised Estimates." The Review of Economics and Statistics **75**(2): 331-36..

[6] See also Silvapulle, P., I. Moosa, et al. (2004). "Asymmetry in Okun's Law." Canadian Journal of Economics **37**(2): 353-74..

[7] Hoover, K. (2001). Methodology of Empirical Macroeconomics. Port Chester, New York, Cambridge University Press..

2. The OL specification

Following the criticism of (Plosser and Schwert 1979) the OL has one of the following formulations[8], U_t^c and y^c represent deviations from trend:

$$U_t - U_t^* = \alpha - \beta \cdot y_t^c + \varepsilon_t \quad (1.2)$$

$$U_t^c = \alpha - \beta \cdot y_t^c + \varepsilon_t \quad (1.3)$$

$$dU_t = \alpha - \beta \cdot dy_t + \varepsilon_t \quad (1.4)$$

Equation (1.2) is derived directly from Okun's formulation; (1.3) is adapted to cyclical fluctuations; and (1.4) overcomes some of the drawbacks of (1.3). Obviously, in terms of estimation, every equation has to be adapted to the dynamic aspects of the data generating process under study. The β coefficient is very sensitive to model specification, to the dynamic structure, to the method of estimation and to the estimation of cyclical values, when it is the case[9]. The first two formulations of the OL, (1.2) and (1.3), introduce an extra element of uncertainty: the presence of cyclical values of the output and the value of the natural rate of unemployment[10]. Equation (1.4) is the most well-known formulation of OL. But it suffers from a problem of efficiency in the case of co-integration between U and y. In addition to these problems another one should be considered, all the three equations ignore what happens to the unemployment rate as a result of the evolution of the participation rate.

3. The OL interdependency and stability

The application of the *mutatis mutandis* principle raises problems of interdependency between variables that affect the stability of the relation. Let us analyse some of these relations, sometimes omitted, which affect the value of β: relations between employment, labor force and unemployment, in a dynamic context, characterizing the supply and demand of labour - (Weber 1995) and (Sögner and Stiassny 2002); variations of the use of productive capacity - (Watts and Mitchell 1991) and (You 1979); variations in productivity - (Altig, Fitzgerald et al. 1997) and (Kahn 1996) and his cyclical fluctuations - (Hultgren 1960); changes of human capital and working time - (Farsio and Quade 2003) and (You 1979); the cyclical evolution in the participation rate - (Thirlwall 1969)[11]; the labour hoarding phenomenon - (Thirlwall and Ireland 1970) and (Sögner and Stiassny 2002); different protection policies of employees (Sögner and Stiassny 2002), (Blanchard 1999), (Moosa 1997), (Kaufman 1988), (Weber 1995) and (Apergis and Rezitis 2003) - which ultimately may be the result of increased international competition or the attempt to reduce the gap between *insiders* and *outsiders* in the labour market; and the evolution of unemployment rate *hysteresis* - (Sögner and Stiassny 2002).

Another problem concerns the behaviour of OL during the cycle. The idea that the contraction phase of the cycle could be more abrupt that the expansion was raised by (Keynes 1936). (Neftci 1984), (Rothman 1991) and (Brunner 1997) have confirmed this idea for output and unemployment.

---

[8] Where β is designated by "Okun coefficient".

[9] See Moosa, I. A. (1997). "A Cross-Country Comparisom of Okun's Coefficient." Journal of Comparative Economics **24**(3): 335-56.. But You, J. K. (1979). "Capital Utilization, Productivity, and Output Gap." The Review of Economics and Statistics **61**(1): 91-100. arrives to the conclusion that the method used to obtain the potential output is not relevant for model results.

[10] Or a benchmark value, as we saw above. See the long paper of Franke, R. (2006). Themes on Okun's Law and Beyond. New York, Bernard Schwartz Center for Economic Policy Analysis..

[11] See the reply and the counter-reply about the proposal of this author, Monhollon, J. R. and W. E. Cullison (1970). "Okun's Law and the Natural Rate of Growthh: Comment." Southern Economic Journal **37**(2): 231-2. and Thirlwall, A. and N. J. Ireland (1970). "Okun's Law and the Natural Rate of Growth: Reply." Southern Economic Journal **37**(2): 232-4..

While empirical research in terms of dynamic processes can solve some problems[12], we can not expect that all of them are miraculously solved. When a relevant interdependence is not explicitly considered we can have a serious problem of instability of our estimations because there is no reason for a stable relation over time[13]. The instability argument is a matter for empirical refutation. If instability is rejected we can conclude either the omission of those relations does not affect the value of the OL parameters; or they are not important; or their effects are offset.

4. Econometric estimation: non-stationary variables

Current econometric methodology allows the rejection of spurious regressions[14]. The first of the above relations, (1.2), is likely to be a regression of a variable with unit root, I (1), on a stationary variable, I (0). In this case the coefficient of Okun can not be stable or constant. The last equation, (1.4), should lead to include an error term, ECM, in the event of *U* and *y* being CI (1,1)[15] in levels. For (Gordon 1984) one of the problems is that usually we search a value for β with inadequate models. If *U* and *y* are co-integrated, the interpretation of Okun coefficient is not obvious. (Attfield and Silverstone 1998) propose to interpret the original β in terms of the error (ECM) coefficient. But this proposal ignores the interdependence between the variables of a VECM (vector auto-regressive error-correction mechanism) model as (Johansen 1995) advocates it. That is, in these cases we should study the effects of shocks on the variables involved in order to be able to take into account short and long term interdependence.

II. AN APPLICATION TO THE PIGS AREA

1. The evolution of unemployment and production

Data was obtained from AMECO, European Commission data base, on 23rd October 2009. U is the rate of unemployment in percentage; YR, the output, is the log of GDP, at 2000 prices in Mrd Euros; and LPR is the log of the participation rate in percentage. We identify the variables with the suffix of the countries name, POR for Portugal, ITA for Italy, GRC for Greece, SPA for Spain and USA for USA.

2. Stationarity analysis of the variables

We have applied the ADF type tests proposed by (Elliott, Rothenberg et al. 1996), DFGLS, and (Elliott 1999), DFGLSu. The search for the deterministic variables and for the lags was

[12] See Kahn, G. (1996). "New Estimates of the U.S. Economy's Potential Growth Rate." Contemporary Economic Policy **14**(4): 1-16. about the question of productivity evolution.

[13] Several authors identified different forms of the instability of β coefficient: temporal, long term and cross-country. See Okun, A. (1980). Postwar Macroeconomic Performance. The American Economy in Transition. M. Feldstein. Chicago, University of Chicago Press., Davenport, P. (1982). "Technical Change and Unemployment." Journal of Post Keynesian Economics **1**: 34-50., Thurow, L. (1983). Dangerous Currents: The State of Economics. Oxford, Oxford University Press., Gordon, R. (1984). "Unemployment and Potential Output in the 1980s." Brookings Papers on Economic Activity **2**: 537-64., Nguyen, D. T. and A. M. Siriwardana (1988). "The Relationship Between Output Growth and Unemployment: a re-examination of Okun's Law in Australia." Australia Economic Review **81**: 16-27., Adams, C. and D. Coe (1989). "A Systems Approach to Estimating the Natural Rate of Unemployment and Potential Output for the United States." IMF Working Paper **89**., Lee, J. (2000). "The Robustness of Okun's Law: evidence from OECD countries." Journal of Macroeconomics **22**(2): 331-56. and Farsio, F. and S. Quade (2003). "An Empirical Analysis of the Relationship Between GDP and Unemployment." Humanomics **19**(3/4): 1-6..

[14] For a list of empirical studies see Apergis, N. and A. Rezitis (2003). "An examination of Okun's law: evidence from regional areas in Greece." Applied Economics **35**: 1147-51. and also Silvapulle, P., I. Moosa, et al. (2004). "Asymmetry in Okun's Law." Canadian Journal of Economics **37**(2): 353-74..

[15] With a unit root and being co-integrated. See Watts, M. and W. Mitchell (1991). "Alleged Instability of the Okun's Law Relationship in Australia: An Empirical Analysis." Applied Economics **23**(12): 1829-38 .

made by previous estimation of an ADF equation with rejection of errors auto-correlation of order 1. Taking into account the Cochrane precaution, (Cochrane 1991), and the variance ratio test he has proposed, (Cochrane 1988) and (Hamilton 1994), we present also the values of V(k) and A1 for k=10. We have chosen k=10 because we have only 49 observations for each variable. Our results are in Table 1 to 3. In what follows we interpret the results for Spain. In the case of LPR for Spain, for the first two tests, we observe that for the variable and its first difference, the presence of a constant and a trend are justified and for the case without transformation a lag of 1 is included and in first differences no lag is necessary. By DFGLS and also DFGLSu the unit root are not rejected for SPA_LPR but are rejected for the first difference, SPA_d_PR. The variance ratio is obviously above 1 for SPA_LPR and less than unity, with a tendency to zero, for SPA_d_LPR. For this variable a shock of 1% will persist 10 years after at a level of 0.52 (for SPA_LPR the shock was multiplied by more than 3).

Table 1

Countries	Det	Lags	DFGLS	DFGLSu	V(10)	S.D.	A1(10)
ITA_LPR	C, T	1	-1.77	-2.38	3.08	1.7	2.14
ITA_d_LPR	C	0	-3.15***	-3.55***	0.21	0.47	0.12
GRC_LPR	C, T	0	-0.91	-1.63	2.97	1.64	1.77
GRC_d_LPR	C, T	0	-6.58***	-6.65***	0.14	0.08	0.43
SPA_LPR	C, T	1	-0.43	-0.71	6.61	3.65	3.24
SPA_d_LPR	C, T	0	-4.76***	-4.75***	0.23	0.13	0.52
POR_LPR	C	1	-1.61	-1.61	2.83	1.57	1.79
POR_d_LPR	C	0	-4.73***	-4.71***	0.19	0.11	0.49
USA_LPR	C	1	-0.28	-1.4	3.61	1.99	2.16
USA_d_LPR	C	0	-3.43***	-4.04***	0.22	0.12	0.49

Table 2

Countries	Det	Lags	DFGLS	DFGLSu	V(10)	S.D.	A1(10)
ITA_U	C	1	-1.36	-1.65	1.94	1.07	1.52
ITA_d_U	C	0	-4***	-4.41***	0.21	0.12	0.46
GRC_U	C, T	1	-2.27	-2.42	2.77	1.53	1.99
GRC_d_U	C	0	-3.56***	-3.58***	0.3	0.17	0.55
SPA_U	C	1	-1.13	-1.63	2.92	1.62	1.97
SPA_d_U	C	0	-3.68***	-3.67***	0.28	0.16	0.54
POR_U	C	1	-1.72*	-2.35	1.63	0.9	1.47
POR_d_U	C	0	-3.86***	-3.9***	0.26	0.14	0.52
USA_U	C	1	-2.98***	2.99**	0.6	0.33	0.79
USA_d_U	C	0	-4.21***	5.45***	0.11	0.06	0.35

Table 3

Countries	Det	Lags	DFGLS	DFGLSu	V(10)	S.D.	A1(10)
ITA_YR	C	0	1.37	-2.92	3.9	2.15	2.18
ITA_d_YR	C, T	0	-6.05***	-6.42***	0.11	0.06	0.35
GRC_YR	C, T	0	-1.21	-1.85	5.05	2.79	2.4
GRC_d_YR	C	0	-3.07***	-4.95***	0.2	0.11	0.51
SPA_YR	C	1	0.81	-0.06	4.48	2.48	3.18
SPA_d_YR	C	0	-1.43	-3.3***	0.42	0.23	0.65
POR_YR	C	1	0.21	-1.13	2.73	1.51	1.88
POR_d_YR	C, T	0	-4.93***	-4.96***	0.21	0.12	0.46
USA_YR	C, T	1	-2.55	-3.27**	0.58	0.32	0.79
USA_d_YR	C	0	-4.98***	-5.2***	0.12	0.07	0.36

With these results we are in conditions to retaining our variables as having a unit root, with its first differentiation being stationary, so they are I(1).

3. The presence of co-integration and the development of a VAR model

We have not found any interesting C-I relations between unemployment, output and the participation rate, U, YR and LPR, for all of the countries of the area and for the USA. The last variable was taken as endogenous and also as exogenous. So, we decided to study the OL with a VAR model. We know from (Sims 1980) and (Sims, Stock et al. 1990) that it is not correct to differentiate unit roots variables for the estimation of a VAR. Differentiation means that we lost information about co-movements in the data. But using levels doesn't allow the application of exogeneity tests because we might have a non-standard F distribution. Another problem is that the impulse responses for long time horizons are inconsistent estimates of the true process ((Enders 2004)). So the stability of the VAR model has to be addressed and we have to confirm if the responses decay to zero. Nevertheless we propose to simulate only 20 periods of responses.

We will use the Choleski decomposition so for the understanding of different shocks it is important the order of the variables. The unemployment rate is a lagging variable in terms of the output evolution, so our first variable will be unemployment rate and the second output. A shock on U will affect immediately U and YR and a shock on YR will affect immediately YR but only a period later will affect U.

3.1. Variables and the order of the VAR models

We have chosen the variables of our models by the exclusion of the null hypothesis. In the case of endogenous variables by an F test, even knowing that its distribution in presence of unit roots is not the conventional one. The exogenous variables were chosen by the usual t tests. We briefly review as an example the case of Portugal. We have important information on Tables 4 and 5. For Portugal we have a VAR of order 2 with deterministic variables a constant and two lags of LPR, POR(U,YR;C,LPR). We retain 2 lags because the ratio test[16], with a chi-square distribution of order 4 the nullity of the coefficients of order 3 is not rejected but the null is rejected in the case of 2 lags against 1 lag, at the level of 1% of probability. At the same time the Akaike information criteria is better for 2 lags than for 1, and better for 2 than for 3. The observations retained have changed accordingly when we change the maximum lag presented in the estimations. The maximum modulus of the eigenvalues of the matrix associated to the VAR is less than 1, what is required for the stability of the model. We cannot reject at the level of 10% the presence of auto-correlation on the unemployment equation. We have no problem of auto-correlation for the output regression and we cannot reject the normality of errors of the two equations as well as the absence of an ARCH process also for the two equations.

Table 4[17].

	POR(U,YR;C,LPR)		ITA(U,YR;C)		GRC(U,YR;C)		SPA(U,YR, LPR;C)		
Order	3/2	2/1	3/2	2/1	3/2	2/1	3/2	2/1	
LR(4)	6.58	24.7***	1.5	21.8***	2.4	23.9***	12.2	36.8***	
AIC	-2.7/-2.8	-2.8/-2.4	-3.7/-3.9	-3.9/-3.6	-2.4/-2.6	-2.3/-2.0	-9.4/-9.6	-9.7/-9.2	
MxMod	0.96		0.96		0.97		0.99		
	U	YR	U	YR	U	YR	U	YR	LPR
AR(1)	3.7*	0.01	0.34	0.35	1.8	0.03	0.002	1.6	0.07
N	1.5	2.5	2.5	2	0.16	12***	0.22	1.5	17.3***
ARCH(1)	0.23	1.2	0.005	3.3*	2	0.21	0.07	1.4	0.0009

[16] With the corrections suggested by Sims, C. (1980). "Macroeconomics and Reality." Econometrica **48**: 1-49. for small data samples.

[17] The AR test is an LM test with F values; the normality is a chi-squared of order 2; and the ARCH is also an LM test with F values. The auto-correlation tests as the ARCH test are of order 1.

Table 5

	USA(U,YR,LPR;C)		
Order	3/2	2/1	
LR(4)	13.1	42.7***	
AIC	-10.9/-11.2	-11/-10.4	
MxMod	0.99		
	U	YR	LPR
AR(1)	0.02	0.39	13.9***
N	9.6***	10.6***	11.0***
ARCH(1)	0.18	0.23	0.05

Table 6

	PIGS(U,YR,LPR;...)		
Order	3/2	2/1	
LR(4)	10.5	104.6***	
AIC	-10.2/-10.24	-10.2/-9.7	
MxMod			
	U	YR	LPR
AR(1)	2.3	2.5	4.2**

In general we have no serious problems in all the VAR's estimated. But some problems remain and are worth mention. The auto-ÿÿgressive nature of the errors in the case of the USA, for LPR, at the level of 1%. For Spain we have imposed an order of 4 and not 2 as was given by the ratio test and the Akaike information criteria. The reason is simple, with 2 lags, even with 3, but not with 4 lags, the model was instable. The modulus of the eigenvalue of the matrix associated was greater than 1 in those cases. The VAR estimated for the group of PIGS is a fixed-effects panel VAR, PIGS(U,YR,LPR;…). The deterministic parts of this last model are dummies, one for each country. A LR test, with a $\chi^2_{(12)}$=53.3***, rejects the null of these deterministic variables. For this model only the LPR equation has a problem of auto-correlation of order 1 on the errors.

3.2. Stability of models and parameters

On Table 4 and 5 MxMod values indicate that all models are stable. Following the suggestion of (Doornik and Hendry 2001), pp. 222-3, we have applied to our models single equation Chow tests and system Chow tests[18]. In both cases, 1 step F test, break-point F test and forecast F test. The figures are on the Appendix for a significance level of 1%. We have no problems concerning the stability of the parameters in terms of single equation and system for Portugal and Spain. In the case of Italy we have one problem of instability in one of the three tests for the system. For Greece we have problems of instability in one of the equation tests for two equations and one problem in one test for the system. For Greece the problems existed at the end of the 90's. For the USA we have, at the end of the 80's, a problem with one of the test for the system.

We considerer our models are reasonably good in terms of stability for the simulation we proposed.

3.3. Simulation of the responses on employment to a shock on output

On the Appendix we include the figures of the responses from shocks (orthogonal impulses) with the plus and minus the equivalent to the standard deviation measured by the 16% and 84% fractiles (see (Sims and Zha 1999)). The responses must be seen in the column associated with YR. For Portugal, Italy and Greece the effects on U vanish at the end of 5 years, for the USA 3 years and for Spain and the PIGS 9 years.

We simulate the current crisis through the existence of an output shock, by assumption of 10%. We must bear in mind that this is only a simulation exercise. To be more realistic it would be necessary to suppose not only an isolated shock but a succession of shocks. But this could be seen as an addition of lagged responses. We are interested in a comparison between the countries of the PIGS area and the PIGS and a benchmark economy, the USA, so we can compare a single output shock on the output and measure the real consequences in terms of unemployment.

[18] For Spain the system was initialized with 32 observations, the USA with 20, Italy and Greece with 14, Portugal with 18, the PIGS with 26.

The first result is already known; the persistence of a shock on the output is greater for the PIGS than for the USA, the persistence for the first is three times that of the second. In the PIGS area we must distinguish Portugal, Italy and Greece from Spain where the persistence is almost the double.

Table 7

	Output Shock -10%			
	POR	ITA	GRC	SPA
2	1.23	0.87	0.76	3.04
3	**1.54**	**1.15**	0.98	4.26
4	1.21	1.02	**1.01**	**5.78**
5	0.75	0.77	0.9	5.09
6				3.71
7				3.1
8				2.91
9				2.32

Table 7 and 8 is a summary of our results. Bold values represent the maximum effects in terms of the unemployment rate. The effects are more important in Portugal than in Italy or Greece. The most relevant fact is the high level of unemployment created in Spain. The higher value in Spain is nothing more than five times that of Italy and Greece and more than three and half the value in Portugal. At the end of the adjustment, year 9, the value is still the double of the maximum registered for the other PIGS economies, 5 and 6 years before. The PIGS are not homogeneous in this aspect of real functioning of their economies.

Table 8

	Output Shock -10%	
	USA	PIGS
2	**3.89**	2.13
3	2.61	3.17
4		**3.44**
5		3.24
6		2.77
7		2.19
8		1.6
9		1.04

The response on unemployment in the USA is strong but disappears rapidly. In the case of the PIGS, as an area, the effects continue 6 years after they have disappeared in the USA. The consequences in terms of loss of output and welfare are much more important in the PIGS area. But as we have mentioned this area is not homogeneous and this results must be interpreted carefully.

CONCLUSION

We consider that the idea behind the OL continues to be useful in empirical macroeconomics even if its deduction is limited to a *mutatis mutandis* application. We have to handle very carefully stability problems caused by interdependencies excluded explicitly of the relation and also about the ignored problems of non-stationarity of the variables involved. We addressed those drawbacks in our paper. We have included the participation rate in order to take into account the interdependence between the labour supply and the macro conditions of the

economy. And we have analysed the time series characteristics of the data used. Our research in terms of co-integration and the associated VECM has not produced interesting results from a statistical and economic point of views, so we have developed VAR models. We have inspected the issue of parameter instability due to interdependency problems beyond the imposition that the eigenvalues of the VAR companion matrix are in the unit circle. We have obtained a VAR model for each of the economies under study using a ratio test and also the Akaike information criteria for the determination of the model's order. For the case of Spain we have used the first number of lags, after the optimal number, conducting to the maximum modulus of the eigenvalue of he matrix lower than 1. Using those VAR models we simulated an output shock of 10% and we studied the responses in terms of the rate unemployment. The PIGS countries can be separated in two groups: Portugal, Italy and Greece, on the first group, where the consequences in terms of unemployment are not very serious and the adjustment can be accomplished in 5 years; and Spain in the other group where the consequences are very severe and the adjustment is slow and it is accomplished in 9 years. Our benchmark economy has a serious impact on unemployment but the adjustment is fast. The PIGS area, as an entity, has a serious impact on unemployment and a slowly adjustment process like the Spanish case. In terms of the results measured by the unemployment rate there is no reason to identify an economic area like the PIGS.

This paper is part of a wider research that we intend to develop in a near future. In order to ascertain PIGS existence we have to study saving imbalances in the area, within each country, and in a benchmark economy.

REFERENCES:

ADAMS, C. AND D. COE (1989). "A Systems Approach to Estimating the Natural Rate of Unemployment and Potential Output for the United States." IMF Working Paper **89**.

ALTIG, D., T. FITZGERALD, ET AL. (1997). "Okun's Law Revisited: should we worry about low unemployment." Economic Commentary **15**: 1-4.

APERGIS, N. AND A. REZITIS (2003). "An examination of Okun's law: evidence from regional areas in Greece." Applied Economics **35**: 1147-51.

ATTFIELD, C. AND B. SILVERSTONE (1998). "Okun's law, cointegration and gap variables." Journal of Macroeconomics **20**(3): 625-37.

BLANCHARD, O. (1999). Macroeconomics. New Jersey, Prentice Hall International.

BLINDER, A. (1997). "A Core of Macroeconomics Beliefs ?" Challenge **40**(4): 36-44.

BRUNNER, A. D. (1997). "On the dynamic properties of asymmetric models of real GNP." Review of Economics and Statistics **79**(2): 321-326.

COCHRANE, J. (1988). "How Big Is the Random Walk in GNP?" Journal of Political Economy **96**(5): 893-920.

COCHRANE, J. (1991). "A Critique of The Application of Unit Root Tests." Journal of Economic Dynamics and Control **15**: 275-284.

DAVENPORT, P. (1982). "Technical Change and Unemployment." Journal of Post Keynesian Economics **1**: 34-50.

DOORNIK, J. AND D. HENDRY (2001). Modelling Dynamic Systems Using PcGive, Volume II. London, Timberlake Consultants Ltd.

ELLIOTT, G. (1999). "Efficient Tests for a Unit Root When the Initial Observation is Drawn from its Unconditional Distribution." International Economic Review **40**(3): 767-83.

ELLIOTT, G., T. ROTHENBERG, ET AL. (1996). "Efficient Tests for an Autoregressive Unit Root." Econometrica **64**(4): 813-36.

ENDERS, W. (2004). Applied Econometric Time Series. Hoboken, John Wiley & Sons.

EVANS-PRITCHARD, A. (2009). "Euro suffering from 'reserve currency curse' as investors pull out." The Daily Telegraph(28 February).

FARSIO, F. AND S. QUADE (2003). "An Empirical Analysis of the Relationship Between GDP and Unemployment." Humanomics **19**(3/4): 1-6.

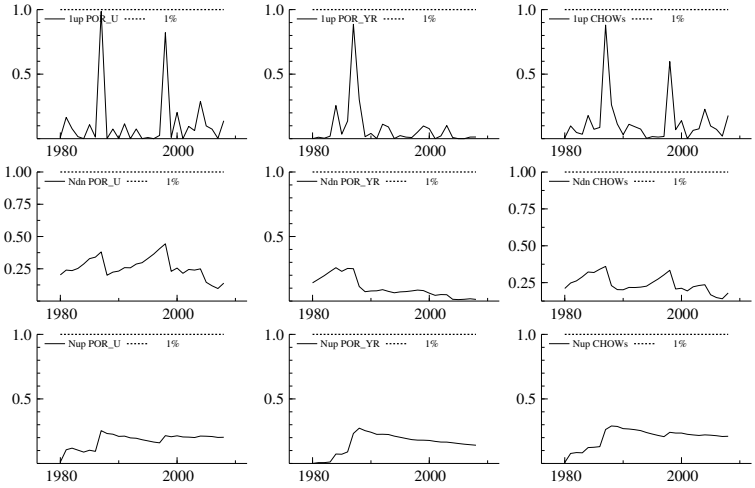
FRANKE, R. (2006). Themes on Okun's Law and Beyond. New York, Bernard Schwartz Center for Economic Policy Analysis.

GORDON, R. (1984). "Unemployment and Potential Output in the 1980s." Brookings Papers on Economic Activity **2**: 537-64.

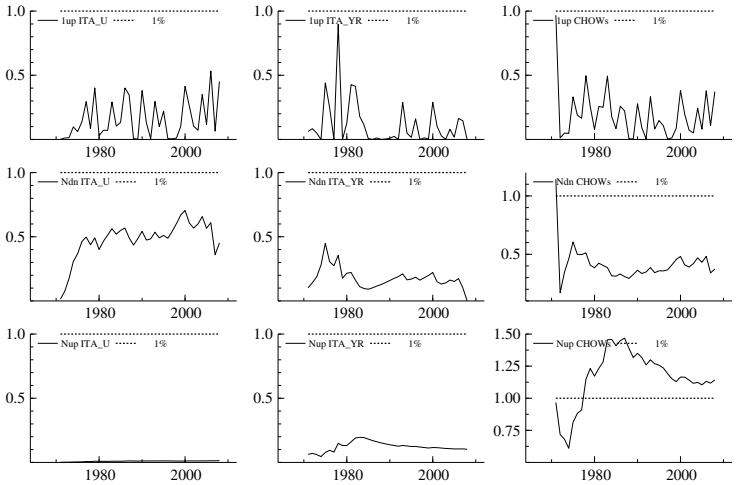
GORDON, R. (1990). Macroeconomics. Glenview, Scott, Foresman/Little, Brown Higher Education.
HALL, R. AND J. TAYLOR (1988). Macroeconomics. New York, Norton & Co.
HAMILTON, J. (1994). Time Series Analysis. Princeton, Princeton University Press.
HOOVER, K. (2001). Methodology of Empirical Macroeconomics. Port Chester, New York, Cambridge University Press.
HULTGREN, T. (1960). "Changes in Labor Cost During Cycles in Production and Business." National Bureau of Economic Research **Occasional Paper 74**.
JOHANSEN, S. (1995). Likelihood-based Inference in Cointegrated vector Auto-Regressive Models. Oxford, Oxford University Press.
KAHN, G. (1996). "New Estimates of the U.S. Economy's Potential Growth Rate." Contemporary Economic Policy **14**(4): 1-16.
KAUFMAN, R. (1988). "An International Comparison of Okun's Law." Journal of Comparative Economics **12**(2): 182-203.
KEYNES, J. M. (1936). The General Theory of Employment, Interest and Money. London, Macmillan.
KNOTEK, E. (2007). "How Useful is Okun's Law?" Economic Review Federal Reserve Bank of Kansas City **Fourth Quarter**: 73-103.
LEE, J. (2000). "The Robustness of Okun's Law: evidence from OECD countries." Journal of Macroeconomics **22**(2): 331-56.
LOVELL, M. C. (2004). Economics with Calculus. River Edge, World Scientific Publishing Company.
MONHOLLON, J. R. AND W. E. CULLISON (1970). "Okun's Law and the Natural Rate of Growthh: Comment." Southern Economic Journal **37**(2): 231-2.
MOOSA, I. A. (1997). "A Cross-Country Comparisom of Okun's Coefficient." Journal of Comparative Economics **24**(3): 335-56.
NEFTCI, S. (1984). "Are economic time series asymmetric over the business cycle?" Journal of Political Economy **92**(2): 307-28.
NGUYEN, D. T. AND A. M. SIRIWARDANA (1988). "The Relationship Between Output Growth and Unemployment: a re-examination of Okun's Law in Australia." Australia Economic Review **81**: 16-27.
OKUN, A. (1962). "Potential GNP: its measurement and significance." American Statistical Association, Proceedings of the Business and Economic Statistics Section, pp.98-103, reprint in The Political Economy of Prosperity, W. Norton, New York, 1970, pp.132-45.
OKUN, A. (1980). Postwar Macroeconomic Performance. The American Economy in Transition. M. Feldstein. Chicago, University of Chicago Press.
PLOSSER, C. AND W. SCHWERT (1979). Potential GNP: Its Measurement and Sgnificance. A Dissenting opinion. Three Aspects of Policymaking, knowledge, data and institutions. K. Brunner and A. Meltzer. Amsterdam, Carnegie-Rochester Conference Series on Public Policy, Morth-Holland**:** 179-86.
PRACHOWNY, M. F. J. (1993). "Okun's Law: Theoretical Foundations and Revised Estimates." The Review of Economics and Statistics **75**(2): 331-36.
ROTHMAN, P. (1991). "Further evidence on the asymmetric behaviour of unemployment rates over the business cycle." Journal of Macroeconomics **13**(2): 291-8.
SILVAPULLE, P., I. MOOSA, ET AL. (2004). "Asymmetry in Okun's Law." Canadian Journal of Economics **37**(2): 353-74.
SIMS, C. (1980). "Macroeconomics and Reality." Econometrica **48**: 1-49.
SIMS, C., J. STOCK, ET AL. (1990). "Inference in Linear Time Series Models with Unit Roots." Econometrica **58**: 113-44.
SIMS, C. AND T. ZHA (1999). "Error Bands for Impulse Responses." Econometrica **67**(5): 1113-55.
SÖGNER, L. AND A. STIASSNY (2002). "An Analysis on the Structural Stability of Okun's Law - a cross-country study." Applied Economics **14**: 1775-87.
THIRLWALL, A. (1969). "Okun's Law and the Natural Rate of Growth." Southern Economic Journal **36**(1): 87-9.
THIRLWALL, A. AND N. J. IRELAND (1970). "Okun's Law and the Natural Rate of Growth: Reply." Southern Economic Journal **37**(2): 232-4.
THUROW, L. (1983). Dangerous Currents: The State of Economics. Oxford, Oxford University Press.
WATTS, M. AND W. MITCHELL (1991). "Alleged Instability of the Okun's Law Relationship in Australia: An Empirical Analysis." Applied Economics **23**(12): 1829-38
WEBER, C. E. (1995). "Cyclical Output, Cyclical Unemployment, and Okun's Coefficient: A New Approach." Journal of Applied Econometrics **10**(4): 433-45.
YOU, J. K. (1979). "Capital Utilization, Productivity, and Output Gap." The Review of Economics and Statistics **61**(1): 91-100.

APPENDIX

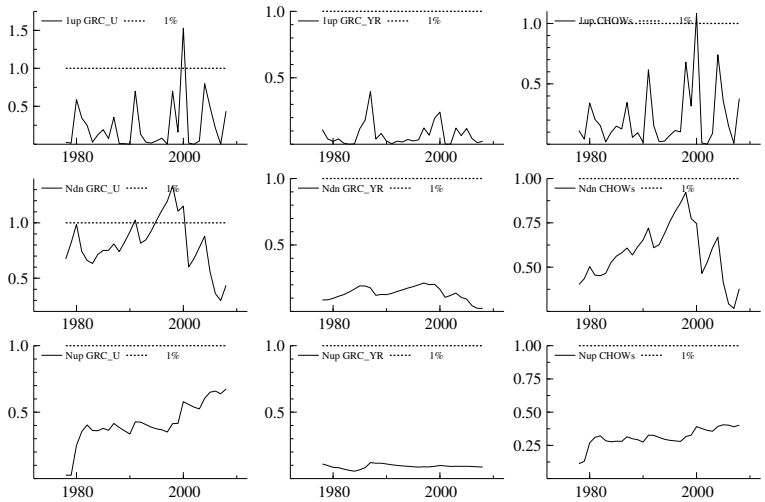
A.1 Parameters stability – Portugal



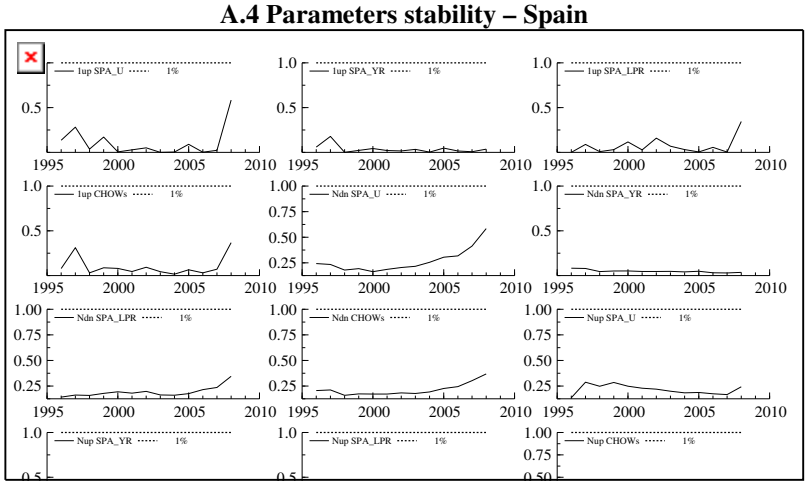
A.2 Parameters stability – Italy



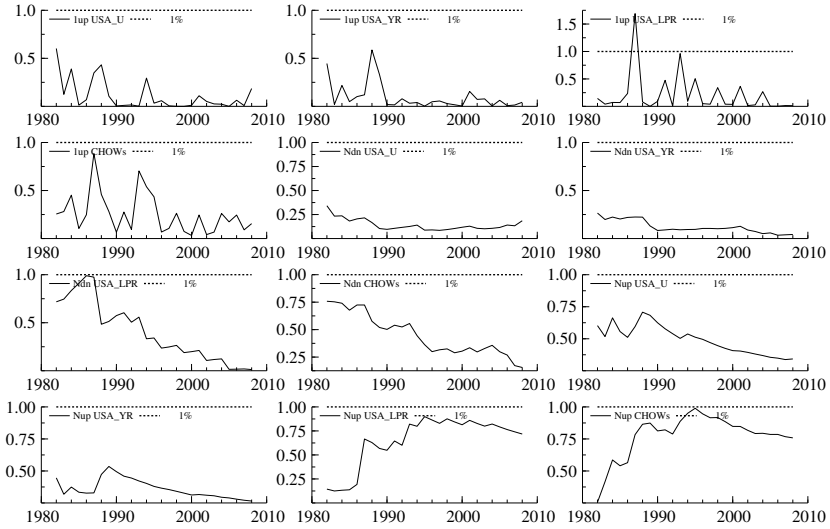
A.3 Parameters stability – Greece



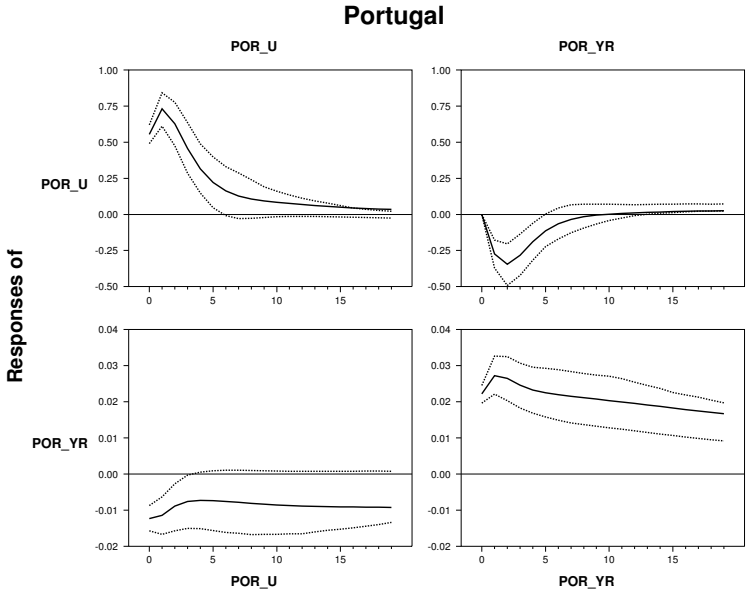
A.4 Parameters stability – Spain



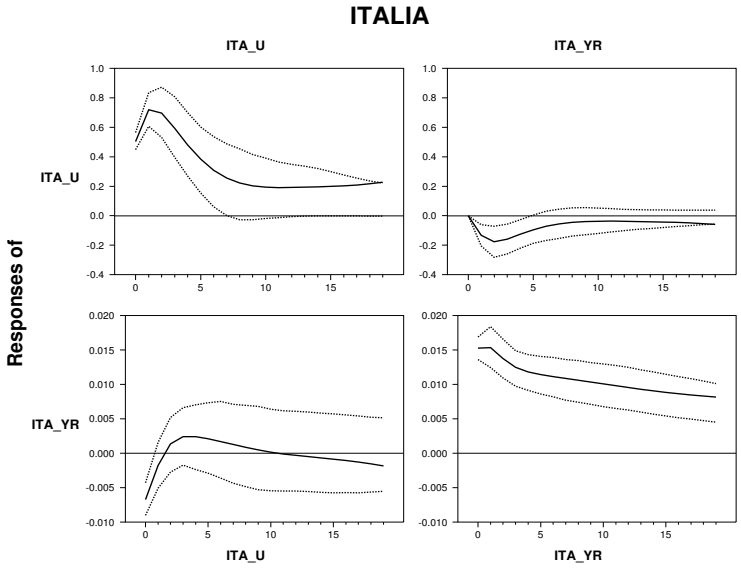
A.5 Parameters stability – USA



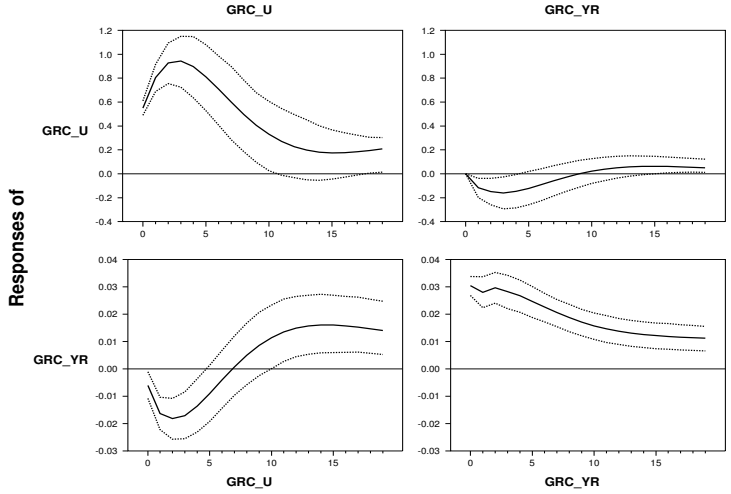
B.1 Orthogonalized Shocks – Portugal



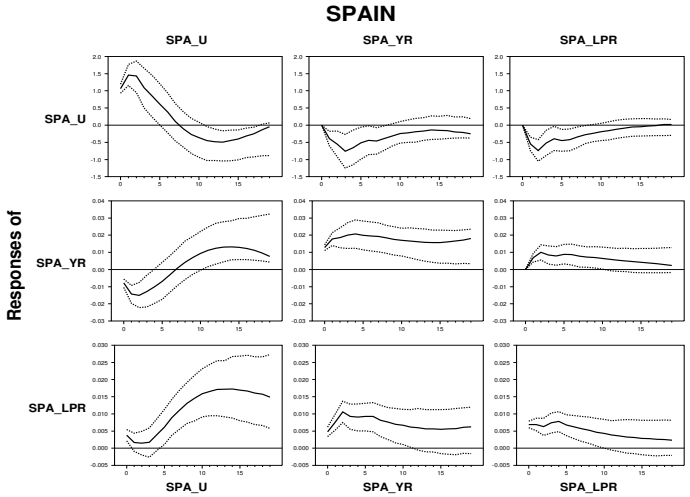
B.2 Orthogonalized Shocks – Italy



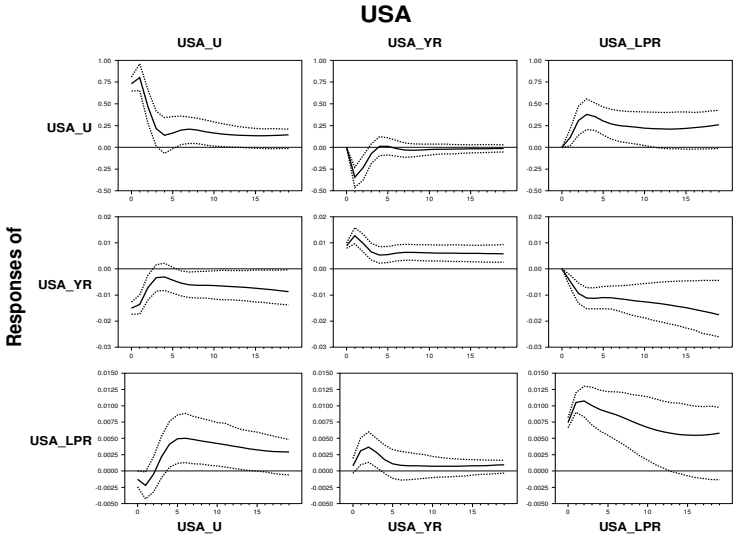
B.3 Orthogonalized Shocks – Greece



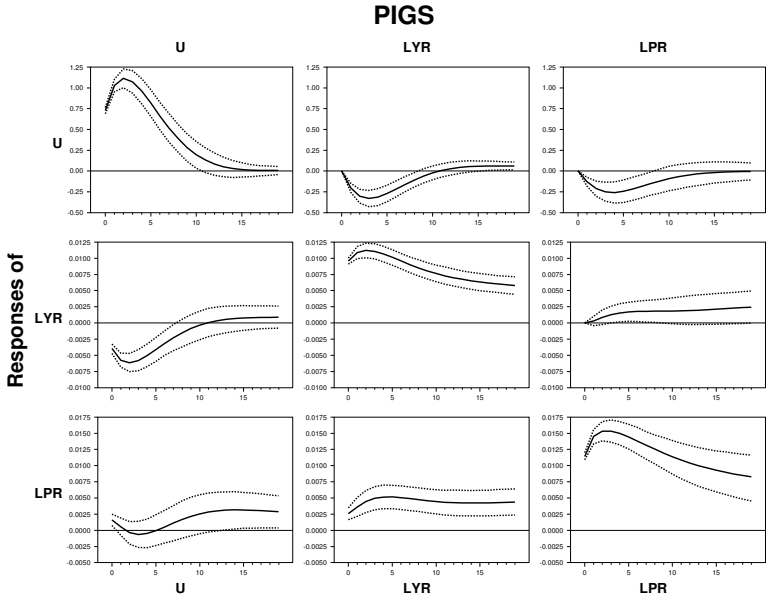
B.4 Orthogonalized Shocks – Spain



B.5 Orthogonalized Shocks – USA



B.6 Orthogonalized Shocks – PIGS



MODELAÇÃO DA PROCURA TURÍSTICA EM PORTUGAL : COINTEGRAÇÃO, MODELOS COM CORRECÇÃO DE ERRO E MODELO ESTRUTURAL BÁSICO

ANA C. M. DANIEL
E-mail: adaniel@ipg.pt
ESTG – Instituto Politécnico da Guarda
CASEE – Centre for Advanced Studies in Economics and Econometrics
PAULO M. M. RODRIGUES
E-mail: pmrodrigues@bportugal.pt
Banco de Portugal
Universidade Nova de Lisboa
CASEE – Centre for Advanced Studies in Economics and Econometrics

Resumo

Neste artigo desenvolvem-se modelos de longo prazo (cointegração) para modelar a procura turística em Portugal e efectuar a posterior ligação com os modelos de curto prazo - Modelos com Correcção de Erro (MCE). No sentido de analisar a performance de previsão destes modelos considera-se ainda um Modelo Estrutural Básico (MEB).
O interesse neste tipo de estudo prende-se com o facto de o turismo ser uma das principais actividades económicas do país, tendo um papel de destaque na criação de emprego, investimento e desenvolvimento de outras actividades económicas.
Os países alvo deste estudo são os principais emissores de turistas para Portugal, ou seja, Alemanha, Espanha, França, Holanda e Reino Unido e ainda Portugal representativo da procura turística interna.
A principal vantagem destes modelos é a de permitirem verificar o impacto na variável dependente ou explicada quando se efectuam alterações nas variáveis causais - o designado *What if Forecasting*.

Palavras-chave: Procura, Turismo, Modelos, Previsão.
Área Temática: Economia Monetária e Financeira.

Abstract

In this paper we develop long-run models (cointegration) to model tourism demand in Portugal and make a later connection with short-term models – Error Correction Models (ECM). In order to analyse the forecasting performance of these models a Basic Structural Model (BSM) is also considered.
The importance of this study is related to the fact that tourism is one of the main economic activities in the country, taking a prominent role in job creation, investment and development of other economic activities.
The countries under study are the main inbound tourism countries to Portugal, *i.e.*, Germany, Spain, France, the Netherlands and the United Kingdom. Portugal is also considered as a representative of the domestic tourism demand.
The main advantage of these models is that they allow us to evaluate the impact on the dependent or explained variable when changes in causal variables occur – the designated *What if Forecasting*.

Key words: Demand, Models, Tourism, Forecasting.
Thematic Area: Monetary and Financial Economics.

1. INTRODUÇÃO

Os modelos desenvolvidos neste artigo são modelos causais ou econométricos. A principal vantagem destes modelos é a de permitirem verificar o impacto sobre a variável dependente

quando se efectuam alterações nas variáveis causais - o designado *What if Forecasting* a que muitos autores se referem. Makridakis (1986) citado por Witt e Witt (1992: 4) e Witt e Witt (1995: 471) salienta nesse sentido:

> "The purpose of econometric models is not purely forecasting. Instead, they attempt to explain economic or business phenomena and increase our understanding of relationships between and among variables. To this direction econometric models provide unique information not available by time series methods".

Até finais da década de 70, os economistas e mais especificamente aqueles preocupados em desenvolverem modelos econométricos baseavam-se no pressuposto de que as séries económicas e o implícito processo gerador de dados[1] era estacionário. No entanto é hoje reconhecido que a maioria das variáveis económicas não é estacionária. No sentido de ultrapassar este problema, duas soluções têm sido propostas. Por um lado, trabalhar apenas com variáveis em diferenças (Metodologia Box-Jenkins), e por outro reter as variáveis originais (em níveis e não em diferenças) nos modelos – cointegração/MCE.

Este é pois um dos principais objectivos deste artigo, ou seja, desenvolver modelos de longo prazo (cointegração) para a procura turística em Portugal e posterior ligação com os modelos de curto prazo (os MCE).

No sentido de analisar a performance de previsão destes modelos considerar-se-á ainda um modelo estrutural básico (MEB). O interesse pelos modelos estruturais de séries temporais tem-se verificado uma vez que se tem obtido melhores resultados em termos de desempenho de previsão quando comparados com outros modelos. Pretendemos verificar se neste caso isso também se passará ou não.

Dado o grande número de dados a utilizar e sua importância neste artigo, far-se-á primeiro uma descrição das variáveis**,** suas principais fontes e características no ponto 2. Segue-se um ponto onde se descreve a metodologia da cointegração e MCE no ponto 3 e posteriormente aplicar-se-á às séries em análise em 4. No ponto 5 apresenta-se a metodologia MEB e aplica-se às séries em estudo. Em 6 analisa-se a precisão de previsão dos modelos desenvolvidos e no ponto 7 finalizaremos com uma conclusão onde se apresentam os pontos-chave do artigo.

2. DADOS

Os países que se considerarão neste estudo são os principais países emissores de turistas para Portugal, ou seja, Alemanha, Espanha, França, Holanda e Reino Unido. Este conjunto de países é responsável por mais de quatro quintos das entradas totais de turistas. Para além da procura externa considerar-se-á também a procura turística interna. De facto, são poucas as investigações que se preocupam em estudar a procura turística dos portugueses pelo próprio país, daí ser esse também um dos nossos objectivos.

Na literatura econométrica de modelação da procura turística, salvo algumas excepções, quase todos os métodos utilizam as mesmas variáveis explicativas: rendimento, preços (custo de vida no país de destino e custo de viajar para o destino), efeitos qualitativos, entre outros.

De salientar no entanto que devido ao facto de algumas das variáveis poderem influenciar uns países e não outros, principalmente a procura turística dos portugueses relativamente à procura turística dos outros países, não se considerarão sempre as mesmas variáveis para todos os países em estudo.

Os dados utilizados neste artigo são trimestrais e referem-se ao período de 1979 a 2006. Os dados de 1979:01 a 2003:04 serão utilizados para se estimarem os modelos e os de 2004:01 a 2006:04 servirão para analisar a precisão de previsão dos modelos estimados.

[1] Processo estatístico pelo qual os dados são gerados (Harris, 1995: 2).

Como medida de procura turística optámos pelo "Número de dormidas nos estabelecimentos hoteleiros incluindo os aldeamentos e apartamentos turísticos". Os dados[2] foram retirados de uma das principais publicações da ex-Direcção Geral do Turismo e ainda do Instituto Nacional de Estatística (INE) e denotar-se-ão por TURALEMANHA, TURESPANHA, TURFRANCA, TURHOLANDA, TURPORTUGAL e TURRUNIDO de forma respectiva para representar a procura turística da Alemanha, Espanha, França, Holanda, Portugal e Reino Unido.

2.1 Variáveis explicativas

De acordo com o que já foi referido anteriormente, são vários os determinantes da procura turística. De uma forma geral, considerar-se-ão os principais e comummente aceites na literatura.

Variável Rendimento

Neste estudo considera-se o Produto Interno Bruto (PIB) a preços constantes para medir o nível de rendimento dos países emissores de turistas (incluindo a procura turística interna). O PIB é medido nas moedas dos países de origem e como os preços a considerar são constantes ter-se-ão em conta os IPC (base 2002) de cada país de origem turística.

Os dados relativos ao PIB da Alemanha foram fornecidos pelo Banco da Alemanha (*Deutsche Bundesbank*) para o período de 1979 a 1991. De 1992 a 2006 retiraram-se os dados das estatísticas da Eurostat. Para Espanha, França[3], Holanda e Reino Unido obtiveram-se os dados desta última fonte. No caso de Portugal os valores foram retirados do Boletim Económico de Verão de 2007 – uma publicação do Banco de Portugal e no qual estão publicadas diversas séries trimestrais para a Economia Portuguesa desde 1977 a 2006 (Banco de Portugal, 2007). Os IPC foram obtidos através das estatísticas da Eurostat.

As séries correspondentes aos rendimentos a preços constantes (2002=100) da Alemanha, Espanha, França, Holanda, Portugal e Reino Unido denotar-se-ão respectivamente por RENDALEMANHA, RENDESPANHA, RENDFRANCA, RENDHOLANDA, RENDPORTUGAL e RENDRUNIDO, sendo cada uma destas variáveis obtidas da seguinte forma:

Rendimento de cada país de origem:

$$\frac{PIB_{PaísOrigem}}{IPC_{PaísOrigem}} \qquad (1)$$

onde $PIB_{PaísOrigem}$ representa o Produto Interno Bruto de cada país de origem (milhões de libras para o Reino Unido e de euros para os restantes países) e o $IPC_{País\ de\ Origem}$ representa o Índice de Preços no Consumidor de cada origem (base: 2002).

Variáveis Preços

No que se refere à variável preços consideram-se dois tipos de preços:
custo de vida dos turistas no país de destino;
custo de viajar para o país de destino.

[2] As séries serão disponibilizadas caso sejam solicitadas.
[3] Também se obtiveram estes dados através do *Institut National de la Statistique et des Études Économiques (INSEE).*

Relativamente ao primeiro custo consideram-se aqui duas variáveis. Uma das variáveis representativas do custo de vida em Portugal é obtida da seguinte forma:

$$CV_{Portugal} = \frac{IPC_{Portugal}}{IPC_{PaísOrigem}} \quad (2)$$

onde $CV_{Portugal}$ representa o Custo de Vida em Portugal em relação ao País de Origem (2002=100); $IPC_{Portugal}$ o Índice de Preços no Consumidor em Portugal (2002=100) e $IPC_{PaísOrigem}$ representa o Índice de Preços no Consumidor do País de Origem (2002=100), para a Alemanha, Espanha, França, Holanda e Reino Unido. Seguindo esta linha de raciocínio para Portugal apenas se terá em conta o seu IPC.

Os dados relativos aos índices de preços no consumidor foram retirados das estatísticas da Eurostat e as séries obtidas e que representam o custo de vida em Portugal para os turistas alemães, espanhóis, franceses, holandeses, portugueses e ingleses denotar-se-ão de forma respectiva por: CVALEMANHA, CVESPANHA, CVFRANCA, CVHOLANDA, CVPORTUGAL e CVRUNIDO.

Uma outra variável que na nossa opinião é importante, que se pretende destacar dos restantes custos, e que grande parte dos estudos não considera mas que representa um custo que é tido em conta pelos turistas aquando da sua procura pelo país de destino, são os preços da hotelaria. Para medir esta variável teve-se em conta os preços em quarto duplo na hotelaria geral[4], cujos dados foram recolhidos no Centro de Documentação do "Turismo de Portugal, I.P". Denota-se esta variável por PHPORTUGAL e será utilizada para todos os países e para os anos em estudo.

No que diz respeito ao custo de viajar para o destino considerar-se-ão dois tipos de custos. São eles o custo de transporte terrestre e o custo de transporte aéreo. Estas variáveis não serão consideradas no caso dos turistas portugueses uma vez que estes custos representam o quanto custa viajar até Portugal e estes são importantes na escolha de Portugal como país de destino turístico, mas isto para quem vem do exterior.

No que se refere ao custo de transporte aéreo, não se utilizará esta variável para o modelo a desenvolver no caso dos turistas oriundos de Espanha uma vez que depois de analisada a composição das entradas de turistas segundo a via (estatísticas do INE e dados fornecidos pelo "Turismo de Portugal, I.P"), verificou-se que a percentagem de turistas deste país entrados por via aérea é bastante reduzida (entre 2004 e 2006 essa percentagem variou entre 14 e 19%), daí que não se justifique a sua inclusão.

Relativamente ao custo de transporte terrestre e seguindo o mesmo raciocínio não se considerará esta variável para os turistas vindos do Reino Unido uma vez que segundo as mesmas fontes mais de 90% destes turistas utilizam o transporte aéreo, daí que não se incluirá no modelo a desenvolver para este país.

Os custos de transporte terrestre assumem neste estudo uma grande importância principalmente no caso da Espanha onde mais de 80% dos turistas utilizam a via terrestre para se deslocarem para o nosso país, seguido da França embora em menor percentagem (60% a 65%). Os turistas alemães e seguidamente os holandeses utilizam também esta via, apesar de a via aérea ser a preferida, sendo cada vez mais utilizada. No entanto e apesar disso, não se pode ignorar a sua influência, daí ter-se optado por considerá-la.

Para o cálculo destes custos têm-se em conta três factores distintos:

- os preços de combustíveis;
- as distâncias em quilómetros entre as capitais dos países emissores e Lisboa;

[4] Todos os preços estão em euros.

- o consumo médio de combustível.

Os custos pretendidos resultam do produto destas variáveis.

No que se refere aos preços dos combustíveis, utilizam-se os preços de venda (dados trimestrais), medidos na moeda de cada país, da gasolina super desde os anos de 1979-1994 e os preços da gasolina super sem chumbo (95) de 1995-2006. Esta opção deve-se ao facto de para os primeiros anos os preços da gasolina sem chumbo não estarem disponíveis uma vez que era um combustível não utilizado, assim como a gasolina super para os tempos mais recentes (para alguns países). Estas séries foram obtidas através da *International Energy Agency* (IEA) e da Direcção Geral de Energia e Geologia pertencente ao Ministério da Economia e da Inovação.

Relativamente às distâncias em quilómetros entre as capitais dos países emissores de turistas e Lisboa, os dados foram fornecidos pelo Instituto Geográfico do Exército.

No que diz respeito ao consumo médio de combustível considera-se o consumo médio de combustível de um veículo ligeiro de passageiros (litros/100kms)[5]. Para a Alemanha os dados foram obtidos através da ADAC - Allgemeiner Deutscher Automobil-Club sedeada em Munique. Para a França a variável foi fornecida pelo *Ministére de l' Économie, de l'Índustrie et de l' Emploi.* No caso da Holanda foram as Estatísticas holandesas a forneceram a informação (CBS – *Central Bureau of Statistics*). No caso da Espanha, os dados foram conseguidos através do *Ministerio de Industria y Energía* (1979-1997)[6] e também pelo *Ministerio de Fomento*. Para os anos mais recentes os dados foram conseguidos através do *Ministerio de Fomento* e do *Instituto para la Diversificación y Ahorro Energético* – IDAE.

Os custos de transporte terrestre entre Portugal e os quatro países europeus em causa, obtêm-se, multiplicando essas três variáveis. Denotar-se-ão estes custos por CTERRESTREALEMANHA, CTERRESTREESPANHA, CTERRESTREFRANCA e CTERRESTREHOLANDA de forma respectiva para a Alemanha, Espanha, França e Holanda. Estas variáveis encontram-se deflaccionadas pelo IPC (base: 2002) pelo que as séries obtidas são a preços constantes.

No que diz respeito ao custo de transporte aéreo consideraram-se as tarifas IATA (International Air Transport Association) e foram retiradas de manuais APT (Air Passenger Tariff) publicadas por este organismo. Estas tarifas têm como referência de partida e de chegada as capitas[7] dos países em análise e foram fornecidas pela IATA da sua sede na Suiça. Todas as séries foram convertidas para a actual moeda em vigor nos países em estudo (euros) para a Alemanha, França, e Holanda e libras no caso do Reino Unido, tendo ainda sido deflacionadas pelo IPC (base 2002) e denotar-se-ão respectivamente por CTAEREOALEMANHA, CTAEREOFRANCA, CTAEREOHOLANDA e CTAEREORUNIDO.

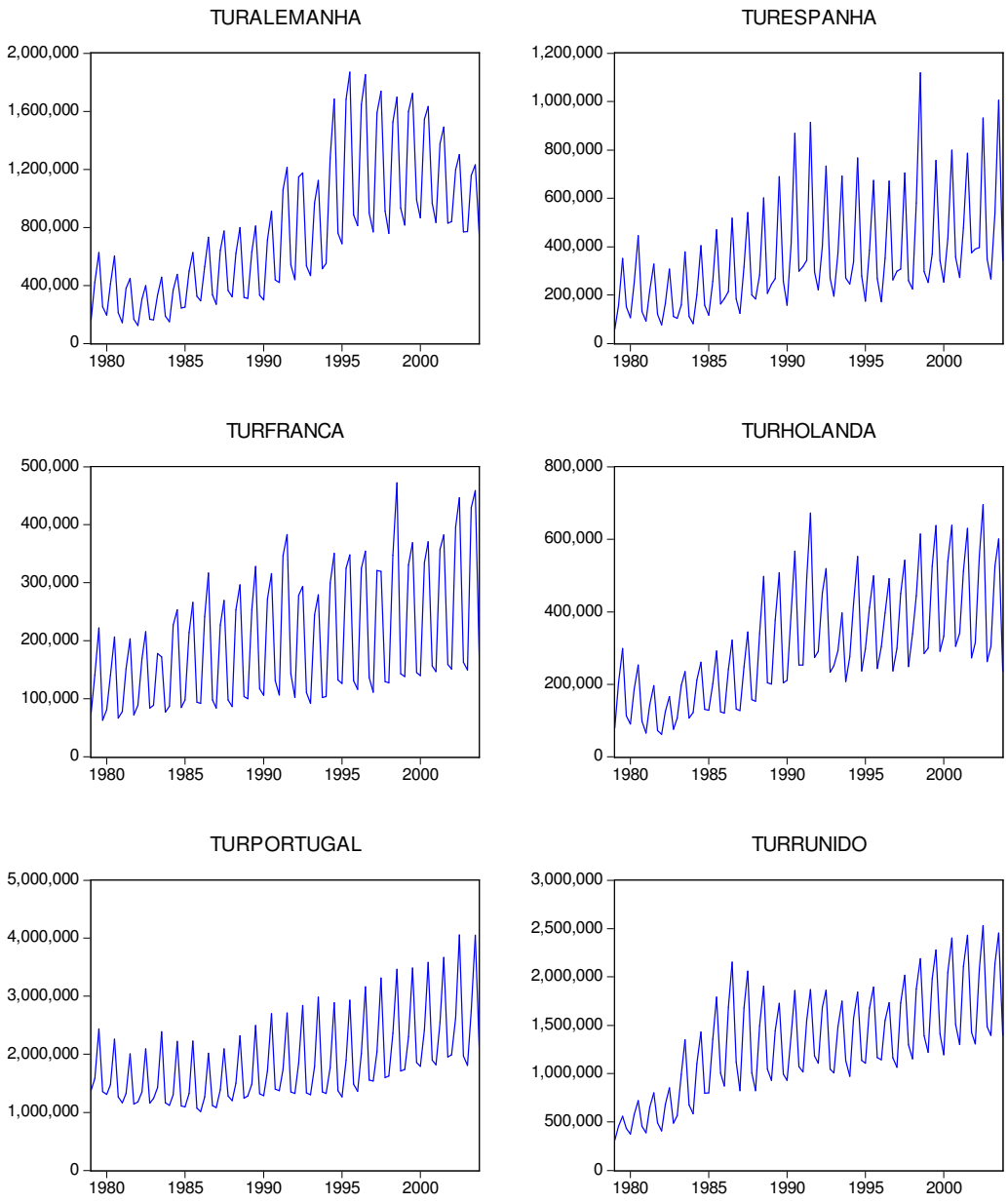
Variável Qualitativa

No que se refere a variáveis qualitativas considerar-se-á uma *dummy* – representativa da EXPO em 1998 que se denotará por $VQ_{1t} = 1$ se t = 1998 e 0 em caso contrário.

A representação gráfica das variáveis encontra-se nas Figuras 1 a 6.

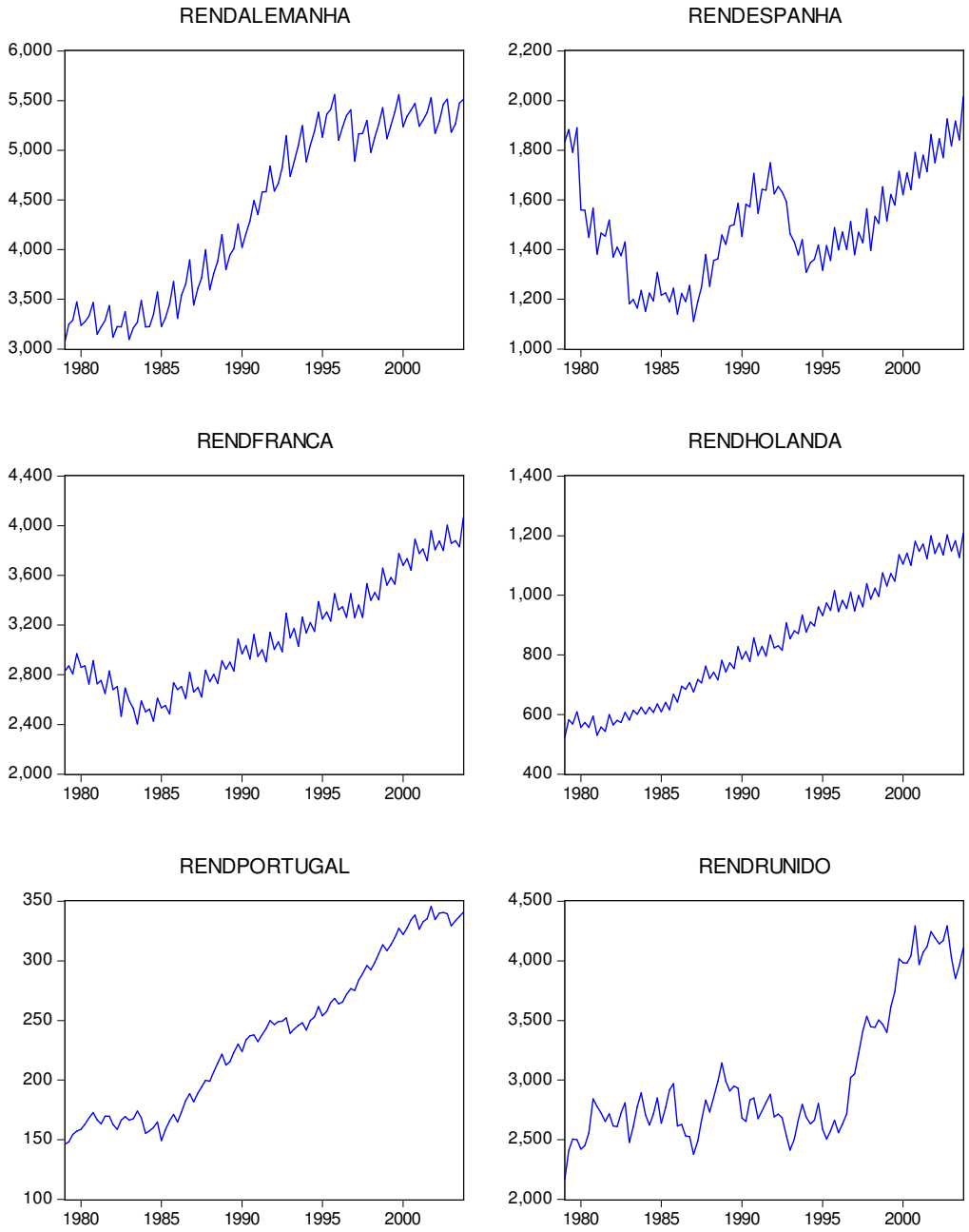
[5] Nesta série os dados são anuais – não existem valores trimestrais para estes dados, pelo que para os quatro trimestres se considera sempre o mesmo consumo, durante o ano.
[6] Actual Ministério de Industria, Turismo y Comercio.
[7] Apenas a opção "*one-way*" origem-destino.

Figura 1 – Procura Turística dos Principais Países Emissores e Procura Interna



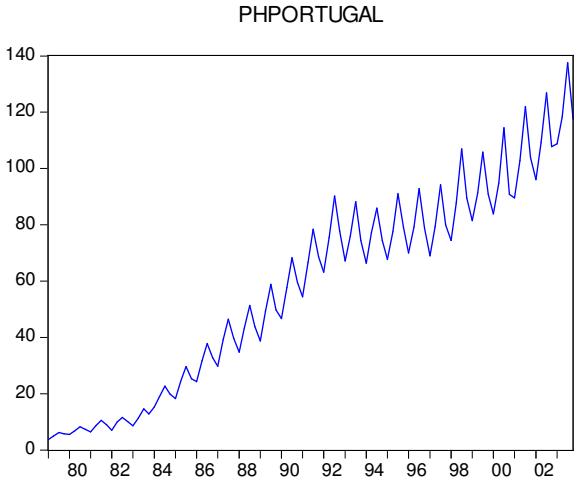
Fonte: Elaboração própria com base nos resultados obtidos pelo programa informático *Eviews (6.0)*.

Figura 2 – Rendimento dos Turistas dos Principais Países Emissores e Portugal



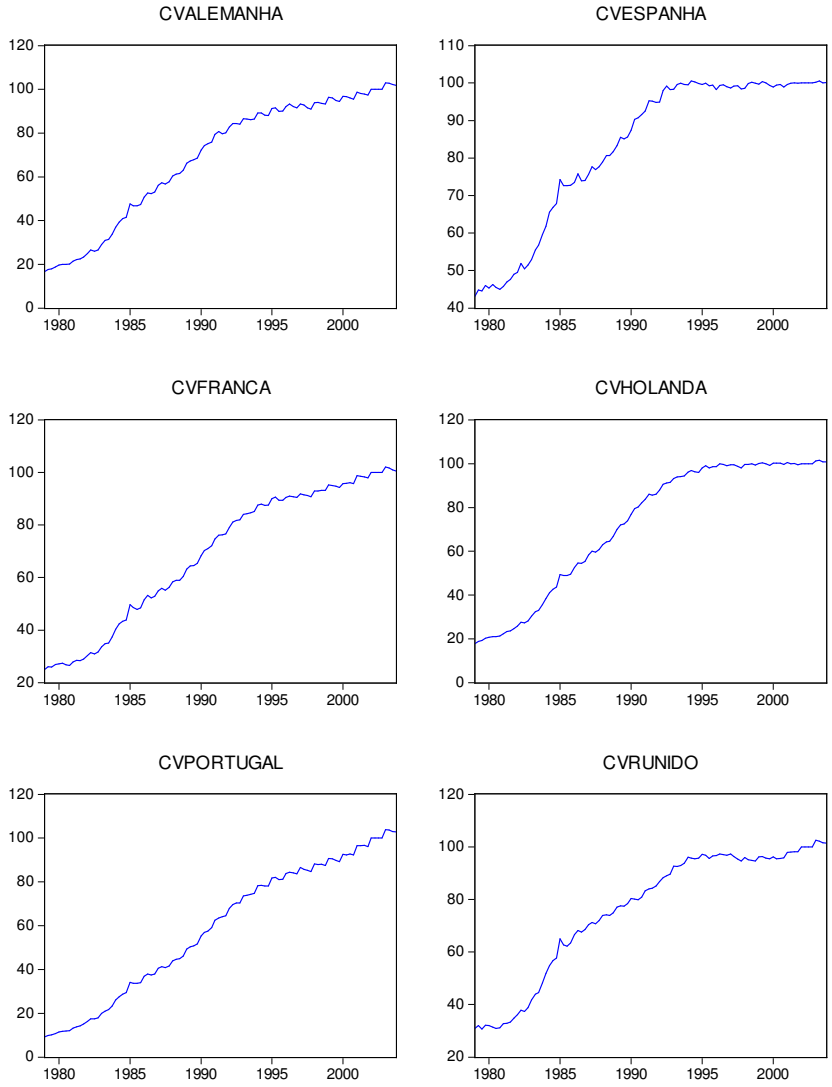
Fonte: Elaboração própria com base nos resultados obtidos pelo programa informático *Eviews (6.0).*

Figura 3 – Preços da Hotelaria em Portugal



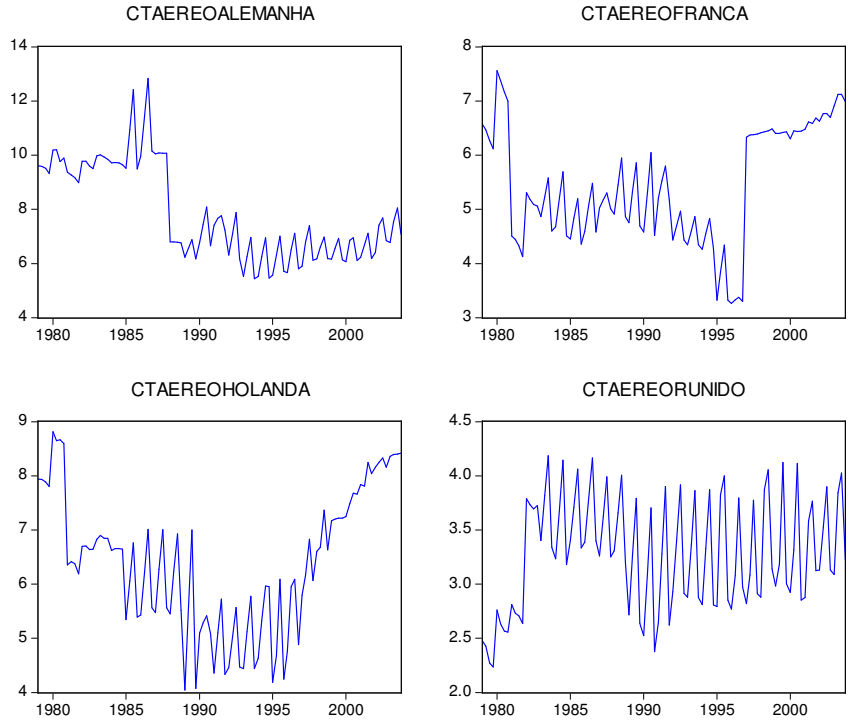
Fonte: Elaboração própria com base nos resultados obtidos pelo programa informático *Eviews (6.0).*

Figura 4 – Custo de Vida em Portugal para os Turistas dos Principais Países Emissores e Portugal



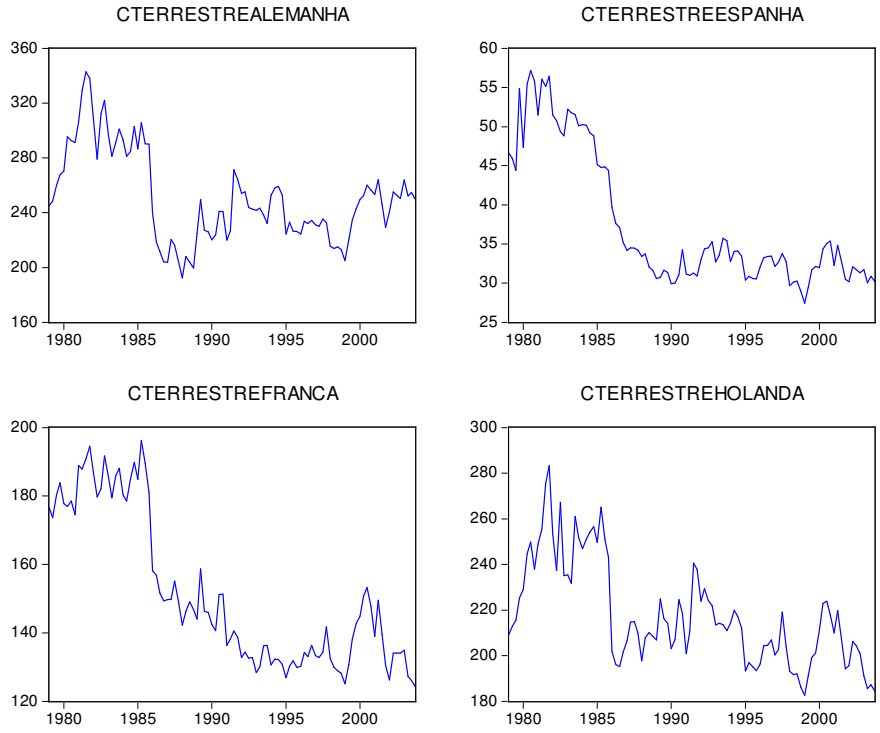
Fonte: Elaboração própria com base nos resultados obtidos pelo programa informático *Eviews (6.0).*

Figura 5 – Custo de Transporte Aéreo para os Turistas da Alemanha, França, Holanda e Reino Unido



Fonte: Elaboração própria com base nos resultados obtidos pelo programa informático *Eviews (6.0).*

Figura 6 – Custo de Transporte Terrestre para os Turistas da Alemanha, Espanha, França e Holanda



Fonte: Elaboração própria com base nos resultados obtidos pelo programa informático *Eviews (6.0).*

Depois da apresentação dos dados descreve-se no próximo ponto a metodologia da Cointegração e MCE.

3.COINTEGRAÇÃO E MODELO CORRECTOR DE ERRO

Dentro do conjunto de variáveis não estacionárias existe um grupo específico que se designa por variáveis integradas. Uma variável X_t é integrada de ordem (d,D) e denota-se por X_t~I(d,D) se tiver que ser diferenciada d e D vezes para se tornar estacionária representando cada uma delas de forma respectiva a diferenciação regular e sazonal. Diz-se também que a série terá d ou D raízes unitárias de ordem regular e de ordem sazonal.

O conceito de cointegração surgiu no sentido ultrapassar os problemas que poderão advir do facto de as séries serem não estacionárias.

3.1 Cointegração

O conceito de cointegração foi introduzido por Granger e Weiss (1983) e Granger (1981). No entanto, é com Engle e Granger (1987) que este é desenvolvido em pormenor.

Se duas ou mais séries não estacionárias estiverem ligadas por uma combinação linear para que haja uma relação de equilíbrio de longo prazo, então mesmo que isoladamente sejam não estacionárias, elas irão ter um percurso bastante próximo ao longo do tempo e a diferença entre elas será estacionária. Esta é a interpretação económica de cointegração.

De acordo com a definição de Engle e Granger (1987), diz-se que as variáveis são cointegradas se:

1º Todos os elementos de X_t onde Xt é um vector de n variáveis, forem integrados da mesma ordem.

2º Existe um vector $\alpha = (\alpha_1, \alpha_2, ...,\alpha_n)$, tal que a combinação linear $Z_t = \alpha' X_t$ é integrada de ordem inferior, e $\alpha \neq 0$.

O vector α designa-se por vector de cointegração. Se no entanto existir um sistema de r vectores α_i (i = 1,...,r), então a matriz de vectores de cointegração (n×r) designa-se por matriz de cointegração. O número de vectores de cointegração é definido como a ordem de cointegração de X_t, isto é, se X_t for constituído por n variáveis então poderão existir até n-1 vectores de cointegração.

O conceito de cointegração está relacionado com o de equilíbrio de longo prazo. Uma definição de equilíbrio entre um conjunto de variáveis X_t será uma relação da forma:

$$\alpha' X_t = 0 \qquad (3)$$

Esta relação define-se como relação de cointegração – relação de equilíbrio ou de longo prazo entre as variáveis. No entanto, uma vez que essa situação é difícil de ocorrer,

$$Z_t = \alpha' X_t \qquad (4)$$
$$(r\times 1)\ (r\times n)\ (n\times 1)$$

mede o grau de desequilíbrio do sistema, ou seja a distância a que o sistema se encontra da situação de equilíbrio representando portanto o seu erro de equilíbrio.

Para além de se poder estabelecer uma relação de equilíbrio ou relação de longo prazo entre duas ou mais variáveis é possível existir ainda uma compatibilização entre o curto e o longo prazo traduzido no Teorema de Representação de Granger.

3.2 Cointegração e MCE – Teorema da Representação de Granger

Se duas ou mais variáveis forem cointegradas estas podem ter uma representação de curto prazo através de um mecanismo que se designa por mecanismo corrector de erro – o modelo MCE. A relação de equilíbrio atrás referida é introduzida como uma variável adicional na forma de termo corrector de erro. Por outro lado o reverso é também possível, ou seja se um vector de variáveis está representado através de um MCE então essas variáveis estão cointegradas.

Apesar de estes modelos terem aparecido antes, é apenas com Engle e Granger (1987) que é apresentada a relação entre estes modelos e o conceito de cointegração. Esta relação é conhecida como Teorema da Representação de Granger[8] que consiste no seguinte: Seja X_t um vector de n variáveis (n×1), cointegradas e com r vectores de cointegração ($0 < r \leq n-1$). Este vector terá então a seguinte representação correctora de erro:

$$A(B)(1-B)X_t = -\gamma Z_{t-1}+u_t \quad (5)$$

onde: $Z_t = \alpha' X_t$

(r×1) (r×n) (n×1)

u_t – vector de variáveis dos resíduos estacionários

A(B) – matriz de polinómios no operador lag B.

γ - coeficientes do termo corrector do erro ou coeficientes de ajustamento, isto porque $Z_{t-1} = \alpha' X_{t-1}$ conhecido precisamente por termo corrector de erro, representa o desequilíbrio entre as variáveis X_{it} no período t-1. Quanto mais elevados forem estes coeficientes, maior a resposta de X_{it} ao desvio do período anterior em relação ao equilíbrio de longo prazo (erro de equilíbrio do período anterior – Z_{t-1}). O sinal negativo, indica que a próxima variação em X_{it} será de sinal contrário a Z_{t-1}.

Os modelos ECM podem ser interpretados como o mecanismo do desequilíbrio que dirigem a economia para uma situação de equilíbrio (Peixe, 1998). Significa isto que no período t os agentes económicos corrigem uma parte do desequilíbrio detectado no período t-1.

Por outro lado, a equação anterior separa os efeitos de longo prazo ($\gamma\alpha'$) dos efeitos de curto prazo (A(B)),

$$A(B)(1-B)X_t = -\gamma Z_{t-1}+u_t = -\gamma\alpha' X_{t-1}+u_t \quad (6)$$

Note que,

$Z_{t-1}=\alpha' X_{t-1}$

Johansen e Juselius (1990) e Johansen (1988) propõem duas estatísticas para determinar o número de vectores de cointegração.

1º - Teste do Traço

Este teste consiste em calcular a estatística

$$\lambda traço(r) = -T\sum_{i=r+1}^{n}\ln(1-\hat{\lambda}_i) \quad (7)$$

onde T é o número de observações e $r = 0,\ldots,n-1$.

sob a hipótese nula de que existem no máximo r vectores de cointegração (ou n-r) raízes unitárias), ou seja: $H_0 : \lambda_i = 0 \;\; i = r+1,\ldots,n$, onde apenas os *r* primeiros valores são diferentes de zero. O número de graus de liberdade do teste é igual ao número de restrições (n-r). O procedimento do teste consiste em calcular sucessivamente a estatística (7) para $r = 0$ (não existência de vectores de cointegração), $r \leq 1$, $r \leq 2$ até que não se rejeite H_0. O correspondente valor de r coincide com o número de vectores de cointegração.

2º - Teste do Máximo Valor Próprio

Este teste ensaia a hipótese de que existem no máximo r vectores de cointegração contra a alternativa de que existem *r+1*. O teste surge de forma complementar ao anterior e a estatística é a seguinte:

$$\lambda \max(r, r+1) = -T\ln(1-\hat{\lambda}_{r+1}) \quad (8)$$

$$com, r = 0,\ldots,n-1$$

Tanto o primeiro teste como o segundo têm uma distribuição assimptótica cujos valores críticos foram obtidos por Johansen e Juselius (1990) e Johansen (1988).

Depois desta descrição passaremos agora à análise de resultados destes modelos.

[8] Johansen (1988) e Engle e Granger (1987).

4. RESULTADOS DA COINTEGRAÇÃO E MCE

Uma das primeiras condições para que duas ou mais variáveis sejam cointegradas, é a de serem integradas da mesma ordem. Nesse sentido inicia-se esta análise por verificar o nível de integração das séries em estudo.

4.1 Análise do nível de integração das séries

Existem diversos testes para analisar a presença ou não de raízes unitárias nas séries. Dado o vasto número de variáveis em análise e a necessidade de as séries serem integradas da mesma ordem, utilizar-se-á neste capítulo o teste HEGY de Hylleberg *et al.* (1990).

O teste HEGY permite determinar o número de raízes unitárias nas frequências anuais e semi-anuais (sazonalidade estocástica) assim como na frequência zero (tendência temporal estocástica). Dada uma série x_t, o teste consiste em testar separadamente a significância dos parâmetros (testes t e F) da regressão que se segue,

$$y_{4t} = \pi_1 y_{1t-1} + \pi_2 y_{2t-1} + \pi_3 y_{3t-2} + \pi_4 y_{3t-1} + \mu_t + \varepsilon_t \quad (9)$$

onde:

$$y_{1t} = (1 + B + B^2 + B^3)x_t$$

$$y_{2t} = -(1 - B + B^2 - B^3)x_t$$

$$y_{3t} = -(1 - B^2)x_t$$

$$y_{4t} = (1 - B^4)x_t$$

$$Bx_t = x_{t-1}$$

μ_t é uma componente determinística que poderá incluir *dummies* sazonais, tendência temporal e um termo constante e ε_t é o erro aleatório. Este é um teste que pode ser estimado via mínimos quadrados com desfasamentos adicionais de Y_{4t}, no sentido de branquear a componente do erro. No que se refere às hipóteses do teste, a hipótese nula $H_0 : \pi_1 = \pi_2 = \pi_3 = \pi_4 = 0$, e que se denota por I(1,1,1), significa que existe uma raíz unitária na frequência zero $(\pi_1 = 0)$, na frequência sazonal semi-anual $(\pi_2 = 0)$ e na frequência sazonal anual $(\pi_3 = \pi_4 = 0)$. As hipóteses devem ser testadas separadamente. Se os testes t individuais $(\pi_1 = 0)$ e $(\pi_2 = 0)$ forem significativamente diferentes de zero (rejeição de H_0), e se o teste F conjunto $(\pi_3 = \pi_4 = 0)$ mostrar que, cada um ou ambos, também o são, então pode-se concluir que a ordem de integração da série é I(0,0,0), o que significa que a série é estacionária. Se a hipótese $H_0 : \pi_1 = 0$ não for rejeitada, mas as restantes sim, então a ordem de integração da série é I(1,0,0), indicando que a série contém uma raíz unitária na frequência zero, mas não tem raízes sazonais. Uma série só será I(1,1,1) se nenhum dos coeficientes for significativamente diferente de zero. Os autores deste teste apresentam ainda no seu artigo (Hylleberg *et al*, 1990) os valores críticos para as raízes unitárias nas frequências zero e sazonais.

O teste foi aplicado aos logaritmos (de base natural) das séries e a equação (9) foi testada considerando duas situações diferentes, (i) com constante e *dummies* sazonais, (ii) com constante, *dummies* sazonais e tendência temporal. Aquando da aplicação dos testes t e F, utilizou-se o teste de autocorrelação de Lagrange (ordem quatro) e adicionaram-se termos desfasados de y_{4t} no sentido de remover a autocorrelação dos erros quando necessário. Os resultados do teste encontram-se na Tabela 1[9].

[9] Para se testar $H_0 : \pi_3 = \pi_4 = 0$, considerou-se o teste F.

Tabela 1 – Ordem de Integração das Séries: Resultados do Teste HEGY

País					
Variável	logTUR	logREND	logCV	logCTerrestre	logCTAereo
Alemanha	I(1,0,1)	I(1,1,1)	I(1,0,1)	I(1,0,0)	I(1,0,1)
Espanha	I(1,0,1)	I(1,1,1)	I(1,0,1)	I(1,0,0)	-
França	I(1,0,1)	I(1,1,1)	I(1,0,1)	I(1,0,0)	I(1,0,0)
Holanda	I(1,1,1)	I(1,1,0)	I(1,0,1)	I(1,0,0)	I(1,0,1)
Portugal	I(1,1,1)	I(1,0,1)	I(1,1,1)	-	-
Reino Unido	I(1,0,1)	I(1,1,1)	I(1,0,1)	-	I(1,1,1)

Fonte: Elaboração própria com base nos resultados obtidos pelo programa informático *Eviews (6.0).*
Notas: 1) ln denota os logaritmos das variáveis em estudo.
2) A variável lnPHPortugal que vai ser introduzida em todas as equações dos países em estudo mostrou ser I(1,1,1).

De acordo com o teste HEGY, todas as séries possuem raízes unitárias na frequência zero. Algumas não possuem raízes unitárias sazonais, como é o caso das variáveis representativas do custo de transporte terrestre sendo I(1,0,0). Apenas a variável rendimento cujo país de origem é a Holanda, é I(1,1,0), o que significa que tem uma raíz unitária na frequência sazonal semi-anual, mas não frequência sazonal anual. Com excepção das I(1,0,0), as restantes séries são sempre não estacionárias (ordem 1) nesta última frequência, sendo que possuem uma raíz também na frequência sazonal semi-anual e na frequência zero, ou seja são I(1,1,1), ou na frequência zero I(1,0,1).

Para se poder prosseguir este estudo e determinar as equações de longo prazo, é necessário aplicar determinados filtros, ou seja remover as raízes unitárias sazonais, para que todas as variáveis sejam integradas da mesma ordem, ou seja I(1,0,0). Nesse sentido, aplicou-se o filtro $S_1(B) = (1 + B + B^2 + B^3)$ às séries I(1,1,1), $S_2(B) = (1 + B^2)$ às séries I(1,0,1), e $S_3(B) = (1 + B)$, à variável rendimento holandesa[10].

4.2 Determinação do número de vectores de cointegração e equações de longo prazo

Neste ponto apresentam-se o número de relações de longo prazo, ou seja, o número de vectores de cointegração, para cada país em estudo e respectivas estimativas. Para tal, utilizou-se o teste do máximo valor próprio, cujos resultados se encontram na Tabela 2.

Tabela 2 – Número de Vectores de Cointegração

País	Nº de vectores *r*	Estatística Máximo Valor Próprio	Valores Críticos (5%)
Alemanha	r = 0	101.9053	36.63019
	r = 1	45.34616	30.43961
	r = 2	22.19302 (*)	24.15921
Espanha	r = 0	61.99244	33.87687
	r = 1	25.95463 (*)	27.58434
França	r = 0	67.30865	36.63019
	r = 1	21.89712 (*)	30.43961
Holanda	r = 0	49.37843	43.41977
	r = 1	30.12483 (*)	37.16359
Portugal	r = 0	51.15891	27.58434
	r = 1	20.72340 (*)	21.13162
Reino Unido	r = 0	42.30022	30.43961
	r = 1	26.35335	24.15921
	r = 2	17.39153 (*)	17.79730

Fonte: Elaboração própria com base nos resultados obtidos pelo programa informático *Eviews (6.0).*

[10] Ver Kulendran e Wong (2005), Kulendran e Witt (2001), Kulendran (1996) e Engle, Granger e Hallman (1989).

Nota: (*) Não é possível rejeitar Ho ao nível de 5%. Relembrando, o teste considera como hipótese nula de que r=q vectores contra a alternativa de ter um vector adicional (r=q+1).

Como se pode verificar, foi possível encontrar dois vectores para a Alemanha e Reino Unido e para os restantes países apenas um. As estimativas de longo prazo correspondentes a esses vectores de cointegração são as que se seguem.

ALEMANHA :

$$S_2(B)\log TurAlemanha_t = 1.0727\, S_1(B)\log \mathrm{Re}ndAlemanha_t + 1.3471\, S_2(B)\log CVAlemanha_t$$
$$- 0.3474\, S_1(B)\log PHPortugal_t - 1.3221\log CTerrestre\ Alemanha_t$$
$$- 2.1508\, S_2(B)\log CTAereoAlemanha_t + \hat{\mu}_{1t} \qquad (10)$$

$$S_2(B)\log TurAlemanha_t = 0.0141\, S_1(B)\log \mathrm{Re}ndAlemanha_t + 4.334\, S_2(B)\log CVAlemanha_t$$
$$- 1.0933\, S_1(B)\log PHPortugal_t + 1.59761\log CTerrestre\ Alemanha_t$$
$$- 0.2793\, S_2(B)\log CTAereoAlemanha_t + \hat{\mu}_{2t} \qquad (11)$$

ESPANHA :

$$S_2(B)\log TurEspanha_t = 0.6859\, S_1(B)\log \mathrm{Re}ndEspanha_t + 1.7582\, S_2(B)\log CVEspanha_t$$
$$- 0.1827\, S_1(B)\log PHPortugal_t - 1.0480\log CTerrestre\ Espanha_t + \hat{\mu}_{3t} \qquad (12)$$

FRANÇA :

$$S_2(B)\log TurFranca_t = 0.5013\, S_1(B)\log \mathrm{Re}ndFranca_t + 2.0973\, S_2(B)\log CVFranca_t$$
$$- 0.4223\, S_1(B)\log PHPortugal_t - 0.4466\log CTerrestre\ Franca_t$$
$$- 0.3366\log CTAereoFranca_t + \hat{\mu}_{4t} \qquad (13)$$

HOLANDA:

$$S_1(B)\log TurHolanda_t = 8.7857\, S_3(B)\log \mathrm{Re}ndHolanda_t + 2.7475\, S_2(B)\log CVHolanda_t$$
$$- 0.8409\, S_1(B)\log PHPortugal_t - 10.0594\log CTerrestre\ Holanda_t$$
$$- 0.6308\, S_2(B)\log CTAereoHolanda_t + \hat{\mu}_{5t} \qquad (14)$$

PORTUGAL :

$$S_1(B)\log TurPortuga\ l_t = 2.3294\, S_2(B)\log \mathrm{Re}\, ndPortugal_t - 0.0624\, S_1(B)\log CVPortugal_t$$
$$- 0.0927\, S_1(B)\log PHPortugal_t + \hat{\mu}_{6t} \quad (15)$$

REINO UNIDO :

$$S_2(B)\log TurRUnid_t = 0.4655\, S_1(B)\log \mathrm{Re}\, ndRUnido_t + 2.3293\, S_2(B)\log CVRUnido_t$$
$$- 0.3135\, S_1(B)\log PHPortugal_t - 0.4687\, S_1(B)\log CTAereoRUn\ ido_t + \hat{\mu}_{7t} \quad (16)$$

$$S_2(B)\log TurRUnido_t = 0.4109\, S_1(B)\log \mathrm{Re}\, ndRUnido_t + 2.0870\, S_2(B)\log CVRUnido_t$$
$$- 0.3871\, S_1(B)\log PHPortugal_t + 0.6848\, S_1(B)\log CTAereoRUn\ ido_t + \hat{\mu}_{8t} \quad (17)$$

Na presença de dois vectores de cointegração, como é o caso da Alemanha e Reino Unido, é normal que uma das equações apresente pelo menos um sinal contrário ao esperado. Nestes casos considerar-se-á o primeiro vector como sendo o correcto, obedecendo às leis económicas.

Todas as variáveis apresentam os sinais esperados em termos de significado económico, ou seja sinal positivo no caso da variável rendimento e sinal negativo para as restantes variáveis, com excepção da variável representativa do custo de vida em Portugal de todos os países representativos da procura externa. O facto de para todos os países[11], o coeficiente associado à variável representativa do custo de vida ser positivo significa que o país de origem não é considerado um país substituto no caso de o custo de vida no país de destino (neste caso Portugal) aumentar.

Os valores apresentados para a variável rendimento e sempre que o coeficiente seja superior a um, significa que o turismo é considerado um produto de luxo. Tal acontece com a Alemanha, Portugal e Holanda. Os valores apresentados para esta variável no caso de Espanha, França e Reino Unido levam a concluir que o turismo é antes considerado como uma necessidade e não um produto de luxo.

Destaque ainda para o peso dos custos de transporte quando comparados com os preços da hotelaria no caso da Alemanha, Espanha, Holanda (para os custos de transporte terrestre) e de forma menos acentuada Reino Unido. Relativamente à procura turística de França, não existem diferenças muito significativas entre os coeficientes associados ao custo de transporte terrestre e os custos hoteleiros. Os custos de transporte aéreo têm também um maior peso na procura turística quando comparados com o custo de transporte terrestre. À mesma conclusão se pode chegar no caso da Alemanha. Relativamente a Portugal é de destacar o peso da variável rendimento quando comparado com o peso das outras variáveis.

Apresentam-se de seguida as equações de curto prazo com os respectivos resíduos desfasados de cada uma das equações de longo prazo $\hat{\mu}_{it-1}$.

4.3 Resultados do MCE

No desenvolvimento destes modelos considerou-se inicialmente para cada país o quarto desfasamento fazendo uma abordagem do geral para o particular para todas as variáveis em estudo, mas apenas os termos significativos e com sinais "correctos" foram retidos. Foi

[11] Com excepção de Portugal cujo sinal negativo neste caso é considerado o correcto. Para Portugal a variável é apenas representada pelo IPC e para os restantes resulta do quociente entre este e o IPC de cada país de origem.

considerada a variável qualitativa já mencionada anteriormente. Considerou-se ainda a variável dependente desfasada e que representa os hábitos de persistência dos turistas no país de destino. Δ = (1-B) representa o filtro de primeiras diferenças [12]. $\hat{\mu}_{it-1}$ é o termo corrector de erro e como já se referiu corresponde ao resíduo desfasado das correspondentes equações de longo prazo. Mede o quanto se está longe do equilíbrio de longo prazo e o coeficiente associado deve ter sinal negativo. Os desvios-padrões encontram-se dentro de parêntesis. Sempre que os coeficientes não se mostraram significativos ou apresentaram sinais contrários ao esperado, retiraram-se dos modelos sendo novamente estimados sem as variáveis associadas. As estatísticas apresentadas são de forma respectiva o coeficiente de correlação corrigido, a estatística de Durbin-Watson, o erro padrão da regressão e o teste de correlação de Breusch-Godfrey.

ALEMANHA :

$$\Delta S_2(B)\log TurAlemanha_t = \underset{(0.0784)}{-0.1550}\Delta S_2(B)\log TurAlemanha_{t-1} + \tag{18}$$

$$\underset{(0.1020)}{0.6823}\Delta S_2(B)\log TurAlemanha_{t-2} - \underset{(0.0932)}{0.2631}\Delta S_2(B)\log TurAlemanha_{t-4} - \underset{(0.0118)}{0.0318}\hat{\mu}_{1t-1} - \underset{(0.0130)}{0.0505}\hat{\mu}_{2t-1}$$

$\overline{R}^2 = 0.5509$ DW = 2.3843 SER = 0.0937 BG - LM(4) = 2.6950

ESPANHA :

$$\Delta S_2(B)\log TurEspanha_t = \underset{(0.4972)}{-2.0335} - \underset{(0.0962)}{07413}\Delta S_2(B)\log TurEspanha_{t-1} +$$

$$\underset{(0.0656)}{0.1363}\Delta S_2(B)\log TurEspanha_{t-3} + \underset{(0.3095)}{0.7681}\Delta S_1(B)\log \mathrm{Re}\, ndEspanha_t -$$

$$\underset{(0.4938)}{-1.1919}\Delta\log CTerrestreEspanha_t + \underset{(0.1061)}{0.2050}VQ1 - \underset{(0.1325)}{0.5476}\hat{\mu}_{3t-1} \tag{19}$$

$\overline{R}^2 = 0.8671$ DW = 1.9107 SER = 0.2046 BG - LM(3) = 1.4793

FRANÇA :

$$\Delta S_2(B)\log TurFranca_t = \underset{(0.0844)}{-0.3943}\Delta S_2(B)\log TurFranca_{t-1} +$$

$$\underset{(0.1019)}{0.6966}\Delta S_2(B)\log TurFranca_{t-2} - \underset{(0.0828)}{0.2788}\Delta S_2(B)\log TurFranca_{t-4} - \underset{(0.0255)}{0.0590}\hat{\mu}_{4t-1} \tag{20}$$

$\overline{R}^2 = 0.7884$ DW = 1.9883 SER = 0.0924 BG - LM(4) = 1.5506

[12] Necessária uma vez que todas as séries são nesta fase I(1,0,0).

HOLANDA:

$$\Delta S_1(B)\log TurHolanda_t = \underset{(0.0725)}{0.7211}\Delta S_1(B)\log TurHolanda_{t-1} -$$

$$\underset{(0.0708)}{0.2416}\Delta S_1(B)\log TurHolanda_{t-4} + \underset{(0.2279)}{0.7742}\Delta S_2(B)\log CVHolanda_{t-3} - \underset{(0.4938)}{0.4066}\Delta S_1(B)\log PHPortugal_{t-4}$$

$$\underset{(0.0007)}{-0.0011}\hat{\mu}_{5t-1} \qquad (21)$$

$\bar{R}^2 = 0.5921$ DW = 1.9557 SER = 0.1090 BG - LM(4) = 1.8968

PORTUGAL:

$$\Delta S_1(B)\log TurPortugal_t = \underset{(1.0154)}{3.4732} + \underset{(0.0953)}{0.3398}\Delta S_1(B)\log TurPortugal_{t-1} +$$

$$\underset{(0.0944)}{0.2374}\Delta S_1(B)\log TurPortugal_{t-2} + \underset{(0.1425)}{0.3232}\Delta S_2(B)\log \mathrm{Re}ndPortugal_t - \underset{(0.0294)}{0.1006}\hat{\mu}_{6t-1} \qquad (22)$$

$\bar{R}^2 = 0.4236$ DW = 1.9854 SER = 0.0475 BG - LM(2) = 0.3471

REINO UNIDO:

$$\Delta S_2(B)\log TurRUnido_t = \underset{(0.0822)}{-0.3774}\Delta S_2(B)\log TurRUnido_{t-1} +$$

$$\underset{(0.0856)}{0.3652}\Delta S_2(B)\log TurRUnido_{t-2} - \underset{(0.0285)}{0.0895}\hat{\mu}_{7t-1} - \underset{(0.0155)}{0.0871}\hat{\mu}_{8t-1} \qquad (23)$$

$\bar{R}^2 = 0.6218$ DW = 1.9716 SER = 0.0686 BG - LM(2) = 0.4421

Todos os termos correctores de erro apresentam sinais negativos como esperado, sendo também todos significativos, com excepção da Holanda cujo desvio-padrão se mostrou um pouco elevado. Destaque para a variável *dummy* representativa da EXPO98 e ainda da variável representativa dos custos de transporte terrestre no caso de Espanha. As variáveis que representam o custo de vida em Portugal são de realçar relativamente à procura turística holandesa. No caso da procura turística portuguesa, novamente se destaca a variável rendimento. No entanto, a maioria das variáveis que afectam a procura no longo prazo não a afectam no curto prazo.

Apresenta-se agora o modelo estrutural básico e sua aplicação à procura turística em Portugal.

5. MODELO ESTRUTURAL BÁSICO – APLICAÇÃO À PROCURA TURÍSTICA EM PORTUGAL

Este modelo pode ser escrito da seguinte forma[13]:

$$y_t = \mu_t + S_t + x'_t\,\delta + \varepsilon_t \qquad t = 1,\ldots\ldots T \qquad (23)$$

onde:

yt – variável dependente
μt – tendência temporal
St – variáveis sazonais
xt – vector de variáveis independentes
δ - vector de parâmetros desconhecidos
εt – erro do processo

As componentes sazonais *St* e a tendência temporal μt, são dadas por:

$$D_{jt} = \begin{cases} 1, & t = j, j+s, j+2s,... \\ 0, & t \neq j, j+s, j+2s,... \\ 1, & t = s, 2s, 3s,... \end{cases} \qquad j = 1,....,s-1$$

Neste caso s = 4 (trimestres).

Os resultados destes modelos apresentam-se a seguir (Tabela 3). D_1, D_2, D_3 são as componentes sazonais e D_4 corresponde à tendência temporal. As restantes variáveis são as já apresentadas anteriormente. Os desvios-padrões encontram-se dentro de parêntesis. Também aqui sempre que os coeficientes não eram significativos ou apresentavam sinais contrários ao esperado, retiraram-se dos modelos e foram novamente estimados sem as variáveis associadas.

Tabela 3 – Estimativas do Modelo Estrutural Básico

Variável Explicativa (Log)	País de Origem					
	Alemanha	Espanha	França	Holanda	Portugal	R.Unido
Dependente (Desfasada)	0.6571 (0.0742)	-		0.7241 (0.0687)	0.4405 (0.0801)	0.6918 (0.0754)
Rendimento	0.7917 (0.1731)	0.9662 (0.0373)	1.0145 (0.1581)	0.4065 (0.1318)	0.6473 (0.1038)	0.1026 (0.0594)
CVida	-	1.1528 (0.0617)	-	-	-	0.2582 (0.0754)
CTTerrestre	-0.3216 (0.0923)	-	-0.7855 (0.1543)	-	-	-
CTAéreo	-	-	-	-	-	-
PHPortugal	-	-	-	-	-0.0738 (0.0181)	-
VQ1	-	0.1723 (0.0853)	-	-	-	-
D1	-0.4852 (0.0524)	-	7.3481 (1.9217)	0.6925 (0.0612)	4.6516 (0.7107)	2.2204 (0.6686)
D2	0.3165 (0.0630)	0.5124 (0.0468)	8.2525 (1.9255)	1.1722 (0.0618)	4.9241 (0.7095)	2.7732 (0.6615)
D3	-	1.2342 (0.0468)	8.4523 (1.9215)	1.0494 (0.0379)	5.2565 (0.7285)	2.6253 (0.6919)
D4	-0.9069 (0.0358)	0.1582 (0.0470)	7.3721 (1.9322)	-	4.3717 (0.7604)	1.9661 (0.7033)
$\overline{R}^2$	0.9673	0.9301	0.9368	0.9479	0.9761	0.9720
SER	0.1247	0.1652	0.1396	0.1317	0.0543	0.0787
F	587.65	264.76	294.53	456.00	579.83	573.85

Fonte: Elaboração própria com base nos resultados obtidos pelo programa informático *Eviews (6.0).*

[13] Ouerfelli (2008).

Neste caso, pode-se verificar a importância e significância de praticamente todas as variáveis representativas da componente sazonal e tendência temporal. A variável dependente desfasada representativa do "*Word of mouth effect*" é de realçar com excepção de Espanha e França. A variável rendimento é significativa para todos os países e com maior peso na variável dependente quando comparada com as restantes variáveis explicativas no caso da Alemanha, França e Portugal. A variável *dummy* mais uma vez apenas se mostrou significativa no caso de Espanha. Também aqui a variável representativa dos preços da hotelaria não se desatacou a não ser no caso de Portugal, mas mesmo assim com um peso muito inferior quando comparada com as restantes variáveis.

Depois de apresentados os modelos analisa-se agora a precisão de previsão[14] dos mesmos.

6. PRECISÃO DAS PREVISÕES DOS MODELOS

Para esta análise consideram-se as medidas de precisão MAE, MSE e RMSE aplicadas ao período de 2004:01 a 2006:12 cujos resultados se encontram na Tabela 4.

Tabela 4 – Precisão das Previsões dos Modelos MEB e MCE

Países / Medidas / Modelos	MAE	MSE	RMSE
Alemanha			
MEB	0.05923	0.00569	0.07547
MCE	*0.04106*	*0.00226*	*0.04757*
Espanha			
MEB	*0.20996*	*0.06675*	*0.25837*
MCE	0.31599	0.14984	0.38709
França			
MEB	0.27338	0.08192	0.28621
MCE	*0.06928*	*0.00725*	*0.08517*
Holanda			
MEB	*0.06259*	*0.00509*	*0.07136*
MCE	0.06438	0.00536	0.07320
Portugal			
MEB	*0.04395*	*0.00284*	*0.05332*
MCE	0.07017	0.00611	0.07818
Reino Unido			
MEB	0.07608	0.00765	0.08747
MCE	*0.03909*	*0.00193*	*0.04399*

Fonte: Elaboração própria com base nos resultados obtidos pelo programa informático *Eviews (6.0)* e *Excel (2007).*

Pela análise da tabela pode-se concluir que, seja qual for a medida de precisão utilizada, o modelo MCE revelou-se o melhor modelo de previsão para prever a procura turística da Alemanha, França, e Reino Unido. No caso da Holanda apesar do modelo estrutural apresentar valores inferiores, as diferenças entre este modelo e o modelo MCE não são significativas. O modelo MEB mostrou ser o melhor modelo de previsão no caso de Espanha e Portugal.

7. CONCLUSÃO

[14] Previsão estática – "*1 period ahead*".

Neste artigo apresentaram-se modelos causais de previsão aplicados à procura turística em Portugal cuja principal vantagem quando comparados com os modelos univariados é o designado "*What if Forecasting*".

Relativamente à análise da cointegração e das correspondentes equações de longo prazo podem-se retirar as seguintes principais conclusões: o turismo é considerado um produto de luxo para a Alemanha, Portugal e Holanda. No caso de Espanha, França e Reino Unido o turismo é antes considerado como uma necessidade e não um produto de luxo.

Destaque ainda para o peso dos custos de transporte quando comparados com os preços da hotelaria no caso da Alemanha, Espanha, Holanda (para os custos de transporte terrestre) e de forma menos acentuada Reino Unido. Os custos de transporte aéreo têm também um maior peso na procura turística quando comparados com o custo de transporte terrestre no caso da França e Alemanha. No caso de Portugal é de destacar o peso da variável rendimento quando comparado com o das outras variáveis.

Relativamente aos modelos de curto prazo (MCE), todos os termos correctores de erro apresentaram sinais negativos como esperado, sendo também todos significativos, com excepção da Holanda cujo desvio-padrão se mostrou um pouco elevado. Destaque para a variável *dummy* representativa da EXPO98 e ainda dos custos de transporte terrestre no caso de Espanha. As variáveis representativas do custo de vida em Portugal são de realçar relativamente à procura turística holandesa. No caso da procura turística portuguesa, novamente se destaca a variável rendimento. No entanto, a maioria das variáveis que afectam a procura no longo prazo não a afectam no curto prazo.

Os modelos MEB permitiram ainda concluir pelo peso e significância de praticamente todas as variáveis representativas das componentes sazonal e tendência temporal. A variável dependente desfasada representativa do "*Word of mouth effect*" é de realçar com excepção de Espanha e França. A variável rendimento é significativa para todos os países e com maior peso na variável dependente quando comparada com as restantes variáveis explicativas no caso da Alemanha, França e Portugal. A variável *dummy* mais uma vez apenas se mostrou significativa no caso de Espanha. Também aqui a variável representativa dos preços da hotelaria não se desatacou a não ser no caso de Portugal, mas mesmo assim com um peso muito inferior quando comparada com as restantes variáveis.

Finalmente e depois de analisada a performance de previsão dos modelos foi possível concluir que seja qual tenha sido a medida de precisão utilizada, o modelo MCE revelou-se melhor modelo de previsão para prever a procura turística da Alemanha, França, e Reino Unido. No caso da Holanda apesar do modelo estrutural apresentar valores inferiores, as diferenças entre este modelo e o modelo MCE não se mostraram significativas. O modelo MEB mostrou ser melhor modelo de previsão no caso de Espanha e Portugal.

BIBLIOGRAFIA

BANCO DE PORTUGAL (2007) *Boletim Económico de Verão de 2007 – Séries Trimestrais para a Economia Portuguesa* 1977-2006, Lisboa, Banco de Portugal.

ENGLE, R.F. e C.W.J. GRANGER (1987) Co-Integration and Error Correction: Representation, Estimation, and Testing*, Econometrica*, 55, 251-276.

ENGLE, R.F., C.W.J. GRANGER e J.J. HALLMAN (1989) Merging Short and Long-run Forecasts: An Application of Seasonal Cointegration to Monthly Electricity Sales Forecasting, *Journal of Econometrics*, 40, 45-62.

GRANGER, C.W.J. (1981) Some Properties of Time Series Data and their use in Econometric Model Specification, *Journal of Econometrics*, 16, 121-130.

GRANGER, C.W.J. e A.A. WEISS (1983) *Time Series Analysis of Error-Correction Models*, New York, Academic Press.

HYLLEBERG, S., E.F. ENGLE, C.W.J. GRANGER e B.S. YOO (1990) Seasonal Integration and Cointegration, *Journal of Econometrics*, 44, 215-238.

JOHANSEN, S. (1988) Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control*, 12, 231-254.

JOHANSEN, S. e K. JUSELIUS (1990) Maximum Likelihood Estimation and inference on Cointegration with Applications to the Demand of Money, *Oxford Bulletin of Economics and Statistics*, 52, 169-210.

KULENDRAN, N. (1996) Modelling Quarterly Tourist Flows to Australia Using Cointegration Analysis, *Tourism Economics*, 2, 3, 203-222.

KULENDRAN, N. e S.F. WITT (2001) Cointegration Versus Least Squares Regression, *Annals of Tourism Research*, 28, 2, 291-311.

KULENDRAN, N. e K.F. WONG (2005) Modelling Seasonality in Tourism Forecasting, *Journal of Travel Research*, 44, 163-170.

OUERFELLI, C. (2008) Co-integration analysis of Quarterly European Tourism Demand in Tunisia, *Tourism Management*, 29, 127-137.

PEIXE, F.P.M. (1998) A Procura de Moeda em Portugal segundo a Abordagem de Johansen, *Estudos de Economia*, XVIII, 407-422.

WITT, S.F. e C.A. WITT (1992) *Modeling and Forecasting Demand in Tourism*, San Diego, Academic Press Limited.

WITT, S.F. e C.A. WITT (1995) Forecasting Tourism Demand: A Review of Empirical Research, *International Journal of Forecasting*, 11, 447-475.

TRENDS IN FOREIGN EXCHANGE MARKETS: AN ANALYSIS BASED ON TAYLOR´S METHODOLOGY

ADRIÁN FERNÁNDEZ-PÉREZ
E-mail: adrian.fernandez102@doctorandos.ulpgc.es
FERNANDO FERNÁNDEZ-RODRÍGUEZ
E-mail: ffernandez@dmc.ulpgc.es
Departmento de Métodos Cuantitativos
Facultad de Ciencias Económicas y Empresariales
Universidad de Las Palmas de Gran Canaria
SIMÓN SOSVILLA-RIVERO
E-mail: sosvilla@ccee.ucm.es
Departmento de Economía Cuantitativa
Facultad de Ciencias Económicas y Empresariales
Universidad Complutense de Madrid

Resumen

Este artículo se dedica a probar la hipótesis de camino aleatorio frente la existencia de tendencias en las series de tipos de cambio para 95 monedas frente al US Dollar. Con ese fin, hacemos uso del Modelo de Tendencia de Precios de Taylor [Taylor, S., 1980, Conjectured models for Trends in financial prices, tests and forecasts, Journal of the Royal Statistical Society, Series A, 143, 338-362] que, en lugar de centrarse en el comportamiento de reversión a la media de los tipos de cambio durante un largo horizonte, se centra en patrones a corto plazo de la evolución de los precios. Se utiliza un método de máxima verosimilitud para estimar los parámetros del modelo por medio de un algoritmo genético. Se derivan las predicciones óptimas de los rendimientos one-step-ahead. Se construye una regla de negociación basada en estas predicciones y se asimila a la popular regla de negociación basada en las medias móviles, superando a la estrategia buy-and-hold en 25 de 39 casos donde detectamos la tendencia.

Palabras Claves: Tipos de Cambio, Modelo de Tendencia de Precios, Algoritmos Genéticos, Reglas de Negociación
Clasificación JEL: C53, F31, G14.
Area Temática: Economía Monetaria y Financiera.

Abstract

This paper is devoted to testing for the random walk hypothesis against the existence of trends in exchange rate series for 95 currencies against the US Dollar. To that end, we make use of Taylor's price trend model [Taylor, S., 1980, Conjectured Models for Trends in Financial Prices, Tests and Forecasts, Journal of the Royal Statistical Society, Series A, 143, 338-362] that, instead of focusing on the mean reverting behaviour of exchange rates measured over a long horizon, focuses on the short-term pattern of the price trend. A maximum likelihood method is used to estimate the model parameters by mean of a genetic algorithm. Optimal one-step-ahead forecasts of returns are derived. The trading rule based on these forecasts is constructed and is found to bear similarity to a popular trading rule based on moving averages, overcoming the buy-and-hold strategy in 25 out of 39 cases where trends are detected.

Thematic Area: Monetary and Financial Economics.

1. INTRODUCTION

The efficient market hypothesis (EMH) is a theme long discussed in financial literature. In its weak form, the EMH establishes that current prices reflect all available public information in the past and investors are only compensated by taking risks. It means that the new information arriving on the market is instantaneously translated to prices and employing any technical

trading strategy it is impossible to obtain an abnormal profit above the market. In an alternative way, the defenders of technical analysis maintain that prices move following trends. It means that when new information arrives on the market it does not immediately translate into prices and a certain amount of time is necessary until the market incorporates this information. This situation will reflect that the market will move through trends which may be used in a profitable way using a technical strategy based on the correlations of past returns.

There was a seminal paper by Taylor (1980) casting doubt over the random walk hypothesis and introducing a price trend model which provided profitable rules in commodity and currency markets. Until the end of the eighties, the literature defended the EMH which supports that no technical trading rule may be able to make extra profits over the naïve buy and hold strategy, taking into account transaction costs. Nevertheless, recent studies reveal that there are situations where future returns are predictable from past returns. So, Lo and MacKinlay (1988) found positive autocorrelations of weekly returns on portfolios of NYSE stocks. Fama and French (1988) discovered negative serial correlation in returns of individual stocks and various portfolios of small and large firms. Brock, Lakonishok and LeBaron (1992) reported that most common technical trading rules as moving average and trading rank break have predictive power in the Dow Jones index. Similar conclusions have been reached by Gencay (1996) who found strong evidence of nonlinear predictability in daily returns of the Dow Jones index. Finally, Kwan et al. (2000) found predictability and profitability considering the price trend model by Taylor (1980) in the Hang Seng Index Futures in Hong Kong.

The paper is organised as follows. Section 2 presents the econometric methodology used for testing for the random walk hypothesis against the existence of trends in financial time series. Section 3 describes the data set and reports our empirical results. In Section 4, we assess the economic value of technical trading strategies based on the trend model. Finally, Section 5 provides some concluding remarks.

2. ECONOMETRIC METHODOLOGY: TAYLOR'S TREND MODEL

Following the weak EMH, the random walk model represents the movement of financial market asset returns

$$\begin{aligned} &x_t = \log(P_t) - \log(P_{t-1}) = \mu + \varepsilon_t, \\ &E(\varepsilon_t) = E(\varepsilon_t \varepsilon_{t+i}) = 0 \;\; i \neq 0 \end{aligned} \tag{1}$$

where P_t is the price of an asset in the instant t, μ is the expected change of the process, called the drift of the process, and the increments of daily returns $\{\varepsilon_t\}$ are IID with zero average.

In a seminal paper, Taylor (1980) introduces a trend model allowing μ to be variable with time being so a factor causing trends in prices, developing a statistical hypothesis framework to test whether the random walk models faithfully reflect the data generating process of the financial asset prices or, on the contrary, whether the prices have trends.

The trend model for a prices time series P_t is defined as

$$\begin{aligned} &x_t = \log(P_t) - \log(P_{t-1}) = \mu_t + \varepsilon_t, \\ &E(\varepsilon_t) = E(\varepsilon_t \varepsilon_{t+i}) = 0, \;\; i \neq 0, \operatorname{cov}(\mu_s, \varepsilon_t) = 0 \;\; \forall s,t \end{aligned} \tag{2}$$

where the drifts μ_t are uncorrelated with white noise series ε_t. In this case, μ_t is a stochastic process representing the trend in the model and it is interpreted as the answer to anticipated

changes in the supply and demand of the assets. This μ_t may be positive or negative giving rise to increasing or decreasing price trends

In what follows we call σ^2 to the variance of ε_t, v^2 to the variance of μ_t and $\overline{\mu}$ to the expectation of μ_t.

The trend models rests in five basic assumptions:

1) The trend values are determined by the actual information of supply and demand arriving on the market.
2) The new information arrives randomly on the market.
3) There is new information in the proportion of $1-p$ trading days, where $0 \le p \le 1$.
4) The trend values change only when the new information arriving on the market is available.
5) When the trend values change, the new value is independent of all past values.

So, the trend model may be formulated with probability as

$$\mu_t = \begin{cases} \mu_{t-1} & with\ probability \quad p \\ \overline{\mu} + \eta_t & with\ probability \quad 1-p \end{cases} \tag{3}$$

where η_t is a white noise with mean zero and independent of the past trend values μ_s for $s < t$.

In order to find out the number of days that the duration of the trend is expected, it is defined a parameter *m* which is called the mean trend duration. This parameter averages the different durations of possible trends

$$m = \sum_{i=1}^{\infty} i(1-p)p^{i-1} = (1-p)^{-1} \tag{4}$$

For instance, if *m* were equal to 5 days, we can say that, on average, the asset would move with the same trend μ_i, positive or negative, for 5 days until new information arrived to the market and the trend changed in μ_{i+6}.

The aforementioned trend model is not very realist because it is very well known that the variance of daily returns is time changing, that is, $\text{var}(x_t) = \Sigma_t^2$.

Furthermore, it is reasonable to assume that both $\text{var}(\varepsilon_t)$ and $\text{var}(\mu_t)$ are time depending quantities. So, as a time varying variance causes serious problems in obtaining the sample correlations, Taylor and Kingsman (1979), Taylor (1980) and Taylor (1986) developed a new methodology with the end of dealing with time varying variance. In this methodology it is necessary to introduce the additional assumption that the ratio $R = \text{var}(\mu_t)/\text{var}(\varepsilon_t)$ in (2) is roughly constant in the time. So, Taylor rescales the trend values in the way μ_t / Σ_t. In this case, denoting the average $E(\mu_t / \Sigma_t)$ as $\overline{\mu}$, we have a trend model with fluctuating variance,

$$\mu_t = \begin{cases} (\Sigma_t / \Sigma_{t-1})\mu_{t-1} & with\ probability \quad p \\ \overline{\mu}\Sigma_t + \eta_t \Sigma_t & with\ probability \quad 1-p \end{cases}, \tag{5}$$

From equation (5) it follows that the variance of daily returns are roughly constant, which facilitates the empirical implementation of the statistical tests.

In order to estimate Σ_t and given that its relation with the mean absolute deviation a_t is $a_t = E\left|X_t\right| = \Sigma_t$ multiplied by a constant, Taylor prefers to estimate a_t rather than Σ_t. Therefore, a_t is estimated using an exponential weighted moving average of the past absolute price changes:

$$\hat{a}_t = \gamma \sum_{i=0}^{\infty} (1-\gamma)^i \mid x_{t-1-i} \mid = (1-\gamma)\hat{a}_{t-1} + \gamma \mid x_{t-1} \mid, \tag{6}$$

concluding that the parameter γ, obtained by the maximum likelihood method, is equal to 0.04 for stock prices and indexes and 0.1 for the remainder series (exchange rates, commodities, etc.).

The base of the price trend test is the existence of positive correlations between daily returns with several lags. On the contrary, in the random walk model, all correlations will be zero for any lag. For technical reasons the correlation employed in the test are the correlations between the rescaled returns $x_t / \hat{a}_t$ from (6).

The correlations of daily rescaled returns are defined as $\rho_i = cor(x_t / \hat{a}_t, x_{t+1} / \hat{a}_{t+1})$. For the model (1) of random walk the autocorrelations are zero for all lags. On the contrary Taylor shows that the model (2), (3) and (5) of series trends provides the following correlation expression

$$\rho_i = \frac{p^i v^2}{v^2 + \sigma^2} = Ap^i, \tag{7}$$

where $A = v^2 / (v^2 + \sigma^2)$.

So Taylor (1980) formulates a hypothesis test where the null corresponds to the random walk:

$$H_0 : \rho_i = 0, \ \textit{for each } i>0 \tag{8}$$

while the alternative hypothesis to random walk model is:

$$H_1 : \rho_i = Ap^i, \ \textit{for some } A \geq 0, \ 0 \leq p \leq 1, \textit{ for each } i>0 \tag{9}$$

The parameter A is a measure of information that is not instantaneously reflected in the prices, meanwhile p measures the speed at which the information is reflected in prices. If both A and p were very close to zero, the information would be used perfectly by the market. But when the trend is accepted, A has a small value, around 3%, and p is close to 1. It means that the market has a slow interpretation of the relevant information that arrives. The additional hypothesis that the ratio $R = \text{var}(\mu_t) / \text{var}(\varepsilon_t)$ in (2) is almost constant is necessary in order to permit a fluctuating variance in the model. As aforementioned, in this case, the time varying problems are avoided using rescaled returns $y_t = x_t / \hat{a}_t$, where $\hat{a}_t$ is defined in (6). This y_t has a variance approximately constant.

Although the trend model is nonlinear by nature, its autocorrelations resemble to the *ARMA(1,1)* model

$$x_t - px_{t-1} = \xi_t - q\xi_{t-1}, \quad \xi_t \approx IID(0, \sigma_\xi^2) \tag{10}$$

because its autocorrelations has also the form $\rho_i = Ap^i$ when *q* verifies the equation

$$q^2 - q\left\{\frac{1+(1-2A)p^2}{(1-A)p}\right\}+1=0 \ for\ 0 \le q \le 1 \tag{11}$$

Therefore, as far as the autocorrelation functions are concerned, there exists a one to one correspondence between the class of price trend models and the *ARMA(1,1)* verifying (11). This does not mean that the two models are equivalent because their fourth or higher-order moments are different in general. Nevertheless, this correspondence may be used for forecasting purposes and forecasts of the future returns are generated under the price trend model by using the forecasts under the corresponding *ARMA(1,1)* model.

As Taylor (1980) observed, the previous tests employed in literature in order to refuse trends in time series are badly specified. The standard test used is the *Q*-test by Box-Pierce. This statistic doesn't offer any specific form to the alternative hypothesis. It has two serious shortcomings when prices have a trend as in (2) and (3). On the one hand, *Q* doesn't distinguish between positive and negative values of ρ_i, meanwhile Taylor's H_1 says that all ρ_i are positives. On the other hand, *Q* emphasis in the same way each one of the *k* first autocorrelation; on the contrary, in Taylor's H_1 a decreasing values of autocorrelations are expected.

In order to reject the presence of trends in the financial series Taylor (1980) proposes a statistic *T* based on the likelihood ratio, using the sample autocorrelations of rescaled returns $(r_1, r_2, ..., r_k)$ in (10).

$$T_{k,\phi} = \sum_{i=1}^{k} \phi^i r_i, \quad (0 < \phi < 1) \tag{12}$$

If H_0 is accepted, the statistic $T_{k,\phi}$ has *N(0,1)* asymptotic distribution. This statistic has only one tail, for which we reject the null hypothesis of random walk in favour of a trend with a significance level of 5% when T^* is higher than the critical value of 1.65. $T_{k,\phi}$ has the inconvenience that is not very robust in facing data errors in the price series. The first effect of a data error is the reducing of the first autocorrelation coefficient r_1 for which Taylor designed another statistic $U_{k,\phi}$ ignoring r_1:

$$U_{k,\phi} = \sum_{i=2}^{k} \phi^i r_i \ , (0 < \phi < 1) \tag{13}$$

$U_{k,\phi}$ is also normally *N(0,1)* asymptotically distributed. For both statistics it is necessary to choose *k*, ϕ and the significance level α. Taylor (1980) recommends using as better values *k = 30* and ϕ *= 0.92*. When H_0 is true the statistical U^* is

$$U^* = \frac{\sum_{i=2}^{30} 0.92^i r_i}{\sum_{i=2}^{30} (0.92^{2i} n^{-1})^{\frac{1}{2}}} = 0.4649\sqrt{n}\sum_{i=2}^{30} 0.92^i r_i \tag{14}$$

Once the trends were detected by the U^* statistic, the trend parameters *A*, *p, q* and *m* are going to be estimated in all series. As the parameter $A = v^2/(v^2 + \sigma^2)$, it is necessary to estimate the variances v^2 and σ^2 in (2).

In order to estimate the trend parameters it is possible to use several methods. So Taylor (1980) employed the generalized method of moments. On the other hand Kwan *et al.* (2000) used the quasi-maximum likelihood in order to estimate the trend parameters in daily returns for Hang Seng Index Futures.

In this paper we will employ the maximum likelihood method in estimating the trend parameters. Following Taylor we try to match the theoretical Ap^i and the observed r_i autocorrelations, assuming the differences between them is $N(0,\sigma_r^2)$, that is

$$\begin{aligned} &r_i = Ap^i + \varepsilon_i \,,\ \varepsilon_i \approx N(0,\sigma_i^2)\,,\ i = 1,...,n_r \\ &E(r_i) = Ap^i,\ \mathrm{var}(r_i) = \sigma_i^2, \end{aligned} \tag{15}$$

where n_r is the number of simple autocorrelations r_i and σ_i^2 is the variance of the sample autocorrelations which following Barlett (1946) is given by the expression

$$Var(r_i) = \sigma_i^2 \approx \frac{1}{n}\left(1 + 2\sum_{k=1}^{i-1} r_k^2\right)$$

and *n* is the sample size of the training period.

In carrying out estimations, 200 sample autocorrelations of the rescaled returns are employed. Assuming that the residues $\varepsilon_i = r_i - Ap^i$ are independent, the likelihood function of the n_r residuals are

$$L(A, p / r_1, r_2, ..., r_{n_r}) = \prod_{i=1}^{n_r} \frac{e^{-\frac{1}{2\sigma_i^2}(r_i - Ap^i)^2}}{\sqrt{2\pi\sigma_i^2}} = \frac{e^{-\frac{1}{2\sigma_i^2}\sum_{i=1}^{n_r}(r_i - Ap^i)^2}}{(2\pi\sigma_i^2)^{n_r/2}} \tag{16}$$

Due to the complexity of function (16), in order to estimate the parameters *A* and *p* by maximizing the likelihood function, a genetic algorithm is employed.

A genetic algorithm (GA, hereafter) is a class of optimization technique, based on principles of natural evolution developed by Holland (1975) which try to overcome problems of traditional optimization algorithms, such as an absence of continuity or differentiability of the loss function. A GA starts with a population of randomly generated solution candidates, which apply the principle of fitness to produce better approximations to optimal solution. Promising solutions, as represented by relatively better performing solutions, are selected and breeding them together through a process of binary recombination referred to as crossover inspired by Mendel's natural genetics. The objective of this process is to generate successive populations solutions that are better fitted to the optimization problem than the solutions from which they were created. Finally, random mutations are introduced in order to avoid local optima [see Dorsey and Mayer (1995) for the use of genetic algorithms for optimizing complex likelihood functions in econometrics. Also see Haupt and Haupt (2004) as a simple introduction to genetic algorithms].

3. DATA AND EMPIRICAL RESULTS

In this paper the study of the existence of trends is carried out using daily data of nominal exchange rates against the US dollar for 95 countries from 4/01/1993 to 8/08/2008[1] taking from Reuters´ EcoWin Pro.

In order to evaluate the capability of Taylor's price-trend model to exploit slight dependence among returns, it is necessary to subdivide each series into two parts: a training period and a prediction period. The training period is the first part of a series and inside it the parameters *A, p* and *q* are estimated. These parameters will be employed for trading in the predicting period which is the second part of the series. The training period used to test for random walk hypothesis against trend ranks from the beginning of the series recorded by EcoWin Pro until 31-12-2007. The prediction period ranks from 01-01-2007 until 27-09-2008. For the series where the trend is accepted the characteristic parameters of the trend model are estimated. Finally, in the series where the mean trend duration is longer than 2, predictions are carried out in the prediction period.

Given that the countries in our sample present different exchange rate regimes that could affect the existence of trends, we have use the "natural fine classification" of Reinhart and Rogoff (2004), updated until December 2007 by Ilzetzki, Reinhart and Rogoff (), to distinguish between a wide range of *de facto* regimes:

1. No separate legal tender
2. Pre announced peg or currency board arrangement
3. Pre announced horizontal band that is narrower than or equal to +/-2%
4. De facto peg
5. Pre announced crawling peg
6. Pre announced crawling band that is narrower than or equal to +/-2%
7. De factor crawling peg
8. De facto crawling band that is narrower than or equal to +/-2%
9. Pre announced crawling band that is wider than or equal to +/-2%
10. De facto crawling band that is narrower than or equal to +/-5%
11. Moving band that is narrower than or equal to +/-2% (i.e., allows for both appreciation and depreciation over time)
12. Managed floating
13. Freely floating
14. Freely falling
15. Dual market in which parallel market data is missing.

As the tables in Reinhart and Rogoff (2004) provide monthly data, we can not know the exact date of the change of regime. Therefore, we assume the regime change take place in the last day of the month.

Given that there are changes in the exchange rate regime in the 95 currencies examined in this paper, Table 1 provides an overview of the evolution of the statistical U* within each regime for each currency. This table reports for each individual currency and exchange rate regime, if at any time within that combination the statistic U* accepts the trend or not. To that end, we use the following codes:

0: the currency is not under a given exchange-rate regime

-1: the currency is under a given exchange-rate regime, but there is not evidence of the presence of a trend

+1: the currency is under a given exchange-rate regime and there is evidence of the presence of a trend at some point.

[1] This period differs between series depending on data availability.

Table 1: Summary of the evidence on the presence of exchange-rate trenes using the U* statistic (sample until 31 December 2007)

Currency\Regime	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Euro (from 1999)	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0
Algeria Dinar	0	0	0	0	0	0	0	-1	0	-1	0	-1	0	-1	0
Angola Adjusted Kwanza	0	0	0	1	0	0	0	0	0	0	0	0	-1	-1	0
Argentina Peso	0	-1	0	0	0	0	0	-1	0	0	0	0	0	-1	-1
Australian Dollar	0	-1	0	0	0	0	0	1	0	0	0	-1	-1	0	0
Bangladesh Taka	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
Barbados Dollar	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0
Belize Dollar	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0
Buthan Ngultrum	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Bolivia Boliviano	0	0	0	0	0	0	-1	-1	0	-1	0	-1	0	-1	0
Brazil Real	0	-1	0	0	0	-1	0	0	0	-1	0	1	0	-1	0
Brunei Darussalem Ringgit	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Burundi Franc	0	0	0	0	0	0	0	1	0	-1	0	0	0	-1	0
Cambodia Riel	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
Canada Dollar	0	0	0	0	0	0	0	1	0	-1	0	0	0	0	0
Cape Verde Escudo	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
Chile Peso	0	-1	0	0	-1	0	0	0	1	1	0	1	0	1	0
China Yuan Renmimbi	0	0	0	-1	0	0	0	-1	0	0	0	-1	0	0	0
Colombia Peso	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0
Congo Democratic Republic Franc	0	0	0	0	0	0	0	0	0	-1	0	-1	-1	-1	0
Costa Rica Colon	0	-1	0	0	0	0	1	1	0	-1	0	-1	0	-1	0
Dominican Republic Peso	0	0	0	0	0	0	0	-1	0	-1	0	-1	0	-1	0
Ecuador Sucre (until 2001)	0	-1	0	0	0	0	0	-1	0	-1	0	-1	0	1	-1
Egypt Pound	0	0	0	-1	0	0	0	0	0	-1	0	0	0	0	0
El Salvador Colon	-1	0	0	-1	0	0	0	-1	0	0	0	-1	0	0	0
Equatorial Guinea Epkwele	0	-1	0	0	0	0	0	-1	0	0	0	0	0	0	0
Ethiopia Birr	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
Fiji Dollar (USD per FD)	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0
Gambia Dalasi	0	1	0	0	0	0	0	1	0	0	0	0	-1	-1	-1
Ghana New Cedi	0	-1	0	0	0	0	0	1	0	0	0	-1	-1	-1	0
Guinea Franc	0	-1	0	0	0	0	-1	1	0	-1	0	-1	0	-1	1
Guinea-Bissau Escudo/Peso (until 1997)	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	-1
Guyana Dollar	0	0	0	0	0	0	-1	0	0	0	0	0	0	-1	0
Haiti Gourde	0	-1	0	0	0	0	0	-1	0	-1	0	1	-1	1	0
Honduras Lempira	0	-1	0	0	0	0	-1	0	0	-1	0	0	0	-1	0
Hong Kong Dollar	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
India Rupee	0	-1	0	-1	0	0	1	1	0	0	0	-1	0	0	0
Indonesia Rupiah	0	0	0	0	0	0	1	0	0	-1	0	1	0	1	0
Israel New Sequel	0	0	0	0	0	-1	0	1	-1	1	0	-1	0	-1	0
Jamaica Dollar	0	-1	0	-1	0	0	1	-1	0	-1	0	0	0	-1	0
Japan Yen	0	-1	0	0	0	0	0	1	0	0	0	-1	1	0	0
Jordan Dinar	0	-1	0	-1	0	0	-1	0	0	-1	0	0	0	-1	0
Kazakhstan Tenge	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0

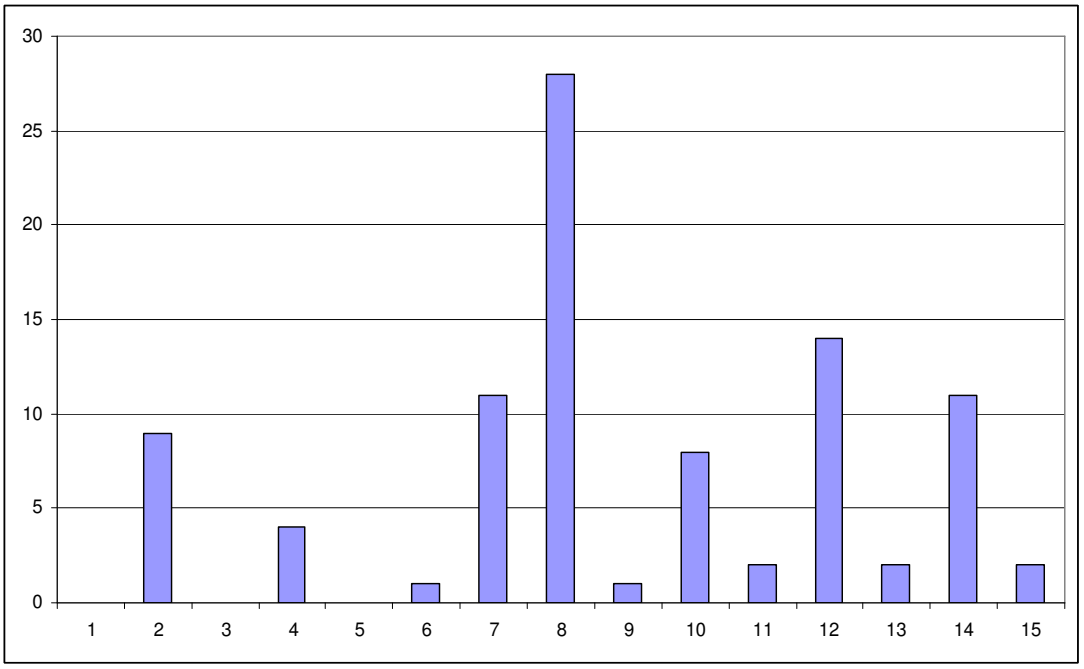
Table 1 (continued)

Currency\Regime	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Kenya Shilling	0	-1	0	0	0	0	0	1	0	0	0	-1	0	-1	0
South Korea Won	0	-1	0	0	0	-1	1	0	0	1	0	1	0	-1	0
Kuwait Dinar	0	-1	0	-1	0	0	-1	0	0	0	0	0	0	0	0
Kyrgyzstan Som	0	0	0	0	0	0	0	1	0	0	0	0	0	-1	0
Lebanon Pound	0	-1	0	0	0	0	-1	1	0	1	0	0	0	1	0
Leshoto Loti	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Madagascar Ariary	0	1	0	0	0	0	0	-1	0	0	0	1	0	-1	-1
Malawi Kwacha	0	1	0	0	0	0	-1	0	0	0	0	1	-1	-1	0
Malaysia Ringgit	0	1	0	0	0	0	0	1	0	0	0	0	-1	0	0
Maldive Islands Rufiyaa (until 1984)	0	-1	0	-1	0	0	0	0	0	0	0	0	0	0	0
Mauritania Ougiyaa	0	1	0	-1	0	0	-1	1	0	-1	0	0	0	0	0
Mauritus Rupee	0	0	0	0	0	0	0	1	0	-1	0	0	0	0	-1
Mexico New Peso	0	-1	0	-1	-1	0	-1	0	-1	0	0	-1	0	1	0
Moldova Leu	0	0	0	-1	0	0	0	1	0	0	0	0	0	1	-1
Mongolia Tugrik	0	0	0	-1	0	0	0	-1	0	0	0	0	-1	-1	0
Morocco Dirham	0	0	0	0	0	0	-1	-1	0	0	0	0	0	0	0
Mozambique New Metical	0	0	0	-1	0	0	0	-1	0	0	0	0	0	0	0
Mianmar (Burma) Kyat	0	-1	0	0	0	0	0	0	0	1	-1	-1	0	1	-1
Namibia Dollar	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0
Nepal Rupee	0	-1	0	0	0	0	0	-1	0	-1	0	0	0	0	0
New Zealand Dollar	0	-1	0	0	0	0	0	1	0	0	0	1	0	0	0
Nicaragua Cordoba Oro	0	-1	-1	0	0	0	-1	-1	0	0	0	0	-1	-1	0
Nigeria Naira	0	-1	0	0	0	0	0	0	0	0	0	1	-1	-1	0
Pakistan Rupee	0	-1	0	0	0	0	-1	-1	0	0	0	0	0	0	-1
Papua New Guinea Kina	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Paraguay Guarani	0	0	0	0	0	0	1	-1	0	1	0	-1	0	-1	0
Peru New Sol	0	0	0	0	0	0	0	1	0	0	0	0	0	-1	-1
Philippines Peso	0	0	0	-1	0	0	-1	1	0	-1	0	-1	0	-1	0
Qatar Ryal	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0
Sao Tome and Principe Dobra	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0
Saudi Arabia Rial	0	-1	0	1	0	0	0	0	0	0	0	0	0	0	0
Seychelles Rupee	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Sierra Leone Leone	0	0	0	-1	0	0	0	0	0	0	0	0	-1	0	0
Singapore Dollar	0	-1	0	0	0	0	0	0	0	0	1	0	0	0	0
South Africa Rand	0	-1	0	0	0	0	0	0	0	0	0	1	1	0	1
Sri Lanka Rupee	0	-1	0	0	0	-1	1	-1	-1	1	0	-1	0	0	0
Sudan Pound	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
Suriname Dollar	0	-1	0	0	0	0	0	-1	0	-1	0	-1	0	-1	0
Swaziland Lilangeni	-1	-1	0	1	0	0	0	0	0	0	0	0	0	0	0
Syria Pound	0	0	0	0	0	0	0	-1	0	-1	0	-1	0	0	0
Tajikistan Somoni	0	-1	0	0	0	0	1	0	0	0	0	0	0	-1	0
Tanzania Shilling	0	-1	0	-1	0	0	0	1	0	-1	0	-1	0	1	0
Thailand Baht	0	-1	0	1	0	0	0	0	0	0	1	-1	0	-1	0
Tonga Pa'anga	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0
Trinidad and Tobago Dollar	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
Tunisia Dinar	0	-1	0	0	0	0	0	1	0	0	0	0	0	0	0
United Arab Emirates Dirham	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0
British Pound	0	1	-1	0	0	0	0	0	0	0	-1	1	0	0	0
Uruguay Peso	0	0	0	0	-1	-1	0	1	0	0	0	1	0	-1	0
Venezuela Bolivar Fuerte	0	-1	0	0	0	1	0	0	0	0	0	1	0	-1	-1
Viet Nam Dong	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Zambia Kwacha	0	-1	0	0	0	0	0	0	0	0	0	-1	-1	1	0

For example, for the Australian Dollar we observe that it was never under regime 1 (no separate legal tender); it was under regime 2 (pre announced peg or currency board arrangement) in some subperiods, but we could not accept the presence of a trend; and it was under regime 8 (*de facto* crawling band that is narrower than or equal to +/-2%) in some subperiods and we find evidence of a trend.

As can be seen, we find evidence of the presence of a trend in the exchange rate in 93 cases. As one could expected, these episodes are more frequent the more flexible the exchange-rate is (Figure 1). In addition, we find that the existence of a trend is generally accepted most frequently for currencies of developed countries and less frequently for currencies of developing countries. A reason for this finding could be that the latter have more efficient markets or because the former are more likely to have less flexible exchange-rate regimes.

Figure 1: Detected exchange-rate trends and exchange-rate regimes



Additionally, we repeated the study using all the available data, therefore ending in 27 September 2008. Given the absence of information in Reinhart and Rogoff about the exchange-rate regimen in 2008, we assume that the countries maintain the same exchange-rate that was previously effective. The results for this experiment are presented in Table 2. As can be observed, several countries that did not accept the trend with data until December 2007 for their last known exchange rate regime, do accept the trend if we extend the sample until 27 September 2008. It is interesting to note that among these countries we find the Euro area and the United Kingdom.

Table 2: Summary of the evidence on the presence of exchange-rate trends using the U* statistic (sample until 27 September 2008)

Currences	last regime	data until 12/31/2007	data until 09/27/2008
Euro (from 1999)	14	-1	1
Algeria Dinar	8	-1	-1
Angola Adjusted Kwanza	4	1	1
Argentina Peso	8	-1	1
Australian Dollar	13	1	1
Bangladesh Taka	7	-1	-1
Barbados Dollar	2	-1	-1
Belize Dollar	2	-1	-1
Buthan Ngultrum	2	1	1
Bolivia Boliviano	7	-1	-1
Brazil Real	12	1	1
Brunei Darussalem Ringgit	8	1	1
Burundi Franc	8	1	1
Cambodia Riel	7	-1	-1
Canada Dollar	10	-1	-1
Cape Verde Escudo	7	-1	-1
Chile Peso	10	1	1
China Yuan Renmimbi	4	-1	-1
Colombia Peso	10	1	1
Congo Democratic Republic Franc	13	-1	-1
Costa Rica Colon	7	1	1
Dominican Republic Peso	12	-1	-1
Ecuador Sucre (until 2001)	14	1	1
Egypt Pound	4	-1	-1
El Salvador Colon	1	-1	-1
Equatorial Guinea Epkwele	2	-1	-1
Ethiopia Birr	7	-1	-1
Fiji Dollar	10	-1	-1
Gambia Dalasi	8	1	1
Ghana New Cedi	8	1	-1
Guinea Franc	10	-1	-1
Guinea-Bissau Escudo/Peso (until 1997)	15	-1	-1
Guyana Dollar	7	-1	-1
Haiti Gourde	12	1	1
Honduras Lempira	7	-1	-1
Hong Kong Dollar	2	1	1
India Rupee	8	1	1
Indonesia Rupiah	12	1	1
Israel New Sequel	10	1	1
Jamaica Dollar	7	1	-1
Japan Yen	13	1	1
Jordan Dinar	4	-1	-1
Kazakhstan Tenge	8	1	1
Kenya Shilling	8	1	1

Table 2 (continued)

Currences	last regime	data until 12/31/2007	data until 09/27/2008
South Korea Won	12	1	1
Kuwait Dinar	4	-1	1
Kyrgyzstan Som	8	1	1
Lebanon Pound	2	-1	-1
Leshoto Loti	2	1	1
Madagascar Ariary	12	1	1
Malawi Kwacha	7	-1	-1
Malaysia Ringgit	8	1	1
Maldive Islands Rufiyaa (until 1984)	4	-1	-1
Mauritania Ougiyaa	7	-1	-1
Mauritus Rupee	8	1	1
Mexico New Peso	12	-1	-1
Moldova Leu	8	1	1
Mongolia Tugrik	4	-1	1
Morocco Dirham	7	-1	1
Mozambique New Metical	8	-1	-1
Mianmar (Burma) Kyat	15	-1	-1
Namibia Dollar	2	-1	-1
Nepal Rupee	8	-1	1
New Zealand Dollar	12	1	1
Nicaragua Cordoba Oro	7	-1	-1
Nigeria Naira	12	1	1
Pakistan Rupee	7	-1	-1
Papua New Guinea Kina	7	1	1
Paraguay Guarani	10	1	1
Peru New Sol	8	1	1
Philippines Peso	8	1	1
Qatar Ryal	2	-1	-1
Sao Tome and Principe Dobra	10	-1	-1
Saudi Arabia Rial	4	1	1
Seychelles Rupee	8	1	1
Sierra Leone Leone	4	-1	-1
Singapore Dollar	11	1	1
South Africa Rand	13	1	1
Sri Lanka Rupee	7	1	1
Sudan Pound	7	-1	-1
Suriname Dollar	2	-1	-1
Swaziland Lilangeni	2	-1	-1
Syria Pound	10	-1	-1
Tajikistan Somoni	7	1	1
Tanzania Shilling	10	-1	-1
Thailand Baht	11	1	1
Tonga Pa'anga	8	-1	-1
Trinidad and Tobago Dollar	7	-1	-1
Tunisia Dinar	8	1	1
United Arab Emirates Dirham	2	-1	-1
British Pound	11	-1	1
Uruguay Peso	8	1	1
Venezuela Bolivar Fuerte	15	-1	-1
Viet Nam Dong	7	1	1
Zambia Kwacha	13	-1	-1

Table 3 reports the results of the U^* test as well as other important parameters in the trend model as it is the probability p of maintaining the trend, the parameter A of the correlation function in (7) and the mean trend duration obtained for the last known regime for each currency. As mentioned, all parameters were obtained by maximum likelihood employing a GA.

As a general comment, it is possible to observe in Table 3 that the series where the statistic U^* accepts the trend predominate values of A which are less than the values corresponding to the series where U^* accepts the null of random walk. The parameter p is usually higher than 0.5 in the series where the trend is accepted, which means that the new information needs more than one day to be incorporated into the prices. Note that for the series where the trend is not accepted we have not estimated the parameters *A, p, q* and *m*, so we have filled with zeros the corresponding columns.

With respect to U^* statistic, the results shown in Table 3 point out the following conclusions:

- In 56 out of a total of 95 exchange rates cases the U^* statistic accepts the null hypothesis of random walk ($U^* < 1.65$, in a one-tail N(0,1) test with 5% of confidence)
- Trends are detected in 39 out of 95 exchange rates ($U^* > 1.65$), being trends more frequent in intermediate exchange-rate regimes (Figure 2).
- The mean trend duration is always higher than one day ($m>1$) when a trend is detected.

Figure 2: Detected exchange-rate trends and exchange-rate regimes

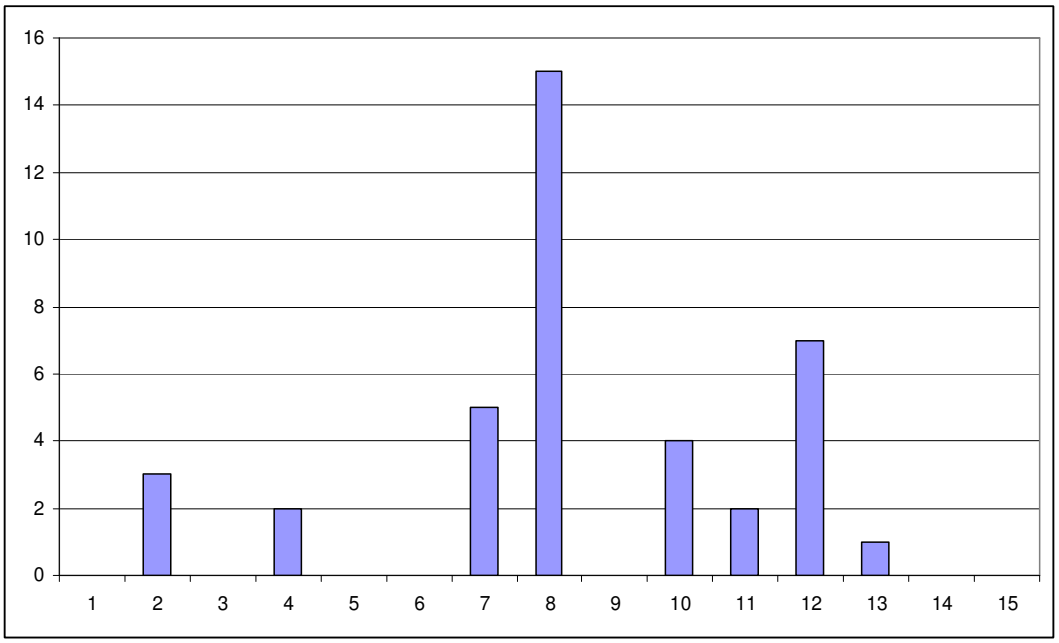


Table 3: Taylor's statistics and trend parameters

Currencies	Last regime	Initial date	Final date	U*	A	p	m
Euro	14	11062007	12312007	-0.3778	0	0	0
Algeria Dinar	8	11062007	12312007	-0.5152	0	0	0
Angola Adjusted Kwanza	4	7202005	12312007	1.6615	0.1216	0.7072	3
Argentina Peso	8	11062007	12312007	-1.1554	0	0	0
Australian Dollar	13	11062007	12312007	-0.2062	0	0	0
Bangladesh Taka	7	11062007	12312007	-0.4555	0	0	0
Barbados Dollar	2	11062007	12312007	0.0000	0	0	0
Belize Dollar	2	11062007	12312007	0.0000	0	0	0
Buthan Ngultrum	2	8112004	12312007	2.2324914	0.02638808	0.95237738	20.9984245
Bolivia Boliviano	7	11062007	12312007	0.0000	0	0	0
Brazil Real	12	10011999	12312007	2.1350	0.0134	0.9858	70
Brunei Darussalem Ringgit	8	2011990	12312007	4.3782	0.0134	0.9911	113
Burundi Franc	8	3242006	12312007	2.568965893	0.0560163	0.90872455	10.9558484
Cambodia Riel	7	11062007	12312007	0.2519	0	0	0
Canada Dollar	10	11062007	12312007	-0.7328	0	0	0
Cape Verde Escudo	7	11062007	12312007	-0.5032	0	0	0
Chile Peso	10	2012002	12312007	2.9770	0.0489	0.8809	8
China Yuan Renmimbi	4	11062007	12312007	-0.3220	0	0	0
Colombia Peso	10	1021985	12312007	23.6481	0.0607	0.9979	465
Congo Democratic Republic Franc	13	10292007	12312007	0.0611	0	0	0
Costa Rica Colon	7	2012002	12312007	2.0314	0.0138	0.9980	512
Dominican Republic Peso	12	11062007	12312007	0.0391	0	0	0
Ecuador Sucre	14	9102001	11022001	0.0000	0	0	0
Egypt Pound	4	11222007	12312007	-0.5040	0	0	0
El Salvador Colon	1	11062007	12312007	0.0000	0	0	0
Equatorial Guinea Epkwele	2	9171986	11111986	-0.6719	0	0	0
Ethiopia Birr	7	11062007	12312007	0.0000	0	0	0
Fiji Dollar (USD per FD)	10	11062007	12312007	-0.5813	0	0	0
Gambia Dalasi	8	5112007	12312007	1.6774	0.0464	0.9389	16
Ghana New Cedi	8	6042007	12312007	4.170444834	0.10777243	0.95284192	21.2052717
Guinea Franc	10	11062007	12312007	0.0588	0	0	0
Guinea-Bissau Escudo/Peso	15	11062007	12312007	0.0000	0	0	0
Guyana Dollar	7	11062007	12312007	0.3133	0	0	0
Haiti Gourde	12	4052007	12312007	1.705477566	0.04337436	0.94006856	16.6857328
Honduras Lempira	7	11062007	12312007	-0.0925	0	0	0
Hong Kong Dollar	2	4162007	12312007	1.7206	0.0721	0.8715	8
India Rupee	8	2012005	12312007	2.470463417	0.02465232	0.95318002	21.3584013
Indonesia Rupiah	12	5031999	12312007	3.3993	0.0527	0.8815	8
Israel New Sequel	10	3011991	12312007	3.4449	0.0223	0.9198	12
Jamaica Dollar	7	6142007	12312007	1.7064	0.2987	0.6303	3
Japan Yen	13	1031978	12312007	5.0840	0.0221	0.9373	16
Jordan Dinar	4	11062007	12312007	-0.6463	0	0	0
Kazakhstan Tenge	8	8012005	12312007	3.6718	0.0663	0.9390	16
Kenya Shilling	8	2011996	12312007	3.9789	0.1732	0.5176	2

Notes:

a. The training period used in the calculations spans from that indicated in the column "initial date" to that in the "final date". The prediction period spans from the day after that indicated in the column "final date" to 27 September 2008

b. The parameters A and p [m=1/(1-p)] were obtained through maximizing the logarithm of likelihood function by a genetic algorithm.

c. In blue, the U^* statistic rejects the null in favour of trend at the 5% confidence level.

Table 3 (continued)

Currencies	Last regime	Initial date	Final date	U*	A	p	m
South Korea Won	12	8031998	12312007	2.7173	0.0353	0.8536	7
Kuwait Dinar	4	11222007	12312007	-0.6618	0	0	0
Kyrgyzstan Som	8	8072007	12312007	2.5968	0.0736	0.9713	35
Lebanon Pound	2	11062007	12312007	0.0205	0	0	0
Leshoto Loti	2	5082000	12312007	1.7028	0.0106	0.9883	85
Madagascar Ariary	12	2021999	12312007	2.0820	0.0080	0.9738	38
Malawi Kwacha	7	11062007	12312007	-0.2968	0	0	0
Malaysia Ringgit	8	11062007	12312007	-1.0430	0	0	0
Maldive Islands Rufiyaa	4	11062007	12312007	-0.6685	0	0	0
Mauritania Ougiyaa	7	11062007	12312007	-0.2448	0	0	0
Mauritus Rupee	8	4082002	12312007	3.2323	0.0090	0.9966	290
Mexico New Peso	12	11062007	12312007	-0.1758	0	0	0
Moldova Leu	8	4032000	12312007	13.4773	0.1224	0.9445	18
Mongolia Tugrik	4	11062007	12312007	-0.1417	0	0	0
Morocco Dirham	7	11062007	12312007	-0.4420	0	0	0
Mozambique New Metical	8	11022007	12312007	-0.5163	0	0	0
Mianmar (Burma) Kyat	15	11062007	12312007	0.0000	0	0	0
Namibia Dollar	2	11062007	12312007	-1.1625	0	0	0
Nepal Rupee	8	11062007	12312007	-0.2180	0	0	0
New Zealand Dollar	12	3221994	12312007	1.6545	0.0122	0.9821	56
Nicaragua Cordoba Oro	7	11062007	12312007	0.0430	0	0	0
Nigeria Naira	12	10031996	12312007	1.7528	0.0003	0.9980	494
Pakistan Rupee	7	11062007	12312007	0.1947	0	0	0
Papua New Guinea Kina	7	2011990	12312007	1.8081	0.0070	0.9735	38
Paraguay Guarani	10	4142005	12312007	1.6760	0.1127	0.7389	4
Peru New Sol	8	12011993	12312007	3.9348	0.0160	0.9936	156
Philippines Peso	8	1032000	12312007	3.6923	0.0172	0.9933	150
Qatar Ryal	2	11222007	12312007	-0.7693	0	0	0
Sao Tome and Principe Dobra	10	11062007	12312007	-0.3100	0	0	0
Saudi Arabia Rial	4	11171980	12312007	2.0420	0.0069	0.9887	89
Seychelles Rupee	8	12222006	12312007	1.6681	0.0664	0.8413	6
Sierra Leone Leone	4	11062007	12312007	-0.6095	0	0	0
Singapore Dollar	11	9171973	12312007	1.7808	0.0008	0.9860	71
South Africa Rand	13	4031995	12312007	2.3190	0.6210	0.1273	0
Sri Lanka Rupee	7	6012001	12312007	2.9840	0.9999	0.1601	0
Sudan Pound	7	11022007	12312007	-0.7243	0	0	0
Suriname Dollar	2	10292007	12312007	0.0000	0	0	0
Swaziland Lilangeni	2	11062007	12312007	-1.1661	0	0	0
Syria Pound	10	11062007	12312007	0.0000	0	0	0
Tajikistan Somoni	7	1272005	12312007	3.9395	0.0390	0.9531	21
Tanzania Shilling	10	11062007	12312007	0.0326	0	0	0
Thailand Baht	11	11011999	12312007	4.0850	0.0238	0.9653	29
Tonga Pa'anga	8	11062007	12312007	-1.2339	0	0	0
Trinidad and Tobago Dollar	7	11062007	12312007	-0.3101	0	0	0
Tunisia Dinar	8	3311977	12312007	2.4425	0.0027	0.9197	12
United Arab Emirates Dirham	2	10292007	12312007	0.3601	0	0	0
British Pound	11	11062007	12312007	-0.6628	0	0	0
Uruguay Peso	8	9252006	12312007	1.8378	0.1185	0.7975	5
Venezuela Bolivar Fuerte	15	10292007	12312007	0.0000	0	0	0
Viet Nam Dong	7	2012002	12312007	2.6170	0.0078	0.9904	105
Zambia Kwacha	13	11062007	12312007	-0.7285	0	0	0

Notes:

d. The training period used in the calculations spans from that indicated in the column "initial date" to that in the "final date". The prediction period spans from the day after that indicated in the column "final date" to 27 September 2008

e. The parameters A and p [m=1/(1-p)] were obtained through maximizing the logarithm of likelihood function by a genetic algorithm.

f. In blue, the *U** statistic rejects the null in favour of trend at the 5% confidence level.

5. ECONOMIC EVALUATION OF TRENDS: TAYLOR'S STRATEGY

Once the parameters associated with the trend model have been estimated, it is possible to construct technical trading strategies in order to beat the market. We will employ the strategy developed by (Taylor, 1986) aimed to profit from substantial trends in either direction. This strategy is compounded by three control parameters k_1, k_2 and k_t where $k_1 > k_2$. The parameter k_1 controls the commencement of trades, telling us when to change a short position for a long position. The parameter k_2 controls the conclusion of the trades, telling us when to change a long position for a short position.

Trading decisions depend on a standardized forecast k_t calculated by assuming the trend model, that is

$$k_t = \frac{f_{t-1,1}}{\hat{\sigma}_{F,t-1}} \tag{17}$$

where

$$f_{t-1,1} = (\hat{a}_t / \hat{a}_{t-1})\{(p-q)x_{t-1} + qf_{t-2,1}\} \tag{18}$$

$$\hat{\sigma}_{F,t-1} = \hat{a}_t \{Ap(p-q)/(1-pq)\}^{1/2} \tag{19}$$

with $t = 21,...,n_{rend}$, being n_{rend}, the total number of returns. In the recursion (18), $f_{t,1}$ is the *ARMA(1,1)* prediction made in the instant *t* of the return *t+1*, $\sigma_{F,t}$ is its standard deviation, x_t is the no rescaled return of the series in the instant *t* and $\hat{a}_t$ is the estimated mean absolute deviation obtained in (6) with $\gamma = 0.1$ for exchange rate series and $\gamma = 0.04$ for stock prices and indexes.

The Taylor strategy is as follows: we need 20 returns before the beginning in order to estimate the mean absolute deviations ($\hat{a}_t$). The values of $f_{t,1}$ and $\sigma_{F,t}$ are assumed to be zero for $t \leq 20$, and for $t \geq 21$ are estimated recurrently in (18) and (19). After $t \geq 21$, we begin with no market position until $k_t > k_1$ (start a long position) or $k_t < k_2$ (start a short position).

When we are inside the market, if we are in a long position we change to a short position when $k_t < k_2$; if we are in a short position we change to a long position when $k_t > k_1$. For $k_t \in [k_1, k_2]$ don't change the position in any case. When we change our position from long to short or vice versa, a transaction cost of 0.05% is subtracted from the total return. Besides, in order to compute total returns, we assume that, when we are in a short position, the proceeds are invested in a money market account with a risk-free rate of 4% per annum (a year of 252 days is assumed).

In order to select the control parameters k_1 and k_2 an optimization process is carried out. So, k_1 and k_2 are selected, maximizing the Sharpe ratio of the Taylor strategy in the training period. With that end a genetic algorithm is also employed.

Once the control parameters are estimated they are employed, together with the trend parameters (*A*, *p* y *q*) obtained in the training period, in the prediction period. The net return obtained in the period *t* to the series *i* is the following

$$R_i^t = \sum_{t=21}^{N_{rend}} (x_t \, buy_t) + \sum_{t=21}^{N_{rend}} \left[\left(x_t - riskf_i \right) sell_t \right] - c_i mov_t \tag{20}$$

where x_t is the no rescaled return, buy_t stands for a buy signal in the instant *t* (equal to 1 when we are in a long position and equal to 0 when we are in a short position or we take no market position), c_i is the transaction cost (0.05%), mov_t is the number of times that we change from a short to a long position and vice versa, $riskf_i$ is the risk-free return (4% per annum), and $sell_t$

stands for the sell signals (equal to -1 when we are in a short position and equal to 0 when we are in a long position or we take no market position).

In order to compare the mean net return of the Taylor strategy with the mean net return of the buy and hold strategy the Sharpe ratio is employed. It divides the net return by its standard deviation, which for the series *i* in the period *t* is defined as

$$Sharpe_i^t = \frac{R_i^t / N_{return}}{\sigma_{R_i^t}} \tag{21}$$

where N_{return} represents the number of returns considered in the period.

The buy and hold strategy returns are obtained by adding the returns of the series from the first to the last, and subtracting two transaction costs corresponding with a buy in the first return and a sale in the last return.

Table 4 reports the values of parameters q, k_1 and k_2 for the training period used in Table 3 and the returns, obtained in the prediction period (01-01-2008 until 27-09-2008), by both, the B&H strategy and Taylor's strategy whose parameters are obtained by means of a GA. The Sharpe ratio of both strategies is also reported.

Table 4: Parameters of Taylor's strategy and prediction performance statistics

Currencies	q	k1	k2	B&H	Sharpe B&H	Taylor	Sharpe Taylor
Euro	0	0	0	0.0081	0.0077	0	0
Algeria Dinar	0	0	0	-0.0826	-0.0856	0	0
Angola Adjusted Kwanza	0.6288	1.0973	-1.4009	-0.0010	-0.0183	-0.0007	-0.0123
Argentina Peso	0	0	0	-0.0412	-0.1350	0	0
Australian Dollar	0	0	0	-0.0218	-0.0168	0	0
Bangladesh Taka	0	0	0	0.0031	0.0306	0	0
Barbados Dollar	0	0	0	0.0040	0.0651	0	0
Belize Dollar	0	0	0	-0.0015	-0.0084	0	0
Buthan Ngultrum	0.9316	1.6548	-1.4455	0.1000	0.1710	0.0286	0.0503
Bolivia Boliviano	0	0	0	-0.0729	-0.1619	0	0
Brazil Real	0.9759	0.6231	-0.1461	-0.0930	-0.0689	0.1034	0.0896
Brunei Darussalem Ringgit	0.9822	1.2265	-0.0517	-0.0215	-0.0415	0.0169	0.0351
Burundi Franc	0.8666	0.7009	-0.1473	0.0581	0.0948	-0.0313	-0.0656
Cambodia Riel	0	0	0	0.0387	0.0587	0	0
Canada Dollar	0	0	0	0.0528	0.0488	0	0
Cape Verde Escudo	0	0	0	-0.0099	-0.0103	0	0
Chile Peso	0.8431	0.1757	-0.7009	0.0396	0.0271	0.0838	0.0619
China Yuan Renmimbi	0	0	0	-0.0670	-0.3056	0	0
Colombia Peso	0.9833	0.0500	-0.0252	-0.0610	-0.0371	0.1024	0.0650
Congo Democratic Republic Franc	0	0	0	0.0080	0.0181	0	0
Costa Rica Colon	0.9924	0.0653	-1.6908	0.1110	0.1989	0.1093	0.2023
Dominican Republic Peso	0	0	0	0.0504	0.1180	0	0
Ecuador Sucre	0	0	0	0.0000	0.0000	0	0
Egypt Pound	0	0	0	-0.0272	-0.0667	0	0
El Salvador Colon	0	0	0	-0.0007	-0.1559	0	0
Equatorial Guinea Epkwele	0	0	0	-0.0258	-0.0608	0	0
Ethiopia Birr	0	0	0	0.0564	0.1722	0	0
Fiji Dollar (USD per FD)	0	0	0	-0.0147	-0.0176	0	0
Gambia Dalasi	0.9044	1.2286	-0.3230	0.0461	0.0301	0.1944	0.1831
Ghana New Cedi	0.8884	1.1623	-1.6024	0.1150	0.3954	0.1053	0.3615
Guinea Franc	0	0	0	0.0588	0.0720	0	0
Guinea-Bissau Escudo/Peso	0	0	0	0.0000	0.0000	0	0
Guyana Dollar	0	0	0	0.0062	0.0143	0	0
Haiti Gourde	0.9075	0.6771	-0.9062	0.0708	0.1696	0.0482	0.1417
Honduras Lempira	0	0	0	0.0038	0.0105	0	0
Hong Kong Dollar	0.8182	1.9221	-0.8663	0.0000	0.0002	0.0144	0.2278
India Rupee	0.9336	0.4619	-0.0105	0.1010	0.1394	0.0681	0.0965
Indonesia Rupiah	0.8412	0.6069	-0.0118	-0.0263	-0.0608	0.0415	0.1058
Israel New Sequel	0.9012	0.8238	-0.1280	-0.0762	-0.0495	-0.0207	-0.0142
Jamaica Dollar	0.4626	1.9914	-0.6054	0.0174	0.0658	0.0143	0.1173
Japan Yen	0.9191	0.6894	-0.0358	-0.0195	-0.0155	-0.0865	-0.0730
Jordan Dinar	0	0	0	-0.0006	-0.0039	0	0
Kazakhstan Tenge	0.8926	0.4669	-0.2301	-0.0084	-0.0680	-0.0019	-0.0164
Kenya Shilling	0.4322	1.2435	-0.4127	0.0716	0.0333	-0.0728	-0.0483

Notes:

a. The training period used in the calculations spans from that indicated in the column "initial date" to that in the "final date". The prediction period spans from the day after that indicated in the column "final date" to 27 September 2008

b. The parameters of Taylor's strategy were obtained through maximizing the Sharpe ratio by a genetic algorithm.

c. In blue, the U* statistic rejects the null in favour of trend at the 5% confidence level, but Taylor's strategy is not able to improve the B&H strategy.

d. In orange, the U* statistic rejects the null in favour of trend at the 5% confidence level, but Taylor's strategy is not able to improve the B&H strategy and Taylor's strategy overcomes the B&H strategy.

Table 4 (continued)

Currencies	q	k1	k2	B&H	Sharpe B&H	Taylor	Sharpe Taylor
South Korea Won	0.8258	1.6384	-0.0429	0.1460	0.1229	0.0842	0.0719
Kuwait Dinar	0	0	0	-0.0232	-0.0704	0	0
Kyrgyzstan Som	0.9287	0.6811	-0.1067	-0.0215	-0.0313	0.0508	0.0774
Lebanon Pound	0	0	0	-0.0042	-0.0504	0	0
Leshoto Loti	0.9804	0.5634	-0.0718	0.1361	0.0672	-0.0714	-0.0421
Madagascar Ariary	0.9669	0.8699	-0.1247	-0.0896	-0.1461	0.0925	0.1574
Malawi Kwacha	0	0	0	0.0222	0.0558	0	0
Malaysia Ringgit	0	0	0	0.0206	0.0305	0	0
Maldive Islands Rufiyaa	0	0	0	0.0124	0.0750	0	0
Mauritania Ougiyaa	0	0	0	-0.0907	-0.1858	0	0
Mauritus Rupee	0.9914	1.2347	-0.7250	0.0318	0.0355	0.0015	0.0022
Mexico New Peso	0	0	0	-0.0729	-0.1160	0	0
Moldova Leu	0.8707	1.0659	-0.1055	-0.1620	-0.4139	0.1689	0.4505
Mongolia Tugrik	0	0	0	-0.0155	-0.2092	0	0
Morocco Dirham	0	0	0	5.3454	0.0808	0	0
Mozambique New Metical	0	0	0	-11.2791	-0.0491	0	0
Mianmar (Burma) Kyat	0	0	0	0.0000	0.0000	0	0
Namibia Dollar	0	0	0	0.1273	0.0678	0	0
Nepal Rupee	0	0	0	0.0980	0.1516	0	0
New Zealand Dollar	0.9725	0.8684	-0.0735	-0.0907	-0.0671	0.0268	0.0236
Nicaragua Cordoba Oro	0	0	0	-3.3268	-0.0796	0	0
Nigeria Naira	0.9977	0.8403	-0.1327	-0.0047	-0.0327	0.0163	0.1715
Pakistan Rupee	0	0	0	0.2013	0.1632	0	0
Papua New Guinea Kina	0.9674	0.5295	-0.3155	0.0839	0.2011	-0.0080	-0.0193
Paraguay Guarani	0.6637	1.0996	-0.1296	-0.1766	-0.2685	0.1715	0.2703
Peru New Sol	0.9843	1.3296	-0.0820	-0.0158	-0.0151	0.0148	0.0148
Philippines Peso	0.9834	0.2385	-0.0077	0.1000	0.1100	0.0865	0.1015
Qatar Ryal	0	0	0	-0.0010	-0.0115	0	0
Sao Tome and Principe Dobra	0	0	0	0.0362	0.0720	0	0
Saudi Arabia Rial	0.9832	1.3994	-0.0360	-0.0007	-0.0097	0.0342	0.4572
Seychelles Rupee	0.7919	0.0165	-0.0198	0.0009	0.0125	0.0059	0.0855
Sierra Leone Leone	0	0	0	0.0081	0.0400	0	0
Singapore Dollar	0.9852	0.6078	-0.4699	-0.0159	-0.0271	0.0537	0.0993
South Africa Rand	0.0486	0.7426	-0.0711	0.1227	0.0620	0.0852	0.0464
Sri Lanka Rupee	0.0000	1.4485	-0.8164	-0.0091	-0.0692	0.0008	0.0067
Sudan Pound	0	0	0	0.0234	0.0432	0	0
Suriname Dollar	0	0	0	-0.0065	-0.0444	0	0
Swaziland Lilangeni	0	0	0	0.1274	0.0665	0	0
Syria Pound	0	0	0	-0.0018	-0.0413	0	0
Tajikistan Somoni	0.9242	0.7146	-0.5408	-0.0144	-0.1930	0.0290	0.4286
Tanzania Shilling	0	0	0	0.0077	0.0082	0	0
Thailand Baht	0.9470	1.0524	-0.1202	0.1284	0.0985	0.0399	0.0326
Tonga Pa'anga	0	0	0	0.0232	0.0261	0	0
Trinidad and Tobago Dollar	0	0	0	-0.0073	-0.0111	0	0
Tunisia Dinar	0.9173	1.3829	-1.7782	-0.0015	-0.0017	0.0353	0.0462
United Arab Emirates Dirham	0	0	0	-0.0011	-0.0699	0	0
British Pound	0	0	0	-0.0804	-0.0881	0	0
Uruguay Peso	0.7158	0.1159	-1.1663	-0.1131	-0.2646	0.0222	0.0634
Venezuela Bolivar Fuerte	0	0	0	0.0000	0.0000	0	0
Viet Nam Dong	0.9845	0.1441	-0.0734	0.0348	0.0878	0.0278	0.0707
Zambia Kwacha	0	0	0	-0.0610	-0.0439	0	0

Notes:

a. The training period used in the calculations spans from that indicated in the column "initial date" to that in the "final date". The prediction period spans from the day after that indicated in the column "final date" to 27 September 2008

b. The parameters of Taylor's strategy were obtained through maximizing the Sharpe ratio by a genetic algorithm.

c. In blue, the U* statistic rejects the null in favour of trend at the 5% confidence level, but Taylor's strategy is not able to improve the B&H strategy.

d. In orange, the U* statistic rejects the null in favour of trend at the 5% confidence level, but Taylor's strategy is not able to improve the B&H strategy and Taylor's strategy overcomes the B&H strategy.

As can be seen in Table 4, for the exchange rates series where the U^* statistic accepted the null hypothesis of random walk at a significant level of 5%, the return obtained by B&H strategy is higher than Taylor's strategy. This lack of predictive power is also confirmed by comparing Sharpe's ratios which are lower for the B&H strategy. Note that for the series where the trend is not accepted we have not applied Taylor's strategy.

The countries where the U^* statistic rejects the null in favour of trend may be divided into two groups:

- Currencies where Taylor's strategy is not able to improve the B&H strategy, neither in return nor in Sharpe ratios. This happens in 14 out of the 39 cases. For these currencies although, in theory, the trends detected could be employed to beat the market, in practice it doesn't, at least not in the prediction period considered. Taking into account that sufficient large and long-life trends in prices will make a market inefficient, such markets were probably inefficient during the years studied. However, Taylor's strategy is not able to exploit these inefficiencies with predicting purposes during the period considered in 2008.
- Currencies where Taylor's strategy overcomes the B&H strategy, as much in returns as in Sharpe ratios. This happens in 25 out of the 39 cases and, as can be seen in Figure 3, this behaviour is more frequent in intermediate exchange-rate regimes. These exchange markets were probably inefficient during the years studied and, it was even possible to exploit slight dependence between returns using Taylor's trend model during 2008.

5. CONCLUDING REMARKS

In this paper we have tested for the random walk hypothesis against the existence of trends in 95 exchange rate series against the US Dollar. To that end, we have applied Taylor's (1980) trend price model and Taylor's U^* statistic. The parameters defining the trend were estimated by maximum likelihood by mean of a genetic algorithm. Finally, a technical strategy, proposed by Taylor, devoted to obtaining extraordinary profits in the case of trends in the financial prices, was implemented.

The main results are as follows. First, when using Taylor's U^* statistics, in 39 of the 95 cases considered we find evidence in favour of the presence of trends in exchange rate series for the last known regime, being trends more frequent in intermediate exchange-rate regimes.

Secondly, when deriving optimal one-step-ahead forecasts of returns based on these trends and constructing a trading rule based on these forecasts, we find that Taylor´s strategy overcomes the buy-and-hold strategy in 25 out of 39 cases where trends are detected.

Therefore, this paper has showed the potential usefulness of Taylor's price trend model for technical trading rules to forecast daily exchange data when the model parameters are estimated by maximum likelihood using genetic algorithms.

REFERENCES

BROCK, W., LAKONISHOK, J. and LeBARON, B. (1992): "Simple Technical Trading Rules and Stochastic Properties of Stock Returns", *Journal of Finance*, Vol. 15, pp.1731-1764.

BARTLETT, M. S. (1946): "On the Theoretical Specification of Sampling Properties of Auto correlated Time Series", *Journal of Royal Statistical Society B*, Vol. 8, pp. 27-41.

CHAN, K. (1992): "A Further Analysis of the Lead-lag Relationship between the Cash Market and Stock Index Futures Market", *The Review of Financial Studies,* Vol. 1, pp. 123-152.

DORSEY, R. D. and MAYER, J. M. (1995): "Genetic Algorithms for Estimation Problems with Multiple Optima, Nondiferentiability, and other Irregular Features", *Journal of Business & Economic Statistics*, Vol. 13, pp. 53-66.

FAMA, E. F. and FRENCH, K. R. (1988): "Dividend Yields and Expected Stock Returns", *Journal of Financial Economics*, Vol. 22, pp. 3-25,

GENÇAY, R. (1996): "A Statistical Framework for Testing Chaotic Dynamics via Lyapunov Exponents", *Physica D*, Vol. 89, pp. 261-266.

HAUPT, R. L. and HAUPT, S. E. (2004): *Practical Genetic Algorithms* (New Jersey: Wiley), 2nd edition.

HOLLAND J. (1975): *Adaptation in Natural and Artificial Systems* (Ann Arbor: The University of Michigan Press).

ILZETZKI, E. O., REINHART, C. M and ROGOFF, K. S. (2008), "Exchange Rate Arrangements Entering the 21st Century: Which Anchor Will Hold?", forthcoming.

LO, A. W. and A. C. MacKINLAY, (1988): "Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test", *Review of Financial Studies*, Vol. 1, pp. 41-66.

KWAN, J. W. C., LAM, K., SO, M. K. P. and YU, P. L. H. (2000): "Forecasting and Trading Strategies Based on a Price Trend Model", *Journal of Forecasting*, Vol. 19, pp. 485-498.

REINHART, C. M. and ROGOFF, K. S. (2004): "The Modern History of Exchange Rate Arrangements: A Reinterpretation," *Quarterly Journal of Economics*, Vol. 119, pp. 1-48.

STOLL, H. R. and WHALEY, R. E. (1990): "The Dynamics of Stock Index and Stock Index Futures Returns", *Journal of Financial and Quantitative Analysis*, Vol. 4, pp. 441-468.

TAYLOR, S. (1980): "Conjectured Models for Trends in Financial Prices, Tests and Forecasts", *Journal of the Royal Statistical Society. Series A*, Vol. 13, pp. 338-362.

TAYLOR, S. (1982): "Tests of the Random Walk Hypothesis against a Price-trend Hypothesis", *The Journal of Financial and Quantitative Analysis*, Vol. 17, pp. 37-61.

TAYLOR, S. (1985): "The Behaviour of Futures Prices over Time", *Applied Economics*, Vol. 17, pp. 713-734.

TAYLOR, S. (1986): *Modelling financial time series* (New York: Wiley).

TAYLOR, S. (1988): "Forecasting Market Prices", *International Journal of Forecasting*, Vol. 4, pp. 421-426.

TAYLOR, S. and KINGSMAN, B. G. (1978): "Non-stationarity in Sugar Prices", *The Journal of the Operational Research Society*, Vol. 29, pp. 971-980.

EXEMPLO DA UTILIZAÇÃO DO MODELO DA ARBITRGE PRICING THEORY NA ANÁLISE DE INVESTIMENTOS SUJEITOS A DIVERSAS FONTES DE RISCO.

FRANCISCO JOSÉ SANCHES TOMÉ
E-mail: tome@ipg.pt
Escola Superior de Tecnologia e Gestão
Instituto Politécnico da Guarda

Resumo

Nos últimos tempos, a incerteza e volatilidade dos cenários macroeconómicos não tem parado de surpreender os analistas de investimentos, exigindo uma grande flexibilidade no ajustamento de parâmetros e critérios. Muitos autores têm estudado, criticado e revisto as técnicas de avaliação, sobretudo devido ao facto de não incluírem o factor risco da estrutura do capital, utilizarem uma taxa mínima da rendibilidade única, e pelo facto de tratarem o risco isoladamente do contexto patrimonial da empresa. Dos vários métodos de avaliação de investimentos, a Arbitrage Pricing Tjeory é um modelo multifactorial que permite estimar o rendimento esperado, em função de diferentes graus de aversão ao risco por parte de investidores, diferentes fontes de informações e diferentes cenários económicos. Neste trabalho, mostramos que a aplicação da análise factorial permite detectar os factores explicativos e a sensibilidade dos investimentos a essas fontes de risco, com vantagens assinaláveis na análise de investimentos.

Palavras-chave: Arbitrage Pricing Theoy, Análise de Investimentos, CAPM, Risco-Rendimento.
Área Temática: Economia Monetária e Financeira .

Abstract

Recently, the uncertainty and volatility of macroeconomic scenarios has not stopped surprising the analysts of investment, requiring great flexibility in adjusting the parameters and criteria. Many authors have studied, criticized and revised the technical evaluation, mainly due at failure to include venture capital, using a minimum rate of return only, and the fact that treating the risk of isolated context of the company's assets The Arbitrage Pricing Tjeory is a multifactorial model that allows to estimate the expected returns, for different degrees of risk aversion by investors, different sources of information and different economic scenarios. In this paper, we show that the application of factor analysis to detect the explanatory factors and the sensitivity of investment to these sources of risk, with significant advantages in the analysis of investments.

Key words: Arbitrage Pricing Theoy, Investment Analysis, CAPM, Risk-Return.
Thematic Area: Monetary and Financial Economics.

1 – INTRODUÇÃO

A mudança do paradigma económico tradicional, e a intensa competitividade exigem uma constante procura de novos métodos e técnicas de gestão, em particular na análise de investimentos. Neste campo usam-se frequentemente indicadores, tais como a taxa interna de rendibilidade (TIR), Valor actualizado Liquido (VAL) e o prazo médio de recuperação do capital investido (*payback);* todavia com resultados pouco satisfatórios para ambientes macroeconómicos instáveis. Muitos autores têm criticado estas técnicas, procurando novos métodos.

Os modelos de equilíbrio de activos financeiros, que tiveram a sua origem com Markowitz (1952) com base no pressuposto da aversão ao risco e do facto da rentabilidade seguir uma função distribuição normal, foram posteriormente actualizados por Sharpe (1964), Linter (1965) e Mossin (1966). Estes autores conceberam um modelo a que chamaram Capital Asset Pricing Model (CAPM) com a finalidade de avaliar activos em mercados financeiros e que marca uma nova era na teoria financeira.

A partir de 1976, como base na crítica a este modelo, Ross (1976) desenvolveu a teoria Arbitrage Pricing Theory (APT), inspirada no modelo CAPM mas de aplicação muito mais genérica. Esta teoria supõe que o rendimento dos activos seja gerado por um conjunto diversificado de k factores de risco. Na APT, assim como o CAPM, a taxa de rendimento de um investimento é função da exposição ao risco sistemático dos activos deste investimento e não do risco total. Contrariamente ao modelo CAPM, a APT não requer que todos os investidores se comportem de maneira semelhante (hipótese da expectativas e funções utilidade homogéneas), aproximando-se mais da realidade pelo facto da aversão ao risco para cada investidor provir de diferentes fontes de informação. Além disso a APT é um modelo multifactorial que permite estimar para diferentes projectos, quer sejam investimentos financeiros ou empresariais o rendimento esperado, considerando diferentes graus de aversão ao risco por parte de investidores, diferentes fontes de informações e diferentes cenários económicos. Segundo Ross (1995) a metodologia a aplicar seria a análise factorial, que permite detectar os ***k*** factores explicativos e assim determinar a sensibilidade dos investimentos a diversas fontes de risco. A exposição a estes riscos é medida pelo factor *beta*, assim como no CAPM. Goetzmann (1997) justifica que a diferença essencial entre o modelo CAPM e a APT está em que o primeiro usa uma variável de risco normalmente associada ao mercado, enquanto que a APT usa várias variáveis fontes de risco. A principal vantagem deste método consiste no facto da determinação dos factores de risco ser uma atitude dinâmica que varia no tempo e na pessoa e por conseguinte, arbitrariamente definida de acordo com os conhecimentos e sensibilidade dos especialistas que o concebem.

O momento actual estes métodos que tiveram a sua aplicação privilegiada nos mercados financeiros, poderão ser aplicados também com vantagens assinaláveis, na avaliação de projectos de investimento, o que obriga a uma compreensão, discussão e teste de métodos de analise de investimentos de forma a responder às mudanças do meio envolvente de grande incerteza e risco.

Perante este contexto, o presente estudo propõem-se analisar o modelo de equilíbrio de activos financeiros CAPM e a teoria de arbitragem APT e avaliar a sua aplicação a analise de projectos de investimento em ambientes macroeconómicos dinâmicos apresentado um confronto com os métodos tradicionais de avaliação de investimento.

2 - OS MODELOS DE EQUILÍBRIO DE ACTIVOS FINANCEIROS

Antes de aplicar os novos modelos financeiros e mostrar a sua utilidade na gestão a projectos de investimento, convém fazer uma revisão bibliográfica dos mesmos. Na era da globalização de mercados, onde a informação é um recurso de importância capital, constatamos que os investimentos estão cada vez mais sensíveis a variados e muitas vezes desconhecidos factores de risco e cuja previsão é muitas extremamente difícil. Então, qual será a medida mais apropriada para medir o risco dum investimento, quando não se conhece a distribuição estatística dos fluxos de caixa esperados? Com origem nos mercados financeiros, desenvolveram-se nos últimos anos duas teorias, com um corpo de conhecimentos suficientemente estruturado, para responder a esta questão: o CAPM e a APT.

O modelo CAPM atesta que existe um factor, não diversificável, que afecta os rendimentos esperados das acções, a que se chamou factor de risco de mercado. Uma década depois do CAPM ter sido divulgado, Ross (1976) formulou um modelo mais generico que o CAPM, pelo facto de assinalar várias fontes de risco para os títulos. Neste modelo o rendimento de um activo é calculado como sendo a soma do rendimento actual de um activo isento de risco (rf), mais o factor beta do activo , multiplicado pelo *prémio de risco* , que é a diferença entre o retorno

histórico médio desta carteira, menos o retorno histórico do activo isento de risco, tal como se expressa na equação (1).

$$r_i = r_f + \beta_i (r_m - r_f) \text{ (1)}$$

O parâmetro mais importante deste modelo é o coeficiente beta e a sua interpretação, como indicador de nível de risco sistemático. A contribuição do titulo i para o nível de risco da carteira de investimentos é medida pela covariância entre o rendimento do activo e o retorno desta carteira. Esta contribuição dividida pela variância do retorno da carteira é chamado de coeficiente beta, conforme a equação 2.

$$\beta_i = \frac{\sigma_{i,c}}{\sigma^2} \text{ (2)}$$

Onde: $\sigma_{i,c}$ é a covariância do retorno do activo *i* com a carteira e σ^2 é a variância do retorno da carteira.

Na prática, a estimativa de beta de um activo é frequentemente feita através de uma simples regressão linear da rendibilidade de um dado activo, em função rendibilidade de um índice representativo do mercado. A inclinação da recta, então obtida é entendida como o coeficiente beta do activo, que permite classificar o risco dos títulos. De acordo com este modelo, não faz sentido estimar o risco isolado de um projecto mas sim utilizar a taxa de rendibilidade da carteira de investimentos para descontar os fluxos estimados de cash-flows do investimento. No entanto, refere a este propósito Goetzmann (1997), que tal projecto deveria ser suficientemente comparável a todos as outras componentes da carteira de investimentos, de forma a este não influenciar o risco global da carteira da empresa. Além disso, deveriam verificar-se todas as hipóteses do modelo CAPM para que os resultado fosse conforme à teoria (expectativas dos investidores homogéneas, carteira diversificada, mercados eficientes, entre outros). A este propósito, encontramos em Ross (1995), um exemplo da utilidade do coeficiente beta. Este coeficiente ao apresentar-se positivo, mostra que quando há subida inesperada da taxa de inflação, a cotação no mercado aumenta. Para empresas cujos custos e receitas crescem e decrescem na mesma proporção, ou seja que ficam imunes à inflação, os beta seriam nulos ($\beta = 0$). No entanto Fama e French (1992) defendem que o coeficiente beta não explica correctamente o desempenho dos títulos no mercado, dado que há outros factores específicos à empresa que justificam mais eficazmente essa variação. Além, disso Roll e Ross (1992) alertam para o facto dos índices de mercado utilizados em muitos estudos com o CAPM podem não reflectir com exactidão a rendibilidade da verdadeira carteira de mercado.

As leituras exploratórias realizadas sobre o modelo CAPM permitem-nos afirmar que, entre outros, Sharpe (1964), Elton e Gruber (1995), Lehmann e Modest (1985), Serra (1996), confirmam o modelo, enquanto os estudos de Fama e Macbeth (1973) o rejeitam abrindo caminho a desenvolvimentos posteriores cada vez mais complexos e à discussão das suas hipóteses. Foi nesta corrente de contestação crítica ao modelo CAPM, que surgiu o trabalho de Ross (1977), dando origem à Teoria de Avaliação por Arbitragem. Desde então vários estudos procuraram testar este modelo, de entre os quais podemos citar os trabalhos de Roll e Ross (1980, 1995), Ross , Roll e Chen (1983), Chen (1983), Cheng (1995), Burmeister, Roll e Ross (1997), Goetzmman (1997).

A Teoria de Avaliação por Arbitragem, utiliza um modelo económico, que parte de um conjunto de hipóteses básicas, comuns ao modelo CAPM; todavia, enquanto o CAPM se centra no processo de construção de carteiras eficientes, a APT subalterniza esse conceito para assumir que a cotação de cada título está depende de diversas influências ou factores, independentes umas das outras. Apoiando-nos no estudo de Roll e Ross (1995), podemos afirmar que o rendimento actual de um activo pode ser decomposto em três componentes: o rendimento esperado, o rendimento devido à sensibilidade de variação aos factores de risco sistemático e o rendimento devido ao factor residual ou idiossincrático. A relação de base que define a rendibilidade de um activo *i* , num dado momento *t*, influenciado por *k* factores de risco (F), é dada por:

$$R_i = E(R_i) + \sum_{k=1}^{m} b_{i\,k}\, F_k + e_i, \text{ (3)}$$

onde:

$E(R_{it})$ = é a rendibilidade esperada para o activo i;

F_k= valor actual do factor k com impacto sobre o rendimento do título i;

b_{ik}= coeficiente de sensibilidade do activo i associado ao factor sistemático k;

e_i = termo residual, com média zero.

È necessário e recomendável, verificar se F_k e e_i são variáveis independentes, e que o termo de erro não se encontre correlacionado com os factores de risco, nem que os factores de risco entre si, para qualquer período.

As investigações empíricas realizadas sobre a APT têm vindo a reforçar a ideia de ser consistente na explicação dos fenómenos financeiros reais. A utilização no modelo da arbitragem de rendimentos históricos é fácil, no entanto, a sua aplicabilidade é em muito limitada pelas dificuldades em fazer previsões acerca de rendimentos futuros, devido sobretudo à instabilidade temporal dos factores. Se há autores que confirmam que os factores "são estacionários a longo prazo" (Weston e Copeland, 1988), outros, pelo contrário, afirmam que ao utilizar sub-períodos de análise, os factores económicos se alteram. Como exemplo destes últimos podemos citar Aby, Upton e Hays (1997), que confirmam a "instabilidade para todos os factores da APT", no mercado norte americano. Crítica esta que não pode ser ignorada, tal como a suposição de que a matriz de variâncias-covariâncias seja constante ao longo do tempo. O estudo de Guibons (1996), confirma precisamente que "esta hipótese não é consistente com a realidade dos dados".

Se bem que o número de factores do modelo seja variável de autor para autor, Ross e Roll (1984) sugerem, que os factores mais apropriados são as variações inesperadas da taxa de inflação, as alterações da produção industrial, os desvios dos prémios pelo risco e os movimentos não previstos na estrutura das taxas de juro. Evidentemente que quer o numero quer a identificação dos factores é uma questão em aberto, e a sua identificação é muitas vezes arbitrária. Os resultados obtidos parecem ser pouco probatórios e a teoria funciona em alguns casos, ou em subgrupos de acções, no entanto parece longe de ser aceite de forma generalizada e inequívoca.

Uma das vantagens frequentemente apontadas à APT é poder incluir múltiplos factores que captam melhor a realidade macroeconómica, e assim permitir o cálculo da taxa de desconto para projectos de investimento. No entanto, a sua utilização em empresas não financeiras não se generalizou. Os métodos tradicionais que utilizam as técnicas de taxa interna de rendibilidade, payback e valor actual, continuam a ser os mais frequentemente utilizados. Khan e Fiorino (1992), Macinnes e Carleton (1982), confirmam que a popularidade destes métodos deriva da preferência dos gestores pela simplicidade e robustez destes, ao invés da maior exigência de informação, sofisticação de cálculos dos modelos financeiros. Os gestores, preferem tomar decisões rápidas que são prováveis de levar a resultados satisfatórios numa faixa mais ampla de cenários possíveis e que comprometam o mínimo de recursos no presente, ao ter maior flexibilidade futura (MacInnes e Carleton, 1982).

3 - UTILIZAÇAO DA APT NA AVALIAÇÃO DE PROJECTOS DE INVESTIMENTO

Nesta secção apresenta-se um exemplo da utilidade do CAPM e da APT no apoio à tomada de decisões na avaliação de quatro projectos considerados independentes e não mutuamente exclusivos. Os dados aqui utilizados partem de um estudo realizado por Khan e Fiorino (1992), citado por Miranda (1998). Para o modelo multifatorial da APT seleccionaram-se três factores: a taxa de inflação, o crescimento do produto nacional bruto (PNB) e a taxa de juros dos bilhetes de tesouro norte-americano, ao invés do índice de mercado utilizados no modelo CAPM. Este projectos apresentam uma vida útil de 8 anos sujeitos a riscos sistemáticos tais como as taxas de câmbio, preço das matérias-primas, taxas de juros, ciclo económico, acordos de comércio, desregulamentação, reformas tributárias, etc.. Além disso considera-se três cenários económicos possíveis equiprováveis, associados a diferentes preços de petróleo, taxas de inflação e de crescimento económico. Um cenário macroeconómico caracteriza-se por apresentar preços internacionais do petróleo abaixo da média prevista e por conseguinte menos inflação e maior

crescimento económico (cenário 1), outro cenário com preços de petróleo mais elevados e menor crescimento económico (cenário 2) e outro cenário intermédio (cenário 3).

Quadro n. 1 - Análise de projectos por métodos tradicionais

Projectos	Cenários	TIR (%)	*payback* (anos)
1	1	16,3	4,43
	2	17,1	4,35
	3	17,5	4,31
2	1	28,9	3,17
	2	29,7	3,12
	3	30,2	3,10
3	1	18,9	4,44
	2	19,3	4,41
	3	19,6	4,38
4	1	18,8	4,14
	2	19,5	4,07
	3	19,9	4,04

Da análise dos projectos pelos métodos tradicionais (taxa interna de rendibilidade - TIR e o prazo médio de recuperação do capital investido -*payback*), sobressai a particularidade destes projectos (de eficiência energética) se correlacionarem negativamente com o crescimento da economia. Assim, os investimentos de eficiência energética são mais rentáveis quando os preços do petróleo estão abaixo da média do preço do petróleo esperado. Como se pode observar no quadro n.1, para qualquer um dos quatro projectos, a TIR é sempre mais baixa para o cenário 1.

Ao aplicar o método CAPM, relaciona-se o rendimento de um projecto em função do nível de risco sistemático a que este seja submetido, conforme a equação 1. Os parâmetros de mercado associados a cada um dos projectos, bem como as covariâncias entre eles e os betas são apresentados nos quadros *2 e 3*; para cada um dos cenários possíveis.

Quadro n. 2 - Parâmetros do modelo CAPM para cada um dos projectos

Cenários	Projectos			
	1	2	3	4
	$(r_1 - \bar{r}_1)$	$(r_2 - \bar{r}_2)$	$(r_3 - \bar{r}_3)$	$(r_4 - \bar{r}_4)$
1	-0,67	-0,70	-0,37	-0,60
2	0,13	0,10	0,03	0,10
3	0,53	0,60	0,33	0,50

Quadro n. 3 - Coeficientes e estatísticas para cada projecto

Projectos	σ_i	$\bar{r}_i$	$\sigma_{i,m}$	β_i
1	0,50	16,97	-4,39	-0,0553
2	0,53	29,60	-4,68	-0,0589
3	0,30	19,27	-2,47	-0,0311
4	0,45	19,40	-3,97	-0,0499

A rendibilidade média esperada para cada um dos projectos, $(\bar{r}_i)$ é dada pela média ponderada dos valores apresentados no quadro *2,* ponderada pela probabilidade de ocorrência (1/3). Os valores $(r_i - \bar{r}_i)$ correspondem ao desvio do rendimento de cada projecto, relativamente à média ponderada destes. Os valores negativos obtidos para covariâncias e betas traduzem uma correlação negativa entre o rendimento dos projectos e o mercado, dada a particularidade destes projectos serem também, negativamente correlacionados com o desempenho da economia.

Verifica-se que quando o risco sistemático é representado por um beta negativo, o retorno exigido para o risco associado ao projecto é menor que a taxa isenta de risco. Além disso para projectos com maior variância, e logo maior risco, exige-se menores taxas de retorno (caso do projecto 2), que com o maior desvio padrão entre todos, 0,53, exige a menor taxa de retorno de 5,51. Como seria de esperar o excesso de retorno é maior, quanto maior for o risco do projecto, como se pode confirmar no quadro 4, relativo ao CAPM.

Quadro n. 4- Parâmetros do CAPM para cada projecto

Projectos	$\bar{r}_i$	r_f	$\bar{r}_m$	β_i	r_i	e_i
1	16,97	6	14,36	-0,0553	5,54	11,43
2	29,60	6	14,36	-0,0589	5,51	24,09
3	19,27	6	14,36	-0,0311	5,74	13,53
4	19,40	6	14,36	-0,0499	5,58	13,82

Onde:

$\bar{r}_i$ é o retorno médio esperado para cada projecto;

r_f é a taxa livre de risco;

$\bar{r}_m$ é o retorno médio esperado de mercado, considerando os três cenários;

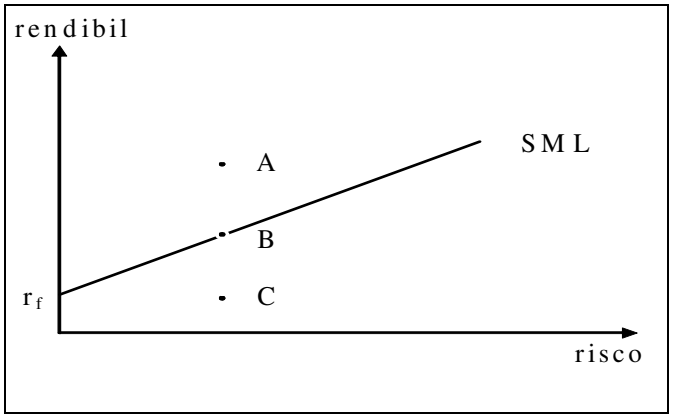
β_i são os betas de cada projecto em relação ao mercado;

r_i é o retorno de cada projecto, calculado pela equação do CAPM; e

e_i é o excesso de retorno de cada projecto.

Neste caso o termo e_i é consistente com o risco sistemático de cada projecto, pois um valor positivo significa que há recompensa pelo risco e portanto, deveria ser aceite como viável. Podemos visualizar num gráfico risco-rendimento o excesso de rendimento, dado pela recta SML do CAPM.

Figura n. 1 – Recta SML do modelo CAPM



A inclinação da SML (Rm-Rf) mede o preço do risco, avaliado pelo mercado, para cada investimento, isto é o valor mínimo exigido para determinado risco, o que nos vai permitir aceitar ou rejeitar os projectos, consoante apresentem um excesso de rendimento negativo (projecto inviável), ou positivo (projecto viável). Observamos que o projecto representado pelo ponto *B* está sobre a recta SML, apresentando um valor correcto para o nível de risco de acordo com a relação risco-rendimento do modelo CAPM. Já o projecto representado pelo ponto *A*, apresenta maior retorno relativamente ao seu risco.; e o projecto *C*, apresenta um rendimento abaixo daquele que seria esperado, sendo considerado economicamente inviável. Face aos dados do quadro anterior todos os quatro projectos são considerados viáveis, por apresentarem um excesso de rendibilidade positiva.

No quadro 5 verificamos que enquanto que para uma taxa mínima de rendibilidade de 20 %, apenas o projecto 2 seria aceite utilizando o método TIR, nenhum dos quatro projectos seria aceite pelo método *payback*, pois o prazo de recuperação é superior a 3 (valor máximo admitido).

Quadro n. 5 - Comparação de métodos TIR, CAPM e Payback

Projectos	CAPM		TIR		*Payback*	
	e_i	decisão	taxa (%)	decisão	anos	decisão
1	11,43	A	17,1	R	4,35	R
2	24,09	A	29,7	A	3,12	R
3	13,53	A	19,3	R	4,41	R
4	13,82	A	19,5	R	4,07	R

Utilizando o modelo multifactorial APT, para cada um dos projectos apresentados neste trabalho, e os três factores sistemáticos seleccionados: inflação ($\bar{r}_{INF}$), produto nacional bruto ($\bar{r}_{PNB}$) e a taxa de juros da economia ($\bar{r}_J$), o rendimento dos projectos analisados; é dado pela *equação 4* :

$$r_i = \bar{r}_{INF} + \beta_{i,INF}(r_{INF} - \bar{r}_{INF}) + \bar{r}_{PNB} + \beta_{i,PNB}(r_{PNB} - \bar{r}_{PNB}) + \bar{r}_J + \beta_{i,J}(r_J - \bar{r}_J) \qquad (4)$$

onde:

$\bar{r}_{INF}$, $\bar{r}_{PNB}$ e $\bar{r}_J$, são os factores de risco;

$\beta_{i,INF}$, $\beta_{i,PNB}$ e $\beta_{i,J}$, representam os coeficientes de sensibilidade

$(r_{INF} - \bar{r}_{INF})$, $(r_{PNB} - \bar{r}_{PNB})$ e $(r_J - \bar{r}_J)$, representam a variação inesperada

Os valores obtidos para cada um dos parâmetros permitiram escrever as equações 5 a 8:

$$r_1 = 3{,}815 + 0{,}169(3{,}871 - 3{,}815) + 3{,}658 + 1{,}904(5{,}489 - 3{,}658) + 3{,}324 - 0{,}609(2{,}338 - 3{,}324) = 14{,}893 \qquad (5)$$

$$r_2 = 3{,}815 + 0{,}185(3{,}871 - 3{,}815) + 3{,}658 + 2{,}039(5{,}489 - 3{,}658) + 3{,}324 - 0{,}722(2{,}338 - 3{,}324) = 15{,}252 \qquad (6)$$

$$r_3 = 3{,}815 + 0{,}100(3{,}871 - 3{,}815) + 3{,}658 + 1{,}089(5{,}489 - 3{,}658) + 3{,}324 - 0{,}392(2{,}338 - 3{,}324) = 13{,}183 \qquad (7)$$

$$r_4 = 3{,}815 + 0{,}156(3{,}871 - 3{,}815) + 3{,}658 + 1{,}733(5{,}489 - 3{,}658) + 3{,}324 - 0{,}609(2{,}338 - 3{,}324) = 14{,}579 \qquad (8)$$

Foi testada a estabilidade do modelo face a alterações das probabilidades de ocorrência dos três cenários apresentados, Constatou-se novamente os resultados obtidos de acordo com a abordagem de métodos tradicionais, onde o maior preço do petróleo tornam os projectos de energias alternativas mais rentáveis.

4 – ANÁLISE DE RESULTADOS

A utilização do método APT, assim como o CAPM, viabiliza todos os projectos, contrariamente aos métodos tradicionais que originalmente, não incorporam o risco na análise de investimentos, conforme resultados do quadro 6.

Quadro n. 6- Comparação de métodos

Projecto	APT		CAPM		TIR		*Payback*	
	e_i	decisão	e_i	decisão	taxa (%)	decisão	anos	decisão
1	2,08	A	11,43	A	17,1	R	4,35	R
2	14,35	A	24,09	A	29,7	A	3,12	R
3	6,09	A	13,53	A	19,3	R	4,41	R
4	4,82	A	13,82	A	19,5	R	4,07	R

Nota: A = Projecto Aprovado, R = Projecto Recusado

A decisão de investimentos baseada na APT aprova a execução de todos os projectos, enquanto a decisão baseada nos métodos tradicionais, reprova a maioria (usando a TIR), ou a totalidade deles (utilizando o *payback*).

5 - CONCLUSÃO

Ao se comparar os resultados da análise pela APT com a análise pelo CAPM, observa-se que o excesso de retorno do primeiro é inferior ao do segundo método. Isto pode ser justificado à maior precisão na captura do risco da APT. Verificamos ainda que esta teoria criada inicialmente para ser utilizada em mercado de capitais pode ser empregue na análise de alternativas de investimentos não financeiros, com vantagens acrescidas e permitir monitorizar a execução de projectos além de permitir fazer simulações a qualquer momento. Apesar da necessidade de realizar mais estudos para comprovar a eficácia dos métodos CAPM e APT, recomenda-se a sua utilização como auxiliares na tomada de decisão, pois caso contrário, pode-se rejeitar projectos economicamente viáveis por não relacionarem correctamente a rendibilidade com o seu nível de risco.

BIBLIOGRAFIA

BARATA, JOSÉ M. (1998), Moeda e Mercados Financeiros, Centro de Estudos sobre Economia Financeira, 2ª edição, Lisboa

BRANDÃO, ELÍSIO (2002) – Finanças, Porto Editora, Porto

BURMESTEIR, E., ROLL, R. E ROSS, S. (1997), A Practitioner's Guide to Arbitrage Pricing Theory, A Practitioner's Guide to Factor Models, Institute of Chartered Financial Analysts, Charlottesville.

CHEN, N.. (1983), Some Empirical Tests of The Investigation of the Theory of Arbitrage Pricing, Journal of Finance, pg. 1393-1414.

CHEN, N., ROLL R. AND ROSS R.(1986), Economic Forces and the Stock Market, Journal of Business, 59(3), p. 383-404.

ELTON, E. E GRUBER, M:. J. (1995), Modern Portfolio Theory and Investment Analysis, 5ª ed. New York, J. Wiley & Sons

FAMA, EUGENE; (1992), French, Kenneth R. The cross section of expected stock returns. Journal of Finance 47, p 427-66.

FAMA, E. F. E FRENCH, K. R. (1988), Multifactor Explanations of Asset Pricing Anomalies, Journal of Finance, Vol. 51, p. 55-84.

GIBBONS, M.; ROSS, S.; SHAKEN, (1989) J. A test of the efficiency of a given portfolio. Econometrica , 57, p. 1121-1152.

GOETZMANN, W.N. (1997) An introduction to investment theory. Yale School of Management. http://viking.som.yale.edu/will/finman540/classnotes.

LEHMANN, B. E MODEST, D. (1988), The Empirical Foundations of the Arbitrage Pricing Theory I: The Empirical Tests, Journal of Financial Economics, 21, p. 213-254

KHAN, A.M.; FIORINO, D.P. (1992) The Capital Asset Pricing Model in Project Selection: A Case Study. The Engineering Economist. v.37, n.2, p.145-60.

MACINNES, J.M.; CARLETON, W.J. (1982) Theory, Models and Implementation in Financial Management. Management Science. p. 957-78.

MARKOWITZ, H. (1952) Portfolio selection. Journal of Finance. p.44-58.

MIRANDA, VINÍCIUS A. (1998), A abordagem do risco em análise de investimentos com a utilização da Arbitrage Pricing Theory (APT). Itajubá: EFEI.

ROLL, R.; ROSS, S.A. (1992), On the Cross-Sectional relation between expected returns and betas. University of California (UCLA) - working paper. Los Angeles. July, 1992.

ROSS S. (1976), The Arbitrage Pricing Theory of Capital Asset Pricing, Journal of Economic Theory, nº 13, p. 344-360.

ROSS S. (1977), Return, Risk and Arbitrage, in Risk and Renan in Finance, ed. by I.Friend and J. Bicksler, Cambridge. pp. 189-218,

ROSS, S.A.; ROLL, R. (1984), The Arbitrage Pricing Theory Approach to Strategic Portfolio Planning. Financial Analysts Journal. p.14-26.

ROSS, S.A.; WESTERFIELD, R.W.; JAFFE, J.F. (1995), Administração Financeira, Corporate Finance. Trad. de Antonio Z. Sanvicente. São Paulo, Atlas.

SHARPE, W.F. (1964), Capital asset prices: A theory of market equilibrium under conditions of risk. The Journal of Finance. v. XIX, n.3, p.425-42.

SHARPE, W.F. (1984), Factor models, CAPMs, and the ABT: Linking them together provides a valuable framework. The Journal of Portfolio Management. p.21-5.

WESTON,J.F.; COPELAND,T.E.(1986) Managerial Finance, Dryden Press, 8 ed., New York.

DOES OWNERSHIP CONCENTRATION, THE IDENTITY OF OWNERS AND THE LEGAL PROTECTION OF SHAREHOLDERS HAVE AN IMPACT ON COMPANY VALUE? EMPIRICAL EVIDENCE FROM WESTERN EUROPE

MÁRIO SACRAMENTO SANTOS
E-mail: msantos@iscac.pt
Instituto Politécnico de Coimbra

Abstract

We confirm recent empirical evidence that indicates that the largest public traded companies throughout the world have a concentrated ownership. Our empirical evidence supports a quadratic relationship between company value and ownership concentration. This result confirms not only the monitoring but also the expropriation effect for the very highest ownership concentration values in European companies. The results also support the hypothesis that the influence of blockholders depends on their identity, on the specific agency problems that companies face, and on the legal context in which they operate.
Concerning the relationship of the second large shareholder and firm value, the results suggest a more complex relationship than that stressed until now. Although the existence of a second large shareholder has a positive influence on the behavior of the first shareholder, what happens is that the dimension of the second larger shareholder's (continuous variable) has a non-linear relationship with the value of companies.
Finally, according to recent results, this paper highlights the superiority of family companies in the control of the agency problems which contrasts with the results of institutional shareholders. Moreover, in family companies, the presence of another individual as the second large shareholder has a negative impact on the company value, whereas there is a positive effect if the second is an institutional investor

*Key words: O*wnership concentration, agency costs, monitoring, expropriation, company value.
Thematic Area: Monetary and Financial Economics.

1. INTRODUCTION

Recent studies sharply contrast with Berle and Means' (1932) model of widely diffused corporate ownership and suggest that the presence of large owners is widespread throughout the world (Franks and Mayer, 1995; La Porta, Lopez-de-Silanes and Shleifer, 1999; Barca and Becht, 2001; and Faccio and Lang, 2002), often the founding families or their offspring.

Empirically, there is no dispute that in (continental) Europe, Asia and Latin America, the vast majority of large, publicly traded companies are family owned (see, for instance, Claessens et al., 2000; Thomsen and Pedersen, 2000; Faccio, Lang, and Young, 2001; Gedajlovic and Shapiro, 1998; Franks and Mayer, 2001; Miguel, Pindado, and de la Torre, 2004). Particularly in European countries, corporations exhibit a large shareholder who hires and closely monitors corporate managers, as documented by Barca and Becht (2001). Faccio and Lang (2002) carried out a study of the ownership and control for 5,232 corporations in 13

Western European countries and found that 86% of the companies have an owner with more than 10% of the voting rights.

This pattern of concentrated ownership is nowadays a fact even in developed countries, including the US (see, e.g., Shleifer and Vishny, 1997; Holderness et al. 1999; La Porta, Lopez-de-Silanes and Shleifer, 1999). For instance, Anderson and Reeb (2003) estimate that families own an average of 18% of equity in the S&P 500. Gaghoum, Lang, and Young (2005) find that 60% of US corporations have controlling shareholders who hold at least 10% of the shares.

Bebchuk, Kraakman, and Triantis (1999), Shleifer and Vishny (1997) and others argue that the primary agency conflict within a context of concentrated ownership is not between owners and managers, but between the dominant shareholder and the rest of minority shareholders. While concentrated ownership helps mitigate the vertical agency problem (Chen and Yur-Austin, 2007), it is possible that a controlling shareholder will extract private benefits of control by forcing decisions which expropriate minority shareholder wealth (Grossman and Hart, 1980; Dyck and Zingales, 2004; Gilson and Gordon, 2003).

Traditional theories argue that concentrated ownership is critical for effective governance, since only large investor have incentives to monitor the manager and, if necessary, intervene to correct value-destructive actions (Shleifer and Vishny, 1986). Although concentrated ownership mitigates managerial expropriation, it engenders a new form of agency problem, the potential expropriation of minority shareholders and other stakeholders by blockholders. These dominant shareholders usually exert full control over managers and frequently hold control power in excess of their cash flow rights, providing them with strong incentives to extract private benefits at the expense of minority shareholders (La Porta, Lopez-de-Silanes and Shleifer, 1999).

The large shareholders expropriation of others investors is limited by two principal mechanisms. First, the legal shareholder protection in a country, as well as the enforcement of these laws functions as a remedy to agency problems (La Porta et al., 1998). Second, the financial incentives associated with cash-flow ownership reduce the owner interest in expropriating other investors, Jensen and Meckling (1976). Empirical research shows that companies operating in countries lacking suitable legal investor protection often have more concentrated ownership structures (La Porta et al., 1998), suggesting that large shareholders have a pronounced role when investors are less protected.

Additionally, in the context of companies with dominant shareholders one of the key issues of corporate governance is the question regarding why control sharing is better than absolute control over a firm when such shareholding structure generates free-rider problems in monitoring? Most large companies in reality have multiple blockholders (see, e.g., Zwiebel, 1995; Barca and Becht, 2001; Faccio and Lang, 2002; Maury and Pajuste, 2005; Holderness, 2007; Gregoric et al., 2008; and Laeven and Levine, 2008). Such a structure appears to be suboptimal for governance, as splitting equity between numerous shareholders leads to a free-rider problem: each investor individually has insufficient incentives to bear the cost of monitoring, and shareholders cannot co-ordinate to share this cost.

One of the most accurate explanation wich research identifies as a positive effect of the second large shareholder remark that the investors face a trade off between holding the optimal portfolio and having little influence on the firm's decisions, or buying more shares, holding a suboptimal portfolio, but influencing the firm's decisions. Depending on the amount invested in the firm investors have different preference for the risk/return characteristic of the project. A potential second blockholder with fewer shares of the initial owner prefers a level of risk and return that lies between the choice of liquidity shareholders and the one of the initial owner. Hence he can influence the voting outcome and mitigate the conflicts of interests between large e minority shareholders.

The aim of this paper is to explore the role of large shareholders in corporate governance. One central theme is how cash flow rights and the type of the large owners influence the value of the company, and how multiple large shareholders interact with each other. We intend to explore the impact of large shareholders in corporate governance by addressing six questions.

First, are large shareholders important for corporate value? Second, what is the relationship between onwnerhip concentration and minority shareholder protection? The previous empirical

evidence is not clear whether ownership concentration and minority protection laws shareholder substitute or complement each other. Third, is the major shareholder's identity important in the market assessment of companies? In particular, do family companies outperform non-family ones? Fourth, do other large shareholders matter for corporate value? What is the empirical relation between the contestability of the first large owner and the value of the company? Fifth, how does the type of shareholders affect this contestability? We believe that homogeneous owners have similar interests and incentives. We anticipate greater probability of collusion if the two largest shareholders belong to the same group. Sixth, can the characteristics of companies (size, debt, profitability, growth opportunities, and assets tangibility) moderate the importance of large shareholders?

In a context of theoretical ambiguous predictions, concerning the impact of ownership structure on company value, which is supported by mix empirical evidence, we decided to use a methodology in which there is broad consensus on its ability to overcome problems of this area of research. It is often considered (see for instance De Miquel et al., 2001 and Himmelberg et al., 1999) that the ambiguity of the results may be due to two main problems that plague the empirical analysis: unobservable heterogeneity and potential endogeneity of ownership. Unobservable heterogeneity means that there are firm specific characteristics that are difficult to measure or difficult to obtain and which do not enter the model. If individual characteristics are not taken into account, the results may be biased. The endogeneity problem might appear when ownership structure is affected by firms' performance (Himmelberg, Hubbard and Palia, 1999; Demsetz and Villalonga, 2001). This problem is exacerbated in our model because the lagged dependent variable is also included as an explanatory variable. If we ignore this potential effect we may get a spurious correlation between ownership structure and firm value. We propose examining the relationship between firm value and ownership concentration using a general method of moments (GMM) dynamic panel estimator that eliminates the major sources of endogeneity inherent in the estimation of the relationship between ownership concentration and value. This estimator, originally developed by Arellano and Bond (1991), and further developed by Arellano and Bover (1995) and Blundell and Bond (1998), provides an excellent econometric specification for dealing with the endogeneity issues we are likely to encounter in the estimation of the ownership/value relationship. The basic steps underlying this estimation strategy is as follows. First, the regression equation of performance on corporate governance is written as a dynamic model that includes lagged value as an explanatory variable. Next, we can take first-differences and carry out GMM estimation using lagged values of the ownership variables, as well as lagged values of value as GMM instruments. First differencing eliminates unobserved heterogeneity and allows arbitrary correlation between unobserved firm effects and our ownership variables. The choice of instruments enables us to control for both dynamic endogeneity and simultaneity.

To achieve the goals of our investigation, we first develop five empirical models that allow us to determine whether there is a stronger positive impact of ownership concentration on company value. These models explore the relationship of the cash flow rights in the hands of the three major shareholders with the company value either when analysed alone or when analysed together.

Six additional models will be proposed to analyse whether the relationship between ownership concentration and company value is moderated by the identity of the first and of the second large shareholder.

An additional set of models (3) explore what is the relationship between ownership concentration in the hands of the second large shareholder and the company value.

Finally, a last set of five models will explore not only the moderating role of age and the opportunities for investment in the relationship between ownership concentration and company value, but also the substitutability between ownership concentration and external protection of minority shareholders' rights.

We contribute to the literature in various ways. First, we will contribute to the emerging literature that investigates the link between ownership structures and value, considering not only the level of ownership concentration but also the identity of the two major shareholders. We

will also contribute to literature on the behaviour of blockholders, showing that blockholders cannot be treated as a single block: different economic incentives of different blockholders produce different types of behaviour and impacts. We find a complex story where a large blockholder has a statistically more significant effect in low investors' protection regimes.

Second, we will contribute to the existing controversy about the benefits and costs of blockholders, giving a particular emphasis to the fact that they are families exploring and acknowledging the existence of nonlinearities relations between ownership concentration and firm value.

Third, we will try to analyse more deeply the influence of ownership concentration control on the ownership value relationship by accounting for firm age, the opportunities of investment (proxy for growth revenues) and by comparing family companies with non-family corporations.

Fourth, we encounter empirical evidence of the importance of the third larger shareholders.

Fifth, our estimation method will explore in a more efficient way than previous studies the impact of ownership concentration as the identity of large shareholders on value of the corporations.

Last, and extremely important, we think that this paper will make us think more deeply about the role and the importance of the second large shareholder and the complex ownership structures incorporate governance. The results suggest that although the existence of a second large shareholder has a positive influence on the behavior of the first shareholder, what happens is that the dimension of the second larger shareholder's (continuous variable) has a non-linear relationship with the value of companies.

This paper proceeds as follows. Section 2 describes the theoretical framework, including the related literature and the unique corporate governance of family companies. The testable hypotheses are then developed. Following this, we describe sample selection and research methodology in Section 3. Section 4 presents the empirical results. The last section highlights the main conclusions of the paper.

2. THEORIES, HYPOTHESES AND MODELS

In this section, we will summarize the key arguments that link largers shareholders with the valuation of the companies. Based on the previous research and on our readings on this field, some hypotheses will be proposed.

2.1. The impact of the major shareholder's ownership on company value

The first question is this: do large shareholders matter for corporate performance? In standard neoclassical models, one shareholder, whether large or small, is not any different from any other shareholder, i.e. all shareholders are homogeneous.

As already previously we underlined, the theoretical literature has extensively modeled the monitoring role of large shareholders as a possible solution to agency problems[1] that arise from the separation of ownership and control in public corporations (e.g., Grossman and Hart, 1980; and Shleifer and Vishny, 1986). Shareholders with large stakes have strong incentives to monitor management and are less inclined to dilute company value given that they bare a substantial share of the cost. In conclusion, ownership by large blockholders addresses the

[1] Although shareholders and management interests can deviate in a number of ways, much of the literature (e.g., Baumol (1959), Donaldson (1984), and Jensen (1986)) focus on what is generally referred to as "empire-building." Managers can be thought of as empire-builders when they invest in negative NPV investment projects in order to pursue growth as an objective rather than value maximization.

agency problem by providing a large enough interest in profit maximization and the power to ensure that their interests are pursued (Holderness and Sheehan, 1988).

Additionally, as large owners can only divest gradually, and may want to return to the capital market later, they care about a reputation for the protection of minority investors. Consequently, to raise finance and to cash in they should be expected to introduce better corporate governance voluntarily at the company level, even if regulation and enforcement lag behind (Durnev and Kim, 2007).

However, a large stake by the largest shareholder is not purely beneficial, it may cause negative effects on company value. A large stake shareholder may exploit minority shareholders (*shareholders expropriation hypothesis*), e.g. see, Burkart and Panunzi (2001). As already previously underlined, dominant shareholders have an incentive and, in the absence of a countervailing force, the ability to divert corporate resources from other shareholders to themselves for personal consumption. Such diversion reduces the observed market value of the company. In addition to agency relations, there may be other reasons why concentrated ownership can raise or lower company value. For example, it may reduce the liquidity and therefore the value of a share to minority investors.

Another cost of large shareholders, from the perspective of small shareholders, is that they may have preferences that diverge from company value maximization. Fama and Jensen (1983) show that large shareholders, because they are undiversified, may favor investment choices that differ from value maximization. Similar arguments can be made regarding the choice of capital structure.

Gorton and Schimid (2000) show that the value of German companies improves as the ownership concentration rises. The findings in Lehmann and Weigand (2000) confirm the benefits of large shareholders in Germany, but only in the case of banks, since the presence of non-financial large owners negatively affects company profitability. Mork, R., M. Nakamura, and A. Shivdasani (2000), find that Japanese companies average q ratios rise monotonically with corporate blockholders. They conclude that the positive relation between company value and corporate block holdings is consistent with the hypothesis of Shleifer and Vishny (1986) that large blockholders are a way of overcoming the free-rider problems in shareholder monitoring associated with dispersed ownership.

Leech and Leahy (1991) showing a negative relation between ownership concentration and company value and profitability for British company. Mudambi and Nicosia (1998) confirm the negative relation between ownership concentration and performance for the British case.

Holderness (2003) provides a summary of the "current learning" on this relationship as follows: "First, it has not been definitively established whether the impact of blockholders on company value is positive or negative. Second, there is little evidence that the impact of blockholders on company value whatever that impact may be is pronounced." Since theory cannot specify the relative importance of these costs and benefits, the shape of the relation between concentration and performance must be determined empirically.

Based on the existence of these two potential and opposing effects of ownership on value, non-linear effects are not unlikely. It may be that the monitorinf effect dominates for small levels of blockholder ownership, whereas expropriation effects set in at higher levels.

The effect of ownership structure on firm's market value has been extensively studied. In the U.S., Morck, Shleifer and Vishny (1988) fit a piecewise linear regression of Tobin's Q on controlholders ownership. Company valuation increases for management holdings of 0% to 5%, decreases in the range of 5% to 25%, and increases for management holdings greater than 25%. McConnell and Servaes (1990) fit a quadratic relation between Q and insider ownership. Q increases with insider ownership, peaks at ownership levels of 40% to 50%, and then slightly decreases with insider ownership.

Recent European studies also document significant relations between firm's market value and its ownership structure. In Sweden, Cronqvist and Nilsson (2003) find a negative relation between Q and controlholders' vote. In Norway, Bohren and Odegaard (2005) report a quadratic relation between Q and insider ownership – firm's Q increases up to insider ownership of about 60%; then it decreases. A quadratic relation is observed in Swiss companies too – see Beiner,

Drobetz, Shmid and Zimmermann (2004). It appears that the quadratic (inverted U) pattern of the relation between Q and insider ownership, first observed in the U.S. by McConnell and Servaes (1990), emerges in European economies as well. In Spain, Pindado et. al. (2006) report that the value of Spanish companies rises as ownership concentration increases from 0% to 87%, and that beyond this breackpoint company value is negatively affected by ownership concentration.

In search of one functional form to answer both monitoring and expropriation hypothesis, Gedajlovic and Shapiro (1998) find evidence of a non-linear relationship between ownership concentration and profitability suggesting both the monitoring and expropriation effects in German and US companies. However, they did not find any relationship between ownership concentration and company performance among UK, France and Canada. Éric Séverin (2002) from a sample of French companies obtained results consistent with the findings of Morck, Shleifer and Vishny (1988), McConnell and Servaes (1990) and Short and Keasey (1999) and suggest a non-linear relation between ownership structure and performance.

In order to capture the non-linear relationship between value and ownership concentration, as the result of joint consideration of the monitoring and expropriation effects, our first hypothesis predicts:

Hypothesis 1a: Company value increases with ownership concentration in the hands of the principal shareholder at low levels (as a result of the monitoring effect), and decreases at high levels (as a consequence of the expropriation effect).

In order to understand the role played by the second and third large shareholders it seems important to place the following hypotheses:

Hypothesis 1b: Company value increases with ownership concentration in the hands of the two largest shareholders at low levels (as a result of the monitoring effect), and decreases at high levels (as a consequence of the expropriation effect).

Hypothesis 1c: Company value increases with ownership concentration in the hands of the three majort shareholders at low levels (as a result of the monitoring effect), and decreases at high levels (as a consequence of the expropriation effect).

Finally, in order to realize if the existence of a third blockholder affects the behavior of the two largest shareholders, and also with the aim of understanding if, contrary to what has been emphasised until now, there is a significant (and thus positive, we suppose) effect of cash flow rights of the third large shareholder with the value of companies. Consequently, we decided to formulate the following hypothesis:

Hypothesis 2: The existence of a third blockholder and the size of his cash flow rights have a positive relationship with the value of the companies.

2.2 The impact of the identity of major shareholders on value

Another streem of literature that in them he seems interesting we investigate is if the identity of the major shareholder is important in the evaluation that the market makes of the companies. While the literature has gone far in showing the prevalence of large shareholders in many countries and stressing the importance of the legal investor protection for the governance companies, a much lesser part of the literature has considered the role of different types of large shareholders and the interaction between such blockholders.

As mentioned above, concentrated shareholders have strong economic incentives to monitor managers and decrease agency costs (Demsetz and Lehn, 1985). In the case of founding-family ownership, this incentive should be particularly strong since families usually have invested most of their private wealth in the company and are not well-diversified. Therefore, families are a

unique type of investor who has exceptional concerns over company survival and strong incentives to monitor management closely. If the monitoring activity requires the knowledge of a company or market-specific technology, families might also have an advantage due to their long-term presence in the company. The long-term nature of family shareholdings also suggests that family companies build a reputation that affects their relationship with customers and external suppliers of capital.

Anderson and Reeb (2003) sustained that the presence of family owners in the company and their longer investment horizons are likely to reduce managerial myopia, thus leading to better company performance. Furthermore these autores found that because of a family firm's long-term presence they could avail themselves of cheaper debt financing in comparison with non-families.

One of the main arguments for the existence of the family company advanced by earliest extant literature is the coincidence between control and management decisions (Jensen and Meckling, 1976; Dalton and Daily 1998), which supposedly reduces the agency costs of separation of ownership and control.

Agency relationships create information asymmetries which may ultimately harm the welfare of owners and pose a moral hazard to agents. In family companies, information is gathered and shared through the communication channels among members whose special relationships to one another allows agency problems to be controlled without separation of management and control decisions. So, family management should further decrease agency costs.

In addition to the possible benefit of reduced agency costs due to incentive alignment, family companies may gain from employment relationships based on altruism and trust (Karra, Tracey and Phillips, 2006). Altruism within the family could lead to superior employment contracts by reducing the necessity of excessive monitoring and incentives-based pay, and by offering credible threat of sanctions from other family members (Chami and Fullenkamp, 2002; De Paola and Scoppa, 2001; Randøy and Nielsen, 2002).

On the other hand, there is also what may be called the family's "dark side" which, through its power position, could use various mechanisms and opaqueness in the firm's organization to expropriate cash flows from the company and direct them into its own pockets or use them for "pet projects". This behavior should lead to an increase in debt agency costs. The combination of management and control might lead to sub-optimal investment decisions since the interests of the family are not necessarily in line with those of other shareholders (Fama and Jensen, 1985).

The family's role in selecting managers and members of the supervisory board will also increase entrenchment and may lower company value since external parties can hardly capture control over the company. In line with these arguments, the high family stake reduces the probability of bidding by other external investors and should lead to a lower market valuation (Barclay and Holderness, 1989).

Family owners with large shareholdings can have a negative impact on non-family employees working in the company. Shleifer & Vishny (1997) suggest that large shareholders insist on being active in management even if they are not competent or qualified as compared to a professional manager. This gives rise to big losses. Burkart, Gromb, & Panunzi (1997) point out that when managers are closely monitored by large shareholders, they can lose their initiative to search for new investment projects, as there is a possibility of shareholder interference. Managers who have full knowledge of the best opportunities for a company have effective control but with monitoring, large shareholders can gain access to this information and block private, discretionary benefits that managers could otherwise gain. Therefore concentrated ownership coupled with monitoring can have a negative impact on managerial initiative.

A number of studies have documented the prevalence of family ownership around the world. Morck, Wolfenzon and Yeung (2005) suggest that family ownership may lead to entrenchment of poor managers, expropriation of resources from minority shareholders, capital misallocation, reduced or inefficient investment, and lower growth. But studies of the relationship between family ownership and performance have produced mixed results.

Holderness and Sheehan (1988) and Morck, Shleifer and Vishny (1988) find that family companies have lower valuations than other companies.

In contrast to this prediction, recent empirical evidence suggests that founding-family ownership is associated with superior company performance when compared to widely-held companies, both in terms of accounting performance and market valuation, Anderson and Reeb (2003) and Villalonga and Amit (2006) Baronti and Caprio, 2006). Additionally, Martikainen, Nikkinen and Vahamaa (2008) find that family firms are more productive than non-family ones.

In terms of the European evidence, Sraer and Thesmar (2007) show that for a sample of French stock-market listed companies, family companies outperform widely-held corporations. Finally, in a cross-country study of Continental European companies, Barontini and Caprio (2006) confirm the finding that market valuation and operating performance are higher in founder-controlled corporations and at least not worse in descendantcontrolled companies.

In this same line of investigation, Thomsen and Pedersen (2000), Lehmann and Weigand (2000), McConaughy et al. (1998), Barth et al. (2005) among others, find equally that the identity of owners may be more important in explaining company performance than the level of ownership itself. Following this line of literature, we are concern about the identity of the shareholders. We expected, in line with previous investigations, that it is not indifferent the identity of shareholders in value terms. In order to evaluate the impact of the identity of the largest owners, we categorize them into several investor groups. First, we look at the identity of the first largest shareholder and distinguish between: *(i) state*, when the ultimate owner is the state or a public authority; *(ii) family*, when the ultimate owner is either a family or an individual; *(iii) industrial* company; *(iv) institutional investors*, when the ultimate owner is either a financial company, an insurance company, mutual a pension fund/trust an private equity company or a bank; and *(v) other.*

Consistent with the potential benefits of family control, we have initially proposed a positive relationship between ownership concentration and company value in the case of family companies.

Hypothesis 3: Family ownership concentration has a positive impact on company value.

Taking into account the potential cost of family control we propose that family ownership may be detrimental to the value maximization objective of the company when the stake of the family in the company is too high. Therefore, we posite the following hypothesis:

Hypothesis 4: As ownership concentration rises in the hands of the family shareholders, company value increases first and, after a certain breakpoint, it decreases.

Given the importance of holding by institutional investors and their particular features, we decided to give them special attention. First, institutional investors are characterized as myopic owners who sell in response to unfavorable short-term results. Certainly, there is evidence of institutional positive-feedback trading, window-dressing and high turnover (Lakonishok *et al.,* 1991). Some executives argue that this myopia makes them peripheral shareholders unentitled to a voice in corporate governance; short-term shareholders are likened to (voteless) optionholders—focused on short-term price moves—and, accordingly, entitled to no voice in corporate governance. Since a key ingredient in effective monitoring is gaining influence with management, a short-term focus will hinder institutional activism. Further, a short time horizon reduces incentives to monitor since it reduces the period over which to amortize start-up costs.

Second, diversification and liquidity are key reasons individuals use institutions to invest. Liquidity needs reinforce the need to diversify, so both diversification and liquidity encourage small position sizes. This makes monitoring difficult as the relatively fixed monitoring costs are spread over fewer shares held. Institutions may choose to avoid the excess risk of non-diversification (Demsetz and Lehn, 1985).

Finally and importantly, institutional investors may lack sufficient expertise at monitoring corporate management. They may lack expertise at any of the three stages of monitoring I

discussed above—effort, influence and results. They may lack the ability to identify good monitoring candidates, may lack the credibility or negotiation skill necessary to get management's ear, and may lack useful value-enhancing advice or the skill to get management to commit to their own value-enhancing actions. If institutional

managers lack expertise in their primary investment management role and systematically underperform the market, one should not expect aptitude in a role ancillary to their primary duty. The fourth hypothesis is as follows:

Hypothesis 5: Institutional ownership is negatively associated with company value.

2.3. Do other large shareholders matter?

The effects of ownership concentration on company performance may depend on how the blockholders interact among each other, namely the contestability of control.

The possibility of shareholder cooperation in monitoring is more likely when the coalition is sufficiently stable over time, when their stakes are sufficiently large and when the legal environment facilitates collective action in exerting voting rights (Pagano and Roel, 1998). Coalitions of large owners are more likely to exist in companies with large financing requirements and with investment opportunities that are difficult for outsiders to evaluate e.g. in non-listed companies (Gomes and Novaes, 2006). Moreover, coalitions are more likely to occur in companies with homogeneous shareholders.

Pagano and Roell (1998), Edwards and Weichenrieder (1999) and Bloch and Hege (2001) they converge in the conclusion of that the net effect of an incremental holding by the second largest shareholder is advantageous to minority shareholders.

Bennedsen and Wolfenzon (2000) and Gomes and Novaes (2006), theoretically show that multiple large shareholders may limit expropriation of minority shareholders.

In a study of the German listed companies, Lehmann and Weigand (2000) show that the existence of a second large owner increases the profitability of German companies.

According to Volpin (2002) in Italy, when blockholders form syndicates the company market value is higher than when there is a single blockholder. In Spain the number of blockholders is positively related to a better performance of private companies (Gutierrez and Tribo, 2004).

Recently, Maury and Pajuste (2005) studied the effect of having multiple large shareholders on company valuation for a sample of Finnish listed companies. Their main result is that as control contestability, measured by different proxies, increases company value also increases[2]. Maury and Pajuste (2005) show that the the incentives of the second large shareholder to collude with or to monitor the controlling shareholder are affected by the type of the individual owner.

Finally, Laeven and Levine (2008) find that the 35% of all companies, and 41% of companies with one large owner, have two or more owners with more than 10% of the voting rights each one. These authors find that the cash-flow rights of the second owner are associated with higher valuations.. These findings confirm previous research by Volpin (2002), who finds that the market value of Italian listed companies is higher for companies with a voting syndicate than for companies with a single large shareholder.

However, the presence of multiple large shareholders may harm company value. Large outside shareholders do not always effectively monitor a firm's controlling shareholder. In fact, these blockholders may collude with the controlling shareholder to divert corporate resources for their own interests. Burkart et al. (1997) argue that having a very large shareholder is value

[2] Maury and Pajuste (2005) assume that the assumption that the marginal cost of stealing increases with the number of coalition partners implies that the simple presence of multiple large shareholders should have a positive effect on firm value. But Benjamin Maury, Anete Pajuste (2005) argue that certain coalitions can actually reduce the marginal cost of stealing either by increasing the voting power of the coalition, or by adding extra knowledge and resources for hiding the diversion of profits. For these two reasons, Benjamin Maury, Anete Pajuste (2005) assume that the marginal cost of stealing is lower in the multiple blockholder case as compared to the one blockholder case.

increasing if he or she performs an effective function in monitoring and disciplining management. This shareholder might, nevertheless, collaborate with managers in expropriating outside shareholders. In companies with several large shareholders, the same problem could arise if these shareholders delegate their voting rights to one shareholder. That is, the delegated shareholder may collude with managers and then share private benefits between the whole controlling group and management.

We now examine whether the presence, size and identity of other big owners – other owners with 10% or more of the cash-flow rights – influence the relationship between value and the stake of the largest shareholder.

Hypothesis 6: The presence of a second large shareholder positively affects the company`s value.

Now, we examine whether the identity of the second owner relative to the identity of the largest owner matters for corporate valuation. When control is in the hands of a single person or family, the control group is, most probably, more cohesive and more cooperative in extracting private benefits, relative to companies that are controlled by several business partners. Thus, all other things equal, family companies should exhibit higher private benefits. Volpin (2002) and Bennedsen and Nielsen (2007), show that family ownership discounts firm's relative valuation (Tobin's Q), presumably because of relatively large private benefits consumption.

The intuition is as follows. Large owners may collude, or the second largest owner may exert a check on expropriation by the largest owner. Does the impact of the cash-flow rights ($BLOCK_{2,it}$) depend on the identity of this second owner and whether the identity is different from the largest controlling owner?

As Hansmann (1996) showed, homogeneity of the interest of the owners is one of the main determinants of the efficient ownership structure, since it aligns the interest and incentives of the owners in achieving the company's goal. Moreover, the homogeneity of interests decreases the costs of collective decision-making and makes the formation of coalition more likely since owners are pursuing the same goal while investing in the company. Maury and Pajuste (2005) show that the incentives to collude with or to monitor the controlling shareholder are affected by the type of the individual owner.

Hypothesis 7: The effect of the largest shareholders ownership on company value is moderated by the shareholder identity and the homogeneity of shareholder interests.

$$VALUE_{it} = \alpha_0 + \beta_1 VALUE_{i,t-1} + (\beta_2 + \square_1 HOMOGEN)BLOCK_{1it} + \beta_3 (BLOCK_{1it})^2 + \beta_4 BLOCK_{2,it} + \beta_5 DEBT_{it} + \beta_6 ROA_{it} + \beta_7 SIZE_{it} + \beta_8 TANG_{it} + \zeta_t + \lambda_i + \theta_p + \varepsilon_{it}$$

HOMOGEN = 1 when the first and the second largest shareholder are of the same investor group and 0 otherwise.

Exploring this issue in family companies we decided test results of Pajuste and Maury (2005) who suggest that the control contest often provides beneficial effects on corporate value, however, if both the largest shareholder and all other reference shareholders are family members, the effect on company value is negative. The possibility of agreement among family shareholders is more feasible because these coalition faces lower costs of extracting private benefits. On the contrary, when the others shareholders are non-family members (we explore institutional investors) the connivance of interests becomes more difficult as they are under stricter supervision by regulatory authorities that, consequently, increase the costs of extracting private benefits.

We decided explore the following hypotheses:

Hypothesis 8a: In family companies, a family as a second shareholder has a negative influence on the company value

Hypothesis 8b: In family companies, an institutional investor as a second shareholder has a positive influence on the company value

2.4. Legal system

The relation between *country-level* (or external) governance and *company-level* (or internal) governance is a subject that is much discussed in the current academic literature. On one hand, some company-level governance mechanisms might be *substitutes* for weak country-level governance. La Porta et al. (2002) contend that in countries with weak investor protection rights, stockholders may have to rely on other means of protection. Research suggests that strong shareholder protection laws increase company valuations (Claessens et al., 2000; La Porta et al.,2002).

When investor protection is imperfect, a conflict of interest arises between the controlling shareholder and outside shareholders. The controlling shareholder values private benefits more under weaker investor protection and is able to derive greater private benefits in larger companies (Baumol, 1959; Williamson, 1964; and Jensen, 1986).

A widespread view on outside blockholders as a mechanism of corporate governance is that they are particularly important when legal protection of shareholders is weak (Berglöf and Pajuste, 2003; Berglöf and Von Thadden, 2000; Shleifer and Vishny, 1997). When legal protection is bad a large outside shareholder is a very important (if not the only) instrument of restricting managerial opportunism. This result is further supported in recent studies (e.g., Volpin, 2002; Lemmon and Lins, 2003; and Nenova, 2003).

Durnev and Kim (2007) and La Porta et al. (2002) find evidence of higher company valuations in countries with better protection of minority shareholders and in companies with higher cash-flow ownership by the controlling shareholder. Lins (2003) finds that the way blockholders impact company valuation is significantly influenced by the type of shareholder protection rules in different countries.

Bearing in mind that the conflict between large shareholders is more severe in civil law countries than in common law countries, we formulate our eighth hypothesis as follows:

Hypothesis 9: The effect of the largest shareholder's ownership on value is moderated by the extent of legal protection

Finally, we go to test the tesis defended for Pagano and Roell (1998). These authors argue that the feasibility of corporation behaviour between several large shareholders depends on the regulatory environment. They claim that a higher legal investor protection hinders collusion between the entrepreneur and potencial monitoring shareholders. Thus the presence of large monitoring shareholders may benefit dispersed shareholders in environments with better legal protection, and harm dispersed shareholders in countries with weaker shareholder protection.

Bennedsen and Wolfenzon (2000), for the opposite, argue that control structures with multiple shareholders may be the most efficient ownership structure in environments with poor shareholder protection. Specifically, they state that "the greater the cash flow possessed by the controlling coalition, the more this coalition internalizes the costs of its actions".

We anticipate that the presence of the second large shareholder is more positive in the environment with better legal protection. This leads us to formulate the following hypothesis.

Hypothesis 10: The effect of the second largest shareholder's ownership on company value is moderated by the extent of legal protection.

2.3. Company characteristics

Finally, we will also evaluate if the relationship between cash flow rights of the largest shareholder and companies` value identified in the literature depends or not on the company characteristics.

Several empirical studies that model company performance confirm the existence of potential interactions between governance mechanism and growth opportunities[3]. That variable controls for firm's investment opportunities. Klapper and Love (2004), Black et al (2003), Himmelber et al (2002), and Durnev and Kim (2007) among others use some type of average growth rate of sales.

In this paper, we argue that as investment opportunities exist, cash flow rights do not increase company value. When investing resources into positive NPV projects rather than transferring resources, the controlling shareholder has chances to increase future wealth in proportion to future cash flows. Thus, the controlling shareholder will wait for future benefits and expropriate less now. Moreover, the investment opportunity can be treated as an opportunity cost of expropriation. When an investment opportunity is present, the cash flow right does not matter so much because the controlling shareholder has more opportunity costs in expropriation,[4] [5] Lemmon and Lins (2003). Therefore, the positive relationship between cash flow rights and company value exists only when investment opportunities are not available.

Given the different magnitude and types of agency costs between high-growth and low-growth companies[6], we expect the effectiveness of ownership concentration to vary with growth opportunities. In particular, if agency problems are associated with greater underinvestment or information asymmetry (a common problem in high-growth companies), we expect ownership mechanisms that mitigate these kinds of problems to be more effective in high-growth companies (Smith and Watts, 1992; and Gaver and Gaver, 1993). However, if, as argued by Jensen (1986), agency problems are associated with conflicts over the use of free cash flow (a common problem in low-growth companies), we expect ownership mechanisms that mitigate such problems to play a more important role in low-growth companies (Jensen, 1986).

As for Doidge et al. (2009) has done, we control for growth opportunities using as proxi: sales growth. We expect that controlling shareholders will be more likely to forgo private benefits of control if the need for external financing to fund growth opportunities is greater.

We apply a dummy variable for companies with highly growth when the value is higher than the industrial asset mean at each 3-digit code, each year. The dummy variable will be interacted with the ownership structure variable in order to ascertain whether the controlling owner has a more positive effect on value. We use sales growth rather than earnings growth or total assets growth because earnings and assets growth are more volatile and sensitive to accounting manipulation (Rajan and Zingales, 1995; Booth et al. 2001).

In this sense, our next hypothesis is:

Hypothesis 11: The importance of the owndership structure as a company value's conditionant depends on the type of the agency conflicts.

[3] For example, McConnell and Servaes (1995) find that the relationship between firm value and leverage is negative for high-growth firms and positive for low-growth firms. Lasfer (2002) points out high-growth firms (low-growth firms) rely more on managerial ownership (board structure) to mitigate agency problems.

[4] Treating the Asian financial crisis as a negative impact on investment opportunities, they find that cash flow rights did not affect stock returns before the crisis, but did have a strong positive impact during the crisis for East Asian countries.

[5] Chan, Hu, and Wang (2006) document (Taiwanese companies) a non-linear relationship between cash flow rights and firm value. When firms have positive NPV projects and cash flow rights are low, cash flow rights positively effects firm value. For example, for industries with high sales growth, the Tobin`s q is not associated with cash flow rights. On the contrary, for industries with low sales growth, the Tobin`s q increases with cash flow rights. Besides, for industries with high sales growth, the positive relationship between Tobin`s q and cash flow rights exists when cash flow rights are relatively low, but disappears when cash flow rights are relatively high.

[6] In the underinvestment problem, managers may decide to pass up positive net present value projects since the benefits would accrue to debt-holders. This is more severe for firms with more growth-options (Myers, 1977). Asset substitution problems, which occur when managers opportunistically substitute higher variance assets for low variance assets, are also more prevalent in high-growth firms due to information asymmetry between investors and borrowers (Jensen and Meckling, 1976). High-growth firms, though, face lower free cash-low problems, which occur when firms have substantial cash reserves and a tendency to undertake risky and usually negative NPV investment projects (Jensen, 1986).

DUMMY = 1 if the company as a growth rate above de median of the sector and zero otherwise.

Finally we aim to study whether the impact of ownership concentration on company value is different depending on the age of the company was already suggests by Morck, Shleifer and Vishny (1988). The reason to classify companies according to company age is associated with two recently investigated issues, i.e. the sucession decision inside family corporations and the generation of the family controlling or running the company. Overall, the aforementioned investigations and results indicate the importance of taking into account company age when studying the ownership-value relationship, particulary in the case of family companies. Therefore, we formulate the following hypothesis:

Hypothesis 12: The relationship between ownership concentration and company value is stronger in older companies than in younger ones

3. DATA, ESTIMATION METHOD, AND VARIABLES

In this section our objective is to specify an appropriate functional form for the relationship between the ownership structure of companies and their value, which will enable us to test the hypotheses.

3.1 Data

To examine the relationship between ownership structure and valuations, we use data on 1079 companies across 13 countries in Western Europe (Belgium, Denmark, Finland, France, Germany, Greece, Italy, Netherlands, Portugal, Spain, Sweden, Switzerland, and the United Kingdom) extracted from an international database, Amadeus. The AMADEUS database is published by Bureau van Dijk Electronic Publishing (BvDEP), one of the world`s leading electronic publishers of business information. This database gives us three types of information: First, the number of outstanding shares and its market price are needed to calculate the market value of the company (i.e., the dependent variable of our models). Second, the distribution of the firm's equity among its shareholders, this information allows us to determine the level of ownership concentration and the identity of the largest shareholder, which we need to test our hypotheses. And third, the companies' financial statements are needed to calculate a set of control variables that will enter the right-hand side of our models.

Additionally, international data such as the growth of capital goods prices, the rate of interest of short term debt, and the rate of interest of long term debt, have been extracted from the Main Economic Indicators published by the Organization for Economic Cooperation and Development (OCDE).

For each country we constructed an unbalanced panel of non-financial companies whose information was available for five consecutive years from 2002 to 2006. This strong requirement is a necessary condition since we lost one-year data in the construction of some variables (see, for instance, Appendix C), we lost another year-data because of the estimation of the model in first differences, and four consecutive year information is required in order to test for second-order serial correlation, as Arellano and Bond (1991) point out. We need to test for the second-order serial correlation because our estimation method, the Generalized Method of Moments (GMM) is based on this assumption.

In order to isolate or separate/disconnect the impact of large shareholders in the value of the companies, we will consider some control variables. According to literature, the factors that affect companies' value and simultaneously ownership structure include industry, size, leverage, profitability, and intangibility assets.

The time period of our study is also restricted by the type and the quality of information needed to test the hypotheses proposed. Particularly, our study period ranges from 2002 to 2006

since these are the years for which we were able to obtain sufficient ownership data from AMADEUS.

We restrict our analysis to non-financial companies because financial companies have their own specificity. The structure of the samples by number of companies (the number of observations per country is proportional) in Table 1. As shown in Table 1, the resultant unbalanced panel comprises 1,079 companies and 5,395 observations.

3.2 Estimation method

Our models have been estimated by using the panel data methodology so as to avoid several problems emphasized by previous literature. Specifically, we use the Generalized Method of Moments which allows us to eliminate unobservable heterogeneity and to control for potential endogeneity.

We therefore resort two alternative methods to address the endogeneity problem. First we include the lagged dependent variable on the right hand side. The lagged dependent variable filters out most of the effect of missing variables, which will affect *yt* and *yt-1* in equal measure.

Our preferred estimator is the Generalized Method of Moments (GMM) system estimator (Arellano and Bover, 1995; Blundell and Bond, 1998). The GMM system estimator combines a set of first differenced equations with equations in levels. Instrument for endogenous variables are internally generated: lagged first differences of the endogenous variables are used as instruments in the levels equations while lagged levels dated t-1 and earlier are used as instruments in the first-differenced equations. This gives consistent parameters in short panels, provided the instruments used are valid. It also displays good finite sample properties even when variables show persistence (Blundell and Bond, 1998).

The consistency of the GMM estimates depends upon the assumption that there is no serial correlation in random disturbance ε_{it} so that $\Delta\ \varepsilon_{it}$ should indicate significant negative first order serial correlation and no second order serial correlation.

We checked for the potencial misspecification of the models. First, we used the m_2 statistic, developed by Arellano and Bond (1991), in order to test for the lack of second-order serial correlation in our models. Second, we used the Hansen *J* statistic of over-identifying restrictions in order to test for the absence of correlation between the instruments and the error term. Tests for serial correlation in the first differenced residuals, denoted as m1 and m2, are reported in the results. The values obtained allow us to confirm the validity of the instruments used and the absence of second order correlation.

Since both governance and performance may reflect industry factors, we include industry dummies based on 3-digit (different industries were in different business cycle positions, faced different regulatory frameworks and, in turn, had very different growth opportunities). The regressions include dummies to account for: (i) time effects, $\boldsymbol{\zeta_t}$ control for common macroeconomics shocks; (ii) $\boldsymbol{\lambda_i}$ control unobserved time-constant individual heterogeneity of companies (since companies are heterogeneous, there are always characteristics influencing company value that are difficult to measure or hard to obtain, Himmelberg, Hubbard, and Palia, 1999); (iii) θ_p are country dummy variables standing for the country-specific effect, which are necessary in that our models are estimated using data from several countries; finally $\boldsymbol{\varepsilon_{it}}$ is a white-noise error.

In our basic empirical model the market value of company is regressed against ownership concentration of the two largest shareholders.

$$VALUE_{it} = \alpha_0 + \beta_1 VALUE_{i,t-1} + \beta_2 BLOCK_{1it} + \beta_3 (BLOCK_{1it})^2 + \beta_4 BLOCK_{2it} + \beta_5 DEBT_{it} + \beta_6 ROA_{it} + \beta_7 SIZE_{it} + \beta_8 TANG_{it} + \zeta_t + \lambda_i + \theta_p + \varepsilon_{it}$$

3.3 Variables

Dependent Variable

The value is measured as the ratio of market value of equity to book value of total assets. The appeal of this measure is that it reflects the current value of expected future performance. The natural log transformation was performed in order to reduce problems of heteroscedasticity???) OR To correct for a risk-skewed distribuition of the company value variable, we use log values. To reduce the impact of extreme values, I censor the value at the 99th percentile (McConnell and Servaes, 1990 delete companies with Q ratios larger than 6; Lyns (2003) censors Q at the 1st and 99th percentiles. VILLALONGA

Control variables

GMM estimation, which includes the lagged dependent variable and the addition of fixed time and company effects, filters out an unusually high amount of company-specific heterogeneity that may influence both blockholder ownership and company value, and this is a key advantage of the present study. However, to ensure the robustness of the results, I include additional control variables similar to those adopted by Morck, Shleifer and Vishny (1988); McConnell and Servaes (1990, 1995); Laeven and Levine (2008); Maury and Pajuste (2005) and Pindado and Chabela (2004) and control for several factors: company size, asset intangibility, profitability and financial leverage.

Other than ownership structure a firm's value may be affected directly or indirectly by other external factors. These factors may be specific to the nature of company, the industry it belongs to or the market it operates in. Omission of these factors may lead to spurious relation between company value and ownership structure.

Industry

Leibenstein (1966) and Hart (1995) argue that product market competition has a disciplinary effect on managerial behavior. Shleifer and Vishny (1977) suggest further that product market competition is perhaps the most effective mechanism to eliminate managerial inefficiency. Moreover, as argued by Carlin and Mayer (2003) different ownership structures may be appropriate for different types of activities. They suggest that companies in high technology industries may require more dispersed ownership structure while concentrated ownership ensuring commitment and long-term investment may be suited to more 'mature' industries.

Size[7]

The effect of this variable in the value of the companies is not clearly. On the one hand, larger companies have a greater ability to absorb risks and can take advantage of economies of scale. Bernard and Wagner (1998) argue that larger companies generally have lower average or marginal costs, suggesting that profitability increases with company size. A positive relation between value and dimension would be therefore waited. On the other hand, company size is expected to have a negative effect on company value as larger companies are, presumably, in a mature stage of their life cycle.

Furthermore in the existing empirical studies ownership concentration tends to be negatively affected by company size (Demsetz and Lehn, 1985, Himmelberg et al. 1999)[8]. This result reflects probably wealth limitations and the concern with risk diversification[9]. But size is also

[7] Sales and assets have been extensively used as measures of firm size. Some authors provide good insights why this variable is a good control one. For instance, Klapper and Love (2004) argue "the effect of size is ambiguous (on governance) as large firms may have greater agency problems and, therefore need to compensate with stricter governance mechanisms." Durnev and Kim (2007) suggest, "Because larger firms tend to attract more attention and may be under greater scrutiny by the public, size may affect governance structure." Himmelberg, Hubbard, and Palia (1999) state "Firm size has an ambiguous effect a priori on the scope of moral hazard." They argue that large firms are more prone to monitoring and agency costs so (managerial) ownership should be greater. But, on the other hand, large firms may have greater economies of scale in monitoring by large stockholders (management in their case) and by rating agencies that may lead to a lower need of large ownership stakes in the firm.

[8] This negative effect of firm size on inside and outside ownership concentration has been widely supported in, for instance, Holderness and Sheehan (1988), Bergström and Rydqvist (1990), Lange and Sharpe (1995), Cho (1998), Denis and Sarin (1999), Holderness, Kroszner and Sheehan (1999) and Demsetz and Villalonga (2001).

[9] A dispersed structure with many shareholders can better share the risk of a large sized firm and, vice versa.

sometimes considered as a proxy for managerial discretion (Himmelberg et al., 1999); in that case we expect that size will positively affect ownership concentration. Size may also be viewed by potential shareholders as a proxy for reputation. Demsetz (1983) suggests that small companies are more likely to be closely-held suggesting a different governance structure than large companies. Following Harvey et al. (2004), we also include a company size.

Leverage

A large capital structure literature emphasizes the role of debt in reducing agency conflicts between managers and shareholders (i.e. Grossman and Hart (1982), Jensen (1986), Stulz (1990), Hart (1995), Hart and Moore (1995)). Jensen and Meckling (1976) put forward one of the earliest analyses in which debt refrains from insiders' opportunistic behavior by compelling them to meet debt servicing requirements. Jensen (1986) further suggests that the fixed payments associated with debt reduce a firm's free cash flow, and effectively limit management's ability to waste corporate resources to benefit themselves. Moreover, the issuance of external debt also results in monitoring by bondholders, other lenders, investment bankers, and bond rating agencies.

Finally, Faccio, Young and Lang (2002) develop a view of debt where it acts as an expropriate device. For a given amount of assets, debt leverage limits the number of voting rights necessary to achieve control. Hence, we do not have a clear prediction on the relation between leverage and company value.

Consistent with Bebchuk, Cohen and Ferrell (2004), Graham, Lang, and Shackleford (2004), and Khanna and Tice (2005) we compute leverage as non current liabilities divided by total assets. We use in the numerator the long term debt since most of the arguments in financial theory are related to this type of debt (see, for instance, Miguel and Pindado, 2001 and, more recently, DeAngelo and DeAngelo, 2006).

Intangible assets

Asset intangibility ($INTANG_{it}$) is the ratio of intangible assets divided by total assets. Companies' characteristics, such as intangible assets, provide opportunities for rent seeking especially through excessive leverage (Williamson, 1988). In this context, Williamson (1988) suggests that the uncertainty of intangible assets provide opportunities to controll owners to expropriate companies' value.

In contrary direction, Klapper and Love (2004)[10] have argued that corporate governance is more relevant in companies that employ significant amounts of intangible assets. In these companies the paper of monitoring of the large shareholders would be still more important.

Companies with lower asset tangibility presumably have a higher proportion of intangible assets (e.g., human capital) generating the cash-flows. Therefore, we expect a positive relation between asset intangibility and value.

Profitability

*A*s a measure of profitability, I include the return on assets, labeled *ROA*, which is calculated as Earnings before interest, taxes, depreciation and amortization (EBITDA) divided by the total assets (TA).

4. RESULTS

4.1 Descriptive results

A first approach to the results can be achieved from the preliminary descriptive analysis of the variables. To address the wide scope of international differences, table 1 reports the mean

[10] Klapper and Love (2004) argue that the composition of a firm's assets will affect its contracting environment. Intuitively, it is easier to monitor and harder to steal fixed assets (e.g., machinery and equipment) than "soft" capital (e.g., intangibles and R&D capital).

value, the standard error, and the maximum and minimum values of the variable value. As can be seen, the 13 countries show significant differences in terms of company`s valuation.

Table 2 presents the mean value, the standard error, and the maximum and minimum values of the main variables.

Concerning the ownership structure (values that consider all the shares without restrictions on their size), civil law companies exhibit higher ownership concentration, consistent with Franks and Mayer (1997) and Oscar Foronda et al. (2008). Thus, while the main shareholder in these countries owns on average 41.18% of the equity, this proportion falls to 16.96% in The United Kingdom. This same pattern is observed for the proportion of the second and the third largest shareholders: for the second, a proportion of 17.9% versus 9,67% and for the third a proportion of 8.36% versus 6,97%.

To account for the importance of controlling shareholders, table 3 provides a more detailed analysis of the identity of the two largest shareholders. Table 3 shows that only in 23.34% of the company-years observations the controlling shareholder is a family. Thus, there is a clear preference for indirect holdings or through non-financial companies (31.01% company-years observations) or by institutional investors (29.97% company-years observations). Only in 13.3% of the sample the controlling shareholder is the state.

Tables 4 and 5 provide insights into the ownership structures and the distribution of shares among large shareholders. First of all, and consistent with numerous previous studies (Bloch and Kremp, 2001; Faccio and Lang, 2002), Western Europe appears as a concentrated ownership region. Only 10.94% of the observations are widely held companies (i.e. companies without any large shareholder holding at least 10% of the cash flow rights). Secondly, we notice a complex ownership structure (with at least two large blockholders) in 63.74% of the sample. Excluding widely held companies, the probability that the first shareholder is alone is 28.43%. That is to say that a second large shareholder is a possible monitor of the main shareholder in 71.57% of the company-years observations. Complex ownership structures are mainly structures with two large blockholders (in 55.15% of the cases). Complex ownership with three shareholders it is not so common, only 44.85% of the cases.

It seems also important to emphasize that the two largest shareholders have, on average, over 50% of cash flow rights. In this context, it is not strange that the third shareholder rarely appears as significant in the regressions. As we will see later in our case, the shares of the third largest shareholder are statistically significant and positive.

4.2 Regression results

Our regressions models explain company value as a function of ownership concentration and other control variables. We will first comment on the evidence obtained from the estimation of the value models specified to study the relationship between ownership concentration and value (i.e. models (1), (2) and (3).

The results[11] (model 1) confirms hypothesis 1a and provide support for the general functional form of the relationship between company value and ownership concentration as suggested by McConnell and Servaes (1990); Gedajlovic and shapiro (1998); Thomsen and Pedersen (2000); Short and Keasey (1999), and Pindado et al. (2004).

A nonlinear relationship between ownership concentration in the hands of the first large shareholder and company value indicates that, at high levels of ownership concentration, the benefit of concentrated ownership (as a clear consequence of the more efficient monitoring provided by concentrated shareholding) is outweighed by the negative effects. Among the negative effects, the expropriation of small shareholders by large shareholders is noteworthy.

[11] The ownership concentration ($Block1_{it}$, $SBlock1_{it}$ and $Block2_{it}$) has statistically significant coefficients in column I. The coefficient on the variable $Block1_{it}$ is positive while that on the variables $SBlock1_{it}$ and $Block2_{it}$ are negative.

Our results of models (2) and (3) support the hypotheses 1b and 1c, that the concentration of ownership has the same relationship with the market value of the companies, regardless of considering only the stakes of first large shareholder or the joint participation of two or even the three main blockholders. We can conclude that for high stakes of the first larger shareholder, the market did not anticipate that the second or even the second and third together, exercise control of their behavior before they advance together in negative coalitions for the interests of minority shareholders.

Model (4) relates to Hypothesis 2 by testing the effect of the ownership in the hands of the first, second and third largest shareholder. The results confirm the coefficients for the first and second large shareholders. The results also indicate a positive impact for the stake of third largest shareholder (the continuous variable). The explanation perhaps lies in the fact that whenever there is a third shareholder, the potential negative coalitions internalize such a large portion of the costs of expropriation that these behaviors become uninteresting.

Furthermore, the control variables included in the analysis enter the regression significantly. Focusing on the control variables, we obtain a negative influence of the leverage, and the company size, on value. Therefore, companies with lower leverage and smaller companies have higher value. Maury and Pajuste (2005); Laeven and Levine (2008) and Demsetz and Villalonda (2001) find a negative effect of the debt on value. The negative influence of company size is also observed by Maury and Pajuste (2005); Laeven and Levine (2008) Hermalin and Weisbach (1991), Agrawal and Knoeber (1996) and Bhagat and Black (2002), among others. The coefficient of TANG has the sign expected and confirms evidence from Maury and Pajuste (2005) and Laeven and Levine (2008). Finally, the coefficient of ROA is positive and it only emphasizes that analyzing all the regressions of this paper is the coefficient with a statistically more significant with value.

Our next purpose is to isolate the cases of majority ownership (i.e., Block1 $\geq$ 50 per cent) in which the contest from the other reference shareholders is nearly irrelevant. When we explore what the relationship between the value of companies and a majority shareholder (Model 5), our results reveal a positive relationship. This result indicates that only for very high levels of cash flow rights, the shares of capital the first largest shareholder have a negative effect on the market value of the companies. Our results line up with, for example, Volpin (2002) for the Italian listed companies. This author documents that when the first largest shareholder owns less than 50% of the firm's cash flow rights, the Q ratio tend to be lower and the CEO turnover is less sensitive to performance.

The literature suggests that the relation between concentrated ownership and company value may be sensitive to the type of the ultimate controlling shareholder. The importance of the identity of the large shareholder is also recognized in the earlier generation of ownership structure research. As it is pointed out in Demsetz and Lehn (1985), different owners vary in their objective functions and in their ability to monitor managers.

The companies in our sample have several different types of large owner. The behaviour of largest owners, that are organizations run by agents, may differ from that of largest owners that are families, since there is a much clearer relationship between the market value of equity and the wealth of the largest owner in the latter case. To test whether the effects of largest owners' cash-flow rights differ by type of owner, we use interactive variables to identify the largest owner as either, a family, an institutional investor, a non-financial company or the state.

One of the main objectives of this paper was to find evidence that helps to clarify whether family companies differ from the non-family companies. Empirical evidence is ambiguous. Renneboog (2000) counters that large shareholding interests in the hands of individual and family investors are linked to a greater probability a board may be restructured, which could be indicative of more effective control. Cronqvist and Nilsson (2003) find out that the presence of family ownership in a company has negative effects on its value. Daily and Dollinger (1992) evidenced that family-owned and managed companies display higher performance due to the unification of ownership and control. Daily and Thompson (1994) found that family-managed companies behave no differently than other professionally managed companies with respect to strategic posture. Anderson and Reeb (2002) show, using Standard & Poor's 500 companies,

that family companies do outperform other non-family companies. Anderson, Mansi and Reeb (2002) argue that the desire of family managers to preserve their reputation gives them incentive to enhance the firm's performance.

The involvement of the controlling shareholder in the company management, however, may not necessarily be favorable to minority shareholders. The large leeway of the owner-manager over the company's decision may lead to the firm's objectives' subordination to family's objectives, and hence may hamper firm's performance. Maury and Pajuste (2002) found that a family member in a CEO position has a negative effect on the company value.

The results presented in Columns 1,2,3,4 and 5 of Table 9 suggest that the concentration of family ownership, as opposed to institutional ownership, caused a positive statistically significant impact on the level of value. Focusing on column 1 of the table mentioned above, we can see that the coefficient of the family is positive and significantly related to value (β = 0.00459, t = 0.041). The estimated coefficients provide evidence that the family companies are more valuable than non-family companies. The empirical evidence supports the hypothesis 3 and the potential benefits characteristics of family shareholders, pointed out in section 2. The results indicate that the impact of ownership concentration on company value is stronger when the largest shareholder is a family than when it is not. Our results are consistent with previous empirical evidence, particularly with the evidence from Western Europe (Maury, 2006; Barontini and Caprio, 2006, and Andres, 2008) and suggest that the potential benefits associated to family control exceed its potential costs regardless of the blockholder effect.

To investigate whether a similar curvilinear relationship between Value and Blockholders, as reported by McConnell and Servaes (1990), is also inherent to our sample, we reestimate the regression equation reported in Column 4 by including a quadratic term of $FD*Block1_{it}$, denoted as $SFD*Block1_{it}$. To avoid problems of multicollinearity, companies are divided between family and non-family companies; for this non-family companies, the variables related to the first large shareholder are $NFD*Block1_{it}$, and $SNFD*Block1_{it}$. The results in Column 4 confirm a curvilinear relation between $FD*Block1_{it}$, and value (hypothesis 4). Company valuation first increases and then decreases as the fraction of cash flow rights owned by the families become larger. We can confirm that the same behavior is observed for the largest shareholders of non-family companies.

The effects of institutional ownership and banks *per se* on company performance have also been shown to be mixed. Morck, Nakamura and Shivdasani (2000), and Goergen, Renneboog and Correia da Silva (2005) find a negative relationship. McConnel and Servaes (1990) Boehmer (2000), and Gorton and Schmid (2000) find a positive relationship. Prowse (1995); Zoido (1998) and Cronqvist and Nilsson (2003) find no systematic relationships. This leaves the subject still open to academic debate. Pedersen and Thomsen (2003) document a positive effect on company when the large owner is a financial institution or another corporation.

Ours results strongly contradict the preponderance of the prior evidence that institutional investors are effective monitors of corporate managers. Additionally, our regression results document a negative and significant contemporaneous association between institutional ownership and company value (table 9, column 5).

Institutions are not a monolithic body. The incentives, culture and regulations governing banks, insurance companies, mutual funds, and pensions vary widely. Here we explore these differences in regressions not presented in the tables and the results show that even restricting the type of institutional investors to the Banks, (among the institutional investors are those with stronger ties with affiliated companies and soon more and better information management of these as they often are among the lenders of debt capital, and because they have greater powers to assess the quality of the decisions of managers) it still continues to be observed a negative relationship between this type of investor and value.

As far as the second shareholder effect is concerned, recent research has shown that a firm's value is positively affected by the ability to challenge the largest shareholder (Bloch and Hege, 2001; Gomes and Novaes, 2005; Edwards and Weichenrieder, 2004). The expropriation of minority shareholders is less common in companies where control is most contested. Shleifer and Vishny (1997) also comment that monitoring by large minority shareholders is effective

only in countries with good investor protection. In countries with poor investment protection, only majority ownership would be effective.

The impact of shareholders' alliances on company value is an empirical challenge. In order to address this issue, we build our analysis on previous literature that connects company value and the cash flow rights of the second large shareholder.

Contrary to most models that predict a positive effect of a second large shareholder on the value (contestability), our results begin to suggest a negative effect. When we explore a linear relationship, the coefficients are always statistically significant and negative. These results suggest an inverse relationship between the size of the stakes of the second large shareholder and the value.

In order to deepen the evidence and trying to reconcile these results with the bulk of previous empirical evidence that stresses a positive effect of the second large shareholder, we propose (column 1 of table 8) at first a variable of interaction between the shares of the first shareholder and the fact that there is a second large shareholder with the aim of realizing the existence of the latter expands or reduces the previously highlighted positive relationship between the cash flow rights of the largest shareholder with the value. Second, we introduced in isolation a dummy variable (dummy = 1 if the company has a second large shareholder with cash flow ownership above 10% and zero otherwise) in the regression (column 2 of table 8).

In any of these regressions, the coefficients are positive and statistically significant and thus support our hypothesis 6. We can therefore state that the positive relationship between the cash flow rights of the first large shareholder and company value is higher when there is a second large shareholder, and the presence of the second large shareholder contributes positively to the value of companies. Our results tell us that the mere existence of a second reference shareholder allows him to control the largest shareholder. Therefore, the largest shareholder's capacity to extract private benefits will be reduced, as also emphasized by Bennedsen and Wolfenzon, 2000; Bloch and Hege, 2001; Edwards and Weichenrieder, 2004; and Gomes and Novaes, 2005.

It is now important to try to find out the key to this puzzle. Since the existence of a second shareholder has a positive effect on the value of companies, how can we explain that the size of these holdings has a linear inverse relationship with the same variable value? With this aim, we will explore if the stakes of the second large shareholder also have a nonlinear relationship with the value of companies. It is worth noting that, as far as we know, there is no prior evidence supporting this view, and providing empirical support to this issue is thus one of the major contributions of this paper. The results of the regression in column 3 (table 8) show empirical evidence that the stakes of the second large shareholder have a negative effect on value for the reduced cash flow rights and become positive on high cash flow rights.

This might be our puzzle's key. For small stakes of the second shareholder, the market anticipates that the presence of a second blockholder allows controlling coalitions with lower stakes to be formed. This could stimulate blockholders expropriating actions as the controlling coalition internalizes a lower proportion of the expropriation costs and thus, could negatively impact on the company`s performance. For stakes higher of the second largest shareholder, the coalitions internalize a greater proportion of the costs of expropriation and thus their interests tend to align with the interests of the minority shareholders and thus producing a positive impact on the value of companies.

To examine to what extent the type of the second large shareholder makes the existence of contestability or collusion between the two large shareholders, we differentiate between family and Institutional investors.

We find (table 9, column 2) that a higher cash flow rights held by another family owner penalizes company value in family-controlled companies, whereas a higher stake held by institutional investor improves company value in family-controlled companies (table 9, column 3). These results support the hypotheses 8a and 8b and suggest that the incentives to collude with the controlling shareholder or to monitor the controlling shareholder are affected by the type of the individual owner. Moreover, these results suggest that different types of the individual shareholders seem to have different costs for extracting private benefits. These results, allow us to posit that, all else equal, the family-family coalition may plausibly be

associated with an alignment of interests of the two largest shareholders (i.e., families), who might find it mutually valuable to collude to extract divisible private benefits of control. Overall, these results confirm the evidence from Maury and Pajuste (2005), and Félix Iturriaga and Óscar Foronda (2008).

In an attempt to confirm these results the model 5 of Table 8 explores whether the fact that the two main shareholders are the same type (any of the type) influences positively or negatively the market value. The results show a negative relationship and thus confirm the results found for the family companies and the hypothesis 7. These results strongly confirm that the coalition formed by two shareholders of the same group is more likely to reduce the marginal cost of private benefit extraction than a coalition formed by different types of shareholders.

Using a dummy variables approach, we explore if ownership concentration, as a mechanism of corporate governance, is particularly important when legal protection of shareholders is weak. We investigate the interaction between the cash flow rights of the largest and the second largest shareholders and the variable Anglo by estimating the models presented in Column 4 and 5 of table 7. The results obtained for the estimated coefficients confirm that ownership concentration exercises a distinct influence on a firm's value according to the particular institutional environment. As can be seen in table 7, the coefficient of the cash flow rights of the first large shareholder is lower for companies from the UK ($\beta_2 + 1 = 0.02137$), significantly different from zero, see t) than for companies from other countries ($\beta_2 = 0.05507$). Therefore, it seems that monitoring of the largest shareholders as agency-cost control mechanism, is more important in the context of weak legal environment. The results confirm the evidence presented by Shleifer and Vishny, 1997; Berglöf and von Thadden, 2000; Berglöf and Pajuste, 2003; and Dahya, Dimitrov and McConnell, 2008, in a legal context, unfavorable to minority shareholders, the stake of shares of the first large shareholder are very important, if not the only, instrument of restricting managerial opportunism. Our results confirm the effect substituion between ownership concentration and minority shareholder protection suggested in recent studies such as Denis and McConnel (2003). This result is further supported in recent studies (e.g., Volpin, 2002; Lemmon and Lins, 2003; Nenova, 2003; and Sánches-Ballesta and García-Meca, 2007).

Finally, we explore if the relationship of the second large shareholder and value depends on the legal context by providing a variable of interaction between the dummy variable Anglo and the cash flow rights of the second large shareholder. Empirical evidence suggests that this relationship is positive in the United Kingdom. As can be seen in table 7 column 5, the coefficient of the cash flow rights of the second large shareholder is positive for companies from the UK ($\beta_2 + 1 = 0.00051$), significantly different from zero, see t) being negative for companies from other countries ($\beta_2 = - 0.01182$). The results confirm the emphasis of Vishny and Shleifer (1997) that the effect of monitoring by large minority shareholders is effective only in countries with good investor protection.

The results suggest that the minority shareholders anticipate different behaviors for the second large shareholder if their interests are defended by the legal system. Under a week legal system, the minority shareholders anticipate, as most likely, the possibility that of two major shareholders participate in negative coalitions according to their interests, than the possibility that the second large shareholder has the role to control the behavior of the first shareholder. This result reflects the fact that a weak legal system is less able to determine the second large shareholder to contest the decisions of the largest shareholder which turn apart from the purpose of maximizing the wealth of minority shareholders.

A competing explanation arises from differences in concentration of ownership evidenced in both realities. In the UK, the first shareholder has on average 20.53% to the cash flow rights, while in Continental Europe the average of these stakes is 43.19% (it should also be noted that in the UK it is more prevalent the lack of a second large shareholder than in other countries). This evidence suggests that the first large shareholder in Continental Europe is more sovereign in the conduct of the companies. We can therefore conclude that in a context where the capacity of the second large shareholder to challenge the decisions of the former is much smaller because of the legal system, and in a context where the first large shareholder's stakes are much larger,

this clearly determines that the probability of the formation of the coaligation with different interests of the minority shareholders.

Once the importance of ownership concentration has been corroborated by our results, we go a step forward and investigate whether or not certain company characteristics moderate this relation. We then propose a test of the moderating role played by two features – age and growth. In table 7 we provide evidence of the moderating role played by company age and growth in the ownership-value relationship. The negative coefficient of the interaction between the age of the company and ownership of the first large shareholder support Hypothesis 12 (those can be seen - Table 7 column 2 - this effect is observed both in family companies and in non-family companies). We can therefore state that the relationship between ownership concentration and company value is weaker in new companies than in older ones. This result supports the role of the company age as a moderating variable when analyzing the effect of ownership on value. In other words, the empirical evidence obtained is that the role of the large shareholder is more important in older companies. We think that a plausible explanation might be the fact that although the existence of blockholders is assessed positively by the market, since it has a greater capacity to monitor the managers (the managers are often the family), this capacity is especially appreciated when the companies have more potential agency problems. Older companies tend to have more problems of free cash flow as they have fewer opportunities for investment.

In the case of family companies, such finding is consistent with the argument that when controlling families are directly involved in the company `s management they contribute to solve the classic agency conflict between owners and managers. As companies grow older, this coincidence of interests will be progressively smaller with the arrival of professional managers. In these circumstances, the monitoring blockhoklder effect gets a renewed importance, strengthening the positive relationship between concentration of ownership and the value of companies.

In order to highlight the role played by the investment opportunity, we provide evidence in table 7 about the moderating role played by company growth in the ownership-value relationship. The negative coefficient of the interaction between the growth of the company and ownership of the first large shareholder support Hypothesis 11. We can therefore state that the relationship between ownership concentration and company value is weaker in growth companies, i.e. in companies with more investment opportunities and therefore those companies where the minority shareholders anticipate as less likely the arising problems of free cash flow identified by Jensen (1986).

5. CONCLUSIONS

This study provides important contributions to the literature and policy debates concerning corporate governance implications of various ownership patterns.

The ultimate purpose of this research is to investigate whether the presence, the cash flow rights, and the identity of the two largest shareholders play an effective role on corporate governance. We develop the scope of analysis further than the largest shareholder to the second and third largest shareholders and suggest that the cash flow rights of these last and the type of shareholder justify a deeper analysis (Laeven and Levine, 2005).

This paper also examines how particular investors control (family and institutional investors) impacts on the market company value. Further, we explore the effect of second shareholder identity in family-controlled companies.

Overall, our results give new evidence to the studies commented in Section 2, which in general point out a monotonic relationship between value and the concentration of ownership. Our results highlight the differential role played by the ownership concentration depending on the legal and institutional setting. In civil law countries, the concentration of cash flow rights in the hands of the first large shareholder has a positive relationship with the value stronger than that seen in the UK. In contrast, while in the UK the existence of a second blockholder has a positive effect on the value of companies, there is a symmetric effect in civil law countries. We

can conclude that the legal system provide, in these latter countries, a lower capacity to challenge the decisions of the first, which causes a greater likelihood of negative coalitions for the interests of minority shareholders.

We find robust evidence that the market value increases not only in the presence but also with the cash flow rights of the largest shareholder. Nevertheless, a more accurate analysis reveals that when ownership concentration is too high, company value decreases. This reduction may stem from the risk of expropriation of minority shareholders.

In additional tests, we find that the age and growth of the company moderate the relationship between value and ownership concentration. The ownership concentration is more important in the companies with weaker opportunities of investment and older, i.e., companies with major problems with the free cash flow.

Our results also suggest that the type of the controlling shareholder affect the company value. When the largest shareholder is a family, the ownership concentration has a stronger impact on company value than in both companies with a dispersed shareholder structure and in other companies with a controlling shareholder. This leads to the conclusion that families successfully balance the two agency problems that minority shareholders are exposed to (owner-manager conflicts on the one hand and minority shareholder expropriation by a controlling shareholder on the other hand). When sub-divided into family companies and non-family companies, we find evidence of a nonlinear relationship between ownership concentration in these two groups. Anderson & Reeb (2003) also found that company performance increases until a family has about a third of the ownership; beyond this, performance tends to decrease.

This paper also shows that "who monitors the family" (La Porta, Lopez-de-Silanes and Shleifer, 1999), in terms of external governance, is a crucial issue in determining the effective impact of family blockholdings. Within family companies, the existence of a second blockholder of the same type causes a reduction of the positive effect previously shown for the family business, while the existence of a second large shareholder of institutional investor type reinforces the positive effect of family companies.

According to previous research, we corroborate the rise of conflicts between family shareholders and minority shareholders (Maury, 2006) and show that the contestability of the control of the largest shareholder increases the value of family-owned companies (Lehmann and Weigand, 2000; Volpin, 2002; Maury and Pajuste, 2005). We also find that the formation of a controlling coalition of shareholders can exacerbate the expropriation of minority shareholders if these controlling shareholders are families; on the other hand, the participation of nonfamily shareholders in the controlling coalition of family companies increases the firm's value.

Table1 - Summary statistics of Value (V_{it}) = Market Equity / Total assets

Country	Number of companies	% Companies	Mean	Standard deviation	Minimum	Maximum
Belgium	11	1.0	0.66248	0.38247	0.14160	1.55475
Denmark	40	3.7	0.73302	0.62378	0.00990	3.56807
Finland	45	4.2	0.81251	0.61438	0.08850	3.01186
France	277	25.7	0.76550	0.65485	0.01000	5.20054
Germany	112	10.4	0.70657	0.62473	0.02879	4.92813
Greece	79	7.3	0.73918	0.70390	0.05945	5.47941
Italy	23	2.1	0.68639	0.43931	0.02032	2.36930
Netherlands	60	5.6	0.81420	0.61820	0.05102	5.03526
Portugal	19	1.8	0.47162	0.38247	0.14160	1.55475
Spain	14	1.3	0.67143	0.58756	0.16856	4.48485
Sweden	63	5.8	0.92470	0.68033	0.07365	4.84869
Switzerland	67	6.2	0.84988	0.68663	0.02025	4.34174
U. Kingdom	269	24.9	0.90865	0.65856	0.03694	4.84207
Total	1079	100	0.80201	0.65272	0.00990	5.47941

Table 2 - summary statistics of the variables used in our models.

Variable	Observations	Mean	Standard deviation	Minimum	Maximum
BLOCK1$_{it}$	4805	38.5499	22.6637	10.00000	100.000
BLOCK2$_{it}$	3442	21.0239	12.3045	10.00000	50.0000
BLOCK3$_{it}$	1545	15.2840	7.2922	10.00000	32.1500
BLOCK12$_{it}$	4808	53.5736	29.1366	10.00000	100.000
BLOCK123$_{it}$	4808	58.4882	31.2956	10.00000	100.000
DEBT$_{it}$	5395	0.27056	0.20535	7.49e-06	0.98879
ROA$_{it}$	5395	0.10617	0.08759	-0.40345	32.1500
INTANG$_{it}$	5395	0.20515	0.18423	0.00000	0.98589
SIZE$_{it}$	5395	12.2784	2.78203	5.81711	20.4333
GROWTH$_{it}$	5395	0.06557	0.45460	-0.97896	16.7172
AGE$_{it}$	5395	3.39410	0.95409	0.00000	6.44254

$BLOCK1_{it}$, $BLOCK2_{it}$ and $BLOCK3_{it}$ is the percentage of common shares held respectively by the largest the second largest and the third largest shareholder of the company; $BLOCK12_{it}$ is the percentage of common shares held jointly by the two major shareholders of the company. $BLOCK123_{it}$ is the percentage of common shares held jointly by the three major shareholders of the company. $DEBT_{it}$ is the ratio of the book value of the long term debt divided by the total assets; ROA_{it} is the return on assets; $TANG_{it}$ stands for the intangible assets; $SIZE_{it}$ is the size of the company; $Growth_{it}$ is the sales growth; AGE_{it} is the company`s age.

Table3 - The characteristics and ownership of the two largest (controlling) shareholders

	1º Large Shareholder	**2º Large Shareholder**
Family	**23,34%**	**23,74%**
Corporate	**31,01%**	**24,28%**
Institutional investor	**29,97%**	**36,16%**
State	**13,30%**	**12,50%**
Others	**2,35%**	**3,34%**

Table4 – Ownership concentration:

	BLOCK1	BLOCK2	BLOCK3	BLOCK12	BLOCK123
Cash-flow rights > 10%	89,58%	66,01%	31,51%	89,12%	89,12%
Cash-flow rights > 20%	68,84%	31,10%	4,93%	80,07%	80,07%
Cash-flow rights > 50%	28,16%	------	------	49,69%	55,51%

Tabela5 – Ownership concentration

OWNERSHIP STRUCTURE	PREVALENCE (Observations) N	% Sample	BLOCK1	BLOCK2	BLOCK3
WIDELY HELD	590	10,94%	_____	_____	_____
ONE CONTROLLING SHAREHOLDER	1366	25,32%	37,48%	_____	_____
MULTIPLE LARGE SHAREHOLDERS	3434	63,74%			
Of which 2 Large	1894	35,11%	42,71%	21,12%	
Of Which 3 Large	1545	28,63%	34,39%	20,81%	15,29%
TOTAL	5395	100,00%	38,55%	21,03%	15,29%

Table 6

	I – FLS	II -FASLS	III - FSATLS	IV-TLS	II - MAJORITY
$V_{i,t-1}$	0.38056***(0.05456)	0.19446***(0.06868)	0.42968***(0.05782)	0.36448***(0.07019)	0.45853***(0.06056)
$BLOCK_{1,it}$	0.03461***(0.01282)			0.02683***(0.01052)	0.03214**(0.01481)
$SBLOCK_{1,it}$	-0.00042***(0.00015)			-0.00034***(0.00013)	-0.00056***(0.00021)
$BLOCK_{2,it}$	-0.01036**(0.00457)			-0.01319**(0.00457)	-0.00993**(0.00461)
$BLOCK_{12,it}$		0.02118***(0.00666)			
$SBLOCK_{12,it}$		-0.00024***(0.00007)			
$BLOCK_{123,it}$			0.02481***(0.00789)		
$SBLOCK_{123,it}$			-0.00025***(0.00008)		
$BLOCK_{3,it}$				0.01065*(0.00580)	
MJ_{it}•$BLOCK_{1,it}$					0.01057*(0.006345)
$DEBT_{it}$	-1.05521***(0.25468)	-1.99316***(0.31443)	-0.94782***(0.26017)	-1.35737***(0.31257)	-0.91354***(0.24718)
$SIZE_{it}$	-0.11592***(0.04676)	-0.10776***(0.03384)	-0.10525***(0.03525)	-0.05934***(0.02048)	-0.05885**(0.02566)
$INTANG_{it}$	0.47896*(0.26055)	0.81424***(0.21784)	0.57062***(0.22509)	0.48579***(0.15096)	0.47396*(0.26531)
ROA_{it}	0.99781***(0.307279)	0.68626**(0.31844)	1.18495***(0.33124)	1.77222***(0.35665)	2.18425***(0.79827)
Constant	1.3749*(0.82867)	1.76756**(0.75050)	1.31372*(0.69963)	0.62075*(0.37408)	0.75287**(0.37656)
AR(1)	z = -6.23 *Pr* > *z* = 0.000	z = -4.35 *Pr* > *z* = 0.000	z = -6.43 *Pr* > *z* = 0.000	z = -6.21 *Pr* > *z* = 0.000	z = -7.63 *Pr* > *z* = 0.000
AR(2)	z = -1.11 *Pr* > *z* = 0.267	z = -0.93 *Pr* > *z* = 0.354	z = -1.55 *Pr* > *z* = 0.121	z = -0.91 *Pr* > *z* = 0.364	z = -0.97 *Pr* > *z* = 0.334
Hansen	chi2 (32) = 33.41 Pr > chi2 = 0.399	chi2 (31) = 17.57 Pr > chi2 = 0.975	chi2 (33) = 29.26 Pr > chi2 = 0.654	chi2 (42) = 50.17 Pr > chi2 = 0.181	chi2 (29) = 28.25 Pr > chi2 = 0.505

$BLOCK1_{it}$, $BLOCK2_{it}$ and $BLOCK3_{it}$ is the percentage of common shares held respectively by the largest the second largest and the third largest shareholder of the company; $BLOCK12_{it}$ is the percentage of common shares held jointly by the two major shareholders of the company. $BLOCK123_{it}$ is the percentage of common shares held jointly by the three major shareholders of the company. $DEBT_{it}$ is the ratio of the book value of the long term debt divided by the total assets; ROA_{it} is the return on assets; $INTANG_{it}$ stands for the intangible assets; $SIZE_{it}$ is the size of the company; $Growth_{it}$ is the sales growth; MJS_{it} is a dummy variable indicating whether the company has a majority shareholder.

Table 7

	I - AGE	II– AGE2	III - GROWTH	IV - ANGLO	V - ANGLO
$V_{i,t-1}$	0.47832***(0.05683)	0.47518***(0.04897)	0.29433***(0.02681)	0.45871***(0.07562)	0.24042***(0.05855)
$BLOCK_{1,it}$	0.04504***(0.015768)	0.03038**(0.01369)	0.03925***(0.01399)	0.05507***(0.01973)	0.02477***(0.00949)
$SBLOCK_{1,it}$	-0.00056***(0.00019)	-0.00037**(0.00016)	-0.00046***(0.00016)	-0.00075***(0.00024)	-0.00031***(0.00012)
$BLOCK_{2,it}$	-0.01129**(0.00476)	-0.00869**(0.00418)	-0.00866***(0.00308)	-0.02481**(0.01126)	-0.01182***(0.00441)
YD_{it}•$BLOCK_{1,it}$	-0.01186**(0.00575)				
YDF_{it}•$BLOCK_{1,it}$		-0.00748*(0.00406)			
$YDNF_{it}$•$BLOCK_{1,it}$		-0.01127**(0.00464)			
$Growth$•$BLOCK_{1,it}$			-0.00601**(0.00295)		
MJS•$BLOCK_{1,it}$					
$ANGLO$ • $BLOCK_{1,it}$				-0.03370**(0.01725)	
$ANGLO$ • $BLOCK_{2,it}$					0.01233**(0.00554)
$DEBT_{it}$	-0.85083***(0.32600)	-0.98506***(0.27442)	-2.41099***(0.15111)	-0.81387**(0.34006)	-1.52759***(0.25473)
$SIZE_{it}$	-0.05197**(0.02531)	-0.04263*(0.02247)	-0.08487**(0.04004)	-0.10158**(0.05333)	-0.11098***(0.03345)
$INTANG_{it}$	0.56096**(0.27934)	0.35669 (0.24971)	0.79876***(0.28826)	0.83632**(0.41988)	0.78143***(0.25855)
ROA_{it}	1.66257***(0.28979)	1.6436***(0.26829)	1.33448***(0.354588)	1.57267***(0.60484)	1.29328***(0.29907)
Constant	0.80802*(0.48394)	0.71670**(0.36339)	1.28450**(0.66366)	1.89570*(1.08340)	1.27215**(0.58757)
AR(1)	z = -7.08 *Pr > z* = 0.000	z = -7.59 *Pr > z* = 0.000	z = -6.44 *Pr > z* = 0.000	z = -5.56 *Pr > z* = 0.000	z = -4.87 *Pr > z* = 0.000
AR(2)	z = -0.50 *Pr > z* = 0.620	z = -0.19 *Pr > z* = 0.849	z = -1.02 *Pr > z* = 0.309	z = -0.23 *Pr > z* = 0.815	z = -1.12 *Pr > z* = 0.264
Hansen	chi2 (33) = 25.05 Pr > chi2 = 0.838	chi2 (49) = 41.88 Pr > chi2 = 0.755	chi2 (24) = 18.29 Pr > chi2 = 0.789	chi2 (23) = 17.11 Pr > chi2 = 0.804	chi2 (38) = 32.43 Pr > chi2 = 0.724

$BLOCK1_{it}$, $BLOCK2_{it}$ and $BLOCK3_{it}$ is the percentage of common shares held respectively by the largest the second largest and the third largest shareholder of the company; $BLOCK12_{it}$ is the percentage of common shares held jointly by the two major shareholders of the company. $BLOCK123_{it}$ is the percentage of common shares held jointly by the three major shareholders of the company. $DEBT_{it}$ is the ratio of the book value of the long term debt divided by the total assets; ROA_{it} is the return on assets; $INTANG_{it}$ stands for the intangible assets; $SIZE_{it}$ is the size of the company; $Growth_{it}$ is the sales growth; YD_{it}: is a dummy variable that equals one if the company is young and zero otherwise; YDF_{it}: is a dummy variable that equals one for young family companies and zero otherwise; $YDNF_{it}$: is a dummy variable that equals one for young non-family companies and zero otherwise;;ANGLO: is a dummy variable where a value of 1 is assigned for companies from the United Kingdom (Common Law countries), and a 0 for all other companies (Civil Law companies).

Table 8

	I –	II -	III -	IV - RESIDENT	V - IDENTITY
$V_{i,t-1}$	0.45618***(0.05539)	0.47568***(0.05610)	0.16859***(0.05783)	0.15036***(0.04695)	0.46295***(0.04564)
$BLOCK_{1,it}$	0.02162**(0.01045)	0.02240**(0.01122)	0.02695***(0.01481)	0.02424**(0.00983)	0.03468***(0.01331)
$SBLOCK_{1,it}$	-0.00031**(0.00013)	-0.00033***(0.00013)	-0.00031***(0.00009)	-0.00035***(0.00011)	-0.00043***(0.00016)
$BLOCK_{2,it}$	-0.01624**(0.00683)	-0.01531***(0.00568)	-0.01068***(0.00376)	-0.00829**(0.00358)	-0.01038**(0.00464)
SLS_{it}•$BLOCK_{1,it}$	0.00569**(0.00266)				
SLS_{it}		0.18820**(0.08951)			
$SBLOCK_{2,it}$			0.00016*(0.00008)		
$RESID$•$BLOCK_{1,it}$				0.00308**(0.00153)	
STD_{it} • $BLOCK_{1,it}$					-0.00347*(0.00209)
$DEBT_{it}$	-0.81376***(0.26496)	-0.76815***(0.27371)	-2.09352***(0.28806)	-1.99028***(0.23960)	-0.91926***(0.34006)
$SIZE_{it}$	-0.05377***(0.02037)	-0.03824**(0.01941)	-0.08909***(0.03452)	-0.09302**(0.03744)	-0.04879**(0.02383)
$INTANG_{it}$	0.44285** (0.20235)	0.39129*(0.20566)	0.46744*(0.26954)	0.55676***(0.20940)	0.42363**(0.19873)
ROA_{it}	1.84260***(0.41812)	1.96708***(0.49405)	0.47386*(0.27604)	0.60126**(0.25762)	1.45413***(0.26008)
Constant	0.28169 (0.36948)	1.9911 (0.66366)	1.34342*(0.75432)	1.20337*(0.66615)	0.71239*(0.43124)
AR(1)	z = -7.26 *Pr* > *z* = 0.000	z = -8.09 *Pr* > *z* = 0.000	z = -4.66 *Pr* > *z* = 0.000	z = -4.83 *Pr* > *z* = 0.000	z = -7.22 *Pr* > *z* = 0.000
AR(2)	z = -0.73 *Pr* > *z* = 0.468	z = -0.94 *Pr* > *z* = 0.346	z = -1.27 *Pr* > *z* = 0.204	z = -0.79 *Pr* > *z* = 0.428	z = -1.05 *Pr* > *z* = 0.292
Hansen	chi2 (38) = 48.63 Pr > chi2 = 0.116	chi2 (37) = 37.06 Pr > chi2 = 0.466	chi2 (44) = 33.42 Pr > chi2 = 0.877	chi2 (38) = 26.74 Pr > chi2 = 0.914	chi2 (37) = 38.23 Pr > chi2 = 0.413

$BLOCK1_{it}$, $BLOCK2_{it}$ and $BLOCK3_{it}$ is the percentage of common shares held respectively by the largest the second largest and the third largest shareholder of the company; $BLOCK12_{it}$ is the percentage of common shares held jointly by the two major shareholders of the company. $BLOCK123_{it}$ is the percentage of common shares held jointly by the three major shareholders of the company. $DEBT_{it}$ is the ratio of the book value of the long term debt divided by the total assets; ROA_{it} is the return on assets; $INTANG_{it}$ stands for the intangible assets; $SIZE_{it}$ is the size of the company; $Growth_{it}$ is the sales growth; SLS_{it}: is a dummy variable that equals one if the company has a second large shareholder and zero otherwise; $RESID_{it}$: is a dummy variable that equals one if the first large shareholder is a domestic shareholder and zero otherwise; STD_{it}: is a dummy variable that equals one if if the two main shareholders are the same group and zero otherwise

Table 9

	FAMILY	FAMILY/FAMILY	FAMILY/INST	FAMILY/NONFAMILY	INSTITUTIONAL
$V_{i,t-1}$	0.56607***(0.05934)	0.50454***(0.059274)	0.31761***(0.06561)	0.50868***(0.05956)	0.15338**(0.06882)
$BLOCK_{1,it}$	0.03859***(0.01528)	0.04026***(0.01537)	0.04176***(0.01568)		0.02510**(0.01168)
$SBLOCK_{1,it}$	-0.00047***(0.00018)	-0.00052***(0.00018)	-0.00054***(0.00019)		-0.00035**(0.00014)
$BLOCK_{2,it}$	-0.02534***(0.00941)	-0.01181**(0.00510	-0.01918***(0.00718)	-0.02263**(0.01167)	-0.00640**(0.00258)
$FD•BLOCK_{1,it}$	0.00459**(0.00224)	0.00436*(0.00233)	0.00358*(0.00203)	0.03964**(0.01757)	
$FFD•BLOCK_{1,it}$		-0.00366*(0.00199)			
$FID•BLOCK_{1,it}$			0.00335*(0.00202)		
$(FD•BLOCK_{1,it})^2$				-0.00038**(0.00018)	
$NFD•BLOCK_{1,it}$				0.04710**(0.01928)	
$(NFD•BLOCK_{1,it})^2$				-0.00059***(0.00022)	
$IID•BLOCK_{1,it}$					-0.00313***(0.00122)
$DEBT_{it}$	-0.59214**(0.28336)	-0.81856***(0.27544)	-1.37600***(0.31854)	-0.74375***(0.28377)	-2.05088***(0.32557)
$SIZE_{it}$	-0.03855*(0.02163)	-0.05940**(0.02560)	-0.08540***(0.03087)	-0.05002*(0.02941)	-0.10202**(0.04135)
$INTANG_{it}$	0.53262**(0.25932)	0.52013**(0.22727)	0.48799**(0.23204)	0.49138**(0.25107)	0.63011**(0.24727)
ROA_{it}	1.42362**(0.58090)	1.13012**(0.55537)	1.08519***(0.32862)	1.44071**(0.66376)	0.60802*(0.33245)
Constant	0.80236*(0.45204)	0.90309*(0.51704)	1.51454**(0.70744)	0.89398*(0.53395)	1.49373*(0.81674)
AR(1)	z = -6.65 *Pr* > *z* = 0.000	z = -7.13 *Pr* > *z* = 0.000	z = -5.75 *Pr* > *z* = 0.000	z = -5.57 *Pr* > *z* = 0.000	z = -3.74 *Pr* > *z* = 0.000
AR(2)	z = -0.86 *Pr* > *z* = 0.392	z = -0.77 *Pr* > *z* = 0.443	z = -0.80 *Pr* > *z* = 0.423	z = -0.85 *Pr* > *z* = 0.394	z = -0.72 *Pr* > *z* = 0.473
Hansen	chi2 (24) = 24.14 Pr > chi2 = 0.454	chi2 (35) = 27.46 Pr > chi2 = 0.814	chi2 (40) = 38.60 Pr > chi2 = 0.533	chi2 (22) = 17.15 Pr > chi2 = 0.755	chi2 (29) = 24.03 Pr > chi2 = 0.728

$BLOCK1_{it}$, $BLOCK2_{it}$ and $BLOCK3_{it}$ is the percentage of common shares held respectively by the largest the second largest and the third largest shareholder of the company; $BLOCK12_{it}$ is the percentage of common shares held jointly by the two major shareholders of the company. $BLOCK123_{it}$ is the percentage of common shares held jointly by the three major shareholders of the company. $DEBT_{it}$ is the ratio of the book value of the long term debt divided by the total assets; ROA_{it} is the return on assets; $INTANG_{it}$ stands for the intangible assets; $SIZE_{it}$ is the size of the company; $Growth_{it}$ is the sales growth; FD_{it}: is a dummy variable that equals one if the largest shareholder is an individual or a family with at least 10 percent of the cash flow rights and zero otherwise; FFD_{it}: is a dummy variable that equals one if the two largest shareholder are family with at least 10 percent of the cash flow rights and zero otherwise. FID_{it}: is a dummy variable that equals one if the first large shareholder is a family and the second a a institutional investor with at least 10 percent of the cash flow rights and zero otherwise; NFD_{it}: is a dummy variable that equals one if the largest shareholder is not a individual or a family with at least 10 percent of the cash flow rights and zero otherwise; IID_{it}: is a dummy variable that equals one if the largest shareholder is a institutional investor with at least 10 percent of the cash flow rights and zero otherwise.

Appendix
Appendix A

In this appendix we present the definitions and calculations, when necessary, of all variables used in our analysis. Except for the items we point out that come from the *Main Economic Indicators* published by the Organization for Economic Cooperation and Development (OECD), the remaining items have been extracted from AMADEUS.

Dependent variable

– *Company value*: is calculated dividing the market value of equity by the total assets. That is, $V_{it} = MVE_{it} / TA_{it}$
where $MVE_{it} = SO_{it} * MPS_{it}$ being SO_{it} a measure of the company shares outstanding and MPS_{it} the market price of shares.

Ownership variables

– $BLOCK1_{it}$; $BLOCK2_{it}$; and $BLOCK3_{it}$ is the percentage of common shares held respectively by the largest, the second largest and the third largest shareholder of the company. Blockholders are shareholders with cash-flow ownership (La Porta et al. 2002) greater than 10% of the firm's outstanding shares. If there is no large shareholder with at least 10% of the cash flow rights, $BLOCK1_{it} = BLOCK2_{it} = BLOCK3_{it} = 0$

– $SBLOCK1_{it}$ and $SBLOCK2_{it}$ is the square of the percentage of common shares held respectively by the largest and the second largest shareholder of the company.

– $BLOCK12_{it}$ and $SBLOCK12_{it}$ is the percentage of common shares held jointly by the two major shareholders of the company and its square.

– $BLOCK123_{it}$ and $SBLOCK123_{it}$ is the percentage of common shares held jointly by the three major shareholders of the company and its square.

– SLS_{it}: is a dummy variable that equals one if the company has a second large shareholder and zero otherwise.

– STD_{it}: is a dummy variable that equals one if if the two main shareholders are the same group and zero otherwise.

– YD_{it}: is a dummy variable that equals one if the company is young and zero otherwise. We consider a company as being young when the company age is below the mean value of this variable in our sample.

– YDF_{it}: is a dummy variable that equals one for young family companies and zero otherwise. We use the aforementioned criteria to classify a company as being young and family at the same time.

– $YDNF_{it}$: is a dummy variable that equals one for young non-family companies and zero otherwise. We use the aforementioned criteria to classify a company as being young and non-family at the same time.

– FD_{it}: is a dummy variable that equals one if the largest shareholder is an individual or a family with at least 10 percent of the cash flow rights and zero otherwise.

– NFD_{it}: is a dummy variable that equals one if the largest shareholder is not a individual or a family with at least 10 percent of the cash flow rights and zero otherwise.

– FFD_{it}: is a dummy variable that equals one if the two largest shareholders are family with at least 10 percent of the cash flow rights and zero otherwise.

– FID_{it}: is a dummy variable that equals one if the first large shareholder is a family and the second a institutional investor with at least 10 percent of the cash flow rights and zero otherwise.

– IID_{it}: is a dummy variable that equals one if the largest shareholder is a institutional investor with at least 10 percent of the cash flow rights and zero otherwise.

- MJS_{it} : is a dummy variable indicating whether the company has a majority shareholder. This variable takes value of one if the controlling shareholder has a majority stake, defined as the percentage of common shares held by the largest shareholder of the company.

ANGLO: is a dummy variable where a value of 1 is assigned for companies from the United Kingdom (Common Law countries), and a 0 for all other companies (Civil Law companies).

Control variables

– *Debt ratio*: $DEBT_{it} = NCL_{it}/(NCL_{it} + MVE_{it})$ where NCL_{it} is the the book value of the long term debt and MVE_{it} is the market value of equity;

– *Return on assets:* $ROA_{it} = EBITDA_{it} / TA_{it}$ where $EBITDA_{it}$ denotes earnings before interest, taxes, depreciation and amortization

– $Size_{it}$*:* $Ln(REVENUE_{it})$

– *Intangible assets*: $INTANG_{it} = IFA_{it} / TA_{it}$ where IFA_{it} is the book value of the intangible fixed assets.

– *Age*: $LNAGE_{it} = Ln (YEAR_{it} – INC_i)$ where $YEAR_{it}$ is the corresponding period of time and INC_i is the date of incorporation of the company.

– *Growth* we follow Klapper and Love (2004) and use the annual Sales Growth. Growth is calculated as $GROWTH_{it} = [(REV_{it} – REV_{i,t-1}) / REV_{i,t-1}]$

REFERENCES

A. AGRAWAL and C. KNOEBER. "Firm Performance and Mechanisms to Control Agency Problems between Managers and Shareholder" *Journal of Financial and Quantitative Analysis*, September 1996, V. 31(3), 377-397.

ADMATI, ANAT R., PAUL PFLEIDERER, and JOSEF ZECHENER, 1994, Large shareholder Activism, Risk Sharing, and Financial Market Equilibrium, *Journal of Political Economy* 102, 1097-1130.

ANDERSON R.C. and D.M. REEB. 2003. «Founding-Family Ownership and Firm Performance: Evidence from the S&P 500.» *Journal of Finance*, Vol. 58, pp. 1301-1329.

ANDERSON, R.C.; MANSI, S.A. and REEB, D.M. (2003): "Founding family ownership and the agency cost of debt", *Journal of Financial Economics*, 68:2, pp. 263-285.

ANDRES, C. (2008): "Large shareholders and firm performance − An empirical examination of founding-family ownership", *Journal of Corporate Finance*, 14:4,pp. 431-445.

ARELLANO, MANUEL and STEPHEN BOND. 1991. "Some Tests of Specification for Panel Data: Monte

Carlo Evidence and an Application of Employment Equations." *Review of Economic Studies* 58: 277-297.

ARELLANO, MANUEL and OLYMPIA BOVER. 1995. "Another Look at the Instrumental Variable Estimation of Error Component Models." *Journal of Econometrics* 68: 29-51.
BARCA, F., BECHT, M., 2001. The Control of Corporate Europe. Oxford University Press.
BARCLAY, M. and C. HOLDERNESS (1989), 'Private Benefits from Control of Public Corporations', Journal of Financial Economics 25, 371-395.
BARONTINI, R. and CAPRIO, L. (2006): "The effect of family control on firm value and performance: Evidence from continental Europe", *European Financial Management*, 12:5, pp. 689-723.
BARTH, E., GULBRANDSEN, T., SCHONE, P., 2005. Family Ownership and Productivity: the Role of Owner-Management. Journal of Corporate Finance 11, 107-127.
BEBCHUK, LUCIAN, REINIER KRAAKMAN, and GEORGE TRIANTIS, 1999, Stock pyramids and dual class equity: The creation and agency costs of separating control from cash flow rights, NBER Working Paper No. 6951.
BEBCHUK, LUCIAN A., ALMA COHEN, and ALLEN FERRELL, 2004, "What Matters in Corporate Governance?,. Working Paper, Harvard Law School and NBER.
BECHT, MARCO and COLIN MAYER (2001), *"The Control of Corporate Europe"*, in Barca and Becht (eds.), The Control of Corporate Europe
BECHT, MARCO and AILSA RÖELL (1999), "Blockholdings in Europe: An International Comparison", *European Economic Review* (43)4-6, pp. 1049- 1056
BEINER, STEFAN; WOLFGANG; DROBETZ, FRANK, SCHMID, e ZIMMERMANN, Heinz 2004, Is board size an independent corporate governance mechanism, *Kyklos*, 57, 3, 327-356.
BENNEDSEN, M., WOLFENZON, D., 2000. The balance of power in closely held corporations. *Journal of Financial Economics* 58, 113–139.
BENNEDSEN, M.; NIELSEN, K.M.; PÉREZ-GONZÁLEZ, F. and WOLFENZON, D. (2007): "Inside the family: The role of families in succession decisions and performance", *The Quarterly Journal of Economics*, 122:2, pp. 647-691.
BERLE, A., MEANS, G., 1932. The Modern Corporation and Private Property. Macmillan, New York.
BERGLOF, E., PAJUSTE, A. (2003), *Emerging Owners, Eclipsing Markets? Corporate Governance in Central and Eastern Europe*, in P. Cornelius and B. Kogut, *Global Competitiveness and Corporate Governance*, Oxford University Press, pp. 267-302
BERGLÖF, E., VON THADDEN E. (2000), *The Changing Corporate Governance Paradigm*, in Pleskovic and Stiglitz (eds.) *World Development Conference*, The World Bank.
BERNARD, A., WAGNER, J., 1998. Export Entry and Exit by German Firms, NBER WP, No. 6538.
BLACK, B., JANG, H., KIM, W., 2003. Does corporate governance affect firm's market values? Evidence from Korea, Working paper, Stanford Law School.
BLOCH, L., and E. KREMP, 2001, "Ownership and Voting Power in France", In *The Control of Corporate Europe* (F. Barca and M. Becht, eds.), Oxford University Press, 106-127.
BLOCH, F., and U. HEGE, 2001, "Multiple Shareholders and Control Contests," unpublished manuscript.
BLUNDELL, RICHARD and STEPHEN BOND. 1998. "Initial Conditions and Moment Restrictions in DYNAMIC PANEL DATA MODELS." *Journal of Econometrics* 87: 115-143.
BOEHMER, E., 2000. Business groups, bank control, and large shareholders: An analysis of German takeovers. *Journal of Financial Intermediation* 9, 117–148.
BØHREN, Ø. and ØDEGAARD, B. A. (2005). Governance and performance revisited. In Gregouriu, G. and ALI, P. U., editors, International Corporate Goverannce after Sarbanes-Oxley. Wiley
BOOTH, L., AIVAZIAN, V., DEMIRGUC-KUNT, A., & MAKSIMOVIC, V. (2001). Capital structure in developing countries. *Journal of Finance, 56*(1), 87-130.
BURKART, M., GROMB, D., PANUNZI, F., 1997. Large shareholders, monitoring, and the value of the firm. *Quarterly Journal of Economics* 112, 693–728.
BURKART, M., PANUNZI, F. (2001) Agency Conflicts, Ownership Concentration and Legal Shareholder Protection. Financial Markets Group (LSE) Discussion Paper 378.
BURKART, M., PANUNZI, F., SHLEIFER, A., 2003. Family firms. Journal of Finance 58, 2167–2202.
CHAMI, R. and FULLENKAMP, C. (2002). `Trust and efficiency´. *Journal of Banking and Finance*, **26**, 1785-1810.
CHEN, X. and YUR-AUSTIN, J. (2007): "Re-measuring agency costs: The effectiveness of blockholders", *The Quarterly Review of Economics and Finance*, 47:5, pp. 588-601.
CLAESSENS, STIJN, SIMEON DJANKOV, and LARRY LANG, 2000, "The Separation of Ownership and CONTROL IN EAST ASIAN CORPORATIONS," *Journal of Financial Economics* 58(1-2), 81-112.
CLAESSENS, S., DJANKOV, S., FAN, J., LANG, L., 2002. Disentangling the incentive and entrenchment effects of large shareholdings. *Journal of Finance* 57, 2741–2771.

COASE, RONALD H., 1937, "The Nature of the Firm," *Economica* 4, 386-405.
CREMERS KJM, NAIR VB (2005) Governance mechanisms and equity prices. *Journal of Finance* 60:2859–2894
CRONQVIST, H., NILSSON, M., 2003. Agency costs of controlling minority shareholders. *Journal of Financial and Quantitative Analysis* 38, 695–719.
CRONQVIST, HENRIK, and RÜDIGER FAHLENBRACH, 2008, Large Shareholders and Corporate Policies, *Review of Financial Studies* forthcoming.
CUCCULELLI, M. and MICUCCI, G. (2008): "Family succession and firm performance: Evidence from Italian family firms", *Journal of Corporate Finance*, 14:1, pp. 17-31.
DAHYA, J.; DIMITROV, O. and MCCONNELL, J.J. (2008): "Dominant shareholders, corporate boards and corporate value: A cross-country analysis", *Journal of Financial Economics*, 87:1, pp. 73-100.
DALTON DR, DAILY CM, ELLSTRAND AE, JOHNSON JL. 1998. Meta-analytic reviews of board composition, leadership structure, and financial performance. *Strategic Management Journal* 19(3): 269–290.
DAILY, C.M., DOLLINGER, M.J., 1992. An empirical examination of ownership structure in family and professionally managed firms. Family Business Review 5 (2), 117–136.
DAILY, C., THOMPSON, S., 1994. Ownership structure, strategic posture, and firm growth: an empirical examination. Family Business Review 7 (3), 237–249.
DAVIES, J., D.HILLIER, and P. MCCOLGAN, 2002, Ownership structure, managerial behavior and corporate value, Working paper, University of Strathclyde
DEANGELO, H., DEANGELO, L., 2006. The irrelevance of the MM dividend irrelevance theorem. Journal of Financial Economics 79, 293–315.
DE PAOLA, M. and SCOPPA, V. (2001). `The role of family ties in the labour market: An interpretation based on efficiency wage theory´. *Review of Labour Economics and Industrial Relations*, **15**, 603-624.
DEMSETZ, HAROLD, 1983, "The Structure of Ownership and the Theory of the Firm," *Journal of Law and Economics* 26, 375-393.
DEMSETZ, HAROLD, and KENNETH LEHN, 1985, "The Structure of Corporate Ownership: Causes and Consequences," *Journal of Political Economy* 93, 1155-1177.
DEMSETZ, H. and B. VILLALONGA (2001), "Ownership Structure and Corporate Performance", *Journal of Corporate Finance*, 7, 209-233.
DENIS, D.K. and MCCONNELL, J.J. (2003): "International corporate governance", *Journal of Financial and Quantitative Analysis*, 38:1, pp. 1-36.
DLUGOSZ, JENNIFER, RÜDIGER FAHLENBRACH, PAUL GOMPERS, and ANDREW METRICK, 2006, Large Blocks of Stock: Prevalence, Size, and Measurement, *Journal of Corporate Finance* 12, 594-618.
DYCK, A., L. ZINGALES. 2004. Private benefits of control: An international comparison. *Journal of Finance.* 59: 537-600.
DURNEV, A., KIM, E.H., 2005. To steal or not to steal: firm attributes, legal environment, and valuation. Journal of Finance 60, 1461– 1493.
EDMANS, ALEX, 2008, Blockholder Trading, Market Efficiency, and Managerial Myopia. *Journal ofFinance*. Forthcoming.
EDMANS, ALEX and GUSTAVO MANSO, 2008, Governance Through Exit and Voice: A Theory of Multiple Blockholders, Working Paper, Wharton School, University of Pennsylvania.
EDWARDS, JEREMY S.S. and ALFONS J. WEICHENRIEDER. 1999. "Ownership Concentration and Share Valuation: Evidence from Germany." Manuscript, University of Cambridge and CESifo, July.
EDWARDS & WEICHENRIEDER, 2004. "Ownership Concentration and Share Valuation," *German Economic Review*, Blackwell Publishing, vol. 5(2), pages 143-171
ÉRIC SÉVERIN" Ownership structure and the performance of firms: Evidence from France" European Journal of Economic and Social Systems 15 N° 2 (2001) 85-107
FACCIO, M., LANG, L., YOUNG, L., 2001. Dividends and expropriation. *American Economic Review* 91, 54–78.
FACCIO, M., LANG, L., 2002. The ultimate ownership of Western European corporations. *Journal of Financial Economics* 65, 365–395.
FAMA, E. (1980), 'Agency Problems and the Theory of the Firm', Journal of Political Economy 88, 288-298.
FAMA, E. and M. JENSEN (1983), 'Separation of Ownership and Control', Journal of Law and Economics 26, 301-325.
FAMA E.F. et JENSEN M.C., 1985, « Organizational Forms and Investment Decisions », *Journal of Financial Economics*, vol. 14, p. 101-119.

FRANKS, J. and C. MAYER (1995): "Ownership and Control" in H. Siebert (ed) Trends in Business Organization: Do Participation and Cooperation Increase Competitiveness? Mohr: Tubingen.

FRANKS, JULIAN, and COLIN MAYER (1997). "Corporate Ownership and Control in the U.K., Germany and France," *Bank of America Journal of Applied Corporate Governance*, 9, 4, 30-45.

GADHOUM, LANG and YOUNG, 2005, Who Controls US?, *European Financial Management*, Vol. 11, No. 3.

GAVER, J.J., GAVER, K.M., 1993. Additional evidence on the association between the investment

opportunity set and corporate financing, dividend, and compensation policies. Journal of Accounting and Economics 16 (1–3), 125–160.

GEDAJLOVIC, E. and SHAPIRO D. 1998. Management and ownership effects: evidence from 5 countries. *Strategic Management Journal* **1**9(6): 533–555.

GILSON, RONALD and JEFFREY GORDON, Controlling Controlling Shareholders, 2003, working paper, Columbia Law School.

GOERGEN, M., RENNEBOOG, L. and L. CORREIA DA SILVA (2005) When do German firms change their dividends? , *Journal of Corporate Finance*. Vol 11, pp 375-399 .

GOMES, ARMANDO and WALTER NOVAES (2006): "Sharing of Control versus Monitoring as Corporate Governance Mechanisms." Working Paper, Washington University in St. Louis

GORTON, G., and F. SCHMID. 2000. Universal Banking and the Performance of German firms. *Journal of Financial Economics* 58: 29-80.

GRAHAM, J. R., M. H. LANG, and D. A. SHACKELFORD. "Employee Stock Options, Corporate Taxes and Debt Policy." *Journal of Finance* (2004).

GREGORIC, ALEKSANDRA., ARJANA B. MASTEN and KATERINA ZAJC, 2008, Exploring Barriers to Coase:

Blockholder Identity and Path-Dependency as Determinants of Ownership Dynamics in the post-Privatization Period, Working Paper. Copenhagen Business School.

GROSSMAN, SANFORD J. and OLIVER HART, 1980, "Takeover Bids, the Free-Rider Problem, and the Theory of the Corporation," *Bell Journal of Economics* 11, 42-64.

GROSSMAN, S.J., e O.D. HART (1982), Corporate financial structure and managerial incentives, em John J. McCall (ed.): *The Economics of Information and Uncertainty*, University of Chicago Press: Chicago, Ill.

GROSSMAN, SANFORD J. and OLIVER HART, 1986, "The Cost and Benefit of Ownership: A Theory of Lateral and Vertical Integration," *Journal of Political Economy* 94, 691-719.

GROSSMAN, S., HART, O., 1988. One share–one vote and the market for corporate control. *Journal of Financial Economics* 20, 175–202.

GUTIÉRREZ, M., and TRIBÓ, J.A. (2004): "Private benefits extraction in closely-held corporations: The case for multiple large shareholders". European Corporate Governance Institute Finance Working Paper 53/2004

HART, O. D., 1983, The Market Mechanism as an Incentive Scheme. *Bell Journal of Economics*, 14, pp. 366-382.

HART, OLIVER and JOHN MOORE, 1990, "Property Rights and the Nature of the Firm," *Journal of Political Economy* 98(6), 1119-1158.

HART, O., & MOORE, J. (1995). Debt and seniority: An analysis of the role of hard claims in constraining management. *American Economic Review*, *85*, 567–585.

HART, OLIVER, 1995, Firms, Contracts, and Financial Structure, (London: Oxford University Press,UK).

HERMALIN, B.E., WEISBACH, M.S., 1991. The effects of board composition and direct incentives on firm performance. *Financial Management* 20, 101–112.

HIMMELBERG, C., HUBBARD, R.G. and PALIA, D. (1999) Understanding the determinants of managerial ownership and the link between ownership and performance. Journal of Financial Economics 53,353–384.

HIMMELBERG, CHARLES P., R. GLENN HUBBARD, and INESSA LOVE, 2002, "Investor Protection, Ownership, and the Cost of Capital," working paper, Columbia University (available on SSRN).

HOLDERNESS, C., SHEEHAN, D., 1988. The role of majority shareholders in publicly held corporations: An exploratory analysis. *Journal of Financial Economics* 20, 317–346.

HOLDERNESS, CLIFFORD, RANDALL S. KROSZNER, and DENNIS SHEEHAN, 1999, "Were the Good Old Days That Good?: Changes in Managerial Stock Ownership Since the Great depression,"J*ournal of Finance* 54(1), 435-469.

HOLDERNESS, C. G. 2003. A Survey of Blockholders and Corporate Control. *Economic Policy Review* 9(1): 51-63.

HOLDERNESS, CLIFFORD (2007): "The Myth of Diffuse Ownership in the United States." Review of Financial Studies, forthcoming

JENSEN, MICHAEL C. and WILLIAM H. MECKLING, 1976, "Theory of the Firm: Managerial Behavior, Agency Cost, and Capital Structure," *Journal of Financial Economics* 3, 305-360.

JENSEN, M.C., 1986, «Agency Costs of Free Cash Flow, Corporate Finance and Takeovers», *American Economic Review*, Vol. 76, May, pp. 323-329.

HANSMANN, H. (1996) *The ownership of enterprise,* Cambridge, Massachusetts, The Belknap Press of Harvard University Press.

HIMMELBERG, C., HUBBARD, R.G., LOVE, I., 2001. Investor protection, ownership and the cost of capital. Unpublished working paper. Columbia University.

KARRA, N., TRACEY, P. and PHILLIPS, N. (2006). `Altruism and agency in the family firm:Exploring the role of family, kinship and ethnicity´. *Entrepreneurship Theory and Practice,* **30**, 861-877.

KHANNA, N., TICE, S., 2005. Pricing, exit, and location decisions of firms: evidence on the role of debt and operating efficiency. Journal of Financial Economics 75, 397–428.

KIM, K.A.; KITSABUNNARAT-CHATJUTHAMARD, P. and NOFSINGER, J.R. (2007): "Large shareholders, board independence, and minority shareholder rights: Evidence from Europe", *Journal of Corporate Finance*, 13:5, pp. 859-880.

KLAPPER, L., & LOVE, I. (2004). Corporate governance, investor protection, and performance in emerging markets. *Journal of Corporate Finance, 10*(5), 703-728.

LA PORTA, RAFAEL, FLORENCIO LOPEZ-DE-SILANES, ANDREI SHLEIFER, and ROBERT VISHNY, 1998, "Law and Finance," *Journal of Political Economy* 106, 1113-1155.

LA PORTA, R., LOPEZ-DE-SILANES, F., SHLEIFER, A., 1999. Corporate ownership around the world. *Journal of Finance* 54, 471–517.

LA PORTA, RAFAEL, FLORENCIO LOPEZ-DE-SILANES, ANDREI SHLEIFER, and ROBERT VISHNY, 2000, "Investor Protection and Corporate Governance," *Journal of Financial Economics* 58, 3-27.

LA PORTA, R., LOPEZ-DE-SILANES, F., SHLEIFER, A., VISHNY, R., 2002. Investor protection and corporate valuation. *Journal of Finance* 57, 1147–1170.

LEECH, D. and LEAHY, J. (1991) Ownership structure, control-type classifications and the performance of large British companies, *Economic Journa*l, **101**, 1418-1437.

LEHMAN, E., WEIGAND, J., 2000. Does the governed corporation perform better? Governance structures and corporate performance in Germany. *European Finance Review* 4, 157–195.

LEIBENSTEIN, H., 1996, Allocative Efficiency Versus X-Efficiency. *American Economic Review*, **56**, pp. 392-413.

LEMMON, MICHAEL L. and KARL LINS, 2003, "Ownership Structure, Corporate Governance, and Firm Value: Evidence from the East Asian Financial Crisis," *Journal of Finance* 58, 1445-1468.

LINS, KARL, 2003, "Equity Ownership and Firm Value in Emerging Markets," *Journal of Financialand Quantitative Analysis* 38, 159-184.

LUC LAEVEN and ROSS LEVINE , 2008, "Complex Ownership Structures and Corporate Valuations." Review of Financial Studies 21(2):579-604.

MANSI and REEB, 2002, Corporate Diversification: What Gets Disconnected? *Journal of Finance*, 2167-83.

MARTIKAINEN, M.; NIKKINEN, J. and VÄHÄMAA, S. (2008): "Production functions and productivity of family firms: Evidence from the S&P 500", *Quarterly Review of Economics and Finance*, in press.

MAUG E. Large Shareholders as Monitors: Is There a Trade-Off Between Liquidity and Control? Journal of Finance 1998; **53**: 65-92.

MAURY, B. (2006): "Family ownership and firm performance: Empirical evidence from Western European corporations", *Journal of Corporate Finance*, 12:2, pp. 321-341.

MAURY, BENJAMIN and ANETE PAJUSTE, 2005, "Multiple Controlling Shareholders and Firm Value," *Journal of Banking and Finance*.

MCCONAUGHY, D.L., WALKER, M.C., HENDERSON, G.V., MISHRA, C.S., 1998. Founding Family Controlled Firms: Efficiency and Value. Review of Financial Economics 7, 1-19.

MCCONNELL, J., SERVAES, H., 1990. Additional evidence on equity ownership and corporate value. *Journal of Financial Economics* 27, 595–612.

MCCONNEL, JOHN and CHRIS J. MUSCARELLA. 1985. Corporate Capital Expenditure Decisions and the Market Value of the Firm. *Journal of Financial Economics.*Sept. 14(3) 399-422.

MCCONNELL , J., SERVAES, H., 1995. Equity Ownership and the Two Faces of Debt, Journal of Financial Economics 39, 1995.

MIGUEL, A., PINDADO, J. and DE LA TORRE, C. (2004) Ownership Structure and Firm Value: New Evidence from the Spanish Case, *Strategic Management Journal*, 25, 1199–1207.

MILLER, D.; LE BRETON-MILLER, I.; LESTER, R.H. and CANNELLA, A.A. (2007): "Are family firms really superior performers?", *Journal of Corporate Finance*, 13:5, pp. 829-858.

MORK, R., M. NAKAMURA and A. SHIVDASANI, 2000, Banks, Ownership structure, and Firm Value in Japan, Journal of Business 73, 539 – 567.

MORCK, RANDALL, ANDREI SHLEIFER, and ROBERT W. VISHNY, 1988, "Management Ownership and Market Valuation: An Empirical Analysis," *Journal of Financial Economics* 20, 293-315.

MORCK, R. K., D. A. STANGELAND and B. YEUNG. 2000. «Inherited Wealth, Corporate Control and Economic Growth: The Canadian Disease?» in *Concentrated Corporate Ownership, editor*

MUDAMBI, RAM & NICOSIA, CARMELA, 1998. "Ownership Structure and Firm Performance: Evidence from the UK Financial Services Industry" Applied Financial Economics, vol. 8(2), pages 175-80, April

NAGAR, PETRON, and WOLFENZON "Governance Problems in Close Corporations" (2004).

NENOVA, T., 2003. The value of corporate votes and control benefits. *Journal of Financial Economics* 68, 325–351.

PAGANO, M., ROELL, A., 1998. The choice of stock ownership structure: Agency costs, monitoring, and the decision to go public. *Quarterly Journal of Economics* 113, 187–226.

PROWSE, S., 1992. The structure of corporate governance in Japan. Journal of Finance 42, 1121--1140.

PEDERSEN, T., & THOMSEN, S. (2003). Ownership structure and value of the largest European firms: The importance of owner identity. *Journal of Management & Governance*, *7*, 27-55.

PROWSE, S. 1995. Corporate Governance in An International Perspective: A Survey of Corporate Control Mechanism Among Large Firms in the U.S., U.K., Japan and Germany. *Financial Markets, Instituions & Instruments* 4, 1-63.

RAJAN, R., & ZINGALES, L. (1995). What do we know about capital structure? Some evidence from international data. *Journal of Finance, 50*(5), 1421-1441

RAJAN, R.G., ZINGALES, L. (2000) The Tyranny of Inequality. *Journal of Public Economics* 76: 521–558.

RANDAL MORCK R. WOLFENZON D. YEUNG B. Corporate Governance, Economic Entrenchment, and Growth. Journal of Economic Literature 2005.

RANDØY, T. and NIELSEN, J. (2002). `Company Performance, corporate governance, and CEO compensation in Norway and Sweden´. *Journal of Management and Governance,* **6**, 57-81.

RENNEBOOG, L., 2000, Ownership, managerial control and the corporate governance of companies listed on the Brussels Stock Exchange, *Journal of Banking and Finance*, vol. 24,

1959-1995.

SHLEIFER, A., VISHNY, R., 1986. Large shareholders and corporate control. *Journal of Political Economy* 94, 461–488.

SHLEIFER, A., VISHNY, R., 1997. A survey of corporate governance. *Journal of Finance* 52, 737–782.

SHLEIFER, A., WOLFENZON, D., 2002. Investor protection and equity markets. *Journal of Financial Economics* 66, 3–27.

SHORT, H. and KEASEY, K., 1999, Managerial ownership and the performance of firms: Evidence from the UK, *Journal of Corporate Finance* 5, 79-101

SMITH JR., C.W., WATTS, R.L., 1992. The investment opportunity set and corporate financing,

dividend, and compensation policies. Journal of Financial Economics 32 (3), 263–292.

SRAER, D. and THESMAR, D. (2007). Performance and Behavior of Family Firms: Evidence from the French Stock Market. ECGI Finance Working Paper No. 130/2006.

STULZ, R. (1990). Managerial discretion and optimal financial policies. *Journal of Financial Economics*, *26*, 3–27.

THOMSEN, S., & PEDERSEN, T. 2000. Ownership structure and economic performance in the largest European companies. Strategic Management Journal, 21: 689-705.

THOMSEN, S.; PEDERSEN, T. and KVIST, H.K. (2006): "Blockholder ownership: Effects on firm value in market and control based governance systems", *Journal of Corporate Finance*, 12:2, pp. 246-269.

VILLALONGA, B. and AMIT, R. (2006a). `How do family ownership, control and management affect firm value?´ *Journal of Financial Economics,* **80**, 385-417.

VILLALONGA, B. and AMIT, R. (2006b). Benefits and Costs of Control-Enhancing Mechanisms in U.S. Family Firms. ECGI - Finance Working Paper No. 131/2006

VOLPIN, P., 2002. Governance with poor investor protection: Evidence from top executive turnover. *Journal of Financial Economics* 64, 61–90.

WANG, D. (2006): "Founding family ownership and earnings quality", *Journal of Accounting Research*, 44:3, pp. 619-656.

WILLIAMSON, O. (1988). Corporate finance and corporate governance. Journal of Finance, 43, 567-591.

ZOIDO, M.E., 1998. Un estudio de las participaciones accionariales de los bancos en las empresas españolas. Investigaciones Económicas 22 (3), 427--467.

ZWIEBEL, J., 1995. Block investment and partial benefits of corporate control. *Review of Economic Studies* 62, 161–185.

EXCHANGE RATE PASS-THROUGH TO PRICES IN THE HONG KONG

MISZTAL PIOTR
E-mail: misztal@msg.radom.pl
The Radom Technical University

Abstract

The main aim of the article is to present the influence of exchange rate changes on the price dynamic in the Hong Kong. The article consists of two parts. The first part deals with theoretical analysis of the phenomenon of incomplete exchange rate pass-through to prices, including reasons and factors determining the range of this phenomenon. Moreover, in this part, there is made a brief overview of theoretical researches involving this subject. In the next part of article, there is analyzed the range of exchange rate pass-through to prices in the Hong Kong by using the Vector Autoregression Model (VAR). There are estimated coefficients of exchange rate pass-through to import, producer and consumer prices on the base of impulse response function. Then, there is made decomposition of the price index variance in order to assess the degree of price determination by exchange rate changes.

Key-Words: Exchange rate, transmission mechanism, pass-through to prices.
Thematic Area: Monetary and Financial Economics.

1. INTRODUCTION

The phenomenon of exchange rate pass-through to prices has become one of the key research areas within so-called New Open Economy Macroeconomics. An exchange rate pass-through is defined as a total percentage response of prices to a one percent shock in the exchange rate. For a long time the literature on the topic had assumed a complete pass-through of such changes to prices in conformity with the law of one price and purchasing power parity. Only in the 1980s empirical analyses revealed that it is rather an incomplete exchange rate pass-through of into domestic and foreign prices which is a common phenomenon

Understanding the behavior of import prices is a key issue for inflation targeting central banks in small open economies. It is important to know the fraction of exchange rate movements that is passed to the import prices, what depends on the extent to which the "law of one price" holds as well as pass-through rates. Pass-through rates measure the percentage of exchange rate changes that are transmitted to the import prices. If a local currency appreciates against foreign currencies, import prices will no doubt fall. A depreciation of the local currency would have the opposite effect. The degree of exchange rate pass-through to import prices are usually between zero and 100 but can be negative if a change in the exchange rate has the opposite effect on import price indices; these rates can exceed 100 if the import price index changes by a greater extent than the exchange rate. The scales of exchange rate pass-through to import prices are also a significant factor explaining the evolution of domestic inflation. Import prices affect the level of domestic inflation in several ways. Decreasing import prices of final goods directly affect the overall level of prices in the economy. Import prices can also be a good indicator of future inflation in a country where many inputs to domestic production are imported. Declining prices of imported inputs to manufactured goods have an indirect impact on the level of output prices by lowering the cost of production. This means that, if import prices for a product are falling, national producers of that product will have to lower their prices to maintain their competitiveness. Any seasonal or trend information embedded in the time series of import prices would contribute to the knowledge about future price levels.

One of the first economists analyzing the phenomenon of incomplete exchange rate pass-through to prices was Krugman and Dornbush. They came to a conclusion that pricing to market is an essential factor determining deviations from the law of one price (Krugman, 1987; Dornbush, 1987). Furthermore, Taylor indicated also a different factor determining the degree of exchange rate pass-through to prices, which is the inflation rate. He proves that a low inflation rate leads to a lower degree of exchange rate pass-through to domestic prices (Taylor, 2000). The situation is reverse in the countries of a low inflation rate (Choudhri & Hakura, 2001). Among the most crucial causes of the incomplete exchange rate pass-through to prices, economic literature names:

a) occurrence of imperfect competition;
b) tendency to invoicing transactions in local currency;
c) growing marginal costs of goods suppliers;
d) enterprises' strategies orientated to increasing a market share.

Moreover, Kardasz, Stollery (2001) and Feinber (1989) believe that the degree of exchange rate pass-through to domestic prices increases together with the increase in the degree of the substitutability of imported and domestic products. In industries where products are diversified, domestic prices react less to changes in exchange rates. Relative prices of these goods can change only when in the situation of change in the exchange rate consumers will not quickly substitute products manufactured in one country for the products manufactured in another country. This situation is more likely to occur when products on a given market are more diversified (Baldwin&Beiling, 2006). Results of empirical investigations indicate also that the degree of the exchange rate pass-through to prices in foreign trade is positively correlated with the degree of openness of a given economy and the time of the exchange rate change duration. On the other hand, the degree of the exchange rate pass-through is negatively correlated with the exchange-rate changeability (Rowland, 2004).

2. VAR MODEL OF EXCHANGE RATE PASS-THROUGH TO PRICES

This article analyses transmission mechanism of exchange rate to prices in the Hong Kong using the vector autoregression (VAR) model presented by Sims in 1980. A motivation for using the structural VAR approach is that it takes explicit account of the endogenity of the exchange rate and permits the estimation of pass-through to a set of prices, such as import prices, producer prices and consumer prices, simultaneously. This approach was used for the first time by McCarthy in 1999 year who analyzed the phenomenon of exchange rate pass-through to prices in the OECD (Organization of Economic Co-operation and Development) member countries (McCarthy, 1999). In the VAR method the phenomenon of exchange rate pass-through to prices is analyzed with the use of a set of equations, which at the same time eliminates the problem of exogenous explanatory variables (Sims, 1980). Then the estimation of the exchange rate effect on particular price aggregates in the model is isolated from the effect of other factors which the exchange rate may be correlated with.

The point of departure for the model of exchange rate pass-through to prices in Hong Kong is an analysis of the so-called distribution chain proposed by Blanchard in 1982 year. The distribution chain is a series of economic shocks between which a cause and effect relation occurs over the same time unit in which the shock occurred (Blanchard, 1982). Naturally this type of approach must be revised and in the VAR model an appropriate lag length between variables must be taken into account as in economy there is no immediate cause-and-effect relation and the effect always occurs with some lag in relation to the moment at which a given economic shock appears. In the analyzed VAR model, the distribution chain looks as follows.

$$s \rightarrow imp \rightarrow ppi \rightarrow cpi \quad (1)$$

where:

s – exchange rate;
imp – import price;
ppi – producer price index;
cpi – consumer price index.

The next step of the analysis is a measurement of the strength of the exchange rate change pass-through to domestic prices. To this purpose is used the so-called impulse response function, that is a function of a given price aggregate (import prices, producer prices, consumer prices) response to an impulse in the form of the exchange rate change. The coefficient of exchange rate change pass-through to prices after the period t is defined by the following equation [Cholewiński 2008].

$$PT(z)_t = \frac{\sum_{i=1}^{k} \Delta z_{t-i}}{\sum_{i=1}^{k} \Delta s_{t-i}} \tag{2}$$

where;

Δz_{t-i} – change of given price index (import price, producer price, consumer prices), in the period form „*t-i*" to „*t*";

Δs_{t-i} – change of exchange rate, in the period form „*t-i*" to „*t*".

Changes in a given price aggregate equal the values of the impulse response function of the analyzed aggregate to the exchange rate shock, and changes in the exchange rate equal the impulse response function of the exchange rate to the exchange rate shock. Shock occurrence is connected with each of the distribution chain links. However, only in the case of the first chain link (exchange rate) the original shock occurs and in subsequent links the shock results from the transmission of shocks in the former links. Hence, the shock occurring in subsequent chain links can be decomposed into an autonomous part (occurring in a given chain link) and the one transmitted from earlier links. Chain link decomposition is accomplished with the use of Cholesky decomposition matrix. Establishing the strength of the shock transmission is indispensable to analyze the phenomenon of exchange rate pass-through to prices in a more detailed way.

On the basis of the distribution model presented earlier, a VAR model was constructed which analyzes the phenomenon of exchange rate pass-through to prices. This model is a set of 4 equations and it looks as follows:

$$\Delta s_t = \sum_{i=1}^{k} \gamma_{11}^{i} \Delta s_{t-i} + \sum_{i=1}^{k} \gamma_{12}^{i} \Delta imp_{t-i} + \sum_{i=1}^{k} \gamma_{13}^{i} \Delta ppi_{t-i} + \sum_{i=1}^{k} \gamma_{14}^{i} \Delta cpi_{t-i} + \varepsilon_{1t} \tag{3}$$

$$\Delta imp_t = \sum_{i=1}^{k} \gamma_{21}^{i} \Delta s_{t-i} + \sum_{i=1}^{k} \gamma_{22}^{i} \Delta imp_{t-i} + \sum_{i=1}^{k} \gamma_{23}^{i} \Delta ppi_{t-i} + \sum_{i=1}^{k} \gamma_{24}^{i} \Delta cpi_{t-i} + \varepsilon_{2t} \tag{4}$$

$$\Delta ppi_t = \sum_{i=1}^{k} \gamma_{31}^{i} \Delta s_{t-i} + \sum_{i=1}^{k} \gamma_{32}^{i} \Delta imp_{t-i} + \sum_{i=1}^{k} \gamma_{33}^{i} \Delta ppi_{t-i} + \sum_{i=1}^{k} \gamma_{34}^{i} \Delta cpi_{t-i} + \varepsilon_{3t} \tag{5}$$

$$\Delta cpi_t = \sum_{i=1}^{k} \gamma_{41}^{i} \Delta s_{t-i} + \sum_{i=1}^{k} \gamma_{42}^{i} \Delta imp_{t-i} + \sum_{i=1}^{k} \gamma_{43}^{i} \Delta ppi_{t-i} + \sum_{i=1}^{k} \gamma_{44}^{i} \Delta cpi_{t-i} + \varepsilon_{4t} \tag{6}$$

where:

s – seasonally adjusted logarithm of nominal, effective exchange rate index (previous quarter = 100);
imp – seasonally adjusted logarithm of import price index (previous quarter = 100);
ppi – seasonally adjusted logarithm of producer price index (previous quarter = 100);
cpi – seasonally adjusted logarithm of consumer price index (previous quarter = 100);
t – given period;
k – lag length (in quarter).

All the above mentioned time series have a quarterly frequency and cover the period from the first quarter of 1993 to the fourth quarter of 2007. Logarithming of particular model variables aimed at elimination of possible regression between variables. Before the model structural parameters were estimated, it was necessary to isolate a seasonal factor from the time series. The occurrence of the seasonal factor in the time series could lead to difficulties in interpreting changes in a given phenomenon in the analyzed period. To purge the time series from seasonal fluctuations, the X12-ARIMA procedure was applied.

This article analyzes the phenomenon of exchange rate pass-through to domestic prices of imported, production and consumer goods over a short period of time (after 1 quarter) and over

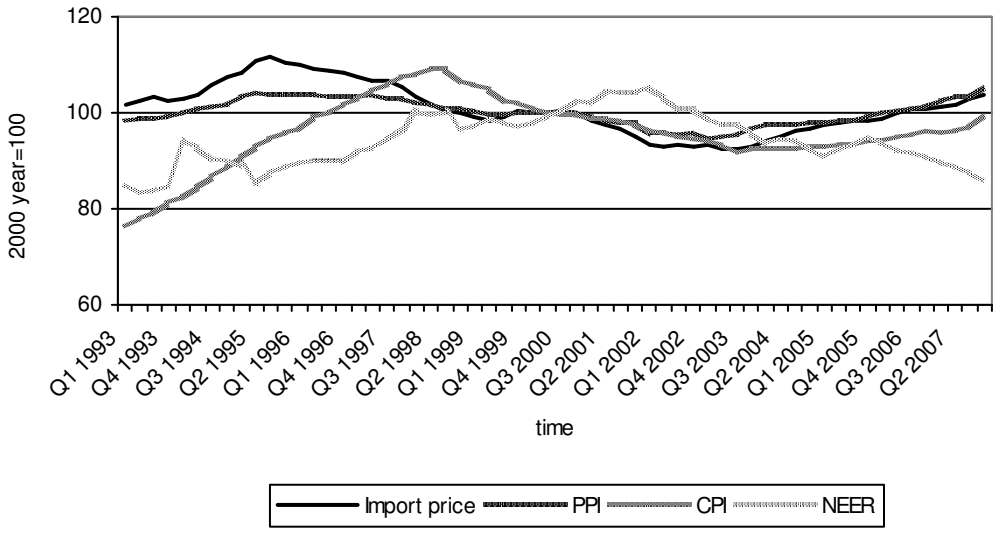
a long period of time (after 4 quarters). For the purposes of the analyses, one lag period (one quarter) between explanatory variables was adopted. The choice of lag lengths is in line with results of the information criteria of the Akaike, Schwartz-Bayesian and the Hannan-Quinn models. According to these criteria, a model with one lag length is characterized by the biggest information capacity.

Before the VAR model estimation it was essential to specify stationarity of the analyzed time series. To this purpose the Augmented Dickey-Fuller Test (ADF) was used. The last step of the analysis of time series was co-integration estimation. Having a set of integrated variables of order 1, a co-integration test was carried out according to the method put forward by Johansen (1988). The choice of the lag lengths for co-integration testing was made on the basis of the earlier mentioned results of the Akaike, Scwartz-Bayesian and Hannan-Quinn information criteria.

3. EXCHANGE RATE PASS-THROUGH TO IMPORT, PRODUCER AND CONSUMER PRICES IN THE HONG KONG

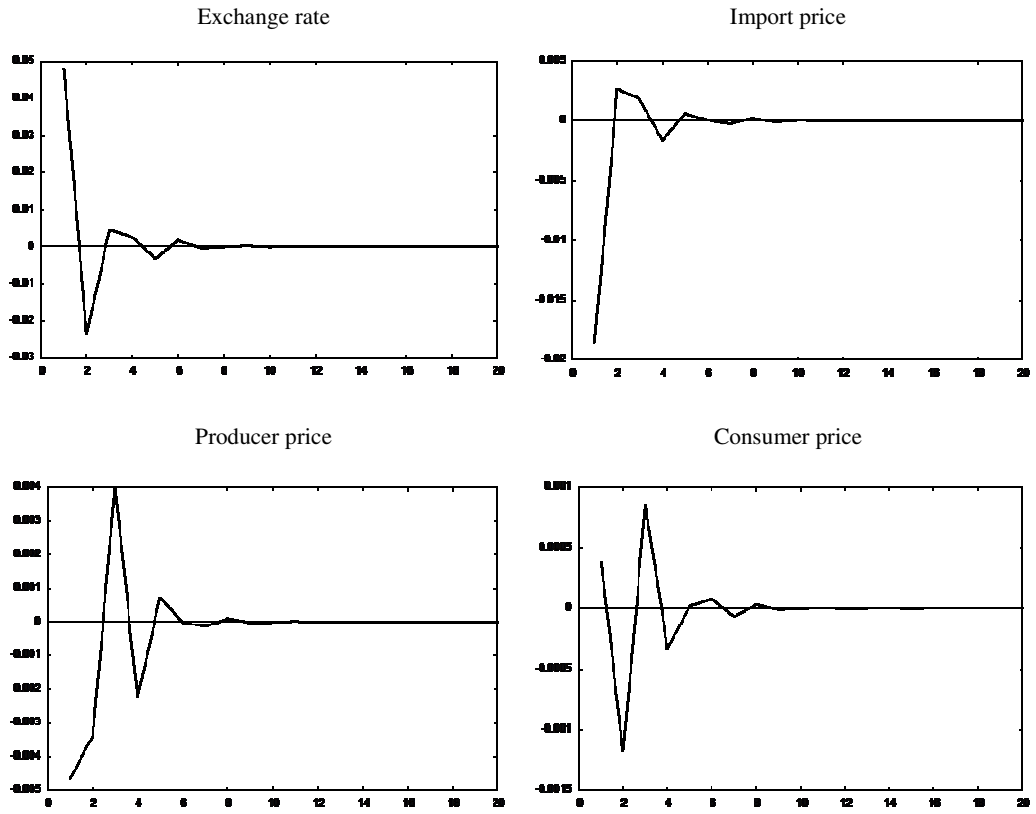
In the period of 1993-2007 took place relatively significant change of nominal, effective exchange rate of Hong Kong dollar and import, producer and consumer prices index. The nominal, effective exchange rate increased above 15% from 1993 year to 2000 year and then decreased to the end of 2007 over 14%. In the same time import prices decreased about 2% and then increased near 4%. On the other hand, producer prices went up about 2% and in the period of 2000-2007 grew up again over 5%, whereas consumer prices increased near 24% to 2000 year and next decreased about 1%.

Figure 1. Exchange rate changes and price dynamics in the period of 1993-2007 [2000 year=100]



The point of departure of the analysis of exchange rate pass-through to prices in Hong Kong was an estimate of structural parameters of the VAR model. Results of the parameter estimate of the VAR model consisting of 4 equations are in the Appendix. Below one can see respective graphs of the impulse response functions of exchange rate, import prices and consumer prices to a one-time unit change of the exchange rate in the Hong Kong.

Figure 2. Impulse response function of exchange rate, import price, producer price and consumer price to a one shock in exchange rate



Source: Own calculations on the basis of *International Financial Statistics*, (2008).

Eventually the magnitude of indices of the exchange rate pass-through to consumer prices is smaller than it is indicated by the analysis of the above drawings because changes in the exchange rate paralleled changes in consumer prices.

Table 1. Indexes of exchange rate pass-through to import prices, producer prices and consumer prices in the Hong Kong

The number of quarter after shock	Import prices	Producer prices	Consumer prices
1	9,6%	9,8%	1,8%
2	26,5%	13,3%	10,5%
3	15,5%	11,4%	7,9%
4	19,2%	12,1%	6,9%
5	18,0%	11,6%	8,2%
6	18,3%	12,0%	7,4%
7	18,3%	11,8%	7,8%
8	18,2%	11,9%	7,6%

Source: Own calculations on the basis of *International Financial Statistics*, (2008).

It is obvious that import prices respond more rapidly to changes in the exchange rates than producer and consumer prices (Hüfner & Schröder, 2002). The average level of exchange rate pass-through to import prices stood at 9,6% in the short-term. However,

after one year only 19,2% of exchange rate changes were passed through to import prices. Producer prices responded less to changes in the nominal, effective exchange rate. The average level of the exchange rate pass-through index to producer prices in the short-term amounted to 9,8%. However, after one year only 12,1% of changes in the exchange rate were passed through. The least affected by exchange rates, both in the short- and long-term, were consumer prices. The average level of exchange rate pass-through index to consumer prices reached 1.8% in the short-term. However, after one year only 6.9% of changes in the exchange rate were passed through to consumer prices. Presented calculations indicate that the unanticipated 10% growth of the nominal effective exchange rate of the Hong Kong dollar, in the course of the first 4 quarters, resulted in an increase of import prices by 1,9%, an increase in the producer price index by 1.2 percentage points and an increase in the inflation rate measured by a consumer price index by 0.7 percentage points.

The last stage of the analysis is the residual component variance decomposition of subsequent price aggregates. This procedure specifies the contribution of the exchange rate shock affecting each of the price variables in accounting for the variances of the individual model variables (price aggregates).

Table 2. The error variance decomposition in the import price equation

The number of quarter after shock	Exchange rate	Import prices	Producer prices	Consumer prices
1	12%	88%	0%	0%
2	16%	78%	1%	4%
3	18%	74%	2%	5%
4	18%	74%	2%	6%
5	18%	74%	2%	6%
6	18%	74%	2%	6%
7	18%	74%	2%	6%
8	18%	74%	2%	6%

Source: Own calculations on the basis of *International Financial Statistics*, (2008).

Table 3. The error variance decomposition in the producer price equation

The number of quarter after shock	Exchange rate	Import prices	Producer prices	Consumer prices
1	17%	2%	82%	0%
2	15%	2%	83%	0%
3	15%	1%	83%	1%
4	14%	1%	83%	1%
5	14%	1%	83%	1%
6	14%	2%	83%	1%
7	14%	2%	83%	1%
8	14%	2%	83%	1%

Source: Own calculations on the basis of *International Financial Statistics*, (2008).

Table 4. The error variance decomposition in the consumer price equation

The number of quarter after shock	Exchange rate	Import prices	Producer prices	Consumer prices
1	0%	6%	2%	92%
2	2%	7%	2%	89%
3	2%	8%	2%	88%
4	2%	9%	2%	87%
5	2%	9%	2%	87%
6	2%	9%	2%	87%
7	2%	9%	2%	87%
8	2%	9%	2%	87%

Source: Own calculations on the basis of *International Financial Statistics*, (2008).

On the basis of the data from the above Tables it can be noted that changes in the nominal effective exchange rate of the Hong Kong dollar accounted for 12% of the price variances in the short-run and for 18% of the price variances in the long-run. However, the effect of changes in the exchange rate on producer price variances in the short- and long-term was similar. In the short-term, ca. 17% of producer price changes could be accounted for by a change in the nominal, effective exchange rate of the Hong Kong dollar and in the long-term – ca. 14% producer price changes.

The role of exchange rate in accounting for consumer price variances in the short- and long-term was even smaller. In the short-term, less than 1% of consumer price changes could be accounted for by changes in the nominal effective exchange rate of the Hong Kong currency. This effect grew insignificantly as the time passed from the moment of change in the exchange rate. In the long-term the change in the nominal effective exchange rate of the Hong Kong dollar accounted for 2% of consumer price changes.

5. CONCLUDING REMARKS

The paper performed an empirical analysis of transmission mechanism of exchange rate changes to import, producer and consumer prices in Hong Kong. The most important findings can be summarized as follows. First, the exchange rate pass-through is incomplete, even in the long-run. Second, the degree and speed of exchange rate pass-through decline along the distribution chain. It means that import prices respond stronger and faster to exchange rate shocks than producer prices and consumer prices. Finally, consumer prices are largely unresponsive to exchange rate shocks. These findings can be interesting from monetary policy point of view. The knowledge that responds of import, producer and consumer prices to changes in exchange rates are relatively low is important for modeling and forecasting the results of changes in exchange rate.)

BIBLIOGRAPHY

BALDWIN J. R., BEILING Y. (2006, *Exchange Rate Cycles and Canada–U.S. Manufacturing Prices*, "Statistics Canada Economic Analysis Research Paper", No 041, June.

BLANCHARD O.J. (1982), *Price Desynchronization and Price Level Inertia*, "NBER Working Paper", No 900.

CHOLEWIŃSKI R. (2008), *Wpływ zmian kursu walutowego na dynamikę procesów inflacyjnych*, "Materiały i studia", nr 226.

CHOUDHRI E.U., HAKURA D.S. (2001), *Exchange Rate Pass-through to Domestic Prices: Does the Inflationary Environment Matter?*, "IMF Working Papers", No.01/194.

DORNBUSH R. (1987), *Exchange rate and prices*, "American Economic Review", No 77.

EDWARDS S. (2006), *The Relationship between Exchange Rates and Inflation Targeting Revisited*, "NBER Working Paper", No. 12163.

HÜFNER F. P., SCHRÖDER M. (2002), *Exchange Rate Pass-Through to Consumer Prices: A European Perspective*, "Center for European Economic Research Discussion Paper", March.

International Financial Statistics, (2008) International Monetary Fund, Washington, July.

KRUGMAN P. (1987), *Pricing to Market When the Exchange Rate Changes*, w: *Real-financial linkages among open economics*, S. W. Arndt, J. D. Richardson (eds.), MIT Press, Cambridge.

McCARTHY J. (1999), *Pass-through of Exchange rates and import prices to domestic inflation in some industrialized countries*, "Bank for International Settlement Working Papers", No 79.

ROWLAND P., Exchange Rate Pass-Through to Domestic Prices: The Case of Colombia, "Banco de la República – ESPE" 2004, No. 47.

SIMS Ch. A. (1980), *Macroeconomics and Reality*, "Econometrica", vol. 1.

TAKAGI S., YOSHIDA Y. (2001), *Exchange Rate Movements and Tradable Goods Prices in East Asia: An Analysis Based on Japanese Customs Data, 1988–1999,* "IMF Staff Papers", Vol. 48, No. 2.

TAYLOR J.B. (2000), *Low inflation, pass-through, and the pricing power of firms*, "European Economic Review", vol. 44.

Appendix

Results of the parameter estimate of the VAR model
Ordinary least squares (OLS) estimates using 40 observations 1993:1-2007:4, 1 quarter lag
Log. likelihood = 829,13768
Covariance matrix = 1,1670925e-017
Akaike Information Criterion (AIC) = -27,1046
Schwartz Bayesian Criterion (BIC) = -24,5180
Hannan-Quinn Criterion (HQC) = -26,8861
Test Portmanteau: LB(15) = 243,082 (df = 244, p-value p 0,181826)

Equation: s

	Coefficient	*Standard error*	*t-ratio*	*p-value*
s_1	-0,359413	0,134493	-2,6723	0,00985
imp_1	-0,140406	0,479185	-0,2930	0,77060
ppi_1	0,255376	0,61223	0,4171	0,67818
cpi_1	0,211752	0,418352	0,5062	0,61473

Equation: imp

	Coefficient	*Standard error*	*t-ratio*	*p-value*
s_1	-0,0679909	0,0367554	-1,8498	0,06962
imp_1	-0,133718	0,130955	-1,0211	0,31160
ppi_1	0,192652	0,167315	1,1514	0,25445
cpi_1	-0,209055	0,114331	-1,8285	0,07280

Equation: ppi

	Coefficient	*Standard error*	*t-ratio*	*p-value*
s_1	-0,0205367	0,0321193	-0,6394	0,52518
imp_1	0,0434247	0,114437	0,3795	0,70578
ppi_1	-0,412912	0,146211	-2,8241	0,00656
cpi_1	0,0578704	0,0999094	0,5792	0,56476

Equation: cpi

	Coefficient	*Standard error*	*t-ratio*	*p-value*
s_1	-0,00810481	0,0403322	-0,2010	0,84146
imp_1	0,339789	0,143699	2,3646	0,02154
ppi_1	0,163215	0,183597	0,8890	0,37781
cpi_1	-0,495473	0,125456	-3,9494	0,00022

Source: Own calculations by using GRETL.

THREE ALTERNATIVE MEASURES OF GREAT MODERATION IN PORTUGAL: THE ROLE OF COMPOSITION VOLATILITY OF GDP

JOSÉ ALBERTO FUINHAS
E-mail: fuinhas@ubi.pt
Department of Management and Economics
University of Beira Interior

Resumo

Neste artigo analisa-se a volatilidade do PIB português na perspectiva da composição do PIB. Investiga-se a volatilidade recorrendo a dados anuais de 1910 a 2007 e a dados trimestrais de 1978:01 a 2007:04. Para testar a robustez da volatilidade, utilizamos três medidas alternativas de crescimento do PIB. Os dados anuais sugerem que se verificaram dois períodos de reduzida volatilidade. O primeiro momento ocorreu no fim dos anos cinquenta e durante os anos sessenta. O segundo momento ocorreu após a segunda metade dos anos oitenta. Os dados trimestrais sugerem uma redução persistente da volatilidade na segunda metade dos anos oitenta e em consonância com o verificado em outras economias. Concluímos que nem a composição, nem as correlações das componentes do PIB tiveram um contributo significativo na volatilidade do PIB. A nossa principal conclusão é que as diferentes medidas da volatilidade utilizadas conduzem a resultados similares.

Palabras-Chave: Volatilidade macroeconómica, Grande Moderação, e economia portuguesa.
Clasificación JEL: C53, F31, G14.
Area Temática: Economía Monetaria y Financiera.

Abstract

This paper provides an analysis of Portuguese GDP volatility. Its focus is on the composition role of GDP components. To understand the nature of Portuguese volatility, the paper recurs to annual data from 1910 to 2007 and quarterly data from 1978:01 to 2007:04. To analyse the robustness of volatility dynamics, we make use of three alternative measures of GDP growth. The annual data, suggest that the phenomena of low volatility occur at last two times. The first episode occurred during later 1950s and 1960s. The second episode occurred after later 1980s. With quarterly data, we detect a persistent fall in output growth volatility after mid eighties in line with evidence for other economies. We conclude that the shares of components neither your correlations had an important contribution to the change of volatility of GDP. Our main conclusion is that alternative measures of volatility send the same picture.

Key words: Macroeconomic Volatility, Great Moderation, Portuguese Economy.
Thematic Area: Monetary and Financial Economics.

1. INTRODUCTION

Since earlier 1980s, numerous countries experienced a considerable decrease in the volatility of aggregate economic activity. The start and the structure of the shrink of volatility were different across countries and have been so widespread, and persistent, that it has been designated the *Great Moderation* [1].

[1] The expression Great Moderation was first conceived by Stock and Watson (2002).

While there are numerous empirical studies on the decline of output volatility for the United States, the corresponding research for other countries is scarce. An apparent consensus still has to emerge, but the review of the literature suggests two main conclusions. First, the decline in output volatility probably reflects a multiplicity of factors. Second, have been ascribed a central role to the *good luck* hypothesis. The prominence of good luck hypothesis in literature leaves relatively little scope for changes in economic policies, particularly monetary policy, and structural changes in the economies.

This paper looks further into the issue by providing a view of the Great Moderation process by analysing the role of composition of GDP of the Portuguese economy. Its main conclusion is that the composition changes have probably played a minor role in the Portuguese Great Moderation. Section 2 introduces the thematic of Great Moderation. Section 3 equates the hypothesis of output composition. Section 4 analyses the different measures of output volatility. Section 5 analyses Portuguese output volatility. Section 6 discusses the steadiness among measures of volatility. Section 7 concludes.

2. THE GREAT MODERATION

The term Great Moderation was first used in United States to describe the break down in volatility of several macroeconomic variables[2], and now is used in the broader context of OECD countries to explain this regime change to a low level of volatility.

It is almost generally accepted that there has been a striking decline in the volatility of aggregate economic activity, at last since the early 1980s[3], in most advanced economies (Davis and Kahn, 2008). Less volatility are notorious for output and employment, at the aggregate level, and verified, as a rule, for expenditure components and industrial sectors. Similar dynamics was verified in inflation and inflation volatility over the same period.

As the Great Moderation has become generally accepted empirically, the phenomenon has begun to be object of major interest, mainly because of its policy implications. The research have two main points of view, namely, whether the fall of volatility result of a long period of luck in the form of soft shocks and could therefore be reversed quickly or is, by other hand, ascribable to changes in the structures of the economies and could be considered permanent.

The most widespread explanations for increased output and inflation stability include better monetary policy, and structural changes in inventory management (Clarida, et al., 2000; Summers, 2005). Other explanations include: 1) financial innovation and increased global integration (Dynan et al., 2006); and 2) evidence that output volatility fell more than sales volatility – interpreted as the result of improved inventory management (Kahn et al., 2002). Some researchers (Stock and Watson, 2002) appoint the hypothesis of "good luck" (smaller exogenous shocks). These approaches are not mutually exclusive how is stressed by Bernanke (2004).

The relative weight of these causes has several implications for economic policy. If the main cause of Great Moderation results of improvements of monetary policy, then policy should be preserved. By other hand, if the low level of volatility happened because of structural changes in the economies resulting, for example, on how inventories of goods and raw materials are managed, then policy must improve the mood of flexibility and innovation. If good luck was the principal source of stability, then policymakers should be prepared to return for the bad past times (Cabanillas and Ruscher, 2008).

---

[2] Some papers have found structural breaks in the volatility of American output (Kim and Nelson, 1999; and McConnell and Perez-Quiros, 2000). Blanchard and Simon (2001) meet a reduction in the volatility of output, but they consider the reduction as a long-term trend rather than as a structural break.

[3] The evidence of an unexpected fall in the volatility of American real GDP growth in the early 1980s initiates the research on the Great Moderation. The early discover of a break in volatility circa 1984 (Kim and Nelson, 1999; McConnell and Perez-Quiros, 2000) inspired the analysis of volatility before and after 1984. Further research show that some economic time series shift volatility dynamics much earlier, some later, and for some, volatility trended down rather than dropped (Davis and Kahn, 2008).

3. THE HYPOTHESIS OF OUTPUT COMPOSITION

There are plentiful evidence that the volatility slowdown. The decline of volatility was explained by five competing theories. First, innovations and improvements in inventory management, such as just-in-time production possibly will have significantly smoothed output (Kahn et al., 2000). The best management of inventories results in a decline in the weight of inventories in sales, principally for durable goods. Second, changes in monetary policy carried out by the Volcker and Greenspan, probably have reduced the effect of economic fluctuations (Boivin and Giannoni, 2006). The lesser Federal Reserve's reaction to output fluctuations relative to inflation may have lead to more stable monetary policy and more stable output growth. Third, the hypothesis of good luck, presumes that the nature of the innovations themselves might have changed, becoming smaller and, in some cases such as oil shocks, less frequent (Ahmed et al., 2004). This explanation assumes a generalized reduction in all sorts of shocks, particularly in high frequency innovations. Fourth, several innovations occurred in financial market regulations (Dynan et al., 2006). Fifth, changes in the composition of the labor market may have reduced economic fluctuations (Jaimovich and Siu, 2009).

The researchers tried to determine the causality behind the decline in macroeconomic volatility. Several papers have investigated the extent to which this phenomenon has encompassed disaggregated data. Specifically, the decrease in volatility is exhibited in both the disaggregated components of output and in industry classifications such as manufacturing (Stock and Watson, 2002; Kim et al., 2004).

A way to learn more about the sources of the decline in output growth volatility is to decompose GDP into its main components. The most obvious starting point to analyse the hypothesis of output composition, was to make a comparison between the actual series of GDP and the series constructed making the shares of their components constant at their value of the first five years of our sample or other span considered of special interest for the analysis pursued. The distance between the recalculated measure of volatility and the actual one gives an indication of the size of the composition effect.

The other possible source of moderation relates to the fact that the volatility of GDP growth depends on the volatility of its individual components, but also on their co-movements. Ceteris paribus, a decrease in the co-movements between GDP components will entail a decrease in the volatility of GDP. To check the magnitude of this correlation effect, we can calculate an estimated measure of the volatility of GDP growth assuming that all pair correlation coefficients between GDP components are held constant at their value of the first five years of our sample or other span considered of special interest for the analysis. The distance between the recalculated measure of volatility and the actual one gives an indication of the size of the correlation effect.

4. THE ALTERNATIVE MEASURES OF OUTPUT VOLATILITY

The issue of the measurement of volatility has received relatively little attention in the empirical literature on the fall in output volatility. The stylised facts are, basically, valid for measures of volatility based on the standard deviation of year on year changes in the variables considered. Other types of indicators could, however, be constructed and stylised facts must be confronted with this alternative volatility indicators and inquired if they pass robustness test of the choice of measurement. By other words, how robust is the analysis to different measures of output volatility?

To analyse the possible sensitivity of the measurement issues, we discusses three alternative measures. The three alternative measures of volatility do not necessarily paint the same picture. The first one, the most used in literature, is the standard deviation of the year on year GDP growth rates. The second one is an obvious simple variation of the basic measure, namely the standard deviation of the quarter on quarter GDP growth rates. The third one is the value of the output gap. The later measure try to capture the idea that a long period of stable, but sluggish

growth away from potential (a rather frequent occurrence in several countries) is not necessarily a sign of low volatility although it will be associated with a low standard deviation of growth. Some authors (Cabanills and Ruscher, 2008) prefer use the absolute value of output gap. Other candidate is the volatility obtained recurring to models of GARCH (generalised autoregressive conditional heteroscedastic) type, but for reasons of scope and space will not be object of analysis.

5. THE ANALYSIS OF PORTUGUESE OUTPUT VOLATILITY

This section presents the analysis of the output volatility process in the Portuguese economy looking at the contribution of GDP components.

We use two frequencies of raw data. Annual series from 1910 to 2007 and quarterly series from 1978:01 to 2007:04. The annual data have fair quality at last until 1952, good quality between 1953 and 1976, and very good quality after that. The quarterly data were subject of several improvements by Portuguese Central Bank (Banco de Portugal). Quarterly data can be considered of high quality.

We appeal to a long period of annual data to give an outlook of the dynamics of Portuguese output growth and yours volatility. The main objective of this approach is to eliminate possible wrong considerations about dampen of volatility that result from the limited span of quarterly data available.

5.1. The data

The raw data came from two main sources: 1) the annual data came from Mateus (1998) until 1994, and Banco de Portugal after that; and 2) the quarterly data came from "Banco de Portugal" (*http:www.bportugal.pt*).

To preserve the additivity property of data, the GDP components were deflated by GDP deflator.

To calculate the potential growth we use Hodrick-Prescott filter with λ=100 for annual data, and λ=1600 for quarterly data.

We identify volatility as the standard deviation of growth rates by one hand, or the value of output gap by other hand. The volatility was calculated to annual or quarterly real output growth or the value of output gap, where real output is chain-weighted GDP. To compute the standard deviations, we use a rolling window of five years for annual data and twenty quarters for quarterly data.[4] So that, the statistic reported for t is the estimated standard deviation over year's t-4 for annual data and quarter's t-19 to quarterly data. We use the same time length to preserve "some" comparability among series.

The effect of output composition in volatility involves what is known by an index-like problem. The volatility dynamics of GDP is the joint result of yours components volatility dynamics and of changes in GDP composition. For this reason, if we make the share of components of GDP fixed, we can identify the contribution of components' volatility to GDP volatility; and if we make components' volatilities fixed this permit identify the contribution of output composition changes to the change in GDP volatility. To separate these two effects we compute a counterfactual series for GDP growth, gotten by holding each components' share constant, and then compare the output volatility schedule obtained using this counterfactual series with the one obtained using the original series.

There are many approaches to compute the hypothesis of constant shares. Specifically, McConnel and Pérez-Quirós (2000) compute their counterfactual series holding each sector's share constant at its sample-wide average, Blanchard and Simon (2001) use the 1947 shares and

[4] In empirical research, most papers use twenty quarters. Some times, for reasons of smoothing volatility, some authors employ forty quarters. For annual data, Daníelsson (2008) makÿÿuse of tÿÿ years.

Stock and Watson (2002) use the 1965 shares. Since the actual and the counterfactual series look very similar to each other, they conclude that composition effects have been of little importance, if any, for the output volatility decline. We opt to use de first five years or twenty quarters of data.

The fixed-weight GDP is calculated as the sum of the year on year growth rates of the GDP components weighted by their average shares in the level of GDP for the first five years available. In the case of inventories, for which the share in GDP is not subject to any drift and is close to zero, and for net exports, the actual contribution to GDP growth is used.

The variance of GDP growth is the sum of the variance of its components (measured in contributions to GDP growth) and of all the pair covariance between these components (again measured in contributions to GDP growth and multiplied by 2). Formally, if $GDP = \sum_{i=1}^{n} Y_i$ then:

$$\sigma_{GDP}^2 = \sum_{i=1}^{n} \sigma_i^2 + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \sigma_{ij}.$$

The covariance between two GDP components depends on their correlation but also on their respective standard deviations. Formally: $\sigma_{ij} = \rho_{ij} \times \sigma_i \times \sigma_j$ (where the σ variables stand for the standard deviations, and ρ variables stand for the correlations between variables). The variance of GDP is recalculated by holding all the pair off ρ_{ij} coefficients constant.

5.2. Annual data analysis

Most studies of the macroeconomic volatility use quarterly data. This option give, when we have a long period, a lot off observations and better statistics resulting from more degrees of freedom. Our employ of annual data result from the shortcoming of time series of quarterly data to be too short for analysis of long run dynamics of volatility, as they start in the first quarter of 1978.

The use of annual rather than quarterly data means that we use much fewer observations on the volatility during each period. In some cases, the analysis and the tests below are based on rather few observations. One consequence of depending on few observations is that observed differences are less likely to be significant at the usual significance levels.

This subsection presents two measures of volatility using annual data, namely the contributions of components to annual GDP growth (in %) and the value of output gap components (in %).

5.2.1. Yearly

The dynamics of the GDP growth changes in the 1970s and began to assuming a clear cyclical pattern (Figure 1). The decline of the growth of GDP started in the 1970s and was noticeable after 1990s.

The decline of volatility started in the early 1930s or at least after the end of Great World War II (Figure 2). This episode was interrupted in mid 1970s with occurrence of 1973 oil crisis and Portuguese Revolution (April 25th, 1974) and the marked contraction happened in the year of 1975 (-4.93%).

The decline of volatility of output growth over the past three decades assumes a marked cyclical shape (Figure 2). Volatility, as measured by the standard deviation of annual GDP growth, fell from 4.26% in the 1970s to 1.62% in the most recent decade (1998-2007). The decline was particularly pronounced in the late 1970s and early 1980s. Since then, output growth volatility has followed a cyclical pattern marked by temporary increases coinciding with periods of cyclical peaks or early phases of downturn. These temporary phases were observed both in the early 1990s and early 2000s. Volatility remained quite low by historical standards in the past few years.

Figure 1: The growth of GDP - 1911 to 2007

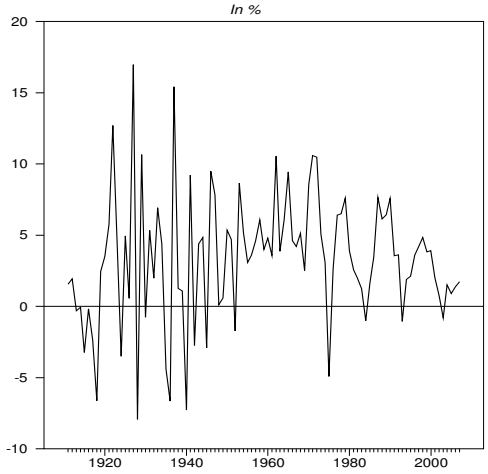


Figure 2: The volatility of GDP - 1915 to 2007

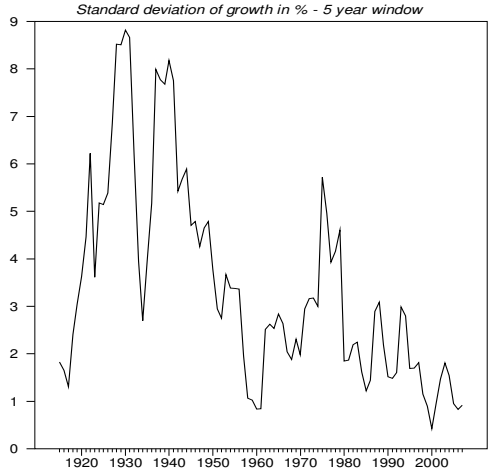


Table 1 displays the standard deviation of the contributions of GDP components to annual GDP growth for the period of 1916 to 2007. The decline in volatility was not broad-based, it result at large extent from the private consumption and at minor amount of the public consumption components. The fall was sharp for private consumption, the most volatile GDP component for the period of 1916 to 1952. In contrast, public consumption does not seem to have played a major role in the moderation process. The pronounced reduction in the volatility of GDP was the result of the outstanding reduction of volatility of the private consumption.

Table 1: Volatility of the GDP components
(Standard deviation of the contributions of components to annual GDP growth – in %)

	(1)	(2)	(3)	(4)	(4) – (1)
Private consumption (a)	6.2324	2.4978	1.6973	2.2549	-3.9775
Public consumption	1.1765	0.4802	0.5955	0.5419	-0.6346
Investment	1.7476	3.0904	2.1032	2.7432	0.9956
Inventories (b)	-	1.8156	1.5864	1.7006	-
Exports	0.9510	1.7246	1.8263	1.7801	0.8291
Imports	1.5426	2.5691	2.5810	2.5944	1.0518
Net exports	1.6257	2.3891	2.3773	2.3925	0.7668
GDP	5.7310	3.1705	2.2638	3.0314	-2.6996

(1) 1916 – 1952 (2) 1953 – 1979 (3) 1980 – 2007 (4) 1953 – 2007

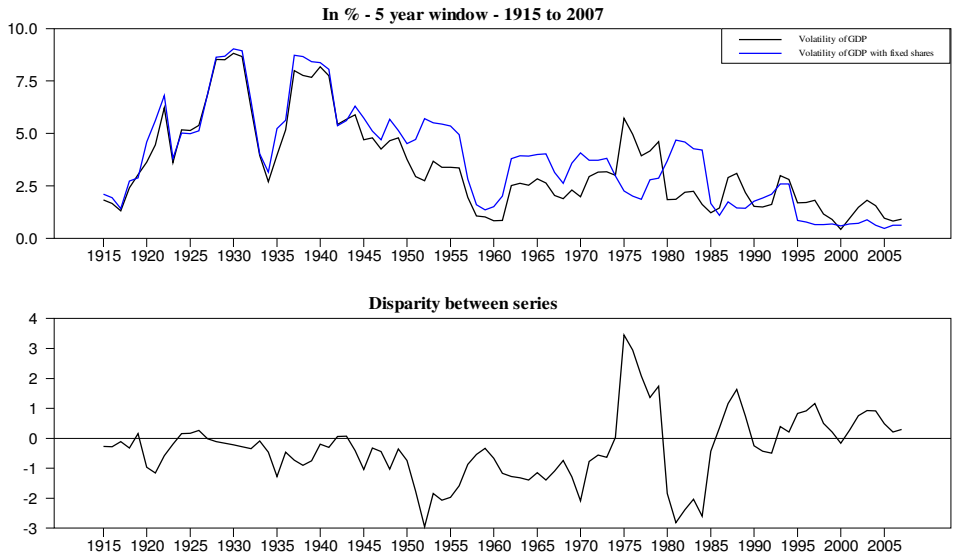
(a) Until 1952 include inventories.

(b) Until 1952 not available.

In addition to the fall in the volatility of its individual components, two other possible sources of moderation in GDP fluctuations should be considered, namely compositional effects and changes in co-movements between components.

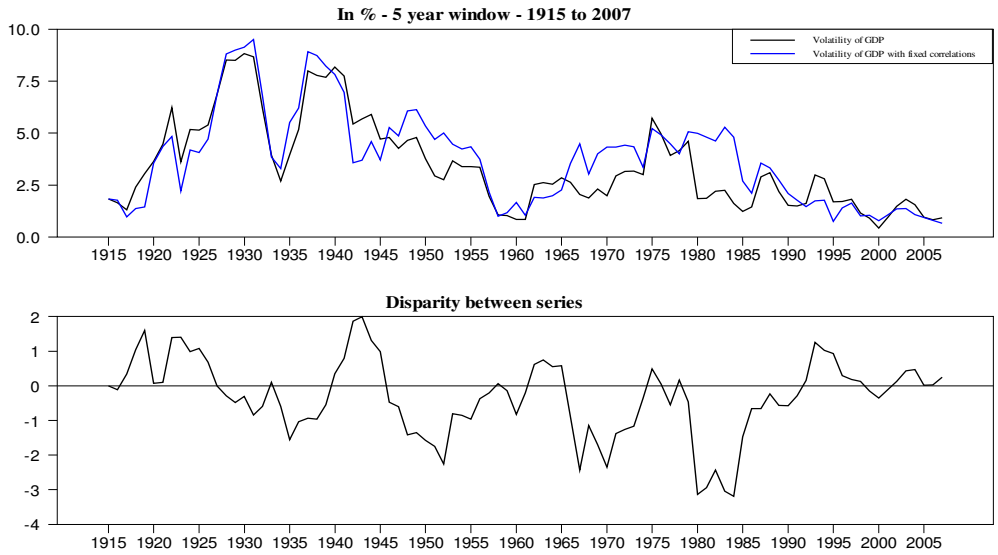
The Figure 3 displays the effects of changes in shares of GDP components on the volatility of GDP growth. The check up of series does not point to any trend of changes in shares of GDP components on the volatility of GDP growth. Although fixed shares components helped to curb GDP volatility temporarily but significantly in the 1950s and 1960s and again in the early 1980s. On all occasions, these dampening effects were subsequently reversed. The two series have a coefficient of correlation of 88.02% and a Theil inequality measure of 0.136972.

Figure 3: Effects of changes in shares of GDP components on the volatility of GDP growth



The Figure 4 displays the effects of changes in the correlations between GDP components on the volatility of GDP growth. The examination of series does not point to any systematic pattern effect although reduced correlation between components helped to curb GDP volatility temporarily but significantly in the late 1960s and again in the early 1980s. On both occasions, these dampening effects were subsequently reversed. The two series have a coefficient of correlation of 87.62% and a Theil inequality measure of 0.135823.

Figure 4: Effects of changes in the correlations between GDP components on the volatility of GDP growth



5.2.2 The value of the output gap analysis

The dynamics of the value of output gap changes in the 1970s, taking an unequivocal cyclical pattern (Figure 5). The decline of the value of output gap began in the 1970s and was clear after 1990s.

The volatility exhibit a U shape between the end of Great World War II and the early

1970s. The decline of volatility started after the end of Great World War II lead to very low volatility in the 1960s. This event was broken up in mid 1970s with occurrence of two shocks: the 1973 oil crisis, and the outcome of Portuguese Revolution.

Clearly, the volatility of output growth declined sharply after later 1970s (Figure 6). Volatility, as measured by the standard deviation of value of output gap, fell from 4.15% in the 1970s to 1.67% in the most recent decade (1998-2007). The decline was particularly pronounced in the late 1970s and early 1980s. Since then, output growth volatility has followed a visible cyclical pattern marked by temporary increases coinciding with periods of cyclical peaks or early phases of downturn. These temporary phases were observed both in the early 1990s and early 2000s. Volatility gives the impression that it is in contraction and stay relatively low by past norm in the last decade.

Figure 5: The value of output gap - 1910 to 2007

in %

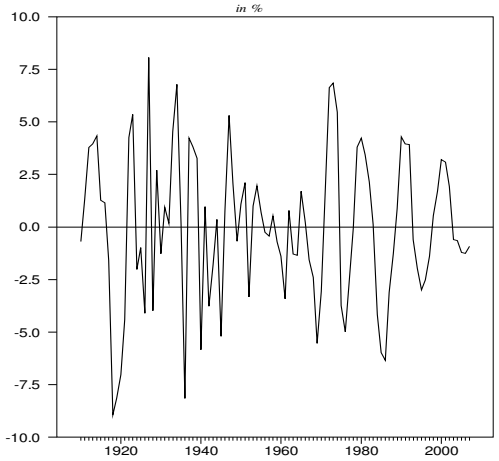


Figure 6: The volatility of GDP - 1915 to 2007

Standard deviation of value of output gap in % - 5 year window

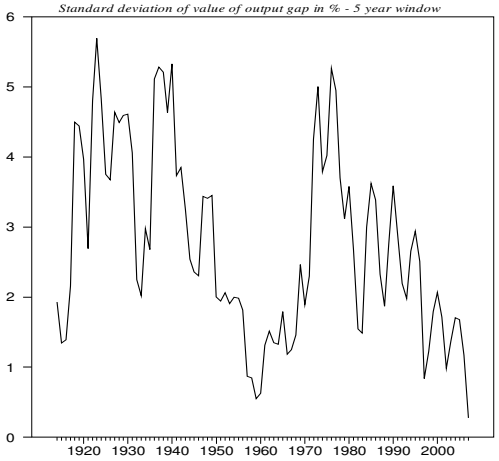


Table 2 displays the standard deviation of the contributions of the value of output gap components for the period of 1915 to 2007. The decline in volatility was restricted to private and public consumption components. The fall was larger for private consumption, the most volatile component of the value of output gap for the period of 1915 to 1952. The volatility of public consumption played a minor role in the moderation process of volatility of output gap. The volatility of investment, inventories, exports, imports, and net exports exacerbated the volatility of output gap.

Table 2: Volatility of the value of output gap components

(Standard deviation of the contributions of the value of output gap components – in %)

	(1)	(2)	(3)	(4)	(4) – (1)
Private consumption (a)	4.9381	1.9203	2.1007	2.0142	-2.9239
Public consumption	1.1495	0.4037	0.7401	0.5992	-0.5503
Investment	1.5992	2.97	2.3063	2.6532	1.0540
Inventories (b)	-	1.5116	1.0891	1.3183	-
Exports	0.7092	1.815	1.7652	1.7915	1.0823
Imports	1.6298	2.3678	2.8098	2.6165	0.9867
Net exports	1.6627	1.662	2.4872	2.1315	0.4688
GDP	4.139	3.0935	2.9371	3.0149	-1.1241

(1) 1915 – 1952 (2) 1953 – 1979 (3) 1980 – 2007 (4) 1953 – 2007

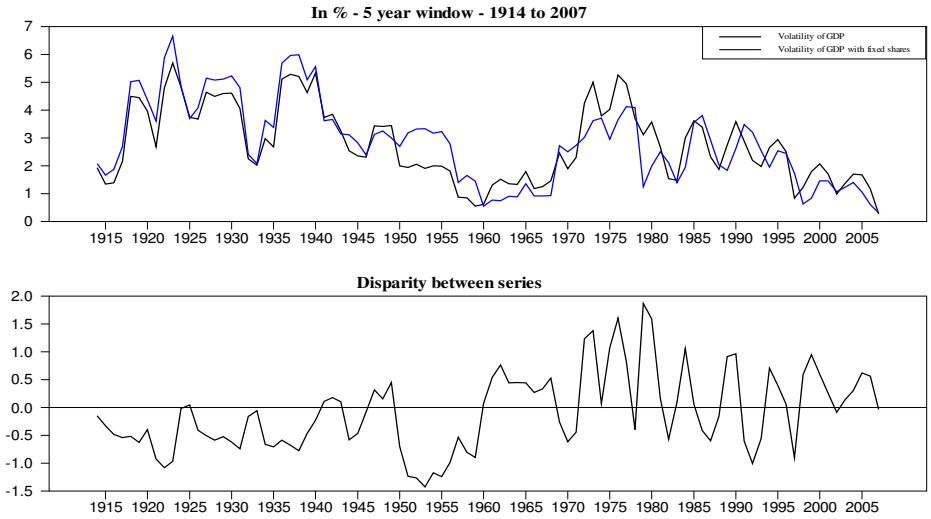
(a) Until 1952 include inventories.

(b) Until 1952 not available.

The Figure 7 displays the effects of changes in shares of GDP components on the volatility of value of output gap. It is clear the marked cyclical dynamics of changes of GDP in the early

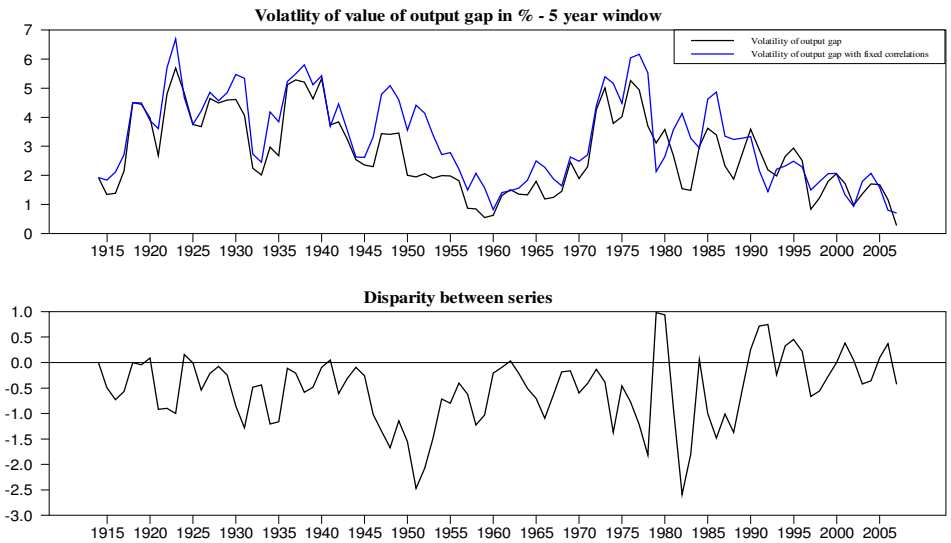
1970s. The two series have a coefficient of correlation of 88.19% and a Theil inequality measure of 0.112717. No trend is detected in disparity between series.

Figure 7: Effects of changes in shares of GDP components on the volatility of value of output gap



The Figure 8 displays the effects of changes in the correlations between GDP components on the volatility of value of output gap. The two series have a coefficient of correlation of 87.63% and a Theil inequality measure of 0.135823. The sample mean of disparity between series is equal to -0.53%, with a t-statistic, for the null hypothesis of mean zero, equal to -7.585012 (significance level lower than 1%). The changes in the correlations between GDP components contribute to lower the volatility of GDP.

Figure 8: Effects of changes in the correlations between GDP components on the volatility of value of output gap



5.3. Quarterly data analysis

We do not make the usual seasonal adjustment of series to preserve yours property of additivity. This is required to build the analysis of constant shares and constant correlations.

This subsection presents three measures of volatility using quarterly data, namely the contributions of components to year on year GDP growth (in %), the contributions of

components to quarter on quarter GDP growth rate annualized (in %), and the contributions of components to annual GDP growth (in %) and the value of output gap components (in %).

5.3.1 Year on year analysis

The volatility of output growth declined substantially in Portugal over the past two decades (Figure 10). Volatility, as measured by the standard deviation of year on year GDP growth, fell from 2.99% in 1980s to 1.76% in the most recent decade (1998:01-2007:04). The decline was particularly pronounced in the late 1990s and early 2000s. Since then, output growth volatility has followed a cyclical pattern marked by temporary increases coinciding with periods of cyclical peaks or early phases of downturn. These temporary phases were observed both in the early 1990s and early 2000s. Volatility remained quite low after mid 1990s.

Figure 9: The growth of GDP - 1979:01 to 2007:04

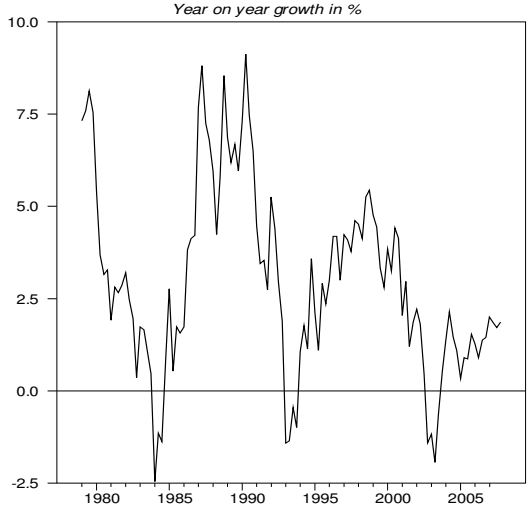


Figure 10: The volatility of GDP - 1983:04 to 2007:04

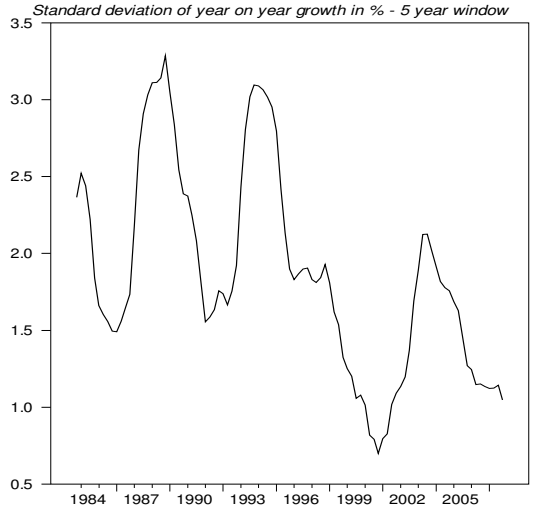


Table 3 displays the standard deviation of the contributions of GDP components to year on year GDP growth for the period 1979:01 to 2007:04. The decline in volatility affects most components. The fall was larger for the most volatile GDP components, particularly inventories, investment, and net exports. Public consumption played an exception worsens the volatility of GDP. The public consumption reflects both the sector's relatively big weight in GDP and its intrinsic instability. The sharp fall in the volatility of exports and imports over the past three decades was not offset by increasing trade openness and the associated rise in the importance the two variables for GDP growth.

As a result, the volatility of the contribution of exports and imports increased significantly over the period although the volatility of the contribution of net trade increased more modestly due to the strong degree of co-movement between exports and imports.

Table 3: Volatility of the GDP components

(Standard deviation of the contributions of components to year on year GDP growth – in %)

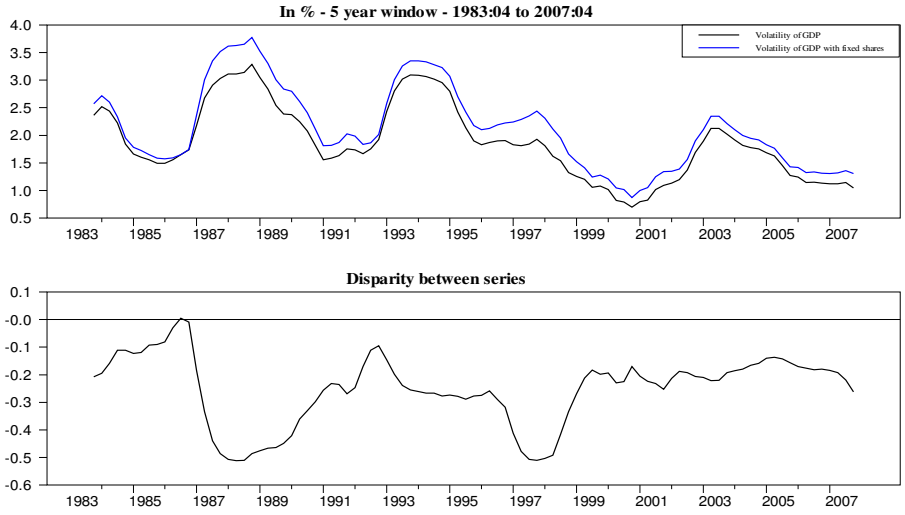
	(1)	(2)	(3)	(4)	(3) – (1)
Private consumption	2.2170	1.7135	1.0190	1.8105	-1.1980
Public consumption	0.5124	0.6175	0.6039	0.6247	0.0915
Investment	3.8261	2.0848	1.3034	2.5132	-2.5227
Inventories	3.3141	1.2221	0.3693	1.8468	-2.9448
Exports	2.7540	2.2047	1.5698	2.3057	-1.1842
Imports	3.5637	2.9560	2.3875	2.9913	-1.1762
Net exports	4.0563	1.9390	1.6465	2.6057	-2.4098
GDP	2.6130	2.4919	1.5613	2.5214	-1.0517

(1) 1979:01 – 1985:04 (2) 1986:01 – 1998:04 (3) 1999:01 – 2007:04 (4) 1979:01 – 2007:04

Increased GDP growth stability could be partly explained by shifts in composition towards more stable GDP components. Figure 11 shows the effect of changes in the composition of GDP by comparing the actual standard deviation of GDP growth to the standard deviation, which would have been obtained if the shares of the various GDP components had been held constant at their average value of the first five years available. The two curves are close, suggesting that composition effects have remained small.

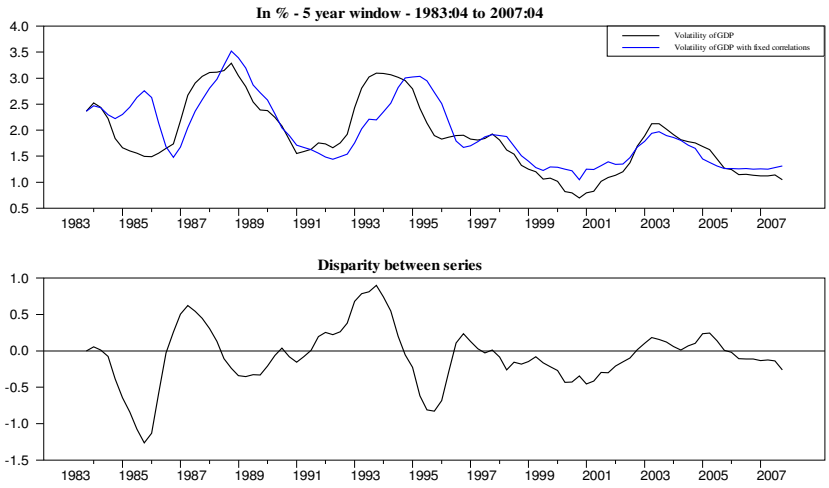
The Figure 11 displays the effects of changes in shares of GDP components on the volatility of GDP growth. The two series have a coefficient of correlation of 98.85% and a Theil inequality measure of 0.066467 (but with a very high bias of 0.804011). The sample mean of disparity between series is equal to -0.25%, with a t-statistic, for the null hypothesis of mean zero, equal to -19.844958 (significance level lower than 1%). The changes in shares of GDP components contribute to lower the volatility of GDP.

Figure 11: Effects of changes in shares of GDP components on the volatility of GDP growth



The Figure 12 displays the effects of changes in the correlations between GDP components on the volatility of GDP growth. The two series have a coefficient of correlation of 81.15% and a Theil inequality measure of 0.099508. No trend or drift is discernible in disparity between series.

Figure 12: Effects of changes in the correlations between GDP components on the volatility of GDP growth



5.4.2 Quarter on quarter analysis

The volatility of output growth declined substantially after later 1980s (Figure 10). Volatility, as measured by the standard deviation of quarter on quarter GDP growth, fell from 4.79% in 1980s to 2.74% in the most recent decade (1998:01-2007:04). The decline was particularly pronounced in the late 1990s and early 2000s. Volatility seems be in fall and reach a historical low mark.

Figure 13: The growth of GDP - 1978:02 to 2007:04

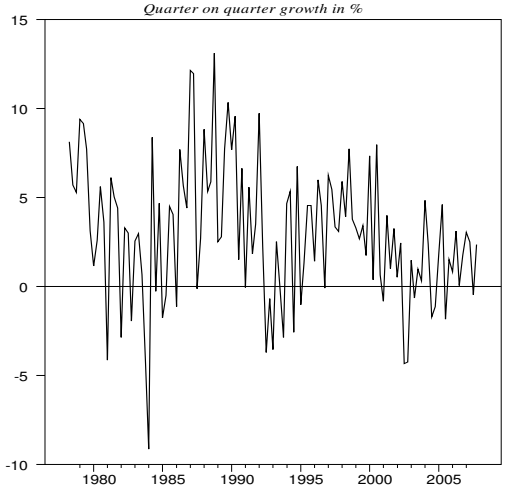


Figure 14: The volatility of GDP - 1983:01 to 2007:04



Table 4 displays the standard deviation of the contributions of GDP components to quarter on quarter GDP growth for the period of 1978:02 to 2007:04. The decline in volatility affects all components. The fall was larger for the most volatile GDP components, particularly inventories and investment. In contrast, public consumption does not seem to have played a major role in the moderation process.

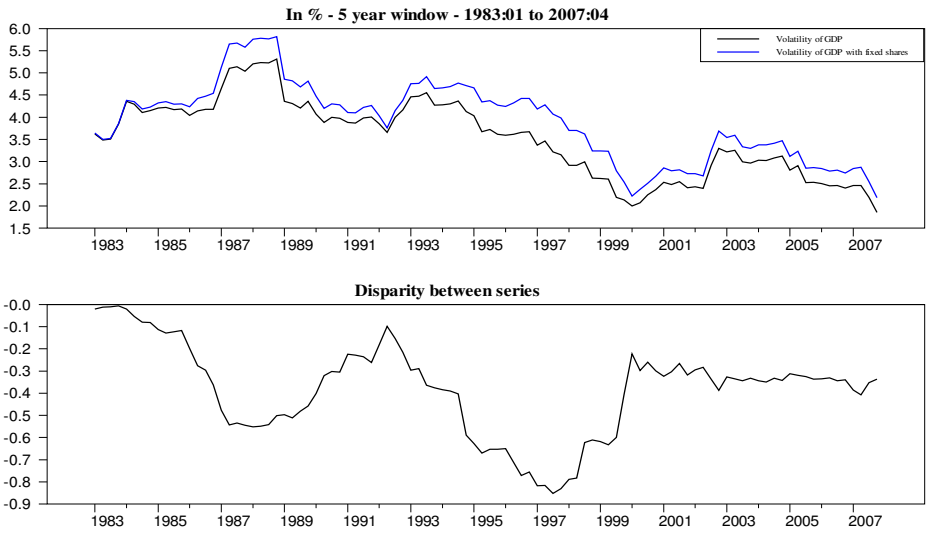
Table 4: Volatility of the GDP components (Standard deviation of the contributions of components to quarter on quarter GDP growth rate annualized – in %)

	(1)	(2)	(3)	(4)	(3) – (1)
Private consumption	3.6568	2.6394	2.1165	2.9026	-1.5403
Public consumption	0.8752	0.7854	0.7518	0.8235	-0.1234
Investment	5.5846	3.0973	2.2868	3.7658	-3.2978
Inventories	4.2441	1.5941	1.1460	2.5047	-3.0981
Exports	3.5509	3.3177	2.8714	3.3260	-0.6795
Imports	6.2173	4.6550	5.0929	5.2421	-1.1244
Net exports	5.5820	3.6793	3.6522	4.2987	-1.9298
GDP	4.2567	3.9725	2.5801	3.8539	-1.6766

(1) 1978:02 – 1985:04 (2) 1986:01 – 1998:04 (3) 1999:01 – 2007:04 (4) 1979:01 – 2007:04

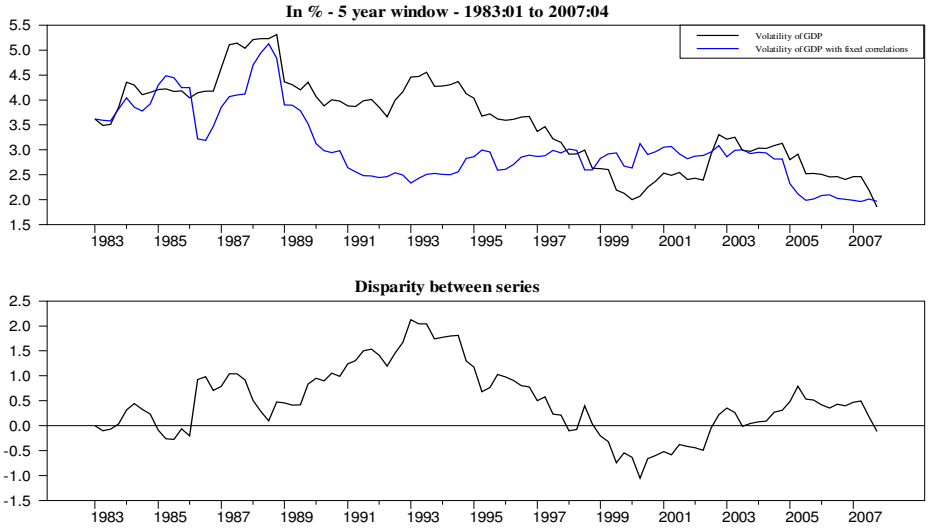
The Figure 15 displays the effects of changes in shares of GDP components on the volatility of GDP growth. The two series have a coefficient of correlation of 97.29% and a Theil inequality measure of 0.056936 (but with a very high bias of 0.778750). The sample mean of disparity between series is equal to -0.38%, with a t-statistic, for the null hypothesis of mean zero, equal to -18.667024 (significance level lower than 1%). The changes in shares of GDP components contribute to lower the volatility of GDP.

Figure 15: Effects of changes in shares of GDP components on the volatility of GDP growth



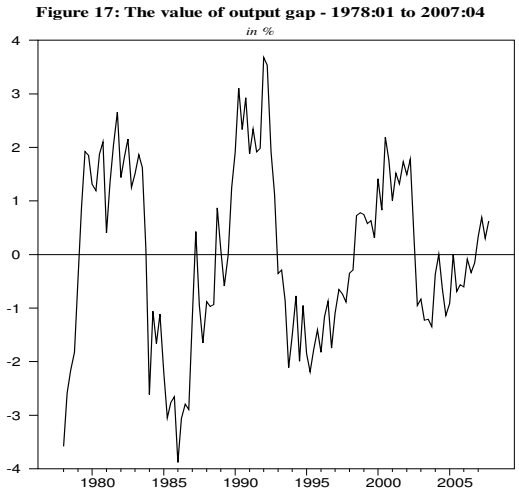
The Figure 16 displays the effects of changes in the correlations between GDP components on the volatility of GDP growth. The two series have a coefficient of correlation of 63.23% and a Theil inequality measure of 0.124336 (but with a high bias of 0.324585). The sample mean of disparity between series is equal to 0.48%, with a t-statistic, for the null hypothesis of mean zero, equal to 6.897572 (significance level lower than 1%). The changes in the correlations between GDP components contribute to higher the volatility of GDP. It is perceptible an apparent stochastic trend in the disparity between series.

Figure 16: Effects of changes in the correlations between GDP components on the volatility of GDP growth



5.4.3 The value of the output gap analysis

The volatility of output growth declined substantially over the past two decades (Figure 18). Volatility, as measured by the standard deviation of quarter on quarter GDP growth, fell from 1.95% in 1980s to 0.95% in the most recent decade (1998:01-2007:04). The decline was particularly pronounced in the late 1990s and volatility remained low in the last decade.



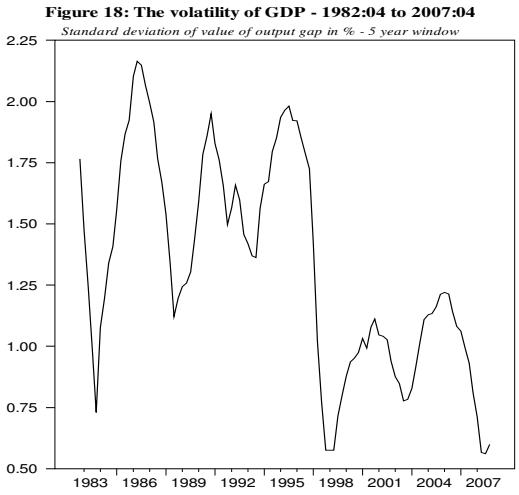


Table 5 displays the standard deviation of the contributions of the value of output gap components for the period of 1978:01 to 2007:04. The decline in volatility was general, affecting all components. The fall was larger for inventories, the most volatile GDP component for the period 1978:01 to 1985:04. In contrast, public consumption does not seem to have played a major role in the moderation process. The sharp reduction in the volatility of GDP was the result of the outstanding reduction of volatility of the net exports, inventories, and investment.

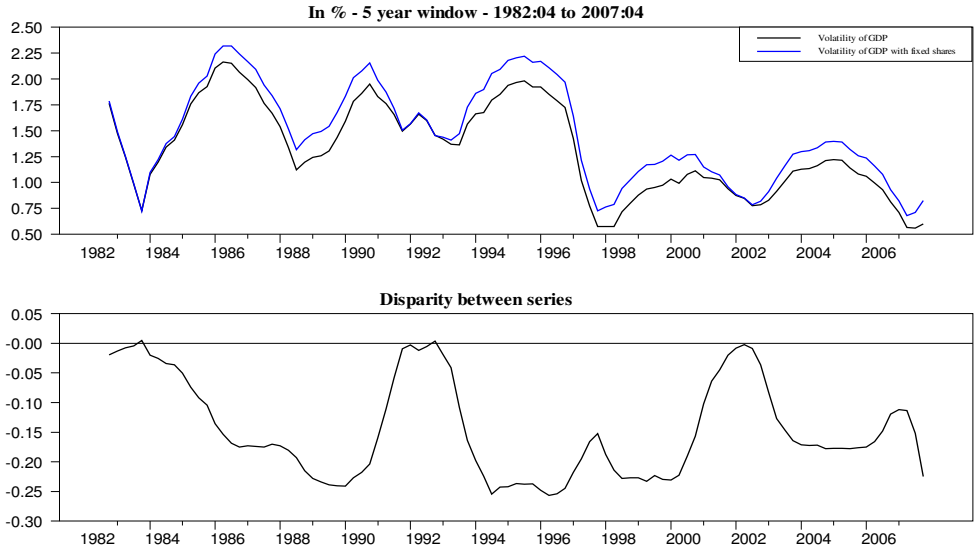
Table 5: Volatility of the GDP components
(Standard deviation of value of the contributions of components to the output gap – in %)

	(1)	(2)	(3)	(4)	(3) – (1)
Private consumption	1,5209	1,2327	0,7049	1,2124	-0,8160
Public consumption	0,3524	0,4495	0,3237	0,3957	-0,0287
Investment	2,2764	1,5791	0,9382	1,6577	-1,3382
Inventories	1,8764	0,7441	0,2588	1,0968	-1,6176
Exports	1,9211	1,5322	1,0308	1,5267	-0,8903
Imports	2,2092	2,1915	1,7876	2,1015	-0,4216
Net exports	2,9056	1,4845	1,2031	1,9164	-1,7025
GDP	1,9425	1,7660	0,9849	1,6407	-0,9576

(1) 1978:01 – 1985:04 (2) 1986:01 – 1998:04 (3) 1999:01 – 2007:04 (4) 1979:01 – 2007:04

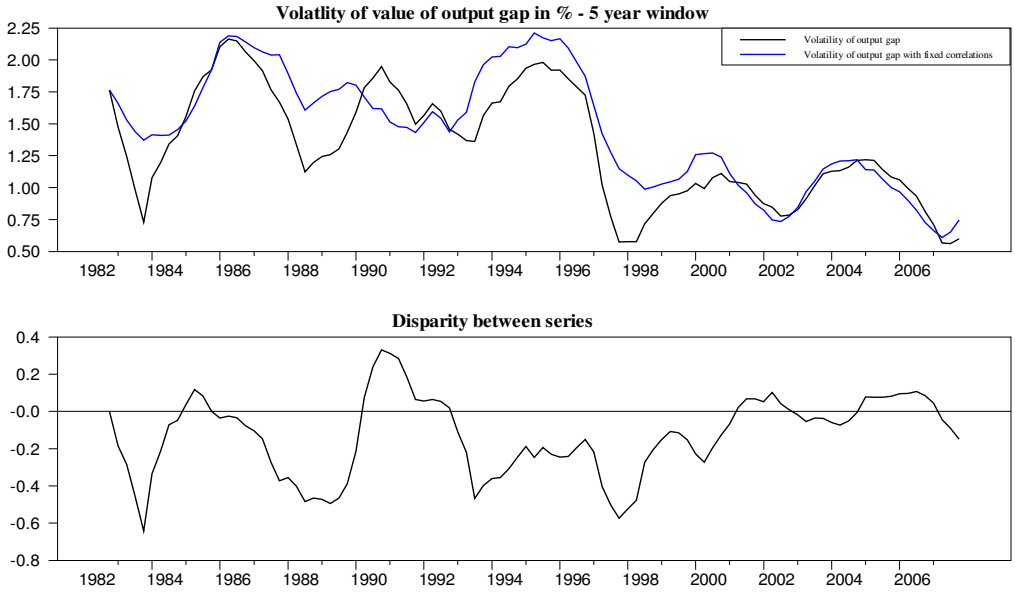
The Figure 19 displays the effects of changes in shares of GDP components on the volatility of value of output gap. The two series have a coefficient of correlation of 98.40% and a Theil inequality measure of 0.056938 (but with a very high bias of 0.757654). The sample mean of disparity between series is equal to -0.15%, with a t-statistic, for the null hypothesis of mean zero, equal to -17.681437 (significance level lower than 1%). The changes in shares of GDP components contribute to lower the volatility of GDP.

Figure 19: Effects of changes in shares of GDP components on the volatility of value of output gap



The Figure 20 displays the effects of changes in the correlations between GDP components on the volatility of value of output gap. The two series have a coefficient of correlation of 88.79% and a Theil inequality measure of 0.084980 (but with a high bias of 0.284661). The sample mean of disparity between series is equal to -0.13%, with a t-statistic, for the null hypothesis of mean zero, equal to -6.308240 (significance level lower than 1%). The changes in the correlations between GDP components contribute to lower the volatility of GDP.

Figure 20: Effects of changes in the correlations between GDP components on the volatility of GDP growth



6. THE STABILITY AMONG MEASURES OF VOLATILITY

The three alternative measures of volatility exhibit diverse graphical dynamics, but conduct to the same conclusions at last: a declining volatility. Table 6 show the correlation matrix of two measures of volatility for annual data from 1915 to 2007. The two alternatives of measure of volatility have the coefficient of correlation of 68.5%. Table 7 display the correlation matrix of three measures of volatility for quarterly data from 1983:04 to 2007:04. The three alternatives measures of volatility have coefficients that indicate a correlation of 84.2% between volatility calculated on year on year and quarter on quarter. The major discordance was between volatility calculated on a basis of year on year and that obtain by the value of output gap that reveal a coefficient of correlation of 39.1%.

Table 6: Correlation matrix of measures of volatility of the GDP
(Annual data from 1915 to 2007)

ANNUAL	yearly	output gap
yearly	1.000000000000	0.684650907612
output gap	0.684650907612	1.000000000000

Table 7: Correlation matrix of measures of volatility of the GDP
(Quarterly data from 1983:04 to 2007:04)

QUARTERLY	year on year	quarter on quarter	output gap
year on year	1.000000000000	0.841573615392	0.391252492591
quarter on quarter	0.841573615392	1.000000000000	0.652190331329
output gap	0.391252492591	0.652190331329	1.000000000000

The three alternative measures of volatility also paint somewhat different pictures when analysing the contributions of GDP components to overall GDP volatility. Inventories appear to play a comparatively more important role in the moderation process in the case of the year on year measure and, to a lesser extent, in the output gap measure. On the other hand, consumption emerges as a more critical component in the case of the quarter on quarter measure.

7. CONCLUSION

The low volatility was not a new phenomenon in Portuguese economy. During 1960s, Portugal experiences what can be considered an exceptional low level of volatility, no matter the measure used.

The moderation of the volatility of output growth can be above all linked to the decline in the volatility of GDP components. There are no persistent and systematic effects of changes in composition (shares) or in correlations among components of GDP in the dynamics of output growth volatility.

When we compare volatility of long run (almost a century of annual data), the private consumption was the major contributor to the decline of volatility. For the quarterly data (nearly three decades), the major contribution was inventories, investment, and net exports.

Public consumption was the minor contributor for the reduction of volatility. When we compare the implications of the three alternative measures of volatility, we find slight differences, but not enough to change our perception of dynamics of volatility phenomenon.

Overall, these examples indicate that caution should be exercised when analysing the reduction of volatility process as, to some extent, conclusions reached may not be fully robust to the type of volatility indicator used.

REFERENCES

AHMED, S., LEVIN, A. AND WILSON, B. A. (2004), "Recent U.S. Macroeconomic Stability: Good Policies, Good Practices, or Good Luck?", *Review of Economics and Statistics*, Vol. 86, No. 3, pp. 824-832.

BANCO DE PORTUGAL (*http://www.bportugal.pt*).

BERNANKE, B. S. (2004), "The Great Moderation", speech at the meetings of the Eastern Economic Association, Washington, February 20.
(*http://www.federalreserve.gov/boarddocs/speeches/2004/20040220/default.htm*)

BLANCHARD, O. AND SIMON, J. (2001), "The Long and Large Decline in US Output Volatility", *Brookings Papers on Economic Activity*, Vol. 2001, No. 1, pp. 135-164.

BODMAN, P. M. (2006), "Output Volatility in Australia", Macroeconomic Research Group, University of Queensland, July.

BOIVIN, J. AND GIANNONI, M. P. (2006), "Has Monetary Policy Become More Effective?", *Review of Economics and Statistics*, Vol. 88, No. 3, August, pp. 445-462

BUCKLE, B., HAUGH, D. AND THOMPSON, P. (2000A), "On Growth and Volatility Regime Switching Models for New Zealand GDP Data", mimeo, New Zealand Treasury.

BUCKLE, B., HAUGH, D. AND THOMPSON, P. (2000B), "Calm After the Storm?: Supply-Side Contributions to New Zealand's GDP Volatility Decline", mimeo, New Zealand Treasury.

CABANILLAS, L. G. AND RUSCHER, E. (2008), "The Great Moderation in the euro area: What role have macroeconomic policies played?", *European Economy*, Economic Papers 331, June.

CLARIDA, R., GALI, J. AND GERTLER, M. (2000), "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory", *Quarterly Journal of Economics*, Vol. 115, No. 1, pp. 147-180.

DANÍELSSON, Á. (2008), "The great moderation Icelandic style", Central Bank of Iceland, Working Papers No. 38, May.

DAVIS, S. J. AND KAHN, J. A. (2008), "Interpreting the Great Moderation: Changes in the Volatility of Economic Activity at the Macro and Micro Levels", *NBER*, working paper, No. 14048.

DYNAN, K. E., ELMENDORF, D. W., AND SICHEL, D. E. (2006), "Can Financial Innovation Help Explain the Reduced Volatility of Economic Activity", *Journal of Monetary Economics*, Vol. 53, No. 1, pp. 123-150.

HAMILTON, J. D., (1989), A new approach to the economic analysis of non-stationary time series and the business cycle, *Econometrica*, Vol. 57, No. 2, pp. 357-384.

JAIMOVICH, N. AND SIU, H. E. (2009), The Young, the Old, and the Restless; Demographics and Business Cycle Volatility, *NBER*, working paper, No. 14063.

KAHN, J. A., MCCONNELL, M. AND PEREZ-QUIROS, G. (2002), "On the Causes of the Increased Stability of the U.S. Economy", Federal Reserve Bank of New York, *Economic Policy Review*, Vol. 8, No. 1, May, pp. 183-202.

KIM, C. J. AND NELSON, C. R. (1999), Has the US Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle, *Review of Economics and Statistics*, vol. 81, No. 4, pp. 608-616.

MATEUS, A. (1998), Economia Portuguesa, Verbo.

MCCONNELL, M., MOSSER, P. AND PEREZ-QUIROS, G. (1999), "A Decomposition of the Increased Stability of GDP Growth", *Current Issues in Economics and Finance*, Federal Reserve Bank of New York, September.

MCCONNELL, M. AND PEREZ-QUIROS, G. (2000), Output Fluctuations in the United States: What has Changed Since the Early 1980s?, *American Economic Review*, Vol. 90, No. 5, pp. 1464-1476.

STEVEN J. DAVIS, S. J. AND KAHN, J. A. (2008), "Interpreting the Great Moderation: Changes in the Volatility of Economic Activity at the Macro end Micro Levels", *NBER*, working paper, No. 14048.

STOCK, J. H. AND M. W. WATSON (2002), "Has the business cycle changed and why?", *NBER*, working paper, No. 9127.

SUMMERS, P., (2005), "What caused the Great Moderation? some cross-country evidence", Federal Reserve Bank of Kansas City, *Economic Review*, Third Quarter, pp. 5-32.

SURPRESAS COM RELAÇÃO À POLÍTICA MONETÁRIA E O MERCADO DE CAPITAIS: EVIDÊNCIAS DO CASO BRASILEIRO

WALTER GONÇALVES JUNIOR
E-mail: walter.goncalves@gmail.com
WILLIAM EID JUNIOR
E-mail: william.eid@fgv.br
Escola de Administração de Empresas de São Paulo

Resumo

Este trabalho avalia os efeitos que as ações de política monetária exercem sobre o mercado acionário brasileiro. Através da definição de medidas para estimar a surpresa causada neste mercado pelas decisões do Comitê de Política Monetária do Banco Central do Brasil (COPOM), constatou-se que a um aumento hipotético de 1% não previsto na taxa básica de juro está associada uma queda de aproximadamente 1,3% no Índice Bovespa. Testes adicionais mostraram que o mercado não reage diferentemente em função da direção tomada pelas decisões e que importantes eventos econômicos recentes ou mesmo o contexto das decisões tomadas não influenciaram as respostas à surpresa.

Palavras-chave: política monetária, Bovespa, surpresas, estudo de evento.
Área Temática: Economia Monetária e Financeira .

Abstract

This article evaluates the effects that monetary policy actions exert on Brazilian stock market. By the measures defined to estimate the surprise caused by Comitê de Política Monetária do Banco Central do Brasil (COPOM) decisions, it was verified that to a hypothetic unexpected 1% increase in the target rate is associated an 1.3% average fall of Bovespa Index. Additional tests did not show distinct reactions caused by direction decisions, neither evidences from relevant recent economic events or decision contexts having influences on the surprise responses..

Key words: monetary policy, Bovespa, surprises, event study.
Thematic Area: Monetary and Financial Economics.

1. INTRODUÇÃO

O crescimento e o desenvolvimento econômico de um país de forma ordenada e sustentável pressupõem uma série de requisitos; o mais evidente deles é a preservação do poder de compra de sua moeda com o controle da inflação. Para tanto, se faz necessária a presença de uma autoridade monetária, figura responsável por atingir este e outros fins ligados à estabilidade monetária.

No Brasil, esse papel institucional é desempenhado pelo Banco Central do Brasil que, dentre os instrumentos disponíveis para tanto, tem empregado mais comumente a taxa básica de juro (e operações de mercado aberto para sua concretização) como forma de obter os efeitos desejados que levam à estabilidade de preços.

Tendo-se em vista os impactos de ampla repercussão na economia brasileira que as decisões sobre a taxa básica promovem, o interesse por elas gerado é crescente, ocupando por vezes local de destaque na imprensa e na literatura acadêmica econômica. Sabidamente, essas mudanças produzem efeitos nas taxas de juros, câmbio, crédito, nas expectativas e também nos preços dos

ativos, esperando-se com isso que, ao afetar a dinâmica destes canais, possa-se influenciar níveis de poupança, investimento e demanda agregada, o que em última instância afetaria a taxa de inflação.

No presente estudo, pretende-se lançar foco no último dos canais mencionados: o preço dos ativos. Atuando através das variações na riqueza dos agentes econômicos, antevê-se que quedas não previstas nas taxas de juros elevem os preços das ações, porque daquelas possivelmente decorreria um crescimento da economia, podendo levar a aumentos nos lucros e dividendos distribuídos pelas empresas, sem mencionar as menores taxas que os descontariam para a apropriação de valores de mercado; trata-se portanto de um canal muito relevante, especialmente em países como Estados Unidos e Inglaterra.

Constitui-se então em um interessante objeto de estudo observar-se como o mercado acionário brasileiro se comporta segundo os estímulos proporcionados pelas decisões do COPOM; pretende-se assim mensurar e analisar as respostas do mercado acionário às decisões do COPOM sobre a meta para a taxa básica de juro, tomando-se para tanto o índice agregado da Bolsa de Valores de São Paulo (BOVESPA).

A relevância do tema é clara: sob a ótica da autoridade monetária, possuir estimativas confiáveis das reações dos mercados em seus diversos instrumentos é uma informação crítica para a formulação de suas políticas, bem como para a definição das decisões; pela perspectiva dos participantes do mercado financeiro, há um interesse igualmente acentuado - dada a influência exercida pela política monetária sobre as expectativas dos mercados, conhecer tais estimativas representaria uma relevante informação para a construção de posições e estratégias, tanto de investimento como de *hedge*, bem como para a gestão dos riscos implícitos às operações normalmente desenvolvidas em seu dia-a-dia.

2. REVISÃO BIBLIOGRÁFICA

Há uma ampla literatura para estudar os efeitos da política monetária, especialmente aqueles intervenientes nas estruturas das taxas de juros de curto prazo; Cook e Hahn (1989), constataram uma resposta intensa das taxas de curto prazo e fraca nas de longo prazo às mudanças nas taxas-alvo dos *Fed Funds* da década de setenta.

Roley e Sellon (1995), analisando as taxas de longo prazo, encontraram evidências que estas se antecipam às mudanças na política monetária e, em Roley e Sellon (1998a), estimaram respostas significativas das taxas de juros de longo prazo.

Há ainda diversos outros trabalhos que tratam da relação política monetária *versus* taxas de juros, freqüentemente se valendo da abordagem por estudo de evento, tal como nos trabalhos retro referidos; em especial, Kuttner (2001) produziu um estudo similar, porém tentando isolar as componentes esperada e não esperada (ambas não diretamente observáveis) das mudanças nas taxas dos *Fed Funds* causadas por decisões do FOMC (*Federal Open Market Committee*). Ele encontrou uma relação robusta entre as surpresas com relação à política monetária e as taxas de juros de diferentes maturidades.

Outros estudos estabeleceram foco nos impactos da política monetária sobre o mercado acionário; Smirlock e Yawitz (1985) verificaram como incrementos exógenos nas taxas de redesconto reduziram os retornos calculados pelo índice NYSE na primeira metade da década de oitenta.

Thorbecke e Alami (1994) constataram uma resposta significativa dos retornos das ações na segunda metade da década de setenta (representadas pelos índices *Dow Jones* e S&P) às mudanças na taxa-alvo pelo FOMC; Thorbecke (1997) ampliou este estudo usando diferentes abordagens (entre elas o estudo de evento) abrangendo o período de 1987 a 1994 (essencialmente quando a direção do *Fed* foi de Alan Greenspan), obtendo resultados qualitativamente similares aos de seu trabalho anterior.

Estudando um período mais restrito (de 1989 a 1992), mas com uma quantidade maior de instrumentos financeiros sob análise, Reinhart e Simin (1997) verificaram também tais

respostas, ressalvando apenas que tais reações pareciam persistir por mais alguns dias além daquele subseqüente à mudança, fato este explicado pela prática de então do *Fed* em não divulgar ao público suas decisões.

Outros trabalhos focando a relação entre mercado acionário e política monetária se seguiram, com abordagens e períodos similares; por vezes, obtiveram-se respostas significativas dos preços das ações à política monetária, como em Bomfim (2003), que também verificou um aumento anormal da volatilidade dos mercados em dias de reunião; por outro lado, como em Roley e Sellon (1998b) e Bomfim e Reinhart (2000), respostas não significativas também foram encontradas.

Artigos mais recentes como o de Rigobon e Sack (2004), que usaram estimadores para explorar a heterocedasticidade introduzida por ações de política monetária, ou o de Craine e Martin (2003), que desenvolveram um modelo de fatores multivariado para respostas dos ativos a choques de informação comuns e não observáveis, resultaram também em respostas significativas do mercado acionário às ações do *Fed*, com conclusões parecidas aos estudos de evento convencionais.

Bernanke e Kuttner (2005), a partir do recurso empregado por Kuttner (2001) para isolar a componente inesperada ou efeito-surpresa nas mudanças da taxa-alvo dos *Fed Funds*, encontraram também respostas significativas e mais intensas dos preços das ações a esta componente – uma tentativa portanto de isolar, dentro das variações nas taxas-alvo, o que o mercado financeiro consegue antecipar (que pela hipótese dos mercados eficientes, não deveria causar respostas), daquilo que não é possível prever.

Neste sentido, Poole, Rasche e Thornton (2002) já haviam observado que estudos de evento como este utilizam especificações sujeitas a erros de medição nos regressores, havendo ainda problemas de correlação entre estes últimos e os erros; eles propuseram então mudanças que transformavam a análise em um problema típico de erros nas variáveis, obtendo respostas parecidas com os trabalhos anteriores e concluindo que tais estudos de evento geram resultados que, embora significativos, possivelmente subestimam a intensidade destas respostas.

Com relação ao observado no Japão, Honda e Kuroki (2006), conduzindo um estudo de evento para o período de 1989 a 2001, verificaram respostas significativas para a surpresa com relação às decisões do BOJ (*Bank of Japan*) sobre política monetária, tanto na estrutura a termo das taxas de juros quanto nos preços dos ativos, portanto resultados similares aos encontrados para os mercados norte-americanos.

No Brasil, através de um estudo de evento para o período de 1996 a 2003, Tabak e Tabata (2004a) não encontraram evidências de que a estrutura a termo de juro brasileira responda a ações de política monetária, medindo-se pelas mudanças na meta da taxa básica de juro do Banco Central; encontraram contudo um certo grau de antecipação dos mercados a tais ações, o que explicaria tais resultados. Posteriormente, ampliaram, o âmbito de tal trabalho, decompondo a meta para a taxa básica de juro em meta antecipada e não-antecipada; assim, Tabak e Tabata (2004b) em um novo estudo de evento para o período de 2000 a 2003, confirmaram o grau de antecipação anteriormente visto e constataram respostas significativas da estrutura a termo à componente não-antecipada.

Por sua vez, Gonçalves e Guimarães (2005), adotando a abordagem sugerida por Rigobon e Sack (2004) de explorar a heterocedasticidade introduzida pelas ações do Banco Central, usaram as taxas de juros anuais para verificar as respostas produzidas nas taxas do mercado de câmbio, encontrando que incrementos inesperados nos juros levaram a depreciações na taxa de câmbio, observando-se dados de 2000 a 2005, sob significâncias moderadas a fracas, conforme as subjanelas históricas selecionadas do período em questão.

Por fim, Robitaille e Roush (2005) empregaram estudos de evento adaptando as metodologias propostas por Kuttner (2001) e Gürkaynak (2005) a partir de dados *intraday* referenciados ao período de 1999 a 2004, obtendo respostas significativas tanto nos *spreads* dos C-Bonds brasileiros como no Índice Bovespa às decisões do FOMC e a alguns anúncios de indicadores macroeconômicos americanos, evidenciando assim o grau de relacionamento entre tais mercados (brasileiro e americano) e seus comportamentos.

3. ESTIMANDO A SURPRESA DOS MERCADOS

De modo geral, os mercados financeiros estão constantemente incorporando novas informações a seus preços; os agentes ditos racionais utilizam toda a informação disponível em suas previsões, tomando valores passados e correntes de variáveis exógenas ou não (Lucas e Sargent, 1996).

Devido a essa característica dos mercados e em se tratando especificamente da Bolsa de Valores de São Paulo (onde Perobelli *et al.* (2000) identificaram uma eficiência informacional semiforte), verificar-se suas respostas às ações de política monetária pode ser uma tarefa particularmente difícil, porque cada decisão tomada pelo COPOM é estudada por esses agentes, que antecipam em alguma medida seus resultados; assim sendo, conforme observou Dewbre (1981), quaisquer mudanças verificadas nas variáveis correntes e futuras através do tempo decorrerão exclusivamente de mudanças não antecipadas.

Isto posto, ao analisarmos uma relação direta entre as variações na meta da taxa de juro e as do Índice Bovespa devemos encontrar baixa significância entre essas variáveis; isto leva, portanto, à necessidade de distinguir na variação da meta suas duas componentes principais: a antecipada e a não antecipada. Uma ferramenta interessante para se atingir tal intento é empregar a metodologia desenvolvida por Kuttner (2001) aplicada aos contratos futuros de taxa média de depósitos interbancários de um dia (referidos comumente por DI 1 dia), negociados na Bolsa de Mercadorias e Futuros de São Paulo (BM&F).

Devemos, portanto, construir uma equação que, além de representar esta propriedade, ao mesmo tempo expresse adequadamente as idiossincrasias deste mercado; uma expressão adequada – a exemplo do adotado em Garcia e La Rocque (1994) - seria:

$$PU_t = E_t\left(\frac{100\,000}{1+i_p}\right),$$

onde: **(1)**

PU_t = Preço unitário do contrato DI 1 dia, em pontos, correspondente a 100.000, descontado pela taxa de juro implícita ao contrato;

i_p = Taxa de juro esperada referente ao período considerado.

Se considerarmos que $(1 + i_p) = (1 + \bar{i}_m)^t$, onde $\bar{i}_m = E_t(i_m)$ é uma taxa média esperada ao dia útil para todo o período (até a liquidação do instrumento) e t é o número de dias úteis até a liquidação, então teremos:

$$PU_t = E_t\left(\frac{100\,000}{1+i_p}\right) = \frac{100\,000}{(1+\bar{i}_m)^t}$$

(2)

Suporemos que no dia **d–1** ocorre a reunião do COPOM e que no dia **útil** seguinte (que chamaremos de **d**) os mercados abrem já cientes da nova meta para a taxa básica de juro, já que em **d** ocorre a implantação da nova meta. Deste modo, em **d–1** espera-se que a taxa média diária passe de $\bar{i}_a$ (antes da reunião) para $\bar{i}_m$ (após a reunião) e que assim permaneça até a liquidação do instrumento; no entanto, em alguma medida eles são surpreendidos pelo COPOM e esta surpresa (que denominaremos Δi^u) passa a ser incorporada ao valor médio diário anteriormente previsto: os mercados se ajustam a eventuais discrepâncias em relação a suas projeções da véspera.

Portanto, empregando os contratos DI 1 dia com duração de 1 mês, teremos:

$$\underset{\text{(expectativas no dia da reunião)}}{\overset{\text{em } d-1}{\Rightarrow}} \frac{100\,000}{PU_{d-1}} = (1+\bar{i}_a)\cdot(1+\bar{i}_m)^D \qquad \textbf{(3)}$$

$$\underset{\text{(expectativas no dia da implantação)}}{\overset{\text{em } d}{\Rightarrow}} \frac{100\,000}{PU_d} = (1+\bar{i}_m)^D \cdot (1+\Delta i^u)^{\frac{D}{252}} \qquad \textbf{(4)}$$

Da relação entre **(4)** e **(3)** obtemos:

$$(1+\Delta i^{u}) = \left[(1+\bar{i}_{a}) \frac{PU_{d-1}}{PU_{d}} \right]^{\frac{252}{D}} \qquad \textbf{(5)}$$

Para a obtenção de uma estimativa para $\bar{i}_a$, é preciso supor-se que na véspera do dia da reunião (antevéspera do dia do evento, portanto) as expectativas sejam similares às do dia em que ela ocorre; isto equivaleria a afirmar que $E_{d-2}(i_m) \cong E_{d-1}(i_m) = \bar{i}_m$ e que $E_{d-2}(i_a) \cong E_{d-1}(i_a) = \bar{i}_a$.

Logo, $\frac{100\,000}{PU_{d-2}} = (1+\bar{i}_a)^2 \cdot (1+\bar{i}_m)^D$, à qual relacionada com **(3)** implica em $\frac{PU_{d-1}}{PU_{d-2}} = (1+\bar{i}_a)$

Substituindo-se este resultado em **(5)**, temos então

$$(1+\Delta i^{u}) = \left(\frac{PU^{2}_{d-1}}{PU_{d-2} . PU_{d}} \right)^{\frac{252}{D}} \qquad \textbf{(6)}$$

Portanto, **(6)** permite que estimemos a componente não prevista ou inesperada Δi^u pelos mercados na variação da meta da taxa básica de juro por decisão do COPOM. Esta medida apresenta, contudo, o inconveniente de necessitar de PUs negociados em até dois dias úteis anteriores ao dia do evento, o que constituiria um problema com dados de decisões em início de mês; esta incongruência, que ocorre em somente duas observações, será tratada em mais detalhes a seguir, na descrição dos dados.

A obtenção da componente esperada, conseqüentemente, virá de:

$$\Delta i^{e} = \frac{(1+\Delta i)}{(1+\Delta i^{u})} - 1 \qquad \textbf{(7)}$$

, onde**:**

Δi = Variação na meta da taxa básica de juro por decisão do COPOM (base 252 dias úteis);

Δi^e = Componente esperada da variação na meta da taxa básica de juro (base 252 dias úteis);

Δi^u = Componente não-esperada da variação na meta da taxa básica de juro (252 dias úteis).

Para conveniência nas especificações e testes a serem conduzidos doravante, os dados podem ser transformados para viabilizar o emprego do regime de capitalização contínua nas taxas. Logo, é possível fazer:

$I^u = \ln(1+\Delta i^u)$: Surpresa ou componente não esperada (regime de capitalização contínua);

$I^e = \ln(1+\Delta i^e)$: Componente esperada (regime de capitalização contínua);

$I = \ln(1+\Delta i)$: Variação na meta da taxa básica (regime de capitalização contínua).

Note-se que, conforme **(7)** e segundo esta forma de tratar as taxas, temos que $\mathbf{I^u + I^e = I}$.

3.1. Respostas do mercado de ações à política monetária

Trata-se portanto de um estudo de evento visando-se a análise do relacionamento entre as componentes citadas da variação da meta da taxa básica de juro e a variação do mercado entre os dois instantes selecionados para a captura da surpresa. Isso pode ser feito com a especificação:

$$\Delta P_t = \alpha + \beta_u . I^u_t + \beta_e . I^e_t + \varepsilon_t \qquad \textbf{(8)}$$

As hipóteses nulas a serem testadas são $\beta_e = 0$ e $\beta_u = 0$. Para tanto, temos que:

ΔP_t = É o logaritmo natural da razão entre cotações de fechamento do Índice Bovespa, tomando-se os dois pregões consecutivos correspondentes à mudança na meta para a taxa básica de juro;

I_t^u , I_t^e = Componentes surpresa e esperada como supra definido;

ε_t = Termo de erro representativo de fatores outros não ligados à política monetária.

Seguindo aqui a mesma premissa adotada por Bernanke e Kuttner (2005), assumiremos a ortogonalidade do termo devido ao erro (ε_t) em relação às mudanças e surpresas assim calculadas, discutindo posteriormente as limitações e a viabilidade de tal assunção.

Será interessante também comparar os resultados da especificação **(8)** com uma outra que analise a resposta do mercado à variação da meta propriamente dita. Esta especificação seria:

$$\Delta P_t = \alpha + \beta . I_t \quad \textbf{(9)}$$

Tendo I_t já sido definida anteriormente.

Por fim, para evitar que os resultados sejam influenciados por observações que tenham impacto excessivamente pronunciado em relação às demais observações da amostra, é imprescindível identificá-las, promovendo-se então os testes a serem conduzidos **com** e **sem** a sua participação na amostra, explorando assim a robustez dos resultados. Bernanke e Kuttner (2005) sugeriram empregar a seguinte estimativa:

$$IE_i = \Delta\beta_i^T \ \Sigma^{-1} \ \Delta\beta_i \quad \textbf{(10)}$$

Onde:

IE_i = Influência estatística da observação i ;

$\Delta\beta_i$ = Mudança no vetor dos coeficientes da regressão **(8)** - ou seja, nos valores α, β_u e β_e - resultante da exclusão da observação i da amostra;

Σ = Matriz Variância-Covariância dos coeficientes da regressão **(8)**.

Desta maneira, as observações cujas influências estatísticas apresentarem uma disparidade clara em relação às demais serão consideradas *outliers*; isso permitirá a realização dos dois ensaios distintos supra mencionados: com e sem os *outliers*.

3.2. Quebras estruturais e assimetria nas respostas

O passo seguinte será investigar como alguns eventos relevantes na história econômica recente, tais como a adoção do câmbio flutuante, as taxas em expressão anual ou mesmo aspectos contextuais das decisões (se extraordinárias, ou se tomadas em reuniões regulares) afetaram estas respostas. Outra possibilidade a ser explorada é se o mercado acionário reage distintamente conforme a direção tomada pela decisão do COPOM (elevação, manutenção, queda ou reversão da tendência anterior).

Estas questões podem ser estudadas a partir da introdução de variáveis *dummy* interativas com a surpresa, definidas para captar mudanças no comportamento das respostas conforme o aspecto a ser focado. A especificação genérica para conduzir tais testes será:

$$\Delta P_t = \alpha + \beta_u . I_t^u + \beta_e . I_t^e + \beta_{ef} . \text{DummyEfeito}_t . I_t^u + \varepsilon_t \quad \textbf{(11)}$$

Onde **DummyEfeito** é a *dummy* interativa que assume valor igual a 1 quando o efeito ou evento estudado se verifica; caso contrário, assume valor zero. No presente trabalho serão utilizadas:

- **Flexibilização do câmbio**: a *dummy* assume valor igual a 1 se a observação for posterior a 15/01/1999; caso contrário, assume valor zero;
- **Adoção de taxas ano-*over***: a *dummy* assume valor igual a 1 se a observação for posterior a 02/01/1998; caso contrário, assume valor zero;
- **Contexto das decisões**: a *dummy* assume valor igual a 1 se a observação for resultante de uma reunião regular; caso contrário, assume valor zero;
- **Aumento na meta**: igual a 1 se a meta foi aumentada; caso contrário, assume valor zero;
- **Redução na meta**: igual a 1 se a meta foi reduzida; caso contrário, assume valor zero;
- **Inatividade na meta**: igual a 1 se a meta foi mantida sem alterações; caso contrário, assume valor zero;

➢ **Reversão na meta**: igual a 1 se houve uma reversão na tendência da meta (uma queda após uma sucessão de aumentos, ou o oposto); caso contrário, assume valor zero.

4. DADOS

O presente estudo depende, para suas estimativas, das variações na meta da taxa básica de juro, extraídas das Atas de Reunião disponíveis no *website* do Banco Central do Brasil; são utilizadas também as variações do Índice Bovespa, calculadas entre os dois instantes usados no cálculo das surpresas, a partir de suas cotações de fechamento diárias compreendidas entre junho de 1996 (início da sistemática de decisões de política monetária pelo COPOM) e março de 2006.

Totalizaram-se assim 125 observações, das quais a 17ª (de 22/10/1997, época na qual se deu a crise dos países asiáticos) não entrou em vigor, tendo sido portanto excluída. Anteriormente a 02/01/1998, a prática do Banco Central era divulgar metas para a TBC em expressão mensal; no sentido de se manter a consistência dos valores na amostra, eles foram anualizados (base 252 dias úteis) a partir do número de dias úteis em que vigoraram.

As surpresas com relação a essas decisões foram calculadas a partir dos PUs de fechamento dos Contratos Futuros de Taxa Média de Depósitos Interfinanceiros de 1 dia com duração de um mês e de seus respectivos dias de saque, disponíveis no *website* da Bolsa de Mercadorias e Futuros de São Paulo (BM&F). As cotações para o Índice Bovespa foram extraídas do banco de dados Economática®, com as correções devidas.

A amostra assim constituída de componente esperada e não esperada para a mudança na meta da taxa básica de juro e das variações do Índice Bovespa possui, no entanto, algumas particularidades que devem ser mencionadas:

I. Há duas observações cujos dias de implantação da nova meta ocorrem nos dois primeiros dias úteis do mês: são as dos dias 05/04/1999 e 01/09/1999 (33ª e 38ª reuniões regulares). Estas são as duas situações que causam os conflitos mencionados em **(6)**, conforme referido no item 3, que podem ser convenientemente contornados pela aplicação de **(5)** a regime de capitalização contínua, ou seja:

$$I^u = \ln(1+\Delta i^u)\text{, onde } (1+\Delta i^u) = \left[(1+\bar{i}_a)\frac{PU_{d-1}}{PU_d}\right]^{252/D} \quad \textbf{(5)}$$

Podemos estimar $\bar{i}_a$ para estes dois casos a partir das taxas CDI – diferentemente de empregar os PUs, como fizemos nos demais casos, nestas duas situações não teremos todos os dados necessários em **(6)** referenciados ao mesmo mês; utilizar as taxas diárias CDI para estimar $\bar{i}_a$ é um recurso que possibilitará a inclusão das observações em questão sem introduzir viés que comprometa os resultados.

II. Houve dez mudanças na meta (com reuniões ou não) consideradas extraordinárias, dado este obtido a partir de consulta ao agendamento pelo COPOM das reuniões seguintes mencionadas nas atas de reunião, como segue na tabela 1 abaixo (as 114 restantes são consideradas regulares):

Tabela 1: Datas das reuniões ou mudanças não previstas em ata, com a variação na meta e o novo valor adotado

Data Reunião	Mudança SELIC	Nova Meta p/ SELIC
10/09/1998	0,00 %	19,00 %
24/03/1999	-3,00 %	42,00 %
05/04/1999	-2,50 %	39,50 %
28/04/1999	-2,00 %	32,00 %
07/05/1999	-2,50 %	29,50 %

Data Reunião	Mudança SELIC	Nova Meta p/ SELIC
12/05/1999	-2,50 %	27,00 %
08/06/1999	-1,50 %	22,00 %
28/03/2000	-0,50 %	18,50 %
07/07/2000	-0,50 %	17,00 %
14/10/2002	3,00 %	21,00 %

III. Em agosto de 1998 ocorreu a crise russa; neste período o Banco Central fazia uso da TBC (taxa básica de juro) e da TBAN (taxa de assistência para redesconto bancário) – a primeira, tida como piso para as taxas (freqüente nas operações de captação do BC) e a segunda, como teto (mais comum na doação de recursos aos mercados).

Desse modo, as taxas variavam de acordo com a instituição na "banda" assim definida, conforme importância, garantias e risco dos tomadores; quando o cenário econômico se deteriorou, o Banco Central implantou, entre outras medidas, o aumento da TBAN com a manutenção da TBC – embora a procura maior por estes recursos tenha se dado (dentro e fora das operações de mercado aberto) a níveis intermediários entre piso e teto, não raro as taxas se situaram mais próximas ao teto, função do ceticismo dos agentes quanto à reversão do quadro.

Iniciada em setembro de 1998, esta situação perdurou até novembro do mesmo ano, com a assinatura de um novo acordo do país com o Fundo Monetário Internacional (FMI) – assim, a tabela 2 traz as quatro observações da amostra sujeitas ao caso:

Tabela 2: Datas das reuniões ligadas à crise russa

Reunião	Data	TBC	TBAN	Taxa adotada
27º	02/09/1998	19,00 %	29,75 %	24,375 %
28º	10/09/1998	19,00 %	49,75 %	34,375 %
29º	07/10/1998	19,00 %	49,75 %	34,375 %
30º	11/11/1998	19,00 %	42,25 %	30,625 %

Portanto, estas quatro observações caracterizaram-se por taxas médias que estiveram acima da TBC e abaixo da TBAN, o que leva a presumir que a simples adoção da TBC (ou da TBAN), nestes casos, não representaria adequadamente a realidade do período. Assim, parece mais verossímil adotar um valor entre estes dois limites – a idéia mais imediata e intuitiva é tomar a média de ambos, como mostrou a tabela 2 – espelhando assim um pouco mais fidedignamente o ambiente econômico da época. Assim, para os casos listados, tomou-se a média entre TBC e TBAN como representativa da taxa básica de juro.

Por fim, todos os erros-padrão são calculados seguindo o procedimento sugerido por Newey e West (1987), para serem robustos à heterocedasticidade e à correlação serial da amostra, cujas presenças são indicadas pelas estatísticas de Durbin e Watson (DW) e também pelos demais testes expressos nas tabelas do item a seguir.

5. RESULTADOS

O primeiro passo é a detecção das observações que possam ser consideradas *outliers*, ou seja, aquelas cuja influência estatística nos resultados da regressão seja excessivamente maior que as demais; para tanto, retomando-se então **(10)** e aplicando-se a **(8)**, conforme descrito em 3.1, encontramos cinco observações que se enquadram ao critério, conforme expresso pela distribuição mostrada no gráfico 1.

Não coincidentemente, tais observações estão ligadas a circunstâncias relevantes da histórica econômica brasileira contemporânea; são elas:

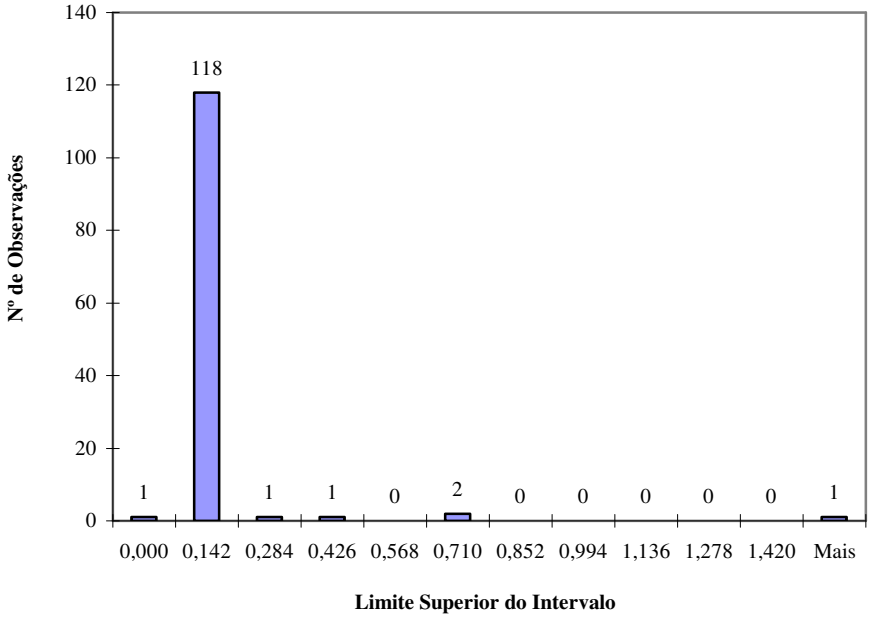
– **30/10/1997**: Nesta ocasião eclodiu a crise nos países asiáticos; em decisão extraordinária, o Banco Central retroagiu na decisão regular anterior (de dias antes) e praticamente dobrou a taxa básica de juro, além de outras medidas objetivando a reversão das expectativas então formadas;

– **02 e 10/09/1998**: em agosto de 1998 iniciou-se a crise russa (que teve como desfecho sua moratória), disseminando assim uma instabilidade internacional que obrigou o Banco Central a implantar um novo pacote de medidas, entre elas a elevação da TBAN. Em particular,

a mudança do dia 10 foi de caráter extraordinário, dada a celeridade dos acontecimentos da ocasião;

– **18/01/1999 e 04/03/1999**: como já mencionado em 3.2, em 15/01/1999 o Banco Central adotou o regime cambial flutuante. Na reunião seguinte, em 04/03/1999, uma forte subida da taxa básica de juro foi implementada, no intuito de se promover uma reversão das expectativas, dada a deterioração do cenário econômico desde o final de 1998.

Gráfico 1: Distribuição das influências estatísticas obtidas para cada uma das 124 observações, conforme equação **(10)** aplicada à regressão **(8)**.



Conhecidas então as observações de grande influência nos resultados, doravante os testes serão procedidos com dois conjuntos de dados: o primeiro, completo, com todas as observações pertinentes ao estudo em foco; o outro, com as mesmas características, porém tendo estas cinco observações previamente removidas. Isso permitirá (como já aludido) verificar-se a robustez das respostas que obtivermos. Esta sistemática será iniciada então com a apresentação da tabela 3.

Tabela 3: Principais estatísticas descritivas das variáveis selecionadas

	Com *Outliers*			**Sem *Outliers***		
	I^e	I^u	ΔP_t	I^e	I^u	ΔP_t
Média	-0,106%	-0,026%	-0,132%	-0,566%	0,013%	-0,207%
Mediana	-0,138%	0,011%	-0,159%	-0,166%	0,011%	-0,211%
Máximo	19,813%	1,488%	12,578%	4,708%	1,470%	4,766%
Mínimo	-6,578%	-4,373%	-9,006%	-6,578%	-1,734%	-5,430%
Desvio Padrão	3,230%	0,660%	2,382%	1,643%	0,405%	1,930%
Assimetria	3,830879	-3,146249	0,655413	-1,08452	-0,313406	-0,272712
Curtose	23,77875	21,76954	9,532345	6,665763	7,204113	2,970585
Nº Observações	124	124	124	119	119	119
Jarque-Bera	2534,038	2024,771	229,3473	89,95685	89,58449	1,479328
ADF	-1,2427 (-5,57)	-1,1158 (-4,72)	-1,3034 (-5,14)	-0,5039 (-3,89)	-0,9715 (-4,05)	-1,0934 (-4,62)
Phillips-Perron	-0,8942 (-9,81)	-1,1772 (-13,1)	-1,2572 (-14,3)	-0,7385 (-8,16)	-1,0474 (-11,24)	-1,0780 (-11,66)

Como se vê, nela estão explicitadas as principais estatísticas descritivas das variáveis explicativas de **(8)**, para ambos os conjuntos de dados (com e sem os *outliers*); conforme esperado, a remoção destes elementos reduz a dispersão dos dados e mitiga valores extremos praticamente nas três variáveis; nota-se ainda uma melhoria nos outros momentos (atenuando a

não-normalidade), bem como a estacionariedade das três séries geralmente a 1%, indicada pelos testes ADF e Phillips-Perron, a despeito da descontinuidade temporal que elas apresentam (fruto da construção para este estudo de evento).

Passando-se então à análise dos resultados das regressões **(8)** e **(9)**, apresentamos a tabela 4, onde se vê primeiramente a baixa significância das variações na meta da taxa básica de juro como variável explicativa para o Índice Bovespa. Nota-se também que o mesmo não se aplica ao teste para as suas componentes: não somente os coeficientes da surpresa são significativos a 1% e apresentam sinal negativo (conforme era inicialmente esperado e alinhado aos resultados de vários trabalhos para a realidade americana), como são robustos aos *outliers*; a componente esperada não possui a mesma característica, não sendo significativa na amostra sem as cinco observações.

Tabela 4: Respostas das variações (a uma taxa de capitalização contínua) do Índice Bovespa às componentes esperada e não esperada (surpresa) das mudanças na meta da taxa básica de juro ou a resposta do mesmo índice às mudanças na meta sem as componentes. Seguem, respectivamente, as especificações **(8)** e **(9)**, erros-padrão não robustos à heterocedasticidade ou correlação serial.

	Com *Outliers*	**Sem *Outliers***		**Com *Outliers***	**Sem *Outliers***
Intercepto	-0,0017 (-0,94)	-0,0017 (-0,91)	**Intercepto**	-0,0012 (-0.57)	-0,0020 (-1.05)
Mudança não esperada	-1,8766 (-6,83)	-1,3145 (-3,06)	**Resposta à variação da taxa básica de juro**	0,0830 (1.24)	0,0183 (0.17)
Mudança esperada	0,1065 (1,9)	0,0438 (0,41)			
Breusch-Godfrey Serial Correlation LM Test	6,6023 *(0,04)*	3,1897 *(0,2)*	**Breusch-Godfrey Serial Correlation LM Test**	10,1002 *(0.01)*	3,2736 *(0.19)*
White Heteroskedasticity Test	17,4896 *(0)*	2,7583 *(0,74)*	**White Heteroskedasticity Test**	9,9587 *(0.01)*	1,5305 *(0.47)*
Nº Observações	124	119	**Nº Observações**	124	119
DW	2,422	2,196	**DW**	2,544	2,139
R^2	0,313	0,081	**R^2**	0,012	0,000

Os testes referem-se à estatística ***nº obs x R^2*** e os respectivos valores entre parênteses referem-se às suas probabilidades para não rejeição da hipótese nula (não haver correlação serial e não haver heterocedasticidade, respectivamente). Os valores entre parênteses referentes aos coeficientes da regressão são as estatísticas ***t*** obtidas a partir de erros-padrão não robustos à heterocedasticidade e correlação serial.

Os testes corroboram, portanto, a hipótese da eficiência informacional desse mercado, explicando também os resultados não significativos para as variações da meta: se os mercados conseguem antecipar uma parte relevante destas variações, então de fato eles deveriam reagir fracamente às mesmas.

Constataram-se, portanto, respostas significativas somente às novas informações; não obstante, há que se ter alguma cautela: como se vê na mesma tabela, os testes para correlação serial e heterocedasticidade dos resíduos indicaram a presença de ambos os problemas – especialmente na amostra completa. Deste modo, relembrando a observação feita ao final do item 4, optou-se por empregar o procedimento sugerido por Newey e West (1987), cujos resultados são apresentados na tabela 5.

A aplicação do procedimento em questão garante erros-padrão robustos à correlação serial e heterocedasticidade dos resíduos, o que permite confirmar as conclusões inferidas a partir da tabela anterior; há que se observar, contudo, que embora a surpresa seja uma variável significativa nas variações do Índice Bovespa, o modelo linear simplificado ora proposto **(8)** tem baixo poder explanatório, especialmente na amostra que desconsidera os *outliers*, como demonstrado pelo respectivo coeficiente de determinação (R^2).

De modo geral, as causas para tal são, possivelmente, uma especificação inadequada (variáveis relevantes ausentes, por exemplo), ou uma não linearidade das variáveis empregadas, ou até mesmo uma não linearidade da própria regressão a ser usada. Não obstante, o interesse central do trabalho é estudar a significância das variáveis independentes para explicar a dependente, e não propriamente promover-se uma otimização ou excelência no ajustamento da equação regredida.

Tabela 5: Respostas das variações (a uma taxa de capitalização contínua) do Índice Bovespa às componentes esperada e não esperada (surpresa) das mudanças na meta da taxa básica de juro, de acordo com a especificação **(8)**.

	Amostra Completa	**Excluindo *outliers***
Intercepto	-0,0017 (-1,14)	-0,0017 (-1,07)
Mudança não esperada	-1,8766 (-4,19)	-1,3145 (-4,32)
Mudança esperada	0,1065 (1,54)	0,0438 (0,45)
Nº Observações	124	119
DW	2,42	2,20
R^2	0,313	0,081

Os valores entre parênteses referentes aos coeficientes da regressão são as estatísticas ***t*** obtidas a partir dos erros-padrão de Newey-West (1987) robustos à heterocedasticidade e correlação serial.

5.1. Quebras estruturais e assimetrias nas respostas

Conforme observado no item 3.2, a adoção da taxa básica de juro expressa ao ano-*over* (252 dias úteis) em 1998, a flexibilização cambial de 1999 e o conjunto de mudanças em caráter extraordinário (em contraposição às oriundas de reuniões regulares) introduziram modificações na dinâmica dos mercados; dessa maneira, o passo seguinte será verificar como tais "perturbações" afetam o objeto deste estudo - é o que mostra a tabela 6, onde as possíveis fontes de quebra estrutural supra citadas e as influências devidas ao contexto das decisões tomadas são testadas individualmente, segundo a especificação **(11)**.

Assim, os resultados mostram que as fontes de quebra ou as influências de contexto são significativas desde que consideradas as amostras completas; para estas últimas em particular, vê-se que as reuniões extraordinárias apresentam respostas mais intensas, enquanto que as respostas das ordinárias são mais brandas, o que é um resultado esperado, dado que mudanças decorrentes de reuniões não previstas deveriam gerar, tal como observado, reações mais pronunciadas.

No entanto, removidos os *outliers*, vê-se que não há robustez nas conclusões supra, exceção feita somente à adoção de taxas em expressão anual – esta *dummy* pareceu absorver a significância da surpresa propriamente dita. No mais, são resultados concordantes com o anteriormente visto: componente esperada com baixa significância e a surpresa, freqüentemente significativa a níveis de 1 a 5%.

Quanto aos resultados obtidos para a adoção de taxas em expressão anual, tendo em vista que se trata de uma variável *dummy* interativa que, aparentemente, deslocou a significância da surpresa propriamente dita para si e que são as 105 observações mais recentes e consecutivas da amostra que respondem por esta característica (das 124 da amostra completa), entendemos que o subconjunto por elas formado pode ser excessivamente representativo do todo – desse modo, a interação entre esta *dummy* e a surpresa poderia, na regressão, representar uma *proxy* da própria surpresa, o que explicaria tais resultados. Logo, no que se refere a este evento e dada esta ambigüidade, é necessário observar-se que, à luz do exposto, não se pode ainda ser conclusivo sobre este tema.

Tabela 6: Respostas das variações do Índice Bovespa (a uma taxa de capitalização contínua) às componentes esperada e não esperada (surpresa) das mudanças na meta da taxa básica de juro. Inclui a introdução de variáveis *dummy* interativas com a surpresa para captar a importância de: mudança cambial de 1999 (após 15/01/1999, *dummy*= 1 ou, caso contrário, zero); adoção de taxas anuais (após 02/01/1998, *dummy* =1 ou, caso contrário, zero) e mudanças regulares (se não são as relacionadas na tabela 1, *dummy*= 1; caso contrário, zero), todas de acordo com o exposto em 3.2 e a especificação **(11)**.

	Amostra Completa	**Excluindo *outliers***	**Amostra Completa**	**Excluindo *outliers***	**Amostra Completa**	**Excluindo *outliers***
Intercepto	-0,0015 (-1,01)	-0,0017 (-1,12)	-0,0024 (-1,62)	-0,0019 (-1,27)	-0,0022 (-1,46)	-0,0018 (-1,12)
Mudança não esperada	-2,2822 (-4,09)	-0,8412 (-1,7)	-2,5123 (-6,59)	-1,8112 (-2,29)	0,3862 (0,82)	0,0252 (0,05)
Mudança esperada	0,0979 (1,22)	0,0462 (0,48)	0,0708 (1,02)	0,0571 (0,62)	0,0410 (0,77)	0,0338 (0,34)
Mudança não esperada x *Dummy* Cambial	0,9096 (1,54)	-0,7168 (-0,96)				
Mudança não esperada x *Dummy* Reuniões Reg.			1,1301 (2,08)	0,7241 (0,72)		
Mudança não esperada x *Dummy* taxa anual					-2,4772 (-3,16)	-1,4466 (-2,47)
Nº Observações	124	119	124	119	124	119
DW	2,36	2,20	2,35	2,20	2,40	2,18
R²	0,329	0,086	0,335	0,086	0,335	0,087

Os valores entre parênteses referentes aos coeficientes da regressão são as estatísticas ***t*** obtidas a partir dos erros-padrão de Newey-West (1987) robustos à heterocedasticidade e correlação serial.

Tabela 7: Respostas das variações do Índice Bovespa (a uma taxa de capitalização contínua) às componentes esperada e não esperada (surpresa) das mudanças na meta da taxa básica de juro. Inclui a introdução de variáveis *dummy* interativas com a surpresa para verificação de possíveis assimetrias (*dummy*= 1 se o movimento ocorre na direção selecionada ou, caso contrário, zero), de acordo com a especificação **(11)**.

	Amostra Completa	**Excluindo *outliers***	**Amostra Completa**	**Excluindo *outliers***	**Amostra Completa**	**Excluindo *outliers***	**Amostra Completa**	**Excluindo *outliers***
Intercepto	-0,0015 (-1,02)	-0,0017 (-1,11)	-0,0013 (-0,85)	-0,0016 (-1,07)	-0,0016 (-1,02)	-0,0014 (-0,9)	-0,0015 (-1,1)	-0,0018 (-1,16)
Mudança não esperada	-1,2674 (-6,16)	-1,3531 (-4,84)	-2,5301 (-5,53)	-1,5731 (-1,79)	-1,8612 (-3,95)	-1,2145 (-3,73)	-1,6156 (-4,66)	-1,1045 (-3,18)
Mudança esperada	0,0954 (1,13)	0,0444 (0,45)	0,0947 (1,13)	0,0441 (0,46)	0,1065 (1,53)	0,0476 (0,49)	0,1333 (1,98)	0,0643 (0,68)
Mudança não esperada x *Dummy* aumento da taxa	-1,2668 (-2,64)	0,2796 (0,21)						
Mudança não esperada x *Dummy* redução da taxa			1,3354 (3,08)	0,3304 (0,36)				
Mudança não esperada x *Dummy* inatividade					-0,5207 (-0,49)	-1,2589 (-1,46)		
Mudança não esperada x *Dummy* reversão							-3,2509 (-3,71)	-1,9443 (-1,42)
Nº Observações	124	119	124	119	124	119	124	119
DW	2,35	2,19	2,36	2,20	2,43	2,20	2,35	2,25
R²	0,344	0,082	0,347	0,082	0,314	0,086	0,368	0,096

Os valores entre parênteses referentes aos coeficientes da regressão são as estatísticas ***t*** obtidas a partir dos erros-padrão de Newey-West (1987) robustos à heterocedasticidade e correlação serial.

No que se refere à hipótese de que os mercados possam reagir diferentemente conforme a direção das mudanças na taxa básica de juro, como se vê na tabela 7, novamente se confirmam resultados previamente conhecidos a respeito da componente esperada e também da surpresa (estes últimos robustos); verificou-se também significância para aumentos, reduções e reversões na taxa desde que considerados na amostra completa – removidos os *outliers*, a significância sucumbe; portanto, dada a falta de robustez dos resultados (indicações contraditórias pelas duas amostras), não se pode afirmar que os mercados respondam assimetricamente às ações de

política monetária, ou seja, que as intensidades das respostas sejam significativamente ligadas às direções das decisões tomadas pelo COPOM para mudanças na taxa básica de juro.

6. CONCLUSÕES

Tendo como base conceitual a hipótese dos mercados eficientes, foi definida uma medida que captasse a surpresa gerada nos mercados a cada decisão do COPOM. Estas medidas são eminentemente baseadas em dados do mercado de futuros DI 1 dia, cuja freqüência, liquidez e similaridade à taxa SELIC fazem desses instrumentos não somente ferramentas úteis, mas também versáteis para explorações sob diferentes horizontes e necessidades.

De posse desse conjunto de informações assim construído – medidas calculadas, mais as variações do Índice Bovespa, formou-se o banco de dados que respaldou o estudo de evento aqui realizado. Assim, como importantes resultados, constatou-se que o mercado acionário reage fracamente às variações diretas na meta da taxa básica de juro – uma conseqüência da capacidade do mesmo em se antecipar às decisões do COPOM; entretanto, verificou-se uma resposta significativa e robusta às variações inesperadas (tal como previsto pela hipótese da eficiência informacional dos mercados), estimando-se que a cada ponto percentual de incremento **não esperado** na meta da taxa básica esteja associada uma queda média de 1,3% do Índice Bovespa.

Uma descoberta adicional foi a de que relevantes eventos contemporâneos de ruptura econômica pareceram não interferir de modo robusto neste padrão de respostas às surpresas, o mesmo se observando com relação à direção tomada pelas mudanças; a exceção a ser citada refere-se à adoção de taxas de juros em expressão anual – como mencionado no item 5, os dados e métodos aqui empregados não permitiram uma posição conclusiva sobre o tema, sendo necessários estudos complementares que ajudem a elucidar a questão.

É importante ressalvar, contudo, que os resultados obtidos originaram-se de um estudo de evento cujo caráter é essencialmente empírico. Assim, é necessário desenvolverem-se modelos nos quais se possa não somente analisar o papel destas variáveis na formação dos preços, como também investigar outras implicações e desdobramentos teóricos; outro fato a se destacar é que estas conclusões repousam em uma diretriz inicialmente assumida: a janela de evento selecionada é de um dia útil – conforme sugerido por Campbell *et al.* (1996), o estreitamento desta janela, desde que conceitualmente factível, é recomendável sempre e pode ajudar a preservar a necessária ortogonalidade entre os resíduos das regressões e outras variáveis, envolvidas ou não no estudo.

Deve, também, ser observado que este estudo tem como premissa implícita que os mercados reagem às decisões do COPOM, e não o oposto (assume-se que política monetária é exógena em relação aos mercados); não obstante tal assunção, ao tratarmos da questão da ortogonalidade supra referida (bem como em 3.1), não se pode descartar a possibilidade de haver eventuais ocorrências de simultaneidade nas respostas de ambos os agentes (governo e mercados) a certos eventos ou tendências verificadas na economia.

Embora o critério adotado quanto à janela do estudo de evento ajude a minimizar sobremaneira tal problema, evidentemente houve situações em que as reações foram conjuntas – não coincidentemente, esta parece ser a característica dos cenários de *stress* anteriormente descritos que deram origem às cinco observações classificadas como *outliers* – note-se como esta constatação ratifica a importância da duplicidade nos testes conduzidos, com e sem tais observações, originalmente visando-se tão somente robustez nos resultados.

De qualquer modo, Bernanke e Kuttner (2005) observaram que esse problema, assim como outras características do estudo de evento, levam a inferências qualitativamente corretas quanto às respostas, a despeito de se ter detectado um viés que tende a subestimá-las em suas intensidades. Assim, estes problemas não invalidam as conclusões obtidas ao longo deste trabalho – embora se deva alertar que os efeitos então estudados podem ser ainda mais pronunciados que os ora estimados.

Por conta disto, recomendam-se algumas sugestões para se lidar com a questão: Gürkaynak, Sack e Swanson (2004), por exemplo, empregaram dados *intraday* como forma de mitigar ainda

mais o problema (embora isto constitua um obstáculo de difícil transposição na realidade brasileira, como já mencionado anteriormente); é possível tentar a abordagem de Rigobon e Sack (2004), a qual se vale de estimadores que exploram a heterocedasticidade introduzida por ações exógenas de política monetária; há ainda a possibilidade explorada por Poole, Rashe e Thornton (2002) dos erros de medidas nos regressores e a proposição de uma especificação modificada para corrigir o problema.

Por fim, entendemos que os próximos passos a serem seguidos consistirão na proposição de modelos que ajudem a expandir os horizontes deste assunto, bem como novos testes econométricos segundo as diferentes abordagens sugeridas (ou outras não tratadas aqui) que podem se constituir em campo fértil para novas contribuições a se somarem ao *mainstream* já estabelecido.

7. REFERÊNCIAS BIBLIOGRÁFICAS

BANCO CENTRAL DO BRASIL. **Atas das Reuniões do COPOM**; disponível em www.bcb.gov.br .

BERNANKE, B.; KUTTNER, K. (2005) What Explains the Stock Market's Reaction to Federal Reserve Policy?, **Journal of Finance**, 60, Nº 3, p. 1221-1257.

BOMFIM, A. N.; REINHART, V. R. (2000), Making News: Financial Market Effects of Federal Reserve Disclosure Practices. **Finance and Economics Discussion, Series Working Paper #2000-14**, Federal Reserve Board of Governors.

BOMFIM, A. N. (2003), Pre-Announcement Effects, News Effects, and Volatility: Monetary Policy and the Stock Market. **Journal of Banking and Finance**, 27, p.133-151.

CAMPBELL, J. Y.; LO, A. W.; MACKINLAY, A.C. (1996), **The Econometrics of Financial Markets**, Cap. 4 - Princeton University, 1ª Ed.

COOK, T.; HAHN, T. (1989), The Effect of Changes in the Federal Funds Rate Target on Market Interest Rates in the 1970s. **Journal of Monetary Economics**, 24, p. 331-352.

CRAINE, R.; MARTIN, V. (2003), Monetary Policy Shocks and Security Market Responses. **Working Paper**, University of California at Berkeley.

DEWBRE, J. H. (1981), Interrelationships Between Spot and Futures Markets: Some Implications of Rational Expectations, **American Journal of Agricultural Economics**, v. 63, p. 925-933.

GARCIA, M. G. P.; LA ROCQUE, E. C. (1994), O Mercado Futuro de Taxas de Juros no Brasil: Especificidades Teóricas e Empíricas do Mercado de DI Futuro, **Texto para Discussão N° 318**, Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia.

GONÇALVES, C. E. S.; GUIMARÃES, B. (2005), Monetary Policy and the Exchange Rate in Brazil, **Working Paper**, disponível em http://personal.lse.ac.uk/guimarae.

GÜRKAYNAK, R. S.; SACK, B. P.; SWANSON, E. T. (2004), The Effects of Monetary Policy on Asset Prices: An Intraday Event-Study Analysis. **Working Paper**, Board of Governors of the Federal Reserve System.

GÜRKAYNAK, R. S. (2005), Using Federal Funds Futures Contracts for Monetary Policy Analysis. **Working Paper**, Board of Governors of the Federal Reserve System (U.S.) Washington, D.C.: Divisions of Research & Statistics and Monetary Affairs.

HONDA, Y.; KUROKI, Y. (2006), Financial and Capital Markets' Responses to Changes in the Central Bank's Target Interest Rate: The Case of Japan. **The Economic Journal**, nº 116, p. 812–842.

KUTTNER, K. N. (2001), Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market, **Journal of Monetary Economics,** 47, p. 523-544.

LUCAS Jr., R. E.; SARGENT, T. (1996), After Keynesian Macroeconomics. In: Miller, P. J. (org.) **The Rational Expectations Revolutions:** Readings from the Front Line. [S.1]: Massachussetts Institute of Technology, p. 05-30.

NEWEY, W.; WEST, K. (1987), A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. **Econometrica**, 55, p.703-708.

PEROBELLI, F. F. C.; PEROBELLI, F. S.; ARBEX, M. A. (2000), Expectativas Racionais e Eficiência Informacional: Análise do Mercado Acionário Brasileiro no Período 1997-1999, **RAC**, v. 4, n. 2.

POOLE, W.; RASCHE, R.; THORNTON, D. (2002), Market Anticipations of Monetary Policy Actions. **Federal Reserve Bank of St. Louis Review**, 84, p. 65–93.

REINHART, V.; SIMIN, T. (1997), The Market Reaction to Federal Reserve Policy Action from 1989 to 1992. **Journal of Economics and Business**, v. 49, n. 2, p. 149-168.

RIGOBON, R.; SACK, B. (2004), The Impact of Monetary Policy on Asset Prices. **Finance and Economics Discussion Series 2002-4,** Board of Governors of the Federal Reserve System.

ROBITAILLE, P.; ROUSH, J. (2005), How do Brazilian Yield Spreads and Stock Prices Respond to FOMC Actions and U.S. Macroeconomic Data Announcements? **Working Paper**, Federal Reserve Bank.

ROLEY, V. V.; SELLON Jr., GORDON H. (1995), Monetary Policy Actions and Long-Term Interest Rates. Federal Reserve Bank of Kansas City, **Economic Review**, 80 p .73-89.

ROLEY, V.; SELLON JR, G. (1998), The Response of Interest Rates to Anticipated and Unanticipated Monetary Policy Actions. **Working Paper**, University of Washington.

ROLEY, V.; SELLON JR, G. (1998), Market Reaction to Monetary Policy Nonannouncements. **Working Paper**, Federal Reserve Bank of Kansas City.

SMIRLOCK, M.; YAWITZ, J. (1985), Asset Returns, Discount Rate Changes, and Market Efficiency. **Journal of Finance,** 40, p.1141-1158.

TABAK, B. M.; TABATA, A. (2004), Surpresas na Política Monetária e a Estrutura a Termo da Taxa de Juro Brasileira. **Economia Aplicada**, v.8, nº 3, p. 383-399.

TABAK, B. M.; TABATA, A. (2004), Testando o Conteúdo Informacional das Decisões de Política Monetária. **Quarto Encontro Brasileiro de Finanças**.

THORBECKE, W.; ALAMI, T. (1994), The Effect of Changes in the Federal Funds Rate Target on Stock Prices in the 1970s. **Journal of Economics and Business**, 46, p.13-19.

THORBECKE, W. (1997), On Stock Market Returns and Monetary Policy. **Journal of Finance,** 52, nº 2, p.635-654.

DIVIDENDS: INSTITUTIONAL FACTORS AND CATERING INCENTIVES

MARIA ELISABETE DUARTE NEVES
E-mail: mneves@iscac.pt
Instituto Superior de Contabilidade e Administração de Coimbra

Resumo

Este trabalho construído de acordo com a teoria de catering dos dividendos, examina como os diferentes ambientes institucionais influenciam o efeito catering.
O argumento principal da análise é que quando as empresas pertencem a diferentes ambientes institucionais e a natureza dos problemas de agência também difere, haverá diferenças na relação entre política de dividendos e efeito catering.
Os nossos resultados mostram que empresas em países da Zona Euro, EUA, REINO UNIDO, Canadá e Japão atendem aos sentimentos dos seus investidores.
Mais interessantes, existe uma interacção entre catering e factores institucionais; particularmente, a protecção legal aos investidores, desenvolvimento de mercados de capitais e orientação dos sistemas financeiros, efectividade do mercado de controlo corporativo, nível de concentração de propriedade e efectividade dos conselhos de administração. Nós corroboramos um efeito substituto dos mecanismos externos de corporate governance no efeito catering. Especificamente, pagadores de dividendos com sistemas de corporate governance fraco pagam mais dividendos.

Palabras clave: Ambientes Institucionais, Dividendos e catering.
Área temática: Economia Monetária e Financeira.

Abstract

This paper is built upon the predictions of the catering theory of dividends, and examines how the different institutional environments impact on catering effect.
The focus of our analysis is the argument that when companies belong to different institutional environments and the nature of existing agency problems also differs, there will also be differences in the relationship between dividend policy and the catering effect.
To achieve this aim, we propose a dividend model that incorporates a variable at a firm-level proxing for the catering effect. The results from the estimation of the model by using the Generalized Method of Moments provide interesting results. Consistent with the predictions of the catering theory, we find that companies in Eurozone countries and US, UK, Canada and Japan cater to their investors' sentiments.
More interesting, our findings show an interaction effect between catering and institutional factors; particularly, the legal protection of investors, development of capital markets and the orientation of the financial systems, the effectiveness of the market for corporate control, the level of ownership concentration and the effectiveness of boards of directors. We find a substitute effect of external corporate governance mechanisms on catering dividends. Specifically, dividend payers with weak governance are significantly more likely to pay dividends than dividend payers with strong governance.

Keywords: Institutional variables, international legal environment, catering incentives, investors' sentiment, payout policy
Thematic Area: Monetary and Financial Economy.

1. INTRODUCTION

One of the most important financial decisions that a firm's managers make is the amount and stability of dividends. That is why the dividend policy has long been an issue of interest in the financial literature and, in spite of the vast research on the topic; it is still an open matter. In fact, dividends have always been a bit of a puzzle in the theory of the firm. According to Baker, Powell, and Veit (2002, p. 255), "despite a voluminous amount of research, we still do not have all the answers to the dividend puzzle".

Since the Miller and Modigliani (1961) irrelevance proposition, according to which dividend policies are all equivalent and there is no a particular policy that can increase shareholder wealth in perfect capital markets, many scholars have offered alternative explanations for dividends in imperfect markets. Within this strand of literature, they are many works that investigate the important role played by a firm's characteristics in shaping its dividend decision; one of the most remarkable is the one by Fama and French (2001). The question as to why companies pay out dividends has given rise to different explanations, the most relevant ones stemming from the agency theory.

The overinvestment problem that is likely to appear in firms with free cash flow is one of the offered explanations for dividends. In this context, the literature shows that more profitable firms pay more dividends[1], while firms with more investment opportunities pay less[2], which is consistent with the propositions of Easterbrook (1984) and Jensen (1986) about the role of dividends in controlling the agency costs of free cash flow. Specifically, Jensen's (1986) free cash flow theory states that higher free cash flow should lead to higher dividend payments in order to prevent firms from overinvesting. Hence there is a positive relationship between a firm's free cash flow and its payout ratio[3] .

Agency problems call not only for higher dividends but for more debt as well. In fact, financial literature documents the role played by debt and dividends as agency-cost control mechanisms (see Ross, 1977, Grossman and Hart, 1980, Jensen, 1986, and Harris and Raviv, 1991 for debt; Lintner, 1956, Bhattacharya, 1979, Rozeff, 1982 and Jensen, 1986 for dividends). This suggests that debt and dividend may be somehow related, although there is not consensus about the way they are related. On the one hand, the substitution hypothesis predicts a negative relationship between debt and dividends based on the minimization of agency conflicts without duplicating efforts (see Rozeff, 1982; Jensen, 1986; Crutchley and Hansen, 1989; Jensen, Solberg, and Zorn (1992), and Chen and Steiner, 1999). On the other hand, the alternative hypothesis points to the complementary use of these two mechanisms and, consequently, to a positive relationship between them (see Jensen, 1989; Eckbo and Verma, 1994; and Zwiebel, 1996).

Lintner (1956) points to another important explanation of dividends. He argues that firms seek to maintain the stability of dividends, and he finds that a firm's earnings is probably the key factor to account for in order to get a stable dividend pattern. Consequently, a positive relationship between a firm's earnings and its dividend payments will exist[4].

The nature of a firm's assets has also been documented as a determinant of dividends (see, for instance, Allen and Michaely, 2003; Aivazian, Booth, and Cleary, 2003). According to Scott (1977), firms with a high proportion of tangible assets are more leveraged, which will in turn affect positively or negative dividend payments depending on whether there is a substitution or a complementary relationship between debt and dividends.

Size has also been traditionally considered among the determinants of dividend policy, and previous evidence seems to agree that larger firms pay higher dividends (see, for instance, Fama and French, 2001, and Denis and Osobov, 2005, 2008).

[1] See, for example, Lintner (1956), Jensen, Solberg, and Zorn (1992), Fama and French (2001), DeAngelo, DeAngelo, and Skinner (2004), and Li and Lie (2006).

[2] See, for instance, Wang, Erickson, and Gau (1993), or Barclay, Smith, and Watts (1995).

[3] The positive relationship between dividends and free cash flow is confirmed by, for instance, Chaplinsky and Niehaus (1993), and Holder, Langrehr, and Hexter (1998). More recently, Gaspar, Massa and Matos (2005), and DeAngelo, DeAngelo, and Stulz (2006) document the role played by dividends in controlling for overinvestment processes in firms with high levels of free cash flow.

[4] More evidence on earnings being a determinant of dividends can be found in, for instance Nissim and Ziv (2001), DeAngelo, DeAngelo, and Skinner (2004), Koch and Sun (2004), Brav, et al., (2005), Denis and Osobov (2008), and Skinner (2008).

More recently, there is the argument that the characteristics of the firms paying dividends (that is, their levels of free cash flow, leverage, earnings, tangible fixed assets and size) should not be separately analyzed from certain psychological components in that an important part of the decision to pay dividends may be due to a firm's desire of satisfying investors' expectations. In fact, in agreement with the recent trends in the theory of the financial behaviour, time-varying catering incentives also appear to shed light on the "disappearance" of dividends by Fama and French (2001)[5]. This new explanation of dividends has its origin in the catering theory proposed by Baker and Wurgler (2004a), according to which the changes in the amount that companies pay on dividends can be explained by what they denominate "catering incentives", that is, a measure of the market desire for dividend-paying stocks. According to this new theory, when investors' demand for payouts increases, firms are more likely to increase payouts (via either dividends or repurchases[6]). Recent studies (see, among others, Brown and Cliff, 2004, 2005; Lai, 2004; Denis and Osobov, 2005; Fairchild and Zhang, 2005; Gemmil, 2005; Hsieh and Wang, 2006; Cohen and Yagil, 2008; Han, 2008; Hoberg and Prabhala, 2008) show the growing interest that is lending to this new theory and they suggest that investors' sentiments can be decisive in the resolution of the dividend puzzle.

There is also literature, despite much scarcer, in what refers to the importance of the different institutional factors for a firm's payout. For example, La Porta, et al., (2000a) obtain empirical evidence that confirms that the corporate dividend policy varies according to the legal system, in such a way that stronger legal protection of investors leads to higher dividend payouts.

This evidence is confirmed by Shleifer and Wolfenzon (2002), who suggest that better-protected investors expect more of their firms' profits to come back to them, instead of being expropriated by the entrepreneur who controls the firm.

There are other institutional factors besides investor protection that have been found to be linked to firms' dividend decision. For instance, in Zwiebel (1996), managers voluntarily pay dividends in order to avert challenges for control, and Pan (2007) shows an association between the propensity to pay dividends and measures of managerial entrenchment based on indices of anti-takeover charter provisions[7]. Aivazian, Booth, and Cleary (2003) find that in countries where the stock market is not so important in financing firms, changes in dividend payouts are more usual, and they documents that firms in emerging markets have more unstable dividend payments than their US counterparts. The financial literature proposes two different ownership concentration systems, revealing a much more concentrated ownership for Continental European countries than for the Anglo-Saxon ones. The results in Schooley and Barney (1994), and Farinha (2003) point out that there is a relationship between dividends and ownership concentration that may change along with the increase of ownership concentration, leading to a non-linear relationship between them. Faccio, Lang, and Young (2001) find evidence on the structure of ownership and control influencing dividend policies in European and East Asian firms, and show that firms that exhibit a wider discrepancy between ownership and control pay higher dividends.

Gugler and Yurtoglu (2003) claim that dividend payouts decrease with the control stake of the largest shareholder, whereas the size of the second largest shareholder is positively related to dividend payouts. Finally, Brennan and Thakor (1990), and Allen, Bernardo, and Welch (2000) find that higher dividends lead to a higher fraction of shares held by institutional investors, whereas Grinstein and Michaely (2005) find that larger institutional ownership does not lead to higher payouts in that institutional investors, in contrast to individual investors, tend to prefer a payout policy based on stock repurchase rather than on dividend payments[8].

[5] For a recent research on disappearing dividends, see for instance, DeAngelo, DeAngelo, and Skinner (2004); Amihud and Li (2006), Denis and Osobov (2008), or Hoberg and Prabhala (2008).

[6] In our work we didn't include repurchases for the difficulty of some markets in this type of transactions that would turn the sample different from our purpose.

[7] See, for example, Claessens, et al., (2002); Farinha (2003); Hu and Kumar (2004); John and Knyazeva (2006), among others, for the relationship between managerial entrenchment and payout policy.

[8] For a review between institutional ownership and corporate dividend policy see, for instance, Han, Lee, and Suk (1999); Short, Zhang, and Keasey (2002), or Ravi (2007).

All this research reveals that institutional characteristics, such as investor protection, development of capital markets, contestability of the market for corporate control, the level of ownership concentration and the effectiveness of boards of directors as well as corporate governance system, are relevant for a firm's dividend decision. This paper tries to provide broader empirical evidence by examining whether these institutional factors shape the implications of the key dividend theories by determining managers' incentives to behave according to their predictions. Note that in respect to the catering theory of the dividends, Baker and Wurgler (2004b) claim that institutional factors cannot explain the decision of dividends through this theory because such factors explain the existence of the clientele effect. More recently, Hsieh and Wang (2006) examine empirically the possible impact of the different legal environment and the investors' preferences on the dividend decision. Their results, however, do not support the catering theory of dividends, although they show a new road in the construction of the dividend puzzle.

To analyze all this implications, the focus of our analysis is to study in depth how the institutional context affects firms' dividend policy. This analysis lies on the premise that the institutional disparity, and the subsequent differences in the nature of existing agency problems, lead to differences on the relationship between dividends and one of their key and more innovative determinants; that is the catering incentives.

In fact, given the growing interest in the new catering theory of dividends, our main purpose in this paper is to investigate how different institutional factors affect the evaluation that investors makes of the firms and of who decides the allotment of dividends, what would lead to differences in the catering effect of dividends in different institutional contexts. Hence we offer a study of the impact of several institutional factors on the investors' sentiment that support the catering theory of dividends. As far as we know our work differs from existing literature in that it tries to answer to several unanswered questions about the dividends policy from the perspective of catering incentives around the world. There is no prior evidence supporting this view, and providing empirical support to this issue is thus one of the major contributions of this piece of this work. The literature more thoroughly known on these subjects it portrays the relationship between dividends and institutional variables and even so almost always for countries usually belonging to common law.

In this context, the aim of our study is to explain how the different institutional factors in different governance systems[9] affect dividends decisions according to a firm's desire of satisfying investor's sentiments. Despite the lack of previous evidence, there are strong arguments that lead us to argue that investors' preference for dividend-paying stocks change according to the above mentioned institutional variables.

To learn which of these institutional variables are more probable of influencing the firm's dividends policy, we examined the payout of the following countries: United States, United Kingdom, Canada, Japan and the Eurozone countries, which represent a great variety of institutional environments.

For continue this approach we use Generalized Method of Moments, and consistent with catering arguments, our findings reveal that investors' sentiments significantly affect a firm's propensity to pay dividends and, as expected, this effect is positive. That is, investors' demand for dividend-paying stocks encourages firms to increase their payouts. Additionally, the analysis performed in this paper, stressing the role played by institutional factors, gives rise to the following conclusions: First, we find that firms with higher levels of ownership concentration and the effectiveness of boards of directors cater to a large extent to their investor's sentiments. This result suggests that the internal mechanisms exert certain pressure on firms to cater to investors demand and that translate into higher dividends payout.

In second place, our results show that, the stronger the legal protection of investors the smaller the extent to which cater to their investor's sentiments and that translate into lower dividends payout what confirms the substitute model of La Porta, et al., (2000a). Third, our study provides further evidence on the moderating role of the development of capital markets

[9] Considering Corporate governance systems as a result of several institutional and legal factors, as: the legal protection of investors, the development of capital markets, the role of the market for corporate control, the level of ownership concentration and the effectiveness of boards.

and the orientation of the financial systems on the relation between dividends and investors' sentiments and specifically, we find that the investors' demand for dividends tends to be "ignored" of firms in market-oriented financial systems. Followed, we investigate the interaction between the catering effect and the contestability of market for corporate control and the results are very similar to the obtained for the previous external corporate governance mechanisms, that is, the more active it is the market for corporate control, the smaller the extent to which firms cater to their investors sentiments.

Our results suggest a substitute effect of external corporate governance mechanisms on catering dividends. Specifically, dividend payers with weak governance are significantly more likely to pay dividends than dividend payers with strong governance.

This way we can say that this results point out that in our sample of fifteen countries, the external mechanisms of corporate governance support the substitute model in opposition to the outcome model. Therefore, we find that firms with weak corporate governance mechanisms[10] are more likely to cater investor's preferences for dividends and pay higher dividends while the relation between catering dividends and governance is stronger for firms with high quality of internal corporate governance mechanisms.

The remainder of this paper is organized as follows: First, we describe the main legal and institutional factors characterizing the corporate governance systems and summarize previous literature and empirical evidence on this matter which leads us to pose our hypotheses. Section 3 describes the data and our model of dividends and discusses the estimation method. The results are discussed in section 4 and, finally, the concluding remarks are presented in Section 5.

2. INSTITUTIONAL FEATURES: PREVIOUS EVIDENCE AND HYPOTHESIS

In what follows, we describe the key institutional factors that may influence a firm's dividend policy and review previous evidence on the matter in order to pose our hypothesis about the role played by the institutional context in moderating the implications of the main dividend theories.

2.1. The legal protection of investors

One of the most widely accepted explanations for the different patterns of corporate finance across countries is based on the role played by laws in protecting investors. The new institutional economics that has come to be called the *Law and Finance* approach (see La Porta, et al., 1998) assumes that the quality of law across countries depends on their legal origin. In this way, two legal traditions are identified: civil law and common law. On the one hand, civil law has its origin in Roman law, and is prevalent in most Continental European countries and in Japan. On the other hand, common law is English in origin and includes the UK, the US and Canada, among others. According to the results in La Porta, et al., (1997, 1998, 2000b, 2002), common law countries protect investors better than those with civil law. This piece of evidence has given rise to an extensive literature on the efficiency of laws in protecting investors (both shareholders and creditors) and on their enforcement across countries (see, for instance, Shleifer and Vishny, 1997; La Porta, et al., 1997, 1998; Johnson and Shleifer, 1999; Demirgüç-Kunt and Levine, 2001).

There are several previous studies revealing that the discussion of the dividend policy cannot be separated from legal features. La Porta, et al., (2000b) provide evidence that the stronger the protection of minority shareholder the higher the dividend payouts. This evidence is consistent with the so called outcome agency model of dividends, according to which firms operating in countries where shareholders protection is weak pay lower dividends because of the higher agency problems between managers and shareholders, whereas in countries where shareholders are more protected more dividends are paid because shareholders are enabled to

[10] Weak corporate governance is understood as a mechanism with smaller protection of investor, bank-oriented systems and weak efficiency of the market of corporate control.

force managers to disgorge cash. This result is corroborated by Shleifer and Wolfenzon (2002). However, the dividend policy can also be seen as a substitute for the legal protection of investors. According to the substitute model, insiders interested in issuing equity in the future pay dividends in order to establish a reputation for a decent treatment of minority shareholders. Supporting this model, Faccio, Lang, and Young (2001) find that in countries with weak investor protection the allotment of dividends is higher, as a way to limit the expropriation of minority shareholders. Dittmar, Mahrt-Smith, and Servaes (2003) also find significantly higher cash disgorgements in countries where shareholders have little legal protection.

In short, the influence of laws on dividends is a matter of fact. However, previous evidence on the role played by the legal protection of investors in shaping a firm's dividend policy is somewhat mixed, or even perhaps confusing. For the whole exposed literature, we know that payout ratio is systematically related to the degree of shareholders' legal power. To shed light on this matter, we analyse the differences in payout ratios across different legal contexts to learn whether the outcome or the substitute model applies. On the basis that the different legal features of a country will shape managers' incentives to accommodate payout ratios to the firm characteristics and the investors' sentiments, we pose our first hypothesis:

Hypothesis 1: The degree of investor's protection will influence the extent which firms cater to their investor's sentiments.

To test this hypothesis, we have constructed several indices. The first one, Legal Origin index, classifies the countries under analysis according to their legal origin, and it takes value 1 if the country is as a common law country and 0 if it is a civil law one. Within this law-driven approach, additional indices have been proposed to assess the effectiveness and quality of enforcement of laws across countries. The second, Anti-director Rights, measures how strongly the legal system favours minority shareholders over managers or dominant shareholders[11]. Like other previous papers, such as Demirgüç-Kunt and Maksimovic (1999, 2002); Beck, et al., (2001) and Leuz, Nanda, and Wysocki (2003), we follow La Porta, et al., (1997, 1998) in the construction of this index, which results from adding up the scores of six indices referring to the protection of minority shareholders[12]. The third index, Creditor Rights, is obtained follow Pindado and Rodrigues (2004), who provide a deeper analysis of the insolvency law than La Porta, et al., (1998) and also correct some of their indices[13]. The fourth index, proxy the degree of enforcement of a country's laws (see La Porta, et al., 1998; Beck and Levine, 2002; Beck, Demirguç-Kunt, and Levine, 2003; Giannetti, 2003, and Leuz, Nanda, and Wysocki, 2003) constructed through the average of Law and Order and Efficiency of Judicial System. The last two indices, protection of investor and effective protection of investor were constructed by us using the previous ones to reinforce the underlying idea[14].

2.2. Capital Markets

The development of capital markets and the orientation of the financial systems are the two main features of capital markets that have been broadly used to establish institutional differences across countries. Rajan and Zingales (1995) were the first in establishing the dichotomy between bank-oriented and market-oriented financial systems and, since then, a lot

[11] For example, Ball, Robin, and Wu (2003), and Hope (2003) find that the presence of strong anti-director rights provides an effective deterrent against the manipulation of financial reports because managers would be aware investors might sue them for losses.

[12] Proxy by mail allowed, shares not blocked before meeting, cumulative voting or proportional representation, oppressed minorities mechanism, pre-emptive rights, and percentage to call an extraordinary shareholders' meeting. Note, however, that the score of Germany in the Proxy by Mail Allowed index and the score of the US in the Cumulative Voting or Proportional Representation index reported by La Porta, et al., (1998) have been corrected following Miguel, Pindado, and de la Torre (2005). First, although German shareholders are not allowed to vote by mail, most of them follow this practice, when necessary; through their bank (see, Vagts, 2002). Second, cumulative voting is not mandatory in Delaware corporate law, and it is rarely observed by American firms (Roe, 2002).

[13] Specifically, the score of the United States in the Absolute Priority and Reorganization with Creditors' Consent indices, and the score of Spain in the Absolute Priority index, have been corrected according to their respective insolvency laws.

[14] The first one is measured for the average of the Anti-director Rights index with Creditor rights index; the effective protection of investor is measure taken by averaging the indices of investor protection and enforcement.

of empirical studies on the matter has been developed (see, for instance, Beck and Levine, 2002; Demirgüç-Kunt and Maksimovic, 2002, and Levine, 2002). Market-oriented systems, such as the ones of the US, the UK, and Canada, are characterized by well-capitalized stock markets. In contrast, the banking sector is of great importance in financing firms, and financial markets are usually small in bank-oriented countries, such as Continental European countries and Japan. The influence of the development of capital markets on the dividend decision has received scarce attention in the literature. On the one hand, according to Demirgüç-Kunt and Maksimovic (1996), in countries with developed stock markets there is a substitution of equity for debt financing. This reliance on equity financing makes managers more concerned with aligning interests with those of shareholders in order to maintain a good reputation in the capital market. Dividend payments may be a solution to the conflict of interests between managers and shareholders (Rozeff, 1982; Easterbrook, 1984), hence firms operating in developed capital markets are expected to pay higher dividends as a way to keep shareholders satisfied and, consequently, have their market value increased. Asymmetric information problems are also critical for firms in developed stock markets. According to signalling theories, higher information asymmetry between managers and shareholders leads firms to be more willing to incur in costs associated with dividend signalling, in that dividends convey management's confidence about the future profitability of a firm and, consequently, about dividend stability (e.g., Bhattacharya, 1979; Talmor, 1981; Miller and Rock, 1985; Bar-Yosef and Huffman, 1986; Nissim and Ziv, 2001). All this suggests that dividend distributions enable firms to obtain equity financing in favourable terms (John and Williams, 1985), and thus higher dividends are expected in market-oriented systems. On the other hand, agency and asymmetric information problems in a bank-oriented system are likely to be resolved by internal channels between the firm and its creditors[15]. In this context, managers' bigger concern is to be on good terms with the firm's creditors to be able to obtain future financing. Given that dividends cause a transfer of wealth from bondholders to shareholders, debt contracts usually restrict the distribution of dividends in the firm, thus downgrading the importance of the dividend decision. Except in La Porta, et al., (1998)[16], previous evidence supports this argument. For instance, Aivazian, Booth, and Cleary (2003) consider dividends as a substitute for the direct communication between the firm and its external investors, and find that institutional structures prevalent in developing countries make corporate dividend policy a less viable mechanism for signalling future earnings and reducing agency costs than in countries with developed capital markets, such as US. Dewenter and Warther (1998) make a similar point by arguing that stable dividend payments may not be as important for Japanese that are part of a *keiretsu*, due to the close ties between managers and investors in this kind of bank-based structures. Similar results are found by Correia da Silva, Goergen, and Renneboog (2005) for the German case. They document a smaller signalling role of dividends in Germany, where firms with banks as their major shareholders are more willing to omit or cut their dividends than firms controlled by other types of shareholders.

In short, dividend payments are expected to be higher in firms operating in market-oriented systems than in firms operating in bank-oriented systems. To go further in the effect of capital markets' development on a firm's dividend policy, we pose our second hypothesis:

Hypothesis 2: Market-oriented systems encourage managers to cater to a larger extent to investors' dividend demand.

Following Beck, et al., (2001), and Demirgüç-Kunt and Maksimovic (2002), we have constructed the Market index that takes value 1 if the country is classified as a market-oriented system and 0 if it is considered a bank-oriented system. We provide four additional indices of capital market development (see, Beck and Levine, 2002). The first one, Stock Market

[15] For example, Roe (1994) finds that in a bank-oriented system, such as the one in Japan and Germany, banks own controlling stakes and play a central role in monitoring and supplying information. Also, Gorton and Schmid (2000) find evidence supporting the notion that banks are an important part of the corporate governance mechanism in Germany.

[16] La Porta, et al., (1998) document that the lack of transparency, inadequate legal infrastructure, and weak investor protection in emerging markets all enhance the role of dividends as a reputation mechanism, and make the payment of dividends a necessary mechanism to attract capital.

Capitalization relative to GDP, captures the importance of stock markets in the financial system. The second index, Total Value Traded to GDP, is a measure of the capital market's liquidity. Note that common law countries are characterized by higher market capitalization and liquidity than those with civil law. The last two indices, index of market development and index of banking development were constructed by us using the previous ones to reinforce the underlying idea[17].

In sum, at the macro level, King and Levine (1993), Levine and Zervos (1998), and Beck, Levine, and Loayza (2000) show that financial development promotes growth and that differences in legal systems could explain the most differences in financial development.

2.3. Contestability of the market for corporate control

The concept of the market for corporate control as a control mechanism is originally due to Marris (1963) and Manne (1965), and subsequently, the financial theory has traditionally held the assumption that the takeover market plays an important role in disciplining management by aligning the interests of owners and managers[18]. First, even the widespread threat of a takeover places the management under greater discipline by institutionalizing an evaluation mechanism of corporate decision-making. Second, when the threat of an acquisition is not enough to guarantee managers' efficiency in the construction of value, the threat is carried out and management in charge is substituted. Consistent with previous arguments, the financial literature supports favourably on the effectiveness of the market for corporate control market in resolving shareholder-manager conflicts (see, among others, Jensen and Ruback, 1983; Brickley, Lease, and Smith, 1988; Jarrell, Brickley, and Netter, 1988; 1994; Franks and Mayer, 1996).

Berglof (1990), and Franks and Mayer (1997) find that one of the biggest institutional differences among G-7 countries is the contestability of the market for corporate control[19].

On the one hand, market-based systems are generally characterized by highly active markets for corporate control. Specifically, a market for corporate control is usually associated with the US and the UK (see, Jensen and Ruback 1983; Jarrell, Brickley, and Netter, 1988, for the US; Franks and Mayer, 1997 for the UK)[20], where firms' stock rights are highly decentralised and shareholders have limited influence over companies' operations and management. Something different occurs in the case of Canada. Attig and Gadhoum (2003) report that more than 80% of Canadian listed firms have controlling shareholders, which make hostile takeovers more difficult to succeed. On the other hand, an active market for corporate control is virtually absent in bank-based systems. As Renneboog (2000, pg 2) argues, whereas in Anglo-Saxon countries managerial performance is maintained by the complementary intervention of both internal and external control mechanisms, the disciplinary function of the takeover market in Belgium, and in most other Continental European countries, is limited. For instance, in countries such as Japan and Germany, the market for corporate control is by no means dynamic since stock rights are highly concentrated. Actually in Germany the market for corporate control has not played a significant role in the post-war period (Höpner and Jackson, 2001), and only three hostile takeovers occurred during the years 1945 and 1995 (Franks and Mayer, 1998). In Japan, the influence of banks and the strength of cross-shareholdings typical of *keiretsus*, represent the main structural barriers to takeovers[21] [22]. In Japan the ownership concentration is plenty high

[17] The index of market development is measured by the average of the market capitalization to GDP with Total Value Traded to GDP; and index of banking development is the average of the ratio of the sum of bank liquid liabilities, bank assets and deposit bank domestic relative to GDP.

[18] See, Becht, Bolton, and Roell (2003) for a comprehensive review of the conventional corporate governance mechanisms include market for corporate control.

[19] Berglof and Burkart (2003, pp. 173) affirm that, " while increased contestability of control is desirable hostile takeovers are a rather blunt instrument for regulation and the market for corporate control is only one of many corporate governance mechanisms to be honed in order to promote corporate restructuring in Europe".

[20] In fact, there is evidence on a substantial number of takeovers in the US and the UK, particularly during the nineties (see, for instance, Conn and Connell, 1990; Hopt, et al., 2000; and Goergen and Renneboog, 2004), and although most of the takeovers in these countries are non-aggressive bids, the fraction of unfriendly bids is not negligible (approximately 47% for the US, see Cottner, Shivdasani, and Zenner, 1997; and 25% for the UK, see Franks and Mayer, 1996).

[21] Holmstom and Kaplan (2001), Kole (2002), and Grullon, Kanatas, and Weston (2004), and Kini, Kracaw, and Mian (2004) provide more detailed discussions on the market for corporate control.

and the banks the main shareholders being also common the big participations inter-companies (Prowse, 1992; Berglof and Perotti, 1994). The market for corporate control is quite inactive in this country. In Spain, a weak market for corporate control is dominant. Traditionally, the market for corporate control among Spanish firms was practically non-existent because of high ownership concentration of quoted firms and poor minority shareholder rights. Perotti and Von Thadden (2006) argue that a society with more diffused financial wealth should exhibit developed equity markets, strong minority protection, and a market for corporate control. He show for a sample of 13 OECD countries that in 1970, stock market capitalization as a percentage of GDP was highest in Britain, Canada, and the U.S., and lowest in Austria, France, Germany, and Italy, closely followed by Belgium. Your results also show that with information for 13 OECD countries that given the available information, it seems that Japan, Austria, and Belgium are close to Germany and France, while Canada are closer to the US and the UK[23].

The influence of the market for corporate control contestability on firms' dividend policy is scarcely documented in the literature. Zwiebel (1996) proposes a model in which managers, voluntarily pay dividends when they are under a constant takeover threat. Pan (2007) shows that firms choose a combination of governance provisions and dividend policy to maximize value, and that dividend payment that reduces a firm's cash holdings can be used to deter hostile takeovers. De Jong (1997) shows empirically, using the 100 larger European companies between 1991 and 1994 that firms in countries with active markets for corporate control tend to pay higher dividends, while companies in countries where hostile takeovers do not exist pay a higher part of net value added in wages.

This previous evidence suggest higher dividends when there is an active market for corporate control, which is consistent with the role played by this market as an external control mechanism, capable of bringing down agency costs and mitigating conflicts between shareholders and management. Our third hypothesis relies on this assumption and predicts the following:

Hypothesis 3: In countries with active markets for corporate control, firms will better fit their dividend payouts to their investors' catering incentives payout ratio.

We have constructed the Corporate Control variable, which accounts for the role played by this market in corporate governance, in order to test this hypothesis. This index takes value 1 in countries where an active market for corporate control exists and 0 otherwise. Note that the coincidence between this index and the Market variable defined in the previous section is complete, with the exception of Ireland. In fact, the usefulness of the market for corporate control is based on the premise that stock prices reflect managerial inefficiencies, thus creating the threat of a takeover.

2.4. The level of ownership concentration

With the increased separation of ownership from control, managers frequently face very little supervision. In this context, a commitment to a high dividend policy attenuates managerial opportunism and forces the firm to frequently intersect with the capital markets.

The separation of ownership and control in the modern corporation has given rise to the well-known principal-agent problem, which is the basis of corporate ownership being a key governance feature. In fact, financial literature proposes ownership structure as one of the main corporate governance mechanisms, especially helpful in solving the conflicts of interests

[22] Morck and Nakamura (1999) point out that the postwar US occupation imposed a dispersed ownership structure on Japanese firms. In response to the hostile takeovers and greenmail payments that took place in the 1960s and 1970s, a rapid growth of intercorporate equity holdings occurred. For evidence on the takeover market in Japan, see Kester (1991), and kaplan and Minton (1994).

[23] For a recent literature, Bebchuk, Cohen, and Allen (2005), and Cremers and Nair (2005) examine one important dimension of corporate governance, namely, the market for corporate control. They show a negative relation between various indices of antitakeover provisions and both firm value and long-run stock return performance. However there are authors as, for instance, Masulis, Wang, and Xie (2007), that question the exact channels of antitakeover provisions that negatively affect shareholder value.

between owners and managers, and in minimizing the associated agency costs. Actually, the interesting point for corporate governance is that in an environment of highly dispersed ownership, the individual shareholder has little or no incentive to monitor management. Hence a concentrated ownership is considered one of the key mechanisms of corporate governance in that larger stakes provide shareholders with enough capability and incentives to undertake monitoring activities (Jensen and Meckling, 1976; Shleifer and Vishny, 1986).

There are significant differences in corporate ownership patterns across countries. These broad differences are documented through different perspectives. For instance, La Porta, et al., (1998) argues that the extent of investors' legal protection is one of the most important determinants of the choice between concentrated and dispersed ownership of corporate shares. In common law countries investor protection is reinforced by stronger law enforcement, whereas in countries with weaker investor rights higher ownership concentration is needed. Therefore, a relatively high ownership concentration in many developed and developing economies may be an equilibrium response to a low level of protection of minority shareholders[24]. The results in La Porta, et al., (1998) support this argument and show that the stronger the legal protection of shareholders rights the lower the ownership concentration**.** The concentration of corporate ownership is also closely related to questions surrounding stock market participation that have attracted considerable interest in the recent literature for example Ferreira, Ornelas, and John (2005) studying the impact of the ownership structure of a corporation on the characteristics and efficiency of the market for corporate control. In this paper those authors explain the operation of the market for corporate control and how the structure of ownership affects efficiency. Perotti and Von Thadden (2006) shows that over the long-term politics drives the development of the equity market. Kaustia and Torstila (2008) show a causal link from increased stock ownership to increased voting for right-of-centre parties, and it is possible that governments use this to manipulate the electorate to their own advantage. In fact, the rate of share ownership has also been subject to a large literature on stock market participation and another example is that by Vissing- Jorgensen (2002) that with your work shows that the limited participation may be a cause of the equity premium puzzle and finds that (nonfinancial) income positively affects the likelihood of participation.[25] As a result, the question as to what facilitates and propels ownership diffusion remains a central research question in corporate finance and dividends policy[26]. Additionally, Pagano and Lombardo (1999), Pagano and Volpin (2001), and Roe (2006), argue that political determinants primarily explain differences in ownership concentration. However, it is clear that significant differences exist in ownership structures within the European Union (Barca and Becht, 2002; Faccio and Lang, 2002, and Gugler and Yurtoglu, 2003). Moreover, those last authors find that the role of financial institutions is scarcely relevant in Spain, France and Italy, because in those countries has the families that usually control most of the firms. In fact, Faccio and Lang (2002), showed that in countries based in Continental Europe (Civil Law tradition) insider ownership is mostly associated with large shareholders who control, through many varied mechanisms such as corporate networks or family links.

The financial literature traditionally distinguishes between two types of ownership concentration systems (see, for instance, Moerland, 1995; Kaplan, 1997; Gedajlovic and Shapiro, 1998; Mayer and Sussman, 2001). On the one hand, the so-called Anglo-Saxon legal system, prevalent in the US, the UK, and Canada among others, is characterized by dispersed shareholdings and a high level of institutional ownership[27]. On the other hand, the Continental European model dominant not only in continental European countries but also in Japan, is characterized by concentrated ownership, which usually belongs to families and banks. For example, Becht and Mayer (2002) report that in more than 50% of European companies there is

[24] Recent theoretical and empirical studies relating ownership and payout include, among others, Eckbo and Verma (1994); Lucas and McDonald (1998); Allen, Bernardo, and Welch (2000); Fenn and Liang (2001); Short, Zhang, and Keasey (2002); Denis and McConell (2003); Farinha (2003); Gugler (2003); Gugler and Yurtoglu (2003), Perez-Gonzalez (2003), Brav, et al., (2005), Grinstein and Michaely (2005), Mancinelli, and Ozkan (2006), or Baker, et al., (2007), among others.

[25] For example, Kothare (1997), and Pham, Kalev, and Steen (2003) find that a less concentrated ownership structure contributes to market liquidity.

[26] See, for instance, Gugler (2003) on the different impact of government-controlled or shareholder- controlled ownership structures on dividend policies.

[27] For instance, Morck, et al., (2004) review the evolution of corporate ownership in Canada during the twentieth century.

a single voting block of shareholders that commands a majority of shares. In contrast, in the UK and US it is less than 3%[28].

Previous research has paid attention to the role of corporate ownership in shaping the dividends decision. For instance, Faccio, Lang, and Young (2001) examine the structure of ownership and control, and they find evidence of systematic expropriation of outside shareholders' interests by controlling owners in European and East Asian firms. They then try to find how this phenomenon is related to firms' dividend behaviour. What they find is that a second large shareholder mitigates agency conflicts in European firms by increasing dividend payments, whereas multiple controlling shareholders intensify the conflicts of interest in East Asian firms, because they tend to collude in expropriating minority shareholders by paying lower dividends. Gugler and Yurtoglu (2003) claim that dividend payouts decrease with the control stake of the largest shareholder, whereas the size of the second largest shareholder is positively related to dividend payouts. More recently, Khan (2006) obtains results consistent with dividends being a substitute for poor monitoring by a firm's shareholders[29].

We focus on the differences in ownership concentration levels across countries, and expect higher dividends in firms with more concentrated patterns, on the basis of ownership concentration being a monitoring device of managerial discretion. This monitoring effect leads us to pose our fourth hypothesis:

Hypothesis 4: The higher a firm's ownership concentration, the better the fit of dividends to investors catering incentives.

In our study we have constructed, following La Porta, et al., (1998), an index measuring Ownership Concentration[30]. This index reveals a high level of ownership dispersion in common law countries, whereas ownership is much more concentrated in those with civil law, except for Japan.

2.5. The effectiveness of boards of directors

The function of the board of directors in the corporate governance is to protect shareholders' interest and discipline management. If the board succeeds in carrying out its implementation and ratification roles, it will ensure that shareholder interests are safeguarded. However, if the functioning of this internal control mechanism is weak or inadequate, shareholder interests will become of secondary importance and managers' discretionary activity will increase. Thus, if boards of directors fail, shareholders suffer because of the combined effects of costly discretionary behaviour, poor financial performance and a falling stock market valuation. In short, this internal control mechanism represents an alternative way of restricting potential conflicts of interests between managers and shareholders. There seems to be no disagreement on the need for monitoring and control by boards of directors; however, their effectiveness differs considerably across countries which allow us to establish institutional differences according to two key features: the composition of the board and its internal structure. In fact, several academics studies shows that board composition varies with firms characteristics and the institutional environment (for example, Brickley and James, 1987; Hermalin and Weisbach, 1988, and Denis and Denis, 1994), and suggest a homogeneous composition, therefore that may be optimal for some firms but not for others[31]. The composition of the board is critical to its efficacy in that the more independent the board members, the greater its effectiveness in monitoring management. In this context, a considerable body of the ongoing debate in the US

[28] These differences between Anglo-Saxon and Continental European countries are corroborated by Franks and Mayer (1997). Faccio and Lang (2002) reveal the same results; moreover, they found that the role of financial institutions is scarcely relevant in Spain, France and Italy, where families usually control most of the firms. See too, Becht and Roel (1999), Mayer and Sussman (2001), Volpin (2002), and Aganin and Volpin, (2003).

[29] In the same vein, Fenn and Liang (2001) report that firms with low managerial stock-option holdings have significantly higher dividend and total payout ratios (including repurchases). This result could be due to the lack of "dividend protection" afforded by most executive stock option contracts (Lambert, Lanen, and Larker, 1989).

[30] Some recent papers by Carlin and Mayer (2003), and Leuz, Nanda, and Wysocki (2003) also used the same index.

[31] See, Bhagat and Black (1999, 2002); Denis and Sarin (1999); Hermalin and Weisbach (2003), and Raheja (2005) for an extensive review about board structure and function.

and the UK deals with the optimal composition of the board of directors[32]. Actually, Anglo-Saxon boards have been generally considered a competent control mechanism because of their independence of management, since the designation of independent or non-executive directors constitutes a widespread practice in these countries[33]. In contrast, the role of boards of directors in most Continental European countries may be questioned given the lack of clear regulation[34].

In fact, given the diversity of boards structure among several countries in our paper, we used the term board in broad sense, in other words, as an internal mechanism of corporate governance with either management, monitoring or supervisory functions. We should enhance that the obvious difficulty of classification of the "boards" in the different countries is arrested with the fact of those not always adjust their characteristics to legally systems or guidelines and principles accepted and globally recognized[35].

Where shareholdings are highly concentrated, as occurs in Continental European countries and Japan, non-executive directors may be considered as a mechanism to control majority shareholders, but without them being able to actively take part in the firm's decision-making process. This is why boards are rarely composed of independent directors in these countries. Although boards in most European counties are evolving towards an effective governance mechanism thanks to the various Codes of Best Practice[36] they are still far from this concept because the greater presence of controlling shareholders there makes it difficult to comply with these Codes' voluntary requirements.

Besides composition, the internal structure of the board is also fundamental for the effectiveness of this mechanism in corporate governance. In this sense, the existence of a one-tier or a two-tier board structure plays a key role in guiding and supervising a company. Firms in Anglo-Saxon countries (specifically the US, the UK and Canada (Hopt and Leyens, 2004; Adams and Ferreira (2007), and Dargenidou, Mcleay, and Raonic, 2007) and in most European countries except Germany, the Netherlands (Renneboog, Franks, and Mayer, 2001, and Chirinko, et al., 2004), and Denmark have adopted the unitary board structure, which implies that all board members are considered to be in the same position since they manage the company and also supervise its activity. There is thus no distinction between managing and supervisory functions. In contrast, the two tier structure is characterized by the existence of two bodies, an executive and a supervisory board, which guarantees that the last is separated from and independent of management. In fact, although in general both the unitary board of directors and the supervisory body in the two-tier structure are elected by shareholders, and that there is usually a supervisory function and a managerial function under both structures, the distinction between these two functions is indeed more formalised in the two-tier structure.

There is not previous empirical evidence, as far as we know, on the structure and composition of boards being determinants of a firm's dividend policy. However, given that there is no disagreement on the key role played by this mechanism in protecting shareholders from managers' abuses; higher dividends are likely to be found in firms with independent and two-tier structures in that they are assumed to better monitor managers in shareholders' interests. This argument leads to pose the following hypothesis:

[32] See, for instance, for a recent research on corporate boards, Dahya, Dimitrov and McConnell (2008), although in another perspective, of value. They find that firm value is positively correlated with the fraction of directors unaffiliated with dominant shareholders especially in countries with weak legal protection for minority shareholders.

[33] More details can be found in Prowse (1995), and Morck (2004).

[34] In a recent and interesting study, Adams and Ferreira (2007) show a theory of friendly boards from cross countries variation on board structure. They argue that shareholders should be allowed to choose between board structures (sole board system or dual board system) and their model illustrates that shareholders are always at least weakly better off if the board has an advisory role.

[35] The diversity of board structures among IOSCO members and the OECD Principles' recognition that there are potentially many different, and in that measured the studies on this matter should cover non-executive board members of companies with unitary boards; members of supervisory (i.e. non-executive) boards of companies with dual board structures; and members of the board of auditors elected by shareholders (which exist, for example, in Italy, Japan and Portugal): See, IOSCO (2002) and OECD (2004).

[36] As documented by Ho Chi-Kun (2005, pg 212), internationally, many corporate governance guidelines and codes of best practices have been published by supra-national agencies being some pointed examples for this author, the following ones : the Commonwealth Guidelines (CACG, 1999); the OECD (1999) Principles; the World Bank Framework for Implementation, 1999), national regulatory bodies (e.g. the French Vienot Commission's Recommendations (AFG-ASFFI, 1998 and amended 2001); the German Code (GPCG, 2000); the Japanese Principles (CGFJ, 1997 and revised 2001); the UK Combined Code of Best Practice (CCGUK, 2000)), and non-regulatory institutions (e.g. the CalPERS (1997 and revised 1999) Principles and Guidelines; the TIAACREFF (2000) policy statement; the European Shareholders Association's Guidelines (ESA, 2000); the International Corporate Governance Network's Statement (ICGN, 1999); the Business Roundtable (1997) Statement). The website of European Corporate Governance Institute (http://www.ecgi.org) provides a full list of these international guidelines, among others.

Hypothesis 5: Independent boards and two-tier boards will lead managers to better fit their dividends to investors catering incentives.

To consider the role played by boards of directors in our analysis, we have constructed the Board variable, a score of 1 is assigned to a country with a predominant two-tier board structure or when non-executive directors represent a significant proportion on boards[37], and 0 otherwise. The control over the behaviour of managers will be more effective when there is a clearer distinction between the 'supervisor' and those being 'supervised'.

2.6. Corporate Governance

Research on corporate governance has identified a number of mechanisms intended to insure that management acts in the best interests of shareholders, and these can includes external and internal corporate governance mechanisms.

In fact, the relationship between corporate governance and dividend policy has recently been emphasized by La Porta, et al., (2000a); Gugler and Yurtoglu (2003), Holmstom and Kaplan (2003); Correia da Silva, Goergen, and Renneboog (2004), among others. La Porta, et al., (2000a) explain for 33 countries the economic basis for testing for a relationship between dividends and quality of governance, highlighting two considerations with opposite implications for the sign of this relationship. On one side, these authors describe an outcome model that leads to the prediction of a positive relationship between dividends and the quality of governance. They interpret their evidence of higher dividends in well-governed firms as a result of effective pressure by minority shareholders on insiders to release cash. On the other hand and in opposition of the outcome model, the substitute view expected a negative relationship that is; weak governance increases the need to payout cash as dividends in order to overcome agency problems[38]. Gugler and Yurtoglu (2003) find large negative effects of announced dividend changes in German companies where corporate insiders have more power. Correia da Silva, Goergen, and Renneboog (2004) find a U-shaped relationship such that dividends first decrease and then increase with the voting share of the largest owner. In fact, most theorical and empirical corporate governance studies use U.S data. We can see an exception in Denis and McConnell (2003) who provide a most comprehensive international literature review on corporate governance, and the lack of cross-country European studies is quit evident. Faccio and Lang (2002) is another exception, as they examine ownership structure throughout Europe.

More recently, John and Knyazeva (2006) show evidence that dividends and total payouts (i.e. the sum of dividends and repurchases) are significantly higher when internal and external governance measures indicate weak governance. So we understood the type of monitoring mechanism is relevant for predicting dividends payouts. Those authors find that firms with weak corporate governance on average pay higher dividends. Managers faced with a high takeover threat (external monitoring) are more likely to repurchase and tend to repurchase more on average. On the other hand, strong internal governance (board) allows more accurate following of managerial actions and is associated with fewer cash distributions. They also find that the incidence of payouts is significantly lower among firms with better governance because these firms face lower agency costs of free cash flows and concludes that in the presence of high agency costs, governance plays a greater role in determining corporate payout.

In the same line of reasoning, Tse (2004) questions the logic of the relation between agency costs and dividend policy in the outcomes model – if well-governed firms are more likely to pay dividends as La Porta, et al., (2000a), then shareholders shouldn't need to rely on the payment of dividends to reduce the agency costs of free-cash-flow because such costs should already be low for well-governed firms.

Other authors as, for instance, Gompers, Ishii, and Metrick (2003); Cremers and Nair (2005), and Yermack (2006), show that companies with weak corporate governance have lower security returns than companies with strong corporate governance. The results of Giannetti and Simonov (2006), also find that investors expect lower returns from companies with weak

[37] Both rules prevent those supervised from being supervisors, and thus lead to independent boards of directors. In fact, an outside-dominated board can be considered as coming close to the two-tier board regime (Wymeersch, 1998; and Hopt, et. al., 2000).

[38] See, for instance, the traditional models of Rozeff's (1982), and Jensen's (1986) with managers versus shareholder agency conflict.

corporate governance. The several arguments given above by the literature described, they take us to pose our last hypothesis:

Hypothesis 6: Governance characteristics will moderate the extent to which firms cater to their investors sentiments.

To test the hypotheses posed we construct a joined index of corporate governance that is measure taken by averaging the indices of ownership by three largest shareholders, corporate control, independent and two-tier index.

3. DATA, EMPIRICAL MODEL AND ESTIMATION METHOD

3.1. Data

To test the hypotheses posed in the previous section, we use data from several Eurozone countries, United States, United Kingdom, Canada, and Japan which represent a great variety of institutional environments.

We have thus used an international database, Worldscope, as our principal source of information. Additionally, international data such as the growth of capital goods prices, the rate of interest of short term debt, and the rate of interest of long term debt, have been extracted from the Main Economic Indicators published by the Organization for Economic Cooperation and Development (OECD).

Since our study is intended to present a wide variety of institutional environments we selected fifteen representative countries and for each country we construct an unbalanced panel of non-financial companies from 1990 to 2003.

Three of the fifteen countries[39] have been excluded from our analysis for different reasons. As occurs in La Porta, et al., (2000b), Luxembourg has been removed from our sample because there are just a few firms listed in Luxembourg's stock exchange, and Greece because dividends are mandatory in this country. Finally, Finland had also to be excluded because no sample with the abovementioned requirement could be selected. The structure of the samples by number of companies and number of observations per country is provided in Table I. As shown in Table II, the resultant unbalanced panel comprises 3000 companies and 20,395 observations.

Table 1-Structure of the samples by countries

Country	Number of companies	Percentage of companies	Number of observations	Percentage of observations
Germany	427	14.23	4,263	20.90
France	391	13.03	3,812	18.70
Netherlands	137	4.57	1,412	6.92
Spain	99	3.3	1,046	5.13
Belgium	83	2.77	841	4.12
Portugal	43	1.43	366	1.80
Ireland	43	1.43	438	2.15
Austria	57	1.9	561	2.75
Italy	135	4.5	1,316	6.45
US	535	17.83	2,140	10.49
UK	560	18.68	2,240	10.98
Canada	79	2.63	316	1.55
Japan	411	13.7	1,644	8.06
Total	3,000	100.00	20,395	100.00

Data of companies in Eurozone countries and US, UK, Canada and Japan were extracted. The resultant samples comprise 427 companies (4,263 observations) for Germany, 391 companies (3,812 observations)

[39] The eurozone currently comprises twelve countries: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxemburg, Netherlands, Portugal, and Spain.

for France, 137 companies (1,412 observations) for the Netherlands, 99 companies (1,046 observations) for Spain, 83 companies (841 observations) for Belgium, 43 companies (366 observations) for Portugal, 43 companies (438 observations) for Ireland, 57 companies (561 observations) for Austria, 135 companies (1,316 observations) for Italy, 535 companies (2,140 observations) for US, 560 companies (2,240 observations) for UK, 79 companies (316 observations) for Canada and 411 companies (1,644 observations) for Japan.

Table 2. Structure of the panel

No. of annual observations per company	Number of companies	Percentage of companies	Number of observations	Percentage of observations
14	327	10.90	4,578	22.44
13	99	3.30	1,287	6.31
12	99	3.30	1,188	5.83
11	93	3.10	1,023	5.02
10	119	3.97	1,190	5.83
9	135	4.50	1,215	5.95
8	159	5.30	1,272	6.24
7	129	4.30	903	4.43
6	124	4.13	744	3.65
5	131	4.40	655	3.21
4	1,585	52.80	6,340	31.09
Total	3,000	100.00	20,395	100.00

Data from firms for which information is available for at least five consecutive years between 1990 and 2003 were extracted. After removing first-year data, used only to construct several variables, the resultant unbalanced panel comprises 3000 companies (20,395 observations).Table III provides summary statistics (mean, standard deviation, minimum and maximum) of the variables used in our analysis.

Table 3. Summary statistics

Variable	Mean	Standard deviation	Minimum	Maximum
FCF_{it}	.04239	.11581	-1.1651	1.9621
D_{it}	.10094	.11518	.0000	.89555
NI_{it}	.02199	.07218	-.84731	.62594
$TANG_{it}$	.27837	.18929	.00008	.99679
SI_{it}	13.0143	1.9605	8.4024	20.3265
CAT_{it}	.0000	.74661	-6.0792	8.8978

The table provides summary statistics (mean, standard deviation, minimum, and maximum) of the variables used in our analysis. FCF_{it} , is the free cash flow, NI_{it} denotes net income, $TANG_{it}$ denotes tangible fixed assets, and SI_{it} is the size

3.2. Empirical model and estimation method

Using the dependent variable, payout ratio, obtained as explained in Neves, Pindado, and de la Torre (2006)[40], and the traditional explanatory variables mentioned in Section 1[41], as well as the catering variable obtained through the value model (explained in the same paper) our basic model is as follows:

$$CPR_{it} = \gamma_0 + \gamma_1 FCF_{it} + \gamma_2 D_{it} + \gamma_3 NI_{it} + \gamma_4 TANG_{it} + \gamma_5 SIZE_{it} + \gamma_6 CAT_{it} + \varepsilon_{it}. \qquad (1)$$

[40] Once the dependent variable is a censured variable in that some companies pay dividends whereas do not, the authors in section 3.2 predicted a Tobit model following Auerbach and Hasset (2003).
[41] For more details about measures used, see once more, Neves, Pindado, and de la Torre (2006) in Section 3.3.

Additionally, and in accordance with the aim of our study, we investigate whether or not several institutional characteristics moderate the catering effect, and for that we propose the following model to be estimated:

$$CPR_{it}=\gamma_0+\gamma_1 FCF_{it}+\gamma_2 D_{it}+\gamma_3 NI_{it}+\gamma_4 TANG_{it}+\gamma_5 SIZE_{it}+CAT_{it}(\gamma_6+\lambda DV_{it})+\varepsilon_{it} \quad (2)$$

where, DV_{it} is a dummy variable constructed according to the institutional characteristics of legal protection of investors; development of capital markets and/or market-oriented financial system and bank-oriented, active markets for corporate control; the level of ownership concentration; effectiveness of boards of directors, and corporate governance. It is worth noting that in all cases whenever the dummy variable equals one and both parameters (γ_6 and λ) are significant, a linear restriction test is needed in order to know whether their sum ($\gamma_6+\lambda$) is significantly different from zero. The null hypothesis to be tested in these cases is the hypothesis of no significance, H_0: $\gamma_6+\lambda=0$.

In fact, in our study we have constructed different indices in accordance with section 2 that we can see in Table IV.

Table 4. Institutional Factors

Country / Variables	Legal Protection of Investors Variables					Development of Capital Markets Variables				Ownership Variable	Corpor. Gover.
	Anti-director Rights	Creditor Rights	Enforcement	Protection Investor	Effective Protection Investor	Market Capitalizat/ GDP	Total Value Traded/GDP	Index of Market Development	Index of Banking Development	Ownership Conc.	Corpo. Gover. Index
Germany	2.00	3.00	7.32	2.50	6.10	0.24	0.28	0.48	0.94	0.48	0.33
France	3.00	0.00	6.89	1.50	3.45	0.33	0.17	0.34	0.85	0.34	0.28
Netherlands	2.00	2.00	8.00	2.00	5.33	0.69	0.43	0.61	0.95	0.39	0.80
Spain	4.00	1.00	5.78	2.50	4.81	0.30	0.23	0.40	0.80	0.51	0.34
Belgium	0.00	2.00	7.85	1.00	2.62	0.36	0.05	0.23	0.81	0.54	0.35
Portugal	3.00	1.00	5.99	2.00	3.99	0.13	0.05	0.17	0.68	0.52	0.17
Austria	2.00	3.00	7.85	2.50	6.54	0.12	0.08	0.25	1.03	0.58	0.53
Italy	1.00	2.00	6.19	1.50	3.10	0.17	0.08	0.22	0.64	0.58	0.36
Japan	4.00	3.00	7.49	3.50	8.74	0.79	0.28	0.55	1.46	0.18	0.23
Ireland	4.00	1.00	6.53	2.50	5.44	0.26	0.14	0.30	0.39	0.39	0.63
U.K.	5.00	5.00	7.29	5.00	12.14	1.13	0.55	0.82	1.09	0.19	0.56
U.S.	4.00	1.00	8.00	2.50	6.67	0.80	0.62	0.75	0.66	0.20	0.57
Canada	5.00	1.00	7.78	3.00	7.78	0.59	0.29	0.49	0.66	0.40	0.63
Sample average	3.00	1.92	7.15	2,46	5.90	0.45	0.25	0.43	0.84	0.41	0.44

The resultant table of institutional factors comprises different index constructed in accordance with section 2 of the paper. Anti-director Rights, measures how strongly the legal system favours minority shareholders over managers or dominant shareholders; Creditor Rights, is obtained follow Pindado and Rodrigues (2004); the third index, proxy the degree of enforcement of a country's laws constructed through the average of Law and Order and Efficiency of Judicial System. The last two indices, protection of investor and effective protection of investor were constructed by us using the previous ones to reinforce the underlying idea. We provide four additional indices of capital market development (see Beck and Levine, 2002). The first one, Stock Market Capitalization relative to GDP, captures the importance of stock markets in the financial system. The second index, Total Value Traded to GDP, is a measure of the capital market's liquidity. The last two indices, index of market development and index of banking development were constructed by us using the previous ones to reinforce the underlying idea. Following La Porta, et al., (1998), we constructed an index measuring Ownership Concentration. Finally we constructed a corporate governance index.

All our models have been estimated by using the panel data methodology. Two issues have been considered to make this choice. First, unlike cross-sectional analysis, panel data allow us to control for individual heterogeneity, and to eliminate the risk of obtaining biased results because of such heterogeneity (Moulton, 1986, 1987). This point is crucial in our study because the dividend decision is very closely related to the specificity of each company. Specifically, we have controlled for heterogeneity by modelling it as an individual effect, η_i, which is then eliminated by taking first differences of the variables. Consequently, the error term in our models, ε_{it}, has been splitted into four components. First, the above mentioned individual or firm-specific effect, η_i. Second, d_t measures the time-specific effect by the corresponding time dummy variables, so that we can control for the effects of macroeconomic variables on the dividend decision. Third, since our models are estimated using data of several countries, we have also included country dummy variables (c_i). Finally, v_{it} is the random disturbance.

The second issue that we can deal with by using the panel data methodology is the endogeneity problem. The endogeneity problem is likely to arise in that the dependent variable (payout ratio) may also explain some of the explanatory variables. Finally, we have checked for the potential misspecification of the models. First, we use the m_2 statistic, developed by Arellano and Bond (1991), in order to test for lack of second-order serial correlation in the first-difference residual. Tables V, VI, VII, and VIII show that there is no a problem of second-order serial correlation in our models (see m_2). Note that although there is first-order serial correlation (see m_1), this is caused by the first-difference transformation of the model and, consequently, it does not represent a specification problem of the models. In second place, our results in Tables V, VI, VII, and VIII provide good results for the following three Wald tests: z_1 is a test of the joint significance of the reported coefficients; z_2 is a test of the joint significance of the time dummies; and z_3 is a test of the joint significance of the country dummies.

Table 5. Estimation results of the basic model

	I
Constant	-.36405* (.01820)
FCF_{it}	.12729* (.02933)
D_{it}	.00847 (.01817)
NI_{it}	.20477* (.05143)
$TANG_{it}$	.03098** (.01202)
S_{it}	.00398* (.00146)
CAT_{it}	.01436* (.00268)
z_1	63.08 (6)
z_2	669.39 (12)
z_3	10.88 (11)
m_1	-5.37
m_2	-0.65

The regressions are performed by using the panel described in Table II. The variables are defined in Table III. The rest of the information needed to read this table is: i) Heteroskedasticity consistent asymptotic standard error in parentheses. ii) *,** and *** indicate significance at the 1%, 5% and 10% level, respectively; iii) z_1, z_2 and z_3 are Wald tests of the joint significance of the reported coefficients, of the time dummies and of the country dummies, respectively, asymptotically distributed as χ^2 under the null of no significance, degrees of freedom in parentheses; iv) m_i is a serial correlation test of order i using residuals in first differences, asymptotically distributed as N(0,1) under the null of no serial correlation.

Table 6. Estimation results of the moderating role of the legal protection of investors

	I	II	III	IV	V	VI
Constant	-.3679* (.01495)	-.3599*(.01563)	-.3859* (.01575)	-.3608* (.01675)	-.3596* (.01615)	-.3557* (.01645)
FCF_{it}	.06793* (.01945)	.12132* (.02411)	.13536 * (.02740)	.13873* (.02679)	.10148* (.02534)	.12884* (.02588)
D_{it}	.03071* (.00987)	.02842*** (.01526)	.01258 (.01574)	.00541 (.01663)	.01507 (.01623)	.00769 (.01637)
NI_{it}	.37866* (.01736)	.29814* (.03813)	.19761 * (.04509)	.25902* (.04492)	.32389* (.04306)	.27232* (.04278)
$TANG_{it}$	.01139 (.00997)	.01225 (.00941)	-.00669 (.00976)	.01887*** (.01663)	.02429** (.01038)	.02441** (.01063)
S_{it}	.00393* (.00118)	.00437* (.00125)	.00320 * (.00126)	.00449* (.00134)	.00446* (.00128)	.00469* (.00131)
CAT_{it}	.01316* (.00234)	.01444 *(.00242)	.00856 ** (.00357)	.03954* (.00420)	.01605* (.00238)	.02157* (.00252)
$CAT_{it}DV_{it}$	-.00794* (.00269)	-.01916* (.00418)	.01219 * (.00426)	-.03863* (.00470)	-.00747** (.00379)	-.02078* (.00377)
t	3.66	-1.34	7.91	.378	2.61	.255
z_1	237.24 (7)	60.77 (7)	50.14 (7)	76.78 (7)	68.34 (7)	67.01 (7)
z_2	1001.58 (12)	776.19 (12)	794.46 (12)	736.00 (12)	748.03 (12)	771.93 (12)
z_3	14.05 (11)	17.15 (11)	10.83 (11)	9.20 (11)	11.23 (11)	8.69 (11)
m_1	-5.33	-5.35	-5.31	-5.36	-5.34	-5.35
m_2	-.66	-.65	-.60	-.58	-.66	-.65

The regressions are performed by using the panel described in Table II. DV_{it} is a dummy variable that takes the following values: a) 1 for common law countries and 0 for civil law countries in Column I; b) 1 if the index of anti-director rights is above the sample mean, and 0 otherwise in Column II; c) 1 if the index of credictor- rights is above the sample mean, and 0 otherwise in Column III; d) 1 if the index of enforcement is above the sample mean, and 0 otherwise in Column IV; e) 1 if the index of protection investor is above the sample mean, and 0 otherwise in Column V; f) 1 if the index of effective protection investor is above the sample mean, and 0 otherwise in Column VI. Note these indexes are defined in the Table V. The remainder of the variables is defined in Table III. The rest of the information needed to read this table is: i) Heteroskedasticity consistent asymptotic standard error in parentheses. ii) *,** and *** indicate significance at the 1%, 5% and 10% level, respectively; iii) t is the t-statistic for the linear restriction test under the null hypothesis of no significance; iv) z_1, z_2 and z_3 are Wald tests of the joint significance of the reported coefficients, of the time dummies and of the country dummies, respectively, asymptotically distributed as χ^2 under the null of no significance, degrees of freedom in parentheses; v) m_i is a serial correlation test of order i using residuals in first differences, asymptotically distributed as N(0,1) under the null of no serial correlation.

Table 7. Estimation results of the moderating role of the development of capital markets and the contestability of market for corporate control

	I	II	III	IV	V	VI
Constant	-.3484* (.01595)	-.3512* (.01587)	-.3613* (.01582)	-.3555* (.01679)	-.3905* (.01657)	-.3377* (.01949)
FCF_{it}	.11946* (.02248)	.12954* (.02284)	.10996* (.02705)	.13621* (.02608)	.15493* (.02696)	.09933* (.03591)
D_{it}	.01376 (.01609)	-.00014 (.01592)	.00484 (.01598)	.00696 (.01632)	.01134 (.01668)	-.00503 (.02066)
NI_{it}	.31572* (.03523)	.27864* (.03616)	.24288* (.04599)	.26625* (.04458)	.27255* (.04340)	.27516* (.04952)
$TANG_{it}$	.02397** (.01066)	.02231** (.01051)	.02082** (.01052)	.02271** (.01077)	.01625 (.01052)	.02174*** (.01316)
S_{it}	.00550* (.00127)	.00549* (.00125)	.00442* (.00126)	.00475* (.00135)	.00222** (.00133)	.00652* (.00157)
CAT_{it}	.01677* (.00322)	.02532* (.00344)	.01989* (.00362)	.02996* (.00380)	.00186 (.00339)	.01379* (.00354)
$CAT_{it}DV_{it}$	-.01241* (.00321)	-.02218* (.00342)	-.01089* (.00409)	-.02529* (.00422)	.02015* (.00414)	-.01087* (.00365)
t	3.52	2.54	3.72	2.12	8.16	1.68
z_1	88.82 (7)	86.41 (7)	50.66 (7)	73.15 (7)	71.31 (7)	35.62 (7)
z_2	729.74 (12)	733.92 (12)	795.53 (12)	746.80 (12)	747.24 (12)	599.22 (12)
z_3	12.81 (11)	12.26 (11)	11.29 (11)	9.78 (11)	14.53 (11)	8.73 (11)
m_1	-5.37	-5.36	-5.29	-5.36	-5.39	-5.30
m_2	-.64	-.61	-.60	-.59	-.73	-.65

The regressions are performed by using the panel described in Table II. DV_{it} is a dummy variable that takes the following values: a) 1 if the country is classified as a market-oriented system and 0 if it is considered a bank-oriented system in column I; b) 1 if the index of market capitalization to GDP is above the sample mean, and 0 otherwise in Column II; c) 1 if the index of total value traded to GDP is above the sample mean, and 0 otherwise in Column III; d) 1 if the index of market development is above the sample mean, and 0 otherwise in Column IV; e) 1 if the index of banking development is above the sample mean, and 0 otherwise in Column V; f) 1 in countries where an active market for corporate control exists, and 0 otherwise in Column VI. Note these indexes are defined in the Table V. The remainder of the variables is defined in Table III. The rest of the information needed to read this table is: i) Heteroskedasticity consistent asymptotic standard error in parentheses. ii) *,** and *** indicate significance at the 1%, 5% and 10% level, respectively; iii) t is the t-statistic for the linear restriction test under the null hypothesis of no significance; iv) z_1, z_2 and z_3 are Wald tests of the joint significance of the reported coefficients, of the time dummies and of the country dummies, respectively, asymptotically distributed as χ^2 under the null of no significance, degrees of freedom in parentheses; v) m_i is a serial correlation test of order i using residuals in first differences, asymptotically distributed as N(0,1) under the null of no serial correlation.

Table 8. Estimation results of the moderating role of certain institutional characteristics (ownership concentration, independent director of boards and corporate governance)

	I	II	III
Constant	-.3672* (.01588)	-.3583* (.01653)	-.3521* (.01572)
FCF_{it}	.14382* (.02653)	.10286* (.02575)	.10385* (.02473)
D_{it}	.01557 (.01538)	.00411 (.01672)	.01277 (.01554)
NI_{it}	.20439* (.04186)	.32892* (.04299)	.28613* (.03674)
$TANG_{it}$	.00343 (.01039)	.03195* (.01079)	.01077 (.01057)
S_{it}	.00444* (.00127)	.00430* (.00131)	.00554* (.00125)
CAT_{it}	.00774* (.00178)	.00329 (.00318)	.01755* (.00287)
$CAT_{it}DV_{it}$	.01097* (.00368)	.01228* (.00370)	-.01451* (.00294)
t	5.62	6.28	2.31
z_1	51.83 (7)	67.20 (7)	60.92 (7)
z_2	746.82 (12)	749.29 (12)	780.82 (12)
z_3	11.39 (11)	10.55 (11)	12.31 (11)
m_1	-5.32	-5.35	-5.32
m_2	-.60	-.66	-.62

The regressions are performed by using the panel described in Table II. DV_{it} is a dummy variable that takes value 1 if the index of ownership concentration is above the sample mean, and 0 otherwise in Column I. In column II, DV_{it} is a dummy variable that takes value 1 for countries with a predominant two-tier board or when non-executive directors represent a significant proportion on boards and 0 otherwise. DV_{it} is a dummy variable that takes value 1 if the index of corporate governance is above the sample mean, and 0 otherwise in Column III. The indexes of the column 1 and 3 are described in the Table V. The remainder of the variables is defined in Table III. The rest of the information needed to read this table is: i) Heteroskedasticity consistent asymptotic standard error in parentheses. ii) *,** and *** indicate significance at the 1%, 5% and 10% level, respectively; iii) t is the t-statistic for the linear restriction test under the null hypothesis of no significance; iv) z_1, z_2 and z_3 are Wald tests of the joint significance of the reported coefficients, of the time dummies and of the country dummies, respectively, asymptotically distributed as χ^2 under the null of no significance, degrees of freedom in parentheses; v) m_i is a serial correlation test of order i using residuals in first differences, asymptotically distributed as N(0,1) under the null of no serial correlation.

4. RESULTS

In this section we first present the results of model in equation (1), which includes the explanatory variables that have been traditionally considered as determinants of a firm's payout ratio at the same time that also incorporates a variable capturing investors' sentiment, that is, the catering variable. We then extend this model, and we test the implications of the catering theory by means some institutional variables, particularly variables capturing investor's protection, development of capital markets and the orientation of the financial systems, contestability of the market for corporate control, the level of ownership concentration, the effectiveness of boards of directors and corporate governance.

4.1. Results of the basic and extended models

The results of the GMM estimation of our basic model in (1) are provided in the Column I of Table V. The level of a firm's free cash flow positively affects its payout ratio, consistent with Jensen's (1986) theory. The coefficient of leverage is not significant. Consistent with Lintner (1956), the positive relationship between a firm's earnings and its payout ratio predicted is confirmed by our results. Regarding the nature of the firm's assets, our results show that firms with more tangible fixed assets have larger payout ratios. Finally, we find a positive coefficient on size according to which larger companies pay higher dividends, consistent, for instance, with Fama and French (2001) or more recently, Denis and Osobov (2008).

Regarding the influence of a firm's investors' sentiments on its payout ratio, the positive coefficient of the catering variable confirms the link between the propensity to pay dividends

and catering incentives, consistent with Baker and Wurgler (2004b). Our result suggests that firms cater to their investors' preferences, so that they are more prone to increase payout ratios when investors exhibit preference for dividend-paying stocks.

4.2. The moderating role of institutional variables

Once the existence of a catering effect has been corroborated by our results, we go a step forward and investigate whether or not the institutional context moderates this effect.

Columns I to VI of Table VI report the results of the model including the interaction of catering with investor protection. Column I show the interaction of catering with a dummy variable which takes value one for common law countries and value zero for civil law countries. As can be seen, the catering effect in civil law countries (γ_6=0.01316) is stronger than the one in common law countries ($\gamma_6+\lambda$=0.0052, significantly different from zero, see t).

This result corroborates that the stronger the legal protection of investors the smaller the extent to which firms cater to their investors sentiments, supporting the substitute model by La Porta, et al., (2000a).

The results in columns II to VI totally confirm this finding by using other investor protection dummies, such as, anti-director rights, creditor's rights, enforcement, protection investor and effective protection investor[42].

The interaction of the catering effect and development of capital markets/market-oriented system is tested in the models presented in Column I to V of Table VII.

In this case, as shown in column I, DV_{it} takes value one if the country is classified as a market-oriented system and zero if it is considered a bank-oriented system. This way, the coefficient of the catering variable is γ_6 for countries considered a bank-oriented system (since DV_{it} takes value zero), and $\gamma_6+\lambda$ for firms considered a market-oriented system (since DV_{it} takes value one). As can be seen, the catering effect in countries considered a bank-oriented system (γ_6=0.01677) is stronger than the one in countries considered a market-oriented system ($\gamma_6+\lambda$=0.0044, significantly different from zero, see t). As shown in the column I of the Table VII, our evidence not supports Hypothesis 2; however, our evidence suggests that in countries considered a bank-oriented system managers are more encouraged to cater to a large extent to investors' demand for dividends, confirming, once more, the substitute model by La Porta, et al., (2000a).

The results in columns II to V of Table VII corroborate the same conclusions by using other dummies of development of capital markets, such as, market capitalization to GDP, total value traded to GDP, market development and banking development.[43]. Followed, we investigate the interaction between the catering effect and the contestability of market for corporate control by estimating the model presented in Column VI of Table VII. In this case, DV_{it} takes value one in countries with effectiveness market for corporate control, and zero otherwise. As can be seen, the catering effect in countries where the market for corporate control is more deficient (γ_6=0.01379) is stronger than the one in countries where an active market for corporate control exists ($\gamma_6+\lambda$=0.0029, significantly different from zero, see t). These results are very similar to

[42] Column II show the interaction of catering with a dummy variable which takes value one if the index of anti-director rights is above the sample mean and zero otherwise; column III show the interaction of catering with a dummy variable which takes value one if the index of creditors rights is above the sample mean and zero otherwise; column IV show the interaction of catering with a dummy variable which takes value one if the index of enforcement is above the sample mean and zero otherwise; column V show the interaction of catering with a dummy variable which takes value one if the index of protection investor is above the sample mean and zero otherwise and finally, the interaction of the catering effect and effective protection investor is tested in the model presented in Column VI of Table VI with a dummy variable which takes value one if the index of effective protection investor is above the sample mean and zero otherwise.

[43] Column II show the interaction of catering with a dummy variable which takes value one if the index of market capitalization to GDP is above the sample mean and zero otherwise; column III show the interaction of catering with a dummy variable which takes value one if the index of total value traded to GDP is above the sample mean and zero otherwise; column IV show the interaction of catering with a dummy variable which takes value one if the index of market development is above the sample mean and zero otherwise; the interaction of the catering effect and development of capital markets is tested in the model presented in Column V of Table 7 with a dummy variable which takes value one if the index of banking development is above the sample mean and zero otherwise.

the obtained for the previous hypotheses and corroborates that the more active it is the market for corporate control, the smaller the extent to which firms cater to their investors sentiments.

Column I of Table VIII reports the results of the model including the interaction of catering with ownership concentration. As we Know ownership concentration may be an instrumental monitoring variable as it can be a bonding device triggering corporate control actions, therefore, as larger the concentration may the wells shareholders' power for they intend high dividends and this is reflected in our results. Therefore, as shown in column I, DV_{it} takes value one if the index of ownership concentration is above the sample mean and zero otherwise. As can be seen, the catering effect in countries with high levels of ownership concentration ($\gamma_6+\lambda=0.01871$, significantly different from zero, see t) is stronger than the one in countries with low levels of ownership concentration. It seems that catering incentives (i.e., investors' preference for dividend-paying stocks) manifest more strongly in firms with more concentrated patterns, corroborating the monitoring effect.

The interaction of the catering effect and independent boards is tested in the model presented in Column II of Table VIII. In this case, DV_{it} takes value one for countries with a predominant two-tier-board or when non-executive directors represent a significant proportion on boards and zero otherwise. As shown in the table, there is no effect of a firm's investors' sentiments on its payout ratio when the firm has poor executive and supervisory board (γ_6 not significantly different from zero). However, the effect is positive and significant for firms with predominant two-tier-board characterized by the existence of two bodies, which guarantees that the supervisory is separated from an independent of management ($\gamma_6+\lambda=\lambda=0.01228$, significantly different from zero, see t), which confirms that, as expected, the catering effect in countries with a predominant two-tier-board or when non-executive directors represent a significant proportion on is stronger than the one in countries where this monitoring device is poor. These results point out that the expected catering effect clearly manifests itself when there are independent boards in the firm.

Finally, we investigate the interaction between the catering effect and the corporate governance by estimating the model presented in Column III of Table VIII. In this case, DV_{it} takes value one if the corporate governance index is above the sample mean, and zero otherwise. As can be seen in the table, the coefficient of the catering variable ($\gamma_6=0.01755$) is stronger for firms with weak corporate governance than the one with stronger corporate governance ($\gamma_6+\lambda=0.00304$, significantly different from zero). Our evidence suggests that investors' demand for dividends translate into higher payout ratios in those firms with weak governance. These results are consistent with the notion that firms adopt a policy of paying dividends when under pressure to reduce agency costs, consistent for instance, with Harford, Mansi, and Maxwell (2008), who report that firms with weak governance (shareholders rights) hold lower cash reserves and therefore are more likely to pay dividends.

Overall, this evidence provide an excellent robustness check for the results of the basic and extended models, since the sign of the coefficients of both the traditional explanatory variables and the catering variable remain identical once we control for the moderating role of certain institutional variables.

5. CONCLUSIONS

This paper provides a test of the predictions of the catering theory of dividends by proposing a new approach for analysing the effect that investors' sentiments exert on corporate dividend policy.

Our results show that investors´ sentiments impact on the payout ratios in Eurozone countries, US, UK, Canada and Japan after controlling for traditional determinants of dividends, such as the free cash flow, leverage, earnings, tangible fixed assets and size. This finding seems to indicate that dividend policies are driven to some extent by investors' sentiments, thus revealing the desire of firms' managers to cater to such preferences. Therefore, our evidence provides empirical support to the existence of a physiological component in the decision to pay, as proposed by the catering theory.

Our analysis has several policy implications that are particularly relevant allowing the possibility to better understand the implications of catering incentives for dividends by examining the moderating role played by certain institutional variables.

This idea has not been accounted for in prior studies, either theoretically or empirically, but our findings corroborate that the way in which investors appreciate dividend payments depends on the internal and external corporate governance mechanisms. In fact, our research makes a further check to see which institutional variables moderate dividend payout to managers' action to cater.

We trace firm-level corporate governance practices in fifteen countries around the world, and in our empirical tests, we find that the higher a firm's ownership concentration and independence of boards of directors, the better the fit of dividends to investors catering incentives. Our evidence also provides empirical support that external mechanisms are important to forcing firms to disgorge cash in a substitute model.

Therefore, our results show that the substitute model can answer to questions around the world about the amount to distribute on dividends, and the external mechanisms aren't so important when we speak about investor's sentiments.

Furthermore, it is substitute model of dividend increases the importance of dividend payout. Dividend payments can be a substitute for other characteristics because poorly-governed firms need an alternate means of establishing a reputation for acting in the interests of shareholders if they intend to raise capital from markets in the future, and therefore a policy of paying dividends is the most valuable at the margin to firms with agency problems. In fact, our results suggest a presence of a more general phenomenon of the catering effect in companies with high quality of internal corporate governance mechanisms.

Therefore, we believe that not only dividend payout are important for us to understanding managers' catering behaviour but also the institutional context.

REFERENCES

ADAMS, RENÉE B., AND DANIEL FERREIRA, 2007, A Theory of Friendly Boards, Journal of Finance 62, 217-250.

AGANIN, ALEXANDER, AND PAOLO VOLPIN, 2003. History of corporate ownership in Italy, *ECGI, Working Paper,* No. 17/2003.

AIVAZIAN, VAROUJ, LAURENCE BOOTH, AND SEAN CLEARY, 2003, Do emerging market firms follow different dividend policies than firms in the U.S: evidence from 8 emerging markets, *Journal of Financial Research* 26, 371-387.

ALLEN, FRANKLIN, ANTONIO BERNARDO, AND IVO WELCH, 2000, A theory of dividends based on tax clientele, *Journal of Finance* 55, 2499-2536.

ALLEN, FRANKLIN, AND RONI MICHAELY, 2003, Payout Policy, in Constantinides, George, Milton Harris and Rene Stulz ed: *Handbook of the Economics of Finance* (Amsterdam: North-Holland) 337-429.

AMIHUD, YAKOV, AND KEFEI LI, 2006, The declining information content of dividend announcements and the effect of institutional holdings, *Journal of Financial and Quantitative Analysis* 41, 637-660.

ARELLANO, MANUEL, AND STEPHEN BOND, 1991, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations? *Review of Economic Studies* 58, 277-297.

ATTIG, NAJAH, AND YOSER GADHOUM, 2003, The governance of Canadian traded firms: An analysis of the ultimate ownership structure, *SSRN Working Paper,* No 434160.

AUERBACH, ALAN J., AND KEVIN A. HASSETT, 2003, On the marginal source of investment funds, *Journal of Public Economics* 87, 205-232.

BAKER, KENT, GARY POWELL, AND THEODORE VEIT, 2002, Revisiting the dividend puzzle: Do all the pieces now fit? *Review of Financial Economics* 11, 241-261.

BAKER, KENT, SAADI SAMIR, DUTTA SHANTANU, AND GANDHI DEVINDER, 2007, The perception of dividends by Canadian managers: new survey evidence, *International Journal of Managerial Finance* 3, 70-91.

BAKER, MALCOLM, AND JEFFREY WURGLER, 2004a, A catering theory of dividends, *Journal of Finance* 59, 1125–1165.

BAKER, MALCOLM, AND JEFFREY WURGLER, 2004b, Appearing and disappearing dividends: The link to catering incentives, *Journal of Financial Economics* 73, 271-288.

BALL, RAY, ASHOK ROBIN, AND JOANNA WU, 2003, Incentives versus standards: properties of accounting income in four East Asian countries, *Journal of Accounting and Economics* 36, 235-270.

BARCA, FABRIZIO, AND MARCO BECHT, 2002, The Control of Corporate Europe, Oxford University Press.

BARCLAY, MICHAEL, CLIFFORD SMITH, AND ROSS WATTS, 1995, The determinants of corporate leverage and dividend policies, *Journal of Applied Corporate Finance* 7, 4-19.

BAR-YOSEF, SASSON, AND LUCY HUFFMAN, 1986, The Information Content of Dividends: A Signalling Approach, *Journal of Financial and Quantitative Analysis* 21, 47-58.

BEBCHUK, LUCIAN, ALMA COHEN, AND FERRELL ALLEN, 2005, What matters in corporate governance? *Harvard Law School John M. Olin Center Discussion Paper 491*.

BECHT, MARCO, AND AILSA ROEL, 1999, Blockholdings in Europe: An International Comparison, *European Economic Review* 43, 1049-1056.

BECHT, MARCO, AND COLIN MAYER, 2002, Introduction to the control of corporate Europe. IN BARCA, F.& BECHT, M. (Eds.) *The Control of Corporate Europe.* Oxford, University Press.

BECHT, MARCO, PATRICK BOLTON, AND AILSA RÖELL, 2003, Corporate Governance and Control, in G. Constantinides, M. Harris and R. Stulz (eds.), The Handbook of the Economics of Finance, North-Holland 1, 1-109.

BECK, THORSTEN, AND ROSS LEVINE, 2002, Industry Growth and Capital Allocation: Does Having a Market- or Bank- Based System Matter? *Journal of Financial Economics* 64, 147-180.

BECK, THORSTEN, ASLI DEMIRGÜÇ-KUNT, AND ROSS LEVINE, 2003, Law, Endowment, and Finance, *Journal of Financial Economics* 70, 137-181.

BECK, THORSTEN, ASLI DEMIRGÜÇ-KUNT, ROSS LEVINE, AND VOJISLAV MAKSIMOVIC, 2001, Financial Structure and Economic Development: Firm, Industry and Country Evidence. In: Demirguç-Kunt, A., Levine, R. (Eds.), Financial Structure and Economic Growth: A Cross-Country Comparison of Banks, Markets and Development. MIT Press, Cambridge, MA.

BECK, THORSTEN, ROSS LEVINE, AND NORMAN LOAYZA, 2000, Finance and the Sources of Growth, *Journal of Financial Economics* 58, 261-300.

BERGLÖF, ERIK, 1990, Corporate Control and Capital Structure, Essays on Property Rights and Financial Contracts, Institute of International Business, (Stockholm).

BERGLOF, ERIK, AND ENRICO PEROTTI, 1994, The governance structure of the Japanese financial Keiretsu, *Journal of Financial Economics* 36, 259-284.

BERGLOF, ERIK, AND MIKE BURKART, 2003, European takeover regulation, *Economic Policy* 36, 173-208.

BHAGAT, SANJAI, AND BERNARD BLACK, 1999, The Uncertainty Relationship Board Composition and Firm Performance, *Business Lawyer* 54, 921-963.

BHAGAT, SANJAI, AND BERNARD BLACK, 2002, The No-Correlation Between Board Independence and Long-Term Firm Performance, *The Journal of Corporation Law* 27, 231-273.

BHATTACHARYA, SUDIPTO, 1979, Imperfect Information, Dividend Policy, and "The Bird In The Hand" Fallacy, *The Bell Journal of Economics* 10, 259-270.

BRAV, ALON, JOHN GRAHAM, CAMPBELL HARVEY, AND RONI MICHAELY, 2005, Payout policy in the 21st Century, *Journal of Financial Economics* 77, 483-527.

BRENNAN, MICHAEL, AND ANJAN V. THAKOR, 1990, Shareholder Preferences and Dividend Policy, *Journal of Finance* 45, 993–1018.

BRICKLEY, JAMES, AND CHRISTOPHER M. JAMES, 1987, The takehover market, corporate board composition, and ownership structure: the case of banking, *Journal of Law & Economics* 30, 161-180.

BRICKLEY, JAMES, RONALD C. LEASE, AND CLIFFORD SMITH, 1988. Ownership structure and voting on antitakeover amendments, *Journal of Financial Economics* 20, 267-292.

BRICKLEY, JAMES, RONALD C. LEASE, AND CLIFFORD SMITH, 1994, Corporate voting: Evidence from charter amendments, *Journal of Corporate Finance* 1, 5-30.

BROWN, GREGORY W., AND MICHAEL T. CLIFF, 2004, Investor sentiment and the near-term stock market, *Journal of Empirical Finance* 11, 1–27.

BROWN, GREGORY W., AND MICHAEL T. CLIFF, 2005, Investor Sentiment and Asset Valuation, *Journal of Business* 78, 405–440.

CARLIN, WENDY, AND COLIN MAYER, 2003, Finance, Investment, and Growth, *Journal of Financial Economics* 69, 191-226.

CHAPLINSKY, SUSAN, AND GREG NIEHAUS, 1993, Do Inside Ownership and Leverage Share and Common Determinants? *Quarterly Journal of Business and Economics* 32, 61-78.

CHEN, CARL, AND THOMAS STEINER, 1999, Managerial Ownership and Agency Conflicts: a Nonlinear Simultaneous Equation Analysis of Managerial Ownership, Risk Taking, Debt Policy and Dividend Policy, *The Financial Review* 34, 119-136.

CHIRINKO ROBERT, HANS VAN EES, HARRY GARRETSEN, AND ELMER STERKEN, 2004, Investor Protections and Concentrated Ownership: Assessing Corporate Control Mechanisms In The Netherlands, *German Economic Review* 5, 119-138.

COHEN, GIL, AND JOSEPH YAGIL, 2008, On the Catering Theory of Dividends and the Linkage between Investment, Financing and Dividend Policies, *International Research Journal of Finance and Economics* 17, 33-39.

CONN, ROBERT, AND F. CONNELL, 1990, International mergers: returns to US and British firms, *Journal of Business Finance and Accounting* 17, 689-711.

CORREIA DA SILVA, LUIS, MARC GOERGEN, AND LUC RENNEBOOG, 2004, *Dividend Policy and Corporate Governance, ECGI*, (Oxford University Press).

CORREIA DA SILVA, LUIS, MARC GOERGEN, AND LUC RENNEBOOG, 2005, When Do German Firms Change Their Dividends*? Journal of Corporate Finance* 11, 375-399.

COTTNER, JAMES, ANIL SHIVDASANI, AND MARC ZENNER, 1997, Do Independent Directors Enhance Target Shareholder Wealth during Tender Offers?, *Journal of Financial Economics* 43, 195-218.

CREMERS, MARTIJN, AND VINAY NAIR, 2005, Governance mechanisms and equity prices, *Journal of Finance* 60, 2859-2894.

CRUTCHLEY, CLAIRE, AND ROBERT HANSEN, 1989, A Test of the Agency Theory of Managerial Ownership, Corporate Leverage, and Corporate Dividends, *Financial Management* 18, 36-46.

DAHYA, JAY, ORLIN DIMITROV, AND JOHN J. MCCONNELL, 2008, Dominant Shareholders, Corporate Boards and CorporateValue: A Cross-Country Analysis, *Journal of Financial Economics* 87, 73-100.

DARGENIDOU, CHRISTINA, STUART MCLEAY, AND IVANA RAONIC, 2007, Ownership, Investor Protection and Earnings Expectations, *Journal of Business Finance and Accounting* 34, 247-268.

DEANGELO, HARRY, LINDA DEANGELO, AND DOUGLAS SKINNER, 2004, Are dividends disappearing? Dividend concentration and the consolidation of earnings*, Journal of Financial Economics* 72, 425-456.

DEANGELO, HARRY, LINDA DEANGELO, AND RENÉ M. STULZ, 2006, Dividend Policy and the earned/contributed capital mix: a test of the life-cycle theory, *Journal of Financial Economics* 81, 227-254.

DE JONG, HENK W., 1997, The Governance Structure and Performance of Large European Corporations, *The Journal of Management and Governance* 1, 5-27.

DEMIRGÜÇ-KUNT, ASLI AND ROSS LEVINE, 2001, Bank Based and Market-Based Financial Systems: Cross-Country Comparisons. In Demirgüç-Kunt, A., Levine, R. (Eds.), *Financial Structure and Economic Growth: Gross-Country Comparisons of Banks, Markets, and Development*. MIT Press, Cambridge, MA, 81-140.

DEMIRGÜÇ-KUNT, ASLI, AND VOJISLAV MAKSIMOVIC, 1996, Stock Market Development and Firm Financing Choices. *World Bank Economic Review*10, 341-369.

DEMIRGÜÇ-KUNT, ASLI, AND VOJISLAV MAKSIMOVIC, 1999, Institutions, Financial Markets, and Debt Maturity. *Journal of Financial Economics* 54, 295-336.

DEMIRGÜÇ-KUNT, ASLI, AND VOJISLAV MAKSIMOVIC, 2002, Funding Growth in Bank-Based and Market-Based Financial Systems: Evidence from Firm-Level Data. *Journal of Financial Economics* 65, 337-363.

DENIS, DAVID, AND DIANE DENIS, 1994, Majority-Owner Managers and Organizational Efficiency, *Journal of Corporate Finance* 1, 91-118.

DENIS, DAVID, AND ATULYA SARIN, 1999, Ownership and Board Structures in Publicly Traded Corporation, *Journal of Financial Economics* 52, 187-223.

DENIS, DAVID, AND IGOR OSOBOV, 2005, Disappearing dividends, catering incentives and agency costs: international evidence, *SSRN Working paper*, No 778024.

DENIS, DAVID, AND IGOR OSOBOV, 2008, Why Do Firms Pay Dividends? International Evidence on the Determinants of Dividend Policy, *Journal of Financial Economics* 89, 62-82.

DENIS, DAVID, AND JOHN MCCONNELL, 2003, International Corporate Governance, *Journal of Financial and Quantitative Analysis* 38, 1-36.

DEWENTER, KATHRYN, AND VINCENT WARTHER, 1998, Dividends, asymmetric information, and agency conflicts: Evidence from a comparison of the dividend policies of Japanese and US firms, *Journal of Finance* 53, 879-904.

DITTMAR, AMY, JAN MAHRT-SMITH, AND HENRI SERVAES, 2003, International corporate governance and corporate cash holdings, *Journal of Financial and Quantitative Analysis* 38, 111-133.

EASTERBROOK, H. FRANK, 1984, Two agency-cost explanations of dividends, *American Economic Review* 74, 650-659.

ECKBO, ESPEN B., AND SAVITA VERMA, 1994, Managerial shareownership, voting power, and cash dividend policy, *Journal of Corporate Finance* 1, 33-62.

FACCIO, MARA, AND LARRY H.P. LANG, 2002, The Ultimate Ownership of Western European Corporations, *Journal of Financial Economics* 65, 365-395.

FACCIO, MARA, LARRY H.P.LANG, AND LESLIE YOUNG, 2001, Dividends and Expropriation, *American Economic Review* 91, 54–78.

FAIRCHILD, RICHARD, AND GANGGANG ZHANG, 2005, Investor Irrationality and Optimal Open-market Share Repurchasing, ICFAI *Journal of Behavioral Finance* 2, 17–31.

FAMA, EUGENE F., AND KENNETH R. FRENCH, 2001, Disappearing dividends: Changing firm characteristics or lower propensity to pay? *Journal of Financial Economics* 60, 3-44.

FARINHA, JORGE, 2003, Dividend Policy, Corporate Governance and the Managerial Entrenchment Hypothesis: An Empirical Analysis*, Journal of Business Finance and Accounting* 30, 1173-1209.

FENN, GEORGE W., AND NELLIE LIANG, 2001, Corporate Payout Policy and Managerial Stock Incentives, *Journal of Financial Economics* 60, 45-72.

FERREIRA, DANIEL, EMANUEL ORNELAS, AND TURNER JOHN, 2005, Ownership Structure and the Market for Corporate Control, *IBMEC RJ Economics Discussion Paper,* No 2005/09 (University of Georgia).

FRANKS, JULIAN, AND COLIN MAYER, 1996, Hostile Takeovers and the Correction of Managerial Failure, *Journal of Financial Economics* 40, 163-181.

FRANKS, JULIAN, AND COLIN MAYER, 1997, Corporate Ownership and Control in the U.K., Germany and France, *Journal of Applied Corporate Finance* 9, 30-45.

FRANKS, JULIAN, AND COLIN MAYER, 1998, Bank control, takeovers and corporate governance in Germany, *Journal of Banking and Finance* 22, 1385-1403.

GASPAR, MIGUEL, MASSIMO MASSA, AND PEDRO MATOS, 2005, Shareholder investment horizons and the market for corporate control, *Journal of Financial Economics* 76, 135-165.

GEDAJLOVIC, ERIC AND DANIEL SHAPIRO, 1998, Management and Ownership Effects. Evidence from Five Countries, *Strategic Management Journal* 19, 533-553.

GEMMILL, GORDON, 2005, Catering for dividends by stripping mutual-fund portfolios, *SSRN Working paper*, No 807904.

GIANNETTI, MARIASSUNTA, 2003, Do Better Institutions Mitigate Agency Problems? Evidence from Corporate Finance Choices, *Journal of Financial and Quantitative Analysis* 38, 185-212.

GIANNETTI, MARIASSUNTA, AND ANDREI SIMONOV, 2006, Which Investors Fear Expropriation? Evidence from Investors' Portfolio Choices, *Journal of Finance* 61, 1507-1547.

GOERGEN, MARC, AND LUC RENNEBOOG, 2004, Shareholder wealth effects of European domestic and cross-border takeover bids, *European Financial Mana*gement 10, 9-47.

GOMPERS, PAUL, JOY ISHII, AND ANDREW METRICK, 2003, Corporate governance and equity prices, *Quarterly Journal of Economics* 118, 107-155.

GORTON, GARY, AND FRANK A. SCHMID, 2000, Universal banking and the performance of German firms, *Journal of Financial Economics* 58, 29-80.

GRINSTEIN, YANIV, AND RONI MICHAELY, 2005, Institutional Holdings and Payout Policy, *Journal of Finance* 60, 1389-1426.

GROSSMAN, SANFORD J., AND OLIVER HART, 1980, Takeover bids, the free – rider problem, and the theory of the corporation, *Bell Journal of Economics* 11, 42-64.

GRULLON, GUSTAVO, GEORGE KANATAS, AND JAMES P. WESTON, 2004, Advertising, breadth of ownership, and liquidity, *Review of Financial Studies* 17, 439-461.

GUGLER, KLAUS, 2003, Corporate Governance, Dividend Payout Policy, and the Interrelation Between Dividends, R&D, and Capital Investment, *Journal of Banking and Finance* 27, 1297-1321.

GUGLER, KLAUS, AND BURCIN YURTOGLU, 2003, Corporate governance and dividend pay-out policy in Germany, *European Economic Review* 47,731-758.

HAN, BING, 2008, Investor Sentiment and Option Prices, *The Review of Financial Studies* 21, 387-414.

HAN, KI, SUK LEE, AND DAVID SUK, 1999, Institutional Shareholders and Dividends, *Journal of Financial and Strategic Decisions* 12, 53-61.

HARFORD, JARRAD, SATTAR MANSI, AND WILLIAM MAXWELL, 2008, Corporate governance and firm cash holdings in US, *Journal of Financial Economis* 87, 535-555.

HARRIS, MILTON, AND ARTUR RAVIV, 1991, The theory of capital structure, *Journal of Finance* 46, 297-355.

HERMALIN, BENJAMIN, AND MICHAEL WEISBACH, 1988, The Determinants of Board Composition, *Rand Journal of Economics* 19, 589-606.

HERMALIN, BENJAMIN, AND MICHAEL WEISBACH, 2003, Boards of directors as an endogenously determined institution: A survey of the economic literature, *Economic Policy* 9, 7-26.

HO, CHI-KUN, 2005, Corporate Governance and Corporate Competitiveness: an international analysis, *Corporate Governance*13, 211-253.

HOBERG, GERARD, AND NAGPURNANAND PRABHALA, 2008, Disappearing dividends, Catering, and Risk, *Review of Financial Studies.* Forthcoming.

HOLDER, MARK, FREDERICK LANGREHR, AND LAWRENCE HEXTER, 1998, Dividend Policy Determinants. An Investigation of the Influences of Stakeholder Theory, *Financial Management* 27, 73-82.

HOLMSTOM, BENGT, AND STEVEN KAPLAN, 2001, Corporate governance and merger activity in the US: Making sense of the 1980's and 1990's, *Journal of Economic Perspectives* 15, 121–144.

HOLMSTOM, BENGT, AND STEVEN KAPLAN, 2003, The State of U.S Corporate Governance: What's Right and what's wrong? *Journal of Applied Corporate Finance*, 8-20.

HOPE, OLE-KRISTIAN, 2003, Disclosure Practices, enforcement of accounting standards and analysts' forecast accuracy: An International Study, *Journal of Accounting Research* 41, 235-272.

HÖPNER MARTIN, AND GREGORY JACKSON, 2001, An Emerging Market for Corporate Control? The Mannesmann Takeover and German Corporate Governance, *MPIfG Discussion Paper* 01/4.

HOPT, KLAUS, AND PATRICK LEYENS, 2004, Board models in Europe: recent developments of internal corporate governance structures in Germany, United Kingdom, France, and Italy, *ECGI, Law Working Paper*, No 18/2004.

HOPT, KLAUS, HIDEKI KANDA, MARK ROE, EDDY WYMEERSCH, AND STEFAN PRIGGE, 2000. Comparative Corporate Governance. The State of the Art and Emerging Research, *The International and Comparative Law Quarterly* 49, 510-511.

HSIEH, JIM, AND QINGHAI WANG, 2006, Determinants of the Trends in Aggregate Corporate Payout Policy, Working Paper, (George Mason University).

HU, AIDONG, AND PRAVEEN KUMAR, 2004, Managerial Entrenchment and Payout Policy, *Journal of Financial & Quantitative Analysis* 39, 759-790.

IOSCO, 2002, Principles of Auditor Independence and the Role of Corporate Governance in Monitoring an Auditor's Independence.

JARRELL, GREGG, JAMES A. BRICKLEY, AND JEFFRY NETTER, 1988, The Market for corporate control: The empirical evidence since 1980, *Journal of Economic Perspectives* 2, 49-68.

JENSEN, GERARD R., DONALD SOLBERG, AND THOMAS S. ZORN, 1992, Simultaneous Determination of Insider Ownership, Debt, and Dividend Policies, *The Journal of Financial and Quantitative Analysis* 27, 247-263.

JENSEN, MICHAEL, 1986, Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review* 76, 323-329.

JENSEN, MICHAEL, 1989, Eclipse of the public corporation, *Harvard Business Review* 89, 61-84.

JENSEN, MICHAEL C., AND WILLIAM MECKLING, 1976, Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure, *Journal of Financial Economics* 3, 305-360.

JENSEN, MICHAEL C., AND RICHARD S. RUBACK, 1983, The market for corporate control: The scientific evidence, *Journal of Financial Economics* 11, 5-50.

JOHN, KOSE, AND ANZEHELA KNYAZEVA, 2006, Payout policy, agency conflicts, and corporate governance, *SSRN Working Paper*, No 841064.

JOHN, KOSE, AND JOSEPH WILLIAMS, 1985, Dividends, dilution, and taxes: A signalling equilibrium, *Journal of Finance* 40, 1053-1070.

JOHNSON, SIMON, AND ANDREI SHLEIFER, 1999, Coase v. The Coasians, *Harvard Institute of Economics Research*, No 1885.

KAPLAN, STEVEN, 1997, Corporate Governance and Corporate Performance: a Comparison of Germany, Japan and The U.S, *Journal of Applied Corporate Finance* 9 (4), 86-93.

KAPLAN, STEVEN N., AND BERNADETTE A. MINTON, 1994, Appointments of Outsiders to Japanese Boards. Determinants and Implications for Managers, *Journal of Financial Economics* 36, 225-258.

KAUSTIA, MARKKU, AND SAMI TORSTILA, 2008, Political Preferences and Stock Market Participation, *SSRN Working Paper*, No 966254.

KHAN, TEHMINA S., 2006, Company Dividends and ownership structure: evidence from UK panel data, *The Economic Journal* 116, 172-189.

KESTER, CARL W.,1991, Japanese takehovers: the global contest for corporate control, Journal of International Business Studies 36, 42-61.

KING, ROBERT, AND ROSS LEVINE, 1993, Finance and Growth: Schumpeter Might Be Right, *Quarterly Journal of Economics* 108, 717-38.

KINI, OMESH, WILLIAM KRACAW, AND SHEHZAD MIAN, 2004, The Nature of Discipline by Corporate Takeovers,*Journal of Finance* 59,1511-1552.

KOCH, ADAM S., AND AMY X. SUN 2004, Dividend changes and the persistence of past earnings changes, *Journal of Finance* 59, 2093–2116.

KÖKE, JENS, 2002, The Market for Corporate Control in a bank-based economy: A Governance Device? *Journal of Corporate Finance* 10, 53-80.

KOTHARE, MEETA, 1997, The effects of equity issues on ownership structure and stock liquidity: A comparison of rights and public offerings, *Journal of Financial Economics* 43, 131-148.

LAI, RICHARD, 2004, A catering theory of analyst bias, *SSRN Working Pape*r, No.548582.

LAMBERT, RICHARD, WILLIAM LANEN, AND DAVID LARCKER, 1989, Executive stock option plans and dividend policy*, Journal of Financial and Quantitative Analysis* 24, 409-425.

LA PORTA, RAFAEL, FLORENCIO LOPEZ-DE-SILANES, ANDREI SHLEIFER AND ROBERT VISHNY, 1997,Legal Determinants of External Finance, *The Journal of Finance* 52, 1131-1150.

LA PORTA, RAFAEL, FLORENCIO LOPEZ-DE-SILANES, ANDREI SHLEIFER, AND ROBERT VISHNY, 1998, Law and Finance, *Journal of Political Economy* 106, 1113-1155.

LA PORTA, RAFAEL, FLORENCIO LOPEZ-DE-SILANES, ANDREI SHLEIFER, AND ROBERT VISHNY, 2000a, Agency problems and dividend policies around the world, *Journal of Finance* 55, 1-33.

LA PORTA, RAFAEL, FLORENCIO LOPEZ-DE-SILANES, ANDREI SHLEIFER, AND ROBERT VISHNY, 2000b, Investor Protection and Corporate Governance, *Journal of Financial Economics* 58, 3-28.

LA PORTA, RAFAEL, FLORENCIO LOPEZ-DE-SILANES, ANDREI SHLEIFER, AND ROBERT VISHNY, 2002, Investor Protection and Corporate Valuation, *Journal of Finance* 57, 1147-1170.

LEUZ, CHRISTIAN, DHANANJAY NANDA, AND PETER D. WYSOCKI, 2003, Earnings Management and Investor Protection: an International Comparison, *Journal of Financial Economics* 69, 505-527.

LEVINE, ROSS, 2002, Bank-based or market-based financial systems: which is better? *National Bureau of Economic Research* 9138, 1-27.

LEVINE, ROSS, AND SARA ZERVOS, 1998, Stock markets, banks, and economic growth. *American Economic Review* 88, 537-558.

LI, WEI, AND ERIK LIE, 2006, Dividends changes and catering incentives, *Journal of Financial Economics* 80, 293-308.

LINTNER, JOHN, 1956, Distribution of incomes of corporations among dividends, retained earnings, and taxes, *American Economic Review 46*, 97-113.

LUCAS, DEBORAH, AND ROBERT MCDONALD, 1998, Shareholder heterogeneity, adverse selection, and payout policy, *Journal of Financial and Quantitative Analysis* 33, 233-254.

MANCINELLI, LUCIANA, AND AYDIN OZKAN, 2006, Ownership structure and dividend policy: Evidence from Italian firms, *European Journal of Finance* 12, 265,282.

MANNE, HENRY, 1965, Mergers and the Market for Corporate Control, *Journal of Political Economy* 73, 110-120.

MARRIS, ROBIN, 1963, A Model of the "Managerial" Enterprise, *Quarterly Journal of Economics* 77, 185-209.

MASULIS, RONALD, CONG WANG, AND FEI XIE, 2007, Corporate Governance and Acquirer Returns, *Journal of Finance* 62, 1851-1889.

MAYER, COLIN, AND OREN SUSSMAN, 2001, The Assessment: Finance, Law and Growth, *Oxford Review of Economic Policy* 17, 457-466.

MIGUEL, ALBERTO, JULIO PINDADO, AND CHABELA DE LA TORRE, 2005, How do entrenchment and expropriation phenomena affect control mechanisms? *Corporate Governance: An Internacional Review* 13, 505-516.

MILLER, MERTON, AND FRANCO MODIGLIANI, 1961, Dividend Policy, Growth and the Valuation of Shares, *Journal of Business* 34, 411-433.

MILLER, MERTON, AND KEVIN ROCK, 1985, Dividend policy under asymmetric information, *Journal of Finance* 40, 1031-1051.

MOERLAND, PIETER W., 1995, Corporate Ownership and Control Structures: An International Comparison, *Review of Industrial Organization* 10, 443-464.

MORCK, RANDALL, 2004, Behavioral finance in corporate governance: Independent directors and non-executive chairs, *Harvard Institute of Economic Research Discussion Paper*, No 2037.

MORCK, RANDALL, AND MASAO NAKAMURA, 1999, Banks and corporate control in Japan, *Journal of Finance* 54, 319-339.

MORCK, RANDALL, PERCY MICHAEL, GLORIA TIAN, AND BERNARD YEUNG, 2004, The rise and fall of the widely held firm: A history of corporate ownership in Canada, *NBER Working Paper*, No W10635, (University of Chicago).

MOULTON, BRENT, 1986, Random Group Effects and the Precision of Regression Estimates, *Journal of Econometrics* 32, 385-397.

MOULTON, BRENT, 1987, Diagnostics for Group Effects in Regression Analysis, *Journal of Business and Economic Statistics*, 5, 275-282.

NEVES, ELISABETE, JULIO PINDADO, AND CHABELA DE LA TORRE, 2006, Dividends: new evidence on catering theory, Working Paper 14/06 (Salamanca University).

NISSIM, DORON, AND AMIR ZIV, 2001, Dividend Changes and Future Profitability, *Journal of Finance* 56, 2111-2134.

OECD, 2004, Principles of Corporate Governance.

PAGANO, MARCO, AND DAVIDE LOMBARDO, 1999, Law and Equity Markets, *CEPR Discussion Papers*, No 2276.

PAGANO, MARCO, AND PAOLO VOLPIN, 2001, The political Economy of Finance, *Oxford Review of Economic Policy* 17, 502-519.

PAN, CARRIE, 2007, Why are firms with entrenched managers more likely to pay dividends? *SSRN Working Paper*, No 905816.

PEREZ-GONZALEZ, FRANCISCO, 2003, Large shareholders and dividends: evidence from U.S. tax reforms, Working paper (Columbia University).

PEROTTI, ENRICO, AND VON THADDEN, 2006, The Political Economy of Corporate Control and Labor Rents, *Journal of Political Economy* 114, 145-174.

PHAM, PETER, PETKO KALEV, AND ADAM STEEN, 2003, Underpricing, stock allocation, ownership structure and post-listing liquidity of newly listed firms, *Journal of Banking and Finance* 27, 919-947.

PINDADO, JÚLIO, AND LUIS RODRIGUES, 2004, Parsimonious Models of Financial Insolvency in Small Companies, *Small Business Economics* 22, 51-66.

PROWSE, STEPHEN, 1992, The structure of corporate ownership in Japan, *Journal of Finance* 47, 1121-1140.

PROWSE, STEPHEN, 1995, Corporate governance in an international perspective: a survey of corporate control mechanisms among large firms in the US, UK, Japan and Germany, *Financial Markets Institutions & Instruments* 4, 1-63.

RAJAN, RAGHURAM, AND LUIGI ZINGALES, 1995. What Do We Know about Capital Structure? Some Evidence from International Data, *Journal of Finance* 50, 1421-1460.

RAJAN, RHAGURAM AND LUIGI ZINGALES, 1998, Financial dependence and growth, *American Economic Review* 88, 559-587.

RAHEJA, CHARU G., 2005, Determinants of board size and composition: A theory of corporate boards, *Journal of Financial and Quantitative Analysis* 40, 283,306.

RAVI, JAIN, 2007, Institutional and individual investor preferences for dividends and share repurchases, *Journal of Economics and Business* 59, 406-429.

RENNEBOOG, LUC, 2000, Ownership, managerial control and the governance of poorly performing companies listed on the Brussels stock exchange, *Journal of Banking and Finance* 24, 1959-1995.

RENNEBOOG, LUC, FRANKS, JULIAN, AND COLIN MAYER, 2001, Who disciplines management in poorly performing companies? *Journal of Financial Intermediation* 10, 209-248.

ROE, MARK, 1994, Srong managers weak owners: the political roots of American Corporate Finance (University Press, Princeton, NJ).

ROE, MARK, 2002, Corporate Law's Limits, *Journal of Legal Studies* 31, 233-272.

ROE, MARK, 2006, Political Determinants of Corporate Governance- Political Context, Corporate Impact. Oxford and New York: Oxford University Press.

ROSS, STEPHEN, 1977, The determination of financial structure: the incentive signalling approach, *Bell Journal of economics* 8, 23-40.

ROZEFF, MICHAEL, 1982, Growth, beta and agency costs as determinants of dividend payout ratios, *Journal of Financial Research* 5, 249-259.

SCHOOLEY, DIANE, AND DWAYNE BARNEY, 1994, Using Dividend Policy and Managerial Ownership to Reduce Agency Costs, *Journal of Financial Research* 17, 363-73.

SCOTT, JAMES, 1977, Bankruptcy, secured debt and optimal capital structure, *Journal of Finance* 32, 1-19.

SHLEIFER, ANDREI, AND DANIEL WOLFENZON, 2002, Investor protection and equity markets, *Journal of Financial Economics* 66, 3–27.

SHLEIFER, ANDREI, AND ROBERT VISHNY, 1986, Large Shareholders and Corporate Control, *Journal of Political Economy* 94, 461-488.

SHLEIFER, ANDREI, AND ROBERT VISHNY, 1997, A Survey of Corporate Governance. *The Journal of Finance* 52, 737-783.

SHORT, HELEN, HAO ZHANG, AND KEVIN KEASEY, 2002, The Link between Dividend Policy and Institutional Ownership *Journal of Corporate Finance* 8, 105-122.

SKINNER, DOUGLAS, 2008, The Evolving Relation between Earnings, Dividends, and Stock Repurchases, *Journal of Financial Economics* 87, 582-609.

TALMOR, ELI, 1981, Asymmetric Information, Signaling, and Optimal Corporate Financial Decisions, *Journal of Financial and Quantitative Analysis* 16, 413-435.

TSE, CHIN-BUN, 2004, A critique of the logic of the agency costs theory for dividends in the finance literature, *Journal of Interdisciplinary Economics* 15, 25-40.

VAGTS, DETLEV, 2002. Comparative Corporate Law- The New Wave. In Schweitzer, R. and U. Gasser, (Eds). *Festschrift for Jean-Nicolas Druey,* Zurich: Schlthess Juristische Medien.

VISSING-JORGENSEN, ANNETTE, 2002, Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures, *NBER Working Paper,* No 8884.

VOLPIN, PAOLO, 2002, Governance with Poor Investor Protection: Evidence from Top Executive Turnover in Italy, *Journal of Financial Economics* 64, 61-90.

WANG, KO, JOHN ERICKSON, AND GEORGE W. GAU, 1993, Dividend Policies and Dividend Announcement Effects for Real Estate Investment Trusts, *Journal of the American Real Estate and Urban Economics Association* 21, 185-201.

WYMEERSCH, EDDY, 1998, A Status Report on Corporate Governance Rules and Practices in Some Continental European States, *In Comparative Corporate Governance. The State of the Art and Emerging Research, Hopt, K., Kanda, H., Roe, M. Wymeersch, E., Prigge, S. Eds. Oxford: Clarendon Press*.

YERMACK, DAVID 2006, Flights of Fancy: Corporate Jets, CEO Perquisites, and Inferior Shareholder Returns, *Journal of Financial Economics* 80, 211-242.

ZWIEBEL, JEFFREY, 1996, Dynamic capital structure under managerial entrenchment, *American Economic Review* 86, 1197-1215.

ÁREA VII-IX

ECONOMÍA EUROPEA, INTERNACIONAL Y DE LA INTEGRACIÓN

ECONOMIA EUROPEIA, INTERNACIONAL E DA INTEGRACÃO

UN ANÁLISIS SVAR DE LA EFECTIVIDAD DE LA POLÍTICA FISCAL EN ESPAÑA

JULIÁN RAMAJO HERNÁNDEZ
E-mail: ramajo@unex.es
Department of Economics
University of Extremadura

Resumen

El objetivo básico del trabajo es investigar los efectos dinámicos de la política fiscal sobre la actividad económica en España. En primer lugar se especifican y estiman modelos VAR estructurales que incluyen las principales variables macroeconómicas de la economía española, así como las dos variables fiscales básicas, ingresos y gastos públicos, o su versión agregada, el déficit público. A continuación se calculan las distintas funciones de respuesta al impulso antes shocks imprevistos en las variables fiscales, analizándose los distintos tipos de efectos que se producen a corto o medio plazo sobre las variables de política económica y monetaria consideradas.

Palabras clave: Política fiscal, Shocks fiscales, Modelos VAR estructurales, España.
Área Temática: Economía Europea.

1. INTRODUCCIÓN: CONSIDERACIONES GENERALES ACERCA DE LOS EFECTOS MACROECONÓMICOS DE LA POLÍTICA FISCAL

Las investigaciones acerca de los efectos estabilizadores de la política fiscal en general, y sobre los beneficios y costes de los ajustes fiscales en particular, se han incrementado en los últimos años por varios motivos.

Por un lado, las situaciones de desequilibrio generalizado en las cuentas públicas de las economías occidentales durante la década de los noventa, y el debate posterior sobre la sostenibilidad de los sistemas de protección social vigentes teniendo en cuenta las tendencias demográficas desfavorables, dieron paso a un proceso de reducción paulatina de los niveles de déficit y de endeudamiento. En el caso particular de Europa, se han realizado reformas profundas en las políticas de gasto y de ingresos, y también se han observado intensos procesos de consolidación fiscal, sobre todo en los países de la UME, los cuales debían cumplir los objetivos contemplados en el Tratado de Maastricht.

Por otro lado, en los momentos actuales de crisis financiera y macroeconómica internacional se plantean distintas posibilidades para incrementar la demanda agregada de las distintas economías, siendo una de las más unánimes la implementación de un paquete combinado de medidas fiscales tales como incrementos del gasto público productivo, recortes de impuestos específicos o aumento de las transferencias a las familias y/o empresas (Spilimbergo et al., 2008). Ante este escenario macroeconómico adverso, se pone de manifiesto la necesidad de conocer a priori los efectos dinámicos derivados de shocks fiscales y, si cabe más aún, disponer de una cuantificación adecuada –tanto en términos de magnitud como de duración- de los distintos multiplicadores fiscales (ver la reciente nota de Barro (2009) en la que se discute acerca de los multiplicadores asumidos en el plan de estímulo fiscal que va a implementar el recién elegido equipo de gobierno de Estados Unidos).

Aunque mucha de la literatura existente sobre los efectos de la política fiscal se ha basado en los resultados de simulación obtenidos con modelos macroeconométricos de mediana o gran escala, las investigaciones basadas en modelos a pequeña escala y, en particular, los trabajos basados en los modelos estructurales de vectores autorregresivos (SVAR), se han convertido en

los últimos años en la herramienta básica para este tipo de análisis empíricos.[1] Por este motivo en este trabajo se utilizarán este tipo de modelos para el análisis de la efectividad de la política fiscal en España.

2. ASPECTOS TEÓRICOS: REVISIÓN DE TRABAJOS TEÓRICOS Y EMPÍRICOS QUE ANALIZAN LOS EFECTOS DE LA POLÍTICA FISCAL

A continuación se hace una breve revisión de los trabajos más relevantes que han utilizado la metodología VAR para el análisis de los efectos macroeconómicos de la política fiscal.

Tras los trabajos seminales de Blanchard y Perotti (2002), Edelberg *et al.* (1998), Fatás y Mihov (2000) o Mountford y Uhlig (2002), en los que se analizaban los efectos macroeconómicos dinámicos derivados de shocks en las políticas de gasto o de impuestos en Estados Unidos, han surgido en los últimos años varias investigaciones en las que, a través de modelos SVAR, se ha estudiado el proceso de transmisión de los shocks fiscales en diferentes países europeos, comparándose los resultados en algunos casos con los obtenidos para otros países de la OCDE.

Así, en Dalsgaards y De Serres (2000) se estiman modelos SVAR para once países de la Unión Europea, con el objetivo de estimar lo que los autores denominan "márgenes presupuestarios de prudencia".

Como en García y Verdelhan (2001), en van Aarle *et al.* (2003) se estudia la transmisión de las políticas fiscal y monetaria en la zona Euro agregada, pero además se estiman modelos SVAR individuales para los países de la UME al objeto de analizar las diferencias entre los distintos Estados.

Por otro lado, Muscatelli *et al.* (2002) estiman modelos VAR para cinco países de la OCDE con el objetivo de examinar la interdependencia que existe entre las políticas monetaria y fiscal.

En Canzoneri *et al.* (2003) se emplea un modelo SVAR para explorar la relevancia de los efectos *spillover* que producen los cambios en la política fiscal americana sobre el PIB de Francia, Italia y Reino Unido, encontrando que dichos efectos son significativos y de dimensión considerable.

De Arcangelis y Lamartina (2003) investigan los efectos de shocks en las políticas de gasto e impuestos de varios países de la OCDE (Alemania Occidental, Francia, Italia y Estados Unidos), concluyendo que para los países europeos las decisiones fiscales obedecen básicamente a las distintas políticas de gasto público y que los shocks fiscales, tanto de ingresos como de gastos, tienen efectos Keynesianos sobre la producción, aunque los multiplicadores fiscales son bastante bajos en términos cuantitativos.

En el trabajo de Claeys (2004), al igual que en otros citados anteriormente, también se analiza la interacción existente entre las políticas fiscal y monetaria, en este caso para nueve países de la OCDE, seis de ellos europeos. A diferencia de otras investigaciones, en ésta se utilizan condiciones de largo plazo que imponen restricciones de cointegración en los modelos estimados.

Por otra parte, Perotti (2005) estudia en detalle los efectos la política fiscal sobre el PIB y sus componentes, la inflación y los tipos de interés en cinco países de la OCDE, siendo Alemania y Reino Unido dos de estos países. En este trabajo se analiza la robustez de los resultados frente a distintos enfoques alternativos, llegándose a una serie de conclusiones contrastadas acerca de la intensidad y persistencia de los efectos estimados, así como de la validez en el tiempo de dichos efectos.

Giuliodori y Beetsma (2005) incluyen las importaciones en un modelo SVAR para Alemania, Francia e Italia al objeto de contrastar la relevancia de los *spillovers* comerciales derivados de las diferentes políticas fiscales en Europa. De forma complementaria, en Beetsma *et al.* (2006) se estima primero un modelo SVAR fiscal y a continuación un bloque de

[1] Véanse, por ejemplo, los artículos de Briotti (2005), Capet (2004), Hemming *et al.* (2002) o Perotti (2006), donde se revisa la literatura teórica y empírica sobre la efectividad de la política fiscal; o el trabajo de McKibbin (1997), donde se comparan los resultados obtenidos a partir de varios modelos que cuantifican los beneficios derivados de la coordinación de las distintas políticas macroeconómicas.

ecuaciones sobre exportaciones para un panel de catorce países de la Unión Europea, concluyéndose de nuevo con la relevancia y significación estadística de dichas externalidades de comercio.

Usando datos anuales, Artis y Onorante (2006) investigan el impacto de la última revisión en las normas establecidas en el PEC, examinando los posibles efectos de distintas reglas fiscales alternativas.

En el trabajo de Marcellino (2006) se estiman modelos SVAR que incluyen como variable externa el tipo de interés de Alemania junto con las correspondientes variables internas nacionales de Alemania, Francia, España e Italia, encontrándose diferencias significativas en las respuestas de estos países ante cambios no sistemáticos en la política fiscal, así como un elevado grado de incertidumbre en el tamaño de dichas respuestas.

Bénassy-Quéré y Cimadomo (2006) analizan la variación temporal de los multiplicadores fiscales (internos) de Alemania, Estados Unidos y Reino Unido y, además, investigan la relevancia y variabilidad de los *spillovers* fiscales de Alemania hacia las siete mayores economías de la Unión Europea (Austria, Bélgica, España, Francia, Italia, Holanda y Reino Unido). Sus resultados ponen de manifiesto la débil efectividad doméstica de las políticas fiscales discrecionales y, al nivel externo, el signo positivo de los multiplicadores fiscales cruzados sugiere que el canal del comercio internacional sirve de contrapeso a la transmisión natural de la política fiscal a través de los tipos de interés.

A diferencia de todos los trabajos anteriores, Schalck (2007) proporciona estimaciones de los efectos de la política fiscal en cuatro países europeos usando modelos SVAR no lineales, mostrando que los efectos macroeconómicos de los shocks fiscales difieren no sólo entre países, sino también según el régimen expansivo o recesivo que prevalezca en cada momento.

Finalmente, en Ramajo (2008) se analizan los efectos de la política fiscal en la Unión Europea y, en particular, se explora la posibilidad de que existan externalidades geográficas en el proceso de transmisión de los shocks fiscales nacionales. Para ello, en primer lugar se estiman modelos VAR estructurales –SVAR- para cada país y se analizan los efectos internos derivados de distintos shocks macroeconómicos. A continuación se investiga la existencia de *spillovers* fiscales entre Alemania y los distintos países europeos considerados a partir de modelos SVAR bi-nacionales. Utilizando datos sobre doce miembros de la UE los resultados muestran, en primer lugar, un alto grado de heterogeneidad en la respuesta de las distintas economías en función del tipo de shock considerado y, en segundo lugar, que los *spillovers* sobre el output gap derivados de shocks fiscales en Alemania no son en general estadísticamente significativos, aunque sí que lo son los efectos externos sobre los precios.

Al nivel individual, y para varios países de la Unión Europea, utilizando una metodología similar a la de los trabajos citados anteriormente, también cabe destacar los trabajos para Alemania de Höppner (2001) y Heppke-Falk *et al.* (2006), para Austria de Badinger (2006), para España de De Castro (2006) y De Castro y Cos (2006), para Francia de Biau y Girad (2005) y Creel *et al.* (2005), o para Italia de Giordano *et al.* (2007).

3. FORMULACIÓN DE UN MODELO VAR ESTRUCTURAL PARA LA ECONOMÍA ESPAÑOLA

A) El modelo teórico subyacente

El modelo VARX (modelo de vectores auto-regresivos con variables débilmente exógenas) que se plantea como especificación empírica básica puede ser visto como la forma reducida de un modelo estructural macroeconómico estándar o, de forma más general, como una aproximación a la solución condicionada (y sin las restricciones impuestas por la teoría) de un modelo de equilibrio general dinámico estocástico (MEGDE).

Así, el MEGDE canónico (más concretamente, su versión log-linealizada) que subyace a la especificación de vectores autorregresivos que se utilizará estaría compuesto por las siguientes cuatro ecuaciones:

- La primera ecuación es una curva IS que explica la desviación del (logaritmo) output real respecto de su nivel natural (*output gap*), $\tilde{y}_t = y_t - y_t^0$, a partir de una media ponderada de sus valores pasados y futuros, y del tipo de interés real contemporáneo, $i_t - E_t\pi_{t+1}$, siendo i_t el tipo de interés nominal a corto plazo:

$$\tilde{y}_t = \beta_0 + \beta_1\tilde{y}_{t-1} + (1-\beta_1)E_t\tilde{y}_{t+1} - \beta_2(i_t - E_t\pi_{t+1}) + \varepsilon_t^{\tilde{y}}$$

- La segunda ecuación del sistema es una curva de Phillips neo-Keynesiana que explica la *inflación*, π_t, a partir sus valores pasados y de las expectativas sobre su futura evolución, y del output gap de la economía en cada período:

$$\pi_t = \alpha_0 + \alpha_1\pi_{t-1} + (1-\alpha_1)E_t\pi_{t+1} + \alpha_2\tilde{y}_t + \varepsilon_t^{\pi}$$

La producción potencial de la economía, y_t^0, en general suele aproximarse por una tendencia lineal o por un componente inobservado (estimable) pero, en cualquier caso, depende del nivel de productividad y del grado de actividad exterior, medido por el output externo y_t^* (Galí and Monacelli, 2005); además, dado el elevado nivel de diseminación de la tecnología en la actualidad, la productividad a su vez puede considerarse como una función del nivel de output externo. Entonces, el output gap de la economía, $\tilde{y}_t$, puede considerarse como una función del output exterior, y_t^* o, de forma más general, del output gap externo, $\tilde{y}_t^* = y_t^* - y_t^{0*}$.

- La tercera ecuación es una regla fiscal simple que explica la *posición fiscal* del sector público (superávit o déficit de las AA.PP.), fp_t, en función de su evolución pasada y del output gap contemporáneo:[2]

$$fp_t = \gamma_0 + \gamma_1 fp_{t-1} + \gamma_2\tilde{y}_t + \varepsilon_t^{fp}$$

- Y la cuarta, y última ecuación, es una regla de Taylor estándar que describe cómo se determina el *tipo de interés* nominal a corto plazo en respuesta a su valores pasados, a las desviaciones esperadas de la inflación y el output respecto de sus objetivos, y también en respuesta a la inflación externa esperada:

$$i_t = \delta_0 + \delta_1 i_{t-1} + \delta_2 E_t[\pi_t - \pi^0] + \delta_3 E_t[y_t - y_t^0] + \delta_4 E_t\pi_{t+1}^* + \varepsilon_t^i$$

B) Especificación econométrica del modelo SVAR

Si se asume un proceso VAR(1) de primer orden para la distribución de las variables externas, $x_t^* = e_0 + E_1^* x_{t-1}^* + v_t$, puede demostrarse que la solución al sistema de expectativas racionales anterior, si existe y es única, es del tipo

$$x_t = a_0 + A_1 x_{t-1} + B_0^* x_t^* + B_1^* x_{t-1}^* + e_t$$

donde $x_t = (\tilde{y}_t, \pi_t, fp_t, i_t)$ y $x_t^* = (\tilde{y}_t^*, \pi_t^*)$.

Si se añadiesen más retardos tanto de las variables endógenas x_t como de las débilmente exógenas x_t^*, la solución condicionada del modelo teórico planteado tomaría entonces la forma

$$x_t = a_0 + A_1 x_{t-1} + A_2 x_{t-2} + \ldots + A_p x_{t-p} + B_0^* x_t^* + B_1^* x_{t-1}^* + B_2^* x_{t-2}^* + \ldots + B_q^* x_{t-q}^* + e_t$$

En términos más generales, la expresión formal para un modelo VARX estructural (SVARX) con un vector x_t de $m_x \times 1$ variables endógenas, un vector x_t^* de $m_{x^*} \times 1$ variables externas débilmente exógenas y un vector x_t^{**} de otras $m_{x^{**}} \times 1$ variables estrictamente exógenas (como los precios del petróleo), incluyendo aquellas de tipo determinístico (tendencias temporales, variables estacionales y/o variables ficticias) es la siguiente:

$$x_t = a_0 + A_1 x_{t-1} + A_2 x_{t-2} + \ldots + A_p x_{t-p} + B_0^* x_t^* + B_1^* x_{t-1}^* + B_2^* x_{t-2}^* + \ldots + B_q^* x_{t-q}^* + C x_t^{**} + e_t$$

[2] En el caso de desagregar las cuentas públicas, esta ecuación se desdoblaría en dos similares, substituyendo el indicador de la posición fiscal global, *fp*, por las variables de ingresos (*t*) y gastos (*g*) de las AA.PP.

donde e_t representa un $m_x \times 1$ vector de errores canónicos independientes, no correlacionados en el tiempo, con media cero y matriz de covarianzas constante y definido-positiva.

Esta expresión representa la forma general del modelo SVARX que se estimará en la parte empírica, y debe completarse con la expresión

$$Ae_t = Du_t$$

que relaciona el vector de errores canónicos, e_t, con el $m_x \times 1$ vector de errores estructurales, u_t. Este último está compuesto en nuestro caso por un shock de de oferta agregada (u_t^s), un shock de demanda agregada (u_t^d), un shock fiscal (u_t^f) y un shock monetario (u_t^m), en correspondencia con las variables contenidas en el vector $x_t = (\tilde{y}_t, \pi_t, fp_t, i_t)$ de variables endógenas.

4. LA BASE DE DATOS PARA LA ECONOMÍA ESPAÑOLA

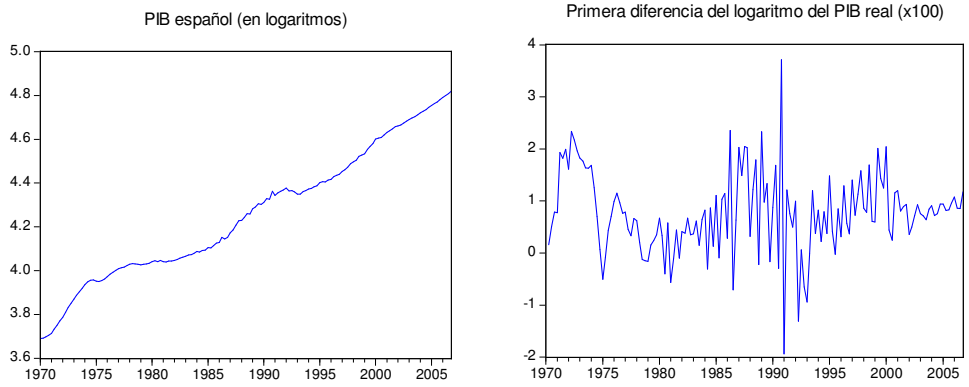
En este apartado se hará una breve descripción de la base de datos construida para poder estimar el modelo SVAR para la economía española con el que se analizarán posteriormente los efectos dinámicos de la política fiscal en España.

El período muestral que se utilizará abarca desde principios de 1970 hasta finales del año 2006. Para este período se han elaborado series temporales de carácter trimestral de las siguientes variables que serán incluidas en los modelos econométricos planteados (cuando ha sido necesario, las series abajo citadas han sido desestacionalizadas previamente a su utilización en las distintas especificaciones):

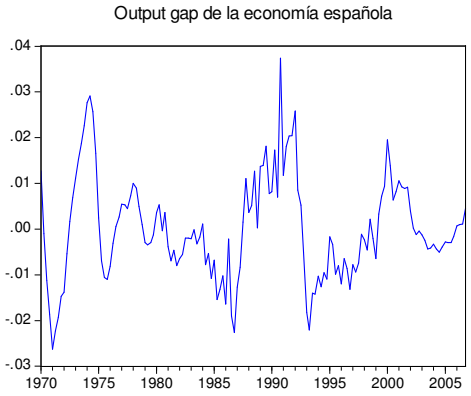
- Producto interior bruto (PIB) en términos reales, también en este caso tanto el nacional como el exterior.
- Índice de precios al consumo (IPC), tanto el interior de España como el exterior correspondiente al entorno europeo de nuestro país.
- Posición fiscal (FP) de la economía, que viene dada por el déficit público total en términos del PIB nacional, definido como la diferencia entre el total de recursos (ingresos) y el total de empleos (gastos) de las administraciones públicas de España, ambos relativos al PIB español.
- Tipos de interés nominales a corto (IS) y largo (IL) plazo, definidos como las tasas de interés a tres meses y los rendimientos de los bonos a diez años españoles.
- Precio del petróleo (POIL), definido como el valor promedio trimestral del crudo tipo Brent en los mercados internacionales.

A) <u>El PIB interno y externo, y la brecha productiva (*output gap*) de ambos</u>

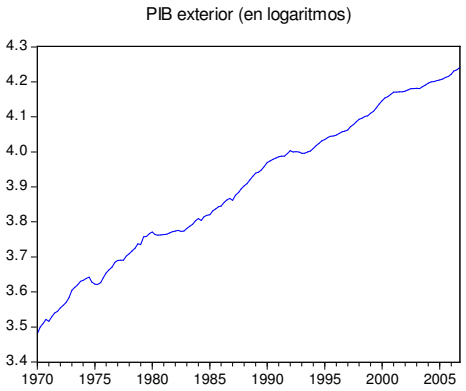
En las dos figuras siguientes se muestran el nivel y las primeras diferencias del logaritmo del PIB real español sobre el período muestral considerado, 1970.1-2006.4.

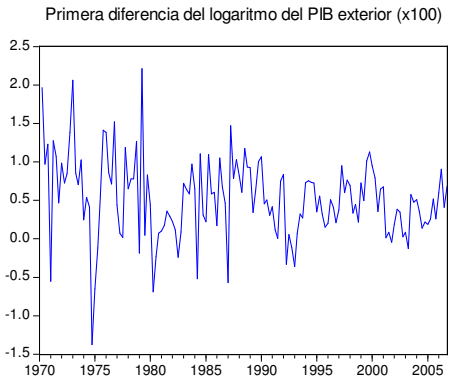


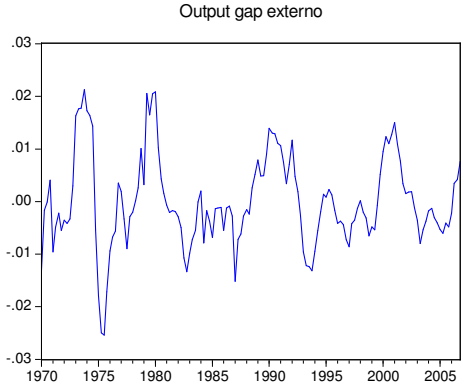
A partir de la variable original en logaritmos, $y_t = \log(PIB_t)$, se ha construido el output gap correspondiente, el cual se obtiene como la diferencia entre la producción real y la potencial, ambas en logaritmos. Ésta última se aproxima en este trabajo como el filtro de Hodrick-Prescott (con $\lambda = 1600$) de la serie $y_t = \log(PIB_t)$. En la gráfica siguiente se presenta el output gap para la economía española, que denotaremos por $\tilde{y}_t$.



Respecto a la producción externa, y_t^*, se ha obtenido como una media ponderada del PIB real (en logaritmos) de las cinco mayores economías europeas (sin incluir a España): Alemania, Francia, Holanda, Italia y Reino Unido. Como ponderaciones se han utilizado las participaciones de cada país en el volumen total de comercio (exportaciones más importaciones) que España ha mantenido con estos países durante el período 1980-2006. En las tres gráficas siguientes se presentan el output externo, y_t^*, su primera diferencia, así como el output gap externo, obtenido igual que en el caso español como la diferencia entre el PIB real y su filtro de Hodrick-Prescott, $\tilde{y}_t^* = y_t^* - y_t^{0*}$.

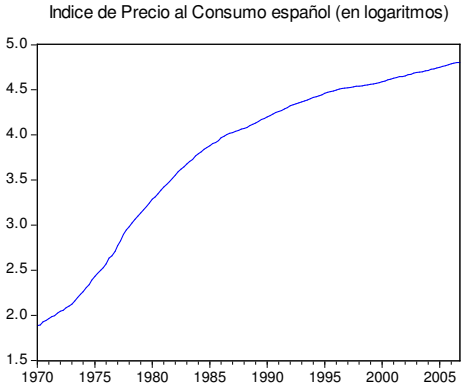


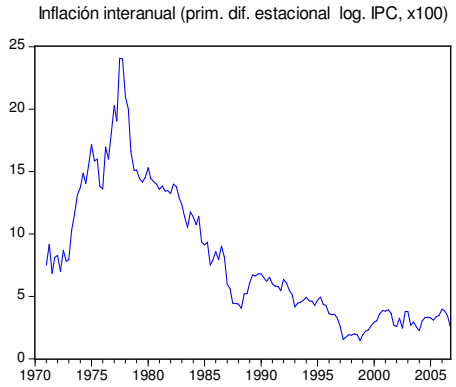


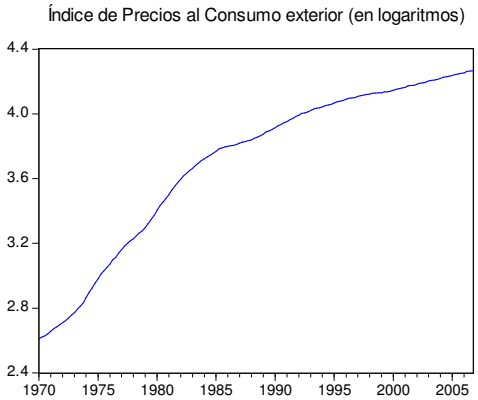


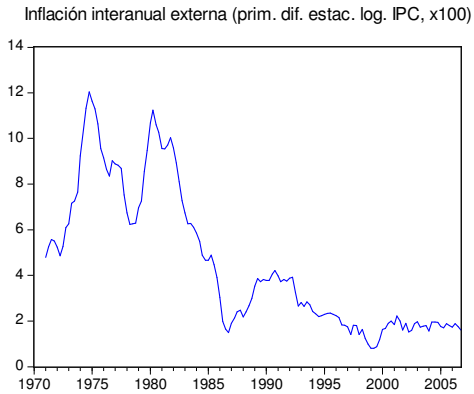
B) Los precios nacionales y externos, y los procesos inflacionarios correspondientes

En las cuatro figuras siguientes se muestran el nivel y las primeras diferencias estacionales del logaritmo (es decir, aproximadamente la inflación interanual) de los índices de precios español y externo sobre el período muestral.



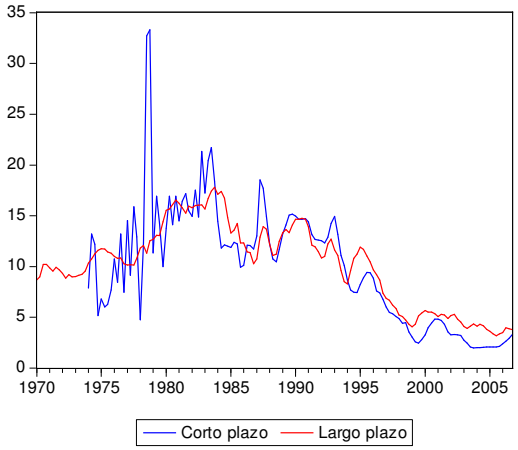






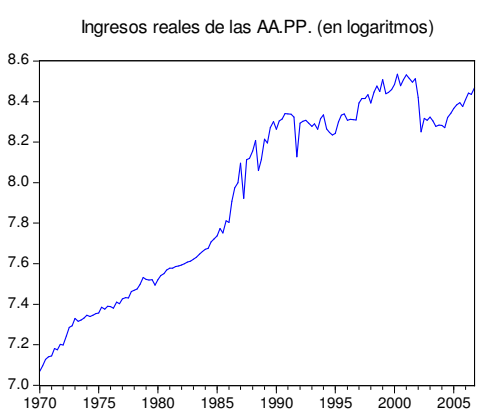
C) Los tipos de interés

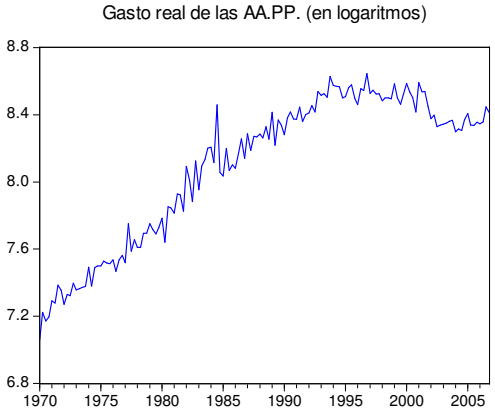
La figura siguiente presenta la evolución de los tipos de interés a corto y largo plazo de la economía española para el período temporal analizado.



D) El comportamiento del sector público español

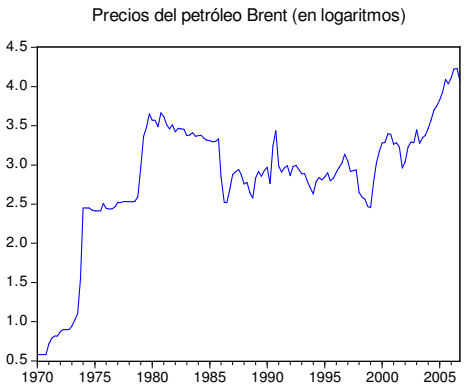
En lo que respecta a la evolución de las cuentas públicas de la economía española, en las dos gráficas siguientes se muestra la evolución de los ingresos y gastos de las administraciones públicas.

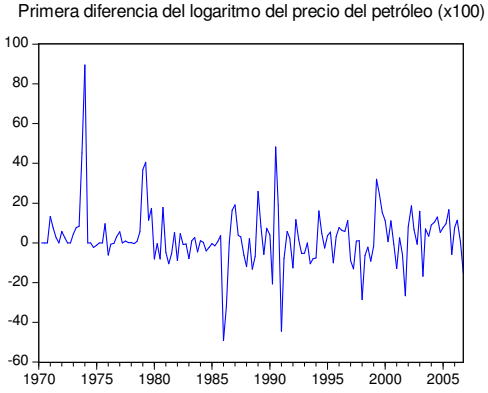




E) El precio del petróleo

Finalmente, se presentan las gráficas correspondientes a los precios del petróleo, variable que puede tener una influencia importante sobre la evolución de algunos agregados macroeconómicos, tanto nacionales como internacionales.





5. ESTIMACIÓN DE LOS MODELOS SVAR PROPUESTOS Y ANÁLISIS DE RESULTADOS

A) Especificación de un modelo SVAR con restricciones de corto plazo y cuentas públicas agregadas

El modelo econométrico que inicialmente se ha estimado, tras aplicar los criterios de información de Akaike, FPE, Hannan-Quinn y Schwarz para decidir el número de retardos en las variables endógenas (p), y suponiendo que sólo se produce un efecto contemporáneo de las variables externas (q=0), ha sido el siguiente:

$$x_t = a_0 + A_1 x_{t-1} + A_2 x_{t-2} + A_3 x_{t-3} + A_4 x_{t-4} + B_0^{*} x_t^{*} + C x_t^{**} + e_t$$

donde el vector x_t^{**} está formado en nuestro caso por el índice de precios del petróleo y por una variable tendencial lineal. Los resultados se muestran en la tabla siguiente:

Tabla 1a. Modelo SVAR con restricciones de corto plazo: parte VAR del modelo

VAR ESTIMATION RESULTS

endogenous variables (x_t): YGAP PI4 FP IS

exogenous variables (x_t^*): YSGAP PIS4

deterministic variables (x_t^{**}): TREND POIL

sample range: [1975 Q1, 2006 Q4], T = 128

Lagged endogenous term:

	YGAP	PI4	FP	IS
CONST	**1.693**	**3.346**	0.733	1.018
	{0.013}	{0.001}	{0.268}	{0.199}
YGAP(t-1)	**0.393**	-0.077	0.127	0.124
	{0.000}	{0.590}	{0.170}	{0.262}
PI4 (t-1)	-0.015	**0.791**	0.037	-0.104
	{0.794}	{0.000}	{0.506}	{0.119}
FP (t-1)	0.013	-0.205	**0.694**	-0.077
	{0.891}	{0.136}	{0.000}	{0.472}
IS(t-1)	0.030	-0.116	-0.097	**0.524**
	{0.719}	{0.344}	{0.225}	{0.000}
YGAP(t-2)	**0.315**	0.076	-0.044	0.097
	{0.001}	{0.606}	{0.644}	{0.395}
PI4 (t-2)	0.067	0.100	0.013	-0.074
	{0.359}	{0.360}	{0.855}	{0.384}
FP (t-2)	**-0.211**	-0.049	**0.396**	0.134
	{0.066}	{0.774}	{0.000}	{0.315}
IS(t-2)	-0.034	**0.262**	-0.111	-0.168
	{0.708}	{0.052}	{0.206}	{0.110}
YGAP(t-3)	0.030	-0.075	-0.006	-0.024
	{0.753}	{0.604}	{0.952}	{0.832}
PI4 (t-3)	-0.046	**0.293**	-0.084	**0.224**
	{0.509}	{0.005}	{0.220}	{0.006}
FP (t-3)	**0.189**	0.081	**0.195**	-0.080
	{0.101}	{0.638}	{0.081}	{0.551}
IS(t-3)	0.002	**-0.246**	-0.065	0.117
	{0.980}	{0.072}	{0.465}	{0.269}
YGAP(t-4)	**-0.231**	**-0.202**	0.035	-0.049
	{0.006}	{0.105}	{0.669}	{0.615}
PI4 (t-4)	-0.058	**-0.402**	0.032	0.001
	{0.315}	{0.000}	{0.567}	{0.992}
FP (t-4)	0.124	**0.229**	**-0.417**	-0.011
	{0.186}	{0.102}	{0.000}	{0.923}
IS(t-4)	-0.049	-0.098	-0.067	0.141
	{0.555}	{0.428}	{0.399}	{0.139}

Current and lagged weakly exogenous terms:

	YGAP	PI4	FP	IS
YSGAP(t)	**0.284**	**0.302**	0.047	0.015
	{0.001}	{0.018}	{0.567}	{0.878}
PIS4 (t)	**-0.105**	**0.180**	-0.060	-0.075
	{0.066}	{0.034}	{0.275}	{0.253}

Deterministic and strictly exogenous terms:

	YGAP	PI4	FP	IS
TREND(t)	**-0.020**	**-0.021**	-0.009	-0.009
	{0.014}	{0.076}	{0.270}	{0.324}
POIL (t)	0.003	-0.001	0.002	0.001
	{0.126}	{0.757}	{0.401}	{0.673}

NOTA: En cada columna, debajo en el coeficiente estimado, se muestran los p-valores de cada coeficiente. Aparecen marcados en negrita aquellos coeficientes estadísticamente significativos a un nivel de confianza superior al 90% (p-valores<0.10).

En cuanto a la ecuación $Ae_t = Du_t$ que relaciona los errores canónicos con los estructurales, la identificación de estos últimos se ha llevado a cabo a través de la imposición de una descomposición recursiva tipo Choleski de la matriz de covarianzas de los shocks canónicos. Es decir, en la ecuación $Ae_t = Du_t$ se ha supuesto que

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a_{21} & 1 & 0 & 0 \\ a_{31} & a_{32} & 1 & 0 \\ a_{41} & a_{42} & a_{43} & 1 \end{bmatrix} \qquad D = \begin{bmatrix} d_{11} & 0 & 0 & 0 \\ 0 & d_{22} & 0 & 0 \\ 0 & 0 & d_{33} & 0 \\ 0 & 0 & 0 & d_{44} \end{bmatrix}$$

lo cual implica las siguientes condiciones sobre las respuestas a corto plazo entre los distintos shocks: las innovaciones de política económica, tanto fiscales y monetarias, se ven afectadas contemporáneamente por los shocks de oferta y de demanda, pero en el mismo trimestre no existe relación en sentido contrario; el indicador de política fiscal puede afectar contemporáneamente al tipo de interés pero no existe respuesta fiscal inmediata ante un cambio en la política monetaria; finalmente, los shocks de demanda no afectan en el mismo trimestre al gap productivo, aunque los shocks de oferta sí que pueden tener un efecto inmediato sobre la inflación.

Los resultados obtenidos bajo esta descomposición han sido los siguientes:

Tabla 1b. Modelo SVAR con restricciones de corto plazo: parte SVAR del modelo

Structural VAR Estimation Results

AB-model (Amisano and Giannini ,1997)
Structural VAR is just identified

Estimated *A* matrix:

1.0000	0.0000	0.0000	0.0000
0.0508	1.0000	0.0000	0.0000
-0.1275	0.0015	1.0000	0.0000
-0.1628	0.0875	0.2006	1.0000

Estimated standard errors for *A* matrix:

0.0000	0.0000	0.0000	0.0000
0.1319	0.0000	0.0000	0.0000
0.0850	0.0569	0.0000	0.0000
0.1004	0.0667	0.1035	0.0000

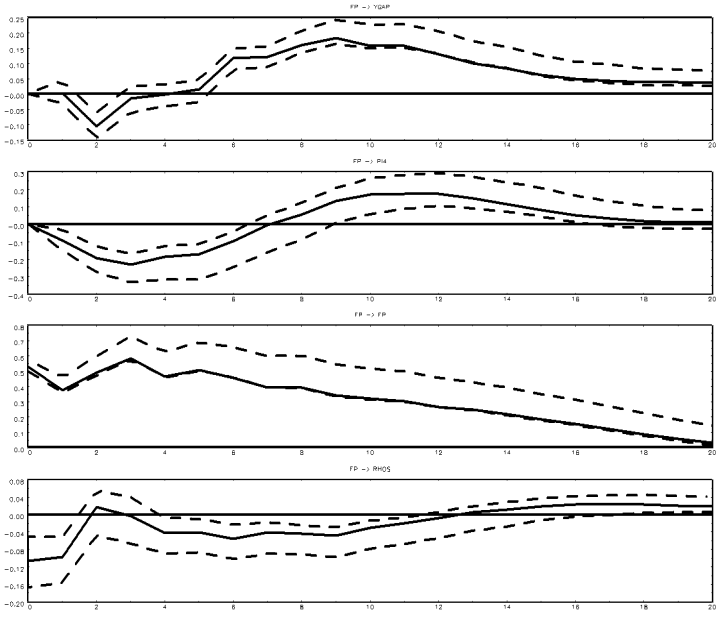
Estimated *D* matrix:

0.5514	0.0000	0.0000	0.0000
0.0000	0.8227	0.0000	0.0000
0.0000	0.0000	0.5299	0.0000
0.0000	0.0000	0.0000	0.6206

Estimated standard errors for *D* matrix

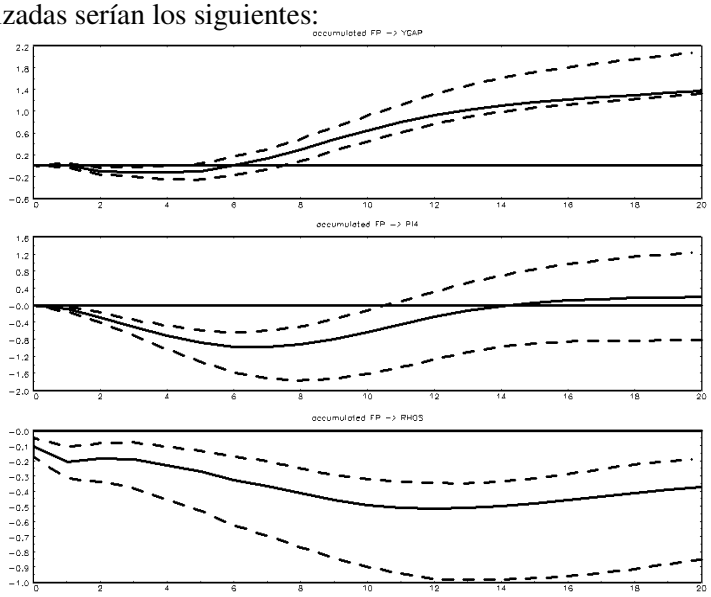
0.0345	0.0000	0.0000	0.0000
0.0000	0.0514	0.0000	0.0000
0.0000	0.0000	0.0331	0.0000
0.0000	0.0000	0.0000	0.0388

En la gráfica siguiente se muestran las funciones de respuesta al impulso estimadas para un shock de una desviación estándar en la variable fiscal (fp_t) del vector x_t (tal como se ha definido esta variable, se trata de un shock fiscal restrictivo, que puede deberse a un aumento de los impuestos o a una reducción del gasto público). Junto a la línea de respuestas, se presenta una banda de confianza del 68% construida a partir de los percentiles correspondientes para cada período, los cuales se han calculado a partir de un método de *bootstraping* usando 1000 simulaciones de Monte Carlo en cada caso.



Por un lado, los shocks fiscales restrictivos parecen tener el efecto negativo Keynesiano esperado a corto plazo, aunque se observa un efecto expansivo sobre la producción en el medio plazo.[3] Igual que para el output gap, en el caso de la inflación se observa un impacto negativo inicialmente, siendo ahora de mayor duración (cerca de dos años), aunque el efecto cambia de signo y se vuelve positivo en el medio-largo plazo. Finalmente, ante un shock fiscal restrictivo, los tipos de interés tienden a bajar en el corto-medio plazo, aunque en el largo plazo el efecto acaba convirtiéndose en positivo, pero marginalmente significativo en términos estadísticos.

En términos acumulados, los efectos del shock fiscal sobre las tres variables macroeconómicas analizadas serían los siguientes:



Como puede apreciarse, el shock fiscal restrictivo tiene un efecto acumulado negativo sobre el output gap hasta el sexto trimestre, comenzando a partir de ese período a mostrar un efecto positivo y estadísticamente significativo. Respecto al efecto sobre la inflación, el efecto acumulado negativo se muestra significativo hasta el décimo trimestre (dos años y medio), perdiendo relevancia a partir de dicha fecha. En último lugar, puede apreciarse que el efecto acumulado sobre los tipos de interés a corto plazo es siempre estadísticamente significativo y mantiene su significación estadística a lo largo de todo el período simulado.

B) Especificación de un modelo SVAR con restricciones de largo plazo y con desagregación de las cuentas públicas

En este apartado se presenta un modelo econométrico que complementa la especificación SVAR presentada en los epígrafes anteriores. Concretamente, el vector de variables macroeconómicas endógenas estará formado ahora por las variables $(y_t, t_t, g_t, i_t, p_t)$, es decir, por el logaritmo del PIB real, los ingresos totales y los gastos totales reales de las AA.PP (ambos en logaritmos), el tipo de interés a corto plazo (en niveles) y el logaritmo del índice de precios al consumo.

Por otro lado, el vector de variables débilmente exógenas (x_t^*) estará formado por las variables externas y_t^* y p_t^*, y el vector de variables determinísticas y estrictamente exógenas (x_t^{**}) por una tendencia lineal y por los precios del petróleo.

El modelo con el que se trabajará en este caso será formulado en primeras diferencias de las variables endógenas y exógenas (salvo la tendencia lineal), con el objetivo de garantizar que se cumpla (puede que de forma aproximada en algún caso) la hipótesis de estacionariedad de las variables del modelo. Entonces, la expresión formal del modelo VAR estimados es la siguiente:

$$\Delta x_t = a_0 + A_1 \Delta x_{t-1} + A_2 \Delta x_{t-2} + A_3 \Delta x_{t-3} + A_4 \Delta x_{t-4} + B_0^* \Delta x_t^* + C \Delta x_t^{**} + e_t$$

[3] Véanse, al respecto, los trabajos de Giavazzi y Pagano (1996), Giavazzi *et al.* (2000), Afonso (2001) o van Aarle y Garretsen (2001).

En consecuencia, las variables endógenas a explicar serán el crecimiento intertrimestral del PIB real ($\Delta \log PIB_t$), los crecimientos intertrimestrales de los ingresos y gastos reales de las AA.PP. ($\Delta \log T_t$ y $\Delta \log G_t$, respectivamente), el cambio en el tipo de interés nominal a corto plazo (Δi_t) y la tasa de inflación intertrimestral ($\Delta \log P_t$).

Respecto al esquema de identificación utilizado, ahora se parte de una especificación del tipo $e_t = Eu_t$, donde $E\left(= A^{-1}D\right)$ representa la matriz de impacto contemporánea, sobre la que es necesario imponer de nuevo cierto número de restricciones para conseguir la identificación de las innovaciones estructurales u_t a partir de las estimaciones de los errores de la forma reducida e_t del modelo VAR. En este caso, el vector u_t está constituido por cinco shocks estructurales: un shock de oferta (u_t^s), un shock de ingresos públicos (u_t^t), un shock de gasto público (u_t^g), un shock monetario (u_t^m) y un shock de demanda agregada (u_t^d).

Las restricciones en nuestro caso vienen dadas por la hipótesis de efectos a largo plazo nulos para algunos shocks. Concretamente, y de forma análoga a como se ha hecho en el caso de restricciones a corto plazo del primer modelo SVAR estimado, ahora se utilizan las siguientes restricciones de identificación: los shocks fiscales, ya sean de ingresos o gastos, no tienen efectos a largo plazo sobre la producción; los shocks de gasto público no tienen efectos permanentes sobre los ingresos públicos; las innovaciones de política monetaria no tienen efectos permanentes sobre el output real, los ingresos o el gasto público; finalmente, los shocks de demanda no tienen efectos duraderos ni sobre el output ni sobre las variables de política económica, ya sean las fiscales o la monetaria.

En términos formales, estas restricciones cero a priori sobre los efectos a largo plazo señaladas implican la siguiente estructura para la matriz de impactos a largo plazo:

$$\Psi = \begin{bmatrix} \psi_{11} & 0 & 0 & 0 & 0 \\ \psi_{21} & \psi_{22} & 0 & 0 & 0 \\ \psi_{31} & \psi_{32} & \psi_{33} & 0 & 0 \\ \psi_{41} & \psi_{42} & \psi_{43} & \psi_{44} & 0 \\ \psi_{51} & \psi_{52} & \psi_{53} & \psi_{54} & \psi_{55} \end{bmatrix}$$

las cuales pueden derivarse de modelos macroeconómicos teóricos similares (o más generales) al expuesto en el apartado A de la parte II (ver, por ejemplo, los trabajos clásicos de Blanchard (1989), Blanchard y Quah (1989), Galí (1992, 1999), King et al. (1991) o los más recientes de Francis y Ramey (2004, 2005)).

Los resultados obtenidos al estimar el modelo VAR para el vector $(\Delta y_t, \Delta t_t, \Delta g_t, \Delta i_t, \Delta p_t)$, junto con las estimaciones de las matrices de impacto E y de largo plazo Ψ del esquema de identificación propuesto se presentan en las tablas siguientes:

Tabla 2a. Modelo SVAR con restricciones de largo plazo: parte VAR del modelo

VAR ESTIMATION RESULTS

endogenous variables (x_t): Y_d1 T_d1 G_d1 IS_d1 P_d1

exogenous variables (x_t^*): YSTAR_d1 PSTAR_d1

deterministic variables (x_t^{**}): TREND POIL_d1

sample range: [1975 Q2, 2006 Q4], T = 127

Lagged endogenous terms:

	Y_d1	T_d1	G_d1	IS_d1	P_d1
CONST	0.857	**6.119**	3.935	-0.733	0.336
	{0.219}	{0.075}	{0.495}	{0.164}	{0.497}
Y_d1 (t-1)	-0.121	0.112	-1.143	0.023	-0.001
	{0.196}	{0.809}	{0.141}	{0.750}	{0.990}
T_d1 (t-1)	0.002	**-0.304**	**0.259**	0.002	-0.006
	{0.924}	{0.002}	{0.108}	{0.872}	{0.654}
G_d1 (t-1)	-0.007	0.048	**-0.831**	-0.001	-0.004
	{0.543}	{0.432}	{0.000}	{0.879}	{0.681}
IS_d1(t-1)	-0.165	0.889	0.784	**-0.305**	0.033
	{0.193}	{0.155}	{0.455}	{0.001}	{0.717}
P_d1 (t-1)	0.023	0.083	0.744	0.045	0.294
	{0.865}	{0.900}	{0.502}	{0.659}	{0.002}
Y_d1 (t-2)	0.143	0.705	0.605	0.057	0.017
	{0.125}	{0.125}	{0.433}	{0.418}	{0.799}
T_d1 (t-2)	0.013	-0.124	0.096	0.001	-0.008
	{0.537}	{0.227}	{0.580}	{0.929}	{0.600}
G_d1 (t-2)	-0.005	0.086	**-0.489**	-0.017	-0.003
	{0.749}	{0.285}	{0.000}	{0.164}	{0.771}
IS_d1(t-2)	**-0.208**	0.238	0.357	**-0.352**	0.127
	{0.103}	{0.705}	{0.735}	{0.000}	{0.160}
P_d1 (t-2)	-0.024	0.328	0.843	**-0.244**	0.060
	{0.852}	{0.605}	{0.429}	{0.012}	{0.509}
Y_d1 (t-3)	0.051	**1.072**	-0.197	0.045	-0.003
	{0.578}	{0.017}	{0.795}	{0.511}	{0.965}
T_d1 (t-3)	0.008	-0.033	0.184	0.004	-0.001
	{0.707}	{0.746}	{0.277}	{0.778}	{0.959}
G_d1 (t-3)	-0.017	0.023	-0.128	-0.018	-0.008
	{0.283}	{0.775}	{0.340}	{0.146}	{0.511}
IS_d1(t-3)	-0.086	0.634	1.250	**-0.224**	-0.029
	{0.481}	{0.294}	{0.218}	{0.016}	{0.742}
P_d1 (t-3)	-0.166	-0.316	-0.867	**0.195**	**0.390**
	{0.201}	{0.621}	{0.419}	{0.047}	{0.000}
Y_d1 (t-4)	**-0.147**	**-1.527**	**-1.510**	0.008	-0.026
	{0.105}	{0.001}	{0.044}	{0.909}	{0.692}
T_d1 (t-4)	0.010	**-0.195**	-0.050	0.001	-0.001
	{0.587}	{0.034}	{0.746}	{0.951}	{0.965}
G_d1 (t-4)	-0.005	-0.005	-0.091	-0.012	0.001
	{0.658}	{0.932}	{0.364}	{0.187}	{0.927)
IS_d1(t-4)	**-0.370**	0.177	-0.300	-0.025	**0.145**
	{0.002}	{0.766}	{0.764}	{0.781}	{0.090}
P_d1 (t-4)	-0.014	-0.107	-0.434	**0.236**	**-0.165**
	{0.922}	{0.875}	{0.704}	{0.024}	{0.092}

Current and lagged weakly exogenous terms:

	Y_d1	T_d1	G_d1	IS_d1	P_d1
YSTAR_d1 (t)	0.305	-0.061	0.746	0.156	0.207
	{0.128}	{0.951}	{0.653}	{0.304}	{0.146}
PSTAR_d1 (t)	0.145	**-1.790**	-0.005	-0.110	**0.585**
	{0.483}	{0.078}	{0.998}	{0.479}	{0.000}

Deterministic and strictly exogeneous terms:

	Y_d1	T_d1	G_d1	IS_d1	P_d1
TREND(t)	0.000	**-0.046**	-0.025	0.003	-0.003
	{0.945}	{0.094}	{0.583}	{0.427}	{0.388}
POIL_d1(t)	0.005	-0.021	0.001	-0.003	-0.004
	{0.412}	{0.510}	{0.991}	{0.535}	{0.415}

NOTA: En cada columna, debajo en el coeficiente estimado, se muestran los p-valores de cada coeficiente. Aparecen marcados en negrita aquellos coeficientes estadísticamente significativos a un nivel de confianza superior al 90% (p-valores<0.10).

Tabla 2b. Modelo SVAR con restricciones de largo plazo: parte SVAR del modelo

Structural VAR Estimation Results

Identified accumulated long run impact matrix Ψ is lower diagonal
Structural VAR is just identified

Estimated *E* contemporaneous impact matrix:

0.7164	0.0125	0.0907	0.4844	0.2243
1.2329	4.1667	-0.3494	-0.7397	0.0146
-0.4985	1.9891	7.0490	-1.0676	-0.3556
-0.2188	0.2185	0.0401	0.5294	-0.2868
-0.2199	0.0021	0.0979	0.2731	0.5233

Bootstrap standard errors for *E* matrix:

0.1576	0.1347	0.1313	0.1737	0.1434
0.6468	0.8083	0.5383	0.6355	0.6678
1.0658	1.0616	1.3154	1.0696	1.1663
0.1131	0.1003	0.0907	0.1035	0.1339
0.1364	0.1243	0.1251	0.1471	0.1127

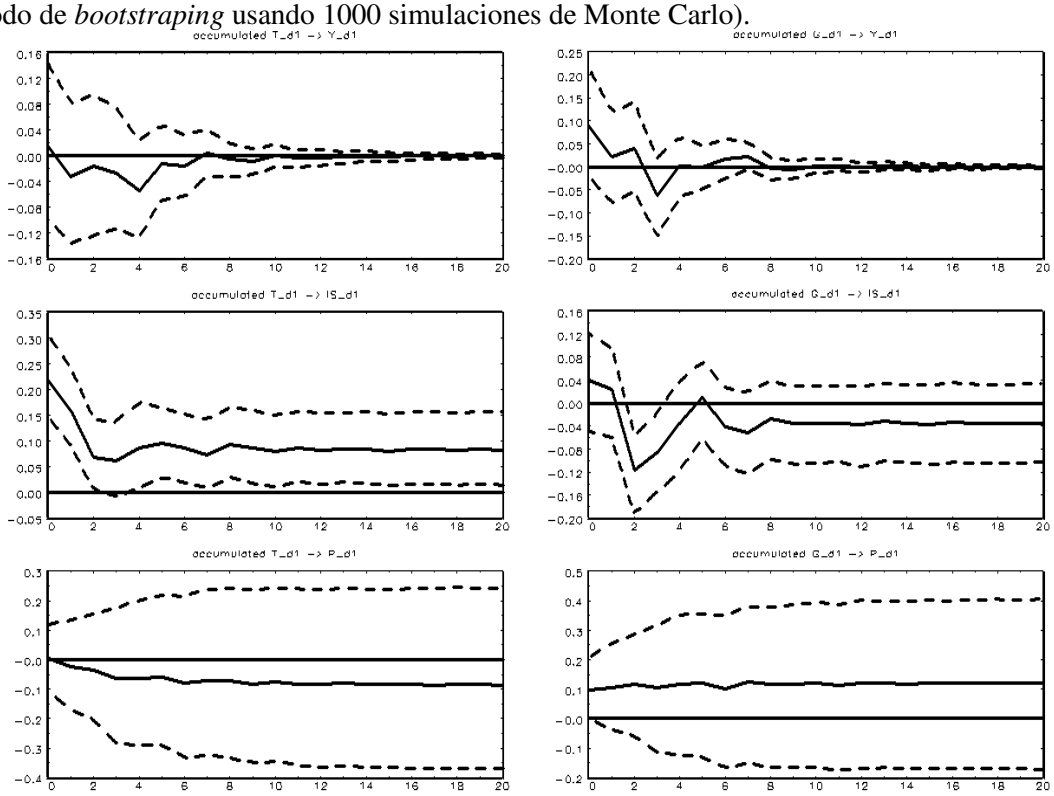
Estimated Ψ long run impact matrix:

0.8986	0.0000	0.0000	0.0000	0.0000
0.7406	2.7381	0.0000	0.0000	0.0000
-0.9964	1.3681	2.7612	0.0000	0.0000
-0.0968	0.0825	-0.0348	0.3871	0.0000
-0.6061	-0.0851	0.1192	0.9023	1.2419

Bootstrap standard errors for Ψ matrix:

0.1837	0.0000	0.0000	0.0000	0.0000
0.6273	0.5264	0.0000	0.0000	0.0000
0.7816	0.5944	0.5229	0.0000	0.0000
0.1032	0.0831	0.0779	0.0955	0.0000
0.4490	0.3536	0.3358	0.4629	0.3643

En la gráfica siguiente se muestran los efectos acumulados[4] de los shocks fiscales (una desviación estándar en las variables t_t y g_t) sobre las tres variables macroeconómicas analizadas, junto con un banda de confianza del 68% construida a partir de los percentiles correspondientes para cada período. De nuevo, estas bandas se han calculado mediante un método de *bootstraping* usando 1000 simulaciones de Monte Carlo).



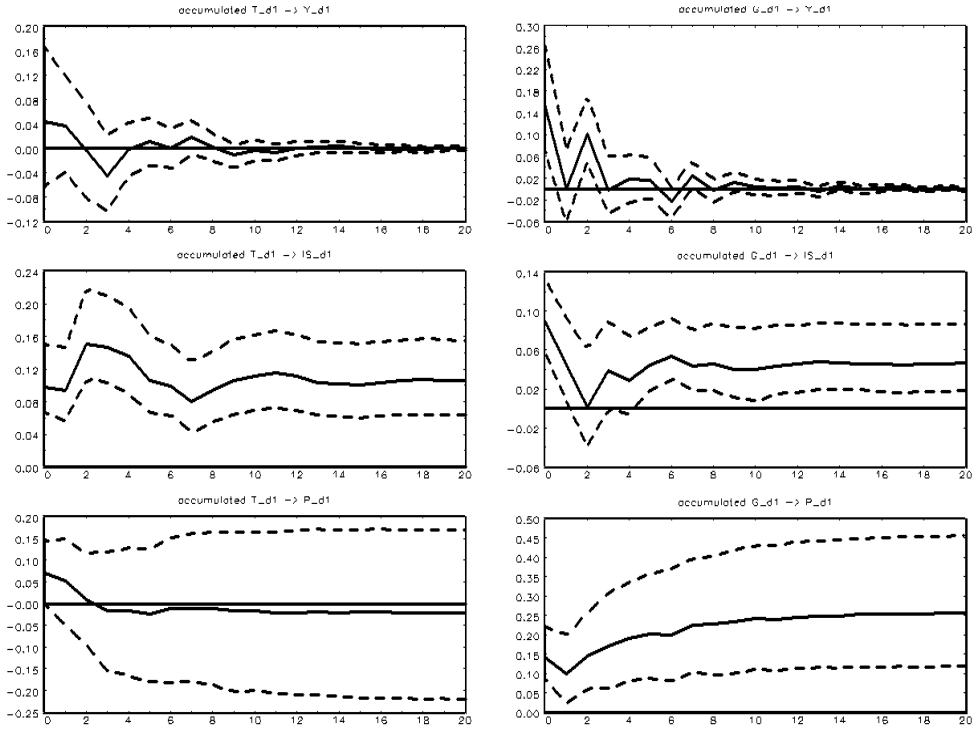
Como puede apreciarse, un incremento no anticipado en los impuestos, es decir, un shock positivo en la recaudación impositiva de las AA.PP., produce un efecto negativo sobre la evolución del PIB real, aunque el mismo no resulta estadísticamente significativo en ningún trimestre. Respecto al impacto sobre los tipos de interés nominales a corto plazo, puede apreciarse un efecto positivo y persistente en el tiempo. Por último, el efecto sobre los precios, siendo negativo, no resulta significativo en ningún trimestre del período simulado.

En lo referente al resultado de un incremento inesperado en los gastos del gobierno, pueden apreciarse unos efectos inicialmente positivos sobre el PIB, que tienden a amortiguarse con el paso del tiempo; sin embargo, dichos efectos no resultan estadísticamente significativos en ningún trimestre. En general, los tipos de interés no parecen reaccionar de forma significativa, salvo en el segundo trimestre tras producirse el impacto, donde se produce un bajada relevante. Finalmente, un mayor gasto de las administraciones públicas conlleva un aumento en el nivel de precios (lo que podría implicar una mayor inflación tras producirse el shock), aunque dicho aumento es insignificante en términos estadísticos.

Para contrastar hasta que punto estos resultados están condicionados por el período muestral utilizado, también se ha re-estimado el modelo SVAR para el período 1980-2006. Con esta sub-muestra los resultados referentes al efecto de shocks en el gasto público varían substancialmente respecto a los recién comentados, no ocurriendo lo mismo en el caso del efecto de shocks de impuestos, donde los resultados no sufren variaciones de relevancia. Así, el shock de gasto positivo tiene un efecto significativo y positivo sobre el PIB hasta los dos primeros trimestres, un impacto también positivo y persistente en el tiempo sobre los tipos de interés, y también un

[4] La acumulación de los efectos sobre las primeras diferencias de las variables endógenas equivale al cálculo de los efectos en cada período sobre dichas variables en niveles. De esta manera se pueden comparar las gráficas de los dos tipos de modelos SVAR estimados (con restricciones de corto y de largo plazo).

efecto persistentemente positivo sobre los precios. Estos efectos se muestran en la gráfica siguiente.



BIBLIOGRAFÍA

AFONSO, A. (2001), "Non-Keynesian Effects of Fiscal Policy in the EU-15", ISEG Working Paper No. 2001/07.

ARTIS, M.J. AND L. ONORANTE (2006), "The economic importance of fiscal rules", CEPR Discussion Paper No. 5684.

BADINGER, H. (2006), "Fiscal shocks, output dynamics and macroeconomic stability: an empirical assessment for Austria (1983–2002)", *Empirica*, 33, 267-284.

BARRO, R.J. (2009), "Demand Side Voodoo Economics", *The Economists' Voice*, 6-2, Article 5.

BEETSMA, R. AND H. UHLIG, 1999, "An Analysis of the Stability and Growth Pact", *Economic Journal*, 109, 546–71.

BEGG, I., HODSON, D. AND I. MAHER (2003), "Economic Policy Coordination in the European Union", *National Institute Economic Review*, 183, 66-77.

BÉNASSY-QUÉRÉ, A., CIMADOMO, J. AND V. MIGNON, 2006, "Time-varying fiscal policy spillovers in the Euro area", mimeo, CEPII.

BÉNASSY-QUÉRÉ, A. AND J. CIMADOMO (2006), "Changing Patterns of Domestic and Cross-Border Fiscal Policy Multipliers in Europe and the US", CEPII Working Paper No. 2006/24.

BÉNASSY-QUÉRÉ, A., 2006, "Short-term fiscal spillovers in a monetary union", CEPII Working Paper No. 2006/13.

BIAU, O. AND E. GIRARD (2005), "Politique budgétaire et dynamique économique en France: l'approche VAR Structurel", *Economie et Prévision*, 169/171, 1-24.

BLANCHARD, O., 1989, "A traditional interpretation of macroeconomic fluctuations", *American Economic Review*, 79, 1146-1164.

BLANCHARD, O., AND R. PEROTTI (2002), "An empirical characterization of the dynamic effects of changes in government spending and taxes on output", *Quarterly Journal of Economics*, 117 (4), 1329-1368.

BLANCHARD, O. AND D. QUAH (1989), "The dynamic effects of aggregate demand and supply disturbances", *American Economic Review*, 79, 655-673.

BRIOTTI, M.G. (2005), "Economic reactions to public finance consolidation: a survey of the literature", ECB Occasional Paper No. 38.

BRUNEAU, C., AND O. DE BANDT (2003), "Monetary and fiscal policy in the transition to EMU: what do SVAR models tell us?", *Economic Modelling*, 20, 959-985.
BRUNILA, A. (2002), "Fiscal policy: Coordination, discipline and stabilization", Bank of Finland Discussion Papers No. 7-2002.
BRUNILA, A., BUTI M., AND D. FRANCO (2001), *The Stability and Growth Pact*, Palgrave, Basingstoke.
CANZONERI, M.B., CUMBY R.E. AND B. DIBA (2002), "The need for international policy cooperation: What's old, what's new, what's yet to come?", mimeo, Georgetown University.
CANZONERI, M.B., CUMBY R.E. AND B. DIBA (2003), "New views on the transatlantic transmission of fiscal policy and macroeconomic policy coordination", in M. Buti (ed.), *Monetary and Fiscal Policies in EMU: Interactions and Coordination*, Cambridge University Press, Cambridge.
CAPET, S. (2004), "The Efficiency of Fiscal Policies: a Survey of the Literature", CEPII Working Paper No. 2004/11.
CLAEYS, P. (2004), "Monetary and Budgetary Policy Interaction: An SVAR Analysis of Stabilization Policies in Monetary Union", EUI Working Paper ECO No. 2004/22.
CREEL, J., MONPERRUS-VERONI, P. AND F. SARACENO (2005), "Discretionary Policy Interactions and the Fiscal Theory of the Price Level: A SVAR Analysis on French Data", mimeo, OFCE.
DALSGAARD, T. AND DE SERRES, A. (2000), "Estimating prudent budgetary margins for EU countries: A simulated SVAR model approach", *OECD Economic Studies*, 30(1), 115–147.
DE ARCANGELIS, G. AND S. LAMARTINA (2003), "Identifying fiscal shocks and policy regimes in OECD countries", ECB Working Paper No. 281.
DE CASTRO, F. (2006), "The macroeconomic effects of fiscal policy in Spain", *Applied Economics*, 38, 913-924.
DE CASTRO, F. AND P. HERNÁNDEZ (2006), "The economic effects of exogenous fiscal shocks in Spain: a SVAR approach", ECB Working Paper No. 647.
EDELBERG, W., EICHENBAUM, M. AND J.D.M. FISHER (1998), "Understanding the effects of a shock to government purchases", NBER Working Paper No. 6737.
EUROPEAN COMMISSION (2002a), "Coordination of economic policies in the EU: a presentation of the key features of the main procedures", DG for Economic and Financial Affairs, Euro Papers No. 45.
EUROPEAN COMMISSION (2002b), "Strengthening the coordination of budgetary policies", Communication from the Commission to Council and the European Parliament 668, November.
FATÁS, A. AND I. MIHOV (2000), "Fiscal policy and business cycles: an empirical investigation", mimeo, INSEAD.
FATÁS, A. AND I. MIHOV (2003), "On constraining Fiscal Policy Discretion in EMU", *Oxford Review of Economic Policy*, 19(1), 112-131.
FRANCIS, N. AND V.A. RAMEY (2004), "The source of historical economic fluctuations: an analysis using long-run restrictions", NBER Working Paper No. W10631.
FRANCIS, N. AND V.A. RAMEY (2005), "Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited", *Journal of Monetary Economics*, 52, 1379-1399.
GALÍ, J. (1992), "How well does the IS-LM model fit post-war U.S. data", *Quarterly Journal of Economics*, 107, 975-1009.
GALÍ, J. (1999), "Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations?", *American Economic Review*, 89, 249-271.
GALÍ, J. (2004), "On the role of technology shocks as a source of business cycles: some new evidence", *Journal of the European Economic Association*, 2, 372-380.
GALÍ, J. AND T. MONACELLI (2005), "Monetary policy and exchange rate volatility in a small open economy", Review of Economic Studies, 72, 707-734.
GALÍ, J. AND R. PEROTTI (2003), "Fiscal policy and monetary integration in Europe", *Economic Policy*, October 2003, 533-572.
GARCIA S. AND A. VERDELHAN (2001), "Le policy-mix de la zone euro : une évaluation de l.impact des chocs monétaires et budgétaires", *Économie et Prévision*, 148, 23-40.
GIAVAZZI, F., AND M. PAGANO (1996), "Non-Keynesian Effects of Fiscal Policy Changes: International Evidence and the Swedish Experience", *Swedish Economic Policy Review*, 3 (1), 67-103.
GIAVAZZI, F., JAPPELLI, T. AND M. PAGANO (2000), "Searching for Nonlinear Effects of Fiscal Policy", *European Economic Review*, 44, 1259-1289.
GIORDANO R., MOMIGLIANO, S., NERI, S. AND R. PEROTTI (2007), "The effects of fiscal policy in Italy: Evidence from a SVAR Model", *European Journal of Political Economy*, 23, 707-733.
GIULIODORI, M. AND R. BEETSMA (2005), "What are the trade spillovers from fiscal shocks in Europe? An empirical analysis", *De Economist*, 153, 167–97.

GÓMEZ, V. AND A. MARAVALL (1997), "Programs TRAMO and SEATS. Instructions for the users", mimeo, MEH-BE.

HEMMING, R., KELL, M., MAHFOUZ, S. (2002), "The effectiveness of fiscal policy in stimulating economic activity: A review of the literature", IMF Working Papers No. 02/208.

HEPPKE-FALK, K.H., TENHOFEN, J. AND G.B. WOLFF (2006), "The macroeconomic effects of exogenous fiscal policy shocks in Germany: a disaggregated SVAR analysis", Deutsche Bundesbank Discussion Paper (Series 1: Economic Studies) No. 41/2006.

HÖPPNER, F. (2001), "A VAR Analysis of the Effects of Fiscal Policy in Germany", mimeo, IIE.

JACQUET, P. AND J. PISANI-FERRY (2001), "Economic policy co-ordination in the Eurozone: What has been achieved? What should be done?", mimeo, CER.

MARCELLINO, M. (2006), "Some stylized facts on non-systematic fiscal policy in the Euro area", *Journal of Macroeconomics*, 28, 461-479.

MCKIBBIN, W.J. (1997), "Empirical Evidence on International Economic Policy Coordination", en *Handbook of Macroeconomic Policy in Open Economies*, Freenwood Press, Chapter 6.

MOUNTFORD, A. AND UHLIG, H. (2002), "What are the effects of fiscal policy shocks?", CEPR Discussion Paper No. 3338.

MUSCATELLI, A., P. TIRELLI AND C. TRECROCI (2002), "Monetary and fiscal policy interactions over the cycle: some empirical evidence", CESifo Working Paper No. 817.

PEROTTI, R. (2005), "Estimating the effects of fiscal policy in OECD countries", CEPR Discussion Paper No. 4842.

PEROTTI, R. (2006), "Comparing alternative methodologies to estimate the effects of fiscal policy", mimeo, IGIER-Bocconi University.

RAMAJO, J. (2008), "Asimetrías y efectos desbordamiento en la transmisión de la política fiscal en la unión europea: evidencia a partir de un enfoque VAR estructural", Papel de Trabajo No. 18/08, Instituto de Estudios Fiscales.

SCHALCK, C. (2007), "Effects of Fiscal Policies in Four European Countries: A Non-linear Structural VAR Approach", *Economics Bulletin*, 5 (22), 1-7.

SPILIMBERGO, A., SYMANSKY, S., BLANCHARD, O., AND C. COTTARELLI (2008), "Fiscal policy for the crisis", IMF Staff Position Note, SPN/08/01, International Monetary Fund.

UHLIG, H. (2002), "One money but many fiscal policies in Europe: what are the consequences?", CentER Discussion Paper No. 2002-32.

VAN AARLE, B., AND GARRETSEN, H. (2001), "Keynesian, non-Keynesian or no effects of fiscal policy changes? The EMU case", CESifo Working Paper No. 570.

VAN AARLE, B., GARRETSEN H. AND N. GOBBIN (2003), "Monetary and fiscal policy transmission in the Euro-area: evidence from a structural VAR analysis", *Journal of Economics and Business*, 55, 609–38.

CONVERGENCE IN THE INNOVATIVE PERFORMANCE OF THE EUROPEAN UNION COUNTRIES

AMAIA ALTUZARRA
E-mail: amaia.altuzarra@ehu.es
LEIRE URKOLA
E-mail: leire.urcola@ehu.es
Applied Economics Department
University of the Basque Country

Resumen

The aim of this work is to study whether a process of convergence in innovation in the European countries has occurred or not, by analysing the relative changes in the innovative position of the European countries during period 2002-2006. A special attention is paid to the relative changes that have occurred in the recently acceded countries and to the poorest countries of the UE-15. The data used in this research come from the fourth and fifth wave of the Community Innovation Survey (CIS). The statistical technique used is the Multiple Factorial Analysis. The study has been carried out in two different sets of firms. On the one hand, it has considered all the manufacturing firms (innovative and non-innovative firms) and on the other, only the innovative manufacturing firms. The results show a moderate convergence process in innovation in the manufacturing as a whole, but not in the innovative firms.

Palabras clave: Convergence, CIS, innovation.
Área Temática: Economía Europea.

1. INTRODUCTION

One of the aims of the European Union since the beginning of the process of economic integration has been to reach greater economic convergence between the European countries. This objective was reached until the 70s (Mud, 1991; Fagerberg and Verspagen, 1996), stopping in the 80s and, undergoing an improvement since the 90s (Iammarino and Santangelo, 2001).

The interest in understanding factors that influence the economic convergence in the European countries has given rise to two main research strands. The first one has focused on the study of the factors that affect, in general, the economic convergence. The second research strand has dealt with the analysis of the economic effects that have occurred as a result of the last two enlargements (2004 and 2007) of the European Union.

Within the first research strand, one aspect that has recently received increased attention in academic and policy-making circles is the importance of the innovation in explaining the existing economic differences among countries (Jungmittag, 2006). The existing consensus about the relevance of technical progress and innovation for the economic growth (Aghion and Howitt, 1997) has led to many authors to suggest that the innovative dynamics of countries provide relevant inputs to understand their economic dynamics. Hence, the existence of convergence in the innovative profile of countries would lead one to believe that they advance towards convergence in terms of income.

From a theoretical standpoint, the study of technological and innovative differences across countries has gained attention since the 80s (Fagerberg and Verspagen, 1999), giving rise to the development of two main theoretical approaches. The first one assumes that technological differences across countries open the possibility for countries with low economic and

technological profile to catch up the rest of the countries by means of imitating more productive technologies applied in leader countries (Fagerberg, 1987). In this strand of thought, imitation is seen as a process less expensive than innovation. Technology is not considered as a public good in the sense that it is not available for all the agents at zero cost . The adoption of new technology entails costs and requires the country to have innovation capacity. The second approach considers that the accumulation of knowledge is basically endogenous (Kaldor, 1966). More precisely, it maintains that countries follow their own growth path, generating his own technological capacities, and consequently, the existing margin for convergence is rather limited.

The second research strand as it has been stated, has dealt with the study about the economic effects of the European enlargement on the countries of the previous EU-15, in general, and on the poorest countries like Spain and Portugal,in particular (Monetary Policy and the Economy, 2008; European Commission, 2006, Kutan and Yigit, 2007, Mountain Altuzarra, and, 2008). From this perspective, one of the subjects that has received increased attention is the empirical study of the effects of FDI on both the technological development and the real convergence of the countries. One stylized fact is the important increase of FDI flows that has taken place since the 90s, especially from the EU-15 countries towards the enlargement countries. The increase of FDI flows in the latter countries has been promoted by both the elimination of barriers to FDI inflows and the acceleration of the transition process towards market economies. There exist, however, great differences in the volume of FDI inflows that have received countries. Most of the FDI inflows have gone into Poland, the Czech Republic, Estonia and Hungary which are the largest countries of the last enlargements and that exhibited better initial economic condition (Bevan and Estrin, 2000). These countries are followed, to certain distance, by Slovakia and Slovenia that have also advanced successfully towards the single market. In Romania, the progress towards a market economy has been slower, which has made difficult the FDI inflows until recent dates. In the last years, however, this country has undergone an important increase of FDI inflows. Table 1 shows the FDI stock per capita and per GDP for years 2002, 2006 and 2007.

Table 1. FDI Stock

	Stock per capita			Stock per GDP		
	2002	2006	2007	2002	2006	2007
Czech Rep	3.790	7.836	9,923	51.40	56.50	58,40
Estonia	3.115	9.451	12,427	57.80	78.50	80,30
Hungary	3.564	8.111	9,711	54.30	72.90	71,10
Lithuania	1.148	3.226	4,33	28.20	37.60	38,90
Poland	1.260	3.265	3,732	24.40	36.60	34.00
Portugal	4.314	8.315	10,749	35.00	45.90	52.00
Romania	0.358	2.111	2,842	17.10	37.10	36,80
Slovakia	1.583	4.446	23,2	34.80	68.40	54.10
Spain	6.212	10.049	12,138	37.50	36.00	37,60

Source: UNCTAD(2008), Major indicators. Own elaboration.

The intense FDI growth in the enlargement countries could have consequences for countries such as Spain and Portugal that have been receiving important amounts of foreign capital coming from more developed European countries. Undoubtedly, the latter countries have been losing relative positions as a destination of FDI inflows in favour of the newly acceded countries, that could put in danger the advance of their innovating activity and their economic development (Redness and Damijan, 2008). Table 2 presents the evolution of the FDI inflows as percentage of the GIP. It can be noted that FDI inflows have been greater in enlargement countries than in Spain and Portugal.

Tabla 2. FDI inflows (% GDP)

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	Average of the period
Czech Rep	1,80	2,00	4,60	2,30	2,30	6,00	10,50	8,80	9,10	11,30	2,30	4,60	9,40	4,30	5,30	5,64
Estonia	3,90	5,10	4,50	3,20	5,30	10,20	5,30	6,90	8,80	3,90	9,60	8,30	20,90	10,40	12,00	7,89
Hungary	6,20	2,70	11,10	7,10	8,90	6,90	6,70	5,80	7,40	4,50	2,50	4,40	7,00	6,10	4,10	6,09
Lithuania	0,50	0,50	1,10	1,90	3,50	8,30	4,50	3,30	3,70	5,20	1,00	3,40	4,00	6,30	5,10	3,49
Poland	1,80	1,70	2,60	2,90	3,10	3,70	4,30	5,50	3,00	2,10	2,10	5,20	3,40	5,60	4,20	3,41
Portugal	1,70	1,30	0,60	1,10	2,10	2,50	1,00	5,90	5,40	1,40	4,60	1,10	2,10	5,90	2,60	2,62
Romania	0,40	1,10	1,20	0,70	3,40	4,80	2,90	2,90	2,90	2,50	3,70	8,50	6,60	9,30	5,90	3,79
Slovakia	1,30	1,60	13,10	1,70	1,10	3,20	2,10	9,40	7,50	16,80	6,50	7,20	4,40	7,40	4,30	5,84
Spain	1,90	1,80	1,40	1,50	1,60	2,40	3,00	6,80	4,70	5,70	2,90	2,40	2,20	2,20	3,70	2,95
UE_7	2,4	2,21	5,56	2,94	4,19	6,30	5,26	6,19	6,33	7,67	4,11	6,30	8,20	7,30	6,30	5,42
UE_2	1,8	1,55	1,00	1,30	1,85	2,45	2,00	6,35	5,05	3,55	3,75	1,75	2,15	4,05	3,15	2,78

Source: UNCTAD(2008), Major indicators. Own elaboration

UE_2 includes Portugal and Spain

UE_7 includes Czech Rep., Estonia Hungary, Lithuania, Poland Romania and Slovakia

The aim of this work is to study whether a process of convergence in innovation in the European countries has occurred or not, by analysing the relative changes in the innovative position of the European countries during period 2002-2006. Special attention is paid to the relative changes that have occurred in the recently acceded countries and to the poorest countries of the previous UE-15. Section two describes the methodology and the data used in this study. Section three contains the results and the last section contains the main conclusions drawn from the study.

2. METHODOLOGY, DATA AND VARIABLES

2.1. Methodology and data

The statistical technique used in this study is the Multiple Factorial Analysis (Escofier and Pagès, 1994). This method is a factorial technique adapted to the treatment of data tables where the same set of individuals (countries, in our case) is described through different groups of variables.

The AFM provides results of great interest from the interpretation standpoint. The interpretative wealth reaches its maximum on the factorial planes, where the individuals are represented and seen through each table or group of variables. These factorial planes allow one to detect the relative position of the individuals considering the whole set of variables. Also, and being this the dimension where the statistical technique offers a major interpretative wealth, the factorial planes allow one to detect the relative position of the individuals from the standpoint of each group of variables. Hence, it can be detected the group of variables that defines the tendency of each individual, as well as the differences between the individuals according to the different standpoints, that is to say, according to the different groups of variables. The statistical programme used is SPAD 6.0.

The data used in this research come from the fourth and fifth wave of the Community Innovation Survey (CIS), elaborated by EUROSTAT and gathered in different European states.

The CIS is a survey on innovation activity in firms of the EU states, candidate countries, Iceland and Norway. Data were collected every four years since 1996. However, since 2004 a reduced version is presented every two years. The observation period for the CIS-4 was from 2002 to 2004 (excluding the Czech Republic, where the observation period was from 2003 to 2005). The CIS-5 is the last survey whose observation period was from 2004 to 2006. In the Survey the reference population is firms with 10 or more employees and it is based on a sample stratified by size and sector.

The CIS-5 largely follows the structure of the CIS-4, with minor modifications. This makes it feasible to make comparisons between the two waves. The comparison with other editions of the CIS is complicated from a technical perspective given that some editions have introduced changes in the definition of innovation, in the wording of certain questions of the questionnaire and in the number of participating countries. The difficulty in establishing comparisons over time is also conditioned by the quality of the data provided. It is a necessary condition for the correct application of certain statistical techniques that the selected variables contain data for each individual participating in the sample. The lack of data may invalidate the conclusions reached. Therefore, in this investigation we have decided to make a comparison of the waves that have a high degree of harmonization

The processing of the information provided by Eurostat for the countries participating in the survey is at industry level. However, for the aim of this research, an aggregated treat of data at manufacturing level has been made. On the other hand, the absence of data for some variables in some countries has meant that the study covers a total of eighteen European countries: Belgium, Czech Republic, Denmark, Estonia, Ireland, France, Italy, Lithuania, Luxembourg, Netherlands, Poland, Portugal, Romania, Slovakia and Sweden.

2.2. Selection of variables

The selection of variables has been made taking into account the availability of information and the different dimensions that, according to the theory, affect the

innovative activity. Of all the information contained in the CIS, firstly, a set of variables has been selected and then, variables have been grouped into two groups for each wave of information. The groups of variables refer, on the one hand, to manufacturing as a whole including both innovative and non-innovative firms. And, on the other, to the subset of innovative manufacturing firms[1]

The information provided by each of subsets is different. Innovative activity across the manufacturing provides an idea of the diffusion of the forms of innovation in all firms, but does not capture the intensity with which the innovation takes place in innovative companies. The study of innovation in the innovative firms allows to know the concentration of forms of innovation in such firms, acting as an indicator of innovation intensity. The distinction between the innovative profile of the whole manufacturing and of the innovative firms allows to detect whether there exist countries with a scarce (intense) diffusion of innovation across manufacturing, but where innovative firms perform an intense innovative activity by combining several (few) types of innovation.

By Proceeding in this way, we can carry out two types of analysis. Firstly, considering the information from all groups of variables, the innovative profile of the European countries can be studied. And secondly, the change in the innovative profile that has occurred in manufacturing and in the innovative firms of countries during the two waves of information can be analyzed. Table 3 presents the variables included in each group of variables. All variables are continuous.[2].

Table 3: Variables included in the analysis

GROUP	VARIABLE	LABEL
GROUP 1 Innovation in manufacturing (CIS 5)	Percentage of innovative firms	C5_MANUF_Inno
	Percentage of firms that perform R & D	C5_MANUF_I+D
	Percentage of firms that innovate in product	C5_MANUF_Pd
	Percentage of firms that innovate in process	C5_MANUF_Pc
	Percentage of expenditure in innovation/total income	C5_MANUF_Ginno
	Percentage of expenditure in R&D/total income	C5_MANUF_GI+D
	Percentage of expenditure in machinery/total income	C5_MANUF_Gmaq
GROUP 2 Innovación en la manufactura (CIS 4)	Percentage of innovative firms	C4_MANUF_Inno
	Percentage of firms that perform R & D	C4_MANUF_I+D
	Percentage of firms that innovate in product	C4_MANUF_Pd
	Percentage of firms that innovate in process	C4_MANUF_Pc
	Percentage of expenditure in innovation/total income	C4_MANUF_Ginno
	Percentage of expenditure in R&D/total income	C4_MANUF_GI+D
	Percentage of expenditure in machinery/total income	C4_MANUF_Gmaq
GROUP 3 Innovación en tejido innovador (CIS 5)	Percentage of firms that perform R & D	C5_INNOV_I+D
	Percentage of firms that innovate in product	C5_INNOV_Pd
	Percentage of firms that innovate in process	C5_INNOV_Pc
	Percentage of expenditure in innovation/total income	C5_INNOV_Ginno
	Percentage of expenditure in R&D/total income	C5_INNOV_GI+D
	Percentage of expenditure in machinery/total income	C5_INNOV_Gmaq
GROUP 4 Innovación en tejido innovador (CIS 4)	Percentage of firms that perform R & D	C4_INNOV_I+D
	Percentage of firms that innovate in product	C4_INNOV_Pd
	Percentage of firms that innovate in process	C4_INNOV_Pc
	Percentage of expenditure in innovation/total income	C4_INNOV_Ginno
	Percentage of expenditure in R&D/total income	C4_INNOV_GI+D
	Percentage of expenditure in machinery/total income	C4_CINNOV_Gmaq

We have, therefore, a matrix of 18 individuals (countries) and 26 variables, classified into four groups or tables (two for each round). Table 4 presents a descriptive analysis of the variables included in the study.

[1] The definition of innovation that is used in the two waves of the CIS is harmonized. It is considered that a firm is innovative when the firm has introduced new or improved products on the market or it has implemented new or improved processes. The term includes all types of innovation: product innovation and process innovation.

[2] Note that for continuous variables AFM behaves as an ACP (weighted variables) and for nominal variables behaves like an ACM (weighted modalities).

Table 4: Description of variables

	CIS 5				CIS 4			
	Mean	S.D.	Max	Min	Mean	S.D.	Max	Min
MANUF_Inno	43.31	14.86	21.21	71.16	42.46	14.29	21.19	73.99
MANUF_I+D	14.44	9.37	2.68	36.53	13.91	8.03	2.11	29.19
MANUF_Pd	22.41	8.54	10.60	36.61	21.87	8.69	10.52	42.46
MANUF_Pc	17.80	6.53	6.26	30.61	18.00	6.64	6.41	35.52
MANUF_GInno	2.71	1.35	0.98	6.19	2.83	1.24	1.37	5.86
MANUF_GI+D	1.20	1.10	0.21	3.95	1.41	1.18	0.22	3.89
MANUF_Gmaq	1.18	0.73	0.36	3.38	1.19	0.57	0.35	2.57
INNOV_I+D	31.24	14.02	11.29	61.71	30.56	11.72	8.06	52.00
INNOV_Pd	51.79	8.01	36.17	68.46	51.26	7.44	36.17	69.11
INNOV_Pc	41.92	10.80	23.13	71.97	43.18	10.76	25.93	65.34
INNOV_GInno	3.63	1.49	1.38	7.73	3.74	1.34	2.03	7.16
INNOV_GI+D	1.81	1.44	0.31	4.93	1.81	1.39	0.30	4.75
INNOV_GMaq	1.57	0.97	0.45	4.29	1.66	0.88	0.45	3.59

3. RESULTS

3.1. Results of the afm

This section contains the presentation and interpretation of the most outstanding results of the analysis of the tables above mentioned. Table 5 shows the histogram of the first five eigenvalues of the global analysis of all the juxtaposed tables. The table reveals the existence of a dominant first global axis, in that it incorporates a significant percentage of inertia, 46.08% versus 17.57% and 15.52% of variability that is projected onto the second and third global factor, respectively. In this work the first two axes will be considered, as they summarize properly the major features of the countries, ie, the main differences and similarities in terms of innovative behaviour.

Table 5: Histogram

Number	Eigenvalue	Percentage	Accumulated Percentage
1	3.5848	46.08	46.08
2	1.3670	17.57	63.65
3	1.2071	15.52	79.17
4	0.6570	8.44	87.61
5	0.3832	4.93	92.54

Tables 6 and 7 show the Lg coefficients matrix and the RV coefficients matrix, respectively. These matrices provide us information regarding the relationship between the groups of variables analyzed. The Lg coefficients matrix shows a great similarity between the elements of the main diagonal, which means that the groups have a similar dimensionality, ie, have a similar number of factors with considerable inertia. The RV coefficients matrix provides additional information. In this matrix, coefficients can take values between 0 and 1, and has the advantage that groups are directly comparable. In our case, by using these coefficients, it is found that the groups are closely related in time, ie, the group that expresses the innovative activity in manufacturing in 2006 is highly correlated (0.858) with the same group in 2002. This indicates that there have been some changes, but that these changes have been modest. In the case of the innovative firms, the over time correlation is even higher (0,879) which means that changes have been smaller than in the whole manufacturing. The correlation between the innovative profile of the whole manufacturing and the innovative profile of the subset of innovative firms is similar for the two waves under study (0,681 in 2006 and 0,675 in 2002). This correlation, which is not very high, permits us to work with the hypothesis that there exist differences between the relative innovative position of each country in the manufacturing as a whole and their relative innovative position in the subset of innovative firms.

Table 6. Lg coefficients matrix

	1	2	3	4	AFM
1	1.158				
2	0.986	1.140			
3	0.913	0.766	1.550		
4	0.900	0.874	1.327	1.470	
AFM	1.104	10.050	1.271	1.275	1.311
	1	2	3	4	AFM

Table 7. RV coefficients matrix

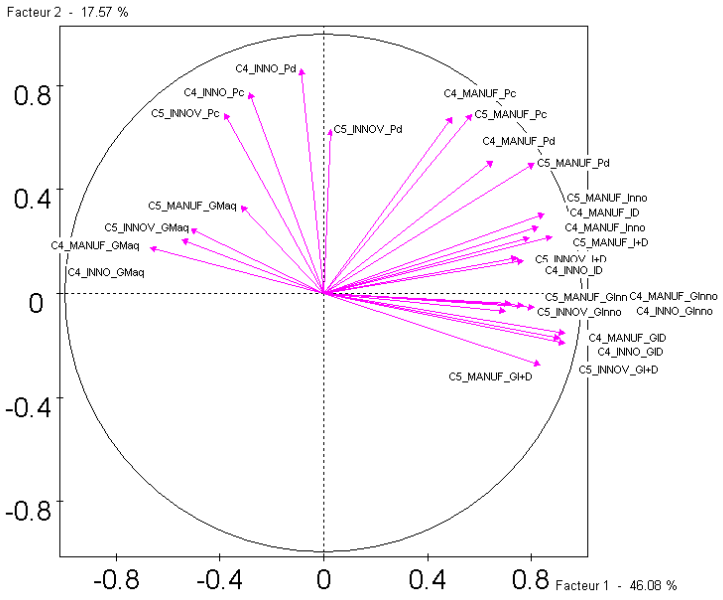
	1	2	3	4	AFM
1	1.000				
2	0.858	1.000			
3	0.681	0.576	1.000		
4	0.690	0.675	0.879	1.000	
AFM	0.896	0.859	0.891	0.918	1.000
	1	2	3	4	AFM

3.2. Analysis of factors and relative position of the countries

The variables that most contribute to the formation of the axes are represented in Figure 1.

The first factor provides information about the behaviour of variables related to R&D activities and expenditure in innovation in both manufacturing and the innovative firms. The main difference among European countries regarding these variables will be reflected through a ranking where countries situated in the right semiplane will be characterized by a high innovative profile in both the whole manufacturing and the subset of innovative firms as well as in the two waves. By contrast, countries situated in the left semiplane will be characterized by a lower innovative profile, in both manufacturing and the set of innovative firms.

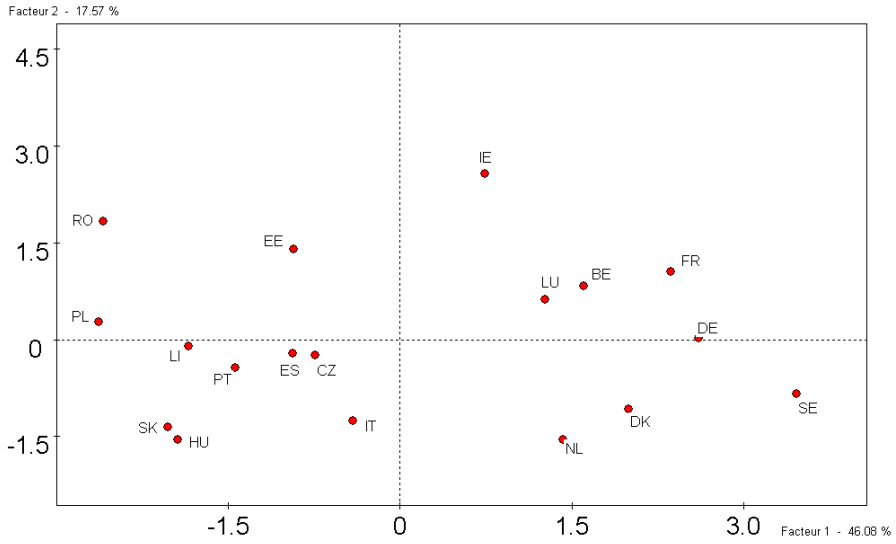
Figure 1. First factorial plane of variables and partial axes



The second axis is a factor that reflects process and product innovation. These variables will be projected on the upper semiplane, for both manufacturing and the set of innovative firms as well as for the two waves of information. Therefore, countries that are characterized by these forms of innovation will be projected in this semiplane. However, it must be noted that in the first quadrant will be projected those countries where these forms of innovation are complementary to R & D and expenditure in innovation, while, in the second quadrant will be screened those countries where these forms of innovation are substitute of R & D and expenditure in innovation.

Based on the distribution of the variables, Figure 2 shows the representation of the mean individuals on the first factorial plane, ie the representation of the countries mean position, taking into account all the views or groups of variables considered. Therefore, this factorial plane provides a very general view through the relative positions of European countries in the first two factors.

Figure 2: Representation of the *mean countries* on the first factorial plane



The horizontal axis (factor 1) can produce an initial ranking of countries as regards the R&D and the expenditure on innovation. Countries that have the most important positions from highest to lowest are: Sweden, Germany, France, Denmark, Belgium, Netherlands and Luxembourg. Ireland and Italy occupy intermediate positions. In the last positions are: the Czech Republic, Estonia, Portugal, Lithuania, Hungary, Lithuania, Slovakia, Romania and Poland.

This classification should be qualified by the information of the second factor. In the positive semiplane of the vertical axis (factor 2) countries with greater process and product innovation are projected. Under this new characterization, a ranking headed by Ireland, Romania, Estonia, France, Belgium and Luxembourg is obtained. In positions close to average are: Germany, Lithuania, Poland, Spain, Czech Republic and Portugal. The countries that occupy less prominent positions are Sweden, Denmark, Netherlands, Italy, Slovakia and Hungary.

By quadrants, in the first one we find countries that are characterized by performing R&D, introducing process and product innovations and spending a large income share in innovation activities. These countries exhibit the highest innovative profile of the European Union. The fourth quadrant represents countries characterized by performing R&D activities and spending a large income share in innovation, but not so much by introducing process and product innovations. In the second quadrant appear countries characterized by innovating in product and process and by spending resources in acquisition of machinery. In the third quadrant are located the least innovative countries of the European Union. These results are consistent with those obtained in other studies (Altuzarra, Puerta, Serrano, 2007; Altuzarra, Serrano, 2008).

3.3. Analysis of convergence in innovation

In the previous section, we have represented the relative mean position for each country on the first factorial plane, taking into account all the information. However, it may happen that the relative position that corresponds to a country in one or more of the viewpoints is significantly far from that mean position. One of the advantages of the statistical technique we use is that it allows us to project each country on the factorial plane, from all the viewpoints that we have been considered. Thus, each country could be represented on the factorial plane through five points: one of them reflects the mean position of the country and the rest, the relative partial position of each country in each of the four viewpoints.

This section seeks to identify changes that have occurred in the period analyzed as regards innovative activity in the manufacturing and in the set of innovative firms. To

carry out this analysis, initially, we have tested whether countries have different relative positions in manufacturing and in the subset of innovative firms. Only if it is, it would be convenient to treat manufacturing as a whole and the set of innovative firms separately to study the changes in the relative positions of countries over time.

Figures 3a and 3b show the representation of the relative partial position of countries in both manufacturing and the set of innovative firms. To ease interpretation, only countries of the last wave (CIS 5) have been represented. Also, for clarity, figures have been divided vertically into two semiplanes, one, with the most innovative countries and the other one, with the least innovative ones.

Figure 3 (a). Relative position of countries in the manufacturing and in the set of innovative firms (right semiplane)

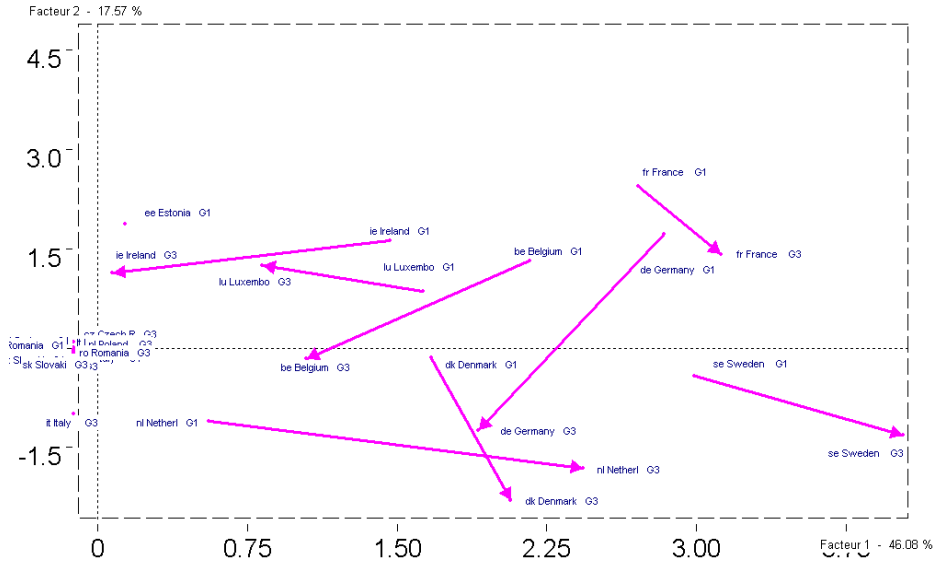
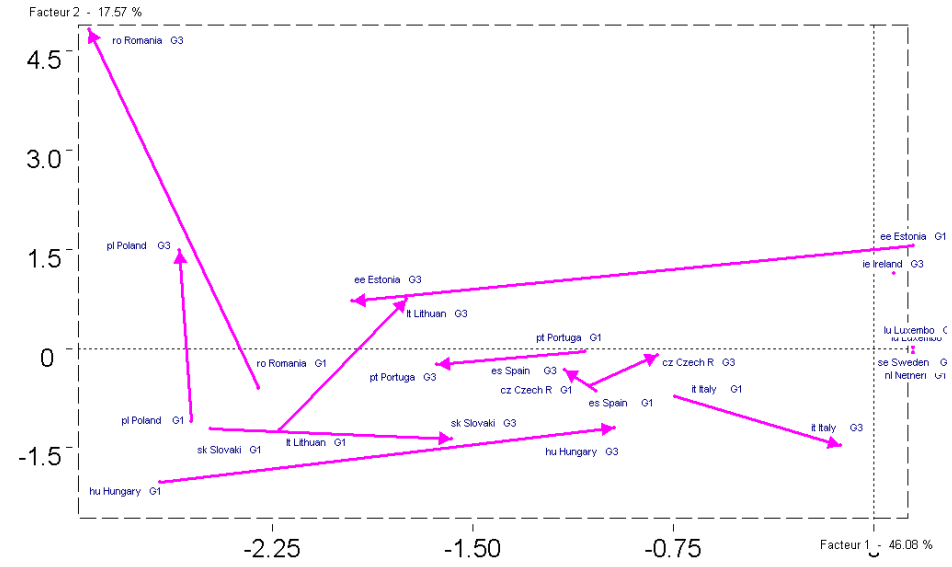


Figure 3 (b). Relative position of countries in the manufacturing and in the set of innovative firms (left semiplane)



Figures show that the relative position of countries in the whole manufacturing differs from the countries relative position in the set of innovative firms. That is, the relative innovative profile for the whole manufacturing in many countries does not

match the relative innovative profile of the set of innovative firms in such countries. The innovative profile of manufacturing, as already mentioned, is related mainly to the diffusion of forms of innovation across manufacturing, while the profile of the set of innovative firms is associated with the concentration of forms of innovation in innovative firms.

The differences in the relative position of the whole manufacturing and the innovative firms have been represented by trajectories. The displacement of the trajectories towards the left means that the country relative position in manufacturing is better than the country relative position in the set of innovative firms. The opposite is true when the trajectories move towards the right.

The results show that countries that exhibit a better relative position in the whole manufacturing than in the set of innovative firms are: Ireland, Belgium, Luxembourg, Germany (right semiplane), Romania, Estonia and Portugal (left semiplane). By contrast, France, Sweden, Netherlands, Denmark (right semiplane), Hungary, Italy, Lithuania and Slovakia (left semiplane) show a better relative position in the set of innovative firms than in the whole manufacturing. Spain, Czech Republic and Poland have no significant deviations between the two sets analysed.

The existing asymmetry in most countries between the innovative profile of the whole manufacturing and of the set of innovative firms makes it interesting to analyse the innovative dynamics of countries in each set separately.

The following figures show the changes in the relative innovative position of European states in the whole manufacturing and in the set of innovative firms. The arrows represent the over time trajectories of the innovative profile of the countries. The displacement of the trajectories towards the right expresses an improvement of the relative position of countries and the opposite happens when the trajectory moves towards the right[3].

The results regarding the whole manufacturing show that, countries that have improved their relative innovative position are: France, Luxembourg, Estonia and Portugal. Ireland, Denmark, Germany, Lithuania, Spain and the Czech Republic have experienced a relative decline. Netherlands, Belgium, Sweden, Poland, Slovakia, Hungary, Italy and Romania have not shown significant changes. The relative decline in the whole manufacturing of some countries with high innovative profile along with the relative improvement observed in manufacturing of countries with low innovative profile suggests that convergence in innovation of European countries has been moderate. Besides, this convergence has not been evenly distributed across European countries. For example, manufacturing in France, despite having a high innovative profile has continued to experience an improvement and, the opposite happens in Spain and the Czech Republic.

Regarding the set of innovative firms, countries that present advances in the relative position are: France, Ireland, Netherlands and Hungary, while Luxembourg, Belgium, Italy, Germany, Estonia, Romania, Spain and the Czech Republic have fallen in its relative position. Sweden, Denmark, Portugal and Lithuania have hardly change. In the case of the set of innovative firms has not been observed a convergence process in innovation. A nunber of high innovative profile countries such as Sweden, France, the Netherlands have shown progress in their relative position, while some lower innovative profile countries such as Netherlands or Estonia have experienced setbacks. This has meant that countries exhibit a greater dispersion in the factorial plane and a reduction of the convergence in innovation.

[3] Intra inertia provides an indicator of the variation that a country shows with respect to its centre of gravity. Countries with the strongest within inertia in the first factor are: Ireland (20.27), Netherlands (16,02), Estonia (13,75), France (10,09), Hungary (10,06), Germany (6,21), Belgium (4,59), Sweden (4,55). Countries with weaker intra inertia in the first factor are: Spain (0,58), Denmark (0,67), Poland (0,80), Czech Republic (0,88), Lithuania (0,88), Portugal (1,04) and Romania (1,28).

Figure 4a. Over time changes in the relative position of countries in the whole manufacturing (right semiplane)

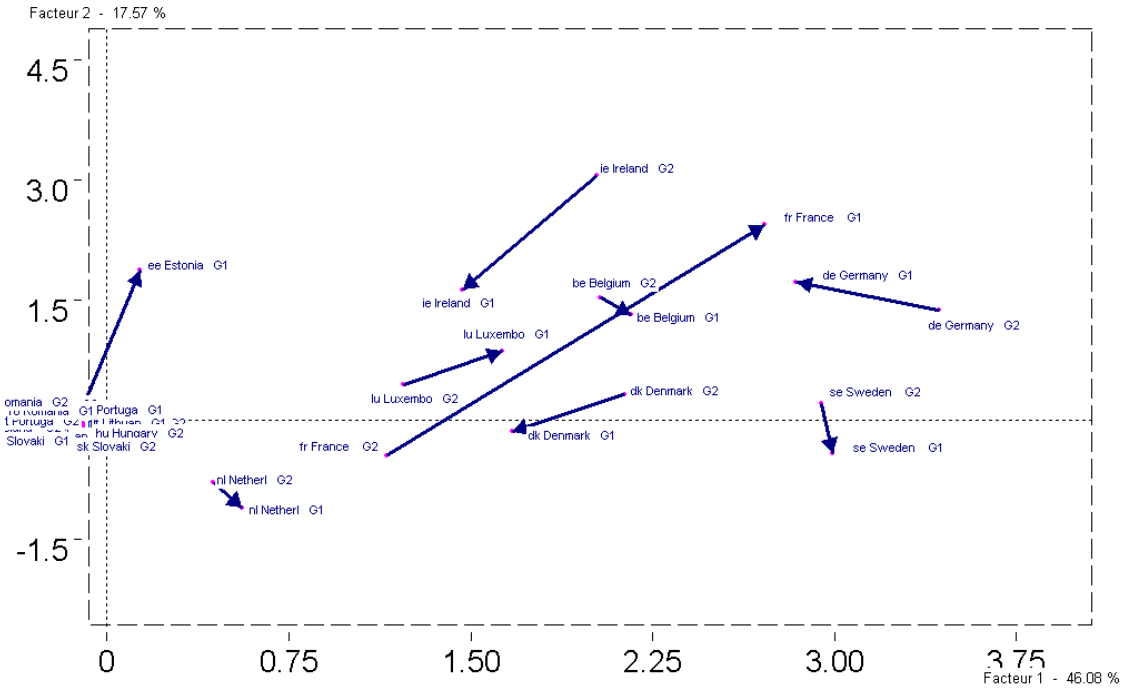


Figure 4b. Over time changes in the relative position of countries in the whole manufacturing (left semiplane)

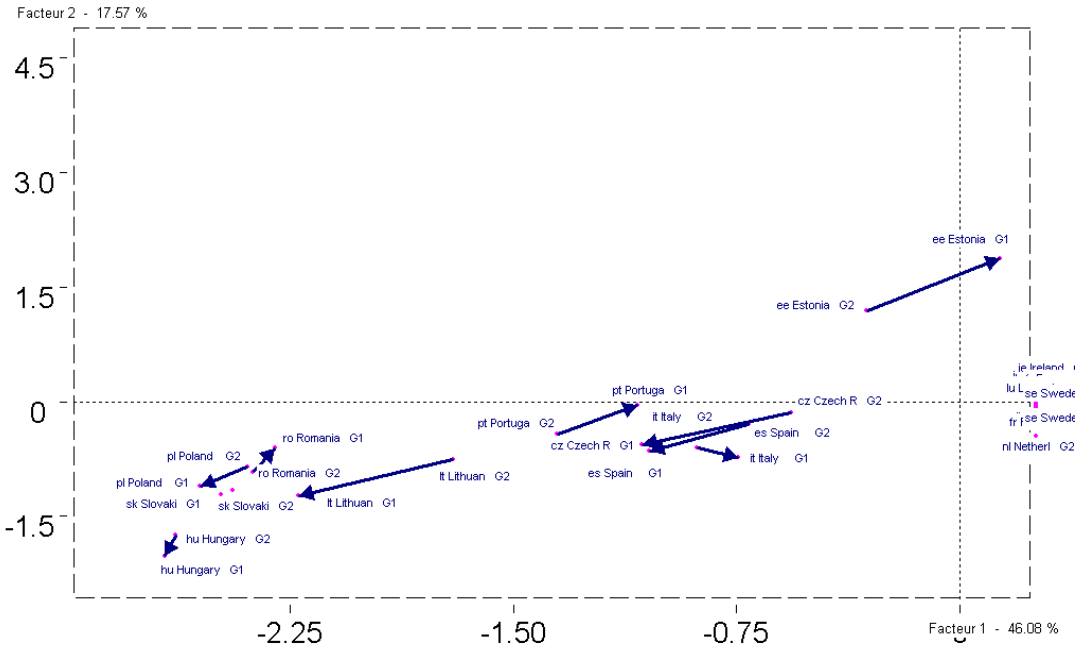


Figure 5a. Over time changes in the relative position of countries in the set of innovative firms (right semiplane)

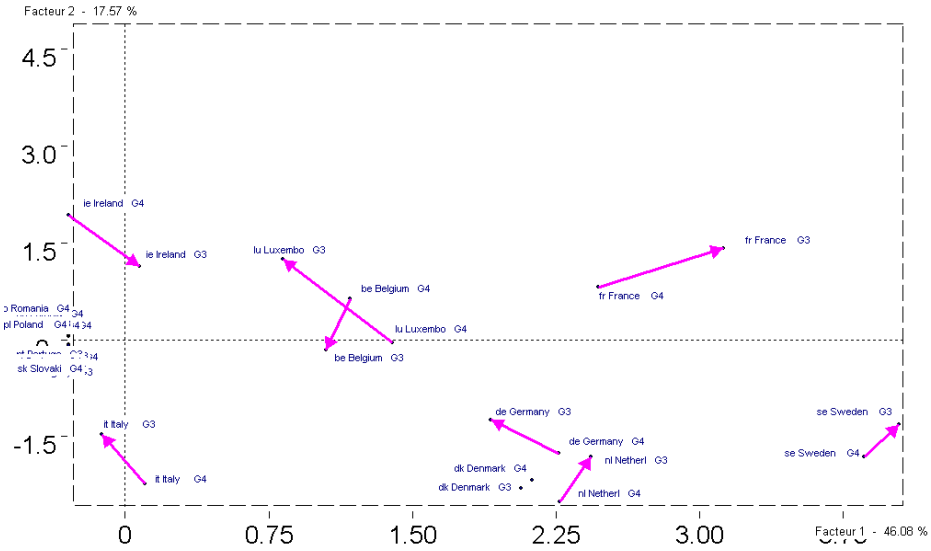
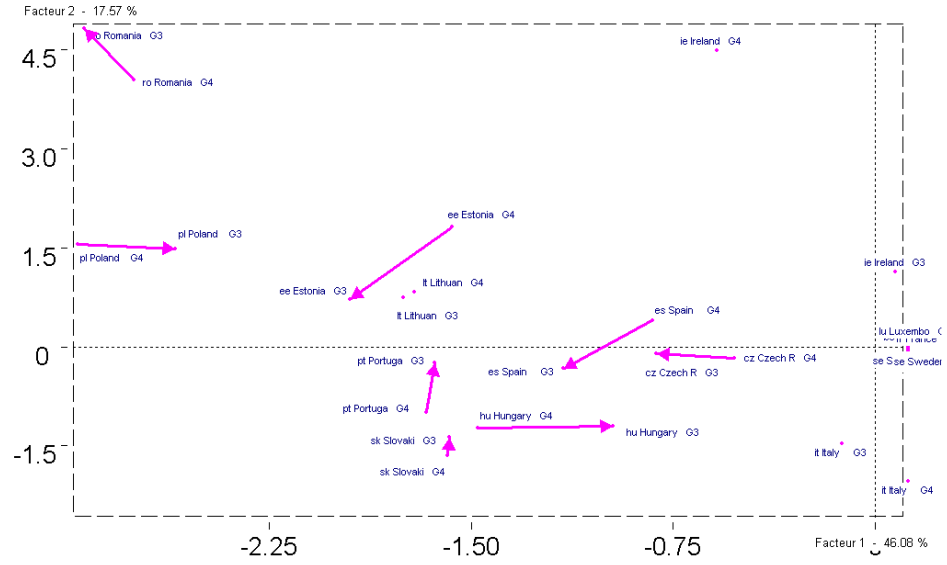


Figure 5b. Over time changes in the relative position of countries in the set of innovative firms (left semiplane)



The results show that some countries as Poland and Hungary have advanced significantly in the relative position of their set of innovative firms despite their non-improvements in manufacturing. Both countries have experienced the largest increase in the stock of direct investment relative to GDP. The Czech Republic, though still has some advantage over other recent acceding countries, this advantage begins to shorten in both manufacturing and the innovative sector. Estonia has experienced a strong relative improvement in manufacturing, but a relative decline in the innovative sector. The innovative strategy in this country seems to have been aimed at the diffusion of innovation among firms, being one of the countries that have had a major relative breakthrough in the European Union. Slovakia, Lithuania and Romania have not had significant improvements in either the manufacture or the set of innovative firms. The latter two countries are those that have the least FID (as percentage of GDP) of the countries of last European enlargements. Spain has lost relative positions in both manufacturing and innovation sector. This fact would support the hypothesis that EU enlargement has had negative

consequences for this country. Portugal, meanwhile, has shortened distance with Spain in manufacturing and in the set of innovative firms.

4. CONCLUSIONS

In this paper we have presented the changes in the innovative position of eighteen EU countries over the period 2002-2006 using the technique of Multiple Factor Analysis. The study has been carried out in two different sets of firms. On the one hand, it has considered all the manufacturing firms (innovative and non-innovative firms) and on the other, only the innovative manufacturing firms.

In terms of convergence in innovation it should be noted that there is a moderate advance in the whole manufacturing in the European countries over the period considered. However, it is not possible to say the same for the set of innovative firms. The results, as regards the innovative profile of manufacturing, show that countries that have improved the relative position are: France, Luxembourg, Estonia and Portugal. Countries that have worsened their relative position are: Ireland, Denmark, Germany, Lithuania, Spain and the Republic Czech. Netherlands, Belgium, Sweden, Poland, Slovakia, Hungary, Italy and Romania have not shown significant changes.

Regarding the set of innovative firms, countries that present advances in the relative position are: France, Ireland, Holland and Hungary, while those who have fallen in their relative position are: Luxembourg, Belgium, Italy, Germany, Estonia, Romania, Spain and the Czech Republic. Sweden, Denmark, Portugal and Lithuania have hardly changed. In the case of the innovative sector one can not assert that there has been a convergence process in innovation. A number of high innovative profile countries such as Sweden, France and Netherlands have shown progress in their relative position, while other low-profile innovative countries such as Romania or Estonia have experienced setbacks.

REFERENCES

AGHION, P., HOWITT, P. (1997), *Endogenous Growth Theory*, MIT Press Cambrige MA

ALTUZARRA, A., PUERTA, C., SERRANO, F. (2007), "Evaluating the innovative position of European Member Countries: an empirical analysis", *International Review of Applied Economics*, Vol. 21, 75-188

ALTUZARRA, A., SERRANO, F. (2008), "Innovation in the EU convergence States and the new Member Countries", *Ekonomia*, 9, 2

BARRO, R. (1991), "Economic growth in a cross section of countries", *Quarterly Journal of Economics*, 106, 407-43

BEVAN, A., ESTRIN, S. (2000), "The determinants of foreign direct investment in transition economies", Discussion paper, 9, *Centre for New and Emerging Market*, London Business School, London

BUCH, C. M., KOKTA, R.M, PIAZOLO, D. (2003), "Foreign direct investment in Europe. Is there redirection from the South to the East?", *Journal of Comparative Economics*, 31, 94-109

CARSTENSEN, K., TOUBAL, F. (2004), "Foreign direct investment in Central and Eastern European countries: a dynamic panel analysis", *Journal of Comparative Economic Studies*, Vol. 32, 1, 3-22

ESCOFIER, B. Y PAGÈS, J. (1994), "Multiple factor analysis (AFMULT package)", *Computational Statistics and Data Analysis*, 18: 121-140, 1994.

EUROPEAN COMMISSION, 2006, "Enlargement, two years after: an economic evaluation", *Occasional Papers*, 24, Directorate-General for Economic and Financial Affaires

FAGERBERG, J. (1987), "A technology gap approach to why growth rats differ", *Research Policy*, 16, 87-99

FAGERBERG, J., VERSPAGEN, B. (1996), "Heading for divergence?, Regional growth in Europe reconsidered", *Journal of Common Market Studies*, 34, 431-48

IAMMARINO, S. AND SANTANGELO, G.D. (2000), "FDI and Regional Attractiveness in the EU Integration Process: some evidence for the Italian regions", *European Urban and Regional Studies,* 7, 1, pp.5-18.

JUNGMITTAG, A. (2006), "Innovation dynamics in the EU: convergence or divergence' A cross-country panel data analysis", *Empirical Economics*, 31, 313-331

KALDOR, N. (1966), *Causes of the slow rate of growth in the United Kingdom*, Cambridge University Press

KUTAN, A. M., YIGIT, T.M. (2007), "European integration, productivity growth and real convergence", *European Economic Review*, 51, 1370-1395

LANDALUCE, M.I. Y APARICIO, S. (2.000), "La importancia de la estructuración en tablas de datos: Una aplicación", *Estudios de Economía Aplicada*, 15, pp. 103-124

MONETARY POLICY AND THE ECONOMY (2008), *The impact of EU enlargement in 2004 and 2007 on FDI and migration flows gravity analysis of factor mobility*, Q2

ROJEC, M, DAMIJAN, J.P. (2008), "Relocation vie foreign direct investment from old to new EU member states. Scale and structural dimension of the process", *Structural Change and Economic Dynamics*, 19, 53-65

UNCTAD (2008), *World Investment Report*, United National, New York.

DETERMINANTS AND EFFECTS OF INCREASING DIVERSIFICATION: NEW EMPIRICAL EVIDENCE FOR THE SPANISH ECONOMY

ALMUDENA MARTÍNEZ CAMPILLO
E-mail: amarc@unileon.es
Mª DEL PILAR SIERRA FERNÁNDEZ
E-mail: pilar.sierra@unileon.es
Facultad de Ciencias Económicas y Empresariales
Universidad de León

Abstract

Previous works that examine the effects of corporate diversification in the Spanish economy are scarce and contradictory, which impedes the generation of widely accepted conclusions. This paper aims to provide new empirical evidence on the nature of the diversification-performance relationship in Spain during a period characterized by important transformations in the competitive environment (1997-2001) and hence in the strategic tendency of firms. After using Heckman´s two-stage method on a sample of 236 large Spanish firms, results indicate that increasing diversification positively and significantly influenced firm profitability over this period. By comparing these results with those previously obtained for the Spanish economy, it may be concluded that the level of diversification and their consequences vary with time and contextual changes.

Key-Words: Corporate diversification, determinants, effects, environmental transformations.
Área Temática: Economía Europea.

1. INTRODUCTION

The diversification strategy is defined as "*the entry of a firm into new lines of activity, either by processes of internal business development or acquisition, which entails changes in its administrative structure, systems, and other management processes*" (Ramanujan and Varadarajan, 1989: 525). Researchers from different areas have sought to explain the antecedents of this strategy and its relationship with firm performance. In particular, this last question is the topic of extensive studies for leading developed countries[1] (Denis *et al*., 1997, 1999; Campa and Kedia, 2002; Kim *et al*., 2004; Desai *et al*., 2005; Miller, 2004, 2006), but evidence hardly exists for other more recently developed countries. Thus, considering the importance that researchers from other nations (such us USA, Canada, UK, France, Germany or Japan) are giving to the study of the performance effects of diversification, and given that previous research analyzing the diversification-performance relationship in the Spanish economy is scarce, contradictory and not too recent (Suárez-González, 1994; Menéndez and Gómez, 2000; Ramírez and Espitia, 2002), it may be interesting to provide new empirical evidence on this issue in Spain, whose business environment is different from the one examined in most early works.

Although Spanish firms have traditionally shown a low level of diversification (Ramírez and Espitia, 2002; Sánchez-Bueno *et al.*, 2006), contextual changes that have taken place in recent

[1] Despite that an extensive research examines the diversification-performance relationship in leading developed countries, the empirical evidence remains inconclusive.

years have led to an extremely competitive environment, in which diversification constitutes one of the strategies that firms may choose to adapt and survive (Simanavičienė and Dagilienė, 2003; Cottrel and Nault, 2004; Colpan and Hikino, 2005). As a result, Spanish firms have indeed increased their levels of diversification in the last decade[2], which in turn may have had varying effects on firm performance. Also, the fact that Spanish companies are smaller in size and less diversified than those in most leading developed countries may lead to insights different from the ones found in previous research in other countries.

For these reasons, the main goal of this study is to examine the nature of the diversification-performance relationship in the Spanish economy in a new context, characterized by important transformations in business environment and the strategic tendency of firms. To that end, since researchers have raised serious questions about the methodology of preceding studies linking diversification strategy to firm performance (Palich *et al.*, 2000; Miller, 2006), this study uses Heckman´s (1979) two-stage method[3]. This statistical technique allows the proper correction of some econometric problems traditionally associated with the estimation of diversification effects, such as sample selection bias and endogeneity.

This work contributes to the existing literature in three important ways: *First*, the study seeks to advance the state of the art on the diversification-performance relationship in Spain by adding to the limited and contradictory empirical research existing to date (Ramírez and Espitia, 2002). Moreover, it provides additional information on this topic by analyzing for the first time the performance effects of the initial decision to diversify; that is, when firms increase their number of business segments from one to two or more. *Second*, the study considers that the nature of the diversification-performance relationship may vary with contextual changes within a country across time (Mayer and Whittington, 2003; Chakrabarti *et al.*, 2007; Singh *et al.*, 2007). For this reason, it attempts to predict the performance effects of diversification over the last years based on the new strategic tendency of Spanish firms derived from changes in business environment. Given that the relationship between diversification and performance seems to be influenced by contextual changes, it may be interesting to provide new empirical evidence for Spain in a different business context. *Third*, the study also contributes from a methodological perspective by applying Heckman´s two-stage method to control for possible endogeneity and sample selection biases in the diversification-performance relationship (Miller, 2006).

The structure of the article is as follows. Section 2 introduces and justifies the hypothesis proposed as regards the impact of diversification on firm performance in Spanish firms. Section 3 contains the data and empirical methodology. Section 4 reports the results. The final section offers some concluding remarks.

2. THEORETICAL FRAMEWORK

2.1 Antecedents about the diversification-performance relationship:

The analysis of diversification effects on firm performance is one of the most widely studied and debated topics in the academic and business world. However, the nature of the diversification-performance relationship remains largely unsolved to date insofar as it depends on the benefits and costs associated with this strategy. It is argued that diversification can be driven by a range of perceived benefits associated with market power advantages (Scherer, 1980), greater efficiency of internal capital markets (Williamson, 1985; Stein, 1997), scope and scale economies from sharing resources and capabilities across different businesses (Fraquelli *et al.*, 2004; Tanriverdi and Venkatraman, 2005), increased debt capacity (Lewellen, 1971), tax

[2] Despite this tendency, Spanish firms still show a significantly lower level of diversification than those from the rest of the European developed countries (Sánchez-Bueno *et al.*, 2006).

[3] To our knowledge, there is not any published study applying Heckman´s method to estimate the nature of the diversification-performance relationship in Spain. However, recent works have used it to analyze such a relationship in the context of the American economy (Campa and Kedia, 2002; Villalonga, 2004; Miller, 2006).

and other financial advantages (Majd and Myers 1987; Berger and Ofek, 1995), and risk reduction (Lubatkin and Chatterjee, 1994). In contrast to diversification, single-business strategy does not provide these benefits.

Nonetheless, as the diversification degree increases the associated costs also escalate. The main potential cost linked with a higher level of diversification is the exacerbated managerial agency problem. Indeed, diversification may allow managers to increase their compensation and status in the business community, to reduce their personal employment risk and to become entrenched by managing this strategy in a way consistent with their own skills (Denis *et al.*, 1997, 1999). In particular, these costs may be described in terms of misallocation of funds due to cross-subsidization of failing businesses (Meyer *et al.*, 1992), internal power struggles generated by the allocation of resources between different divisions (Rajan *et al.*, 2000), inefficient investments in businesses with poor opportunities (Stulz, 1990) and information and incentive problems between corporate headquarters and division managers (Harris *et al.*, 1982), or between managers and external investors (Hadlock *et al.*, 2001). Additional costs may also arise due to the increase in coordination and control problems across businesses (Hoskisson and Hitt, 1988).

Because of these myriad benefits and costs, it is difficult to predict the net impact of diversification strategy on firm performance. This may explain why early research on the diversification-performance relationship does not show a consistent pattern in results. A review of the literature indicates four sets of findings to be prevalent. The first set of studies reports a positive effect of this strategy on firm performance (Rhoades, 1973; Campa and Kedia, 2002; Villalonga, 2004). The second set of findings concludes a negative relationship between diversification and performance (Berger and Ofek, 1995; Lamont and Polk, 2002; Singh *et al.*, 2007). A third set of studies finds no significant relationship between both variables (Gort, 1962; Palepu, 1985; Menéndez and Gómez, 2000). Finally, the fourth set of findings reports that the diversification-performance relationship is not linear, but curvilinear (Grant *et al.*, 1988; Ramírez and Espitia, 2002; Mayer and Whittington, 2003).

In this clearly contradictory analysis framework, a study by Palich *et al.* (2000) brings a degree of clarity by using meta-analytic data drawn from 55 previously published works. This study revises and synthesizes mayor theoretical perspectives and empirical studies on the subject and proposes three theoretical models to explain the nature of the diversification-performance relationship: the linear model and two curvilinear models; the inverted-U model and the intermediate model.

The *linear model* is based on the premise that diversification and performance are linearly and positively related (Arnould, 1969; Markham, 1973; Rhoades, 1973). Thus, under this model it is considered that unrelated diversification outperforms related diversification, while this last strategy performs better than single-business strategy. This position rests mainly upon the assumptions derived from the perspectives of market power and internal market efficiency (Scherer, 1980).

The *inverted-U model* posits that related diversification is better than unrelated diversification and single-business strategies in terms of firm performance (Grant *et al.*, 1988; Palich *et al.*, 2000; Ramírez and Espitia, 2002; Mayer and Whittington, 2003). The main theoretical argument suggesting the superiority of a moderate level of diversification focuses on advantages derived from scope and scale economies, since the firm is able to distribute the cost of an asset already capitalized by spreading its use across multiple linked businesses (Markides and Williamson, 1994; Fraquelli *et al.*, 2004; Tanriverdi and Venkatraman, 2005). However, this model argues that the marginal costs of diversification increase rapidly as the diversification hits high levels because of problems derived from managing an increasingly disparate portfolio of businesses, as well as control and effort losses due to increased shirking, coordination and administrative costs, and internal capital market inefficiencies (Grant *et al.*, 1988). Thus, this model posits that firms experience some optimal level of diversification, with performance decrements to either side of that point of maximization.

Finally, the *intermediate model* notions that diversification yields positive but diminishing returns beyond some point of optimization (Markides, 1992). A diversifying firm will first apply its excess assets in the closest market it can enter. If excess capacity remains, the firm will enter

markets even further afield. But as assets are applied in more and more distant fields, they lose their competitive advantage and thus earn lower profits (Wernerfelt and Montgomery, 1988). This implies that the relation between diversification and its marginal benefits is a decreasing function (Grant and Jammine, 1988). Therefore, under this model it is assumed that diversification performs better than single-business strategy, although both related and unrelated diversification may have a similar impact on firm performance, due to the costs of exploiting relatedness between businesses and the unique benefits derived from unrelated strategy (Nayyar, 1992).

2.2 The diversification-performance relationship in the Spanish Economy

The diversification-performance relationship is one of the issues that is arousing most interest in economic research. Despite this fact, empirical studies in the Spanish economy are limited and their findings are clearly contradictory. Firstly, the study by Suárez-González (1994) examines the impact of diversification on firm performance in a sample of 4,548 firms and concludes that single-business firms outperformed diversified firms over the period 1987-1990. Secondly, Menéndez and Gómez (2000), by using a panel data of 64 non-financial firms quoted on the Madrid Stock Exchange during the period 1990-1994, show that diversification strategy does not seem to significantly influence firm value. Finally, Ramírez and Espitia (2002) demonstrate using a sample of 103 large, non-financial Spanish companies, that the firms with moderate levels of diversification had the best performance, while the firms with low and high levels of diversification showed significantly lower performance over the period 1992 to 1995.

Spanish firms have traditionally been characterized by a low level of diversification (Ramírez and Espitia, 2002; Sánchez-Bueno *et al.*, 2006). However, during the last decade the entry into new lines of activity has become a common practice in business reality, as well as being one of the most important alternatives for corporate growth, along with internationalization and innovation. The main reasons behind this change in the strategic tendency of firms seem to be the following: the significant growth potential of the Spanish economy over this period (Ramírez and Espitia, 2002), and the fact that the competitive environment in which Spanish firms operate has changed dramatically over the last ten years. Indeed, they have been immersed in an extremely fierce world, determined basically by the globalization of markets and a growing emphasis on knowledge and innovation (Sánchez-Bueno *et al.*, 2006). In this new business context, it is of paramount importance that companies understand and adapt to these environmental transformations if they want to survive. In this regard, diversification is a central phenomenon for the adaptation, survival, and growth of firms in a highly competitive environment; that is, it is a critical strategy for continued existence of firms (Simanavičienė and Dagilienė, 2003; Cottrel and Nault, 2004; Colpan and Hikino, 2005).

A recent study by Sánchez-Bueno *et al.* (2006) notes that during the period 1993-2003, in order to adapt to the new business context, large Spanish firms have shown a marked decrease in the single-business strategy (-15 percent), the predominance and stability of the dominant-business strategy (44 percent of them), and a substantial increase in related diversification (+14 percent). Indeed, "*in this period of profound changes in the competitive environment, the most frequent moves are those from single-business strategy to dominant-business strategy and from these two strategies to related diversification"* (Sánchez-Bueno *et al.*, 2006: 28), while unrelated diversification hardly has representation in the whole decade[4]. Thus, the panorama

[4] The study by Sánchez-Bueno *et al.* (2006) uses Rumelt´s (1974) classification of diversification strategies, based on the specialization and related ratios: *Single-business strategy* implies that the proportion of a firm´s revenue attributable to its largest single business is 95% or more in a given year. *Dominant-business strategy* implies that the proportion of a firm´s revenue attributable to its largest single business is 70% or more but less than 95%. *Related diversification* implies that the proportion of a firm´s revenue attributable to its largest single business is less than 70% and majority of businesses are related to each other (the proportion of a firm´s revenue attributable to its largest group of related businesses is 70% or more). *Unrelated diversification* implies that the proportion of a firm´s revenue attributable to its largest single business is less than 70% and majority of businesses are not related to each other (the proportion of a firm´s revenue attributable to its largest group of related businesses is less than 70%).

described reveals that in recent years large Spanish firms have shown a clear tendency towards greater diversification, especially towards moderate levels of diversification[5].

Given that the nature of the diversification-performance relationship varies widely with changes in context over time (Geringer *et al.*, 2000; Mayer and Whittington, 2003; Chakrabarti *et al.*, 2007; Singh *et al.*, 2007), such variation in the strategic tendency of Spanish firms derived from environmental transformations may have had an impact on subsequent performance. In particular, Palich *et al.* (2000: 164) conclude that "*positive effects (on firm performance) occur as firms shift from a single-business strategy to a related diversification strategy"*. Indeed, whichever of the three theoretical models proposed by these researchers - linear, inverted-U and intermediate- allow us to deduce that diversification is positively related to performance across the low to moderate range of diversification. Firms with low level of diversification do not have multiple businesses, so they may not enjoy economies of scope and other important advantages that may accompany diversification, with the consequent negative implications for performance. However, firms with moderate levels of diversification may generate cost advantages derived from sharing resources and activities between businesses. Moreover, these firms may also benefit from learning curve efficiencies, intra-firm product/process technology diffusion, and restricted access to factors of production that are necessary for operations stemming from a specific industry (Markides and Williamson, 1994; Tanriverdi and Venkatraman, 2005).

In addition, the study by Palich *et al.* (2000: 164) adds that these positive effects of moderate levels of diversification on firm performance *"are detectable as a linear effect in samples restricted away from the high end of diversification"* –that is, in samples that hardly include firms with high levels of diversification[6]. With these arguments in mind, considering the strategic tendency shown by large Spanish firms during the last decade, the study presents the following hypothesis in order to test the linear model over the period analyzed[7]:

H_1: *Diversification strategy has a positive linear effect on firm performance.*

3. METHODS

3.1 Population

The population for this study comprises a set of large firms operating in Spain over the period from 1997 to 2001[8]. Specifically, those Spanish public companies with total sales greater than three million euros per annum and more than 100 employees during this time period were chosen. The Dun&Bradstreet Directory yields a total of 3,655 firms fulfilling these criteria. However, the question of evaluating diversification effects is best broached by focusing on firms when they decide to diversify for the first time; that is, when they increase their number of business segments from one to two or more (Miller, 2004; Villalonga, 2004). For this reason, we restrict the research to firms reporting only one business segment at the four-digit SIC level in 1997, excluding all companies specializing in financial services, regulated utilities,

[5] As Palich et al. (2000: 158) note "*level and type of diversification are conceptually distinct, but we do not differenciate them here... In support to our approach, empirical research consistently indicates that type of diversification is strongly associated with continuous data representing level of diversification*". For the purpose of maintaining homogeneity and allowing comparisons, in this paper is also considered that single-business strategy is equivalent to low diversification, the range from dominant-business strategy to related strategy is equivalent to moderate diversification, and unrelated strategy is equivalent to high diversification.

[6] The findings of the study by Palich *et al.* (2000: 164) expressly state that "*Diversification appears to have an inverted-U curvilinear relationship with performance: positive effects occur as firms move from a single-business strategy to a related diversification strategy (detectable as a linear effect in samples restricted away from the high end of diversification), but negative effects occur as firms move from a related strategy to an unrelated strategy (detectable as a linear effect in samples restricted away from the low end of diversification)*"

[7] The other relations that emerge from the review of the literature are considered in the alternative hypothesis.

[8] The analysis of the period between 1997 and 2001 is justified by three main reasons: (1) the lack of studies that examine the impact of diversification on Spanish firm performance since 1996; (2) the public availability, when the population was selected, of data on the number of businesses in which large Spanish firms operate until the year 2001; (3) the fact that important contextual changes with potential to affect the diversification-performance relationship have taken place during such period.

government and non-classifiable establishments. After applying these restrictions, the final population consists of 1,256 single-business firms in 1997. Once the different annual editions of the Dun&Bradstreet Directory were consulted, it was possible to observe that 520 companies made the decision to diversify during the 1998-2001 period and 736 remained specialized[9].

3.2 Sample selection

The measurement of the level of diversification the year in which firms diversify for the first time between 1998 and 2001 takes into account both the number of business segments at the four-digit SIC level in which they operate, which is taken from the Dun&Bradstreet Directory, and the relative importance of each business segment to the firm´s sales, which is not publicly available to Spanish firms. Thus, it was necessary to carry out a survey to collect this information about the year of diversification. Specifically, a questionnaire was sent to the CEOs of the 520 diversifying firms in the population, which was completed between May and July 2003[10]. A valid response rate of 22.7 percent provides a sample of 118 diversifying companies (sampling error was 8.1 percent with a 95 percent confidence level).

The distribution of these 118 firms as regards the diversification strategy followed (dominant-business, related diversification and unrelated diversification) is as follows[11]: Starting from all firms pursue a single-business strategy in 1997, 50.8 percent of them shift to a dominant-business strategy (60 firms), 40.6 percent to related diversification (48 firms), and only 8.4 percent to unrelated diversification (10 firms)[12]. Moreover, it is possible to arrive at the same conclusion as Sánchez-Bueno *et al.* (2006) regarding the strategic changes of specialized firms: the most frequent moves are those from single-business strategy to dominant-business strategy and related diversification; that is, towards moderate levels of diversification. As a result, it possible to conclude that the sample of diversifying firms used in this study does not present a high level of diversification; that is, it is a sample restricted away from the high end of diversification.

Recent literature indicates that "*the confidence with which one can draw conclusions from empirical studies of strategic phenomena is significantly limited if the sample is constructed of firms that have experienced the phenomenon under study*" (Jensen and Zajac, 2004: 512). For this reason, each diversifying firm from the sample was paired with one of the 736 specialized firms from the population to avoid sample selection bias. Matching criteria were proposed by Miller (2004) for a similar purpose: sharing the same principal business at the two-digit SIC code level and having a similar size (within 70-130 percent of sales and/or employees) in the year prior to the diversification event. After applying these criteria, the final sample consists of 236 firms, distributed equally between diversifying and specialized companies.

3.3 Independent variable

The entropy index (Jacquemin and Berry, 1979; Palepu, 1985) was used to measure the level of diversification the year *t* in which firms diversify for the first time between 1998 and 2001. This index is defined as follows:

[9] Dun&Bradstreet Directories annually offer the four-digit SIC codes for the 50,000 largest Spanish firms.

[10] These 520 firms were specialized in 1997 and they moved from a single-business strategy to a multi-business strategy during the period 1998-2001. Given that in this study the level of diversification is measured the year *t* in which the firm diversifies for the first time over this period, four different versions of the questionnaire were mailed: The *version 1* was sent to the firms that diversified for the first time in 1998 and their CEOs had to specify the proportion of sales in each business segment at the four-digit SIC level with respect to total sales in 1998; *versions 2, 3* and *4* were sent the firms that diversified for the first time in 1999, 2000 and 2001, respectively, and their CEOs had to specify the proportion of sales in each business segment at the four-digit SIC level with respect to total sales in 1999, 2000 and 2001, respectively. As well as inquiring about this question, the questionnaire also asked for information required in the framework of other research.

[11] The diversification strategy followed by each firm in the sample have been identified through Rumelt´s classification (see footnote 4), as in the works by Palich *et al.* (2000), Suárez-González (1994) and Sánchez-Bueno *et al.* (2006), for the purpose of maintaining homogeneity. As in these previous studies, the largest group of related businesses has been also determined within a two-digit SIC industry.

[12] As regards unrelated diversifiers, although Rumelt´s related ratios are less than 70%, they are not too distant from this threshold that allows differentiating between related and unrelated diversification (see footnote 4).

where n is number of the firm´s business segments at the four-digit SIC level and P_i is ith business segment´s sales divided by the firm´s total sales. The entropy measure is zero for single-business firms and it increases with greater diversification.

$$DIV = \sum_{i=1}^{n} P_i \ln (1/P_i) \qquad (1)$$

The choice of this index is justified because previous research concludes that "*the high degree of correspondence between the continuous (level of diversification) and categorical (type of diversification) measures can be interpreted as a "plus" for SIC-based diversification measures*" (Montgomery, 1982:305). As a result, on the base of the entropy index it is possible to consider that a low level of diversification is equivalent to a single-business strategy, a moderate level of diversification is equivalent to the range from dominant-business strategy to related diversification, and a high level of diversification is equivalent to unrelated diversification. Also, the entropy index has been reported as superior basically because it combines objectivity, content and construct validity, and simplicity (Kim *et al.*, 2004).

3.4 Dependent variable

The effects of diversification on firm performance were measured in terms of profitability. Specifically, *firm profitability* was measured by the variation in return on assets -ROA- (Jensen and Zajac, 2004; Desai *et al*., 2005); that is, a measure of accounting performance[13]. This variable was quantified as percentage change in average ROA between the three-year pre- and post-diversification periods. The averages for the three years prior and following to the diversification event were used to control for any seasonality in firms´ ROA figures. Performance data were taken from the SABI database.

3.5 Control variables

The analysis included the following control variables that have demonstrated significant effects on firm performance independent of diversification: (a) two variables to identify the corporate governance characteristics of companies the year of diversification (Mayer and Whittington, 2003): first, one to account for the *corporate control*, that takes a value of 1 if firms are owner-controlled (external owners have 5 percent or more of the outstanding shares) or owner-manager controlled (CEOs have 2 percent or more of the outstanding shares), and a value of 2 if firms are manager-controlled (external ownership is diffused -less than 5 percent- and CEO ownership is limited -less than 2 percent-)[14]; and, second, another one to control the *proportion of insiders on the board of directors*; (b) two firm-level economical variables, such as *size* (log of total assets) and *investment* (capital expenditures/sales), to abstract from any performance differences due to these internal factors (Palich *et al.*, 2000; Campa and Kedia, 2002; Villalonga, 2004). We calculate 3-year pre-diversification averages for these variables; (c) a dummy variable that indicates whether firms develop *export activity* the year of diversification, since firms that operate in more than one country are able to reap benefits that are not available to purely domestic firms (Palich *et al.*, 2000); and (d) one industry-level variable such as *profitability* (industry ROA at the four digit SIC level), since prior research has shown industry effects to have an important impact on variation of firm performance (Campa

[13] ROA was chosen as an accounting-based measure of firm profitability because this indicator remains the most widely used performance measure in the strategy literature (Kim *et al*., 2004). Although market value-based measures may be considered more complete indicators of the effects of corporate strategies, their use in this study would have restricted the population to firms that were quoted on the Spanish Stock Market. The limited nature and uneven distribution across industries of our national Stock Market advised against this approach.

[14] This criterion is also followed by Desai *et al.* (2005) to account for the corporate control when analyzing the performance effects of diversification.

and Kedia, 2002; Villalonga, 2004). We calculate 3-year pre-diversification average for this variable. All necessary data to construct control variables were taken from the SABI database.

3.6 Model specification

Standard regression techniques are not able to control for endogeneity bias from self-selection associated with studying the diversification-performance relationship. One solution is to apply Heckman´s (1979) two-stage method. In the first stage, the procedure estimates the selection equation as a maximum-likelihood probit model to analyze the propensity to diversify and calculate the *Inverse Mills Ratio* (λ_i). In the second stage, the corrected regression equation is estimated by OLS regression to examine the effects of diversification on performance.

Heckman´s method eliminates the endogeneity bias from self-selection in the following way (Wooldrigde, 2002): (1) it requires identifying at least one variable that may be a significant regressor in the selection equation but not in the regression equation. While other variables may be available, *the fraction of firms in the main industry that are diversified* (Villalonga, 2004) resulted to be appropriate for such a purpose. Due to data limitations, this variable had to be calculated from our population; and (2) it requires most regressors in the regression equation may be included in the selection equation.

In this case, the study considered the following regression and selection equations:

$$\boldsymbol{Y_i = \alpha + \beta_1 D_i + \beta_2 X_i + \varepsilon_i} \quad \textbf{(regression equation)} \qquad (2)$$

where Y_i is the diversifying firm´s $_i$ performance; D_i a diversification index[15]; X_i a vector of control variables and ε_i a normal error term. This equation uses data from diversifying firms.

$$\boldsymbol{DIV_i^* = \gamma Z_i + \mu_i} \quad \textbf{(selection equation)} \qquad (3)$$

where the latent variable DIV_i^* is observed as:

- $DIV_i = 1$ (the firm i decides to diversify) if $DIV_i^* > 0$,

or as

- $DIV_i = 0$ (the firm i decides not to diversify) if $DIV_i^* \leq 0$;

Z_i is a vector of variables that affect a firm´s propensity to diversify (all control variables from the regression equation since they also relate to the choice to diversify, as well as the fraction of diversified firms in the main industry) and μ_i is a normal error term. This equation uses data from both diversifying firms and matching specialized firms.

The fact that the diversification performance, $Y_{i,}$ is only observed if $DIV_i = 1$ may lead to endogeneity bias from self-selection. Heckman´s method controls for this bias by meeting the two previously cited identification conditions, as well as including the *Inverse Mills Ratio* (λ_i) as an additional regressor in the regression equation. The *Inverse Mills Ratio* approximates the likelihood of diversification in each firm and is calculated by diverse statistical programmes using estimates obtained from the selection equation. After incorporating this correction, the final regression equation is:

$$\boldsymbol{Yi = \alpha + \beta 1 Di + \beta 2\ Xi + \theta \lambda i + \eta i} \quad \textbf{(corrected regression equation)} \qquad \textbf{(4)}$$

4. RESULTS

Table 1 presents means, standard deviations and correlation coeficients for all variables used in this study. As shown, the estimated correlation between diversification and change in ROA is equal to 0.25. Thus, it is possible to conclude that the association between diversification and firm performance is not too strong. Although some variables in regression equations show a high correlation, the examination of variance inflation factors (VIFs) indicates no evidence of multicolineality.

[15] This study also tests for potential curvilinearity in the diversification-performance relationship by including a measure of *Diversification Squared* (D_i^2) in the regression equation.

Table 1
Means, standard deviations and pairwise correlations

Selection equation (*N = 236*)	Mean	S.D.	1	2	3	4	5	6	7
1.Diversification (dummy)	0.50	0.50	1.00						
2.Corporate control	1.21	0.41	0.04	1.00					
3.Insiders proportion	13.5	26.6	0.22***	0.49***	1.00				
4.Log of total assets	17.04	1.23	-0.24***	-0.08	-0.05	1.00			
5.CAPEX/Sales	0.45	0.70	-0.06	0.11	0.05	0.02	1.00		
6.Export activity	0.57	0,49	0.18**	-0.10	0.17**	-0.02	0.05	1.00	
7.Industry ROA	-1.84	22.85	-0.16**	-0.09	-0.07	0.19**	-0.07	-0.05	1.00
8.Fraction diversified firms	6.27	3.19	0.15*	0.03	0.01	0.02	-0.11†	0.02	-0.03

Regression equation (*N =118*)	Mean	S.D.	1	2	3	4	5	6	7	8
1.Sales Growth	0.68	0.99	1.00							
2.Change in ROA	0.17	1.46	0.05	1.00						
3.Diversification	0.75	0.41	0.33***	0.25**	1.00					
4.Corporate control	1.23	0.42	-0.06	-0.20**	-0.14	1.00				
5.Insiders proportion	19.39	31.28	-0.11	0.45***	0.08	0.35***	1.00			
6.Log of total assets	16.74	1.16	-0.29**	-0.05	-0.13	-0.01	0.04	1.00		
7.CAPEX/Sales	0.41	0.57	-0.05	-0.04	-0.03	0.03	-0.07	0.05	1.00	
8.Export activity	0.67	0.41	0.05	0.34***	0.05	-0.09	0.27**	0.09	-0.01	1.00
9.Industry ROA	-5.60	20.94	0.07	0.02	0.01	-0.11	-0.01	0.06	-0.12	-0.03

†p < 0.10; * p < 0.05; ** p < 0.01; *** p < 0.001

This study applies a hierarchical regression analysis in order to test the performance effects of diversification in the second stage of Heckman´s method[16]. Table 2 summarizes regression results on the relationship between diversification and firm profitability. All models indicate that the coefficient of the λ_i variable is not significant, indicating the absence of sample selection bias.

Results for control variables are reported in Model 1. As shown, the coefficients for corporate control, the proportion of insiders on the board of directors, and export activity are significantly associated with change in ROA. Specifically, these variables explain 39.3 percent of the variance (*p>0.001*). Model 2 reflects the linear effect of diversification. As predicted, this strategy by itself appears to have a substantial positive impact on firm profitability (*p>0.05*). In particular, the addition of the diversification variable in Model 2 contributes significantly to the prediction of change in ROA, producing a significant ΔF equal to 3.476 and explaining an additional 1.9 percent of the variance (*p>0.05*). Model 3 includes the quadratic term of diversification as well. As shown, its coefficient is not statistically significant. This means that the diversification-profitability relationship is not curvilinear, but linear. Therefore, although the effect size is not quite as strong as expected, the results obtained offer support for Hypothesis 1 when performance is measured in terms of firm profitability.

[16] Results of first-stage probit regression predicting propensity to diversify are reported in the Appendix I, which reveals significant relationships for several variables with the choice to diversify. The *Inverse Mills Ratio* (λ_i) calculated using estimates obtained from this selection equation is introduced as a control variable in regression equations.

Table 2
Second-stage regression predicting the effect of diversification on firm profitability
(Dependent variable: Change in ROA)

Variables	Model 1	Model 2	Model 3[a]
Constant	2.29 (2.07)	1.29 (2.10)	1.68 (2.06)
Corporate control	-1.25*** (0.30)	-1.14*** (0.30)	-1.14*** (0.31)
Insiders proportion	0.02*** (0.01)	0.02*** (0.01)	0.02*** (0.01)
Log of total assets	-0.07 (0.14)	-0.04 (0.14)	-0.04 (0.14)
CAPEX/Sales	0.06 (0.20)	0.07 (0.20)	0.07 (0.20)
Export activity	0.45† (0.30)	0.44† (0.29)	0.45† (0.30)
Industry ROA	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
Diversification (D_i)		0.52* (0.26)	0.51† (0.28)
Diversification Squared (D_i^2)			0.04 (0.44)
Inverse Mills Ratio (λ_i)	-0.29 (0.72)	-0.34 (0.71)	-0.33 (0.72)
Wald[a]	$\chi2$ (12) = 70.40***	$\chi2$ (13) = 75.31***	$\chi2$ (14) = 75.43***
R^2	0.393	0.412	0.412
ΔR^2	0.393	0.019	0.000
ΔF	10.166***	3.476*	0.015

[a] The diversification variable (D_i) used in the quadratic term (D_i^2) is centered to avoid multicolinearity problems. Results are similar if uncentered.

[b] Wald test is a χ^2 test of all coefficients in the regression model, except the constant, are equal to 0 (Heckman, 1979).

Values are unstandardized coefficients, with standard errors in parentheses
† $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

5. CONCLUSIONS

The main purpose of this study is to provide new empirical evidence on the impact of diversification on firm performance by using a sample of 236 large Spanish firms over the 1997 to 2001 period. The choice of this time period is interesting in that it allows us to consider the new conditions of the environment in which firms have operated over the last decade and, hence, the strategic changes that they have made with the aim of adapting to such conditions. To that end, the study justifies and tests one hypothesis that proposes that the diversification strategy has a positive linear effect on firm profitability during the period of study. This hypothesis is tested by using Heckman´s two-stage method.

Our hypothesis receives total support when performance is measured in terms of change in ROA. On the one hand, results have demonstrated that the impact of diversification on this profitability measure has been positive and significant. Over the period analyzed, in order to adapt to the new business environment, large Spanish firms showed a clear tendency towards greater diversification, especially towards moderate levels of diversification (Sánchez-Bueno *et al*., 2006). Results indicate that, as a consequence of this change in the strategic tendency, Spanish firms were able to improve firm profitability. Previuos studies such as Grant *et al*., (1988), Tallman and Li (1996), and Geringer *et al.* (2000), by using SIC code-based measures

of level of diversification on British, American, and Japanese samples respectively, also find that an intermediate level of diversification is positively related to profitability[17]. The main explanations for our result could be the access to additional profits by exploiting multiple advantages across different businesses, without having to bear the costs of maintaining high levels of diversification (Palich *et al*., 2000), and the fact that in more recently developed economies, such us the Spanish economy, diversifiers may create more value and be more sucessful because of the presence of greater market inefficiencies (Chakrabarti *et al.*, 2007). On the other hand, results have also shown that, according to the study by Palich *et al.* (2000), increasing diversification is associated with proportional increases in firm ROA, when using a sample that hardly includes firms with high levels of diversification.

Overall, it is possible to conclude that the tendency towards greater diversification shown by large Spanish firms during the last decade seems to have favourably affected firm performance. Indeed, results indicate that the strategic change from low to moderate levels of diversification has been translated into profitability benefits over recent years; that is, as long as diversification stays within the scope of the resources and capabilities, it will provide increased benefits for firms in terms of ROA.

This study also shows that the Spanish context has changed over recent years in a way that has positively influenced the diversification-performance relationship. Historically Spanish firms have shown a low level of diversification and previous empirical evidence has hardly been able to demonstrate the positive effects of this strategy. During the period studied, the strategic changes from low to moderate levels of diversification have mainly been a defensive reaction to a perceived threat in a highly competitive environment; that is, the outcome of a rational benefits-seeking behaviour for the adaptation and survival in a new business context. For this reason, benefits derived from this strategy by Spanish firms seem to have been greater than costs associated with its implementation, with the consequent improvement in firm profitability. The results obtained support the notion that as environmental conditions fluctuate, the level of diversification also seems both to vary and to have varying effects on performance[18] (Geringer *et al.*, 2000; Mayer and Whittington, 2003; Singh *et al.*, 2007; Chakrabarti *et al.*, 2007).

The findings of this study may have implications for research. First of all, our results, combined with the results of previous works for an earlier period (Suárez-González, 1994; Menéndez and Gómez, 2000; Ramírez and Espitia, 2002), suggest that the performance effects of diversification in the Spanish economy varies over time as a consequence of possible strategic changes in response to changing environment. Thus, it is possible to deduce that contextual variations within countries may help explain the pattern of inconsistent results in this research stream. Second, our findings suggest that it may have substantial benefits for research on the diversification-performance link in a particular country from systematic replication studies across different periods characterized by different business contexts.

From a practical perspective, the results of this study constitute a new contribution to the debate regarding the effects of diversification on firm performance. When in international economic and business forums, in the presence of an extremely fierce and competitive environment, it is common to hear about the urgent need for firms to engage in strategies to improve their competitiveness, the findings of this study demonstrate that the change from the low to moderate range of diversification may be an appropriate option for attaining such a goal via an increase in firm profitability.

[17] These studies did not directly test for the linear model. In particular, they tested for the curvilinear effect of diversification on firm performance from samples that included firms with different levels of diversification (low, moderate and high), and all of them found that moderate degrees of diversification often predict higher performance in terms of ROA.

[18] As well as the analysis of a different period and business context, other reasons for further explaining the discrepancies in results between this study and previous empirical research in Spain may be the focus on the initial decisión to diversify, the employment of different measures of diversification and performance, the use of different sources of data, the application of different methods, or the unequal distribution of the sample by industries.

REFERENCES

ARNOULD, R. J. (1969) Conglomerate growth and profitability, in *Economics of Conglomerate Growth* (Ed.) L. Garoian, Oregon State University, Corvallis, pp. 72-80.

BERGER, P. AND OFEK, E. (1995) Diversification effect on firm´s value, *Journal of Financial Economic*s, 37, 39-65.

CAMPA, J. M. AND KEDIA, S. (2002) Explaining the diversification discount, *Journal of Finance*, 57, 1731-1762.

CHAKRABARTI, A., SINGH, K. AND MAHMOOD, I. (2007) Diversification and performance: Evidence from East Asian firms, *Strategic Management Journal*, 28, 101-120.

COLPAN, A. M. AND HIKINO, T. (2005) Changing economic environments, evolving diversification strategies, and differing financial performance: Japan´s largest textile firms, 1970-2001, *Industrial and Corporate Change,* 14, 897-940.

COTTRELL, T. AND NAULT. B. (2004) Product variety and firm survival in the microcomputer software industry, *Strategic Management Journa*l, 25, 1005-1025.

DENIS, D. J., DENIS, D. K. AND SARIN, A. (1997) Agency problems, equity ownership and corporate diversification, *Journal of Finance*, 52, 135-160.

DENIS, D. J., DENIS, D. K. AND SARIN, A. (1999) Agency theory and the influence of equity ownership structure on corporate diversification strategies, *Strategic Management Journal*, 20, 1071-1076.

DESAI, A.; KROLL, M. AND WRIGHT, P. (2005) Outside board monitoring and the economic outcomes of acquisitions: a test of the substitution hypothesis, *Journal of Business Research,* 58, 926-934.

FRAQUELLI, G., PIACENZA, M. AND VANNONI, D. (2004) Scope and scale economies in multi-utilities: Evidence from gas, water and electricity combinations, *Applied Economics*, 36, 2045-2057.

GERINGER, M., TALLMAN, S. AND OLSEN, D. (2000) Product and international diversification among Japanese multinational firms, *Strategic Management Journal*, 21, 51-80.

GORT, M. (1962) *Diversification and Integration in American Industry*, Princeton University Press, Princeton.

GRANT, R. M. AND JAMMINE, A. P. (1988) Performance differences between Wrigley/Rumelt strategic categories, *Strategic Management Journal*, 9, 333-346.

GRANT, R. M., JAMMINE, A. P. AND THOMAS, H. (1988) Diversity, diversification, and profitability among British manufacturing companies, 1972-1984, *Academy of Management Journal*, 31, 771-801.

HADLOCK, C., RYNGAERT, M. AND THOMAS, S. (2001) Corporate structure and equity offerings: Are there benefits to diversification?, *Journal of Business*, 74, 613-635.

HARRIS, M., KRIEBEL, C. AND RAVIV, A. (1982) Asymmetric information, incentives and intrafirm resource allocation, *Management Science*, 28, 604-620.

HECKMAN, J. J. (1979) Sample selection bias as a specification error, *Econométrica,* 47, 153-161.

HOSKISSON, R. E. AND HITT, M. A. (1988) Strategic control systems and relative R&D Investment in large multiproduct firms, *Strategic Management Journal*, 9, 605-621.

JACQUEMIN, A. AND BERRY, C. (1979) Entropy measure of diversification and corporate growth, *Journal of Industrial Economics*, 27, 359-370.

JENSEN, M. C. AND ZAJAC, E. J. (2004) Corporate elites and corporate strategy, *Strategic Management Journal*, 25, 507-524.

KIM, H., HOSKISSON, R. E. AND WAN, W. P. (2004) Power dependence, diversification and performance in keiretsu member firms, *Strategic Management Journal,* 25, 613-636.

LAMONT, O. AND POLK, C. (2002) Does Diversification Destroy Value? Evidence from the Industry Shocks, *Journal of Financial Economics,* 63, 51-77.

LEWELLEN, W. (1971) A pure financial rationale for the conglomerate merger, *Journal of Finance,* 26, 521-545.

LUBATKIN, M. AND CHATTERJEE, S. (1994) Extending modern portfolio theory into the domain of corporate diversification: Does it apply?, *Academy of Management Journal*, 37, 109-136.

MAJD, S. AND MYERS, S. (1987) Tax asymmetries and corporate income tax reform, *Journal of Financial Economics*, 12, 147-175.

MARKHAM, J. W. (1973) *Conglomerate enterprise and economic performance*, Harvard University Press, Cambridge.

MARKIDES, C. C. (1992) Consequences of corporate refocusing: Ex ante evidence, *Academy of Management Journal*, 35, 398-412.

MARKIDES, C. C. AND WILLIAMSON, P. J. (1994) Related diversification, core competences and corporate performance, *Strategic Management Journal*, 15, 149-165.

MAYER, M. AND WHITTINGTON, R. (2003) Diversification en context: A cross-national and cross-temporal extension, *Strategic Management Journal*, 24, 773-781.
MENÉNDEZ, E. AND GÓMEZ, S. (2000) La estrategia de diversificación: creación o destrucción de riqueza, *Cuadernos de Economía y Dirección de la Empresa*, 7, 383-400.
MEYER, M., MILGROM, P. AND ROBERTS, J. (1992) Organizational prospects, influence costs, and ownership changes, *Journal of Economics and Management Strategy*, 1, 9-35.
MILLER, D. J. (2004) Firms´ technological resources and the performance effects of diversification: A longitudinal study, *Strategic Management Journal*, 25, 1097-1119.
MILLER, D. J. (2006) Technological diversity, related diversification and firm performance, *Strategic* Management Journal, 27, 601-619.
MONTGOMERY, C. A. (1982) The measure of firm diversification: some new empirical evidence, *Academy of Management Journal*, 25, 299-307.
NAYYAR, P. (1992) On the measurement of corporate diversification strategy: Evidence form large U.S. service firms, *Strategic Management Journal*, 13, 219-235.
PALEPU, K. (1985) Diversification strategy, profit performance and the entropy measure, *Strategic Management Journal*, 6, 239-255.
PALICH, L. E., CARDINAL, L. B. AND MILLER, C. C. (2000) Curvilinearity in the diversification-performance linkage: An examination of over three decades of research, *Strategic Management Journal*, 21, 155-174.
RAJAN, R., SERVAES, H. AND ZINGALES, L. (2000) The cost of diversity: The diversification discount and inefficient investment, *Journal of Finance*, 55, 35-80.
RAMANUJAN, V. AND VARADARAJAN, P. (1989) Research on corporate diversification: A synthesis, *Strategic Management Journal*, 10, 523-551.
RAMÍREZ, M. L. AND ESPITIA, M. A. (2002) The impact of product diversification strategy on the corporate performance of large Spanish firms, *Spanish Economic Review*, 4, 119-137.
RHOADES, S. A. (1973) The effect of diversification on industry profit performance in 241 manufacturing industries: 1963, *Review of Economics and Statistics*, 55, 9-24.
RUMELT, R. P. (1974) *Strategy, structure, and economic performance*, Harvard Business School Press, Boston.
SANCHEZ-BUENO, M. J., GALÁN, J. I. AND SUAREZ-GONZÁLEZ, I. (2006) Evolución de la estrategia corporativa y la estructura organizativa en los grandes grupos españoles, *Universia Business Review*, 11, 22-35.
SCHERER, F. M. (1980) *Industrial Market Structure and Economic Performance*. Rand McNally, Chicago.
SIMANAVIČIENĖ, Z. AND DAGILIENĖ, L. (2003) The benefits of business combinations, *Transformations in Business and Economics*, 2, 105-120.
SINGH, M., NEJADMALAYERI, A. AND MATHUR, I. (2007) Performance impact of business group affiliation: An analysis of the diversification-performance link in a developing economy, *Journal of Business Research*, 60, 339-347.
STEIN, J. (1997) Internal capital markets and the competition for corporate resources, *Journal of Finance*, 52, 111-134.
STULZ, R. (1990) Managerial discretion and optimal financing policies, *Journal of Financial Economics*, 26, 3-27.
SUÁREZ-GONZÁLEZ, I. (1994) Estrategia de diversificación y resultados de la empresa española, *Revista de Economía Aplicada*, 2, 103-128.
TALLMAN, S. AND LI, J.T. (1996) The effects of international diversity and product diversity on the performance of multinational firms, *Academy of Management Journal*, 39, 179-196.
TANRIVERDI, H. AND VENKATRAMAN, N. (2005) Knowledge relatedness and the performance of multi-business firms, *Strategic Management Journal*, 26, 97-119.
VILLALONGA, B. (2004) Does Diversification Cause the 'Diversification Discount'?, *Financial Management*, 33, 5-32.
WERNERFELT, B. AND MONTGOMERY, C. A. (1988) Tobin´s *q* and the importance of focus in firm performance, *American Economic Review*, 78, 246-250.
WILHELMSSON, L. AND MCQUEEN, D. (1999) Product diversity and performance: a macro study of 20 large Swedish industrial corporations, *Scandinavian Journal of Management*, 15, 43-64.
WILLIAMSON, O. E. (1985): *The economic institutions of capitalism*, Free Press, New York.
WOOLDRIGDE, J. (2002): *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge.

APPENDIX I

First-stage probit regression predicting propensity to diversify

Variables	Coefficients	S.E.	z-Statistic
Constant	4.35**	1.38	3.15
Corporate control	-0.37	0.27	-1.39
Insiders proportion	0.01**	0.01	2.99
Log of total assets	-0.28***	0.08	-3.56
CAPEX/Sales	-0.12	0.12	-0.97
Export activity	0.40*	0.18	2.20
Industry ROA	-0.02*	0.01	-2.06
Fraction diversified firms	0.07**	0.03	2.54
Number of total observations			236
Number of censured observations			118
Log-likelihood test statistic			-139.46***
Pseudo-R^2			0.15

† $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

LAS TRANSFERENCIAS COMUNITARIAS A LA AGRICULTURA ESPAÑOLA EN EL CONTEXTO DE LA OCDE

SILVIA ANDRÉS GONZÁLEZ-MORALEJO
e-mail: silangle@upvnet.upv.es

Departamento de Economía y Ciencias Sociales
Grupo de Economía Internacional. Universidad Politécnica de Valencia

Resumen

La política agraria en España ha estado sujeta en los últimos años a cambios sustanciales, que han implicado modificaciones en el nivel y los mecanismos de apoyo al sector. Tras la "Revisión Intermedia", los instrumentos de la Política Agraria Común (PAC) cada vez se aplican más de manera diferenciada a nivel nacional y regional. El presente artículo ofrece una valoración regional de las transferencias comunitarias a la agricultura española, estableciendo un sistema de seguimiento y evaluación del nivel global, la distribución de ese apoyo y la ubicación de España y sus regiones en el marco de la OCDE. El trabajo adopta como base metodológica los avances realizados en la estimación de indicadores de apoyo a la agricultura por la OCDE.

Palabras Clave: apoyo a la agricultura, transferencias regionales, agricultura española, Política Agraria Común.
Área Temática: Economía Internacional y de la Integración.

Abstract

The agricultural policy in Spain has been involved in substantial changes which have implied some adjustments on the levels of support and its instruments. After the "Mid Term Review" of the Common Agricultural Policy (CAP) more and more those instruments are applied with varied intensity at the national and regional level. This paper offers a regional valuation of the European support to the Spanish agriculture by establishing a system of monitoring and evaluation of the support, its distribution and the position of Spain into the OCDE framework. This work makes use of the progress undertaken by OECD in defining a methodology to estimate support to the agricultural sector.

Thematic Area: International and Integration Economics.

1. INTRODUCCIÓN

En los últimos años, las políticas agrarias en los países de la OCDE se han visto sometidas a reformas. En la Unión Europea (UE), los niveles y las formas de apoyo a la agricultura han venido influidos por las reformas recientes de la Política Agraria Común (PAC). Los cambios fundamentales que han tenido lugar, acentuados tras las mencionadas reformas (reforma "MacSharry" en 1992, la Agenda 2000 en 1999 y la "Revisión Intermedia" en 2003), han sido, por un lado, el menor énfasis en las transferencias por la vía del apoyo a los precios, que se reemplazan progresivamente por transferencias directas a los productores; y por el otro, la aplicación de instrumentos de política cada vez más diferenciados a nivel nacional y regional. En otros países de la OCDE se ha dado una transformación similar.

En muchos países del área OCDE, las ayudas a la agricultura se han basado en la superficie de la explotación, o en criterios históricos, como el número de hectáreas, de animales o la renta de los productores. En algunos casos se requiere producción para recibir la ayuda, en otros no se requiere producción o la ayuda se vincula a la obtención de bienes no comerciales. Incluso, se exigen otros criterios que el productor debe reunir para beneficiarse del apoyo, como

restricciones en el uso de inputs o condiciones de carácter medio ambiental. Una tendencia general es el desacoplamiento del apoyo de los niveles de producción, por lo que los agricultores reciben ayudas que no dependen de sus decisiones de cultivo. En consecuencia, las diferencias nacionales en los niveles de apoyo pueden ser apreciables entre los países de la OCDE, como también las especializaciones en los objetivos o en los instrumentos.

Este contexto de aplicación diferenciada de las políticas agrarias y de cambio en los instrumentos elegidos de apoyo al sector suscita la necesidad de disponer de un sistema de información sobre el apoyo a la agricultura, que permita un seguimiento y evaluación de los siguientes aspectos: el nivel global, los mecanismos de apoyo al sector, la distribución de ese apoyo y la ubicación de nuestro país en el marco de la OCDE. La construcción de estos sistemas de información debería suponer una condición previa a la formulación de mecanismos racionales o modelos de toma de decisiones en políticas agrarias, habiendo ya algunos antecedentes en la literatura sobre la economía agraria española (Atance y Gómez-Limón, 2004; Philippidis, 2005).

El presente artículo ofrece una valoración regional de las transferencias comunitarias a la agricultura en España, basada en una adaptación de la metodología de la OCDE para el cálculo de indicadores de apoyo a la agricultura. La idea es suministrar unas estimaciones regionalizadas, a partir de las cuales obtener el total nacional, que sean compatibles con los resultados proporcionados por la citada organización para la UE y para otros países de la OCDE[1]. De esta manera, se propone un sistema de seguimiento de los efectos regionales de las reformas de la PAC en España que sea comparable con las estimaciones realizadas por la OCDE para sus Estados miembros.

Con base en las consideraciones expuestas, el presente artículo persigue los siguientes objetivos:

a. Estimar para España y cada una de las CCAA los principales indicadores de apoyo al productor que utiliza la OCDE, en el periodo 2002-2006: Estimación del Apoyo al Productor, Estimación de los Servicios Generales a la Agricultura y Estimación del Apoyo Total, distinguiendo el cálculo de sus componentes: apoyo al precio de mercado y transferencias presupuestarias.
b. Analizar, para España y cada una de las CCAA, la composición del apoyo al productor por tipos de medidas (sostenimiento de precios, pagos directos derivados de la política de precios y mercados, medidas de desarrollo rural...) y su financiación por consumidores o contribuyentes.
c. Comparar, para el año 2006, la Estimación del Apoyo a la Agricultura española otorgada por las transferencias derivadas de la PAC con el apoyo total a la agricultura en otros países de la OCDE.

2. METODOLOGÍA

2.1 Antecedentes

Como antecedentes de la estimación del apoyo a la agricultura a escala regional podemos citar los trabajos emprendidos por la Comisión Europea tras las reformas de los 90, con resultados para 1994 y 1996 (European Commission, 2001) y para 1991 y 1995 (Tarditi y Zanias, 2001). Estos trabajos utilizan el concepto de Estimación del Apoyo al Productor (*Producer Support Estimate*) introducido por la OCDE (2007) como también lo hace el artículo de Zanias (2002) orientado al cálculo de las transferencias al nivel de Estado miembro. Más recientemente, el proyecto *European Spatial Planning Observation Network* analizó los impactos territoriales de la PAC para 1999, aunque sólo tiene en cuenta los productos continentales y el cálculo de las transferencias del "primer pilar" de la PAC (ayudas a precios y

[1] Es pertinente recordar que los cálculos de la OCDE sólo proporcionan estimaciones del apoyo a la agricultura para la UE en su conjunto y no ofrecen desagregación nacional o regional como lo hace el presente artículo para el caso español. Ahí radica el valor añadido que se aporta en este trabajo.

mercados). Otros estudios han analizado las transferencias en países con organización administrativa federal, como son los casos de Suiza (Walkenhorst, 2003), Alemania (Anders et al., 2004; Hansen, 2005) y Estados Unidos (Sumner y Brunke, 2003).

Para el caso de España, un antecedente del estudio del gasto público en las agriculturas de las CCAA fue realizado en García et al. (1994). En ese mismo año, otro antecedente es el esfuerzo realizado para la elaboración del *Libro Blanco de la Agricultura y el Desarrollo Rural* elaborado por el Ministerio de Agricultura, Pesca y Alimentación (MAPA, 2004). Los resultados de estos estudios han permitido realizar análisis sobre la repercusión de los apoyos públicos de las distintas administraciones en las agriculturas regionales españolas, como es el caso de los efectos de la reforma de la PAC y la Ampliación (Compés y García, 2005).

Un antecedente del presente trabajo se encuentra en García y Andrés (2007), en el que se realiza un ensayo de estimación del apoyo de la PAC a las distintas regiones españolas mediante la aplicación de los indicadores de la OCDE. En este ejercicio, el periodo objeto de estudio se limitó a los años 2002-2003, pero se discutían distintas opciones de cálculo de productos no incluidos en las estimaciones de la OCDE. El presente trabajo ofrece una actualización de la metodología de regionalización de los apoyos públicos a la agricultura, extendiendo la estimación a un periodo más amplio 2002-2006 y revisando el contexto internacional en que dicha estimación se encuadra.

Este artículoÿÿoma como referencia ÿÿs baÿÿs metodolóÿÿÿÿs planteadas por la OCDE (2007) para el cálculo de indicadores de apoyo a la agricultura[2]. Hemos partido de esta base para plantear una estimación del apoyo a la agricultura en las CCAA que sea comparable con las estimaciones realizadas por la OCDE para sus Estados miembros.

2.2. Conceptos de Apoyo al Sector Agrario según la OCDE

La Estimación del Apoyo al Productor es un indicador adoptado por la OCDE para medir la ayuda a la agricultura. Denotaremos el indicador mediante las siglas PSE (*Producer Support Equivalent* sería el término anglosajón). El PSE es un indicador del valor anual de las transferencias brutas a nivel de explotación agraria, originadas por políticas de apoyo a la agricultura, independientemente de su naturaleza, objetivo o impacto en la producción agraria, las rentas o el medioambiente. A los efectos de poder comparar los niveles de la ayuda en diversos países o regiones, el valor monetario total de estas formas de ayuda se expresa normalmente en términos de ratios, como el PSE por Unidad de Trabajo Agrario (UTA), por Hectárea cultivada, o en términos de porcentaje del margen bruto total o del valor bruto de los ingresos de la agricultura. El PSE se calcula para productos individuales y puede ser sumado a los efectos de obtener un PSE nacional o regional.

El PSE es un indicador discutido, cuyo método de cálculo debe ser interpretado con reservas (Berthelot, 2004; Wise; 2004; Tangermann, 2005). Las críticas tienen que ver con tres aspectos, principalmente. En primer lugar, en el plano conceptual el PSE asume que la oferta agraria no varía con respecto a los cambios en los precios. De este modo, el PSE no es una medida exacta de los cambios en el bienestar de los productores (excedente del productor). En segundo lugar, las transferencias derivadas de las políticas de protección comercial dependen de estimaciones de costes de oportunidad del comercio internacional basados en precios de mercados representativos que pudieran, a su vez, estar distorsionados por el efecto de las medidas comerciales de los grandes productores (aranceles, subsidios a la exportación). En tercer y último lugar, el PSE mide la transferencia total a los productores pero no aporta detalles sobre la distribución de dicha transferencia entre los propietarios de los distintos factores de producción. A ese respecto, por ejemplo, una parte de dichas transferencias pudiera capitalizarse en mayores precios de la tierra o concentrarse en determinados sistemas productivos.

No obstante, entendemos que el PSE es un indicador cuya utilidad se basa en su sencillez metodológica y en la posibilidad de realizar comparaciones internacionales. Reconociendo sus

[2] A su vez, los indicadores de la OCDE se basan en el concepto de Equivalente de Subsidio al Productor propuesto por Corden (1971) e introducido por primera vez como medida del apoyo a la agricultura por Josling (1979). Veáse Reig (1991) para una discusión de las aplicaciones del concepto.

limitaciones, es importante observar que el énfasis del PSE radica no sólo en su cuantía total, sino en su evolución y distribución entre los distintos instrumentos de apoyo, cuyo grado de distorsión del mercado es variable.

Tan importante como la magnitud de la transferencia lo son "las formas de transferencia". Hay formas de ayuda que distorsionan más el comercio internacional que otras (Blandford, 2005). En general, se tiende a aceptar que las transferencias vía precios comportan un mayor grado de distorsión. La metodología de la OCDE permite clasificar las ayudas por tipos, considerando por un lado las transferencias derivadas de intervenciones de precios, y por el otro, los grupos de transferencias relacionadas con desembolsos o pagos directos a los productores. Así, el PSE recoge dos tipos de transferencias, cuyo cálculo para este trabajo se explica en los siguientes epígrafes:

- Las que mantienen precios internos a los productores más altos (y ocasionalmente más bajos) que los mundiales, conformando el componente denominado *apoyo al precio de mercado*;
- Las que proporcionan pagos a los productores, basados en criterios como el volumen producido, la cantidad de inputs usada, el número de animales, el área cultivada o la renta recibida, conformando el componente denominado *transferencias presupuestarias.*

La OCDE considera también la Estimación de los Servicios Generales a la Agricultura (GSSE, *General Services Support Estimate* sería el término anglosajón), que recoge el valor monetario de las transferencias anuales a la agricultura (que siendo gasto público no está incluido en el PSE porque no tiene carácter individual) proporcionadas para la provisión de servicios de investigación, desarrollo rural, inspección, marketing, promoción, etc. A partir de los indicadores definidos anteriormente, la OCDE obtiene la Estimación de la Ayuda Total (TSE, *Total Support Estimate* sería el término anglosajón). Se trata de un indicador del valor monetario anual de todas las transferencias, derivadas de medidas de política, que apoyan al sector agrícola. El TSE es la suma de las transferencias a productores (PSE), las actuaciones de carácter estructural (GSSE) y las transferencias de los contribuyentes a los consumidores. El TSE porcentual expresa el apoyo general en porcentaje del PIB. Una descripción de estos indicadores y su aplicación a la UE se encuentran en los documentos que la propia OCDE ofrece en su página web[3].

A continuación ofrecemos un detalle algebraico sobre la metodología de cálculo aplicada para la obtención de los principales indicadores (PSE, GSSE y TSE). En el siguiente apartado se explica con mayor profundidad cómo han sido estimados u obtenidos los ítems MPS y BP.

- **MPS:** Apoyo a Precios de Mercado **(estimación)**

- **BP:** Pagos con Cargo a Presupuestos **(obtención directa a partir del FEGA (varios años) y del MCA[4] 2000-2006)**

  BP_1: Transferencias directas al productor
  BP_2: Transferencias de carácter estructural o GSSE
  BP_3: Transferencias de contribuyentes a consumidores

$$\mathbf{PSE = MPS + BP_1}$$

$$\mathbf{TSE = PSE + BP_2\ (GSSE) + BP_3}$$

La ayuda a la agricultura también se mide en el contexto de las negociaciones comerciales multilaterales. En el marco del Acuerdo sobre Agricultura de la Ronda Uruguay (AsA) toda ayuda interna en favor de los productores agrícolas está sujeta a normas. El AsA implica dos

[3] http//:www.oecd.org. La medición del apoyo recibido por los agricultores se ha institucionalizado en las publicaciones anuales *Agricultural Policies, Markets and Trade. Monitoring and Outlook* y *Agricultural Policies in OCDE countrie. Monitoring and Evaluation.*
[4] Marco Comunitario de Apoyo 2000-2006.

tipos de compromisos en materia de ayuda interna: uno cualitativo y el otro cuantitativo. El compromiso cualitativo establece una definición de las políticas de ayuda interna exceptuadas de los compromisos de reducción, mientras que el compromiso cuantitativo establece listas de compromisos que limitan las subvenciones a los productos agropecuarios. A expensas de los acuerdos que se adopten en la actual Ronda Doha de negociaciones multilaterales de la Organización Mundial de Comercio, el indicador utilizado desde la Ronda Uruguay (RU) ha sido la Medida Global de la Ayuda (MGA). Adicionalmente, existe un tipo de pagos directos exentos de compromisos de reducción, en el marco del llamado "compartimento azul" de la RU[5]. El Recuadro 1 explica con mayor detalle los compartimentos de ayuda interna de la RU. El Recuadro 2 recoge una comparación entre los indicadores PSE y MGA.

Recuadro 1. El apoyo a la agricultura según la OMC

Medidas no sujetas a compromisos de reducción

La Caja Verde. Las medidas de la "Caja Verde " figuran en el Anexo 2 del AsA. El requisito fundamental para que queden exceptuadas de los compromisos de reducción es que no tengan efectos de distorsión del comercio ni efectos en la producción o, a lo sumo, que los tengan en grado mínimo. La ayuda en cuestión debe prestarse por medio de programas gubernamentales financiados con fondos públicos (incluyendo ingresos fiscales sacrificados) que no impliquen transferencias de los consumidores, y no debe tener el efecto de prestar ayuda en materia de precios a los productores. En el marco de la OMC, los desembolsos en estas medidas eximidas pueden ser aumentados sin limitación alguna. La Caja Verde se aplica tanto a los Miembros que son países desarrollados como a los países en desarrollo. La lista de la Caja Verde incluye las siguientes medidas:

- servicios de apoyo, tales como investigación, lucha contra plagas y enfermedades, formación, divulgación, inspección, servicios de comercialización y promoción, y servicios de infraestructura;
- constitución de existencias públicas con fines de seguridad alimentaria;
- ayuda alimentaria interna; y
- pagos directos a los productores tales como la ayuda a los ingresos desconectada, programas de seguro de los ingresos y de red de seguridad de los ingresos, socorro en casos de desastres naturales, programas de retiro de productores y de detracción de tierras y otros recursos, ayudas a la inversión, programas ambientales, y programas de asistencia regional.

Medidas sujetas a compromisos de reducción

La Caja Ámbar. La ayuda prestada en el marco de políticas no exceptuadas está sujeta a los compromisos de reducción. Los compromisos de reducción se expresan en términos de la "Medida Global de la Ayuda Total o MGA Total", que es la suma de los gastos en ayuda interna no exenta correspondiente a todos los productos y las políticas. Para estimar el nivel de la ayuda prestada al sector agropecuario, el AsA menciona cuatro medidas de ayuda diferentes, que son:

- MGA referida a productos específicos: la cuantía total de la ayuda otorgada a cada producto agropecuario de base (por ejemplo, sostenimiento de los precios, pagos directos, etc.).
- MGA no referida a productos específicos: la cuantía total de la ayuda otorgada por políticas dirigidas al sector agropecuario en conjunto, excluida la ayuda a productos específicos.
- Medida de la ayuda equivalente (MAE): ayuda referida a productos específicos cuyo cálculo con arreglo a la metodología de la MGA no es

[5] En las negociaciones actuales de la Ronda de Doha se está considerando otro indicador más general: Ayuda Total con Efectos de Distorsión al Comercio, que englobaría la MGA (incluidas las ayudas denominadas *de minimis*) más la caja azul.

factible.

- MGA Total*:* es la suma de toda la ayuda interna no exenta otorgada a los productores agropecuarios, obtenida sumando todas las MGA correspondientes a los productos específicos, todas las MGA no referidas a productos específicos y todas las medidas de la ayuda equivalentes.

El compartimento azul

La Caja Azul. Los pagos directos efectuados en el marco de programas de limitación de la producción (clasificados a veces como medidas de la "Caja Azul") están eximidos de los compromisos de reducción de la ayuda interna si:

- se basan en superficies y rendimientos fijos; o
- se realizan con respecto al 85 por ciento o menos del nivel de base de la producción; o
- los pagos relativos al ganado se realizan sobre un número fijo de cabezas.

En la Unión Europea, la inclusión en la caja azul de los pagos directos a cereales y oleaginosas, así como de las primas al vacuno, introducidos en la reforma de 1992, permitió que la Medida Global de Ayuda de la Unión Europea se viera muy reducida ya que los nuevos pagos quedaban fuera del compromiso de reducción. Gracias a la definición sectorial de la MGA, esos "ahorros" se han podido utilizar para apoyar otros productos agrarios de la UE, aunque el apoyo a productos concretos no puede superar el concedido en 1992 como condición para que la UE se beneficie de la cláusula de paz.

Fuente: elaboración propia.

Recuadro 2. Comparación entre las medidas del PSE y de la MGA

- El PSE incluye toda transferencia monetaria de los consumidores y contribuyentes a los productores.
- La MGA excluye algunas de las transferencias tanto de los consumidores como de los contribuyentes.
- El PSE incluye todos los pagos directos a los productores (es decir las transferencias monetarias de los contribuyentes).
- La MGA excluye algunos de éstos, por ejemplo los pagos directos (de los contribuyentes) derivados de la Caja Verde y la Caja Azul.
- El PSE incluye las transferencias monetarias implícitas de los consumidores que resultan de los obstáculos a la importación.
- La MGA, no: el sostenimiento de los precios de mercado sólo ocurre cuando existe un precio administrado.
- El PSE utiliza los precios externos corrientes (es decir, los precios vigentes cuando se calcula el PSE).
- La MGA utiliza un período de referencia fijo (1986-1988) para todos los cálculos de la MGA: de Base y Corriente.
- El PSE varía según los cambios de los precios mundiales, los tipos de cambios y los "volúmenes de ayuda".
- La MGA varía sólo según el volumen de ayuda.
- El PSE puede medir transferencias negativas a los productores, y éstas son tomadas en cuenta cuando se calcula el PSE total para el sector agrícola.
- La MGA ignora las transferencias negativas en el cálculo de la ayuda total.
- El PSE permite registrar una ayuda significativa a los cereales forrajeros como ayuda negativa para el ganado.
- La MGA no lo permite.

Fuente: elaboración propia.

2.3. Aplicación empírica para España: metodología y fuentes

Apoyo al precio de mercado (MPS)

Para la estimación del componente de apoyo al precio de mercado (MPS, *Market Price Support*) consideramos el MPS calculado por la OCDE para un conjunto de productos individualizados, que llamaremos *estándar*[6]. Para la regionalización de los apoyos a los precios de mercado hemos tomado inicialmente dichos resultados del componente de apoyo al precio de mercado en la UE calculado por la OCDE en los productos *estándar*. Una vez disponemos de los resultados agregados a nivel de Unión Europea, hemos supuesto que ese apoyo al precio de mercado interno se distribuye entre las regiones europeas de manera proporcional a su participación en la producción final agraria de la Unión[7]. Para el cálculo de dichos coeficientes de participación se requiere conocer el nivel de producción comunitario y regional de cada producto. Para la producción comunitaria hemos tomado los datos empleados por la OCDE procedentes de EUROSTAT (garantizando así la máxima coherencia con los cálculos de la OCDE), y para la producción regional hemos usado los datos ofrecidos en el Anuario de Estadística Agroalimentaria del MAPA (2003, 2004, 2005, 2006). Para las producciones regionales del año 2006 hemos recurrido a las estimaciones que ofrecen los Avances de Superficies y Producciones de Cultivos (2006, 2007) y las Encuestas de Sacrificio de Ganado (2006) del MAPA[8]. Con objeto de armonizar las unidades de medida utilizadas por las distintas fuentes hemos traducido las estimaciones de producción en valores. Para ello, hemos empleado los precios al productor que utiliza la OCDE en sus cálculos, salvo en un producto, flores y plantas, que hemos tenido que recurrir a los precios que publica el MAPA porque la OCDE no suministra el dato[9]. Por producto *estándar* se obtiene el apoyo al precio de mercado de cada CCAA aplicando los coeficientes de participación regional en el apoyo al precio de mercado comunitario calculado por la OCDE. Algebraicamente, para el cálculo del MPS *estándar* regional por producto usamos las fórmulas que siguen:

$$\mathbf{MPS_{EST}\,REG = (CP * MPS_{EST}\,UE)/100}, \text{ siendo}$$

$$\mathbf{CP = VTP\ REG/(VTP\ UE * 10.000)}, \text{ de donde}$$

MPS_{EST} REG: Apoyo al precio de mercado *estándar* regional (millones de euros)
MPS_{EST} UE: Apoyo al precio de mercado *estándar* en la UE (millones de euros)
CP: Coeficiente de participación (en porcentaje)
VTP REG: Valor total de la producción regional (en euros)
VTP UE: Valor total de la producción en la UE (en millones de euros)

La OCDE no suministra ninguna información para aceite de oliva, tabaco, algodón, frutas y hortalizas. Para subsanar este problema, la OCDE realiza una extrapolación para el cálculo del MPS total a través del incremento del MPS *estándar* en virtud de la participación de los productos *no estándar* en la producción final agraria total. Este supuesto no es inmune a la crítica pues se está suponiendo que los productos *no estándar*, típicamente mediterráneos en el caso de España, se benefician de un nivel de apoyo al precio de mercado porcentual idéntico al promedio del conjunto de productos *estándar*. En nuestro caso, y para mantener coherencia con estas orientaciones de la OCDE, hemos optado por calcular el porcentaje que los productos *estándar* representan en el valor de la producción de la rama agraria de cada CCAA, llegando así a unos coeficientes de corrección regionales con el objetivo de tener en cuenta el efecto de la especialización productiva de cada CCAA. Después se ha dividido el apoyo al precio de mercado *estándar* de cada CCAA entre su correspondiente coeficiente de corrección para

[6] Trigo, maíz, cebada, avena, arroz, soja, colza, girasol, azúcar, patatas, tomates, flores y plantas, vino, vacuno, porcino, pollo, ovino, leche y huevos.
[7] Se trata de estimar de qué parte del MPS total que recibe cada producto en la UE se están beneficiando los productores de las distintas CCAA.
[8] Por supuesto, hemos comprobado previamente que las distintas fuentes, OCDE y MAPA, son coherentes con los resultados que EUROSTAT ofrece a nivel nacional.
[9] Un ejercicio exploratorio de cálculo del PSE para frutas y hortalizas puede encontrarse en Nucifora y Sarri (1997).

obtener los valores totales de apoyo al precio de mercado. El detalle algebraico que utilizamos sería:

$$\mathbf{MPS_{TOTAL}\ REG = (MPS_{EST}\ REG * 100)/CC}, \text{ siendo}$$

$$\mathbf{CC = (VTP_{EST}\ REG * 100)/PRA\ REG}, \text{ de donde}$$

MPS_{TOTAL} REG: Apoyo al precio de mercado total en la región (incluyendo todos los productos, en millones de euros)
MPS_{EST} REG: Apoyo al precio de mercado *estándar* en la región (incluyendo los productos *estándar, en millones de euros*)
CC: Coeficiente de corrección regional (en porcentaje)
VTP_{EST} REG: Valor total de la producción regional (incluyendo los productos *estándar, en millones de euros*)
PRA REG: Producción de la Rama Agraria regional (millones de euros)

Pagos con cargo a presupuestos (BP)

Los países miembros de la OCDE acaban de consensuar cambios significativos en la clasificación de las categorías de ayuda que componen el PSE. Con estos cambios se pretende disponer de un desglose del PSE en categorías que sean coherentes con los instrumentos actuales de política agraria, para reflejar mejor la evolución y la reforma de dicha política en los países miembros. Desde septiembre de 2007, la OCDE ha incorporado la nueva clasificación en su informe *Agricultural Policies in OECD Countries: Monitoring and Evaluation.* El concepto y cálculo del PSE permanece idéntico y se mantiene el desglose del resto de indicadores. Los criterios para efectuar la clasificación de las nuevas 7 categorías, que se ordenan según su grado de desacoplamiento, se basan en *cómo están aplicadas* las políticas (¿por kg?, ¿por ha?, ¿requiere producción?...) y no en los objetivos o impacto de dichas políticas, distinguiéndose:

- Transferencias basadas en: apoyos vinculados a la producción (categoría A), a los niveles de inputs (categoría B), al área cultivada/número de animales/ingresos/rentas (categorías C, D y E), y a criterios distintos al producto (categoría F).
- Transferencias vinculadas a valores corrientes (categorías A, B, C y F) o históricos (categorías D, E y también F dependiendo de las condiciones de implementación).
- Transferencias que requieren producción (categorías C y D) o no (categoría E).

Además, la nueva clasificación incluye subcategorías, que serán aplicadas a las políticas individuales para proporcionar una mayor especificación en función de cómo estén implementadas: con o sin límite de producción, con o sin restricciones sobre los inputs, tasas fijas o variables.

En la regionalización de los apoyos de la PAC a las agriculturas regionales, la valoración de las transferencias distintas del MPS se ha abordado a partir de los Informes de Actividad del Fondo Español de Garantía Agrícola (FEGA, varios años), donde se recogen todas las transferencias del FEOGA-Garantía a las CCAA. Como toda la información debe ser suministrada de manera congruente con los conceptos utilizados por la OCDE para el cálculo de los indicadores PSE, GSSE y TSE, hemos clasificado dichas transferencias según se propone en la nueva metodología para el cálculo de los indicadores. Para completar, hemos incorporado las transferencias del FEOGA-Orientación[10] que se asocian fundamentalmente con la financiación de acciones estructurales en el sector[11]. En la estimación de las transferencias del FEOGA-Orientación se empleó el promedio anual del gasto público asociado a dicha sección del FEOGA para el periodo de programación 2000-2006. El Cuadro 1 clasifica este conjunto de transferencias públicas directas de la UE a la agricultura de las CCAA. Siguiendo la clasificación de la OCDE, dichas transferencias del presupuesto público se han estimado y agrupado como:

[10] Hay que apuntar que a partir del 1 de enero de 2007 la clasificación de Fondos agrícolas ha variado con la clasificación de los gastos de política agraria en un Fondo Europeo Agrícola de Garantía (FEAGA) y un Fondo Europeo Agrícola de desarrollo Rural (FEADER).

[11] Los conceptos dentro de la PAC que se financian a cargo de la sección Orientación se relacionan principalmente con la modernización de explotaciones y la instalación de jóvenes agricultores.

a) Apoyos directos a los productores (a incluir en el cálculo del PSE).
b) Servicios generales a la agricultura o GSSE (que son gasto público que no está incluido en el PSE por no suponer una transferencia directa al productor, pero sí en el TSE).
c) Ayudas directas al primer consumidor (que son gasto público que no está incluido en el PSE por no suponer una transferencia directa al productor, pero sí en el TSE).

Debemos advertir que en el presente trabajo estimamos las transferencias derivadas fundamentalmente de la PAC y que el indicador resulta por fuerza incompleto puesto que no recoge algunas transferencias nacionales implícitas como pueden ser las derivadas del coste del factor agua o del régimen fiscal especial para la agricultura. De todos modos, el indicador calculado para el conjunto de la UE por la OCDE tampoco recoge estos elementos por lo que nuestras estimaciones siguen siendo válidas a título de comparación.

Cuadro 1. Clasificación de las transferencias del FEOGA según el nuevo criterio de la OCDE

ESTIMACIÓN DEL APOYO AL PRODUCTOR (PSE)
A. Ayuda basada en la producción
A.1. Apoyo al precio de mercado (MPS)
A.2. Pagos basados en la producción
Ayuda a la producción (aceite de oliva)
Ayuda al almacenamiento privado (aceite de oliva)
Reembolso de gastos de almacenamiento (azúcar e isoglucosa)
Ayuda a la producción de lino textil (lino textil y cáñamo)
Ayuda a la producción de cáñamo (lino textil y cáñamo)
Ayuda a la producción (gusanos de seda)
Ayuda a la producción (tabaco)
Ayuda al almacenamiento privado de vinos y mostos (vinos y alcoholes)
Almacenamiento privado de mantequilla y nata (leche y productos lácteos)
Prima por sacrificio (vacuno)
Almacenamiento privado de carne de porcino (porcino)
B. Pagos basados en el uso de inputs
B.1. Uso de inputs variables
B.2. Formación de capital fijo
Acciones reestructuración reconversión viñedo (vinos y alcoholes)
Inversión en explotaciones agrícolas (desarrollo rural)
Modernización de las explotaciones (FEOGA-Orientación)
Instalación jóvenes agricultores (desarrollo rural)
Instalación jóvenes agricultores (FEOGA-Orientación)
B.3. Servicios agrarios
Formación (desarrollo rural)
C. Pagos basados en A/An/R/I* actuales, producción requerida
C.1. De un único producto
Ayuda suplementaria al trigo duro (cultivos herbáceos)
Prima a la calidad del trigo duro (cultivos herbáceos)
Ayuda por superficie de lino no textil (cultivos herbáceos)
Ayuda por superficie (arroz)

Ayuda al olivar (aceite de oliva)

Ayuda a la producción de uvas para pasificación (frutas y hortalizas)

Medidas especiales de ayuda al espárrago (frutas y hortalizas)

Medidas especiales a favor de la producción de avellana (frutas y hortalizas)

Ayuda por superficie (algodón)

Ayuda a la producción (lúpulo)

Prima láctea (leche y productos lácteos)

Pagos suplementarios a los productores de leche (leche y productos lácteos)

Importe adicional (leche y productos lácteos)

Prima a la vaca nodriza (vacuno)

Prima al ternero (vacuno)

Prima por extensificación (vacuno)

Pagos adicionales de las primas (vacuno)

Apicultura

C.2. De un grupo de productos

Ayuda por superficie de cultivos herbáceos (cultivos herbáceos)

Ayuda por superficie de cereales (cultivos herbáceos)

Ayuda por superficie de proteaginosas (cultivos herbáceos)

Prima a las proteaginosas (cultivos herbáceos)

Ayuda por superficie de oleaginosas (cultivos herbáceos)

Retirada de tierras (cultivos herbáceos)

Ayuda por superficie de leguminosas de grano (leguminosas de grano y forrajes)

Ayuda por superficie (cultivos energéticos)

Ayuda por superficie a los frutos de cáscara (frutas y hortalizas)

Cereales y arroz (semillas)

Gramíneas (semillas)

Leguminosas (semillas)

Oleaginosas (semillas)

Prima a la oveja y a la cabra (ovino y caprino)

Prima en zonas desfavorecidas y de montaña (ovino y caprino)

Pagos adicionales (ovino y caprino)

C.3. De todos los productos

Medidas agri-medioambientales (desarrollo rural)

Zonas desfavorecidas y sujetas a dificultades medioambientales (desarrollo rural)

D. Pagos basados en A/An/R/I* no actuales, producción requerida

E. Pagos basados en A/An/R/I* no actuales, producción no requerida

E.1. Tasas variables

E.2. Tasas fijas

Pago único

F. Pagos basados en criterios distintos al producto

F.1. Retirada de recursos a largo plazo

Prima abandono definitivo de superficies plantadas de vid (vinos y alcoholes)

Abandono definitivo de la producción lechera (leche y productos lácteos)

Selvicultura y repoblación forestal (desarrollo rural)

F.2. Producción de productos no comerciables

Programa LEADER (FEOGA-Orientación)

F.3. Otros criterios diferentes del producto

G. Pagos diversos

Otras medidas a productores (frutas y hortalizas)

Otros gastos (lino textil y cáñamo)

Medidas especiales de apoyo a productores (vacuno)

Ayudas a productores (POSEICAN)

ESTIMACIÓN DEL APOYO A LOS SERVICIOS GENERALES (GSSE)

H. Investigación y desarrollo agrario

I. Centros de capacitación agraria

J. Servicios de inspección

K. Infraestructura

Fomento de la adaptación y desarrollo de las zonas rurales (desarrollo rural)

Programas Operativos Región Objetivo nº 1 (FEOGA-Orientación)

Jubilación anticipada (desarrollo rural)

L. Promoción y comercialización

Programas de actividades de Organismos Operadores (aceite de oliva)

Mejora de la calidad del aceite (aceite de oliva)

Ayuda a la comercialización del plátano (frutas y hortalizas)

Fondos Operativos de las Organizaciones de Productores (frutas y hortalizas)

Ayuda plan de mejora frutos de cáscara (frutas y hortalizas)

Medidas de promoción (frutas y hortalizas)

Ayuda a la utilización de mostos (vinos y alcoholes)

Medidas de promoción (vinos y alcoholes)

Medidas de apoyo al mercado: EEB (vacuno)

Medidas de promoción (vacuno)

Medidas de apoyo al mercado: peste porcina (porcino)

Medidas escepcionales de apoyo al mercado (huevos y aves)

Mejora de la transformación y comercialización de productos agrícolas (desarrollo rural)

M. Almacenamiento público

Gastos de almacenamiento público (cereales)

Gastos de almacenamiento público (arroz)

Gastos de almacenamiento público (aceite de oliva)

Gastos de almacenamiento público (azúcar e isoglucosa)

Gastos de almacenamiento público (vinos y alcoholes)

Gastos de almacenamiento público (leche y productos lácteos)

Gastos de almacenamiento público (vacuno)
N. Otros servicios
Otros programas (desarrollo rural)
ESTIMACIÓN DEL APOYO AL CONSUMO (CSE)
O. Transferencias a los consumidores de los contribuyentes
Ayuda a la producción de fécula de patata (almidón y fécula de patata)
Prima a la fécula de patata (almidón y fécula de patata)
Ayuda a la utilización de almidón y fécula de patata (almidón y fécula de patata)
Ayuda a la producción de aceituna de mesa (aceite de oliva)
Ayuda a la utilización de aceite de oliva en conservas (aceite de oliva)
Ayuda al consumo (aceite de oliva)
Compensación por retirada de cítricos (frutas y hortalizas)
Compensación por retirada de frutas (frutas y hortalizas)
Compensación por retirada de hortalizas (frutas y hortalizas)
Ayuda a la transformación de cítricos (frutas y hortalizas)
Ayuda a la transformación de tomates (frutas y hortalizas)
Ayuda a la transformación de frutas (frutas y hortalizas)
Distribución gratuita (frutas y hortalizas)
Otras medidas a industrias (frutas y hortalizas)
Ayuda a la utilización en industria química (azúcar e isoglucosa)
Ayuda a la transformación de fibras de lino y cáñamo (lino textil y cáñamo)
Ayuda a la producción (algodón)
Ayuda a la producción de forrajes desecados (leguminosas de grano y forrajes)
Destilación obligatoria de subproductos de la vinificación (vinos y alcoholes)
Destilación de vinos de mesa (vinos y alcoholes)
Destilación de vinos no de mesa (vinos y alcoholes)
Ayuda a la mantequilla con destino repostería y otros (leche y productos lácteos)
Ayuda leche desnatada en polvo para alimentación animal (leche y productos lácteos)
Leche para escolares (leche y productos lácteos)
Ayudas a industrias y operadores (POSEICAN)

Fuente: elaboración propia.

*Nota: A (área), An (número de animales), R (ingresos), I (renta).

3. INDICADORES DE APOYO PÚBLICO A LA AGRICULTURA ESPAÑOLA: UNA COMPARACIÓN INTERNACIONAL

A continuación se muestran para España los principales resultados obtenidos por esta investigación. Después, para revisar el contexto internacional en que dichos indicadores se encuadran, hemos tomado los datos relativos a los años 2003 y 2006 que proporciona la propia OCDE[12] para construir gráficos comparativos que ilustren las diferencias entre países. Debemos incidir de nuevo en las limitaciones de los indicadores de la OCDE, señaladas en el apartado

[12] OCDE Database 1986-2007.

2.2, subrayando que su utilidad reside sobre todo en el seguimiento y la evaluación de las políticas, más que en una medida exacta del bienestar de los productores.

3.1. Estimación del Apoyo al Productor (PSE)

El Cuadro 2 muestra los resultados que hemos obtenido para el indicador PSE a lo largo del periodo considerado, su distribución regional promedio y su evolución entre principio y fin del periodo. El valor promedio del PSE en España ascendió a unos 11.350 millones de euros en el periodo 2002-2006. En 2003 experimentó un ligero incremento en términos nominales, para después reducirse paulatinamente hasta el 2006. Así, ha pasado de 11.800 millones en 2002 a menos de 10.500 en 2006, lo que implica una reducción aproximada del 12 por ciento. En cuanto a la distribución espacial del indicador se observa que más de la mitad del apoyo al productor se concentra en sólo tres CCAA, Andalucía, Castilla y León y Cataluña, que aglutinan el 22, 18 y 11 por ciento del mismo respectivamente.

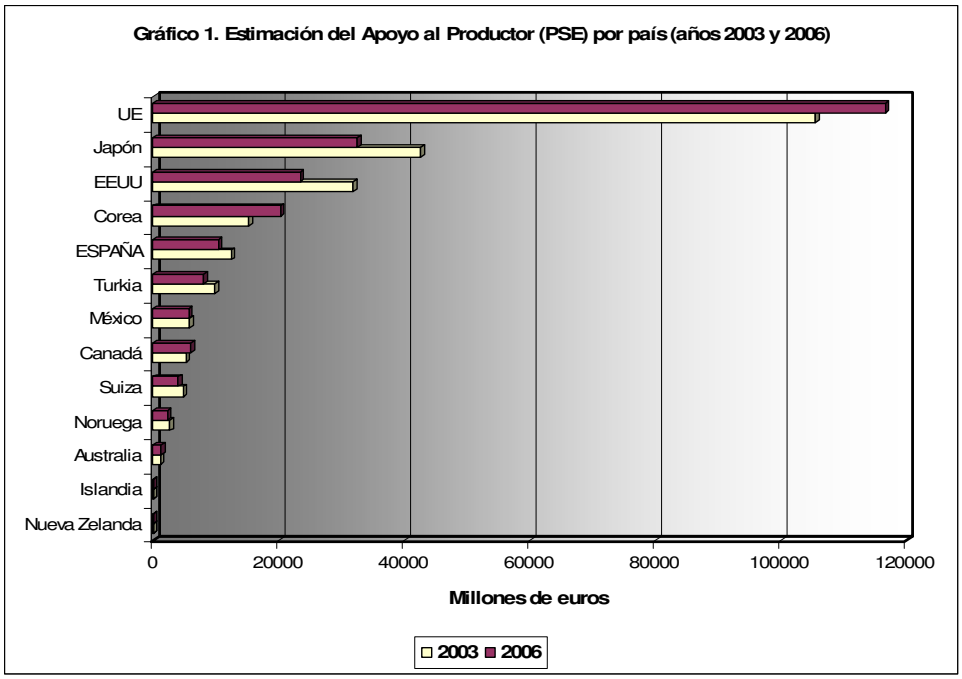
Cuadro 2. Estimación del Apoyo al Productor (PSE) por CCAA (millones de euros corrientes)

	PSE SEGÚN METODOLOGÍA OCDE					MEDIA	%	2006/2002
	2002	2003	2004	2005	2006	2002-2006	2002-2006	2002=100
GALICIA	652,8	688,5	648,7	548,9	582,9	624,4	5,5	89,3
ASTURIAS	245,3	260,0	234,7	197,6	192,3	226,0	2,0	78,4
CANTABRIA	148,3	150,4	140,6	114,4	101,3	131,0	1,2	68,3
PAÍS VASCO	199,8	208,5	174,3	152,1	133,2	173,6	1,5	66,7
NAVARRA	237,3	274,3	240,8	222,2	226,5	240,2	2,1	95,4
LA RIOJA	92,3	87,7	79,2	64,8	76,6	80,1	0,7	83,0
ARAGÓN	894,8	1.040,0	937,8	882,4	929,9	937,0	8,2	103,9
CATALUÑA	1.224,7	1.440,8	1.364,5	1.158,9	1.032,2	1.244,2	10,9	84,3
BALEARES	88,3	87,7	89,6	84,2	86,5	87,3	0,8	98,0
CASTILLA LEÓN	2.110,4	2.234,1	2.067,2	1.935,6	2.012,9	2.072,0	18,2	95,4
MADRID	230,2	230,4	179,2	156,0	163,4	191,8	1,7	71,0
CASTILLA LA MANCHA	1.080,1	1.075,7	1.101,4	959,4	975,3	1.038,4	9,1	90,3
COMUNIDAD VALENCIANA	614,7	632,6	680,6	594,7	538,7	612,3	5,4	87,6
MURCIA	324,0	419,9	378,6	334,1	320,0	355,3	3,1	98,8
EXTREMADURA	633,6	765,5	714,0	705,2	720,8	707,8	6,2	113,8
ANDALUCIA	2.941,1	2.697,6	2.454,6	2.314,3	2.282,6	2.538,0	22,3	77,6
CANARIAS	110,8	139,8	89,0	93,4	90,2	104,6	0,9	81,4
TOTAL ESPAÑA	11.828,5	12.433,5	11.574,8	10.518,2	10.465,3	11.364,1	100,0	88,5

Fuente: elaboración propia a partir de OCDE, MAPA y FEGA.

Si atendemos a la evolución por regiones se puede constatar la tendencia a la reducción en la práctica totalidad de las CCAA. En Navarra, Aragón, Baleares, Castilla y León y Murcia, el PSE no muestra variación apreciable en los últimos cinco años. Con reducción en torno a la media española destacan Galicia, Comunidad Valenciana y Castilla-La Mancha. En el resto de comunidades, la disminución es más acusada. Así, disminuye entre 15 y 20 por ciento en La Rioja, Cataluña y Canarias, entre 20 y 30 en Asturias, Madrid y Andalucía, y entre 30 y 40 en Cantabria y País Vasco. Como caso especial queda Extremadura, que aparenta un aumento de casi un 14 por ciento; sin embargo su MPS presenta una cifra anormalmente baja en el 2002 que estaría provocando el mismo efecto en su PSE, de tal manera que, si esa cifra no se considera Extremadura también habría iniciado su reducción del PSE, si bien es cierto que dicha reducción entre 2004 y 2006 no sería muy notable.

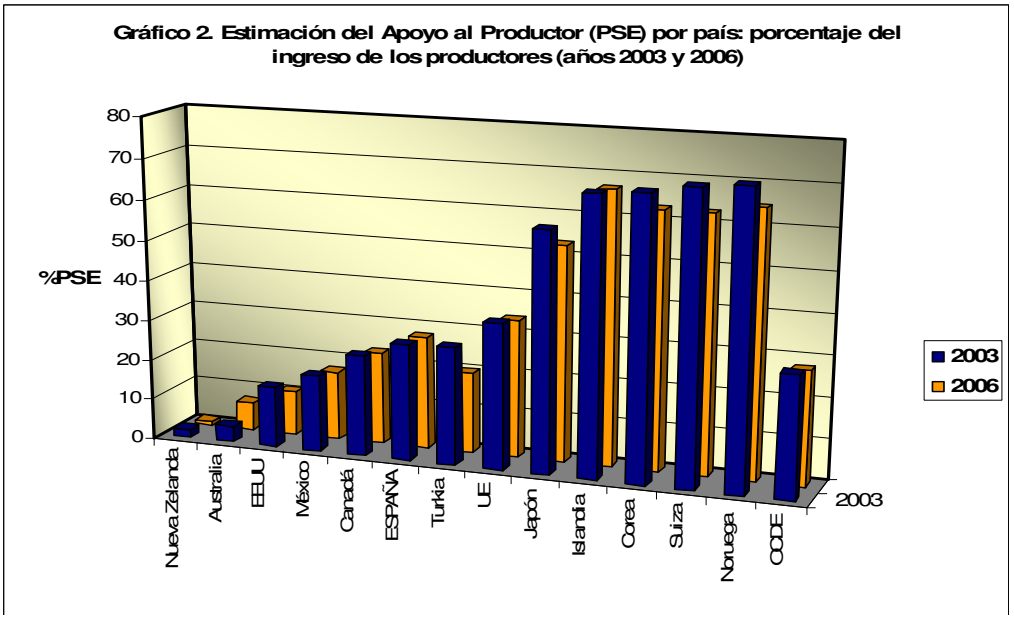
Como acabamos de ver, el valor total de la Estimación del Apoyo al Productor (PSE) originado en la PAC en España ascendió en 2006 a 10.500 millones de euros. Esta cuantía representa el 9 por ciento del PSE total estimado para la UE-25 según informa la OCDE, participación que viene reduciéndose a lo largo de todo el periodo considerado en este trabajo a medida que aumenta el número de socios comunitarios. A parte de otros socios comunitarios, como se ve en el Gráfico 1, sólo Corea, EEUU y Japón aglutinan un PSE superior al español. Destacan los altos valores obtenidos por los dos países asiáticos, 20.300 millones de euros para Corea y 32.400 para Japón, el doble y el triple, respectivamente, que en el caso español. El país más cercano a España respecto al PSE es Turquía, con 8.000 millones de euros. Ya por debajo del nivel español aparecen Canadá y México, con 6.000 millones, Suiza con 4.000, Noruega con 2.500 y Australia con 1.500. Testimonial es el caso de Islandia y Nueva Zelanda, con 180 y 70 millones de euros en concepto de PSE. El conjunto de la OCDE sumó un total de 220.500 millones en ese año.



Fuente: elaboración propia a partir de OCDE Database 1986-2007.

El PSE porcentual para España es de un 28 por ciento. Aquí es interesante observar las diferencias regionales: la parte de los ingresos que procede de transferencias de la política agraria es mucho más elevada para los productores del norte y oeste que para los del sur y este de España, siendo la especialización productiva de cada región la responsable de dichas diferencias.

En el 2006, el PSE porcentual es idéntico al promedio de la OCDE y significativamente inferior al promedio de la UE-25, que fue del 34 por ciento, como se desprende del Gráfico 2. Sobre España y la UE destacan los porcentajes de los países altamente proteccionistas como Japón, Suiza, Corea, Noruega e Islandia, donde más del 50 por ciento de los ingresos de los productores proceden de la política agraria. En México, Turquía y Canadá las ayudas públicas generan una quinta parte del ingreso total. En EEUU el PSE porcentual es tan sólo del 11 por ciento; en Nueva Zelanda y Australia es testimonial, con un 1 y un 7 por ciento respectivamente.



Fuente: elaboración propia a partir de OCDE Database 1986-2007.

3.2. Estimación del Apoyo a los Servicios Generales a la Agricultura (GSSE)

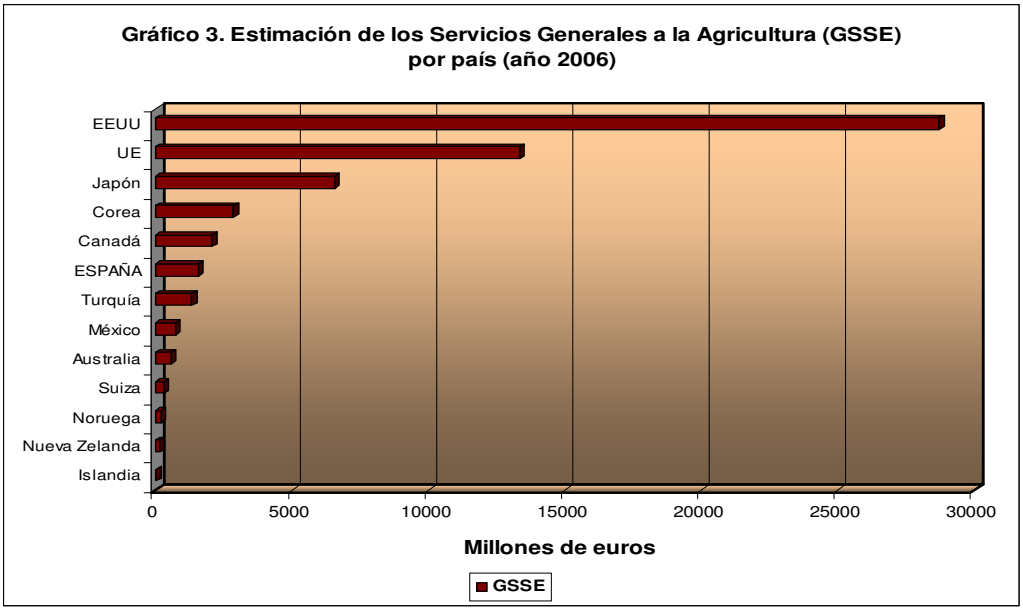
El valor del Apoyo a los Servicios Generales a la Agricultura relacionado con las transferencias de la PAC (GSSE), que se muestra en el Cuadro 3, asciende en España a casi 1.600 millones de euros, como promedio del periodo estudiado. Es una cuantía que no tiende a aumentar (más bien hay una reducción promedio nominal del 1 por ciento), a pesar de que incluye medidas que no distorsionan el comercio, y por tanto, no están cuestionadas en la Agenda Doha para el Desarrollo. Su distribución por regiones es más homogénea que el apoyo directo al productor, apareciendo la mitad del valor total concentrada en cuatro CCAA: Andalucía, Castilla y León, Castilla-La Mancha y Galicia, que acumulan porcentajes superiores al 10 por ciento del total. Las comunidades de la cornisa cantábrica, Baleares y Madrid reciben transferencias vía GSSE que no alcanzan el 2 por ciento del total.

En 2006, las transferencias en materia de Servicios Generales a la Agricultura (GSSE) ascendieron en España a 1.500 millones de euros, lo que representa un 11,6 por ciento de los 13.300 millones que se dedican a este fin en la UE-25. Como vemos en el Gráfico 3, por encima de España destacan en este sentido Canadá con 2.000 millones, Corea con 2.800, Japón con 2.500 y sobretodo EEUU con 29.000 millones, más del doble que en toda la UE, denotando el peso que tiene el gasto federal en políticas de promoción del consumo, investigación y transferencia de tecnología. El país más próximo a España resulta ser de nuevo Turquía, que dedica a servicios a la agricultura 1.300 millones de euros. México, Australia y Suiza invierten 700, 550 y 300 millones respectivamente, Nueva Zelanda y Noruega 140 y 200, y testimonial es el valor en Islandia con sólo 16 millones de euros. El conjunto de la OCDE sumó por este concepto un total de 56.800 millones.

Cuadro 3. Estimación de los Servicios Generales a la Agricultura (GSSE) por CCAA (millones de euros corrientes)

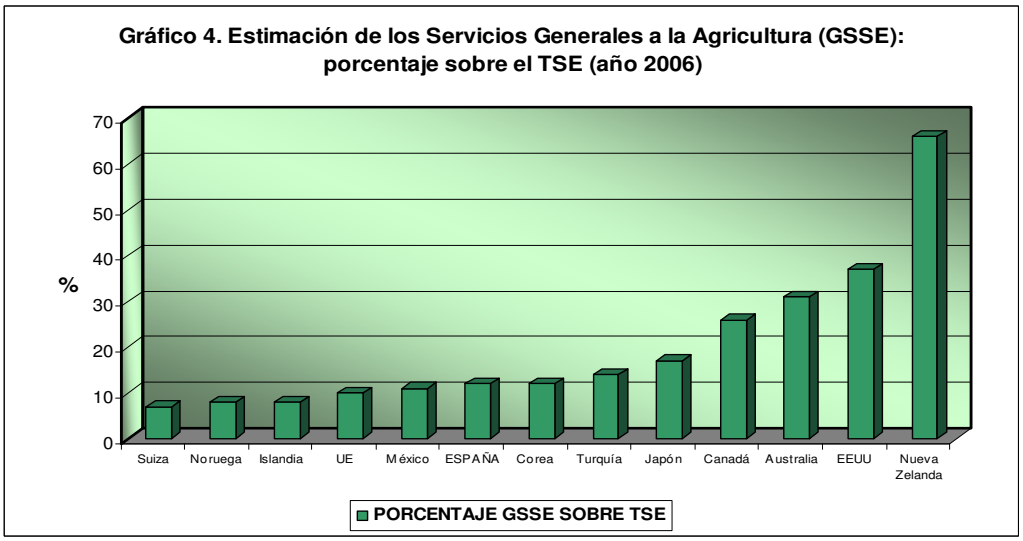
	GSSE SEGÚN METODOLOGÍA OCDE					MEDIA	%
	2002	2003	2004	2005	2006	2002-2006	2002-2006
GALICIA	160,3	157,6	155,3	160,1	154,9	157,6	10,0
ASTURIAS	45,5	44,0	42,6	41,3	44,9	43,7	2,8
CANTABRIA	19,8	19,8	19,5	20,1	19,2	19,7	1,2
PAÍS VASCO	21,5	19,9	21,4	22,4	22,2	21,5	1,4
NAVARRA	29,5	30,2	30,9	30,4	31,5	30,5	1,9
LA RIOJA	16,6	17,8	17,6	16,4	16,3	16,9	1,1
ARAGÓN	111,3	112,2	105,4	104,6	108,4	108,4	6,8
CATALUÑA	88,0	94,0	86,6	95,0	96,7	92,1	5,8
BALEARES	13,2	12,9	12,3	12,8	12,1	12,7	0,8
CASTILLA LEÓN	238,8	241,8	236,6	230,7	227,9	235,2	14,9
MADRID	23,7	25,0	24,0	24,5	24,9	24,4	1,5
CASTILLA LA MANCHA	173,0	169,0	163,3	170,2	162,1	167,5	10,6
COMUNIDAD VALENCIANA	94,5	96,8	96,2	99,2	108,5	99,0	6,3
MURCIA	65,8	66,4	66,7	80,7	90,2	74,0	4,7
EXTREMADURA	108,1	103,3	103,9	107,9	96,7	104,0	6,6
ANDALUCIA	257,1	240,0	257,9	263,0	254,8	254,6	16,1
CANARIAS	146,3	149,8	148,5	84,9	79,6	121,8	7,7
TOTAL ESPAÑA	1.613,0	1.600,5	1.588,7	1.564,2	1.550,9	1.583,5	100,0

Fuente: elaboración propia a partir de OCDE, MAPA y FEGA.



Fuente: elaboración propia a partir de OCDE Database 1986-2007.

Muy interesante es comentar el porcentaje que el GSSE representa sobre el apoyo total (TSE), que aparece en el Gráfico 4. En España es de un 12 por ciento, algo más que en la UE-25, donde sólo alcanza el 10 por ciento. Ambas cifras son inferiores al promedio del área OCDE, que ya alcanza el 19 por ciento. Lo habitual es que los servicios a la agricultura no lleguen a una quinta parte del apoyo total, situándose entre el 7 por ciento de Suiza y el 17 por ciento de Japón. En Canadá ocupan una cuarta parte del apoyo total y en Australia y EEUU algo más de un tercio. Como caso especial hay que citar a Nueva Zelanda, donde sus 140 millones de euros en materia de GSSE son el 66 por ciento del apoyo total.



Fuente: elaboración propia a partir de OCDE Database 1986-2007.

3.3. Estimación del Apoyo Total (TSE)

El TSE, elaborado en el Cuadro 4, arroja para España un valor promedio anual de 13.700 millones de euros entre 2002 y 2006. La mitad se concentra en Andalucía, Castilla y León y Cataluña, que reciben el 22, 17 y 10 por ciento del mismo. La tendencia que experimenta el indicador es a la baja, perdiendo a nivel nacional casi 12 puntos porcentuales, tal y como sucedía con el PSE.

Cuadro 4. Estimación del Apoyo Total (TSE) por CCAA (millones de euros corrientes)

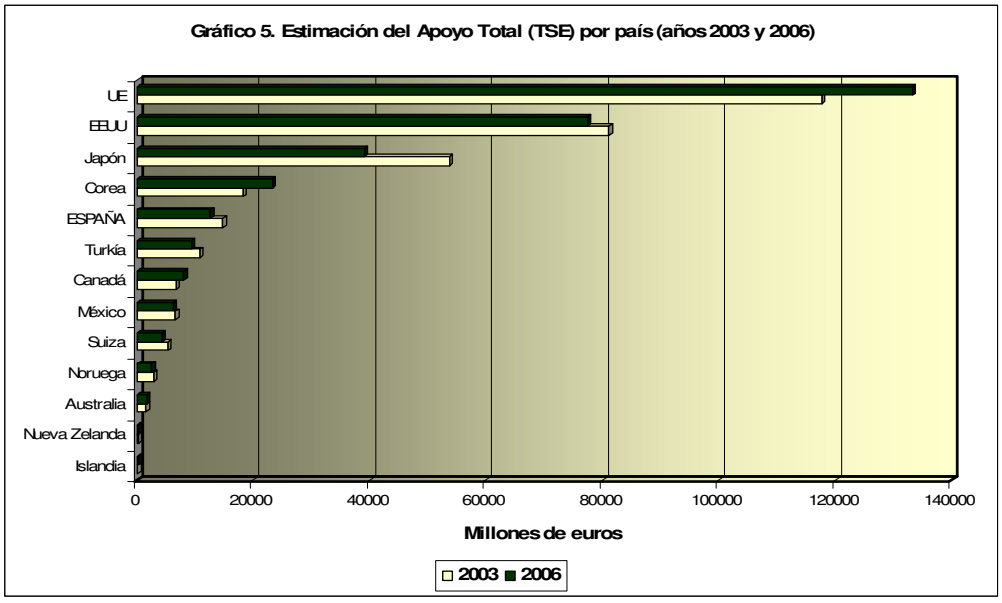
	TSE SEGÚN METODOLOGÍA OCDE					MEDIA	%	2006/2002
	2002	2003	2004	2005	2006	2002-2006	2002-2006	2002=100
GALICIA	824,9	858,7	817,0	719,0	744,1	792,7	5,8	90,2
ASTURIAS	293,2	306,2	280,7	242,0	239,7	272,4	2,0	81,8
CANTABRIA	169,0	171,2	160,9	135,6	121,1	151,6	1,1	71,7
PAÍS VASCO	223,0	229,8	197,1	176,1	156,7	196,5	1,4	70,3
NAVARRA	280,4	316,5	284,8	263,1	277,2	284,4	2,1	98,9
LA RIOJA	111,5	108,2	99,9	84,8	95,8	100,0	0,7	85,9
ARAGÓN	1.086,7	1.235,3	1.131,4	1.028,1	1.108,2	1.117,9	8,2	102,0
CATALUÑA	1.392,6	1.607,3	1.525,6	1.303,3	1.185,9	1.402,9	10,2	85,2
BALEARES	101,9	101,0	102,2	97,2	98,9	100,2	0,7	97,1
CASTILLA LEÓN	2.357,8	2.486,4	2.322,7	2.183,0	2.254,6	2.320,9	16,9	95,6
MADRID	256,2	258,1	206,0	183,0	190,6	218,8	1,6	74,4
CASTILLA LA MANCHA	1.388,3	1.395,8	1.423,6	1.310,6	1.307,0	1.365,1	10,0	94,1
COMUNIDAD VALENCIANA	770,6	793,0	830,4	754,5	700,3	769,8	5,6	90,9
MURCIA	416,5	511,9	465,5	437,1	434,1	453,0	3,3	104,2
EXTREMADURA	806,3	944,5	894,4	900,9	871,3	883,5	6,4	108,1
ANDALUCIA	3.465,9	3.171,7	3.033,9	2.862,1	2.609,4	3.028,6	22,1	75,3
CANARIAS	257,4	289,8	261,8	206,3	199,9	243,0	1,8	77,7
TOTAL ESPAÑA	14.202,2	14.785,4	14.037,9	12.886,7	12.594,8	13.701,4	100,0	88,7

Fuente: elaboración propia a partir de OCDE, MAPA y FEGA.

La estimación del Apoyo Total (TSE) derivado de la PAC, que totalizó para España en el año 2006 los 12.600 millones de euros, representa el 9,4 por ciento del total estimado por la OCDE para la UE-25. Al igual que sucedía con el PSE, si exceptuamos a otros socios europeos, los países que superan el TSE español son Corea y Japón, que doblan y triplican el valor que hemos comentado, y EEUU que alcanza más de 77.000 millones de euros. En torno a España se encuentran Canadá y Turquía, con 8.000 y 9.500 millones. Entre 1.500 y 6.500 millones es lo que dedican a financiar la política agraria en Australia, Noruega, Suiza y México, y unos 200 millones se gastan en Islandia y Nueva Zelanda. El conjunto de la OCDE sumó un apoyo total que supera los 300.000 millones. Esta información se recoge en el Gráfico 5.

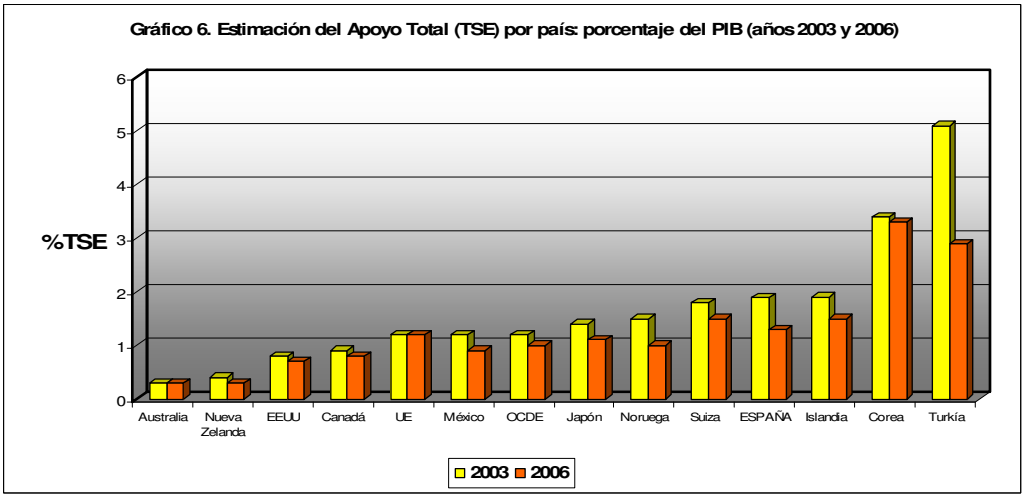
El crecimiento del PIB en España en el periodo de estudio, unido a la reducción del apoyo total en términos nominales, ha conllevado una caída significativa del TSE, en porcentaje del PIB, del 1,9 por ciento al 1,3 por ciento, situándose prácticamente en el nivel promedio de la UE (1,2 por ciento) y sólo un poco por encima de la media de la OCDE (1 por ciento). A nivel regional destacan los altos valores que caracterizan a Extremadura, las dos Castillas, Aragón y Andalucía, mientras que no llega al 1 por ciento en País Vasco, Cataluña, Baleares, Madrid, Comunidad Valenciana y Canarias. Así, las regiones mediterráneas y de la cornisa cantábrica presentan unos valores más bajos que el sur y el centro de la península.

El TSE en porcentaje del PIB, para los países miembros de la OCDE y en el año 2006, se muestra en el Gráfico 6. La mayor parte de países se mueven en el intervalo comprendido entre el 0,3 por ciento que el TSE representa para Australia y Nueva Zelanda y el 1,5 por ciento de Suiza e Islandia. En España es del 1,3 por ciento, un poco por encima del promedio de la Unión que está situado en el 1,2 y del promedio de la OCDE que está en el 1 por ciento. En este sentido, destaca Turquía con el 2,9 y Corea con el 3,3.



Fuente: elaboración propia a partir de OCDE Database 1986-2007.

Del análisis y comparación internacional de los indicadores de la OCDE se deduce la existencia de grupos de países con distintas características en relación a los niveles de protección que conceden a su agricultura. Así, destacan los países altamente proteccionistas encabezados por Japón y Suiza. En el otro extremo encontramos países muy liberales, como Australia y Nueva Zelanda, que basan buena parte de su apoyo al sector en ayuda estructural. Finalmente, entre ambos grupos quedan enmarcados los EEUU y la UE, todavía muy proteccionistas, aunque se encuentran en los últimos años comprometidos con una verdadera reforma de su política agraria. Si bien se observa, en general, una tendencia hacia una menor protección a la agricultura, lo cierto es que en el periodo estudiado los cambios no han sido muy radicales: sigue existiendo un apoyo importante a la agricultura. Pero esta situación podría alterarse sustancialmente a partir del año 2007 como efecto del aumento de los precios mundiales de los principales alimentos.



Fuente: elaboración propia a partir de OCDE Database 1986-2007.

4. COMPOSICIÓN DEL APOYO TOTAL

Una vez concluidos los cálculos del TSE resulta de mucho interés analizar, para España y las CCAA, la composición del apoyo al productor por tipos de medidas y su financiación por

consumidores o contribuyentes, especialmente teniendo en cuenta que los cálculos realizados ya recogen algunos efectos de la "Revisión Intermedia" de la PAC que se empieza a aplicar en España en 2006. Es justo reconocer a la vista de los datos que hemos revisado que la intervención pública en la agricultura sigue siendo apreciable en el conjunto de España. Además, todavía se expresa en gran parte a través de políticas que financian los consumidores y de instrumentos que generan precios internos superiores a los precios internacionales. Una actualización de las estimaciones para 2007 posiblemente recoja los efectos de reformas recientes (azúcar, frutas y hortalizas) y posiblemente también los cambios de tendencia en los mercados mundiales de productos básicos, que apuntan a una elevación de los precios internacionales, y por tanto, a una más que probable reducción del MPS. Además, la liberalización de las intervenciones de precios y la apertura gradual de los mercados exteriores han corrido paralelas a la creciente utilización del presupuesto público como mecanismo de apoyo. Así, los cambios introducidos tras la reforma MacSharry, en 1992, implicaron la utilización de pagos directos a los productores. Tras la Agenda 2000, la proporción del apoyo total a la agricultura en la UE financiada por los contribuyentes ya representaba más del 50 por ciento en 2003 y esa tendencia se ha acentuado tras la aplicación de la Revisión Intermedia, alcanzando el 63 por ciento que vemos en el Cuadro 5.

El cambio en las formas de apoyo también ha sido evidente en España, con apreciables diferencias entre CCAA. Como refleja el Cuadro 5, en España la proporción del apoyo financiada por los contribuyentes representa un 59 por ciento del apoyo total del año 2006, es decir, menor a su participación calculada para la UE en conjunto. Como vemos en el citado cuadro, la financiación del TSE en España se encuentra en la media de la OCDE.

Si atendemos a la situación regional, el predominio de las transferencias de los contribuyentes es patente en la España interior (Navarra, La Rioja, Aragón, Castilla-La Mancha, Andalucía y Extremadura) y Canarias, mientras que las de los consumidores están en torno al 60 por ciento o lo superan en el País Vasco, Madrid y en el litoral mediterráneo (Murcia, Comunidad Valenciana y Cataluña). A pesar de todo, la tendencia a una política agraria cada vez menos financiada por el consumidor persiste y se acentúa en todas nuestras regiones.

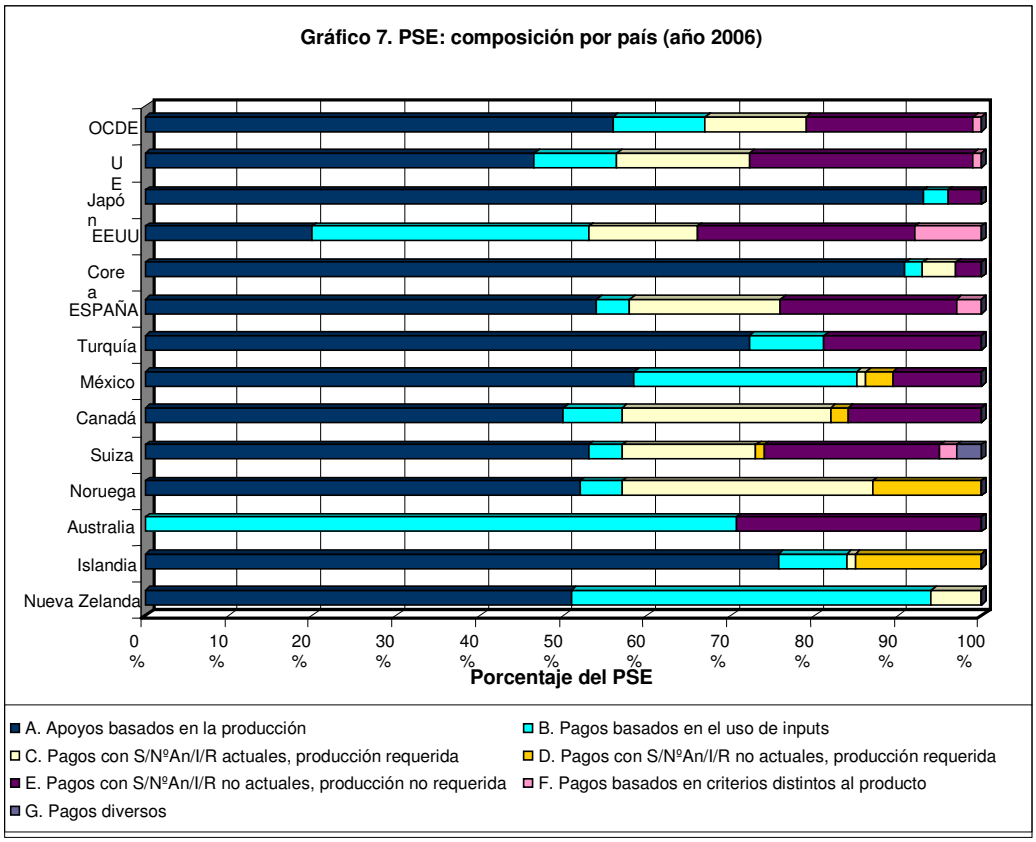
Por otra parte, la clasificación de las categorías de ayudas recientemente adoptada por la OCDE permite ahondar más en la composición del apoyo total por tipo de medidas y reflejar mejor la evolución de las mismas. Este análisis se vuelve especialmente relevante en un contexto de reforma de la PAC, al proporcionar una evaluación de los primeros resultados de la última gran reforma.

Así, dentro del PSE, aún después de las reformas recientes de la PAC, siguen teniendo importancia las políticas de ayuda basadas en la producción, es decir, las medidas de apoyo a los precios más los pagos vinculados a la producción. Como se ve en el Gráfico 7, los apoyos basados en la producción originados en la PAC todavía representan en España el 54 por ciento del PSE, mientras que este porcentaje es del 47 por ciento en la UE, y del 56 por ciento en promedio de la OCDE. Hay que tener en cuenta, por un lado, que el sistema de pago único por explotación está todavía en fase de implantación en nuestro país, y por el otro, el grueso del apoyo de algunos productos importantes (por ejemplo, frutas y hortalizas, lácteos) todavía se basa en la protección en frontera, es decir, en un tipo de apoyo vinculado a la producción.

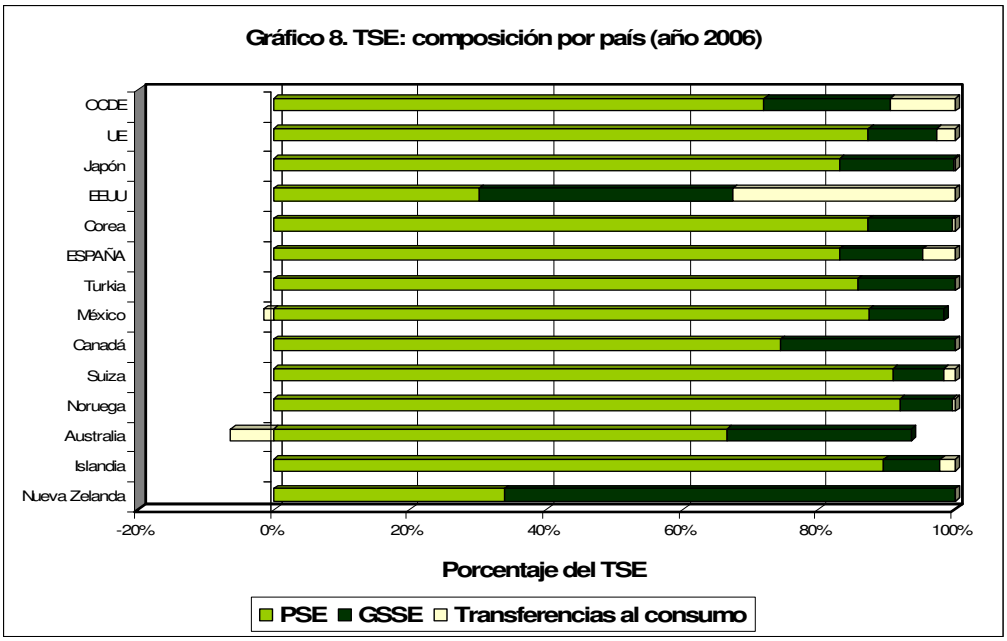
Cuadro 5. La Estimación del Apoyo Total (TSE) y sus componentes en el año 2006 (millones de euros corrientes)

	Estimación del Apoyo Total			
	De los consumidores	% del total	De los contribuyentes	% del total
Galicia	344,4	46,3	399,7	53,7
Asturias	102,7	42,8	137,0	57,2
Cantabria	63,4	52,4	57,7	47,6
País Vasco	93,8	59,8	62,9	40,2
Navarra	109,3	39,5	167,8	60,5
La Rioja	34,7	36,2	61,1	63,8
Aragón	445,2	40,2	662,9	59,8
Cataluña	830,0	70,0	355,9	30,0
Baleares	49,7	50,2	49,3	49,8
Castilla-León	955,0	42,4	1299,6	57,6
Madrid	119,8	62,9	70,8	37,1
Castilla-La Mancha	256,3	19,6	1050,7	80,4
Comunidad Valenciana	446,5	63,8	253,8	36,2
Murcia	259,1	59,7	175,0	40,3
Extremadura	148,4	17,0	722,9	83,0
Andalucía	846,7	32,4	1762,8	67,6
Canarias	51,4	25,7	148,5	74,3
Total España	**5156,3**	**40,9**	**7438,5**	**59,1**
Total UE	**46212,0**	**36,7**	**79724,0**	**63,3**
Total OCDE	**159673,0**	**40,6**	**234055,0**	**59,4**

Fuente: elaboración propia a partir de OCDE, MAPA y FEGA.



Fuente: elaboración propia a partir de OCDE Database 1986-2007.



Fuente: elaboración propia a partir de OCDE Database 1986-2007.

Finalmente, vamos a analizar la composición del apoyo total que recibe el productor (TSE) incluyendo no sólo las transferencias a los agricultores (PSE) sino también la Estimación de Servicios Generales a la Agricultura (GSSE) y las ayudas al consumo. El Gráfico 8 demuestra que el PSE domina con creces por encima de otras ayudas en el TSE. Así, el apoyo al productor originado en la PAC representa en España el 83 por ciento del TSE, mientras que este porcentaje es todavía mayor para el conjunto de la UE, un 87 por ciento, y del 72 por ciento en promedio de la OCDE. Destacan en la OCDE los casos de EEUU y Nueva Zelanda, donde el PSE no supera el 35 por ciento del TSE.

5. CONCLUSIONES

La política agraria en los países de la OCDE ha estado sujeta en los últimos años a cambios sustanciales, que han implicado profundas modificaciones en el nivel y los mecanismos de apoyo al sector, con aplicación diferenciada en las distintas regiones. En este contexto, el presente trabajo ha propuesto la estimación sistemática de indicadores del apoyo comunitario a la agricultura en España y sus CCAA, a través de una adaptación de la metodología OCDE al análisis regional[13]. El periodo considerado ha abarcado los años 2002-2006, pero su actualización es sencilla recurriendo a las mismas fuentes de información. El sistema incluye indicadores como la *Estimación del Apoyo al Productor* (PSE), que agrupa el apoyo al precio de mercado y las transferencias presupuestarias directas a los productores, la *Estimación de los Servicios Generales a la Agricultura* (GSSE), que recoge las medidas de desarrollo rural (que no comportan transferencias a los productores), y la *Estimación del Apoyo Total* (TSE), que suma los dos indicadores anteriores más las subvenciones al consumo.

La PAC supone un flujo de transferencias importante para España y sus CCAA. En concreto, el PSE español ha ascendido en promedio a 11.350 millones de euros anuales, cifra que incluye tanto las transferencias tangibles del FEOGA, como las intangibles asociadas a la protección vía precios. Un 60 por ciento de este apoyo se concentra en cuatro CCAA, relevantes por la dimensión de su sector agrario: Andalucía, Castilla y León, Cataluña y Castilla-La Mancha. Como en otros países de la OCDE y en la propia UE, se puede constatar la tendencia a la reducción del apoyo, situándose el valor español en 2006 en un 28 por ciento de los ingresos

[13] Una información más completa sobre la utilidad de los indicadores planteados, su estimación y su interpretación puede encontrarse en un informe detallado elaborado para el Ministerio de Agricultura, Pesca y Alimentación (MAPA, 2007).

de los productores, cifra idéntica al nivel promedio de la OCDE. Este porcentaje es superior para los productores del norte y oeste que para los del sur y este de España (la especialización productiva importa), y varía, con una tendencia a la reducción, desde el 13 por ciento de La Rioja o Canarias al 50 por ciento que supone en Asturias y Madrid. Las transferencias en materia de Servicios Generales a la Agricultura (GSSE) representan en España un 12 por ciento del apoyo total a la agricultura, suponiendo una transferencia anual de 1.550 millones de euros. Este desequilibrio con respecto a las ayudas a los precios y las rentas es tradicional en la historia de la PAC. Por último, el TSE arroja para España un valor anual promedio de 13.700 millones de euros, lo que equivale a un 1,7 por ciento del PIB. La tendencia del TSE y el TSE porcentual es claramente a la reducción prácticamente en todas las regiones. Puede observarse que el crecimiento del PIB en España entre ambos años, unido a la reducción del apoyo total en términos nominales, ha conllevado una caída significativa del TSE en España, en porcentaje del PIB, del 1,9 por ciento al 1,3 por ciento, situándose prácticamente en el nivel promedio de la UE (1,2 por ciento) y sólo un poco por encima de la media de la OCDE (1 por ciento). En comparación con la UE, España se aproxima mucho a la media europea. En la Europa Ampliada, España representa en torno al 9 por ciento para el PSE y el TSE y el 11,6 por ciento para el GSSE.

Estos resultados sugieren que los efectos de las intervenciones en la agricultura siguen siendo palpables en España, y todavía se expresa en gran parte a través de políticas que financian los consumidores y de instrumentos que generan precios internos superiores a los internacionales. Esto está ocurriendo incluso tras la reforma de la PAC denominada "Revisión Intermedia", aprobada en 2003 y aplicada a partir de 2006 en diez de las CCAA. La proporción del apoyo financiada por los contribuyentes representa ya un 59 por ciento del total en 2006, porcentaje inferior al observado en la UE que alcanza el 63 por ciento y muy próximo al promedio de la OCDE. Por regiones observamos un predominio patente de las transferencias de los contribuyentes en la España interior, con presencia de cultivos herbáceos (Navarra, La Rioja, Aragón, Castilla-La Mancha, Andalucía y Extremadura) y en Canarias (plátano), mientras que las transferencias de los consumidores están en torno al 60 por ciento o lo superan en País Vasco, Madrid y litoral mediterráneo (Murcia, Comunidad Valenciana y Cataluña). En estas regiones, la especialización láctea o mediterránea de la producción conlleva al predominio de medidas de protección en frontera. A pesar de todo, se puede entrever que la tendencia a la pérdida de importancia de los consumidores como financiadores se acentúa a lo largo de los años estudiados en todas las regiones. El alza de los precios de las materias primas agrícolas probablemente intensifiquen este menor papel de la protección vía precios.

En consecuencia, este trabajo evidencia para España unas diferencias regionales en los niveles de apoyo que son apreciables, así como también en las medidas o en los instrumentos que privilegian. La situación es distinta de unas regiones a otras, por tanto los riesgos son diferentes. El conocimiento de una aplicación diferenciada de las políticas agrarias a nivel autonómico, con origen principalmente en la especialización productiva, es potencialmente útil en la racionalización de la toma de decisiones políticas. Por otra parte, la evolución que experimentan las CCAA (aunque con ritmos diferentes) es acorde a lo que acontece en el seno de la UE y responde al desarrollo de las negociaciones multilaterales en la Ronda Doha que actualmente se están llevando a cabo en el seno de la OMC.

En el año 2006, el valor total de las ayudas a los productores de la OCDE ascendió a 220.500 millones de euros. Tal y como se refleja en el porcentaje de subsidios a la agricultura, las ayudas representaron el 28 por ciento de los ingresos de los productores. Incluidas las ayudas a los servicios de la agricultura tales como la investigación, la infraestructura, la inspección, el marketing y la promoción, las ayudas totales al sector agrícola fueron iguales al 1 por ciento del PIB de la OCDE.

En el conjunto de la OCDE, en general ha habido pocos cambios en el nivel de apoyo a los productores en el periodo objeto de estudio. Dichas ayudas a los productores han disminuido cerca de un 5 por ciento entre los años 2003 y 2006. Estas fluctuaciones reflejan fundamentalmente la adopción de medidas políticas tendentes a limitar que los precios del comercio internacional pasen a ser los de los mercados internos. La reforma política se ha centrado en modificar el modo en el que se hace llegar dicha ayuda a los productores, con un cambio importante en las medidas vinculadas a la producción. Mientras que dicho cambio es

probable que continúe en los próximos años, las medidas relacionadas con la producción todavía dominan la ayuda a los productores en la mayoría de los países, fomentando la producción, distorsionando el comercio y contribuyendo a bajar los precios mundiales de los productos agrícolas. Además, algunos sectores productivos se han mantenido relativamente al margen de las reformas hasta la fecha y hace mucha falta subsanar esa deficiencia. A pesar de dicho desplazamiento de las ayudas vinculadas a la producción, el giro hacia las políticas que definen claramente los objetivos y los beneficiarios ha sido débil. Se necesitan aún más esfuerzos para asegurarse de que las políticas sean más transparentes a la hora de distribuirse y destinarse a soluciones específicas, y más flexibles para poder responder a las prioridades cambiantes. El proceso de reforma actual en la Unión denominado *"Health Check"* apunta a esa dirección, aunque todavía de una manera tímida.

Los indicadores de apoyo utilizados y su traducción en medidas de la ayuda a la agricultura española y de las Comunidades Autónomas presentan, como se ha señalado en este artículo, ventajas e inconvenientes. Mientras sería posible encontrar medidas que aproximasen mejor los cambios en los niveles de equilibrio de producción y de precios y en el excedente de los productores como consecuencia de las reformas de las políticas agrarias, resulta difícil su traslación a indicadores sencillos que permitan un seguimiento de dichas reformas y, sobre todo, la comparación internacional de los cambios en las políticas agrarias. Los indicadores presentados cumplen esta función y contribuyen a la evaluación de las políticas agrarias en nuestro país, en un contexto internacional.

REFERENCIAS BIBLIOGRÁFICAS

ANDERS, S.; HARCHE, J.; HERRMANN, R.; SALHOFER, K. (2004): Regional Income Effects of Producer Support under the CAP. *Cahiers d'economie et sociologie rurales*. Vol. 73, No. 4, págs. 103-122.

ATANCE, I.; GOMEZ-LIMÓN, J.A. (2004): Identificación de objetivos públicos para el apoyo al sector agrario, *Revista Española de Estudios Agrosociales y Pesqueros*, nº 203, págs. 49-84.

BERTHELOT, J. (2004): *Ending Food dumping: taking the US and EU through the WTO disputes procedure after the expiry of the "peace clause"*, http://solidarite.asso.fr) May 2004.

BLANDFORD, D. (2005): Imposing WTO disciplines on domestic support: an assessment of the Doha Round Approach, *Economía Agraria y Recursos Naturales*, Vol. 5, nº 9, págs. 3-26.

COMPÉS, R. Y GARCÍA ALVAREZ-COQUE, J.M. (2005): Las reformas de la política agrícola común en la Unión Europea ampliada: implicaciones económicas para España, *Papeles de Economía Española* nº 103, págs. 230-244.

CORDEN W.M. (1971). *The Theory of Protection*, Oxford.

EUROPEAN COMMISSION (2001). *Study on the impact of community agricultural policies on economic and social cohesion*, Directorate- General for Regional Policy, Brussels.

EUROPEAN SPATIAL PLANNING OBSERVATION NETWORK (2005). ESPON Project 2.1.3: *The Territorial Impact of CAP and Rural Development Policy.* Arkleton Institute for Rural Development Research, University of Aberdeen.

FEGA (varios años): Informe de Actividad del Fondo Español de Garantía Agraria. Ministerio de Agricultura, Pesca y Alimentación, www.mapya.es.

GARCIA ÁLVAREZ-COQUE, J.M. et al. (1994): *Gasto Público y Sistema Autonómico. La política agraria autonómica a examen.* Conselleria d'Agricultura, Pesca i Alimentació de la Generalitat Valenciana. Serie Estudios.

GARCÍA ÁLVAREZ-COQUE, J.M. y ANDRÉS GONZÁLEZ-MORALEJO, S. (2007): Estimación del apoyo comunitario a la agricultura de las Comunidades Autónomas, *Revista Española de Estudios Agrosociales y Pesqueros*, nº 213, págs. 179-206.

HANSEN, H. (2005): Effects of agricultural policy measures on gross transfers to farmers: Intertemporal and interregional evidence from the CAP *Paper prepared for presentation at the XIth Congress of the EAAE (European Association of Agricultural Economists) 'The Future of Rural Europe in the Global Agri-Food System' Copenhagen, Denmark August 24-27, 2005.*

JOSLING T. (1979). Agricultural protection and stabilization policies: analysis of current neomercantilist practices, *in : International Trade and Agriculture: Theory and Policy*, Hillman J.A., Schmitz A. (eds), (Westview Special Studies in International Economics and Business), Boulder, págs. 149-162.

MAPA (varios años): Anuario de Estadística Agroalimentaria. Ministerio de Agricultura, Pesca y Alimentación, www.mapya.es.

MAPA (varios años): Avances de Superficies y Producciones de Cultivos. Ministerio de Agricultura, Pesca y Alimentación, www.mapya.es.

MAPA (2004): Territorialización de los apoyos públicos, en Libro Blanco sobre la Agricultura y el Desarrollo Rural, Capítulo 3, Ministerio de Agricultura, Pesca y Alimentación, www.mapya.es.

MAPA (2006): Encuestas de Sacrificio de Ganado. Ministerio de Agricultura, Pesca y Alimentación, www.mapya.es.

MAPA (2007): Observatorio del Apoyo Público a la Agricultura. Estudio del apoyo público, presupuestos autonómicos y sistema de seguimiento, ejercicio 2007. Ministerio de Agricultura, Pesca y Alimentación. Págs. 35-85.

NUCIFORA, A. y SARRI, D. (1997): *Levels of protection for the fruit, vegetables, olive oil and wine sectors of the European Union.* Quaderno per discussione nº 19, febrero. Centro Interdepartamental de Política Agroalimentaria-ambiental, Universidad de Siena.

OCDE (varios años): *Agricultural Policies, Markets and Trade. Monitoring and Outlook*. París.

OECD Database (2007): Agricultural Policies in OECD countries: monitoring and evaluation 2007. http://www.oecd.org/document/59/0,3343,en_2825_494504_39551355_1_1_1_1,00.html

PHILIPPIDIS, G. (2005) Agricultural trade liberalisation in the Doha Round: impacts on Spain. *Economía Agraria y Recursos Naturales*, Vol 5. nº 10, págs. 41-68.

REIG, E. (1991): "Hacia una valoración cuantitativa de las políticas agrícolas: la experiencia de la OCDE". *Información Comercial Española*, Vol. 700. Págs. 61-70

SUMNER, D.A. y BRUNKE, H. (2003). "Commodity Policy and California Agriculture", en Jerome B. Siebert. Forthcoming in *California Agriculture: Issues and Challenges*. University of California, Giannini Foundation.

TANGERMANN, STEFAN (2005), "Is the Concept of the Producer Support Estimate in Need of Revision?", *OECD Food, Agriculture and Fisheries Working Papers,* No. 1, OECD Publishing. doi:10.1787/845314770374

TARDITI S., ZANIAS G. (2001). Common agricultural policy, *in : Competitiveness and Cohesion in EU Policies,* Hall R., Smith A.and Tsoukalis L. (eds.), Oxford University Press, págs. 179-216.

WALKENHORST P. (2003). Regional diversity, policy targeting and agricultural producer support in Switzerland, *Jahrbuch für Regionalwissenschaft*, 23 (2), págs. 141-153.

WISE, T. A. (2004), *The Paradox of Agricultural Subsidies: Measurement Issues, Agricultural Dumping, and Policy Reform*, Global Development and Environment Institute, Tufts University, Working paper N° 04-02, February 2004 (http://ase.tufts.edu/gdae).

ZANIAS G.P. (2002). The distribution of CAP benefits among member states and the impact of a partial re-nationalisation: a note, *Journal of Agricultural Economics*, 53 (1), págs. 108-112.

9. Economía Internacional y de la Integración

ECONOMIA BRASILEIRA E PORTUGUESA: ESPECIFICIDADES DE SUA DEMANDA NA CRISE MUNDIAL

TEREZINHA FILGUEIRAS DE PINHO
E-mail: terezinharr@yahoo.com.br; terezinha@ifrr.edu.br
Instituto Federal de Educação

Resumo

O trabalho objetiva analisar a situação econômica e financeira vivenciada no Brasil, a realidade nos mesmos aspectos presente em Portugal e a posição de ambos diante da atual crise mundial. Esta, tendo as raízes no sistema financeiro, originou-se nos países avançados e atingiu o mundo de forma sincronizada. A importância do tema decorre das vinculações entre essa crise e o processo de globalização, acarretando reflexos diferentes nas economias das nações e, especificamente, de Portugal e do Brasil. O ponto inicial da análise é a interligação entre a globalização e a economia e, para entender o contexto lusitano e brasileiro, aborda a situação nos Estados Unidos. Problemas estruturais, situação presente e perspectivas são consideradas nas reflexões sobre os países irmãos. Possibilidades para atuação conjunta de modo a atenuar as conseqüências internas de ambos os países são discutidas na última parte desta contribuição.

Palavras-chave: Crise, Globalização, Relações Bilaterais.
Área Temática: Economia Internacional e da Integração.

Abstract

The work aimed to analyze the economic and financial situation experienced in Brazil, the reality in the same aspects present in Portugal and the position of both in the face of the current world crisis. This, having roots in the financial system, caused-in advanced countries, has reached the world in a synchronized way. The importance of the theme follows on from the connections between the crisis and the process of globalization, causing different reflexes in the economies of the nations and, specifically, Portugal and Brazil. The initial point of the analysis is the interconnection between globalization and the economy, and to understand the Portuguese and Brazilian background, it addresses before the situation in the United States. Structural problems, the present situation and prospects are considered in the deliberations on the brothers countries. Opportunities for joint action to alleviate the internal consequences are discussed in the last part of this contribution.

Key words: Crisis, Globalization, bilateral relations.
Thematic Area: International and Integration Economics.

1. INTRODUÇÃO

A situação econômica e financeira vivenciada no Brasil, a realidade nos mesmos aspectos presente em Portugal e a posição desses países diante da atual crise mundial é o objetivo deste trabalho. As conseqüências da turbulência internacional nas economias locais são captadas por indicadores, como, por exemplo, contração do Produto Interno Bruto - PIB**,** diminuição do nível de confiança dos consumidores, retracção na concessão de crédito, além da geração de ambiente de incertezas

A importância do estudo decorre de considerar que as mudanças econômicas e outras a que estão sujeitas as nações que praticam a economia de mercado, causam reflexos em sociedades com especificidades e características singulares decorrentes do processo de globalização. Torna-se, portanto fundamental abordar o comportamento da economia dos Estados Unidos neste final

de milênio, haja vista que a referida crise vem do subprime ai iniciado e acabou se disseminando em diversas nações acarretando o enfraquecimento do consumo global.

Segundo Nunes (2002) desde 1994 os avanços na economia norte-americana tornaram-se importantes para os paises emergentes. Este fato é particularmente aplicável ao caso brasileiro como decorrência da consolidação do processo de abertura comercial, quando os mercados funcionavam independentes das decisões de politica monetária como as tomadas em reuniões do Comitê Federal de Mercado Aberto do Federal Reserve em Washington e, também, da implantação do Plano Real[1]. Relativamente a Portugal, a partir da década de 1980 o país iniciou um processo de modernização tendo ingressado na União Europeia - UE no ano de 1986 e experimentado um processo de privatização de empresas até então controladas pelo Estado, especialmente na área de telecomunicações. Isto oportunizou o crescimento da economia lusitana, hoje baseada principalmente em serviços.

Em termos metodológicos, este trabalho deve ser visto como um estudo analítico e comparativo entre as economias brasileira e portuguesa. Amostras de comportamento destas economias foram constituídas buscando compreender os principais reflexos da crise, as mudanças no modo de agir do consumidor e efetuar um entrelaçamento do papel de Portugal e do Brasil no cenário internacional. O intuito é o de viabilizar uma reflexão contemplando as articulações existentes entre as políticas adotadas e o aprofundamento da integração na busca da estabilidade econômica. Com efeito, decorrência de eventuais acontecimentos perversos na capacidade produtiva de cada país, a concepção e determinação das estratégias a serem implantadas pelos governos para atenuar as conseqüências da turbulência, devem ser referendadas dentro do que ocorra no cenário mundial. Na medida em que este experimenta freqüentes mudanças, essas ações necessariamente devem ser examinadas à luz de cada situação. Por esta razão, este estudo antes de considerar os acontecimentos no cenário brasileiro e de Portugal, abordará, também, as relações entre a globalização e a economia e, ainda, a origem e desdobramentos da crise no contexto da nação norte-americana

2. A GLOBALIZAÇÃO DA ECONOMIA

Ao falar em globalização é importante entender a utilização do termo, levando-se em consideração as forças econômicas uma vez que estas impactam diretamente no crescimento ou diminuição do Produto Interno Bruto – PIB, nas elevações e quedas das taxas de juros, no processo inflacionário e na variação cambial. No mercado, essas mudanças podem ser consideradas como oportunidades ou ameaças.

A discussão sobre globalização da economia mundial hoje centra no que se poderia designar um novo nível de complementariedade entre empresas e dentro das empresas internacionais, no âmbito regional e mundial, fruto de uma concorrência mundial extremamente acirrada, que inclui também a disputa dos mercados emergentes pelos grandes grupos industriais que produzem para escalas supra-nacional, regional e mundial (Klagsbrunn, 1999).

De uma forma geral a globalização proporcionou melhorias e inovações cientificas porém o *scanning* ambiental traduz a análise de informações sobre quais são as tendências relevantes. Wright e outros (2000) demonstram algumas forças macroambientais que oportunizam a percepção geral das mudanças, subdivididas quanto à natureza, em político-econômicas, sociais, econômicas e tecnológicas, conforme descrição efetuada no Quadro 1.

[1] O Plano Real constituiu uma tentativa de redefinição da estratégia de desenvolvimento nacional, buscando viabilizar o processo de redistribuição de renda, o controle da inflação, a valorização cambial e a promoção de abertura comercial. É importante ressaltar que o processo de concentração de renda no Brasil, tem origem na fase da colonização do País, consolidando-se a partir da década de 1930.

Quadro 1: Forças Macro ambientais e Mudanças de Cenário

Forças Político-económicas	**Forças Sociais**	**Forças Económicas**	**Forças Tecnológicas**
Leis tributárias	Atitudes em relação a inovação de produtos, estilo de vida, carreiras e activismo dos consumidores	Fornecimento de dinheiro	Gastos em pesquisa e desenvolvimento (do governo e do sector)
Regulamentações do Comercio Internacional	Preocupação com a qualidade de vida	Politica monetária	Enfoque em gastos com pesquisa e desenvolvimento
Regulamentações de empréstimos a Consumidores	Expectativas de vida	Taxa de desemprego	Taxa de introdução de novos produtos
Leis de protecção ambiental	Expectativas em relação ao trabalho	Custos com energia	Automação
Sanções de regulamentações anti trust	Mudança na presença de mulheres na força de trabalho	Renda pessoal disponível	Robótica
Leis de contratação, demissão, promoção e pagamento	Taxas de nascimento	Estágio do ciclo económico	
Controle de preços e salários	Alterações populacionais		

Fonte: Wright, Kroll e Parnell (2000), p. 59

A análise da economia globalizada oferece a idéia de que os mercados estejam aproximados. Entretanto, esta interação econômica crescente deve ser vista com cautela face às crises internacionais estabelecidas nos últimos anos. Dentre essas podem ser citadas:

1929 - A euforia de Wall Street quando em 29 de outubro conhecida como a "terça feira negra", a bolsa de Nova York obteve queda aproximada de 12%, houve declínio da produção e alastrou-se a falência de empresas iniciando um processo de desemprego. Este momento ficou conhecido como a grande depressão.

1939 - Segunda Guerra Mundial fez com que houvesse o declínio da atividade econômica em nações industriais, houve queda na produção e o crescimento do desemprego.

1973 - A guerra entre os países Árabes e Israel permeou a especulação do Petróleo, formalizou-se o cartel da OPEP, interrompendo o fornecimento aos EUA, houve disparada no preço do barril e o impacto entre as economias de países considerados emergentes.

1982 - A Moratória Mexicana que desencadeou uma série entre outros países da America Latina.

1987 - O declínio de cerca de 22% na bolsa de Dow Jones fez com economias do mundo inteiro sofressem o impacto dessa queda.

1994 - Conhecida como o efeito tequila houve queda das reservas internacionais Mexicana, desvalorização da Moeda, obrigando o país a recorrer ao Fundo Monetário Internacional.

1997 - A Crise dos Tigres Asiáticos teve inicio com a desvalorização da moeda Tailandesa, demonstrando como a globalização influenciou rapidamente nos efeitos negativos para os países Asiáticos. Esta conjuntura influenciou a Bolsa de Nova York e outros mercados.

1998 - A Moratória da divida externa russa com a desvalorização de moeda local, a especulação financeira obrigou o Fed a solicitar aporte de bancos privados demonstrando que os EUA também estavam vulneráveis.

2000 - Conhecida como o "estouro da bolha" uma vez que as bolsas Nasdaq e Dow Jones que vinham do boom em anos anteriores tiveram quedas recordes de 9,67% e 5,66% respectivamente. Esse declínio atingiu diversas economias mundiais.

2001 - O ataque terrorista de 11 de setembro provocou queda no mercado financeiro mundial houve perdas com a bolsa de Nova York ficando sem operar por 4 dias e em sua retomada do pregão o índice Dow Jones teve o pior desempenho nos últimos tempos chegando a

casa dos 7,13% Neste mesmo ano houve o escândalo das empresas Enron e Worldcom com prejuízos contábeis não divulgados provocando incertezas e falta de credibilidade

2002 - A crise da economia Argentina, os pacotes econômicos, a desvalorização da moeda, a eleição presidencial e a conturbação social assustaram os mercados da America Latina.

2007 - A forte baixa apresentada pela bolsa chinesa com quedas do Xangai Composto em cerca de 8,9%, a maior baixa diária em mais de dez anos. O Shenzhen Composto recuou 8,5%. Analistas diziam que não havia notícias negativas que justificassem essa queda e a atribuíam à realização de lucros após as contínuas altas do mercado desde a metade de 2005. Essa queda atingiu os mercados Filipinos, de Hong Kong, Taiwan, Sidney e Malásia e Indonésia.

2007 - A economia americana vinha intensificando o estímulo ao mercado imobiliário com juros mais baixos para pessoas com pouca capacidade financeira desde que sofreu o ataque terrorista de 11/09/2001. A crise deste setor intensifica-se quando o nível de emprego e a capacidade de investimento empresarial decrescem. A elevação na taxa de juros faz com que a capacidade de pagamento desses empréstimos não fosse possível, surgindo mais oferta de imóveis e desvalorização dos mesmos, provocando crise no mercado financeiro imobiliário.

Na interligação dos mercados, de acordo com Friedman[2] (2000) citado por Chang[3] (2008) os países do mundo da oliveira não poderão entrar no mundo dos Lexus a não ser que praticassem um conjunto de políticas econômicas muito restritivas (camisa de força). Assim, Chang ao interpretar as idéias de Friedman quando demonstra a ortodoxia econômica neoliberal de hoje e o caminho para o sucesso na nova economia global descreve:

O país precisa privatizar as empresas públicas, manter a inflação baixa, reduzir o tamanho da burocracia do governo, equilibrar o orçamento (até mesmo gerar superávit), liberar o comércio, desregulamentar o investimento estrangeiro e os mercados de capitais, tornar a moeda conversível, reduzir a corrupção e privatizar a previdência. (Chang, 2008)

Ainda segundo Chang (2008) todas as economias bem sucedidas, desenvolvidas e em desenvolvimento, chegaram ao ponto em que estão por meio da integração seletiva e estratégica com a economia mundial, e não pela integração global incondicional. O autor acredita que o livre comércio foi "imposto" e não opcional para os países que estão em desenvolvimento.

A globalização oferece sensibilidade e diversos riscos financeiros. Desta maneira, se faz necessário estar atento a políticas macroeconômicas uma vez que estas determinam as mudanças do comportamento das economias mundiais. Charles Goodhart e Avinash Persaud[4] (2008) consideram que as crises no âmbito do mercado financeiro não surgem de forma aleatória, mas ocorrem após períodos de prosperidade (*boom*). Estes são abastecidos pelas estimativas de que os riscos são baixos. Esse otimismo encoraja procedimentos tais como concessão de empréstimos de forma imprudente, que eventualmente conduzem à próxima crise.

3. ECONOMIA NORTE-AMERICANA

A atual crise que se alastrou por todos os continentes a partir dos Estados Unidos, requer uma reflexão de como a situação estaria se configurando nesse país. Paralelamente, a recuperação da economia mundial passa necessariamente pela economia norte-americana.

Os Estados Unidos, assim como o Reino Unido, a Espanha (tradicional parceira e país vizinho a Portugal) e, em menor escala, o Japão e a Alemanha enfrentam ameaça concreta de deflação e passam por uma destruição de riqueza e impacto no sistema bancário de proporção tal que levou grandes organizações creditícias a depender de socorros governamentais para evitar a quebra. Por outro lado, uma deflação agravaria os problemas dessas instituições e do próprio crédito nesses países, uma vez que aumentaria o valor real das dívidas, impondo uma sobrecarga adicional aos respectivos habitantes angustiados com o crescente desemprego, e as empresas, cujos mercados restringem-se e o faturamento desaba.

---

[2] T. Friedman (2000), "The Lexus and the Olive Tree", Anchor Books, New York, p.31.

[3] Joan Chang Ha (2008), "Maus Samaritanos – O Mito do Livre Comércio e a História Secreta do Capitalismo", Elsevier, Rio de Janeiro, p. 19.

[4] Charles Goodhart e Avinash Persaud (2008). "How to Avoid the Next Crash", *Financial Times*, January, 30, Disponível em http://iepecdg.com/DISK%201/Arquivos/Leiturassugeridas/080130Goodhart_Persaud.pdf . Acesso em 24/03/09

Munhoz (2008) pondera que a crise teve inicio anteriormente a 2008. Desde o início do Século XXI, após os atentados ao World Trade Center no ano de 2001, os Estados Unidos passaram a viver um período de incerteza econômica e de insegurança nacional. Para saná-los, o governo local, entre outras providências, elevou fortemente os gastos em armamentos, reduziu a taxa de juros e cortou a carga tributária das empresas. O país conseguiu sair rapidamente da situação em que se encontrava, mas ampliou as contradições que explodiriam anos à frente. O mercado imobiliário logo entrou em expansão[5]. Comprar casas passou a ser objetivo de quem desejava além do imóvel próprio, fazer algum investimento. Empresas hipotecárias, bancos e financeiras começaram a emprestar e financiar cada vez mais. Surgiram os chamados "subprimes" (ou operações de risco), os clientes de um segmento de renda mais baixa, segmento este, constituído também de mutuários que não conseguiam facilmente comprovar renda e/ou tinham algum histórico de inadimplência. Em 2005 ocorreu o crescimento explosivo de financiamentos de retorno incerto, em que foram protagonistas tais mutuários e outros com tradição de maus pagadores, ou trabalhadores de baixa renda, sem ocupação fixa e, muitas vezes, emigrantes ilegais.

Transcorridos dois anos, esses contratos, facilitados que haviam sido por baixos juros e prestações reduzidas passam a outra fase na qual o aumento dos juros traduzia encargos acima da capacidade de pagamento dos compradores. Conseqüentemente, centenas de milhares de compradores de casa própria através desses financiamentos classificados como de "subprime", deixaram de pagar as prestações acarretando prejuízos aos bancos hipotecários, os quais iniciam um processo de retomada dos imóveis que haviam financiado. Um novo cenário passou a surgir com o trancamento do crédito, diminuição dos meios de pagamento e outros inconvenientes. Quanto aos bancos hipotecários, prossegue Munhoz (2008, p.16), a baixa contábil dos mencionados prejuízos penalizou e aniquilou o capital dessas instituições. Como estas haviam emitido e vendido títulos lastreados nos créditos hipotecários, "o que significa que eram credoras dos compradores de casa própria e devedoras dos aplicadores que lhes forneciam recursos", igualmente enfrentam prejuízos dos agora desvalorizados títulos emitidos pelos referidos bancos hipotecários.

Na seqüência, os imóveis retomados dos devedores inadimplentes se desvalorizaram na mesma intensidade em que se elevavam as execuções das hipotecas. Este fato explica o surgimento de programas emergenciais para a recapitalização de bancos implementados no território norte-americano e em outros países (Alemanha, por exemplo). Infelizmente, nos Estados Unidos se sucedem pacotes incorporando reciclagem das injeções do Tesouro e empréstimos aos bancos, ajuste nas contas do Federal Reserve, aquisição de carteira de títulos de créditos hipotecários não pagos, e aquisição de ações de bancos cujo capital tenha minado na compensação de prejuízos[6]. A cada proposta de ajuda governamental, um novo pacote é requerido e, o mais intrigante, a crise do setor imobiliário não cede. Até o momento, ao contrário do que deveria ter acontecido, não foram concedidos subsídios diretos aos devedores mais frágeis. Com efeito, apenas aconteceram acenos de estímulos, notadamente através dos estados/municípios, no setor de infraestrutura, além de corte de tributos e possível rolagem de débitos hipotecários.

Conforme reconhece Freitas (2008), atrás da bolha[7] do "subprime" estiveram aspectos atrelados a uma inadequada supervisão bancária, a taxa de juros baixa, exportações de manufaturas chinesas a preços reduzidos e financiamento abundante do Banco Central chinês (People's Bank of China).

---

[5] Aproximadamente 80% do aumento do emprego e quase dois terços do incremento do PIB nos Estados Unidos, nos últimos anos, teria se originado, direta ou indiretamente no setor imobiliário [Joseph E. Stiglitz, (2003), "Os exuberantes anos 90: uma nova interpretação da década mais próspera da história", Tradução de Sylvia M.S.C. Santos e Dante M. Aldrighi, Companhia das Letras, São Paulo

[6] Dércio G. Munhoz (2008), "A Crise do subprime com a cor verde-amarela", *Revista de Conjuntura*, Ano VIII, nº. 36, outubro/dezembro, pp. 16-17.

[7] De acordo com Kindleberger (1992), uma bolha especulativa é um aumento significativo do preço de um ativo ou conjunto de ativos, em um processo continuo, no qual o aumento inicial gera expectativas de novos aumentos no futuro, atraindo novos compradores, geralmente especuladores interessados nos lucros que podem usufruir com a venda de tal ativo e não devido ao uso produtivo do mesmo. Aliás, os seres humanos sentem prazer em correr riscos. Mencionado processo costuma ser seguido por uma reversão de expectativas e um grande colapso dos preços, o que leva a uma crise financeira. Isto acontece quando a bolha estoura. Porém, nunca se sabe o momento exato em que haverá a conversão dessa bolha em crise.

Com cerca de 300 milhões de habitantes, os Estados Unidos possuem um PIB de aproximadamente US$13,8 trilhões. Para efeito de comparação, no caso dos países da União Européia, ele é pouco superior, ou US$16,8 trilhões, isto é, sozinho os EUA produzem quase o mesmo que todo o bloco europeu. As importações norte-americanas totalizam US$ 2 trilhões em produtos e serviços do resto do mundo e pela Bolsa de Nova York são negociados US$35 trilhões em ações todo ano, enquanto na Bolsa NASDAQ, voltada para ações de empresas do ramo tecnológico, são outros US$17 trilhões. Com estes números, fica evidente que esse país é o centro do capitalismo e quais os malefícios inerentes a crise atual.

As indicações são no sentido de que os preços dos imóveis continuarão regredindo uma vez que os principais sustentáculos do crescimento da economia norte-americana - gastos dos consumidores, empresas, governo e as exportações - continuam em declínio. Ademais, a demanda externa por bens americanos, que ajudou o setor industrial a evitar uma desaceleração mais profunda em 2008, deve diminuir à medida que as maiores economias mundiais sofrem com a recessão e, também algumas nações de rápida expansão (caso da China e da Índia).

Os dados estatísticos recentes sobre a situação econômica norte-americana não são encorajadores. No quarto e último trimestre de 2008 e ao início de 2009 o PIB e o consumo estão em declínio, com o primeiro tendo sofrido uma retração de aproximadamente 7% nesse período. Enquanto isso houve retração de 21,1% nos investimentos privados, redução de 23,6% nas exportações e, somente em fevereiro/2009, registrou-se a perda de cerca de 700 mil empregos, cifra idêntica às apuradas em dezembro de 2008 e janeiro deste ano, o que eleva a taxa de desemprego para 8,1%, a mais alta dos últimos 25 anos A economia, desde dezembro de 2007, perdeu quase 4,5 milhões de empregos, a maior perda desde a 2º Guerra. Lamentavelmente os trabalhadores lá estariam enfrentando a fome e a degradação das condições de vida. Adicionalmente, houve aumento brutal dos dependentes de cupons alimentares, já usados por 31 milhões de pessoas que passam fome e privações, isto é, um em cada dez estadunidenses (Borges, 2009).

Por último mas não menos relevante, a duração da crise nos Estados Unidos estaria condicionada, além das intervenções pontuais diretas, ao tempo a que os sistemas bancários formal e sombra (representado pelos bancos e fundos de investimento) levar para reduzir a excessiva alavancagem[8], o que significa redução das carteiras de empréstimos (contração de crédito) e aumentos de capital, inclusive mediante a utilização de dinheiro público. Esta contração também está atrelada ao periodo necessário para que as famílias, indivíduos e empresas alcançados pelo empobrecimento restabeleçam o equilíbrio patrimonial e voltem a gastar normalmente.

4. ECONOMIA BRASILEIRA

No Brasil o grande desafio é a discussão profunda para saber o que é desejável na relação Estado/Sociedade. Desta forma, a busca para manter a sustentabilidade econômica envolve dilemas como, por exemplo, a redução da pobreza e da desigualdade social bem como o crescimento sustentado. O grande paradigma para ser quebrado é o desempenho do setor público, onde a máquina administrativa apresenta um padrão de gasto considerado por vários autores, como ineficiente.

Mendes[9] (2007) enumera como condição necessária para a eficência do governo o cumprimento de seis funções:

Garantia do direito de propriedade e demais direitos e liberdades individuais: sistema judicial, policial e prisional;

[8] Alavancagem = captação de depósitos em relação ao capital próprio.

[9] M. J. Mendes (2007), "Ineficiência do Gasto Público no Brasil" in R.B. Bouri. e M.M. Saboya (Org.) (2007), "Aspectos do Desenvolvimento Fiscal". Cap. 11, Instituto de Pesquisa Econômica e Aplicada - IPEA, Brasília, p. 125.

Estabilidade macroeconômica: controle da inflação, crescimento econômico, suavização dos ciclos econômicos, equilibrio do balanço de pagamentos;

Redução da desigualdade e da pobreza: acesso a condições mínimas de saneamento básico, saúde, educação, habitação e assistência social;

Oferta ou subsidio à produção de bens públicos e de bens geradores de externalidades positivas: inovação tecnológica, universalização da educação, proteção e recuperação ambiental, segurança nacional e construção de rodovias;.

Regulação de monopólios naturais e defesa da concorrência; e

Solução de problemas de informação imperfeita: disponibilidade de informações sobre risco de crédito, financiamento educacional, regulação de planos de saúde e dos planos de seguros.

A economia brasileira, considerada por alguns economistas nos últimos cinco anos como de estabilidade macroeconômica e criação de condições de crescimento sustentado, inspira hoje preocupação. Com efeito, em 2008 iniciou um processo de desaceleração do Produto Interno Bruto - PIB, conjugado a aumento do desemprego com demissões em massa em alguns segmentos, redução na concessão de crédito, diminuição de investimentos produtivos e decréscimo do consumo doméstico.

Segundo Almeida (2008) países que praticam economia de mercado possuem interesses contraditórios entre os níveis de risco e de retorno (sempre opostos) associados às várias opções abertas à aplicação das poupanças individuais e coletivas, e, de outro, os diferentes mecanismos de regulação institucional existentes no vasto mercado de gerenciamento dessas poupanças. Assim, quem faz a concessão de crédito ou efetua aplicação financeira dos respectivos ativos espera que a taxa de juros seja a maior possível. Entretanto, os que se encontram na contramão estão ávidos para que esses juros se encontre no menor patamar possível. O autor apresenta, sinteticamente, as grandes crises financeiras ocorridas em países que tem praticado esse tipo de economia nos últimos 80 anos.

Quadro 2: Crises Financeiras Internacionais: Cronologia Últimos 80 Anos

Período	Fatores dominantes	Efeitos	Resultados institucionais
1928-1939	▪ Especulação com ações; ▪ Quebra da Bolsa NY (1929); ▪ Enxugamento da liquidez; ▪ Aumento do protecionismo; ▪ Não-conversibilidade cambial.	▪ Perdas patrimoniais; ▪ Quebra de instituições financeiras; ▪ Cessação de pagamentos; ▪ Controle de capitais; ▪ Recessão e depressão por toda a década.	▪ Ajuste pelo mercado, seguido de medidas keynesianas mais adiante prolongaram o *timing* da recuperação; ▪ Medidas restritivas no setor bancário.
1965-1975	▪ Déficits comerciais e fiscais nos EUA; ▪ Insuficiência de ouro para atender às obrigações do garantidor; ▪ Pressão por troca de ativos (mas apenas França é bem sucedida); ▪ Suspensão do padrão ouro-dólar pelos EUA.	▪ Desvalorização imediata do dólar traz grandes prejuízos à Alemanha e ao Japão; ▪ Anarquia monetária e depreciação das *commodities* levam ao aumento do dólar e a turbulências financeiras; ▪ Primeira alta do petróleo.	▪ Incapacidade do sistema em definir um novo padrão monetário; ▪ Modificação do convênio do FMI e flutuação de moedas; ▪ Acordos do Smithsonian e do Plaza levam à desvalorização administrada do dólar.
1979-1985	▪ Grande déficit comercial e fiscal nos EUA; ▪ Excesso de liquidez traz enorme endividamento na América Latina; ▪ Aumento dos juros nos EUA e crise da dívida no México e no Brasil (1982).	▪ Atração de capitais para os EUA e revalorização do dólar; ▪ Inadimplências na América Latina com processo de transferência de recursos da periferia para o centro; ▪ FMI comanda planos de ajuste.	▪ Bancos constituem comitês assessores para renegociar dívidas dos países afetados; ▪ FMI cria novos programas de ajuste; ▪ Tesouro - EUA oferece planos de trocas da dívida (Baker e Brady).

Quadro 2 (cont.)

Período	Fatores dominantes	Efeitos	Resultados institucionais
1994-2002	▪ Abundância de fluxos inunda os mercados emergentes de capitais de todo o tipo; ▪ Rigidez cambial estimula ataques especulativos às moedas; ▪ Déficits prolongados de transações correntes	▪ Desvalorizações desastradas desde o México (1994) dão início a quebras em série em países da Ásia e na Rússia (1998); ▪ Brasil faz acordos preventivos com o FMI (1998, 2001, 2002).	▪ FMI desiste de liberalizar movimentos de capitais; ▪ Países aderem à flutuação; ▪ G20: Fórum de Estabilidade Global (G8 + 12 países emergentes); ▪ BIS: Basiléia II; ▪ Aumento de reservas.
2005-2008	▪ Juros exageradamente baixos e expansão dos empréstimos hipotecários, transformados em derivativos financeiros; ▪ Crise das subprimes levou à inadimplência generalizada.	▪ Quebra de bancos de investimentos e contaminação de outras instituições; ▪ Governos montam pacotes de resgate e injetam liquidez no sistema.	▪ Processo em curso: provável reforço das regras de controle sobre derivativos, com possível engessamento do sistema; ▪ Estatização temporária de instituições.

Fonte: Adaptação de Paulo R. Almeida (2008)

Atualmente existe certo otimismo por parte do governo. O Ministro de Estado da Fazenda, Guido Mantega[10] afirma que a economia brasileira foi à última a desacelerar o crescimento na atual crise e será a primeira a sair dela. Segundo essa autoridade, o Brasil tem o desafio de impedir que essa turbulência aborte o ciclo de desenvolvimento econômico e social que o país viveu nos últimos anos. Este otimismo foi tema de pronunciamento[11] onde foram elencadas e classificadas 10 (dez) razões.

Reservas de US$ 200 bilhões intocadas após seis meses de crise: a utilização de apenas US$ 5 bilhões das reservas internacionais é um dos principais fatores de estabilidade;

Bancos competentes, regulados, com baixa exposição a riscos: Instituições financeiras brasileiras (as maiores delas) com ativos saudaveis e cautelosas na concessão de empréstimos;

Ausência de bolhas de crédito e imobiliária, com potencial de crescimento real desses setores: O total de crédito disponível hoje na economia brasileira equivale a 40% do PIB, um volume muito inferior àquele observado nas nações desenvolvidas e até mesmo em boa parte das emergentes;

Mercado interno forte, crescendo em poder de compra e em proporção da população: Declínio da pobreza, a chamada classe C (renda familiar entre R$ 1.115,00 e R$ 4.807,00 mensais) é a nova dominante no País formada por consumidores emergentes;

Matriz Energética mais "Verde" do mundo, com Independência do Petróleo Importado: Utilização de combustível originário de fontes renováveis (energia hidráulica, etanol / cana-de-açucar e grande potencial hídrico existente na região Norte, além das megareservas do pré-sal);

Estabilidade política em que a democracia foi entronizada como patrimônio nacional: As previsibilidades econômica e política são alicerces para o desenvolvimento duradouro. O Brasil de hoje é resultado de um processo ocorrido nos últimos quinze anos, nos quais foram estabelecidos poderosas restrições culturais e institucionais ao populismo;

Estabilidade econômica e arcabouço regulatório imperfeito mas previsível: O país completa dez anos sob o mesmo regime de câmbio flutuante e metas de inflação. Esses instrumentos, aliados às metas de superavits primários, asseguram a previsibilidade da gestão econômica, atraindo investimentos produtivos;

[10] Guido Mantega (2009), "Brasil será o primeiro país a sair da crise", Discurso proferido na instalação do Conselho Administrativo de Recursos Fiscais - CARF e posse do primeiro presidente da entidade, Carlos Alberto Barreto, 04 de março, Disponível em www.fazenda.gov.br/audio/2009/marco/a040309.asp Acesso em 06/03/09

[11] "O Brasil e a Crise Mundial - Dez Razões para o Otimismo" (2009), *Veja,* Ano 42, nº. 9, Edição 2102 - Seção Economia, 04 de março, pp 82 - 88.

Maior exportador de alimentos do mundo, o que garante vendas externas volumosas em qualquer cenário: O país possui a maior fronteira agrícola, grande potencial de crescimento do agronegócio com as commodities (soja, carne) sendo responsáveis pelo maior volume de exportação;

Mercado externo diversificado com compradores em todo o mundo e mercadorias de valor agregado crescente: Grande diversificação de produtos bem como de compradores em diversas partes do mundo, permitindo certa flexibilidade mas, não garantindo as possíveis perdas com diminuição da demanda mundial;

As mesmas projeções que apontam estagnação no mundo estimam crescimento do PIB do Brasil em 2009: A expectativa é de que o Brasil cresça em torno de 1,5%.

Apesar do otimismo e a esperança de que o Brasil não sofra tanto com o impacto da crise mundial, é necessário estar alerto pois esta não se restringe ao âmbito financeiro, existindo sinais de desaquecimento em diversos setores: industrial, comercial, exportação entre outros. Decisões foram tomadas pels autoridades governamentais objetivando restaurar a liquidez (diminuição do compulsório incidente sobre depósitos a vista dos bancos), prevenir falências bancárias (os estabelecimentos de crédito oficiais - Banco do Brasil e Caixa Econômica Federal - foram autorizados a absorver bancos menores), aliviar impostos (redução do imposto incidente sobre a aquisição de veículos novos permtindo aos compradores pagarem menos e aos produtores receberem mais) e para desencorajar a dispensa de empregados (incidência de contribuição previdenciária sobre aviso prévio indenizado). Entretanto, esta última providência não está surtindo o efeito desejado pois, segundo estatisticas do Ministério do Trabalho e Emprego inseridas no Quadro 3 e referentes ao mês de fevereiro de 2009, os setores de atividade onde o desemprego continua elevado e sendo sentido de forma mais aguda são a industria da transformação, o comércio e a indústria extrativa mineral.

Por outro lado, deve ser levado em consideração que a contração econômica poderá vir a deflagrar um estado de alerta vermelho para a busca pelo emprego. Como tentativa de minimizar o problema, hoje, grande parte dos próprios trabalhadores pertencentes a setores onde o fantasma da massificação de demissões se faça presente, negocia acordos de redução de jornada de trabalho e de salários buscando assim minimizar os desligamentos. Cabe lembrar que os informes do Quadro 3 não incluem os trabalhadores informais, muitos dos quais fazem parte de uma massa eliminada do emprego formal.

Seja como for, o grande desafio do governo brasileiro é implantar politicas monetária e crediticia e outras que também fomentem as micro, pequenas e médias empresas. Desta forma, caberia repensar a taxa básica de juros - a Taxa Selic - para que haja maior estímulo a economia, principalmente no segmento em referência, o grande responsável, em qualquer pais, pela geração de empregos. Em resumo, neste ambiente, é necessário a sinergia entre governo e sociedade dando à crise global a devida importância e estando atento para as externalidades do mercado.

QUADRO 3: Brasil = Evolução do Emprego por Setor de Atividade Econômica

	FEVEREIRO/2009			NO ANO			EM 12 MESES					
ATIVIDADE ECONÔMICA	TOTAL ADMISSÃO	TOTAL DESLIGAMENTO	SALDO	VARIAÇÃO EMPREGO %	TOTAL ADMISSÃO	TOTAL DESLIGAMENTO	SALDO	VARIAÇÃO EMPREGO %	TOTAL ADMISSÃO	TOTAL DESLIGAMENTO	SALDO	VARIAÇÃO EMPREGO %
Extrativa Mineral	2.882	3.587	-705	-0,41	5.866	7.030	-1.164	-0,67	52.102	46.052	6.050	3,68
Indústria de Transformação	221.522	277.978	-56.456	-0,77	438.321	549.907	-111.586	-1,51	3.367.888	3.406.656	-38.768	-0,54
Serviços Utilidade Pública	6.654	5.847	807	0,23	13.557	12.037	1.520	0,43	72.081	65.060	7.021	2,09
Construção Civil	138.627	135.785	2.842	0,15	290.092	275.926	14.166	0,73	1.859.293	1.713.476	145.817	9,07
Comercio	270.441	280.716	-10.275	-0,15	553.509	614.565	-61.056	-0,86	3.753.728	3.432.228	321.500	4,96
Serviços	482.005	424.487	57.518	0,45	934.147	874.177	59.970	0,47	5.849.931	5.265.220	584.711	5,05
Administração Publica	20.573	6.082	14.491	1,86	32.251	15.526	16.725	2,15	110.112	98.506	11.606	2,10
Agropecuária	90.850	89.893	957	0,06	182.361	193.505	-11.144	-0,72	1358.470	1384.656	-26.186	-1,70
TOTAL	**1.233.554**	**1.224.375**	**9.179**	**0,03**	**2.450.104**	**2.542.673**	**-92.569**	**-0,29**	**16.423.605**	**15.411.854**	**1.011.751**	**3,44**

Fonte: Brasil. Ministério do Trabalho e Emprego - Cadastro Geral de Empregados e Desempregados

5. ECONOMIA PORTUGUESA

Portugal está entre os países na zona do Euro, com renda per-capita e IDH bastante significativos cabendo destacar que a estabilidade da economia vem sendo gradativamente envolvida na crise financeira mundial. Incertezas foram sinalizadas para o mercado, com o aparecimento de desequilibrios potencialmente inflacionários . Assim como em várias partes da União Européia, ela reflete restrição na concessão de crédito, insegurança por parte dos consumidores com temor relativamente ao futuro e declínio do Produto Interno Bruto.

Inicialmente, é importante ressaltar que a União Européia acabou sendo atingida pela crise através de canais distintos. Em primeiro lugar o da bolha imobiliária norte-americana, pois vários bancos da região tinham nas respectivas carteiras títulos derivativos dos créditos hipotecários dos Estados Unidos. Em segundo, está o canal do crédito, uma vez que os bancos em consideração tiveram que correr para cobrir os prejuízos. Este fato, em situações normais, aconteceria através de operações interbancárias, porém o problema foi de que o temor de inadimplências estendeu-se inter bancos e acarretou a necessidade dos bancos centrais dos países europeus injetarem dinheiro no sistema. O crédito acabou totalmente travado. Na

seqüência desses canais, e não menos relevante, está o da economia real com os sinais dados pelo retrocesso na economia norte-americana e significando queda na produção, desemprego, renda e outros malefícios. O continente europeu foi à região mais atingida e para Portugal não houve alternativa a não ser adotar providências de atenuação dos impactos. Em princípio, a eficácia das medidas que tomasse variariam na razão inversa do tempo em que a crise perdurasse.

Segundo Ribeiro(2008)[12] a economia portuguesa atravessa um período de estagnação econômica que se arrasta desde 2002, constituindo, assim, na história local e desde várias décadas, o mais longo espaço de tempo com ausência de progresso econômico. Nesta situação conjugam-se problemas estruturais por resolver, enquadramentos econômicos externos penalizadores da especialização produtiva nacional e erros graves de estratégia e de gestão econômica dos últimos governos.

A análise do cenário português nos mercados financeiros demonstra que os bancos lusitanos desenvolvem atividades em ambiente pouco favorável, principalmente quanto ao respectivo financiamento nos mercados Internacionais. Nessa contribuição obtida externamente, observa-se que a canalização de recursos aos bancos locais deixa a desejar. Por esta razão, na opinião de SANTOS (2009), para atenuar a crise e ajudar a combatê-la, medidas diversas, algumas extrapolando a alçada do governo são requeridas. Em última instância, em face a este quadro, deveria haver um plano global de estabilização financeira e de apoio às economias dos diversos países incluindo Portugal, promoção do comércio exterior e estímulo aos movimentos internacionais de capitais, evitando ações protecionistas. No território português, injeções de liquidez no sistema financeiro com taxas de juros baixas estariam no centro de uma política monetária a ser levada adiante. Além disto, no contexto da União Européia, e de Portugal em particular, seriam recomendáveis iniciativas de reforço da estabilidade financeira (concessão de garantias e aumentos de capital por parte do Estado visando assegurar o financiamento da atividade econômica e garantir a proteção das poupanças) e medidas de apoio à economia real, tais como o estímulo à procura, o aumento seletivo do investimento público, apoio às empresas e ao emprego e ações de natureza social.

Os dados recentes sobre a economia portuguesa mostram que a industria de tranformação estaria condicionada pela redução das exportações de bens e da demanda interna. Consequentemente, no atual contexto do sistema financeiro, seria perfeitamente admissível um reforço do apoio do Estado para a estabilidade do comércio internacional, do país, tanto que na expectativa de favorecer o investimento e o emprego, o governo estaria assegurando uma cobertura adicional do risco de crédito nas vendas para paises componentes ou não da Organização para a Cooperação e Desenvolvimento Econômico - OCDE, ou seja, destinando para essa finalidade um montante de recursos equivalente a 2,5% do PIB, este estimado para 2007 em US$ 227 trilhões.

Quanto a área da construção civil, esta vem apresentando reduções sucessivas nos últimos anos, mas poderá haver alguma recuperação se acontecer investimento em habitação por parte das familias; entretanto, é forçoso reconhecer que a população tem optado fortemente por aplicação em poupança. Por outro lado, no segmento de serviços (expansão de 2% em 2008), a estimativa é de que haja desacelaração particularmente na área de turismo No campo governamental, as contas públicas deverão ter crescimento nulo, tendo em vista a deteriorização das receitas fiscais e como resultado das medidas de apoio econômico e social colocadas em prática, estas certamente vem contribuindo para o deficit do do orçamento[13].

Segundo o Instituto Nacional de Estatística - INE (2009), na União Européia a confiança dos consumidores e os indicadores de sentimento econômico fixaram mínimos históricos para as respectivas séries desde 1995. Com efeito, o nível de atividade e a expectativa dos consumidores continuam com trajetória de declínio afetando o consumo e o investimento, e isto

[12] José Cadima Ribeiro (2008), "A actual crise financeira mundial e a economia portuguesa", *Diário do Minho - Suplemento de Economia,* 02/12, Disponível em http://economiaportuguesa.blogspot.com/2008/12/actual-crise-financeira-mundial-e.html Acesso em 21/03/09

[13] *Boletim Econômico: perspectivas para a economia portuguesa: 2008-2009*, Disponível em http://www.bportugal.pt/publish/bolecon/verao_08/econ_port_verao08_p.pdf Acesso em 22/03/09

desde janeiro de 2008, enquanto as taxas do comércio internacional de bens relativo às importações passaram de - 6,4% para -15,8% (refletindo a evolução do preço do petróleo) e a das exportações de -11,0% para -19,7% (contração econômica observadas nos países clientes). Em fevereiro/2009, houve estabilização da inflação em 0,2% (variação mensal) enquanto a taxa de desemprego se encontrava em 7,8% tendo, no quarto trimestre de 2008, ocorrido redução de 19.500 postos de trabalho relativamente ao trimestre anterior (-0,4%). Ainda segundo essa mesma fonte, no último trimestre de 2008, o PIB português encolheu 2% em relação aos anteriores sendo o ritmo de crescimento nesse ano, considerado o mais baixo desde a recessão de 2003; no ano teria se fixado em zero por cento, depois de uma elevação de 1,9% em 2007.

Faz-se necessário manter a competitividade lusitana no mercado internacional e intensificar particularmente as relações com a Espanha, tradicional parceiro tanto no comercio externo quanto na empregabilidade de portugueses que lá trabalham seja para empresas locais quanto para empresas de compatriotas ali sediadas. Não deve ser esquecido que Portugal figura na listagem dos paises que mais recebe investimentos diretos do exterior e que este costuma ser menos afetado por crises financeiras, além de oferecer, entre outros benefícios, aumentos de produtividade, transferência de tecnologia e economias de escala. Adicionalmente, projetos em curso no país como por exemplo, nos domínios da energia, comunicações e turismo, ao "estarem em linha com as apostas estruturais do governo, constituem oportunidades de negócio a serem propagadas de forma globalizada" (Santos, 2009). Estes projetos poderiam minorar os efeitos da atual crise. Isto é válido inclusive para empresas brasileiras que deveriam ser lembradas da possibilidade de Portugal como uma "rampa de lançamento" para a União Européia. Aliás, vale mencionar que a Embraer - Empresa Brasileira de Aeronáutica que adquiriu a Indústria Aeronáutica de Portugal - OGMA, e a Companhia Siderúrgica Nacional que passou a LUSOSIDER - Aços Planos recentemente lá se instalaram.

6. CONSIDERAÇÕES FINAIS

A medida em que a atual crise, enraizada no setor financeiro, originária dos paises avançados, especialmente os Estados Unidos, atingiu o mundo de forma sincronizada e, também, que a recuperação da situação passa, necessariamente, pela reerguimento da economia norte-americana, uma apreciação desta aconteceu antes de enfocar os desdobramentos causados em Portugal e no Brasil.

Os efeitos da crise foram e continuam sendo sentidos de forma diferente em Portugal e no Brasil, os quais vivenciam momentos diversos no ciclo econômico. No caso brasileiro, o efeito de contágio da turbulência econômica e social demonstra avanços estruturais e uma autonomia frente ao mercado norte-americano bem maior do que no passado. O país prossegue como destino importante para o investimento estrangeiro. Relatório recente da Conferência das Nações Unidas sobre o Comércio e o Desenvolvimento - UNCTAD, o coloca entre as cinco mais atrativas economias do mundo para se investir. Ademais, na qualidade de país emergente, prossegue tendo melhor desempenho do que vários dos países europeus.

Portugal, inserido na União Européia, sofre mais com a crise. As autoridades lusitanas se esforçam para que os efeitos dela sejam minimizados. Não obstante, os indicadores econômico-sociais não se apresentam satisfatórios. A população está insegura com o curso dos acontecimentos e com expectativas pouco alentadoras quanto ao desenrolar dos fatos. Este comportamento é justificável tendo em vista a elevação do desemprego, queda de renda mensal e da própria massa salarial - volume de dinheiro em circulação diminuindo, inadimplência crescente e o crédito bancário, base de sobrevivência no mundo capitalista, cada vez mais difícil e, assim, emperrando as engrenagens do sistema.

Referente ao relacionamento Portugal e Brasil, indicadores preliminares relativos aos primeiros meses de 2009 demonstram que as empresas portuguesas estão resistindo melhor à crise e sofrendo menos o impacto da queda de procura. Considerando a balança comercial, historicamente a importância lusitana para os exportadores brasileiros supera a relevância do mercado no sentido inverso, fato que leva a acreditar que, em princípio a crise financeira estaria afetando mais as empresas vendedoras do Brasil. Alguns entraves a esse comércio bilateral foram equacionados. Exemplo são os acordos aduaneiros celebrados entre a União Européia e o

MERCOSUL diminuindo a parcela da tributação no preço internacional o que pode levar Portugal a tirar vantagens como facilitador de negócios entre o Brasil e a Europa e, paralelamente, utilizando e traduzindo benefícios inversos ao se recorrer do território brasileiro no sentido de "rampa de lançamento" para a América Latina como um todo e ao MERCOSUL em particular.

Cabe ressaltar que o Fundo Monetário Internacional - FMI e o Banco Mundial foram 'incapazes' de prever e controlar a desordem financeira atual. No contexto desta crise, cujo fim não é delineado e para a qual alguns especialistas argumentam não haver "uma luz no final do túnel" deve ser levado a bom senso a identidade existente entre Brasil e Portugal podendo vir a demonstrar que essa turbulência globalizada inaugura um novo período em que haverá mais qualidade nos negócios, nas relações entre pessoas/países e nos produtos.

BIBLIOGRAFIA

ALMEIDA, PAULO R. (1980), "Crises financeiras internacionais: uma síntese histórica dos últimos 80 anos", Texto para apresentação em fase de revisão, Brasília, 26-28/10, Disponível em www.pralmeida.org/05DocsPRA/1944CrisesFinancIntern.pdf Acesso em 24/03/09

ARONE**,** PRISCILA (2007), "Bolsa chinesa registra maior queda em mais de dez anos", Disponível em http://portalexame.abril.com.br/ae/financas/m0123447.html Acesso em 15/03/09

BOLETIM ECONÔMICO: perspectivas para a economia portuguesa: 2008-2009*,* Disponível em http://www.bportugal.pt/publish/bolecon/verao_08/econ_port_verao08_p.pdf Acesso em 22/03/09

BORGES, ALTAMIRO (2009), "A crise mundial está só no começo", São Paulo, 16/03, Disponível em http://feebbase.com.br/site/temas_e_debates/a_crise_mundial_esta_so_no_comeco_.htm Acesso em 28/03/09

BRASIL - MINISTÉRIO DO TRABALHO E EMPREGO (2009), "Evolução do Emprego por Setor de Atividade Econômica", Disponível em
http://estatistica.caged.gov.br/consulta.aspx?mesCPT=02&anoCPT=2009 Acesso em 20/03/09

CHANG, HA JOON (2008). "Maus Samaritanos: O Mito do Livre Comércio e a História Secreta do Capitalismo", Tradução de Celina M. Ramalho, Elsevier, São Paulo

FERRARI, ANDRÉS E CUNHA ANDRÉ M. (2005) "As origens da crise argentina: uma sugestão de interpretação", Disponível em
http://www.anpec.org.br/encontro2005/artigos/A05A023.pdf Acesso em 04/03/09

FREITAS, CARLOS EDUARDO (2008), "Crise Econômica Internacional", *Revista de Conjuntura,*Ano VIII, nº. 36, outubro/dezembro, pp. 9 - 15

GOODHART, CHARLES E PERSAUD, AVINASH (2008), "How to Avoid the Next Crash", *Financial Times*, January, 30, Disponível em http://iepecdg.com/DISK%201/Arquivos/Leiturassugeridas/080130Goodhart_Persaud.pdf Acesso em 24/03/09 http://pt.wikipedia.org/wiki/Economia_de_Portugal Acesso em 04 de março de 2009

INSTITUTO NACIONAL DE ESTATÍSTICA (2009), **"**Fevereiro de 2009: Actividade económica recua fortemente em Janeiro. Importações e Exportações nominais de bens apresentam reduções significativas". *Destaque*, março, Disponível em
http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=46456094&DESTAQUESmodo=2 Acesso em 22 /03/09

_____ (2009), *Principais Indicadores* Disponível em http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_princindic Acesso em 22/03/09

KINDLEBERGER**,** C. (1992) in Eatwell**,** J., Milgate, M. e Newman, P. (Org.),"The New Palgrate Dictionary of Money and Finance", Macmillan, London

KLAGSBRUNN**,** VICTOR H.(1999), Globalização da Economia Mundial e os Mercados de Trabalho", *Tempo Brasileiro*, Rio de Janeiro, v. 139, oututubo-dezembro, pp. 207-228, Disponível em
http://www.reggen.org.br/midia/documentos/globadaeconomiamundial.pdf Acesso 04/03/ 09

MANTEGA**,** GUIDO (2009). "Brasil será o primeiro país a sair da crise". Discurso proferido na instalação do Conselho Administrativo de Recursos Fiscais - CARF e posse do primeiro presidente da entidade, Carlos Alberto Barreto, 04 de março, Disponível em www.fazenda.gov.br/audio/2009/marco/a040309.asp Acesso em 06/03/09

MENDES, MARCOS J. (2007) **"**Ineficiência do Gasto Público no Brasil**"** in Boueri**,** Rogério B. e Saboya, Mauricio M. (Org.), "Aspectos do Desenvolvimento Fiscal", Cap. 11, Instituto de Pesquisa Econômica Aplicada - IPEA, Brasília

MUNHOZ, DÉRCIO G. A (2008), "Crise do subprime com a cor verde-amarela" *Revista de Conjuntura*, Ano VIII, nº. 36, outubro/dezembro, pp 16 a 20

NUNES, JOÃO M. M. (2002) "O vôo da Águia - A economia americana no fim do milênio", Saraiva, São Paulo

"O Brasil e a Crise Mundial - Dez Razões para o Otimismo" (2009), *Veja*, Ano 42, nº. 9, Edição 2102 - Seção Economia, 04 de março, pp. 82 - 88

RIBEIRO, JOSÉ CADIMA (2008***), "***A actual crise financeira mundial e a economia portuguesa***", Diário do Minho - Suplemento de Economia, 02/12, Disponível em* **http://economiaportuguesa.blogspot.com/2008/12/actual-crise-financeira-mundial-e.html** ***Acesso em 21/03/09.***

SANTOS, FERNANDO T. (2009) "O papel do sistema financeiro nas relações econômicas entre Portugal. Brasil e África". República Portuguesa: Ministério das Finanças e Administração Pública, 16/02, Disponível em http://www.mf.gov.pt/discursos/Apres_MF_20090216.pdf Acesso em 10/03/09

STIGLITZ, JOSEPH E. (2003) "Os exuberantes anos 90: uma nova interpretação da década mais próspera da história", Tradução de Sylvia M.S.C. Santos e Dante M. Aldrigh, Companhia das Letras, São Paulo

WRIGHT, PETER, KROLL, MARK J.E PARNELL, JOHN (2000). "Administração Estratégica: Conceitos", Tradução de Celso A. Rinolli e Lenita R. Esteves, Atlas, São Paulo

COMPORTAMENTO DO IDE E CRESCIMENTO ECONÓMICO: UMA AVALIAÇÃO PARA OS PAÍSES DA COESÃO[1]

VERA BARROS
E-mail: vgbarros@uma.pt
Universidade da Madeira
MARIA PAULA FONTOURA
E-mail: fontoura@iseg.utl.pt
Universidade Técnica de Lisboa

Resumo

Este estudo investiga o *Investment Development Path* (IDP), que relaciona a posição das saídas líquidas de IDE de um país com o seu crescimento económico, tomando como referência os países da Coesão. Para o efeito é estimado um modelo para o período 1980-2005, com base em expressões não lineares previamente utilizadas. Os resultados obtidos são posteriormente discutidos à luz de possíveis limitações do referencial do IDP, designadamente no que se refere à utilização da posição do IDE em termos líquidos e ao recurso ao produto *per capita* como indicador de desenvolvimento económico na relação estipulada com o IDE. O IDP de cada país é seguidamente reevaliado tendo em consideração a primeira limitação. Discute-se ainda o possivel comportamento do NOI no caso de se utilizarem indicadores mais adequados relativamente às saídas e entradas de IDE. Os resultados no seu conjunto confirmam as estimações efectuadas para Portugal e Espanha mas apontam para a possibilidade de a Irlanda se encontrar numa fase mais avançada do IDP assim como de a Grécia apresentar uma curva de IDP muito mais suave do que nos restantes casos.

Palavras-chave: investimento directo estrangeiro, *investment development path*, países da Coesão.
Área tematica: Economia Internacional e da Integração.

Abstract

This survey studies the Investment Development Path (IDP), which relates a country´s net outflow of FDI to its growth, taking as a reference the cohesion countries. To that end a model for the period 1980-2005 is estimated, based on non-linear expressions previously used. The results are then analyzed within a context where possible shortcomings to IDP´s framework are brought to the fore, namely for using FDI in net terms and the choice of GDP *per capita* as a proxy for economic development. Each country´s IDP is then reassessed taking into consideration the first of these two shortcomings. The effects on NOI of taking in more adequate indicators for both in and outflows of IDE are also discussed. The results taken as a whole confirm the estimates made for Portugal and Spain but also point to the possibility of Ireland being at the a more advanced stage of IDP, and of Greece showing a much smoother IDP curve than what it might have been should we have used the other assumptions.

Keywords: foreign direct investment, investment development path, Portugal, cohesion countries
Thematic Area: International and Integration Economics.

[1] Agradece-se o apoio financeiro da FCT (Fundação para a Ciência e a Tecnologia). Este artigo é parte do Projecto de Financiamento Multi-anual (POCI/U0436/2006).

1. INTRODUÇÃO

Dunning (1981, 1986) e Dunning e Narula (1996) argumentaram que a posição das saídas líquidas de investimento directo estrangeiro de um país (NOI-*net outward investment*) está sistematicamente relacionada com o seu nível de desenvolvimento económico. Esta relação, formalizada no conceito de *investment development path* (IDP), propõe uma relação em U (ou em J) entre o nível de desenvolvimento económico de um país e a respectiva posição do investimento directo. O modelo IDP sustenta que as características estruturais de uma economia (a sua estrutura produtiva, as infraestruturas técnicas, o seu enquadramento institucional), se vão alterando com o desenvolvimento económico e têm uma relação sistemática com o NOI.

Diversos estudos descritivos e econométricos têm sido desenvolvidos para testar se a versão postulada pelo IDP pode ser confirmada empiricamente. O número total de estudos sobre IDP é mais de 20. Desses estudos, destacamos dois resultados fundamentais.

Em primeiro lugar, embora cada país possa evoluir de acordo com o postulado pelo IDP, este é idiosincrático por natureza. Portanto, cada país segue o seu próprio padrão, determinado pela estrutura dos seus recursos, a dimensão do seu mercado, a estratégia de desenvolvimento prosseguida e o papel do seu governo na organização da actividade económica (Dunning, 1981, 1988a; Narula, 1996; Dunning e Narula, 1996; Lall, 1996). De facto, os testes mais recentes tendem a analisar o IDP para um país particular em detrimento dos estudos seccionais por países, na medida em os últimos não conseguem captar as alterações estruturais inerentes ao processo de desenvolvimento Justifica-se, assim sendo, os estudos de caso.

Em segundo lugar, alguns autores têm sugerido que nem o NOI é o indicador apropriado para captar alterações estruturais na captação/prossecução de investimento directo estrangeiro nem o produto *per capita* evidencia as características estruturais de uma economia que se relacionam com a evolução do NOI. De facto, a evidência empírica mostra que o comportamento do investimento estrangeiro tende a diferir significativaamente para o mesmo nível de PIB per capita (Dunning e Narula, 1996). Como resposta à primeira limitação foi proposta uma representação adicional do IDP baseada na utilização separada dos stocks de entrada e de saída de IDE (Úbeda, 1999 e Durán e Ubeda, 1999). No segundo caso, remete-se para a necessidade de captar com maior rigor a especificidade das estruturas económicas, tendo sido propostas, para o efeito, diversas variáveis estruturais como a formação bruta de capital, o nível de escolaridade ou a capacidade tecnológica e científica do país (Dunning e Narula, 1996).

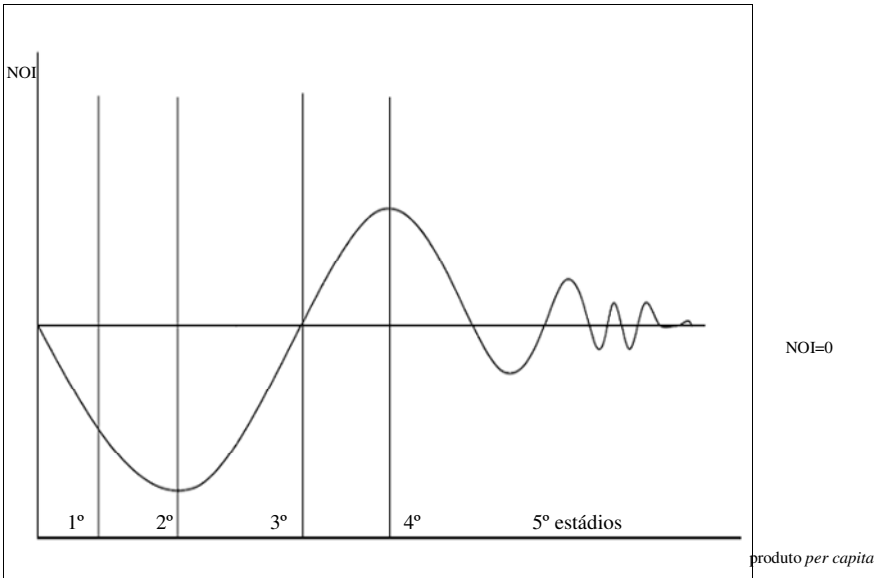
Neste trabalho seleccionámos os países da Coesão entre 1980-2005 (a partir de 1985 no caso da Irlanda e de 1986 no caso da Grécia, por limitações dos dados utilizados) para analisar as propriedades empíricas do IDP, tendo em consideração a versão original (doravante designada por "clássica") e as limitações desta acima referidas. Para este efeito recorremos, por um lado, à reformulação do modelo de IDP sugerida por Úbeda (1999) e Dúram e Úbeda (1999) e, por outro lado, utilizaremos indicadores de entrada e de saída de IDE "potencial", no sentido em que entram em consideração na sua construção, com algumas das especificidades das estruturas económicas acima mencionadas.

O trabalho está organizado nas seguintes secções. Na secção 2 descrevemos sucintamente o modelo do IDP e os problemas metodológicos que têm sido identificados na versão "clássica". Na secção 3, começamos por estimar o modelo do IDP para os quatro países da Coesão segundo as três formas não lineares mais utilizada na literatura. Seguidamente, discutimos os resultados obtidos tendo em consideração (i) a reformulação proposta por Úbeda (1999) e Dúran e Úbeda (1999) e (ii) a substituição dos valores observados para as entradas e saídas de IDE pelos que resultam da consideração das características estruturais da economia que os determinam, com base em indicadores construídos pela UNCTAD. Finalmente, a secção 4 apresenta as principais conclusões.

2. O MODELO DE IDP NA VERSÃO CLÁSSICA

O modelo de IDP identifica 5 estágios de um NOI de um país idealizado, relacionados com o seu nível de desenvolvimento económico (PIB p.c.) (fig. 1).

Figura 1. Estádios do ciclo de investimento directo



Fonte: Narula (1996)

O primeiro estádio está associado a uma fase de pré-industrialização. Sendo reduzido o nível de rendimento *per capita*, a procura interna é fraca e o mercado interno exíguo. Por outro lado, a economia apresenta uma série de problemas, tais como, a inadequação/inexistência de infra-estruturas, a baixa qualificação dos recursos humanos, um quadro económico e legal ineficiente e instabilidade política. Ou seja, nesta fase, são muito diminutas as vantagens de localização; as poucas que existem estão ligadas à posse de recursos naturais, sendo praticamente inexistentes as vantagens criadas. Portanto, o investimento directo do estrangeiro será, também ele, diminuto e quase exclusivamente virado para o aproveitamento de alguns recursos naturais. O investimento no exterior está praticamente ausente, já que as empresas domésticas não possuem vantagens de propriedade que lhes permitam a internacionalização: dedicam-se, maioritariamente, a produções trabalho-intensivas, tecnologicamente pouco avançadas e, geralmente, orientadas para o sector primário. A ausência de investimento no exterior assinalável, combinada com um tímido investimento do exterior, resulta num NOI negativo.

Este padrão de vantagens (ou da ausência delas) pode, porém, ser alterado, sobretudo pela actuação dos poderes públicos, que, neste estádio, toma normalmente duas formas. Compete-lhes melhorar as infra-estruturas, qualificar os recursos humanos e reduzir algumas das falhas de mercado endémicas que entravam o desenvolvimento. Mas é também normal a adopção de políticas de apoio às exportações e protecção contra as importações, que afectam a estrutura dos mercados.

No segundo estádio, o NOI torna-se mais negativo. Em termos de estrutura económica, o aumento do poder de compra e a decorrente expansão do mercado interno reforçam as vantagens de localização. Consequentemente, é viabilizada a produção por parte de empresas estrangeiras. Motivado pela imposição de barreiras às importações, a maioria do investimento estrangeiro destinar-se-á, neste estádio, a substituir aquelas, ou seja, será do tipo market seeking. Pode existir também alguma produção estrangeira do tipo resource seeking no sector exportador, ligada à exploração de recursos naturais ou a bens do sector primário em que seja possível alguma integração vertical a jusante com indústrias trabalho-intensivas, dependendo da capacidade de oferta de infra-estruturas de transporte e comunicação e de trabalho (não-qualificado). Existindo uma força de trabalho motivada e relativamente qualificada, poderá ocorrer também investimento do tipo *efficiency seeking*. Inexistente será o *strategic asset seeking investment*.

Relativamente ao primeiro estádio, aumentam as vantagens de propriedade das empresas domésticas. Surge, assim, o *outward investment*, mas ainda de forma incipiente. Segundo

Dunning e Narula (1996), a combinação de investimento directo estrangeiro com investimento interno induz o aparecimento de economias de aglomeração e a melhoria da produtividade. Mais produtivas, algumas empresas domésticas conseguem mesmo lançar-se no exterior. Fazem-no, normalmente, para adquirir tecnologia estrangeira ou em projectos ligados aos recursos naturais e ao sector primário. Procurando novos mercados, dirigir-se-ão, tendencialmente, para economias próximas, menos desenvolvidas e numa fase anterior do IDP; se o objectivo for a conquista de activos estratégicos que aumentem as vantagens já possuídas, os países de acolhimento serão, certamente, mais desenvolvidos e mais avançados no IDP.

Estas primeiras tentativas de internacionalização das empresas domésticas exigem políticas governamentais, mormente, de promoção das exportações e auxílio à inovação tecnológica. Mas também a captação de investimento estrangeiro deve ser estimulada, já que a fixação de empresas multinacionais contribui para a melhoria geral das infra-estruturas e da produtividade, assinaladamente, pela introdução de novos processos produtivos, transferência de tecnologia ou formação de recursos humanos.

O segundo estádio termina quando o NOI atinge o seu mínimo. O terceiro estádio apresentará, então, um NOI ainda negativo, mas de tendência crescente, resultado de um abrandamento e de uma aceleração nos ritmos de crescimento, respectivamente, do investimento do estrangeiro e no estrangeiro. Com o aumento do produto *per capita* característico de um processo de desenvolvimento, cria-se um mercado interno alargado e gera-se procura destinada a bens de qualidade superior. Outras vantagens de localização incluem os recursos humanos qualificados e a existência de potencial tecnológico, pelo que as economias de escala ganham importância. Estimuladas por tais vantagens e/ou pela presença de multinacionais e, eventualmente, auxiliadas pelos governos, as empresas domésticas aumentam a sua capacidade competitiva. O que, para as empresas estrangeiras, significa a redução das suas vantagens de propriedade iniciais. Mas outras surgem, baseadas em conhecimento e inovação ao nível tecnológico, da gestão e do *marketing*. Até porque o desenvolvimento económico acarreta geralmente uma subida dos salários, que diminui a vantagem comparativa dos sectores trabalho-intensivos. Assim, o *inward investment* será menos do tipo *market seeking* e mais *efficiency seeking* ou até mesmo *strategic asset seeking* nas indústrias em que as empresas domésticas detiverem vantagens. Quanto às empresas domésticas, é neste estádio que se inicia verdadeiramente o processo de internacionalização, possível graças ao incremento das vantagens de propriedade.

Logicamente, a captação de investimento estrangeiro deverá dar-se nos sectores em que as vantagens de localização do país são fortes, mas onde as empresas domésticas não detêm vantagens de propriedade. E aquelas devem expandir-se internacionalmente em indústrias sem vantagens de localização na economia domésticas e onde possam obter vantagens de propriedade. Ou seja, inicia-se a especialização em termos de investimento directo. O que pode levar à deslocalização de empresas, principalmente nos sectores trabalho-intensivos ou que funcionam em condições de grande competição, sendo o bem produzido homogéneo e tendo atingido a fase de maturidade (Wilson, 1978).

Apesar de as vantagens criadas se assumirem como mais importantes que as de origem natural, o papel do governo é, neste estádio, menos preponderante. Dele espera-se, essencialmente, que promova uma maior integração entre empresas nacionais e estrangeiras. E que encoraje o investimento de e no estrangeiro, mas levando em consideração os aspectos referidos.

Quando a taxa de crescimento do investimento no exterior ultrapassa a do investimento do exterior e o país se torna um exportador líquido de capital, o NOI passa a positivo e o país está no quarto estádio. Este envolve uma mudança profunda na estrutura produtiva dos países, a qual está relacionada com a procura de eficiência associada a custos de produção mais baixos, aspecto especialmente relevante nos sectores muito baseados no factor trabalho. Por outro lado, possuidoras de vantagens de propriedade, as empresas nacionais podem competir interna e externamente; para ultrapassarem barreiras comerciais, passam a investir em países estrangeiros, qualquer que seja o seu nível de desenvolvimento. Quanto às vantagens de localização, são formadas maioritariamente por activos criados, o que inclui recursos humanos qualificados,

sistemas de inovação e capacidade de I&D, economias de aglomeração e mercados sofisticados. Neste contexto, o investimento directo do estrangeiro tem origem predominantemente nos países desenvolvidos e é motivado pela procura de activos estratégicos.

Às autoridades cumpre continuar a sua função de supervisão e regulação, de modo a suprimir as falhas de mercado e garantir os níveis de concorrência entre empresas nacionais e estrangeiras. Mas deve dar, igualmente, atenção aos aspectos estruturais: fomentar as capacidades tecnológicas, em especial nas indústrias nascentes, que devem ser protegidas, e minimizar os impactos socio-económicos negativos da extinção de indústrias em declínio. Por outro lado, é de esperar uma atitude mais estratégica nas políticas adoptadas. Não se pretende uma intervenção directa, mas antes medidas que melhorem os recursos e as capacidades domésticas e que assegurem o bom funcionamento do mercado. Isto porque, neste estádio, existe uma acrescida competição entre países semelhantes.

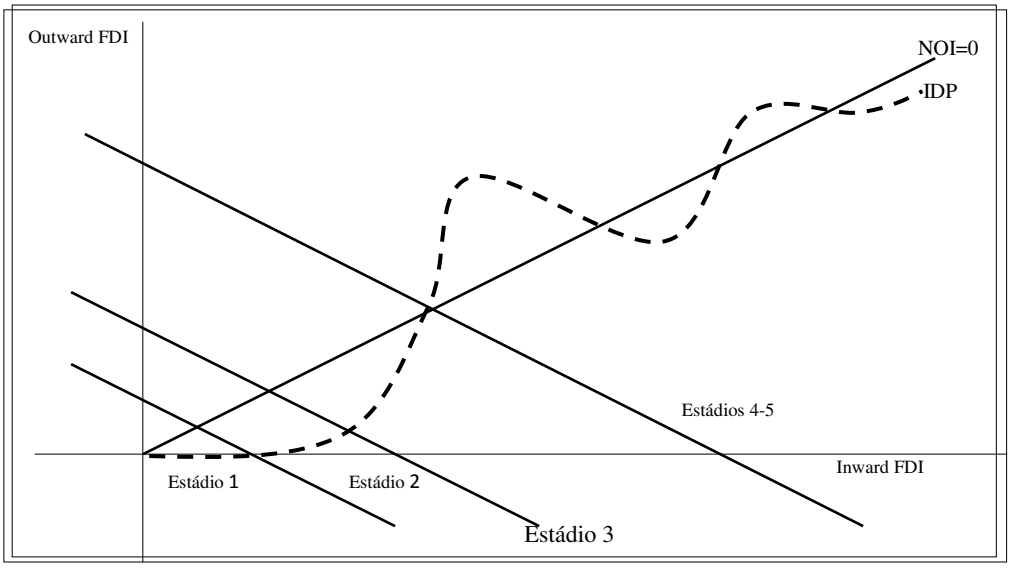
Dados o processo de convergência dos níveis de rendimento e estrutura económica entre as economias mais desenvolvidas e a harmonização das políticas governamentais, os padrões de investimento estrangeiro tornaram-se mais simétricos (Dunning, 1986b, 1988a, 1988b; Narula, 1993, 1996; Dunning e Narula, 1996). Assim, no quinto estádio, após um aumento inicial do investimento directo do estrangeiro, atinge-se um equilíbrio flutuante em que *outward investment* e *inward investment* continuam a crescer, mas assumem valores próximos um do outro, ou seja, o NOI adopta um comportamento errático, com grande volatilidade, mas em média perto do zero. Quanto ao investimento directo do exterior, será proveniente, por um lado, de países em estádios mais baixos que procuram aceder a mercados e à tecnologia e, por outro, de países dos quarto e quinto estádios, cujas empresas investem no estrangeiro por questões de eficiência.

No último estádio, o governo comportar-se-á como um oligopolista, que desenha e implementa as suas políticas estratégias tendo em consideração as decisões de outros governos. Dele se espera, também, um papel pró-activo na prossecução da eficiência de mercado e que coopere com as empresas na redução dos custos de ajustamento estrutural e de transacção. Além de poder actuar de variadas formas na promoção de novas vias de crescimento económico (Dunning e Narula, 1996).

2.1. Reconfigurações do Modelo de IDP

Durán e Úbeda (2001) afirmam que a evolução do NOI não é o indicador apropriado para inferir alterações estruturais na captação/prossecução de investimento directo estrangeiro. Por um lado, porque, assumindo a variável valores próximos de zero, poder-se-á estar perante a primeira fase ou, igualmente, na última. Mas, sobretudo, porque uma subida do *net outward investiment* pode dever-se, tanto a um aumento do investimento no exterior, como a uma situação de desinvestimento, em que o país perdeu competitividade. Este problema pode ser ultrapassado pelo recurso aos valores do investimento directo estrangeiro do e no país, em vez da utilização do saldo líquido, naquilo que configura uma nova representação do *investment development path* (figura 2).

Figura 2 - Nova representação do IDP proposta por Durán e Úbeda (2001)



Fonte: Durán e Ubeda (2001)

Na figura 2, os stocks de IDE aparecem separadamente, para além da posição líquida do stocks de IDE (NOI). Países acima da bissectriz têm um NOI positivo e os que ficam abaixo dessa linha têm um NOI negativo. Um movimento horizontal para a direita significa que se registou uma melhoria nas vantagens de localização e no sentido oposto o contrário. O primeiro caso é entendido com uma perda de competitividade e o segundo como um ganho de competitividade. Um movimento de subida do NOI na vertical identifica uma melhoria de comeptitividade, independentremente do movimento na horizontal.

Quanto à outra variável, o nível de produto per capita, tem sido referido que o produto, seja nacional ou interno, bruto ou líquido, total ou *per capita*, é um indicador com muitas limitações, na medida em que não reflecte as alterações estruturais que o desenvolvimento implica e que possam ter relação directa com as entradas e saídas de IDE. Alterações essas que importam na composição e distribuição geográfica do investimento do e no estrangeiro. Diferentes autores sugerem, por isso, a inclusão, em alternativa ao produto, de variáveis estruturais, de que são exemplo a formação bruta de capital *per capita*, os níveis de escolaridade, as taxas de iliteracia, a despesa em I&D ou o número de patentes (Dunning, 1981b e Dunning 1986a)

Outras críticas têm ainda sido apontadas ao modelo de IDP, abrindo pistas para um aprofundamento da relação que lhe é inerente. Scott-Kennel e Enderwick (2005) apontam o risco de, nos últimos estádios, se sobreestimar a capacidade das empresas domésticas em investir no exterior. Isto porque, ao recorrer aos valores agregados do *outward investment*, se ignora que parte dele pode ser realizado por empresas estrangeiras. Estes autores consideram, assim, ser importante compreender as ligações entre empresas nacionais e estrangeiras e de que modo esta interacção contribui para o incremento das vantagens das empresas locais e, em última análise, para o investimento no exterior. Segundo eles, a ideia central do IDP é intrinsecamente microeconómica; no entanto, a análise tem sido sempre conduzida a um nível macro ou mesoeconómico, nunca se explorando os mecanismos através dos quais o investimento do estrangeiro promove as vantagens de propriedade das empresas domésticas e, consequentemente, o investimento no estrangeiro. Para Scott-Kennel e Enderwick (2005), "these micro-level resource and capability transmission mechanisms constitute the *black box* of the IDP".

Dunning e Narula (1996) reconhecem que a relação entre investimento directo estrangeiro e desenvolvimento económico implica a comparação de dois fenómenos em diferentes níveis de análise económica: enquanto o investimento é, essencialmente, microeconómico e específico da empresa, o desenvolvimento constitui-se como um fenómeno macroeconómico e próprio de cada país. Portanto, estudar o investimento directo como algo associado a determinada localização pressupõe que as actividades das empresas domésticas e estrangeiras possam ser

agregadas. O que só é justificável para países onde a natureza, o modo e a motivação das empresas multinacionais são relativamente homogéneos e o valor acrescentado das suas actividades se mantém a níveis baixos.

Lall (1996) defende que o modo como tecnologia e competências são transferidas pode afectar, em termos quantitativos e qualitativos, a criação de activos no país de acolhimento, ou seja, tem impacto no seu potencial de desenvolvimento a longo prazo. Sugere, por isso, que, eventualmente, o IDP possa ser melhor analisado através da transferência internacional de activos intangíveis, para além do investimento directo. Por outro lado, visto que as consequências do investimento directo sobre a geração de vantagens de propriedade e de localização dependem das políticas adoptadas, estas deveriam ser endogeneizadas pela diferenciação entre diferentes abordagens.

Ainda a propósito da relação entre investimento e desenvolvimento, Moudatsou (2001) nota que elevadas taxas de crescimento e de captação de investimento estrangeiro tendem a ser realidades simultâneas. Mas tal facto nada diz sobre a relação de causalidade. Países em rápido processo de crescimento oferecem maiores oportunidades de lucro e têm uma maior procura de investimento directo estrangeiro. Por outro lado, este aumenta o stock de capital do país, podendo ter efeitos sobre o crescimento económico semelhantes ao do investimento doméstico. Além disso, tem um provável impacto positivo sobre as exportações e pode facilitar o acesso das empresas domésticas aos mercados internacionais. Refira-se, ainda, a melhoria da produtividade das empresas locais através das ligações a jusante e a montante com as filiais estrangeiras, o chamado efeito spillover decorrente da difusão da tecnologia transferida. Ou seja, se, por um lado, um maior desenvolvimento económico se constitui como factor de atracção do investimento estrangeiro, por outro, este é, muitas vezes, impulsionador daquele. A relação de causalidade dá-se, portanto, em ambos os sentidos, surgindo três casos distintos: países em que é o crescimento económico a explicação para a atracção do investimento; países em que foi a captação de capitais estrangeiros um dos propulsores do crescimento económico; e, finalmente, países em que crescimento e investimento se influenciam reciprocamente. Esta pertinente observação sugere rumos frutuosos para uma investigação empírica que procure entrar em consideração com a relação causal bidireccional entre investimento directo estrangeiro e desenvolvimento económico.

3. O IDP NOS PAÍSES DA COESÃO

3.1. Caracterização do NOI nos Países da Coesão

Não obstante Portugal ter participado desde sempre nos movimentos de circulação de capitais e a década de 60 ter conhecido uma maior abertura ao exterior, com o consequente aumento do investimento directo estrangeiro, este é um fenómeno dos anos oitenta. Efectivamente, só a partir de então, ganham os fluxos de investimento uma dimensão considerável. De facto, no período que decorre até 1960 ficou marcado por uma política de cariz marcadamente nacionalista. Se a isto se juntar a exiguidade do mercado, a inadequação das infra-estruturas e a baixa qualificação dos recursos humanos, facilmente se conclui que as vantagens de localização então existentes eram bastante limitadas.

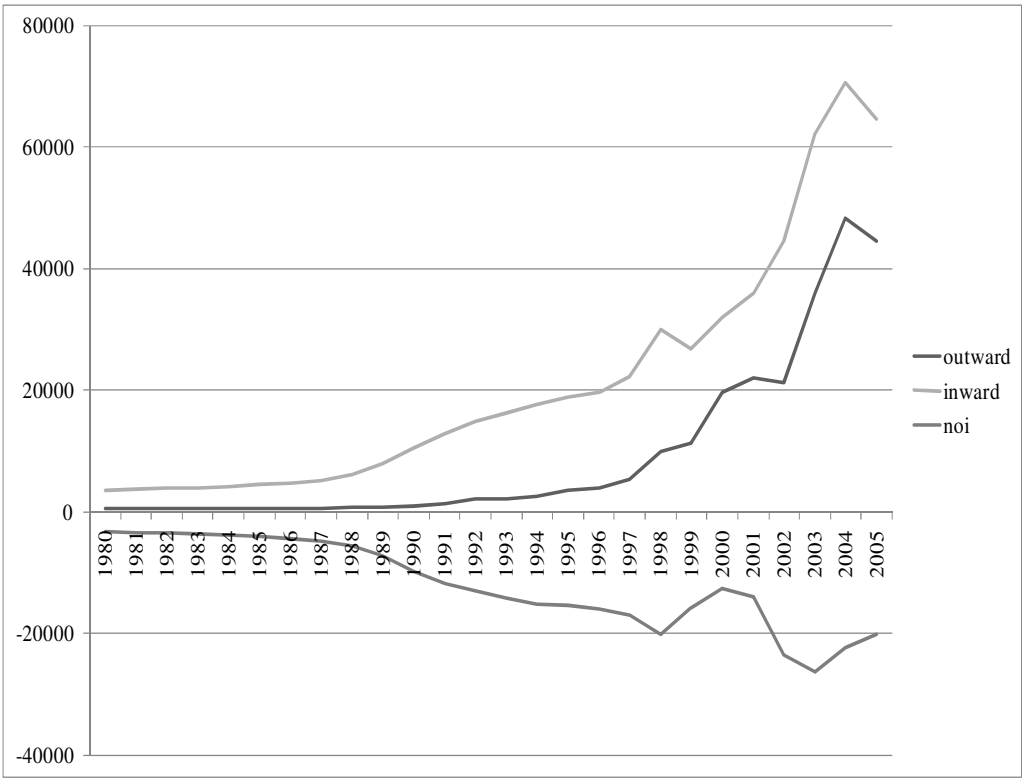
Com a adesão à EFTA, foram introduzidas alterações legislativas que conduziram a uma relativa liberalização dos movimentos de capitais. Isto é, reforçaram-se as vantagens de localização do país, que conheceu um aumento do investimento directo, direccionado, sobretudo, à indústria transformadora trabalho-intensiva. A entrada de capitais experimentou um período de crescimento exponencial até final dos anos oitenta.

Com a década de noventa, abranda o crescimento do investimento do exterior em Portugal; simultaneamente, as empresas portuguesas dão início ao seu processo de internacionalização. A importância deste último advém sobretudo dos interesses das empresas portuguesas no Brasil. Nesta última fase, a análise sectorial revela que os sectores que mais recebem investimento directo estrangeiro são os das actividades imobiliárias, alugueres e serviços prestados às empresas e o das actividades financeiras, e que estes sectores são igualmente aqueles onde as

empresas portuguesas mais investem no exterior. Este facto não surpreende, já que o investimento intra-indústria é uma das grandes tendências actuais ao nível dos movimentos internacionais de capitais. Por outro lado, insere-se no postulado pela terceira fase do IDP: o aumento da importância dos sectores associados a bens de qualidade superior, à inovação e a recursos humanos qualificados e a perda de relevância das actividades trabalho-intensivas.

A evolução do NOI per capita em Portugal está representada no gráfico 3 para o período 1980-2005.

Gráfico 3 – NOI per capita em Portugal, 1980-2005



Fonte: UNCTAD, FMI, cálculos próprios

O percurso espanhol é bastante semelhante ao do seu vizinho Portugal, sendo o seu ciclo de investimento determinado, em boa parte e, de resto, como na maioria dos países, pelo contexto politico-económico nacional e internacional. Os anos que se seguiram a 1891 ficaram marcados pelo regresso ao proteccionismo, pondo fim à tendência mais liberal de meados do século XIX, que havia trazido ao país alguns investimentos franceses, belgas e ingleses em sectores como os caminhos-de-ferro, a extracção mineira, os vinhos, a banca, os seguros ou os serviços públicos. Ainda assim, a década de 1920 conheceu, se bem que com altos e baixos motivados por crises de índole vária, a entrada de capitais americanos, alemães e franceses, associados à maquinaria eléctrica, químicos, automóveis e telecomunicações. À Grande Depressão sucedeu a Guerra Civil e ambos lesaram em muito a capacidade espanhola de atrair investimento. Algo que não se alterou com a instauração de um regime autoritário, de inspiração nacionalista.

Em 1959, com a implementação de algumas reformas económicas, o investimento directo estrangeiro voltaria a fazer parte da realidade espanhola. Por exemplo, a imposição de elevados impostos sobre a importação de bens industriais e de consumo num mercado doméstico com grande potencial de crescimento funcionou como factor de atracção no período 1959-1973. Já o investimento espanhol no exterior continuou a deparar-se com grandes entraves e manteve-se confinado à busca de matérias-primas, criação de canais de distribuição, projectos de construção e engenharia e banca. Com a crise económica mundial de 1973 e a mudança de regime político a partir de 1975, o investimento em Espanha sofre novo revés.

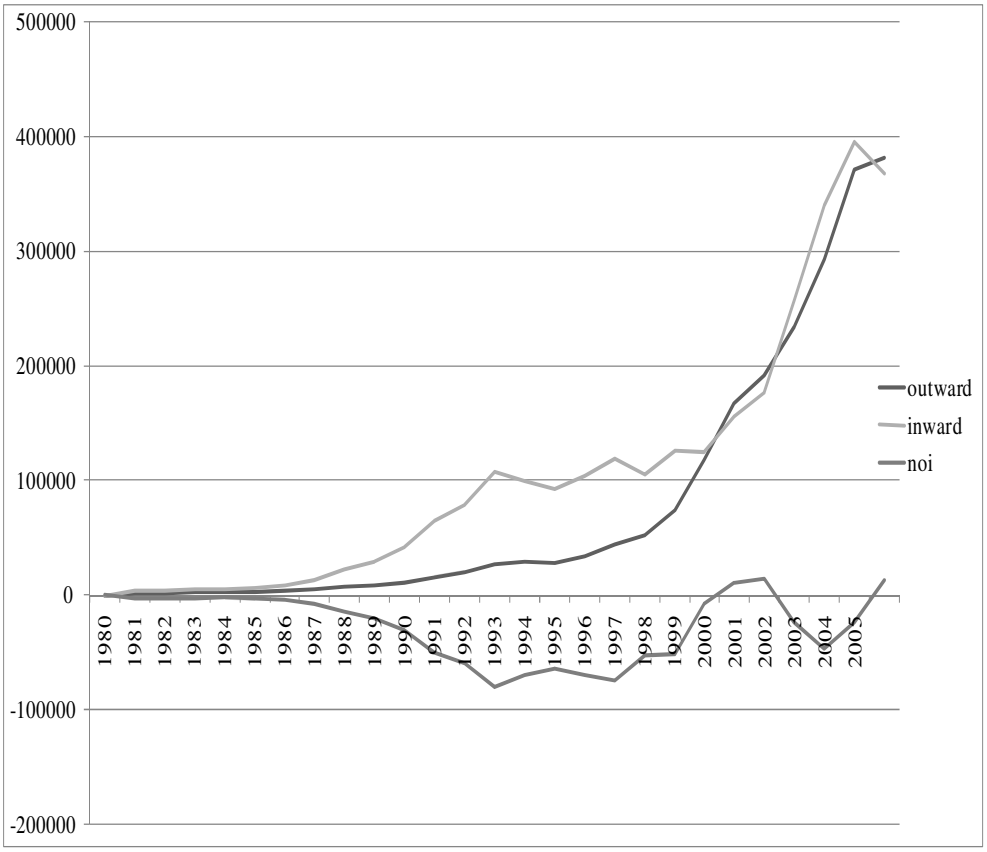
O final da década de 70 conhece alguma recuperação, mas é, sem dúvida, a adesão à CEE, em 1986, que constitui o grande marco impulsionador da internacionalização da economia espanhola. Aumentam as vantagens de localização, como a consolidação da democracia, a

liberalização dos fluxos de capital, o crescimento do mercado interno, a melhoria das infraestruturas ou a maior qualificação dos trabalhadores e, simultaneamente, o investimento passa a dirigir-se cada vez mais a sectores que utilizam intensivamente a tecnologia (Bajo-Rubio et Sosvilla-Rivero, 1994; Campa et Guillén, 1996; Álvarez, 2001; Guillén, 2005; Bajo-Rubio, Díaz-Mora e Díaz-Roldán, 2007).

Deste modo, o país passou por um intenso processo de mudança que incluiu a liberalização da economia, privatizações, maior competição e exposição a importações e ao investimento estrangeiro. Neste contexto, puderam as empresas espanholas incrementar as suas vantagens de propriedade, nomeadamente, um maior nível tecnológico, melhorias ao nível dos recursos humanos e sua organização ou, ainda, criação de imagens de marca. O que se traduziu, a partir dos anos noventa, e sobretudo desde 1992, quando se criou o Mercado Único, numa enorme vaga de internacionalização das empresas espanholas, que passaram a investir fortemente no exterior (Alvarez, 2001; Guillén, 2005).

O gráfico 4 ilustra a evolução do NOI per capita no caso de Espanha entre 1980-2005.

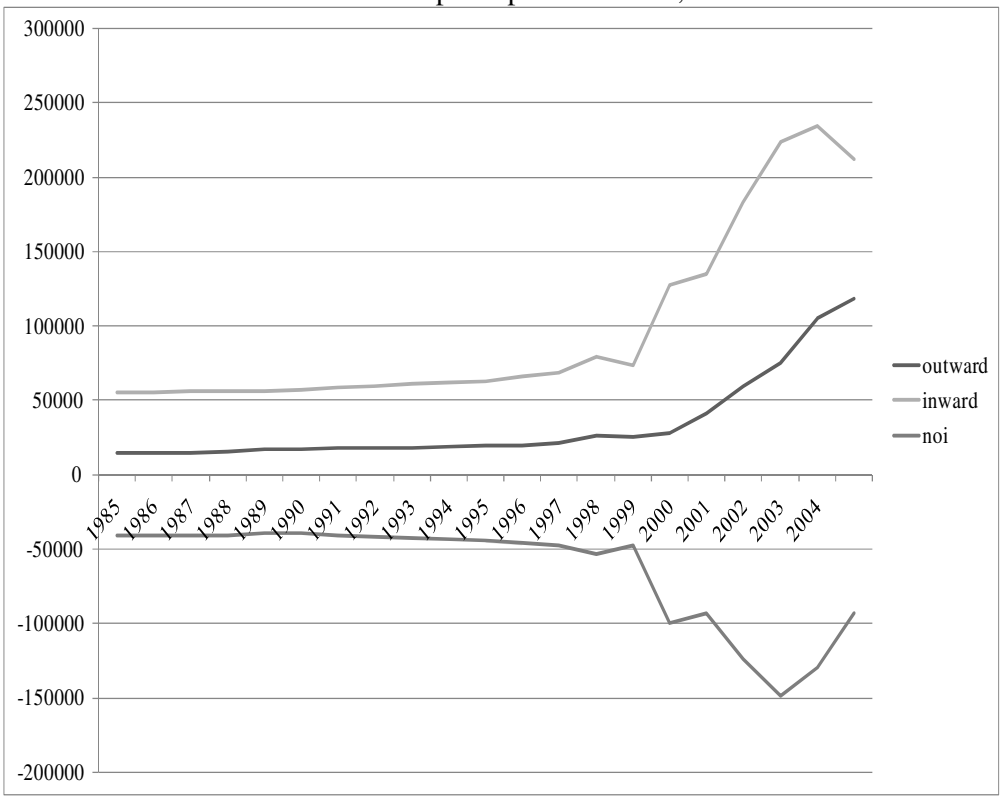
Gráfico 4 – NOI per capita em Espanha, 1980-2005



Fonte: UNCTAD, FMI, cálculos próprios

O caso irlandês reveste-se de especial interesse, dado o seu espectacular crescimento económico num período de tempo muito curto, com a adesão à UE. No final dos anos 90 a Irlanda registava o terceiro stock de entrada de IDE relativamente ao produto na UE, depois da Bélgica/Luxemburgo e da Holanda, enquanto as saídas representavam o terceiro stock mais baixo, depois da Grécia, Portugal e Áustria. Esta tendência começa a alterar-se no final dos anos 90, passando a Irlanda a investir IDE sobretudo nos EUA.

Gráfico 5 – NOI per capita na Irlanda, 1985-2005



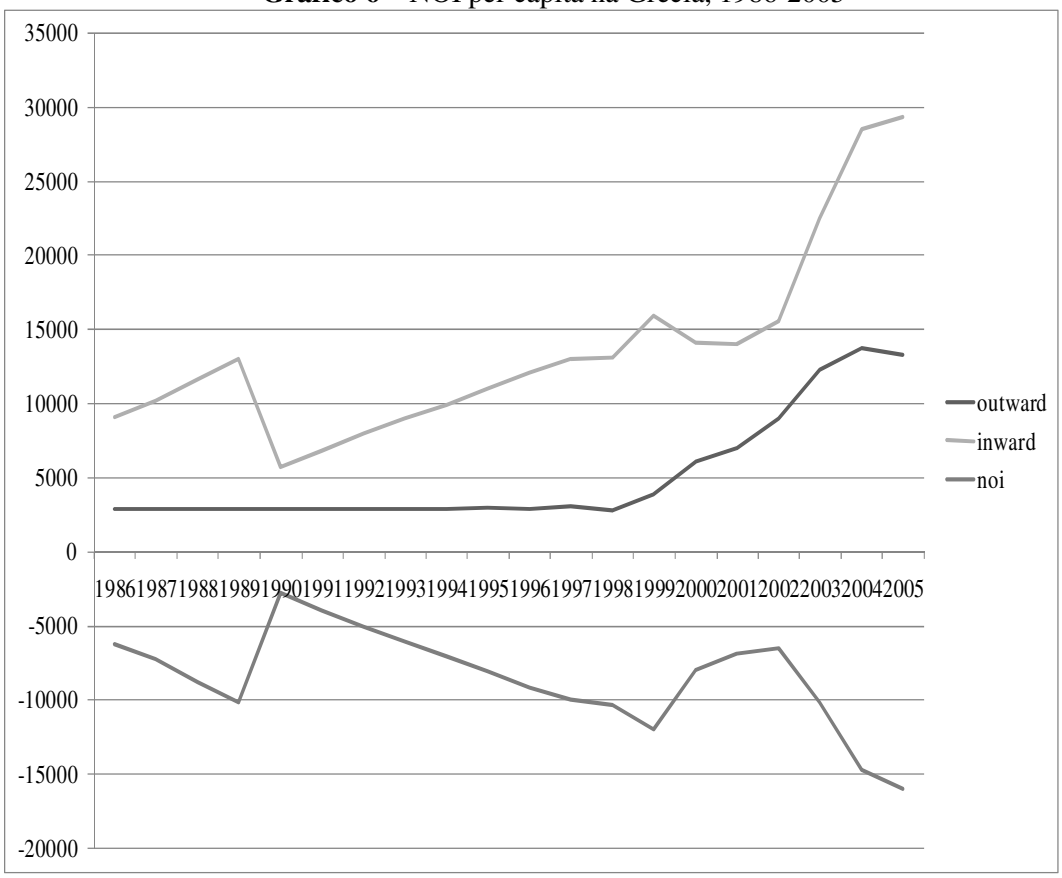
Fonte: UNCTAD, FMI, cálculos próprios

A abertura da economia grega ao investimento directo estrangeiro dá-se nos anos que se seguiram à II Guerra Mundial, num período que foi, para este país, de notável crescimento. Sectores como o dos químicos, da metalurgia ou dos transportes conseguiram captar capitais estrangeiros e constituíram-se como factor de rejuvenescimento e expansão da base industrial do país (Filippaios e Stoian, 2005). Esta tendência viria, contudo, a ser contrariada nas décadas de 70 e 80: demasiada regulação, défices orçamentais,que aceleraram a inflação e obrigaram a desvalorizações cambiais, e instabilidade macroeconómica, criaram um contexto adverso aos investimentos estrangeiros. Perante os desafios da integração europeia, o governo grego adoptou medidas no sentido de reverter este quadro. O programa de austeridade de 1985, combinado com um pacote de incentivos ao investimento lançado em 1988, alcançou relativo sucesso na redução da inflação, no equilíbrio da balança de pagamentos e na recuperação da confiança dos investidores, traduzindo-se numa entrada signficativa de capitais. Estes resultados foram contudo meramente temporários, pois a instabilidade política de 1989 e 1990 trouxe novamente problemas às finanças públicas e ao crescimento económico, com as entradas de investimento directo estrangeiro a estagnarem. Um novo programa de ajustamento foi seguido a partir de 1991. Combinava reformas estruturais, tais como, privatizações, com o estabelecimento de objectivos audaciosos para as políticas monetária e fiscal. Tais medidas não deixaram de contribuir para a melhoria das condições económicas, mas a recuperação fez-se sentir lentamente (OECD, 1994). Assim, a Grécia manteve-se como um receptor líquido, não muito bem sucedido, de investimento directo estrangeiro (Demos, Filippaios e Papanastissou, 2004; Filippaios e Stoian, 2005).

Desde a década de 90, porém, que o investimento directo grego no estrangeiro tem crescido consideravelmente. A queda do Muro de Berlim e a transição das economias soviéticas para um regime de mercado trouxe às empresas gregas a oportunidade de se expandirem e internacionalizarem. Por outro lado, as filiais de grandes multinacionais instaladas na Grécia transformaram-se em centros de decisão regionais, com o intuito de penetrarem nos mercados dos Balcãs e dos países da Europa Central e Oriental (Demos, Filippaios e Papanastassiou, 2004; Filippaios e Stoian, 2005).

O gráfico 6 ilustra o NOI per capita para a Grécia entre 1986 e 2005.

Gráfico 6 – NOI per capita na Grécia, 1986-2005



Fonte: UNCTAD, FMI, cálculos próprios

3.2. Teste do IDP para os Países da Coesão

Diversas funções têm sido propostas para estimar o IDP. Algumas utilizam o PIB (por ex. Buckley e Castro, 1998), outras o PNP (por ex. Tolentino, 1987,1993).

Tolentino (1987, 1993) propôs a seguinte relação quadrática para estimar o IDP:

$$NOI_{PC\,i} = \alpha + \beta_1 GNP_{PC\,i} + \beta_2 GNP^2_{PC\,i} + \varepsilon \text{ (equação 1)}$$

em que as variáveis surgem normalizadas pela respectiva população.

Narula (1993, 1996) propõe uma especificação ligeiramente diferente, na medida em que elimina o termo constante:

$$NOI_{PC\,i} = \beta_1 GNP_{PC\,i} + \beta_2 GNP^2_{PC\,i} + \varepsilon \text{ (equação 2)}$$

Buckley e Castro (1998) argumentam que a forma quadrática pressupõe que o investimento é o motor do desenvolvimento. Ora, segundo eles, espera-se que, no primeiro estádio, a taxa de crescimento do produto seja superior à do investimento líquido, uma tendência que só a partir do segundo estádio se inverterá, o que pressupõe uma equação polinomial:

$$NOI_{PC} = \alpha + \beta_1 GDP^3_{PC} + \beta_2 GDP^5_{PC} + \varepsilon \text{ (equação 3)}$$

Para que a relação entre NOI e GDP descreva a já referida curva em U, os coeficientes β_1 e β_2 têm de ser, respectivamente, negativo e positivo (e, obviamente, estatisticamente significativos). Já o α espera-se que não seja significativamente diferente de zero, de modo a que, a níveis muito baixos de rendimento, não exista nem outward nem inward investment, de onde resulta um NOI igual a zero.

Neste estudo, as três equações são estimadas para Portugal e os outros três países da coesão (Espanha, Irlanda e Grécia). O NOI é calculado pela diferença entre os stocks de outward investment e de inward investment[2]. São várias as organizações que publicam informação sobre estas duas variáveis. Atendendo à necessidade de estabelecer comparações, foram usados os dados da UNCTAD. A informação relativa ao PIB foi obtida no FMI. Os valores são nominais e estão expressos em dólares americanos[3]. Também no FMI foram obtidos os dados para a população total que permitiram calcular os agregados *per capita*.

Os resultados das estimações (através de OLS) estão sumariados na tabela 1.

Tabela 1. Resultados da estimação do IDP para Portugal, Espanha, Irlanda and Grécia

	Equação 1: $NOI_i = \alpha + \beta_1 GNP_{PC\,i} + \beta_2 GNP^2_{PC\,i} + \varepsilon$					
	α	β_1	β_2	R^2	$\overline{R}^2$	F
Portugal	252.8460 (1.518479)	-0.204945* (-4.945131)	3.61E-06 (1.612756)	0.896709	0.887727	99.83592
Espanha	1223.293 (2.385369)	-0.304869* (-3.810071)	9.87E-06* (3.479684)	0.396789	0.344335	7.564625
Irlanda	-3532.003 (-0.746614)	-0.695339 (-1.671062)	2.45E-06 (0.322039)	0.660987	0.623319	17.54763
Grécia	-964.2847 (-2.769527)	0.060390 (1.285351)	-3.11E-06** (-2.134742)	0.550963	0.498135	10.42939
	Equação 2: $NOI_{PC\,i} = \beta_1 GNP_{PC\,i} + \beta_2 GNP^2_{PC\,i} + \varepsilon$					
	β_1	β_2		R^2	$\overline{R}^2$	
Portugal	-0.147506* (-8.483304)	8.49E-07 (0.633570)		0.886354	0.881619	
Espanha	-0.126992* (-4.002833)	4.26E-06** (2.458401)		0.247560	0.216208	
Irlanda	-0.989535* (-7.488817)	7.39E-06*** (2.010606)		0.650488	0.632093	
Grécia	-0.066628* (-5.579898)	6.11E-07 (0.924047)		0.348360	0.312158	
	Equação 3: $NOI_{PC} = \alpha + \beta_1 GDP^3_{PC} + \beta_2 GDP^5_{PC} + \varepsilon$					
	α	β_1	β_2	R^2	$\overline{R}^2$	F
Portugal	-365.8043* (-6.011849)	-1.33E-09* (-13.26855)	3.35E-18* (9.845433)	0.929425	0.923288	151.4466
Spain	-163.7201 (-0.757735)	-4.32E-10* (-3.587644)	6.81E-19* (3.704240)	0.373677	0.319214	6.861125
Ireland	-8748.920* (-7.921585)	-9.61E-10* (-7.925735)	3.49E-19* (6.510523)	0.853405	0.837116	52.39356
Greece	-643.8588* (-6.404985)	-3.04E-11 (-0.515015)	-2.81E-20 (-0.312222)	0.539203	0.484992	9.946313

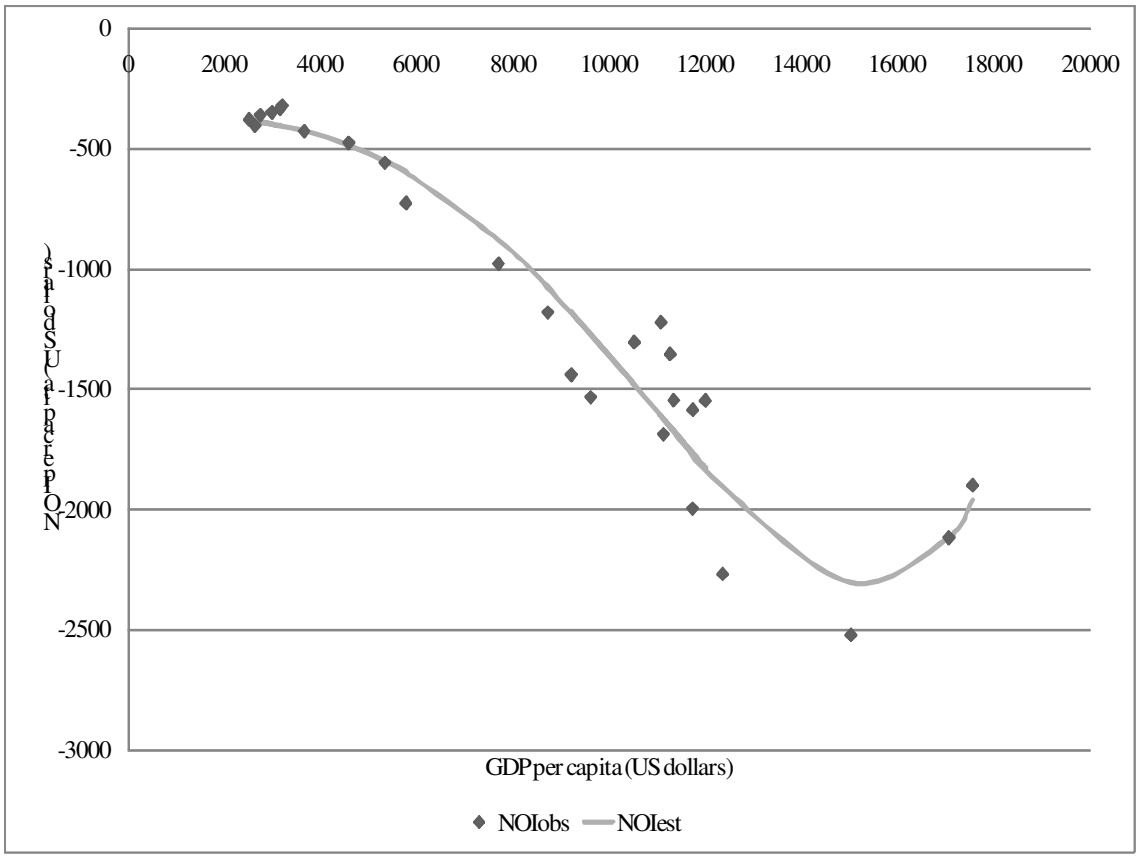
Nota: t-statistics em parêntesis; *significativo a 1%, ** significativo a 5%, * significativo a 10%**

[2] Narula (1993, 1996) critica o recurso a dados respeitantes a fluxos de investmento.
[3] As estimações foram tmabém feitas em termos reais mas não se detectaram diferenças assinaláveis nos resultados.

Os resultados dão suporte ao modelo de IDP no caso de três dos quatro países estudados: Portugal, Irlanda e Espanha, os dois primeiros encontrando-se na 3ª fase o último na 4 ª fase. Estes resultados estão em consonância com outros estudos, designadamente Buckley and Castro (1998), Castro (2000) and Castro and Buckley (2001) para Portugal, Durán e Ubeda (2001) e Álvarez (2001) para Espanha e Gorg (2000) e Barry et al (2003) para a relação entre a Irlanda e os EUA, principal parceiro da Irlanda no que se refere ao IDE. Quanto à Grécia as estimações realizadas não comprovam a relação esperada, na linha dos resultados encontrados para a Áustria e China, respectivamente por Bellak (2001) e McCarthy (2007).

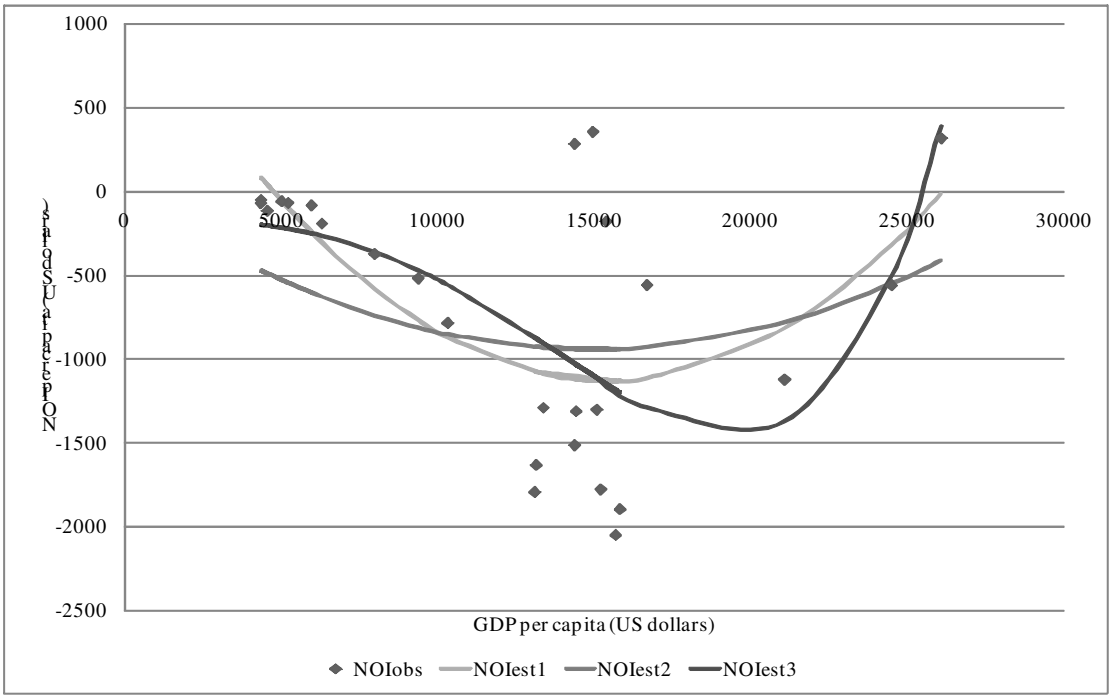
Os valores observados e estimados para Portugal, Espanha e Irlanda podem ser observados nos gráficos 7 a 9, reflectindo o posicionamento de cada um destes países na etapa indicada.

Gráfico 7. Estimação do IDP para Portugal, 1980-2005



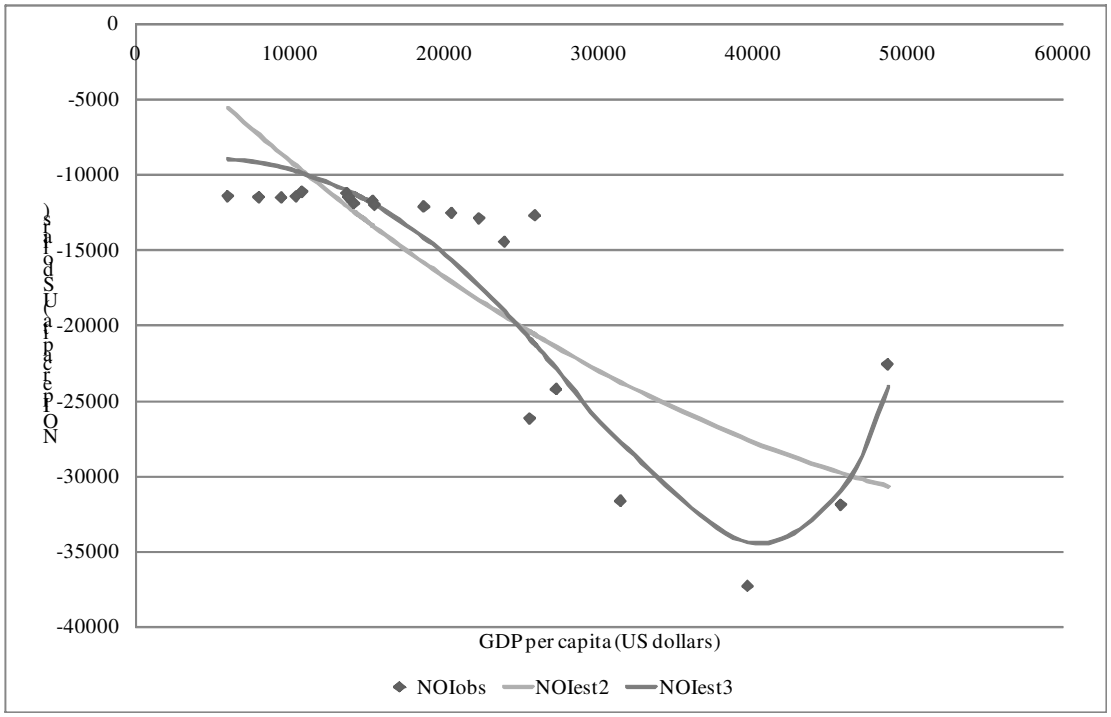
Source: own calculations

Gráfico 8: Estimação do IDP para Espanha, 1980-2005



Source: own calculations

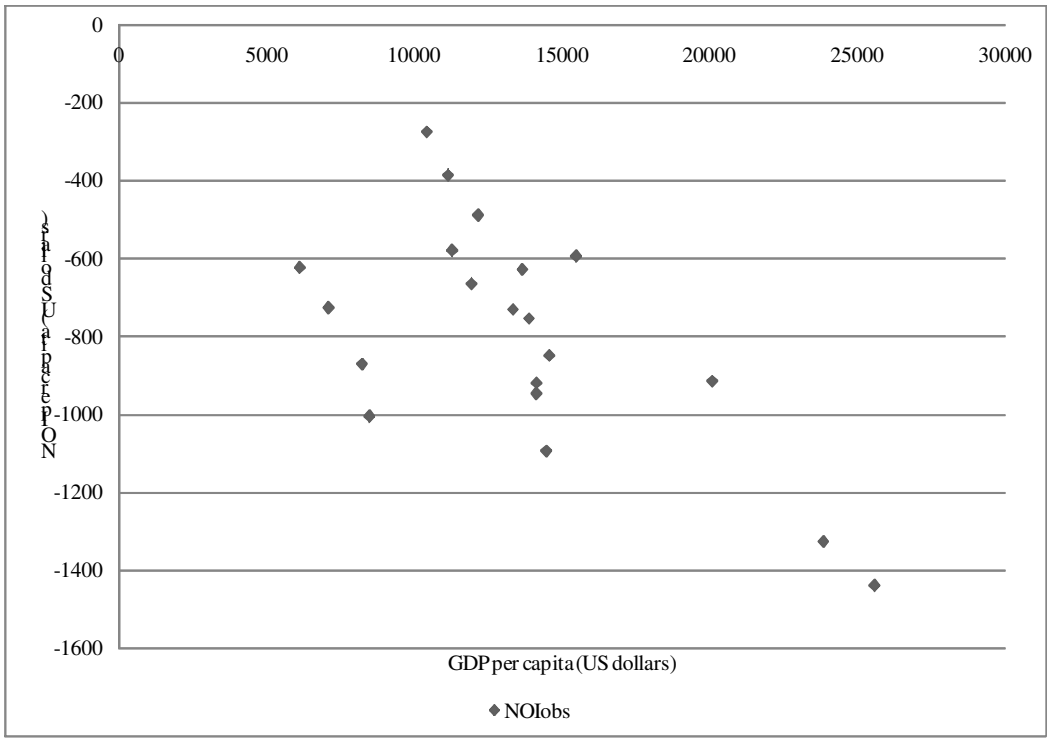
Gráfico 9: Estimação do IDP para a Irlanda, 1985-2005



Source: own calculations

No caso da Grécia, não se detecta a esperada curva em U, conforme se observa no Gráfico 6 e é confirmado pelas estimação efectuadas.

Gráfico 10: NOI e PIB p.c. para a Grécia, 1986-2005

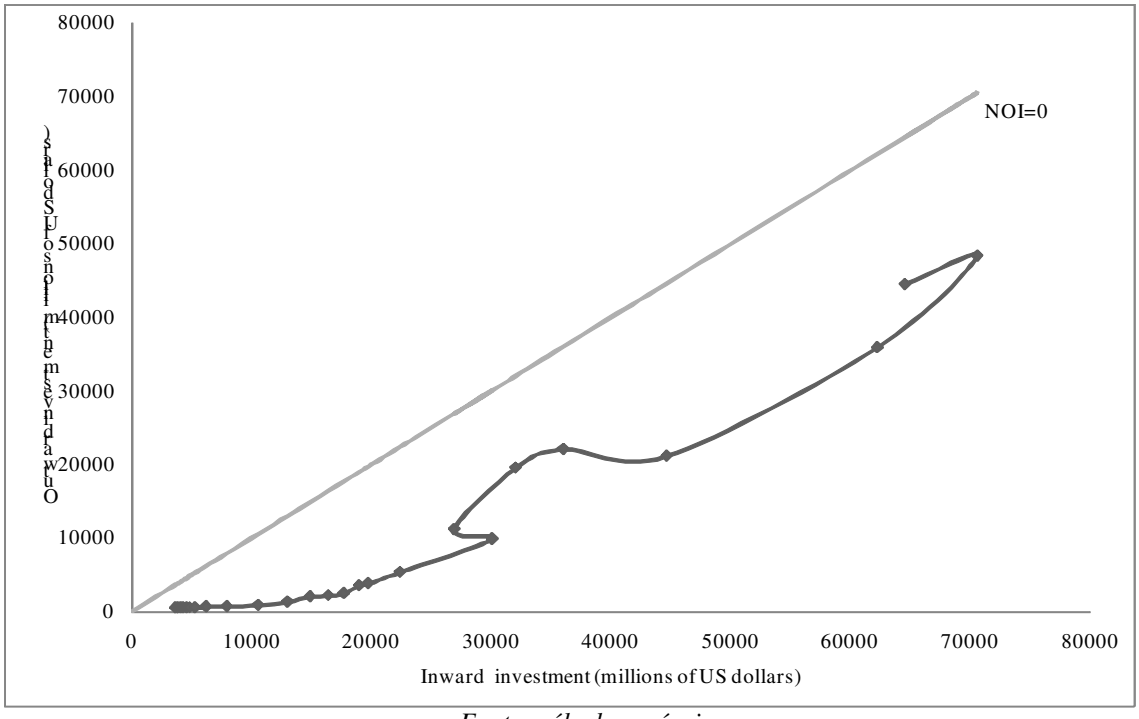


Fonte: UNCTAD, FMI, cálculos próprios

3. 3. Representação alternativa do IDP

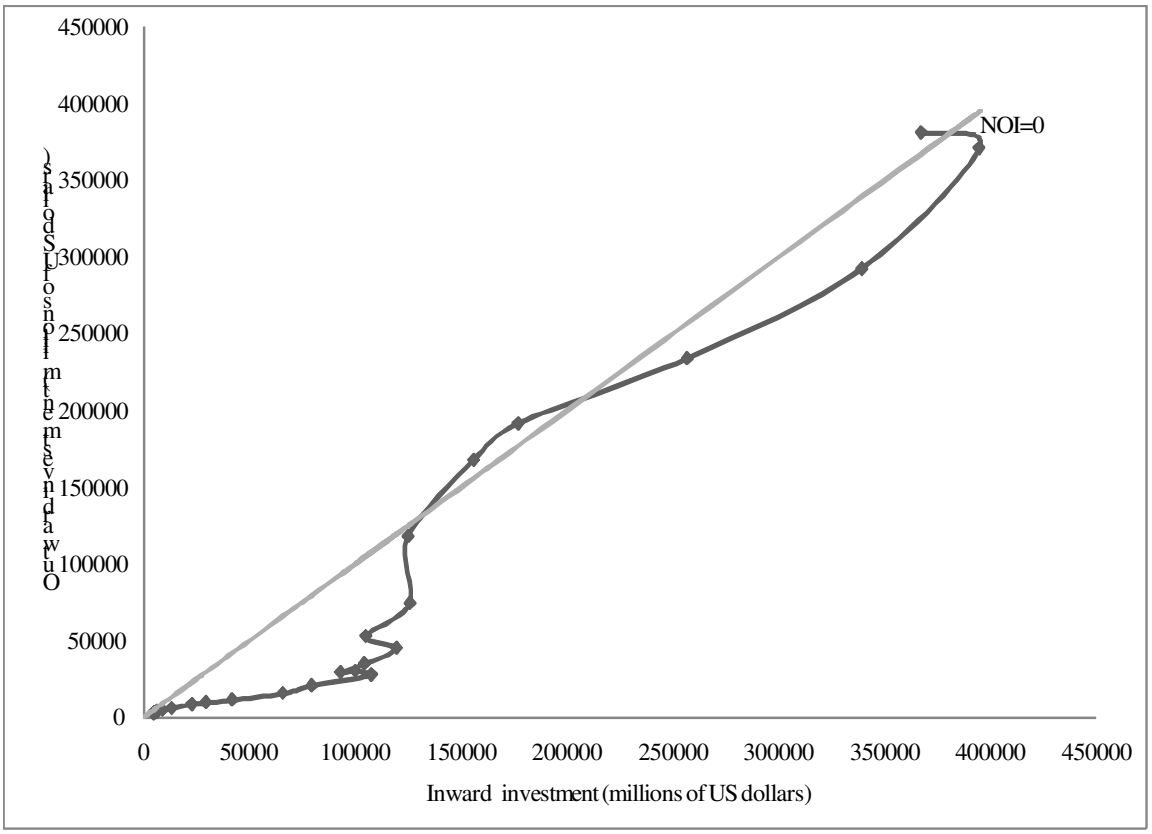
Conforme referido na secção 2, Durán and Ubeda (2001), procederam a uma representação alaterntiva do IDP, desagregando o NOI nas suas duas componentes. Os resultados da aplicação da sua metodologia aos 4 paises estudados podem ser obeservados nos gráficos 11 a 14.

Gráfico 11. Uma representação alternativa do IDP para Portugal



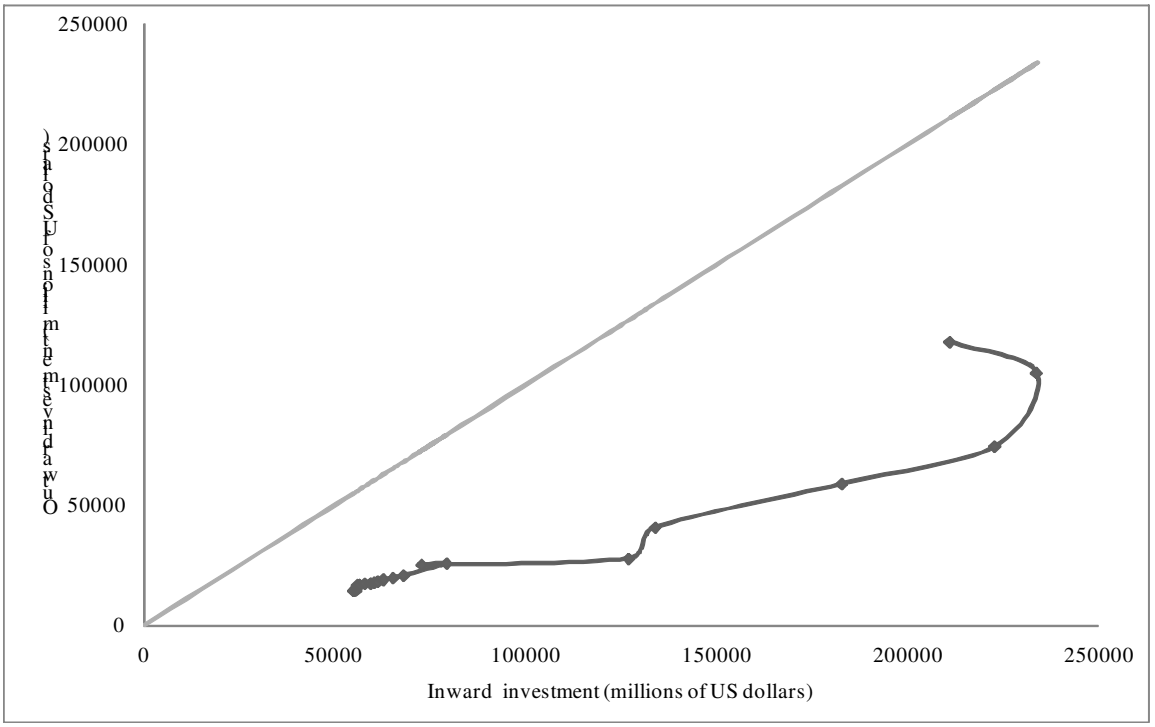
Fonte: cálculos próprios

Gráfico 12. Uma representação alternativa do IDP para Espanha



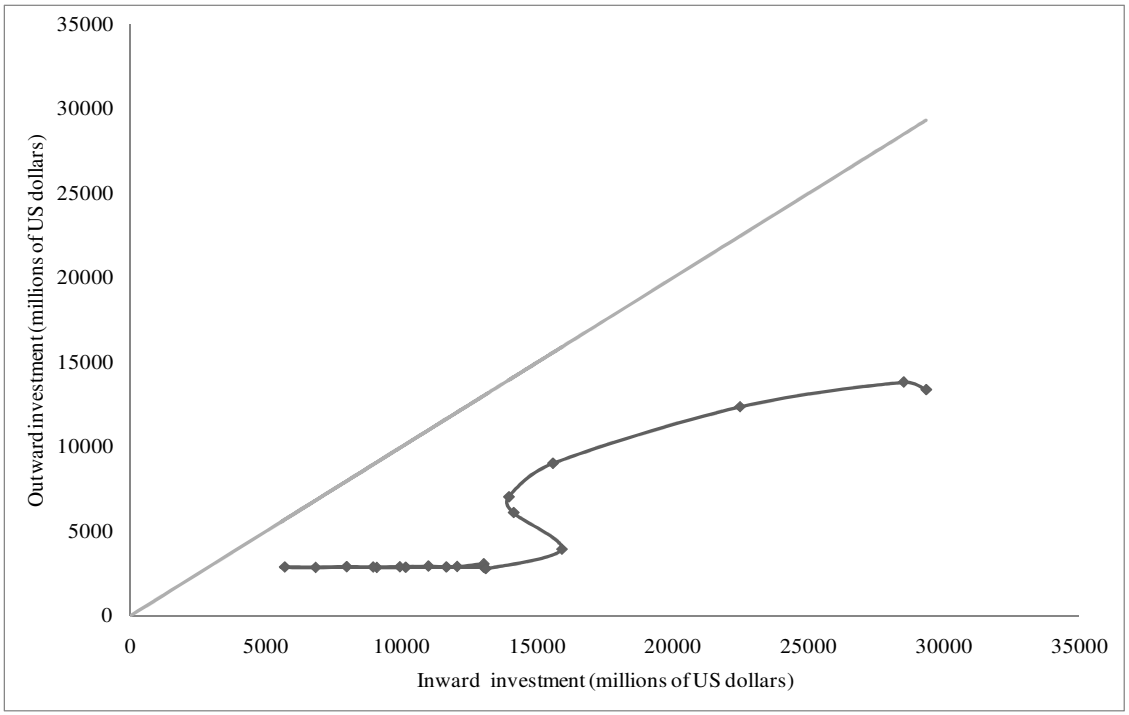
Fonte: cálculos próprios

Gráfico 13. Uma representação alternativa do IDP para a Irlanda



Fonte: cálculos próprios

Gráfico 14. Uma representação alternativa do IDP para a Grécia

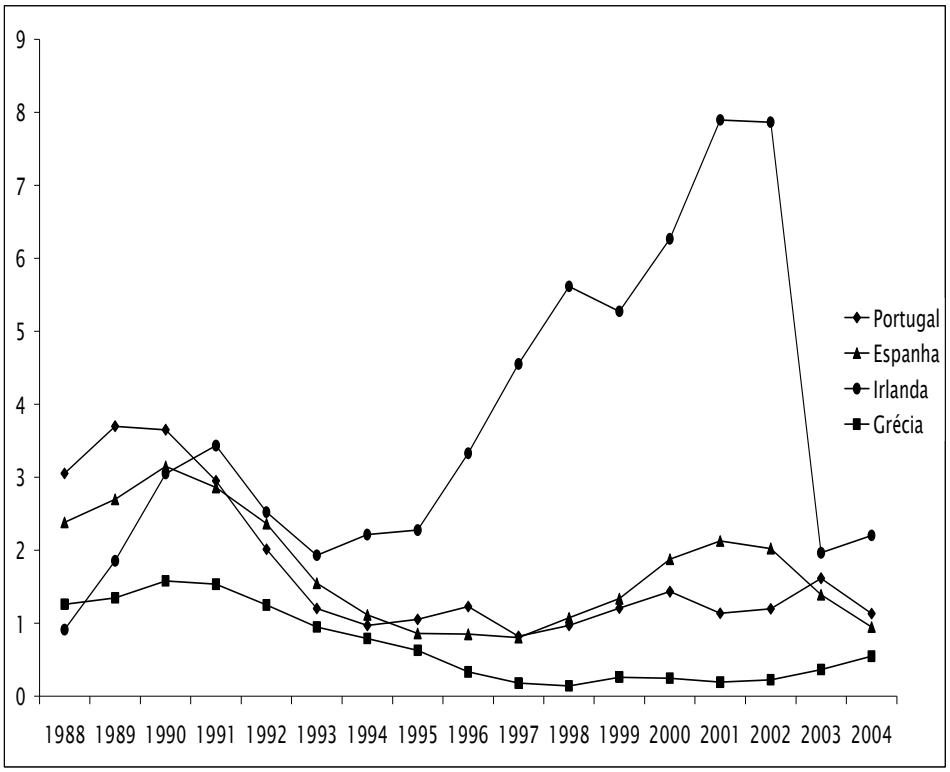


Fonte: cálculos próprios

Os gráficos anteriores permitem confirmar o posicionamento de Portugal no terceiro estádio do IDP, enquanto assinalam o posicionamento de Espanha e a Irlanda na etapa seguinte. Quanto à Grécia a evidência não é clara, na linha do acima detectado.

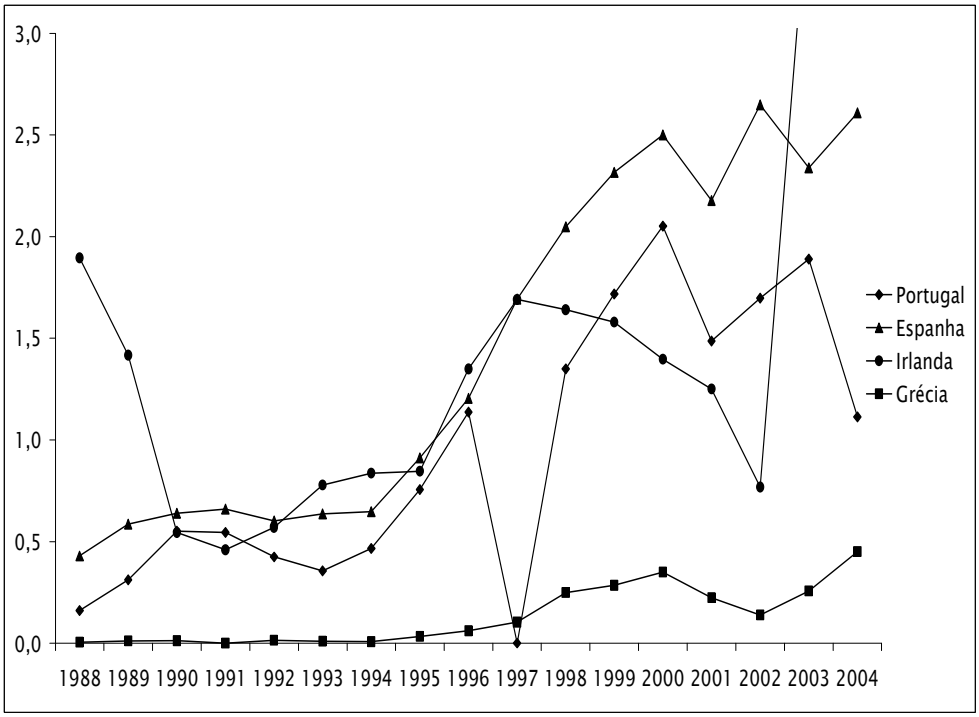
Conforme referimos, uma fragilidade do IDP consiste na utilização do produto *per capita* como indicador das características estruturais das economias que explicam os stocks de IDE. Uma alternativa seria substituir o PIB per capita pelo Indicador de Desenvolvimento Humano (IDH), na medida em que atende a algumas características estruturais (formação bruta de capital *per capita*, níveis de escolaridade, taxas de iliteracia, despesa em I&D, número de patentes), com destaque para o nível educacional. Contudo não dispomos de séries contínuas para este indicador. Outra possibilidade é substituir os stocks de IDE por um valor potencial que incorpore essas características. Para este objectivo dispomos dos indicadores construídos pela UNCTAD que procuram precisamente incoporar algumas dessas características estruturais na captação e saída de (fluxos) IDE: o *Inward FDI Performance Index* e o *Outward FDI Performance Index. A* evolução destes indicadores para cada um dos países estudados consta dos Gráficos 15 e 16.

Gráfico 15 - Inward FDI Performance Index



Fonte: UNCTAD

Gráfico 16 - Outward FDI Performance Index



Fonte: UNCTAD

A comparação dos dois gráficos mostra uma consonância com os resultados obtidos acima nos casos de Portugal e Espanha. Já no caso da Irlanda, identificam uma alteração brusca entre 2002 e 2003, que assinalam a entrada deste país na 4ª fase, em consonância com a versão

reformulada do IDP. No caso da Irlanda estes resultado é particularmente interessante na medida em que alguns autores optaram por excluir a Irlanda por considerá-la um *outlier*, dado identificarem uma 3ª fase do IDP num país com elevados níveis de produto per capita (ver, por ex., Fonseca et al, 2007)

Quanto à Grécia, mostram a possibilidade deste país seguir uma versão "moderada" do IDP, com uma curva em U muito suave, que não é captada pelas expressões analíticas usadas nas estimações anteriores. A razão principal parece ser o facto de a Grécia registar entradas anormais de IDE no início do período que não correspondem às características estruturais desta economia. A utilização dos indicadores "potenciais" corrige estes valores, produzindo uma informação mais adequada sobre o IDP.

4. CONCLUSÕES

O estudo efectuado estima a versão "clássica" do IDP para os países da Coesão e discute os resultados à luz das limitações deste modelo no que diz respeito, por um lado, à utilização conjunta das entradas e saídas de IDE (NOI) e, por outro lado, à ausência de uma adequada captação das características estruturais das economias determinantes da evolução das componentes do NOI.

A última via parece ser particularmente promissora. De facto, apesar do carácter preliminar desta analise, ficou evidente que as principais incongruências detectadas nos países estudados no contexto da versão "clássica"do IDP – o facto de a Irlanda se encontrar na 3ª fase, a par de Portugal, apesar de ter um produto *per capita* muito mais elevado, e o de não se conseguir identificar o IDP no caso da Grécia- poderão ser explicadas e ultrapassadas com a utilização de indicadores que contemplem essas características estruturais. Quanto à Espanha, todos os resultados apontam para o posicionamento deste país na 4ª fase do IDP.

REFERÊNCIAS

ÁLVAREZ, MONSERRAT (2001). "España y la Senda de Desarrollo de la Inversión Directa: Una Aproximación", *Institut d'Economia de Barcelona document de treball* 2001/11.

BARRY, Frank; GÖRG, Holger and MCDOWELL, ANDREW (2003). "Outward FDI and the Investment Development Path of a Late-Industrialising Economy: Evidence from Ireland", *Regional Studies*, 37(4), 341-349.

CASTRO, FRANCISCO BARROS (2000). *Foreign Direct Investment in the European Periphery: The competitiveness of Portugal*, PhD. Thesis, Leeds University.

CASTRO, FRANCISCO BARROS and BUCKLEY, PETER (2001). "Foreign Direct Investment and the Competitiveness of Portugal", paper presented at the *IV Conference on Portuguese Economy*, Lisbon, 4-5 May.

DUNNING, JOHN H. (1981). "Explaining the International Direct Investment Position of Countries: Towards a Dynamic or Developmental Approach", *Weltwirtschaftliches Archiv*, 117(1), pp. 30-64.

DUNNING, JOHN H. (1986a). "The Investment Development Cycle Revisited", *Welwirtschaftliches Archiv*, 122(4), pp. 667-676.

DUNNING, JOHN H. (1986b). "Investment Development Cycle and Third World Multinationals", *in* Kushi Khan (ed.), *Multinationals of the South: New Actors in the International Economy*, New York: St. Martin's Press.

DUNNING, JOHN H. (1988a). *Explaining International Production*, London, Unwin Hyman.

DUNNING, JOHN H. (1988b). "The Eclectic Paradigm of International Production: A Restatement and Some Possible Extensions", *Journal of International Business Studies*, 19(1), pp. 1-31.

DUNNING, JOHN H. and NARULA, RAJNEESH (1996). "The Investment Development Path Revisited: Some Emerging Issues", *in* John Dunning et Rajneesh Narula (eds.), *Foreign Direct Investment and Governments: Catalysts for Economic Restructuring*, London and New York: Routledge.

DURÁN, JUAN J. and ÚBEDA, FERNANDO (2001). "The Investment Development Path: A New Empirical Approach and Some Theoretical Issues", *Transnational Corporations*, 10(2), pp. 1-34.

FILIPPAIOS, FRAGISKOS and STOIAN, CARMEN (2005). "Where to, Odysseus? The Quest of Greek Firms to Expand Abroad", Kent Business School working paper 104.

GÖRG, HOLGER (2000). "Irish Direct Investment in the US: Evidence and Further Issues", *Journal of the Statistical and Social Inquiry Society of Ireland*, 30, pp. 33-64.

LALL (1996). "The Investment Development Path: Some Conclusions", *in* John Dunning et Rajneesh Narula (eds.), *Foreign Direct Investment and Governments: Catalysts for Economic Restructuring*, London and New York: Routledge.

NARULA, RAJNEESH (1993). *An Examination of the Evolution and Interdependence of Foreign Direct Investment and Economic Structure: The Case of Industrialized Countries*, PhD. thesis, Rutgers University.

NARULA, RAJNEESH (1996). *Multinational Investment and Economic Structure*, London: Routledge.

WILSON (1978). "Foreign Disinvestments: Friend or Foe", University of Virginia, Graduate School of Business Administration, working paper, DSWP-78-08.

LA REFORMA DEL SISTEMA BANCARIO CHINO: ESPECIAL ATENCIÓN AL PROCESO DE REGULACIÓN Y SUPERVISIÓN

ANA I. SALVADOR CHAMORRO
E-mail: ana.salvador@unileon.es
CRISTINA GUTIÉRREZ LÓPEZ
E-mail: cristina.gutierrez.lopez@unileon.es
Facultad de CC. Económicas y Empresariales
Universidad de León

Resumen

En 1978 China inició un proceso de reformas que ha transformado por completo su economía, a través de la introducción gradual de los elementos propios de una economía de mercado. Desde mediados de los noventa y, sobre todo, tras su adhesión a la OMC, uno de los grandes objetivos de este proceso es reformar el sistema bancario. En concreto, el reto consiste en pasar de una estructura al servicio de una economía planificada a un sistema propio de una economía de mercado. Los resultados de este proceso (clave para el futuro de la economía china) dependerán crucialmente del establecimiento de un adecuado esquema regulador. En este contexto, este trabajo revisa los aspectos más relevantes de esa reforma bancaria china, poniendo especial interés en la supervisión y regulación de la solvencia y su relación con el Acuerdo de Capitales conocido como Basilea II.

Palabras Clave: reforma económica china, sistema bancario, regulación y supervisión, Basilea II.
Área Temática: Economía Internacional y de la Integración.

Abstract

In 1978 China began a reform process that has completely transformed its economy, through the gradual introduction of the main elements in a market economy. From the middle of nineties and, specially, after its adhesion to the WTO, one of the main aims in this process has been the reform of the banking system. Thus, the challenge is to change from a structure of a planned economy to a market economy. The results of that process (a key factor in the future of the Chinese economy) will depend on a suitable regulatory scheme . In this context, this paper reviews the most significant aspects in this Chinese banking reform, with special interest in solvency´s supervision, in relation to the Capital Accord best known as Basel II.

Key words:. Chinese economic reform, banking system, regulation and supervision, Basel II.
Thematic Area: International and Integration Economics.

1. INTRODUCCIÓN

A finales del año 1978 comenzó en China un proceso de reforma y apertura que constituyó uno de los acontecimientos económicos de mayor relevancia del siglo XX, tanto dentro del país como fuera de él. La meta de este proceso era la construcción de una economía de mercado socialista con miras a modernizar el país más poblado del mundo, es decir, materializar un *socialismo con peculiaridades chinas*. China se embarcó en un proceso de reforma de su economía que, sin haberlo planificado de forma expresa, en la práctica ha ido introduciendo paulatinamente los mecanismos propios de la economía de mercado, presentando, por la estrategia desarrollada, un caso único en el mundo.

Los cambios que se han ido produciendo en la economía china han exigido una profunda reforma del sector financiero que, desde el inicio de la misma, intenta transformarse en un sistema que satisfaga las necesidades de una economía cada vez más orientada al mercado.

Entre las medidas adoptadas para reformar el sistema bancario cobran especial relevancia, tanto por protagonizar la actualidad de la reforma como por su importancia en el desarrollo del sistema, las destinadas a mejorar la regulación y supervisión bancaria.

Una de las principales formas de llevar a cabo ese proceso regulador y supervisor es el control de la solvencia. A nivel internacional, el Acuerdo de Capitales de Basilea (actualmente conocido como *Basilea II*) es el referente para fijar los estándares de solvencia, a través del cálculo de los fondos propios mínimos que las entidades financieras deben mantener, con relación a su nivel de riesgos, cuya medición y gestión cobran especial importancia. La aplicación de este Acuerdo en China es por tanto un indicador del grado de supervisión del sistema financiero, así como de su garantía y solidez.

2.EL PROCESO DE REFORMA ECONÓMICA EN CHINA

Podemos dividir el proceso reformista en dos fases o etapas: las primeras reformas (desde 1978 hasta mediados de los noventa), y las reformas llevadas a cabo desde ese momento y hasta nuestros días, entre las que destacan las aplicadas a partir de la adhesión de China a la Organización Mundial de Comercio (OMC) en diciembre de 2001.

2.1.Las primeras reformas económicas

La reforma que inició Deng Xiaoping en 1978 se distinguió de las anteriores en que por primera vez se antepusieron los objetivos económicos a las cuestiones políticas. De este modo, el gobierno apostó por el crecimiento económico como motor de la reforma sin plantearse un proceso de liberalización en lo político, es decir, sin abandonar el régimen comunista. Además, en contraposición a la opción ortodoxa llevada a cabo por otras economías socialistas−conocida como *big bang* o transición rápida−, China optó por una reforma caracterizada por el gradualismo −tanto geográfico como sectorial− y por la experimentación[1]. Tal y como lo expresa Nolan (1994), la transición gradual consistiría en "pisar las piedras para cruzar el río", en lugar de tratar de atravesarlo de un solo salto.

En la práctica las medidas reformistas de esta primera etapa se concretaron en dos direcciones: las Cuatro Modernizaciones y la política de puertas abiertas. Se denominó las Cuatro Modernizaciones a un conjunto de medidas liberalizadoras en el campo de la agricultura, la industria, el ejército, y la ciencia y tecnología, cuyo principal objetivo era crear progresivamente espacios de actividad económica al margen del sistema de planificación central (Naughton, 2007). Se utilizó una estrategia de doble vía, de manera que durante un tiempo convivieron la planificación y el mercado. Se adoptó un sistema dual de precios, lo que permitió introducir progresivamente los precios libres en sectores que hasta entonces habían estado muy intervenidos. En esta primera reforma las autoridades evitaron la privatización, aplicando en su lugar otras políticas para cambiar el comportamiento de los agentes y el sistema de incentivos; las relaciones entre las autoridades y los diferentes agentes se formalizaron a través de acuerdos particulares, los denominados "sistemas de responsabilidad por contratos" con similares características en el sector agrícola e industrial.

Además, en esta etapa China abandonó su tradicional autarquía y comenzó un proceso de apertura al exterior −la política de puertas abiertas− cuyos objetivos principales eran: atraer capital y tecnología foráneos (para financiar el proceso y para modernizar el sector industrial) y fomentar las exportaciones de manufacturas (para impulsar la entrada de divisas necesarias para financiar las importaciones y la deuda externa).

Así, en 1980 se inició la apertura con la creación de las zonas económicas especiales (ZEE) que se concibieron como áreas geográficas, dentro del territorio del país, donde el gobierno permitía una legislación económica más abierta que en el resto; además el gobierno incentivaba en ellas una serie de actividades productivas, habitualmente orientadas a la exportación, sin que esto fuese generalmente aplicable al resto del país. En estas zonas hubo una sorprendente

[1] Chang y Nolan (1995) analizan los rasgos más significativos de ambas estrategias de transición.

expansión en la actividad económica, así como grandes flujos de capital extranjero y de transacciones comerciales con el exterior, sobre todo exportaciones. Estas ZEE, que fueron uno de los elementos claves de la reforma, funcionaban como "laboratorios" (Ge, 1999); su éxito legitimó a las autoridades para extenderlas por toda la economía de manera que se fueron creando cada vez más zonas de similares características por todo el país.

2.2.- La segunda etapa de reformas

A medida que se iban aplicando las medidas reformistas analizadas en el apartado anterior iba surgiendo la necesidad de nuevas reformas en otros ámbitos. Además, el éxito del programa reformista de Deng Xiaoping hizo que sus sucesores continuasen el proceso. En este contexto, en la década de los noventa se produjo un giro importante en la estrategia de transición, pasando de las reformas escalonadas y cautelosas de los ochenta a un programa de liberalización mucho más profundo a partir de mediados de los noventa. Ese giro se realizó con una relativa facilidad –y con un notable éxito–, lo que, como señala Naughton (2004), constituye uno de los rasgos más llamativos de la transición hacia el mercado de China

En efecto, los resultados de las reformas llevadas a cabo hacían de China una economía cada vez más capitalista, lo que a su vez hacía necesario cambios normativos e institucionales para afrontar ese nuevo contexto. De este modo, las medidas reformistas de esta etapa (que duran hasta hoy) se han centrado en fortalecer –y en su caso crear– las instituciones propias de una economía de mercado, en mejorar el entorno legal y regulatorio y en crear nuevas reglas acordes con su nueva economía.

Al mismo tiempo comienza a reducirse la participación del sector estatal en la economía. En los años ochenta las autoridades intentaron aumentar gradualmente la autonomía empresarial con el fin de fomentar su competitividad y eficiencia –es decir una "mercantilización sin privatización" (Bustelo, 1996)–; sin embargo, a mediados de los noventa se puso en marcha un programa de privatización gradual de empresas. Así, en 1997, y por primera vez en China, la reforma de las empresas estatales se tomó como la pieza central de la siguiente etapa de reformas (Perkins, 1997) y Jiang Zemin (presidente de la República desde 1989 hasta 2003) anunció su determinación de reformar las más de cien mil empresas públicas del país.

Al margen de las empresas estatales, las medidas más importantes en los primeros años de esta segunda etapa reformista se llevaron a cabo en el sistema fiscal e impositivo y en el sistema financiero y bancario (que será analizado en el siguiente apartado). Desde la adhesión de China a la OMC, en diciembre de 2001, las medidas de transición hacia el mercado se han acentuado y se han extendido a la mayoría de los sectores. Como consecuencia de esta adhesión, China ha tenido que asumir la normativa de esta organización[2], es decir, una serie de compromisos de apertura y liberalización de sus relaciones económicas con el exterior, que han dado lugar a nuevas reformas. No obstante, hay que tener en cuenta que el Protocolo de Adhesión de China no establecía el acceso pleno, inmediato y sin limitaciones de productos, servicios o proveedores extranjeros. En su lugar, el Estado chino se comprometió a eliminar gradualmente los obstáculos al comercio y ampliar el acceso a su mercado, pero estableciendo un calendario específico de liberalización para cada tipología de productos y servicios, que además no siempre culmina con la plena libertad de acceso[3].

La entrada en la OMC supuso, y aún está suponiendo, la lenta –pero progresiva– apertura de un gran número de sectores claves de la economía china. Los compromisos asumidos por China en materia de comercio de servicios (GATS)[4] afectan a un número de sectores cuyo acceso a la inversión extranjera estaba hasta el momento de la adhesión prohibido o muy restringido. De ellos se derivan la apertura de la inversión extranjera a actividades como la banca, los seguros,

[2] En ella tienen especial relevancia los siguientes acuerdos: Acuerdo GATT (Acuerdo General sobre Aranceles Aduaneros y Comercio), que abarca el comercio internacional de mercancías, Acuerdo GATS (Acuerdo General sobre el Comercio de Servicios), Acuerdo TRIPS (Acuerdo sobre Derechos de Propiedad Intelectual relacionados con el Comercio) y Acuerdo TRIMS (Acuerdo sobre las Medidas en materia de Inversiones relacionadas con el Comercio).

[3] Véase Protocolo de Adhesión de la República Popular China (documento WT/L/432).

[4] Estos compromisos están recogidos en el Proyecto de Informe del Grupo de Trabajo sobre la adhesión de la República Popular China: Lista de compromisos específicos en materia de servicios. Documento: WT/ACC/CHN/49/Add.2.

la distribución comercial y las telecomunicaciones. En la mayor parte de los sectores la liberalización se está llevando a cabo de forma gradual, tanto en el ámbito geográfico como en los porcentajes de participación extranjera permitidos, cumpliendo el calendario de liberalización previamente acordado. De todos modos, este proceso está todavía inacabado; de hecho, el desarrollo del sector de los servicios figura como prioridad en el Undécimo Plan Quinquenal para el Desarrollo Económico y Social Nacional[5]. En él se expone como propósito que China siga abriendo el sector servicios y eliminando barreras gradualmente a los inversores extranjeros –aunque en algunos sectores todavía no se permite la participación de las empresas de capital totalmente extranjero (WFOE)– y en las empresas mixtas. No obstante, es preciso recalcar que los compromisos asumidos por China con la OMC respecto del sector servicios son muy amplios. En opinión de Mattoo (2004) han representado la reforma más radical de este tipo en la historia de la OMC, aunque están sujetos a reservas o a salvedades más frecuentes que en el caso de otros países.

Esta segunda fase de reformas ha supuesto una transformación fundamental en la economía china, que, aunque en general ha tenido resultados claramente positivos, también ha tenido cargas que soportar. El grupo que más ha perdido su posición "privilegiada" ha sido el de los trabajadores estatales (Naughton, 2004). Algunos de ellos han seguido empleados en compañías de accionariado común, en las cuales el Estado era el accionista mayoritario, o en empresas totalmente privatizadas, pero millones de ellos fueron despedidos, disparando el desempleo urbano desde finales de los noventa. Por todo ello, esta segunda fase de reforma, a diferencia de la primera[6], puede considerarse una "reforma con perdedores" (Naughton, 2007), lo que constata que a medida que China avanza en la transición el reparto de los beneficios de la misma es menos equitativo.

3.- LA REFORMA DEL SISTEMA BANCARIO

3.1.- Características del sistema financiero chino

Hasta muy recientemente el sistema financiero se había quedado detrás del resto de la economía en su proceso de transición. En concreto, el sector bancario ha sido una de las industrias más protegidas de la competencia exterior y ha estado sobrerregulado y dominado por la propiedad estatal. El crecimiento y el desarrollo de la economía china ha ido planteando nuevas demandas al sistema financiero. Aunque se ha reestructurado el sistema bancario y se han producido muchos cambios e innovaciones, las instituciones financieras que operan en China aún no realizan el papel dinámico e independiente propio de estas entidades en una economía de mercado.

Estas características del sistema financiero nos hacen plantearnos dos cuestiones importantes. En primer lugar, suele aceptarse que el papel del sistema financiero es un elemento crucial en el desarrollo de la economía de un país; sin embargo, China ha conseguido unos niveles de desarrollo y crecimiento económico muy altos con un sistema financiero retrasado[7], lo que puede resultar paradógico (Yao y Yueh, 2009).

Por otro lado, como veremos más tarde, el hecho de tener un sistema financiero no sometido a las reglas del mercado permitió, según Naughton (2007), que la primera etapa de reformas fuera una "reforma sin perdedores", dado que el papel de este sistema –ineficiente desde el punto de vista de la asignación de recursos– amortiguó los efectos negativos de esta etapa de cambios[8].

No obstante, el desarrollo de la economía está demandando otro sistema financiero; de hecho, éste se encuentra en medio de un periodo de cambios significativos, y es uno de los

[5] Las líneas generales pueden verse en la información en línea de la Comisión Nacional de Desarrollo y Reforma (NDRC): http://en.ndrc.gov.cn/hot/t20060529 _71334.htm.
[6] Para un análisis detallado de esta cuestión véase Lau, Qian, y Roland (2000).
[7] Allen, Qian *y* Qian (2005) documentan los diferentes aspectos en los que los sistemas financieros y legales en China están subdesarrollados.
[8] El gobierno había usado al sistema bancario para compensar a los grupos que de otra forma habrían resultado muy perjudicados por el proceso de transición, sobre todo a los trabajadores de las empresas estatales no viables.

puntos más importantes de atención del gobierno. Se podría decir que hoy en día el sistema financiero es "profundo pero estrecho" (Naughton, 2007); entendiendo la profundización financiera como la participación de productos financieros en la economía (la cual es cada vez mayor, de manera que la economía cada vez depende más de las finanzas) y la "estrechez" en el sentido de la poca diversidad de servicios financieros (el sector está dominado por el sistema bancario, en concreto por los bancos de propiedad estatal).

Además, como veremos, el sector cada vez está más sometido a la competencia exterior (sobre todo desde diciembre de 2006) como consecuencia de los compromisos de adhesión a la OMC y del cumplimiento del calendario de liberalización del sector. Asimismo, los mercados de capitales, incluyendo los mercados bursátiles, también están inmersos en un proceso de cambios, reformas y de comienzos de regulación (Fung, Pei y Zhang, 2006).

Estos rasgos del sistema financiero conducen a temer que el sistema pueda quedarse por detrás de otros aspectos del desarrollo económico chino y sea, por tanto, una fuente posible de vulnerabilidad económica.

3.2.- El sistema bancario antes de la reforma.

Hasta que comenzó la reforma económica china, el sistema financiero tenía como función atender los planes de producción e inversiones del Gobierno, como un sistema propio de una economía planificada. No existían mercados de capitales, y el Banco Popular de China (PBC) era prácticamente el único banco[9], con lo que asumía la doble función de banco central y de banco comercial. Este sistema es denominado por muchos autores como *monobank system* (monobancario). El PBC era la vía para canalizar los beneficios de las empresas estatales hacia otras empresas o hacia sectores prioritarios, lo cual se realizaba a través de un plan de crédito administrado y de conformidad con el Gobierno (Bustelo y Fernández Lommen, 1996). Es decir, la banca tenía como función principal proporcionar crédito comercial y servicios de pago para facilitar el comercio, pero no financiaba proyectos de inversión a largo plazo, los cuales, cuando era necesario, eran financiados por el gobierno a través de su presupuesto (Naughton, 2007). Además era un sistema en el que las decisiones importantes nunca estaban basadas en consideraciones financieras, sino atendiendo a las necesidades propias de la planificación central.

3.3.- Reforma del sector

A mediados de los ochenta se produce la primera reforma financiera a través de una reestructuración del sistema bancario que fraccionó el sistema monobancario. A partir de ese momento las funciones de banco central seguirían en manos del PBC mientras que para realizar las labores propias de la banca comercial, se crearon cuatro grandes bancos de propiedad estatal (*los Cuatro Grandes*), cada uno de ellos especializado en un sector (He, 2005): el Banco Agrícola de China (ABC)[10], vinculado a las zonas agrícolas y rurales, el Banco Industrial y Comercial de China (ICBC), vinculado a los sectores industriales y comerciales en las zonas urbanas, el Banco de la Construcción de China (CCB), especializado en la financiación de proyectos de construcción e infraestructura, y el Banco de China (BOC), especializado en transacciones internacionales. Con esta medida, el gobierno chino intentó cambiar la forma de financiación de las empresas estatales; estos cuatro bancos asumieron la responsabilidad de aportar capital (a través de préstamos, con un tipo de interés determinado) a las empresas de propiedad estatal, que antes eran financiadas a través de ayudas presupuestarias (sin coste alguno). De este modo estos cuatro grandes bancos se acostumbraron a conceder préstamos a demanda del gobierno, lo que dio lugar a un importante volumen de préstamos fallidos, debido fundamentalmente a que se habían cargado al sistema bancario público grandes deudas contraídas por las empresas estatales (muchas de las cuales operaban con pérdidas).

[9] No obstante, en esta época ya existía una amplia red de bancos comerciales de pequeño tamaño, las cooperativas de crédito rurales, dedicadas a captar el ahorro (cuando éste existía) y a la concesión de créditos al comercio.

[10] Las abreviaturas utilizadas en todo el trabajo corresponden a las siglas en inglés de las distintas instituciones, tal y como las utilizan el Banco Popular de China y la Comisión de Reglamentación Bancaria de China.

Como consecuencia, a mediados de los noventa, las autoridades chinas crearon tres nuevos bancos estatales especializados (denominados bancos de desarrollo o bancos políticos) con el fin de desempeñar funciones de política de préstamos para relevar a los Cuatro grandes en esa función: el Banco de Desarrollo, el Banco de Exportación e Importación y el Banco de Desarrollo Agrícola. Además, en 1995 se promulgaron la Ley sobre el Banco Popular de China −que establecía el marco jurídico para el funcionamiento del PBC− y la Ley sobre bancos comerciales −que creaba las condiciones para formar un sistema de banca comercial−.

A medida que avanzaba la reforma económica los hogares y los nuevos negocios que se iban creando aumentaron su nivel de ahorro; éste se dirigió al sistema bancario, lo cual hizo que el sistema se fuera convirtiendo en un importante canalizador para que el ahorro de los hogares fuera invertido en las empresas, adoptando el rol típico de una economía de mercado. También en este sentido China contrasta con otras economías en transición, sobre todo las europeas, dado que, en muchos casos, el nivel de inflación que soportaron hizo que el sistema bancario no captara el ahorro de las familias.

No obstante, existía un volumen muy alto de préstamos fallidos y el sistema bancario seguía necesitando un cambio drástico. Por eso, desde finales de los noventa, comenzaron a ponerse en práctica una serie de medidas aperturistas encaminadas a crear un sector competitivo y rentable. En aquel momento la participación extranjera era muy pequeña y aún muy restringida (por moneda, tipos de clientes o geográficamente). Para ello se redujo el intervencionismo estatal y el sector comenzó a abrirse gradualmente a la inversión y competencia exteriores y, como veremos más adelante, se introdujeron nuevas medidas para resolver el problema de los créditos fallidos, sobre todo de los Cuatro grandes bancos. En 1999 se creó una compañía estatal de gestión de activos impagados para cada uno de ellos.

Un elemento clave del proceso de liberalización financiera ha sido la adhesión de China a la OMC, según la cual China se comprometió a cumplir un calendario de apertura del sistema bancario a la competencia exterior (Liao, 2006). Desde la adhesión, todos los bancos −incluso los extranjeros− han podido realizar ciertas transacciones en moneda extranjera, particularmente en lo que concierne a las actividades de depósito. Asimismo, desde 2003, los bancos extranjeros han sido autorizados a operar en el mercado mayorista en moneda local, pero con límites geográficos estrictos. Y en 2006 se suprimieron las restricciones geográficas y las aplicadas a los clientes con respecto a las actividades comerciales en renminbis, así como otras restricciones no cautelares aplicadas a las operaciones de los bancos extranjeros[11].

También en 2003 se establecieron nuevas estructuras de regulación; entre ellas, la más importante es la Comisión de Reglamentación Bancaria de China (CBRC), un organismo público de control y regulación que vela por el correcto funcionamiento y administración de todo el sistema bancario chino, para lo cual tiene la capacidad de emitir normativas o de llevar a cabo controles e inspecciones. Éste ha sido un importante paso administrativo necesario para cumplir con la OMC, además de ser esencial para la siguiente etapa de reformas en el sector bancario, dado que éste camina hacia un aumento de la competencia y del capital privado que hace necesaria esa mayor regulación (Naughton, 2007).

3.4.- El sistema bancario en la actualidad

3.4.1.- Composición

El sistema bancario chino está formado por un núcleo de instituciones de propiedad estatal, la mayoría descendientes del sistema de planificación central.

Las tres primeras categorías a), b) y c) acaparan las tres cuartas partes de los activos del sistema y son el núcleo de los esfuerzos de reestructuración del gobierno.

a) Bancos comerciales de propiedad estatal (SOCBs)

Como ya vimos, los Cuatro grandes descienden directamente del sistema bancario de planificación central, y nacieron en los ochenta como subdivisiones del único banco del

[11] Para ver un análisis del papel de la banca extranjera en el sector bancario chino consúltese Santabárbara (2007).

gobierno. Son organizaciones gigantescas, cada uno de ellos con cientos de miles de empleados, y se encuentran entre las cincuenta instituciones financieras más grandes del mundo en términos de activos.

Los cuatro tenían el peso de su legado del sistema antiguo, es decir, estaban sobreempleados y tenían carencia de capacidades y ausencia de orientación de negocios. Pero, como ya vimos, lo más importante es que, dado su objetivo en origen, estaban acostumbrados a conceder créditos a demanda del gobierno, de forma que su subordinación política incrementó considerablemente los créditos incobrables.

Desde mediados de los noventa (segunda fase de la reforma económica) se ha hecho un gran esfuerzo en ellos para colocarlos sobre bases comerciales. Concretamente, a raíz de la aprobación de la Ley de Bancos Comerciales en 1995, se establecieron las bases para transformar el sistema de bancos estatales *especializados* en un sistema de bancos estatales *comerciales*. Se aprobaron nuevas leyes para proporcionarles protección para la toma de decisiones independiente[12] y se tomaron medidas para disminuir la influencia que sobre estos bancos tenían los gobiernos locales. Con estas medidas han conseguido mayor autonomía, aunque dependen más del gobierno central. Además, en el año 2007 la CBRC pasó a incluir formalmente un quinto banco dentro de este grupo: el Banco de Comunicaciones; este banco fue creado originalmente en 1908 y reestablecido como un Banco Comercial de Capital Mixto desde 1986, pero la fuerte presencia del Gobierno Central en su capital y el tipo de actitudes que desarrollaba hacían que fuera considerado frecuentemente como un SOBC.

En la actualidad estos Bancos están inmersos en un gran programa de reestructuración de vital importancia para el gobierno chino. De hecho, probablemente el mayor reto de todo el sistema financiero sea conseguir reestructurar bien estos grandes bancos (Naughton, 2007). El objetivo final de esta reestructuración es transformar estas entidades en Bancos Comerciales de Capital Mixto con presencia en los mercados de valores (a través de una Oferta Pública Inicial), todo ello con el fin de mejorar su gestión.

Actualmente estos cinco bancos representan más del 50% de los activos totales del sistema bancario (véase Tabla nº1) y acaparan más de la mitad de los trabajadores del sector en China (CBRC, 2008).

b) Bancos comerciales de capital mixto (JSCBs)

En la actualidad son 12 y son radicalmente distintos a los anteriores. Se fueron estableciendo entre 1986 y 2001, por lo que tienen menos peso en sus carteras de la era de planificación central, de manera que acumulan una cantidad mucho menor de préstamos fallidos en sus cuentas.

Han introducido un elemento de competencia en el mercado bancario, y han ido ganando cuota de mercado hasta representar en la actualidad el 13,78% del total de activos del sistema, como puede verse en la Tabla nº 1. Varios cotizan en el mercado de valores, de modo que una parte de sus acciones (de momento una minoría) está en manos del público. Sin embargo no están libres de problemas, sobre todo por la influencia que ejercen sobre ellos los gobiernos locales y sus relaciones con estos.[13]

c) Bancos comerciales urbanos (CCBs)

Los bancos comerciales urbanos fueron creados a partir de las cooperativas de crédito urbanas, dedicadas a proporcionar préstamos a pequeñas empresas urbanas. La mayoría (hoy son 124 bancos) se crearon a partir de 1998; generalmente operan a escala local, y son controlados predominantemente por los gobiernos locales. En comparación con los JSCBs, estos son más pequeños, generalmente sus activos son de menor calidad y su personal está menos cualificado.

[12] El plan de crédito nacional (por el que el gobierno tomaba las decisiones sobre préstamos) se abolió definitivamente 1997.
[13] No son propiedad del gobierno nacional, pertenecen a diversos grupos de empresas estatales (SOEs), de agencias del gobierno y de agencias no estatales.

d) Otros bancos

Los datos de China engloban normalmente el resto de instituciones en "otros", categoría que representa un cuarto del total de activos del sistema. No obstante, las instituciones que aquí se incluyen son muy diferentes, por lo que resulta conveniente efectuar, al menos, la siguiente subdivisión:

d.1. Bancos de desarrollo estatales (Policy banks)

En conjunto representaban 8,13% del total de activos en 2007. Como ya vimos se crearon con el fin de liberar a los Cuatro de la presión de llevar a cabo proyectos "políticamente populares". Estos tres bancos de desarrollo han tomado caminos diferentes: el Banco de Desarrollo ha expandido su negocio hacia proyectos de infraestructuras a gran escala, de forma muy agresiva y empresarial; el Banco de Exportación e Importación, se dedica a la promoción de exportaciones tradicional; y el Banco de Desarrollo Agrícola se ha desarrollado en el mundo rural.

d.2. Cooperativas de crédito rurales

Se desarrollaron durante los años cincuenta como parte de un sistema de organización colectiva en el campo. Aunque se integraban débilmente en el sistema financiero estatal, nunca formaron parte del monobanco. Tuvieron un papel muy positivo a la hora de aportar financiación comercial al campo para la adquisición de los inputs modernos cruciales para el desarrollo agrícola. Además, cambiaron mucho durante la primera fase de la reforma agraria. Estas Cooperativas canalizaron el ahorro local hacia el desarrollo de las TVE´s[14], que en aquel momento obtenían grandes beneficios, lo cual hizo que los recursos de las cooperativas aumentaran mucho durante los ochenta. Pero el entorno más competitivo de los años noventa hizo que los beneficios de las TVE´s cayeran y comenzaran a tener problemas para devolver préstamos a las Cooperativas, que empezaron a tener dificultades.

Dados sus malos resultados y su gran volumen de préstamos fallidos, en la actualidad estas Cooperativas también se encuentran inmersas en un programa de reestructuración a través de la inyección de capital y de nuevos programas de préstamos, incluidos microcréditos (Gale y Collender, 2006). El objetivo final es revitalizar este tipo de institución rural, cuya importancia potencial es crucial para que el campo salga adelante.

d.3. Otros bancos

El resto de instituciones que forman parte del sistema bancario tienen un peso mucho más reducido que los anteriores. Entre las instituciones que lo forman, la más importante en volumen es el Sistema de Ahorro Postal (creado a imagen y semejanza del Japonés). También tienen un papel relativamente importante los Bancos Extranjeros, aunque por ahora su participación en el volumen total de activos del sistema sea muy pequeña (alrededor del 2% del total), entre otras razones porque su incorporación al sistema bancario chino es muy reciente (en el futuro se espera que los bancos extranjeros tengan un papel significativo).

En conjunto, el sistema bancario chino aún está dominado por las entidades controladas por el Estado. En 2007 el grupo de los bancos estatales, más los tres Bancos de Desarrollo y las Cooperativas rurales representaban casi el 70% de los activos del sistema. Un 20% de activos está en manos de JSCBs y City Banks, bajo tutela de gobiernos locales, y el 10% de activos en el resto de instituciones.

[14] Las TVE´s (town-ship and village enterprises) son pequeñas empresas que operan normalmente a pequeña escala, exhiben toda una variedad de regímenes de propiedad, siendo predominante la propiedad colectiva por aldeas y pueblos, y trabajan en condiciones muy próximas al mercado. Su importancia y contribución al éxito de la reforma industrial china y a su transición hacia el mercado es resaltada de forma generalizada.

TABLA Nº 1: Activos totales de las instituciones bancarias 2003-2007

	Total Activos (100 millones RMB)					Porcentaje				
	2003	2004	2005	2006	2007	2003	2004	2005	2006	2007
Bancos comerciales de propiedad estatal	160.511,7	179.816,7	210.050,0	242.363,5	280.070,9	58,03	56,91	56,06	55,15	53,25
Bancos comerciales de capital mixto	29.598,6	36.476,0	44.654,9	54.445,9	72.494,0	10,70	11,54	11,92	12,39	13,78
Bancos comerciales urbanos	14.621,7	17.056,3	20.366,9	25.937,9	33.404,8	5,29	5,40	5,44	5,90	6,35
Bancos de desarrollo estatales	21.247,0	24.122,5	29.283,2	34.732,3	42.781,0	7,68	7,63	7,81	7,90	8,13
Cooperativas de crédito rurales	26.509,2	30.767,0	31.426,7	34.502,8	43.434,4	9,59	9,74	8,39	7,85	8,26
Banco Postal de ahorro	8.984,4	10.849,6	13.786,8	16.122,0	17.687,5	3,25	3,43	3,68	3,67	3,36
Bancos extranjeros	4.159,7	5.822,9	7.154,5	9.278,7	12.524,7	1,50	1,84	1,91	2,11	2,38
Otras instituciones bancarias	1.853,1	2.352,2	7.812,0	11.522,4	13.868,2	0,67	0,75	2,08	2,62	2,64
Otras instituciones no bancarias	9.100,0	8.726,8	10.161,9	10.594,1	9.717,0	3,29	2,76	2,71	2,41	1,85
TOTAL	276.585,4	315.990,0	374.696,9	439.499,6	525.982,5	100,00	100,00	100,00	100,00	100,00

Fuente: CBRC (2008), Annual Report 2007.

3.4.2.- Las debilidades del sistema bancario y el programa actual de la reforma: perspectivas y retos

El sistema bancario chino comenzó el siglo veintiuno sin estar ajustado a las demandas de una economía de mercado sofisticada, en parte porque era el heredero de un sistema creado para una economía planificada, que no tenía capacidades, incentivos, o la cultura necesaria para convertirse rápidamente en un sistema bancario comercial efectivo. Como ya vimos, ese sistema amortiguó durante la primera etapa de reformas económicas los efectos de una competencia creciente sobre los trabajadores de las empresas estatales, pero al mismo tiempo implicó continuos préstamos a empresas inviables (de hecho en muchas ocasiones sin esperanza de que fueran devueltos), lo que ocasionó a la larga un gran volumen de préstamos fallidos; además, dada la política del gobierno financiando proyectos no productivos, esa cifra continuó aumentando durante los años posteriores. Esto ha supuesto, además de una asignación ineficiente de recursos, un lastre de mucho peso para el sistema bancario.

A finales de los noventa el gobierno reconoció la necesidad de inyectar recursos al sistema bancario para evitar una crisis económica potencial. Desde ese momento este problema se convirtió en una prioridad para el gobierno chino, que comenzó a tomar medidas para intentar solucionarlo. En realidad ese problema puede subdividirse en dos vertientes (Naughton, 2007): el problema de stock y el de flujo. El problema de stock haría referencia a la dificultad para limpiar el volumen existente de préstamos fallidos en las cuentas de los bancos y vendría a reflejar el legado de los costes pasados. Por otro lado, el problema de flujo se referiría a la necesidad de prevenir que las nuevas decisiones de préstamo cayesen en préstamos de elevado riesgo, en lugar de tomar decisiones en base a criterios comerciales; este problema reflejaría la necesidad de mantener prácticas saludables ahora y en el futuro, para lo cual es particularmente útil una adecuada regulación y supervisión bancaria.

Entre las medidas que el gobierno chino ha introducido para paliar el problema del stock, destaca la creación en 1999 de cuatro compañías de gestión de activos (una para cada uno de los SOCBs) cuya misión fue comprar esos préstamos fallidos para intentar recuperar todo lo posible mediante su venta a terceros. Además desde finales de 2003 se puso en marcha una nueva inyección de capital en diversos bancos a raíz de la creación de la Compañía de Inversión Central Huijin.

Aunque estas medidas de recapitalización y reestructuración han supuesto una importante reducción del stock de préstamos incobrables, han sido muy costosas para el gobierno chino. Asimismo, reducir el problema de stock puede hacer aún más difícil el problema de flujo, porque los salvamentos actuales pueden desincentivar la correcta toma de decisiones sobre préstamos futuros. Todas estas cuestiones hacen muy necesaria la introducción de un sistema de incentivos que responsabilice a los dirigentes de sus decisiones sobre riesgos (Naughton, 2007).

Otra medida para apoyar la recapitalización y reestructuración bancaria ha sido la introducción de inversores extranjeros estratégicos en diversos bancos, lo que constituye uno de los cambios más drásticos en la política de China. En efecto, desde diciembre de 2003 el gobierno chino ha promovido que los bancos lleguen a acuerdos con instituciones extranjeras de acuerdo con el plan de reforma estructural de cada entidad. El objetivo del gobierno no es sólo introducir capital en los bancos chinos, sino también que esos inversores *estratégicos* aporten sistemas y técnicas novedosas de gestión y que ayuden a mejorar el gobierno de los bancos comerciales chinos.

Por último, la reforma de la estructura de propiedad de los bancos estatales ha conllevado también, como ya señalamos, la transformación de su sistema de accionariado y su salida a los mercados de valores. El Informe del gobierno presentado por China a la OMC[15] para su último Examen de Políticas Comerciales manifiesta que se ha logrado un avance decisivo en la reforma del sistema de valores de los bancos comerciales de propiedad estatal. En los últimos cuatro años, el Banco Industrial y Comercial de China, el Banco de China y el Banco de la Construcción de China han completado sucesivamente la reforma de su sistema de acciones, y

[15] Documento de la Organización Mundial de Comercio: WT/TPR/G/199, de mayo de 2008.

han salido a bolsa en los mercados de valores nacionales y/o extranjeros. Con ello, ha cambiado su estructura de propiedad, aunque la parte estatal sigue siendo claramente dominante. Se han producido cambios importantes en la gestión empresarial y en el concepto de administración de estos bancos comerciales. El resultado ha sido una mayor consolidación general y una mejor calidad de activos y aptitudes de gestión.

En definitiva, el panorama está cambiando mucho hacia un sistema con mayor competencia y mayor participación de la banca extranjera, bien a través de sus acciones en los bancos chinos –que hasta ahora ha sido la principal vía de entrada de capital extranjero–, o a través sus propias sucursales y filiales[16], aunque el gobierno ya ha impuesto algunas barreras "tenues" de entrada a extranjeros limitando la extensión de sucursales o filiales a través de requisitos de capital, retrasos reguladores, límites de actividad, etc. (Santabárbara, 2007).

A pesar de todos los cambios, persisten vulnerabilidades importantes y el sistema financiero de China es todavía incapaz de lograr plenamente su tarea central: la asignación de crédito de un modo eficiente a todos los sectores. Además, casi todos los progresos se están centrando en los SOBCs y en unos pocos JSCBs, mientras que las Cooperativas rurales siguen teniendo muchos problemas, así como los Bancos de Desarrollo con los préstamos fallidos.

Además, los niveles de eficiencia del sector bancario todavía son bajos y sus costes son altos, y aún existen decisiones erróneas sobre préstamos, transacciones lentas y procedimientos complicados. Para mejorar la eficiencia es necesario redefinir el alcance del negocio bancario (por ejemplo, aumentar sus ingresos a través de comisiones y de tarjetas de crédito), diversificar el sistema para tener más seguridad (un sistema que estaba centrado en préstamos a las empresas estatales y que hoy camina hacia el aumento de activos financieros como préstamos a los consumidores, hipotecas, empréstitos, etc.) e incrementar la especialización[17].

El sistema también necesita el establecimiento de mecanismos de supervisión y control internos (dado que hasta ahora ha existido mucha discrecionalidad en la concesión de préstamos) y la introducción de incentivos en los contratos del personal bancario para la correcta valoración de riesgos, para la realización de provisiones para riesgos futuros, etc., es decir, desarrollar un adecuado sistema de análisis de control de riesgos (Podpiera, 2006).

Por todo ello, se espera ver un periodo extenso de cambios, en el que muchos bancos chinos quebrarán mientras que otros tendrán éxito; en definitiva, existen riesgos significativos en el proceso de reforma del sistema bancario, pero también alta probabilidad de éxito. En este sentido podríamos decir que el sistema financiero chino encara diversos retos.

El primero de estos retos es evitar la inestabilidad en el proceso de apertura y desarrollo financieros, especialmente si tenemos en cuenta que China tiene mucha profundidad monetaria (más que cualquier otro país desarrollado, incluido Japón). Ese tamaño y la poca diversidad de fórmulas financieras –fundamentalmente son préstamos– son las fuentes principales que generan temores de inestabilidad financiera (Naughton, 2007). Además, es importante la manera en que el sistema bancario chino afronte la entrada en el sector de la competencia extranjera, lo cual resulta difícil incluso después de años preparándose. El previsible aumento de la competencia que encara el sector hace aún más relevante el papel de la regulación y supervisión del sistema bancario.

En segundo lugar, China necesita desarrollar un círculo virtuoso a través de un sistema financiero que funcione bien. La regulación y supervisión financiera son componentes cruciales para mejorar la gestión de las corporaciones, dado que aumentarían la confianza, por un lado, de la población china en su sistema bancario –lo que facilitaría la aparición de nuevos activos financieros– y, por otro lado, de los bancos extranjeros a medida que vayan aplicándose las normas o parámetros internacionales en el sistema chino.

De cualquier forma, algunos autores (Lu y Yao, 2009; Yao y Yueh, 2009) destacan que aún más importante que un buen diseño de la regulación, es que esa regulación sea finalmente efectiva y genere buenas prácticas en el sistema financiero, lo cual no depende solamente del

[16] No debemos olvidar que en los últimos años se está produciendo también una lenta aparición de agentes bancarios extranjeros, con motivo de la finalización del periodo de transición contemplado en el Acuerdo de adhesión de China a la OMC.

[17] He (2005) apunta como a lo largo del proceso de desarrollo financiero se ha producido una diversificación del mercado financiero, pero no una especialización de los agentes que operan en esos mercados, probablemente porque esos agentes proceden, en su mayoría, del sistema gubernamental.

diseño normativo, sino también del resto de cuestiones no formales e instituciones socioculturales que condicionan el comportamiento de los agentes.

4.- LA REGULACIÓN Y SUPERVISIÓN BANCARIA

4.1. -Situación internacional

La regulación bancaria es de carácter singular, dado que su objetivo es evitar el peligro de contagio de las crisis bancarias. Su inicio puede situarse en la Gran Depresión (años 30), a partir del diseño de la Glass-Steagall Act estadounidense, que pretendía restablecer la estabilidad financiera del país, combinando la creación de un seguro de depósitos, la separación de funciones en la banca y la restricción a determinadas actividades operativas de la banca comercial. Lo más significativo, no obstante, fue la toma de conciencia de la necesidad de que las entidades financieras cuenten con un respaldo suficiente para afrontar dificultades, convirtiéndose así sus fondos propios (fundamentalmente, el capital) en el vehículo de control de la solvencia. Esta medida tiene además importantes connotaciones estratégicas, al condicionar la estructura financiera de estas entidades como combinación de financiación propia y ajena.

Uno de los principales mecanismos de control de la solvencia se estableció, por tanto, mediante la fijación de un capital mínimo a soportar que permita la cobertura de los riesgos asumidos. Además, poco a poco se hizo visible la necesidad de fijar algún tipo de normativa internacional, más allá de las regulaciones locales, a medida que la actividad de los bancos se ampliaba y los mercados superaban las fronteras nacionales.

El objetivo de la supervisión bancaria no es, por tanto, evitar el fracaso de las entidades a nivel individual, sino garantizar la estabilidad del sistema, protegiéndolo del peligro de una crisis de enormes dimensiones; lo que distingue a las quiebras bancarias no es sino el temor a que un fracaso individual se extienda a otras entidades, e incluso más allá del sistema bancario hacia el sistema financiero en conjunto, afectando a las variables macroeconómicas.

Debido a la preocupación creciente por la salud financiera de los bancos internacionales y a las quejas sobre competencia desleal, en los años 80 el Comité de Supervisión Bancaria de Basilea (BCBS), uno de los comités permanentes del Banco de Pagos Internacionales (BIS), encargado de mejorar la colaboración entre las entidades de supervisión bancaria, comenzó a considerar propuestas para fijar estándares de capital en estas entidades. Consecuencia de ello es el primer Acuerdo de Convergencia de Capitales (Basilea I) publicado en 1988. Pese a carecer de efectos legales, su aprobación por parte de los países del G-10 (miembros del BCBS) facilitó su aceptación final en más de 100 países, siendo trasladado a las normativas nacionales a fin de mejorar la solvencia de las entidades financieras, especialmente las de actividad internacional, sometidas a un capital mínimo del 8% de sus activos, clasificados en cuatro grandes categorías definidas de mayor a menor riesgo en atención al posible impago de los clientes, esto es, al riesgo crediticio.

El riesgo de crédito era, en ese momento, el más importante en la actividad bancaria, dado que se refiere a la posibilidad de que la contraparte de una operación no cumpla lo acordado y ocasione una pérdida. Surge pues en las operaciones de activo y habitualmente se identifica con el riesgo de insolvencia, presente en la financiación concedida por los bancos.

A pesar de sus indudables bondades, como una creciente prudencia en la actividad de las entidades de crédito, el Acuerdo fue quedándose obsoleto debido a la aparición de nuevos productos financieros y nuevas categorías de riesgo, y a los graves problemas financieros sufridos por algunas entidades, de gran trascendencia mediática.

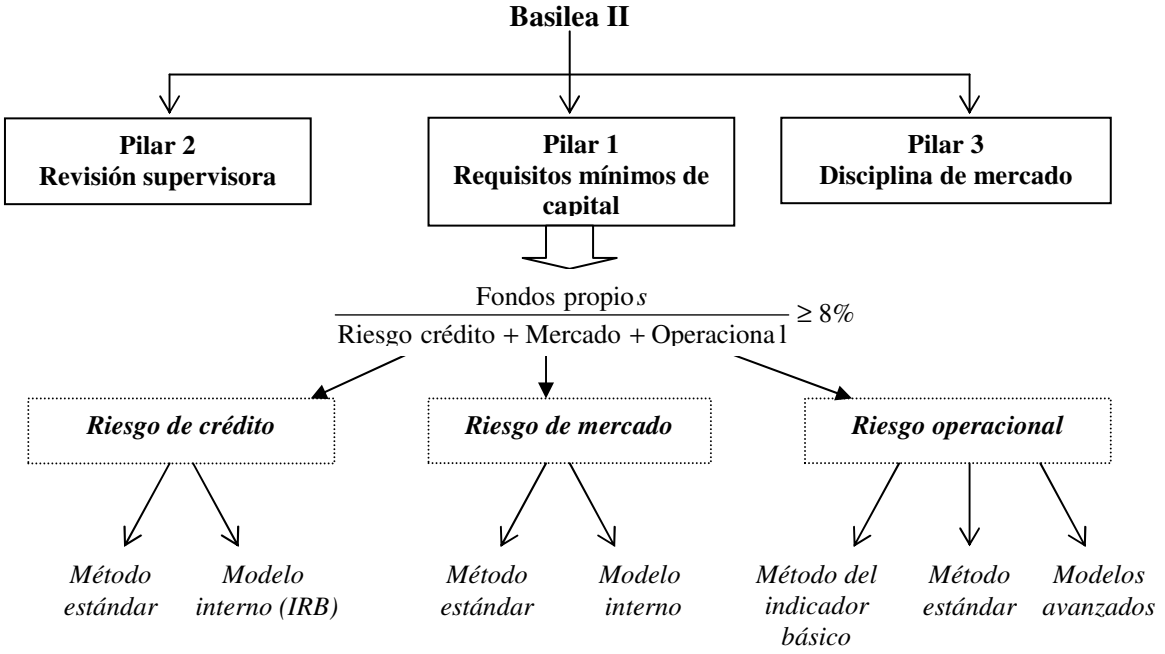
En términos de riesgos, es frecuente identificar, además del riesgo de crédito, riesgos de mercado, consecuencia de posibles pérdidas por cambios adversos en el precio de algún instrumento. Los casos más habituales son el riesgo por tipo de interés, tipo de cambio, operaciones con derivados, pago anticipado o aseguramiento. Además, en los últimos años se ha tipificado una tercera categoría, de carácter más cualitativo, denominada riesgo operacional. En ella se incluyen los riesgos por errores en los procesos (realización de operaciones, problemas

de documentación, contabilidad), fallos humanos (fraude de empleados), fallos de sistemas (equipos informáticos, infracciones de seguridad) o eventos externos (sucesos naturales, ataques terroristas).

Por eso, el Acuerdo fue inicialmente modificado para incluir los riesgos de mercado (BCBS, 1996) y, posteriormente, para añadir modelos más sensibles al riesgo desarrollados por las propias entidades financieras (modelos internos) y que deben contar el visto bueno de los reguladores, cuya complejidad superior a los conocidos como "métodos estándar". Además, la quiebra de algunas entidades en los años 80-90 y los casos de fraude llevaron a incorporar el tratamiento del riesgo operacional.

El proceso de redacción del Nuevo Acuerdo de Capitales (más conocido como Basilea II) comenzó en 1999 con la publicación de un primer documento de consulta (BCBS, 1999), seguido de una segunda versión (BCBS, 2001), una tercera (BCBS, 2003b) y la publicación del Acuerdo definitivo el 26 de junio de 2004 (BCBS, 2004), cuya estructura resume la figura 1.

FIGURA Nº 1: El Acuerdo de Basilea II



Fuente: Elaboración propia.

Basilea II se divide en tres grandes apartados o pilares que corresponden al cálculo de los recursos propios mínimos (referidos a riesgos de crédito, mercado y operacional, cifrados mediante modelos estándar o modelos internos), revisión supervisora (para verificar si la cifra mínima anterior es acorde al perfil de riesgos de la entidad) y disciplina de mercado (reflejada en una exigencia superior de información a suministrar al mercado para fortalecer la transparencia).

En este período superior a 5 años se sucedieron opiniones y consultas del Comité de Basilea con la industria bancaria y los reguladores, así como estudios de impacto para conocer la situación de las entidades afectadas (en principio, las de actividad internacional) y la posible incidencia del cambio normativo.

Partiendo de la aceptación general, el grado de aprobación y previsible aplicación del Acuerdo ha diferido en los distintos países. En la Unión Europea, por ejemplo, la Directiva de Adecuación de Capital (actual 2006/49/EC) se transformó prácticamente de forma simultánea a la publicación de Basilea II, emitiéndose una primera versión de ésta en julio de 2004. Su final transposición a las normativas nacionales ha sido, en el caso español, reflejada en la reciente nueva Circular de Solvencia (CBE/3/2008, de 22 de mayo) que prácticamente reproduce el esquema de Basilea II.

En otros entornos económicos, las modificaciones encaminadas al Nuevo Acuerdo han sido más lentas, debido a las iniciales reticencias, que han ido salvándose con posterioridad, la diferente situación de partida y las características del sector bancario, con menor peso de la actividad internacional.

La evaluación del Acuerdo no ha olvidado, sin embargo, sus posibles inconvenientes y otros efectos, tanto en términos de discrecionalidad supervisora como de acentuación de los ciclos económicos. Desde la perspectiva de las finanzas internacionales, no se descarta que Basilea II dificulte el acceso a la financiación por parte de usuarios de menor dimensión (pequeñas empresas), ya que desincentivará los flujos de capital desde los países más desarrollados a las economías más emergentes. Esto previsiblemente beneficiará los intereses minoritarios, correspondientes a las entidades de mayor tamaño.

Se teme también que el Acuerde afecte negativamente a las posibilidades de crecimiento de esos países emergentes, al suponer una posible restricción o el encarecimiento de la financiación concedida, perjudicando a los grupos bancarios internacionales que se han instalado en estas zonas.

Por estos y otros motivos, es comprensible que en estos países se plantee si es conveniente apostar por Basilea II frente a la permanencia en el esquema de Basilea I (Liebig *et al*., 2007). Para ello, deberán valorarse aspectos como el grado de cumplimiento del pilar 2, la presencia de las agencias de calificación en ese país (si no es considerable, se dificultará la aplicación de varias metodologías de medición de riesgos y cálculo de capital), el nivel de fondos propios necesario en ese momento y la capacidad de las entidades y el supervisor para afrontar los cambios (Powell, 2004).

Dentro del conjunto de efectos que podrían presentarse en países emergentes, deben considerarse tanto los de carácter positivo como negativo (García-Herrero y Gavilá, 2006):

a) Efectos negativos:

- Los modelos de riesgo de Basilea II se han calibrado a partir de datos de bancos del G-10, cuyas calificaciones crediticias (aplicadas al riesgo de crédito) son superiores a las de los países emergentes.
- Si el Acuerdo afecta a la prociclicidad dificultando el acceso al crédito, las consecuencias se dejarán notar.
- La implantación del Acuerdo está sujeta a notable discrecionalidad. Ese trato selectivo puede minar los beneficios que se esperan de la homogeneización de la regulación bancaria a nivel mundial, particularmente necesaria en la actualidad.
- El Acuerdo no recoge los beneficios de la diversificación que aparecen al financiar a estas economías, por lo que las inversiones pueden huir de estos países.

b) Efectos positivos:

- El principal y más significativo es la mejora en la medición y gestión de riesgos, además de la fortaleza de la supervisión bancaria como garantía de seguridad del sistema financiero.
- Previsiblemente, el balance bancario de estas entidades mejorará por la desincentivación del préstamo a corto plazo y la ampliación de los activos que pueden emplearse como colaterales.

Además de estos aspectos generales, interesa evaluar también qué significa este Acuerdo respecto a su precedente. Así, parece que Basilea II conducirá a superiores niveles de exigencia de capital si se aplican los modelos estándar de riesgos; por el contrario, esos fondos propios disminuirán si se aplican modelos internos. Esta menor financiación propia es especialmente atractiva para las entidades y un incentivo para desarrollar metodologías propias, pese a que exijan notables esfuerzos técnicos y humanos.

Los resultados serán, no obstante, muy diferentes según el país emergente del que se trate. Por ejemplo, es posible que para los países emergentes que no pertenecen a la OCDE (como China) el coste de financiación incluso disminuya.

En definitiva, al margen de otros aspectos de difícil anticipación, China debe emprender el camino hacia Basilea II dado que promoverá una gestión bancaria efectiva, con una adecuada regulación y asignación de capital, apoyada por una fuerte supervisión bancaria (Okazaki, 2007). Sin embargo, es preciso ser consciente de que Basilea II no es la panacea en el

tratamiento de riesgos y vigilancia del sistema bancario, ya que puede resultar sólo parcialmente más sensible a los verdaderos riesgos de cada entidad, en tanto genere otras divergencias que puedan ir más allá de los efectos micro hacia el sistema en conjunto.

4.2.- Regulación bancaria en china: la postura frente a basilea ii

China es uno de los países que inicialmente puso más reparos al Acuerdo. En el momento en que se emprendió el proceso de reforma de Basilea, China no había entrado aún en la OMC y la apertura de su sector bancario era casi inexistente[18]. Desde ese momento, partiendo de ratios de solvencia que no alcanzaban el 8% exigido internacionalmente y fiel reflejo de un sector considerado técnicamente insolvente, la CBRC inició un acercamiento a los estándares internacionales en busca del cumplimiento de, al menos, los Principios básicos de Basilea emanados de Basilea I, reflejados en la publicación de la *Ley para la Regulación y Supervisión Bancaria* de 2003, en vigor a partir de febrero de 2004. En paralelo, China adaptó sus estructuras paulatinamente hacia la convergencia con Basilea I, emitiendo también en 2003 el documento Regulación de la Adecuación de Capital en la Banca Comercial, más conocido como *Nuevas reglas de capital*, aplicable desde abril de 2004.

Pese a estos cambios normativos, en el proceso de discusión de Basilea II China se mantuvo parcialmente al margen, manifestando incluso su disconformidad con ciertas partes del Acuerdo. Así, en los comentarios sobre el segundo documento de consulta (BCBS, 2001), China rechazó el tratamiento del riesgo de crédito, incluso desde la perspectiva de los modelos estándar, exigiendo se facilitara un enfoque aún más sencillo y dirigido a bancos más pequeños.

Posteriormente, China (incluido Hong Kong) participó junto con la República de Corea, como parte de la región del este asiático, en el tercer estudio de impacto de Basilea (BCBS, 2003). Sin embargo, con inmediata posterioridad a dicho estudio, rechazó el Acuerdo formalmente, junto con India, argumentando que no consideraba las particularidades de los países en desarrollo y manifestando que sólo adoptaría las reglas concernientes a supervisión y suministro de información; en otras palabras, los pilares 2 y 3 (Anderson, 2004).

Publicado el tercer documento de consulta (BCBS, 2003b), China se reafirmó en su postura, indicando que seguiría los principios de Basilea I al menos hasta 2008. Asimismo insistió en la necesidad de permitir mayor flexibilidad y discreción nacional para recoger las condiciones de los países que no forman parten del G-10. Adicionalmente, China pidió al BCBS la elaboración de una guía de referencia para el desarrollo de un procedimiento de modelos internos de riesgo crediticio en bancos de mercados emergentes y economías menos desarrolladas.

En cuanto a las primeras normativas de regulación y supervisión bancaria de 2003-2004 antes citadas y con relación a Basilea II, la CBRC ya reconocía en 2003 que no sería posible aplicar en los próximos años los modelos internos incluidos en el pilar 1 sobre riesgo de crédito. Más aún, incluso la metodología estándar para este riesgo (relativa continuación del tratamiento de esta categoría según Basilea I) ofrecía dificultades, dada su dependencia del rating crediticio[19] asignado por una agencia de calificación (institución externa de evaluación de crédito), considerando las escasas empresas chinas evaluadas en ese momento. De hecho, en julio de 2003, Standard & Poor´s sólo había asignado rating a 7 bancos y 12 empresas chinas. Además, la CBRC consideraba que, de aplicarse el método estándar, el capital medio mantenido por las entidades financieras chinas se elevaría de forma muy significativa por encima de las cifras actuales.

Respecto a la convergencia entre la normativa de regulación y supervisión china y Basilea II en ese momento, las mayores coincidencias se encuentran en los pilares 2 y 3. El pilar 1, por el contrario, plantea diversos obstáculos.

En el ***pilar 1***, las ponderaciones de riesgo de crédito para algunas categorías no coinciden. Así, según Basilea II ciertas exposiciones catalogadas de alto riesgo pueden requerir un capital

[18] El análisis la eficiencia del sector, los efectos de la reforma bancaria y la prudencia con que actúan las entidades pueden consultarse, entre otros, en Horrocks *et al*. (2008), Fu y Heffernan (2009), Jia (2009) y Berger *et al*. (2009).

[19] Calificación crediticia normalizada según los criterios de una agencia de evaluación, que se mide en una escala numérica o simbólica y que indica el grado relativo de riesgo asociado a la inversión, para permitir su comparación de forma homogénea. En nuestro caso, permite clasificar a los clientes de una entidad financiera según su calidad crediticia, es decir, la probabilidad de su incumplimiento o impago.

del 150%, porcentaje no considerado en la normativa china. Esto, unido a otras diferencias, seguramente derive en requisitos de capital regulatorio inferiores en China en comparación con otros países. En el caso del riesgo de mercado, frente a la alternativa del método estándar o los modelos internos (basados fundamentalmente en el cálculo del valor-en-riesgo[20]) que ofrece el Acuerdo, China sólo considera el primer esquema. Por último, el riesgo operacional no forma parte inicialmente de esta regulación.

En cuanto al paralelismo con el futuro ***pilar 2***, la normativa china de 2004 se muestra de acuerdo en la posibilidad de exigir un ratio de adecuación de capital superior a aquellas entidades cuyo perfil de riesgo y capacidad de gestión del mismo así lo requieran. De hecho, las *Nuevas Reglas de Capital* dividen a los bancos en tres categorías según la revisión supervisora: adecuadamente capitalizados, sub-capitalizados o significativamente sub-capitalizados.

Finalmente, las similitudes también son notables respecto al ***pilar 3***, indicando la CBRC que los bancos comerciales deben revelar información relativa a las áreas de i) objetivos y política de gestión de riesgos, ii) ámbito de aplicación, iii) capital, iv) ratios de adecuación de capital y v) riesgos de crédito y mercado.

A la vista de todo lo anterior, parece que la premisa de los cambios regulatorios de 2003-2004 es mantener a China en los niveles de Basilea I, con la incorporación de los pilares 2 y 3 de Basilea II, creando así un enfoque mixto entre ambos acuerdos (Brehm y Macht, 2004). De hecho, en ese momento en que lograr un coeficiente de solvencia del 8% no estaba al alcance de muchas entidades, el objetivo era mejorar los requisitos a nivel de regulación "local" para lograr una paulatina aproximación a los estándares internacionales.

Dado que el año 2003 marca un punto y aparte en la regulación bancaria, es interesante observar la evolución de las cifras de impagados y de las coberturas de riesgos en bancos comerciales desde esa fecha hasta la actualidad (ver tabla nº 2). En el primer caso, la importancia de los préstamos impagados en el balance ha ido disminuyendo desde un 17,9% en 2003 hasta el 6,7% en 2007; más significativo resulta, si cabe, que dentro de ellos, el porcentaje de fallidos cayó del 5,7 al 3,3%. Simultáneamente, el volumen de provisiones para la cobertura de estas operaciones se ha elevado hasta duplicar las cifras del inicio (del 19,7 al 39,2%). Por último, el número de bancos que han verificado los estándares de capital ha ido creciendo, hasta representar un 79% de los activos bancarios del sistema en 2007.

TABLA Nº 2: Préstamos impagados y provisiones. Bancos comerciales 2003-2007

	Total Activos (100 millones RMB)					Porcentaje				
	2003	2004	2005	2006	2007	2003	2004	2005	2006	2007
Préstamos impagados en balance	21.004,6	17.156,6	12.196,9	11.703,0	12.009,9	17,9	13,2	8,9	7,5	6,7
Substándar	3.201,1	3.074,7	2.949,6	2.270,7	1.844,3	2,7	2,4	2,2	1,5	1,0
Dudosos	11.130,7	8.899,3	4.609,0	4.850,3	4.357,5	9,4	6,8	3,4	3,1	2,4
Fallidos	6.712,8	5.201,6	4.638,4	4.581,9	5.808,1	5,7	4,0	3,4	2,9	3,3
Provisiones/Impagados						19,7	14,2	24,8	34,3	39,2
Bancos que cumplen requisitos de capital (% activos bancarios)						0,6	47,5	75,1	77,4	79,0

Fuente: CBRC (2008), Annual Report 2007.

4.3.- Cambios regulatorios en china para la adaptación a basilea ii

El inicial rechazo de Basilea II, seguido del posterior inicio de la transformación supervisora china, ha culminado en el reciente intento por avanzar a grandes pasos hacia el cumplimiento de las normas internacionales de solvencia.

[20] Pérdida potencial máxima, medida en términos monetarios que, dentro de un intervalo definido de probabilidad, puede experimentar una posición en un período de tiempo determinado, por los movimientos adversos del mercado. Puede entenderse también como la cantidad de fondos propios necesarios para absorber las pérdidas de la cartera. A modo de ejemplo, si se establece que el valor-en-riesgo de una cartera es de 1.000 € con un 99% de confianza, significa que sólo 1 de cada 100 días la pérdida que sufriría esta cartera sería igual o mayor a 1.000 €.

El actual tratamiento de la solvencia parte de las normativas de la CRBC, iniciadas en 2004 y revisadas en 2007, reflejadas en *Normas sobre la suficiencia de capital de los bancos comerciales*, basadas en Basilea I. Últimamente se ha decidido abordar progresivamente los cambios introducidos por Basilea II, distinguiendo dos grandes grupos de bancos comerciales: aquellos sujetos al Nuevo Acuerdo (grandes entidades con actividad en mercados financieros de otros países o regiones, cuya presencia en el exterior es parte importante de su actividad comercial total) y el resto. Los primeros empezarán a aplicar Basilea II a finales de 2010 y nunca después de finales de 2013. Para los otros, Basilea II es un Acuerdo meramente voluntario, manteniéndose pues las reglas de adecuación de capital actuales.

En esta línea, en febrero de 2007 la CRBC publicó el documento *Guía sobre la implantación del Nuevo Acuerdo de Capitales en la industria bancaria China*, que propone una implementación gradual de Basilea II, basada en el caso particular chino. En términos generales, la guía reconoce que los bancos deben asumir requisitos de capital distintos de acuerdo a su tamaño, que el proceso debe ser paulatino, y que se les permitirá afrontar los estándares de Basilea II paso a paso (CBRC, 2008).

Además de la toma de conciencia de la necesidad de aplicar, si bien con cierto retraso, Basilea II, el tratamiento de los riesgos se ha ido considerando lentamente en los últimos años. En concreto, las últimas normativas elaboradas son las siguientes:

a) **Riesgo de crédito**

En julio de 2007 la CBRC revisó la Guía de clasificación de préstamos según riesgo, y emitió la *Guía provisional de clasificación de préstamos por riesgo*, según la cual:

- Se aplica el sistema a las cooperativas rurales y a las cooperativas de crédito rural.
- Se amplía la lista de principios para clasificación de riesgos, enfatizando la exactitud, oportunidad, importancia y prudencia, para reflejar así los riesgos.
- Se añade un método más (fuera de período) para prestamistas familiares rurales y pequeñas empresas.
- Se consideran los riesgos inherentes a los préstamos, además de la capacidad de repago de los prestamistas.

También se publicó la *Guía provisional de clasificación de préstamos a la pequeña empresa según riesgo*.

En cuanto a supervisión, se revisó la *Guía sobre gestión del riesgo de créditos de la banca comercial a grupos de clientes*, para reducir el ratio de crédito a un solo grupo de clientes. En agosto se publicó la *Nota sobre refuerzo de vigilancia de los préstamos no-rentables de gran valor*. Además, se pretende avanzar en las prácticas de gestión de riesgos de crédito. Así, en febrero de 2007 se publica *Reglas sobre supervisión y evaluación de activos no-rentables de cooperativas rurales de crédito*, y en agosto *Guía sobre el negocio del préstamo sindicado*.

b) **Riesgo de mercado**

La normativa de 2007 insta a los bancos a intensificar su infraestructura de gestión de estas incertidumbres, con una adecuada implantación en los bancos comerciales, siendo preciso disponer de un departamento de gestión de riesgos de mercado. Además, se pretende establecer un nivel de referencia para su medición, exigiendo que las entidades monitoricen los movimientos de los precios de los activos, tipos de cambio, tipos de interés, etc.

c) **Riesgo operacional**

Acorde a su definición más global, en mayo de 2007 se emite la *Guía sobre riesgo operacional*, para reforzar su control y promover la adopción de las prácticas que se están llevando a cabo a nivel internacional. Simultáneamente, se publican otros documentos sobre expedición de tarjetas de crédito y el negocio de la banca telefónica. También se ha incidido en reglas para el control de la banca electrónica, servicio cuya prestación requiere de un sólido sistema de control interno y gestión del riesgo. Se pretende pues controlar una de las principales

causas de riesgo operacional. Por tanto, podemos afirmar que en la regulación china el tratamiento del riesgo operacional se ha centrado especialmente en las tecnologías de la información y en casos fraudulentos.

La postura actual de la regulación china sobre los métodos de medición de riesgos incluidos en Basilea II es la siguiente:

- Riesgo de crédito: la CRBS anima a la adopción por parte de los bancos de modelos internos o IRB[21].
- Riesgo de mercado: los bancos deberían adoptar modelos internos. Si una entidad no cumple los requisitos para su aplicación, puede elegir la metodología estándar.
- Riesgo operacional: debería usarse un enfoque de medición de este riesgo.

Adicionalmente, convendría establecer un sistema seguro de rating interno y mecanismos de revelación de información. Todo ello sin olvidar los efectos sobre un posible riesgo de liquidez, que aumentó en 2007 tanto por al continuo incremento del precio de los activos y la subida de los requisitos de reservas (que llevaron a depósitos a más corto plazo), como por los préstamos a largo plazo.

A finales de 2008, la CBRC ha emitido un primer paquete de regulaciones supervisoras bajo Basilea II, que incluyen los modelos IRB y otros estándares técnicos. El objetivo es que a finales de 2010 los bancos nacionales puedan operar bajo el enfoque de Basilea (previa solicitud a la CBRC al menos 6 meses antes), aunque en algunos casos esta opción se ha adelantado a octubre del pasado año, superando incluso a la banca estadounidense. Otras entidades pueden solicitar su aplicación a partir de 2011.

En cuanto al calendario establecido, la CBRC espera publicar sucesivas normativas referentes a la transposición de Basilea II en el período 2008-2009, llevando a cabo un estudio de impacto.

5.- CONCLUSIONES

El sistema financiero ha sido uno de los sectores de último desarrollo en la economía china. Así, la industria bancaria, tradicionalmente muy protegida y sobreregulada, ha estado dominada por la propiedad estatal, en un esquema inicial monobancario que comenzó a reestructurarse en los años 80 en medio de las reformas financieras. Por eso, en los años recientes se han producido importantes cambios tendentes a reducir el intervencionismo estatal, controlar el volumen de préstamos impagados e introducir competencia exterior, sobre todo como consecuencia de los compromisos derivados de la adhesión de China a la OMC.

Este proceso de cambio requiere aún importantes acciones para mejorar la actividad financiera en China, como la entrada de más banca extranjera, la inversión en mejores servicios a los clientes, favorecer la actividad de la banca regional y analizar las oportunidades para competir con entidades internacionales.

En este contexto, China se ve afectada también por los cambios en la regulación y supervisión bancaria a nivel internacional, reflejados actualmente en el Acuerdo de Basilea II publicado en 2004, que tiene por objeto reforzar la solvencia de las entidades financieras y favorecer un clima de cooperación y armonización de la supervisión bancaria. Es por ello que desde el 2003 se han llevado a cabo significativos cambios regulatorios en la normativa bancaria china, emanada de la CBRC, en un intento por adaptar los esquemas del país a las exigencias de Basilea II.

El impacto de Basilea II sobre los países emergentes es aún incierto. Analizando su predecesor, Basilea I tuvo consecuencias positivas generales sobre la estabilidad de los sistemas bancarios, al introducir mecanismos de control y vigilancia. No existe consenso, sin embargo, sobre sus implicaciones hacia un aumento o restricción del crédito, o lo que es lo mismo, sus efectos procíclicos.

[21] Modelos internos o *Internal Rating Based*, en los que es la propia entidad quien determina la probabilidad de fallido de la contraparte. Según el nivel de desarrollo de este esquema, el resto de datos del análisis podrá calcularse también internamente o vendrá dado por el supervisor. Para la aplicación de este esquema, es precisa la previa evaluación y aprobación del supervisor. Lo mismo ocurre en caso de aplicación de modelos internos para riesgos de mercado y/u operacionales.

Basilea II adolece para estos países del inconveniente previo que se deriva de la no participación de éstos en el Comité que elabora el Acuerdo. Además, los modelos de cálculo de riesgos se han previsto para entidades correspondientes a sistemas bancarios más desarrollados. Pese a ello, no cabe duda de que la mera coordinación entre supervisores que se pretende a partir de ahora debe tener un impacto favorable también en estos entornos emergentes, gracias a un efecto aprendizaje de los métodos que ya se están desarrollando a nivel internacional.

Con todo esto, China es consciente de las consecuencias que Basilea II puede tener para su sistema financiero. Dada su estructura, los niveles de capital requeridos por riesgo de crédito aumentarán de forma considerable, tanto por los bajos niveles de partida como porque el Acuerdo considera categorías de alto riesgo. Pese a estos y otros obstáculos de adaptación, China apoya firmemente la filosofía que subyace en los pilares 2 y 3, y coincide en la necesidad de mejorar la gestión de riesgos. Sus esfuerzos por modernizar estos procesos de control apoyarán el proceso de reforma, aperturismo y mejora de la eficiencia y competitividad de la banca china.

BIBLIOGRAFÍA

ALLEN, F., QIAN, J., QIAN, M. (2005): "Law, finance and economic growth in China", Journal of Financial Economics, Vol 77, No 1, pp. 57-116.

ANDERSON, L. (2004), "Basel II: the path to promoting financial stability in the Asian and Pacific region?" en Bulletin on Asia-Pacific Perspectives 2003/2004, Development research and policy analysis division, ESCAP.

BANCO DE ESPAÑA (2008), Circular 3/2008, de 22 de mayo, del Banco de España, a entidades de crédito, sobre determinación y control de los recursos propios mínimos.

BASEL COMMITTEE ON BANKING SUPERVISION (BCBS) (1988), "International convergence of capital measurement and capital standards", Bank for International Settlements, July, Basel.

BASEL COMMITTEE ON BANKING SUPERVISION (1996), "Amendment to the capital accord to incorporate market risks", Bank for International Settlements, January, Basel.

BASEL COMMITTEE ON BANKING SUPERVISION (1999), "A new capital adequacy framework. Consultative paper", Bank for International Settlements, June, Basel

BASEL COMMITTEE ON BANKING SUPERVISION (2001), "Consultative document. Operational risk. Supporting document to the New Basel Capital Accord", Bank for International Settlements, January, Basel.

BASEL COMMITTEE ON BANKING SUPERVISION (2003A), "The loss data collection exercise for operational risk: survey of the data collected", Risk Management Group of the BCBS (BIS), March.

BASEL COMMITTEE ON BANKING SUPERVISION (2003B), "The New Basel Capital Accord. Consultative Document", Bank for International Settlements, April, Basel.

BASEL COMMITTEE ON BANKING SUPERVISION (2004), "International convergence of capital measurement and capital standards: a revised framework", Bank for International Settlements, June, Basel.

BERGER, A.N., HASAN, I., ZHOU, M. (2009), "Bank ownership and efficiency in China: what will happen in the world´s largest nation?", Journal of Banking & Finance, Vol. 33, pp. 113-130.

BREHM, S., MACHT, C. (2004), "Banking supervision in China: Basel I, Basel II and the Basel Core Principles", Banking Supervision.

BUSTELO, P. (1996), "Progress without Privatization in China: The Reform of State-owned Enterprises in the 1990s", Comunicación presentada en la 11ª Conferencia de la European Association of Chinese Studies, Barcelona, septiembre de 1996. (Versión en castellano en Revista de Estudios Asiáticos (1996), pp. 117-127).

BUSTELO, P., FERNÁNDEZ LOMMEN, Y. (1996), La economía china hacia el siglo XXI. Veinte años de reforma. Editorial Síntesis, Madrid.

CHANG, H.–J, NOLAN, P (1995), "Europe versus Asia: Contrasting Paths to the Reform of Centrally Planned Systems of Political Economy", The transformation of the Communist Economies against the Mainstream, pp. 3-45. (Versión en castellano en Revista de Estudios Asiáticos (1996), pp. 11-34).

CHINA BANKING REGULATORY COMMISSION (CBRC) (2008), Annual Report 2007. URL: http://www.cbrc.gov.cn/english/home/jsp/index.jsp

DIRECTIVA 2006/49/CE DEL PARLAMENTO EUROPEO y del Consejo, 14 de junio de 2006, sobre la adecuación del capital de las empresas de inversión y las entidades de crédito.

FU, X., HEFFERNAN, S. (2009), "The effects of reform on China´s bank structure and performance", Journal of Banking & Finance, Vol. 33, pp. 39-52.
FUNG, H-G., PEI, C., ZHANG K.H. (EDS.) (2006), China and the Challnge of Economic Globalization: the impact of WTO membership. M.E.Sharpe, New York.USA.
GALE, F., COLLENDER R. (2006), "New directions in China´s agricultural lending". USDA WRS-06-01. www.ers.usda.gov/publications/WRS0601/WRS0601.pdf
GARCÍA-HERRERO, A., GAVILÁ, S. (2006), "Posible impacto de Basilea II en los países emergentes", Documentos ocasionales del Banco de España nº 0606.
GE, W. (1999), "Special Economic Zones and the Opening of the Chinese Economy: Some Lessons for Economic Liberation". World Development, Vol. 27, No. 7, pp. 1267-1285.
HE, L. (2005), "Evolution of Financial Institutions in Post-1978 China: Interaction between the State and Market", China & Word Economy, Vol 13, No. 6, pp. 10-26.
HORROCKS, M.; RAMASWAMY, V.; RUPP, K. (2008), "Financial reforms in Chine banking: the impact on personal lending and operational efficiency", Business Horizons, No. 51, pp. 511-517.
JIA, C. (2009), "The effect of ownership on the prudential behaviour of banks – The case of China", Journal of Banking & Finance, Vol. 33, pp. 77-87.
LAU, L.J., QIAN, Y., ROLAND, G. (2000), "Reform without Losers: An Interpretation of China's Dual-Track Approach to Transition" Journal of Political Economy, Vol. 108, No. 1.
LIAO, M (2006), "Institutional Reform in the Chinese Banking System and China´s Implementation of Commitments to the WTO". En FUNG, H-G., PEI C. y ZHANG K.H. (eds.) (2006), pp. 207-234.
LIEBIG, T., PORATH, D, WEBER, B., WEDOW, M. (2007), "Basel II and bank lending to emerging markets: evidence from the German banking sector", Journal of banking & finance, Vol. 31, pp. 401-418.
LU, S., YAO, Y. (2009), "The effectiveness of Law, Financial Development, and Economic Growth in an Economy of Financial Repression: Evidence from China", Word Development, Vol. 37, No. 4, pp. 763-777.
MATTO, A. (2004), "The services dimension of China´s accession to the WTO". En Bhattasali, D., Li, S., Martin, W. (eds.) (2004), pp. 117-140.
NATIONAL DEVELOPMENT AND REFORM COMMISSION. URL: http://en.ndrc.gov.cn/
NAUGHTON, B. (2004), "La transformación de China" Economía Exterior, No. 30, pp. 120-143.
NAUGHTON, B. (2007), The Chinese Economy: Transitions and Growth. The MIT Press, Cambridge, MA.
NOLAN; P (1994), "The China Puzzle: Touching Stones to Cross the River". Challenge, January – February, pp. 25-31.
OKAZAKI, K. (2007), "Banking system reform in China. The challenges of moving toward a market-oriented economy", Occasional Paper, National Security Research Division.
ORGANIZACIÓN MUNDIAL DE COMERCIO. URL: http://www.wto.org.
ORGANIZACIÓN MUNDIAL DE COMERCIO (2008). Documentos WT/TPR/S/199 y WT/TPR/G/199.
ORGANIZACIÓN MUNDIAL DE COMERCIO. Documento WT/ACC/CHN/49/Add.2.
ORGANIZACIÓN MUNDIAL DE COMERCIO. Protocolo de Adhesión de la República Popular China. WT/L/432, de 23 de noviembre de 2001.
PERKINS, D. H. (1997), "Have China´s Economic Reforms Stalled?" Development Discussion Paper, No. 613. Harvard Institute for International Development.
PODPIERA, R. (2006), Progress in china's banking sector reform: Has bank behaviour changed?, IMF Working Paper WP/06/71.
POWELL, A. (2004), "Implicaciones de Basilea II para América Latina", Revista de Estabilidad Financiera, No. 6, pp. 149-172.
SANTABÁRBARA, D. (2007), "El proceso de apertura el sector bancario chino y el papel de la banca extranjera Situación y perspectivas", Revista de Estabilidad Financiera, No 13. pp. 103-122.
YAO, Y., YUEH, L. (2009), "Law, Finance, and Economic Growth in China: An Introduction". Word Development, Vol 37, No. 4, pp. 753-762.

¿SE DESTINA LA AYUDA OFICIAL AL DESARROLLO EN MATERIA DE SERVICIOS SOCIALES BÁSICOS A AQUELLOS PAÍSES QUE MÁS LA NECESITAN?

YOLANDA MUÑOZ OCAÑA
e-mail: ymuñoz@etea.com
MERCEDES TORRES JIMÉNEZ
e-mail: mtorres@etea.com
Departamento de Métodos Cuantitativos y Gestión empresarial
Facultad de Ciencias Económicas y Empresariales-ETEA
Institución Universitaria de la Compañía de Jesús adscrita a la Universidad de Córdoba

Resumen

La comunicación propuesta analizará la orientación de la ayuda oficial al desarrollo de los países y organismos multilaterales donantes del Comité de Ayuda al Desarrollo (CAD) a Servicios Sociales Básicos (SSB) durante los últimos años (2000-2007). Tras una breve introducción sobre los principales foros internacionales así como de los compromisos adquiridos por los países donantes y socios en esta materia, se efectuará un análisis estadístico descriptivo de la ayuda (su evolución temporal, identificación de los principales donantes internacionales, etc.). Para finalizar se analizará si la ayuda a SSB va destinada realmente a aquellos países en desarrollo que más la necesitan, seleccionando algunos indicadores que hacen referencia al desarrollo logrado en la cobertura de los servicios sociales básicos por parte de los países receptores de la ayuda. De este modo podremos comprobar si la "necesidad" es el criterio a seguir en la asignación de la ayuda o si prevalecen otros criterios.

Palabras Clave: Servicios sociales básicos, Desarrollo, Ayuda oficial al desarrollo, Orientación de la ayuda.
Área Temática: Economía Internacional y de la Integración.

Abstract

This study analyses the orientation of Official Development Aid in Basic Social Services (BSS) during the last few years. It begins by analyzing the main international forums and commitments adopted for these ends by different countries (donating and receiving), based on the quantitative evolution of this aid by countries, sectors…Finally through the use of some development indicators related to basic social services, it will be analyzed if the countries more needed in SSB are those that more aid receive, or if others criterions are considered in the aid assignation. .

Thematic Area: International and Integration Economics.

INTRODUCCIÓN

Se puede definir el desarrollo humano como el proceso de ampliar las opciones a las personas. El objetivo del desarrollo debe ser permitir a las personas disfrutar vidas prolongadas, saludables y creativas. El principio básico que guíe la cooperación debe ser siempre colocar a personas en el centro del desarrollo y concentrarse en sus necesidades y en su potencial.

El análisis de las políticas de desarrollo y los resultados obtenidos a lo largo de la última mitad de siglo ponen de manifiesto que, en muchas ocasiones, un aumento del Producto Interior Bruto sólo va acompañado de una distancia mayor entre los ricos y los pobres, ya que la distribución de este crecimiento suele ser desigual, sin mencionar los desequilibrios medioambientales así como el agotamiento progresivo de cada vez más recursos.

En esta línea, los objetivos del milenio reflejan cómo la dimensión de la pobreza no es únicamente el bajo ingreso, sino que forma parte de ella el analfabetismo, la mala salud, la desigualdad de género… pero mientras que aparecen pequeños avances en el crecimiento económico, estos no se traducen en avances en desarrollo humano, ya que existen grandes diferencias en salud y educación entre los ricos y los pobres de cualquier país del mundo, lo que pone de manifiesto que el crecimiento económico no es suficiente para alcanzar las cotas deseadas en educación y salud, lo que exigirá más recursos y un uso más eficaz de dichos recursos, lo que se traducirá en mejores servicios (agua, saneamiento, energía, transporte, salud y educación) que contribuyan a elevar el nivel educativo y de salud de los distintos pueblos que habitan la tierra. Estos servicios no fallan en general, sino que suele faltar a los pobres, ya que además de que son claramente insuficientes, son de baja calidad, lo que se manifiesta en un alto grado de absentismo laboral por parte de los profesionales de la salud y de la educación.

En la actualidad la medición del Desarrollo Humano se realiza a través del Indicador de Desarrollo Humano (IDH), que aunque no exento de carencias, permite medir, evaluar y comparar. Es un índice compuesto a partir de indicadores parciales de longevidad (esperanza de vida al nacer, en función de la nutrición y la salud), educación (porcentaje de población escolarizada y promedio de años de escolarización) y el ingreso real per cápita. Desde esta visión del desarrollo se puede afirmar que con políticas adecuadas se puede conseguir un desarrollo humano razonable con moderados niveles de PIB per cápita.

La medición del desarrollo humano con estos nuevos criterios revela datos más reales de las distintas zonas del planeta, aunque el concepto de desarrollo es infinitamente más amplio que lo recogido por el índice. El desarrollo humano abarca todo el espectro de las necesidades y ambiciones humanas, por tanto, es un concepto amplio e integral, y aunque es cierto que esta forma de medir el desarrollo posee muchas limitaciones, es fácil advertir que en este desarrollo juegan un papel fundamental la cobertura de las necesidades sociales básicas.

A pesar de que con el tiempo el IDH se ha ido perfeccionando, existen aspectos como la libertad, que son difíciles de cuantificar. Pero sin duda, la principal carencia del índice es la ausencia de medición de las sostenibilidad del desarrollo, porque no basta con garantizar la satisfacción de las necesidades básicas en la actualidad sino que es necesario garantizar el acceso a dichos bienes a las generaciones venideras.

Una de las principales amenazas contra el desarrollo humano y económico sostenible proviene de la espiral descendente de pobreza y degradación ambiental que amenaza a la generación actual y a las futuras. Los peligros ambientales y riesgos contra la salud planteados por la contaminación, la vivienda inadecuada, las condiciones sanitarias deficientes, el agua contaminada y la falta de otros servicios básicos, constituyen una amenaza desproporcionada para los pobres. Muchas de estas personas, ya desprovistas de necesidades básicas, también viven en las áreas más vulnerables desde el punto de vista ecológico.

Además del IDH existen otros indicadores globales y gran número de indicadores parciales que evalúan la situación de cada país respecto a distintos aspectos del desarrollo.

Esta comunicación comienza realizando un recorrido por las principales cumbres y foros internacionales que se han celebrado en materia de Necesidades Sociales Básicas (NSB). A continuación se efectúa un análisis cuantitativo de la Ayuda Oficial al Desarrollo (AOD) de los donantes internacionales (bilaterales y multilaterales) para Servicios Sociales Básicos (SSB) durante el período 2000 a 2007 (último año con datos disponibles). En dicho análisis se estudia la evolución temporal de la ayuda, los principales donantes internacionales, su distribución sectorial y geográfica. Para finalizar se analiza quienes son los principales países receptores de AOD para SSB y cual es su posición en los indicadores previamente seleccionados para medir su desarrollo en dicha materia.

PRINCIPALES EVENTOS Y ACUERDOS INTERNACIONALES EN COBERTURA DE NECESIDADES SOCIALES BÁSICAS.

En la Cumbre sobre Desarrollo Social, celebrada del 6 al 12 de Marzo de 1995 en Copenhague, los gobiernos alcanzaron un nuevo consenso sobre la necesidad de asignar a las personas el papel principal dentro del desarrollo

No es la primera vez que se hablaba de necesidades sociales básicas en el marco del desarrollo humano y la erradicación de la pobreza. Son muchas las voces que han dejado constancia, en su mayoría en sus escritos, de la necesidad de satisfacer las necesidades básicas para conseguir el desarrollo, aunque es verdad que estas voces difieren en cuanto a cuales son estas necesidades y el método para llevar a cabo dicha satisfacción. Incluso existen claras diferencias en la nomenclatura utilizada hablándose deÿÿecesidades bÿÿÿÿas, necesidades humanas, necesidades sociales y algún otro.

Es de resaltar que la Cumbre Social constituyó la mayor reunión de dirigentes mundiales hasta el momento. Allí expresaron la voluntad de considerar la erradicación de la pobreza, el objetivo del pleno empleo y el fomento de la integración social como las metas más importantes del desarrollo. Cinco años después, se reunieron de nuevo en Ginebra en junio del 2000 para evaluar lo que se había conseguido y se comprometieron a adoptar nuevas iniciativas.

En esta Cumbre, los jefes de Estado participantes adoptan una resolución por la que se pretende establecer un contrato entre países donantes y receptores por el cual, el 20% del gasto público del país receptor y el 20% de la ayuda al desarrollo del país donante irán destinados a servicios sociales básicos. Con esta iniciativa se pretende, por un lado, reconducir la cooperación internacional hacia la cobertura de necesidades sociales básicas, y por otro, conseguir la implicación del presupuesto de los países en desarrollo, con el fin de conseguir sinergias derivadas del objetivo común de donantes y donados.

Esta iniciativa es fruto del intento, por parte de la comunidad internacional, de poner al ser humano en el centro de las políticas de desarrollo, y conseguir, como objetivo amplio y ambicioso, no exento de muchas dificultades, pero en cualquier caso irrenunciable, erradicar la pobreza en el mundo que vivimos.

Aunque ya años antes muchas voces había clamado sobre la necesidad de conseguir erradicar la pobreza a través de la cobertura de las necesidades sociales básicas, son los Informes de Desarrollo Humano del Programa de Naciones Unidas para el Desarrollo (PNUD) de principios de la década de los noventa, los que sientan las bases de lo que sería el acuerdo final de la Cumbre de Copenhague. En el Informe de 1991 se hace una recomendación a los gobiernos para que dirijan la mitad de sus gastos sociales hacia las prioridades humanas, llegando incluso a concretar dichos gastos: educación básica, salud básica, acceso a agua potable, eliminación de la malnutrición y acceso a la planificación familiar. Puesto que el global de gastos sociales se estimaba en un 40% del presupuesto público, la recomendación suponía invertir, por lo menos, 20% del presupuesto público en servicios sociales básicos, lo que se concreta en el Informe del año siguiente (1992), en el que se recomienda comprometer al menos el 20% del presupuesto de la ayuda en prioridades humanas.

La iniciativa 20/20 pretende movilizar los recursos necesarios a nivel nacional para lograr los objetivos perseguidos internacionalmente, a las cuales se comprometió la comunidad mundial durante la década de los 90: en la Conferencia Mundial sobre Educación Básica para Todos (1990), en la Cumbre Mundial sobre la Infancia (1990), en la Conferencia de las Naciones Unidas sobre Medio Ambiente y Desarrollo (1992), la Conferencia sobre Derechos Humanos (1993), la Conferencia Internacional sobre Población y Desarrollo (ICDP), la Cumbre Mundial para el Desarrollo Social (1995), la Cuarta Conferencia Mundial sobre la Mujer (1995) y la Conferencia Habitat (1996), en las cuales se habían puesto de manifiesto la necesidad de alcanzar ciertos logros en materias sociales.

Las metas sociales formuladas en estas conferencias incluyen:

- ✓ Acceso universal y ciclo completo de la educación primaria.

- ✓ Reducción a la mitad de los niveles de analfabetismo de 1990, con énfasis en reducir la disparidad entre las tasa de analfabetismo para hombres y mujeres.
- ✓ Reducción de la malnutrición tanto severa como moderada, entre niños menores de 5 años, a la mitad de los niveles de 1990.
- ✓ Reducción a la mitad de la tasa de mortalidad materna de 1990.
- ✓ Reducción de la mortalidad de lactantes y menores de cinco años en un tercio en relación con su nivel de 1990, o hasta entre 50 y 70 por mil nacidos, si esta cifra es más baja.
- ✓ Atención a la salud reproductiva, incluyendo información y servicios de planificación familiar, al alcance de todos los individuos y parejas.
- ✓ Acceso universal al agua potable y saneamiento seguro.

A pesar de que la iniciativa 20/20 concreta la orientación de parte de la ayuda al desarrollo, su implementación no es fácil. Este motivo lleva a los representantes gubernamentales a tener un encuentro con el fin de establecer las concreciones prácticas de la iniciativa. Aunque no todos los países presentan el mismo grado de entusiasmo a la hora de acoger la iniciativa, ésta supone un punto de arranque en la movilización de recursos hacia un objetivo común, que, aunque no de forma directa ni exenta de controversias, se puede cuantificar en cuanto a su consecución y por tanto, evaluable en su implantación.

Este encuentro tiene lugar en Oslo, un año después de la Cumbre de Copenhague, en abril de 1996, con la presencia de 16 países donantes y de 22 países receptores, así como varios organismos de las Naciones Unidas, obteniendo como fruto el llamado "Consenso de Oslo sobre la iniciativa 20/20". En dicha reunión se especifica lo que se entiende por servicios sociales básicos, a saber: atención primaria de salud, programas de nutrición, población, salud reproductiva, salud ambiental con agua potable, saneamiento y educación básica, y también la capacidad institucional para proporcionar estos servicios. El acceso a estos servicios debe ser universal, y al mismo tiempo, dirigirse a los sectores más pobres y vulnerables.

No hubo acuerdo general en otros campos como la ayuda de emergencia o la vivienda, lo que ha conducido en algunos casos, a que la medición de la consecución de los objetivos por parte de los países donantes difiera en cuanto a las magnitudes que se incluyen. Por otro lado, la cifra 20/20 es una cifra estimada grosso modo, como orientación para los países, con el fin de aunar esfuerzos por parte de ambos lados de la ayuda y con el fin de ser prácticos. Si se consiguiera alcanzar las cifras, pero no se consigue el acceso universal a los servicios sociales básicos, no se habrá conseguido el fin de la iniciativa. Es más, aún cuando se consiguiera el acceso universal a servicios sociales básicos, el objetivo de erradicar la pobreza supone promover un desarrollo humano sostenible, por lo que la iniciativa 20/20 forma parte de un objetivo mucho más amplio que los países, tanto donantes como receptores, no pueden olvidar.

Con la iniciativa 20/20 se confirma la orientación de la política internacional hacia la erradicación de la pobreza, reconociendo que sólo se podrá conseguir este objetivo de forma duradera en el tiempo consiguiendo que los hombres y mujeres puedan desarrollar sus capacidades básicas, a través del acceso universal a los servicios básicos. Esta iniciativa ha tenido concreciones posteriores, como en el documento de la Comisión de Asistencia al Desarrollo (CAD) de mayo de 1996, titulado "Shaping the 21st Century: the Contribution of development cooperation", donde se enumeran una serie de metas para la inversión en desarrollo social a conseguir antes del año 2015, entre las cuales destacan: la reducción a la mitad de la proporción de personas que viven en extrema pobreza; la educación primaria universal, el progreso hacia la igualdad de género y empoderamiento de las mujeres; la reducción a dos tercios de las tasas de mortalidad para lactantes y niños menores de 5 años; la reducción a tres cuartos de mortalidad materna; acceso a servicios de salud reproductiva.

En 1998 se celebra la Conferencia de Hanoi, con una participación que supera a las del consenso de Oslo, con 31 países en desarrollo y 19 países donantes, bajo el auspicio de 10 organismos de las Naciones Unidas, el Banco Mundial y 9 organizaciones no gubernamentales representadas, aparte de las organizaciones no gubernamentales locales.

En Hanoi, tanto los países en desarrollo como los países donantes dialogan abiertamente sobre los esfuerzos realizados para cumplir con su parte en la iniciativa 20/20. También se

introdujeron las primeras estadísticas sobre los aspectos cualitativos y cuantitativos de los servicios sociales básicos. Se discutió de la iniciativa 20/20 no sólo de forma aislada, sino también en relación con otros ámbitos de la política: en primer lugar, se analizaron los aspectos negativos de la crisis económica y la disponibilidad de fondos para servicios sociales básicos; en segundo lugar, destaca la falta de progreso en la cancelación de la deuda, sobre todo para los PPME (Países Pobres Altamente Endeudados), iniciativa del Fondo Monetario Internacional y el Banco Mundial, representando un grave obstáculo para muchos de los países en desarrollo en sus esfuerzos por proporcionar servicios sociales básicos; en tercer lugar, el objetivo 20/20 debe estar vinculada a la estrategia de desarrollo orientada a la OCDE se describe en el informe Shaping the 21st Century y los resultados de las principales conferencias internacionales.

En Hanoi se logró un consenso sobre la iniciativa 20/20. Se produjo un avance significativo de la consideración teórica de la iniciativa a la fase operativa, logrando por primera vez estadísticas de los logros conseguidos en materia de necesidades sociales básicas. Las estadísticas no son sólo una fuente de información en términos absolutos, sino que también proporcionan la base para comparaciones entre sectores y entre países.

Cinco años más tarde de la Cumbre Mundial sobre el Desarrollo Social, se celebra en Ginebra en el año 2000 "Copenhague+5", un período extraordinario de sesiones de la Asamblea General de la ONU como seguimiento de la Cumbre (el vigésimo cuarto período extraordinario de sesiones de la Asamblea General), donde se reafirmó la importancia de crear sociedades más equitativas, justas desde el punto de vista social y centradas en el ser humano. También se constató que desde la celebración de la Cumbre de Copenhague se ha asignado más prioridad al desarrollo social en los objetivos de las políticas nacionales e internacionales, aunque no ha habido uniformidad en las medidas adoptadas.

En esta resolución se indica que la Declaración de Copenhague y el Programa de Acción seguirán siendo el marco básico para el desarrollo social en los próximos años, indicando asimismo que para alcanzar estos objetivos son fundamentales el acceso universal a una educación de gran calidad, acceso a los servicios de salud y otros servicios sociales básicos, y la igualdad de oportunidades para participar activamente en el proceso de desarrollo y recibir sus beneficios.

En la Cumbre se fijaron y se reafirmaron objetivos cuantitativos en la esfera de los servicios sociales básicos y la asistencia oficial al desarrollo. De trece objetivos, los nueve siguientes debían alcanzarse en el año 2000 como muy tarde: el fomento de la educación; la reducción de la tasa de analfabetismo en los adultos; la mejora del acceso al agua potable y al saneamiento; la reducción de la malnutrición en los niños menores de cinco años; el aumento de la esperanza de vida; la reducción de la mortalidad y la morbilidad debidas al paludismo y la mayor disponibilidad de viviendas económicas y adecuadas para todos.

Los datos obtenidos indican que los avances en esas esferas no eran satisfactorios. El documento indica que los progresos en la erradicación de la pobreza han sido desiguales, ya que en muchos países han aumentado desde 1995 las personas que viven en la pobreza. En muchos países en desarrollo han disminuido los créditos para servicios sociales y, debido a ello, muchas personas se han quedado sin acceso a servicios sociales básicos.

Del mismo modo, desde la perspectiva de género, tanto en la Cumbre como en la Cuarta Conferencia Mundial sobre la Mujer, celebrada en Beijing, la comunidad internacional reconoció expresamente que la mujer y el hombre vivía la pobreza de forma distinta y desigual y se empobrecían de manera diferente, y que es necesario que se tomaran en consideración esas diferencias para poder comprender las causas de la pobreza o para que el Estado pudiera adoptar medidas para eliminarlas.

Se acordaron nuevas iniciativas como la creación de un entorno económico, político, social, cultural y jurídico que permita lograr el desarrollo social; erradicar la pobreza en el mundo, fijando por primera vez un objetivo mundial de reducir a la mitad el número de personas que vive en la extrema pobreza, antes de fines de 2015; elaboración de una estrategia internacional coordinada sobre el empleo; promover la integración social fomentando sociedades estables, seguras y justas.

Debemos destacar el sexto compromiso, ya que hace referencia directa a la promoción y logro de objetivos de acceso universal y equitativo a una educación de calidad, el nivel más alto

posible de salud física y mental y el acceso de todas las personas a la atención primaria de la salud, reconociendo la responsabilidad de los gobiernos de prestar servicios sociales básicos para todos o, por lo menos, garantizar el acceso a ellos.

En el mismo año que se celebra "Copenhague+5", en el año 2000 se celebra la Cumbre del Milenio, quizá la cumbre más importante en la historia de la cooperación internacional al desarrollo, donde los189 países firmantes aprueban la Declaración del Milenio. Fruto de esta declaración surgen los conocidos por su enunciado (lejos de su consecución) Objetivos de desarrollo del Milenio, en los que tienen peso específico la cobertura de necesidades sociales básicas.

De hecho, de los 8 objetivos, 5 de ellos tienen relación directa con la cobertura de necesidades sociales básicas, a saber:

- Objetivo 1: Erradicar la pobreza extrema y el hambre
- Objetivo 2: Lograr la educación primaria universal
- Objetivo 4: Reducir la mortalidad de los niños menores de 5 años.
- Objetivo 5: Mejorar la salud materna
- Objetivo 6: Combatir el VIH/SIDA, el paludismo y otras enfermedades.

Cada uno de estos objetivos se concreta en unas metas a alcanzar, con unos plazos de ejecución, que según los seguimientos realizados, parecen cada vez más difíciles de conseguir.

En diciembre del año 2002, la Asamblea General de la ONU adopta una resolución por la que se insta a las instancias pertinentes, países y organismos a la "Aplicación de las decisiones adoptadas en la Cumbre Mundial sobre Desarrollo Social y el vigésimo cuarto período extraordinario de sesiones de la Asamblea General" y pide un informe sobre la cuestión, en que se tenga en cuenta la necesidad de un seguimiento integrado y coordinado de todas las principales conferencias, períodos extraordinarios de sesiones y cumbres de las Naciones Unidas.

ANÁLISIS DESCRIPTIVO DE LA AOD INTERNACIONAL DESTINADA A SSB EN LOS ÚLTIMOS AÑOS.

El análisis cuantitativo que se ha efectuado sobre la AOD internacional destinada a SSB se ha basado en la definición acordada por el "DAC Working Party on Statistics (WP-STAT)" el 14 de Junio de 1999, según la cual se incluirían los siguientes sectores como SSB:

- Educación básica (qué incluye educación primaria, capacitación básica de jóvenes y adultos y educación primera infancia);
- Salud básica (que incluye atención sanitaria básica, infraestructura sanitaria básica, nutrición básica, control de enfermedades infecciosas, educación sanitaria y formación del personal sanitario);
- Políticas y programas de población y salud reproductiva (que incluye política sobre población, atención salud reproductiva, planificación familiar, lucha contra enfermedades de transmisión sexual incluido el SIDA y formación de personal para población y salud reproductiva);
- Abastecimiento y depuración de agua, sistemas menores;
- Ayuda multisectorial para SSB.

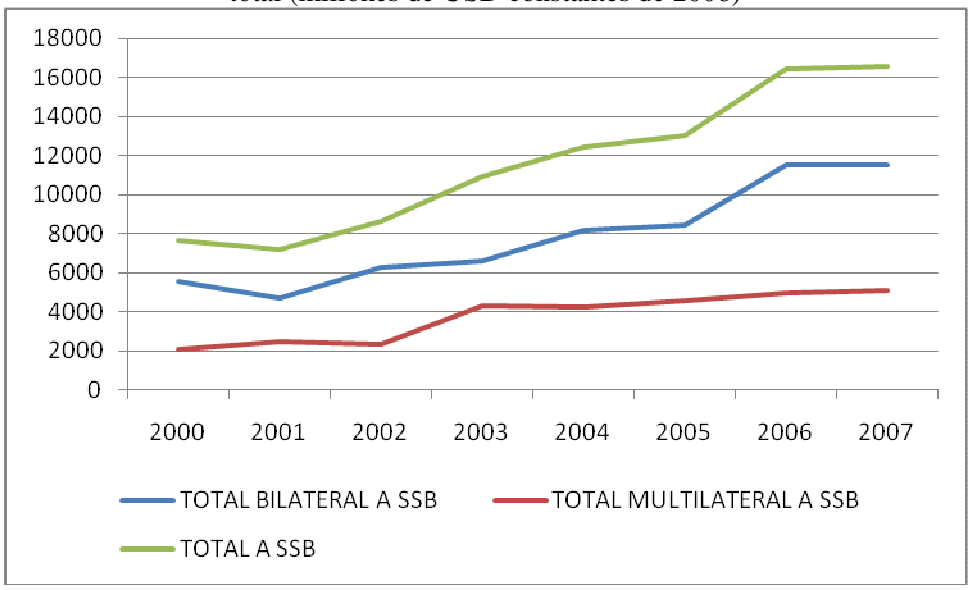
Se ha utilizado la información contenida en la base de datos CRS de la OECD disponible on line www.oecd.org/dac/stats/crs correspondiente al período 2000 a 2007 (último año disponible). La unidad de medida utilizada ha sido "compromisos" de AOD total (bilateral y multilateral) expresada en millones de USD constantes de 2006[1].

[1] Los compromisos de AOD que adquieren los donantes pueden ser llevados a la práctica en varios años consecutivos por ello puede ser que en un determinado período compromisos y desembolsos no coincidan.

3.1. Evolución de la AOD total destinada a SSB

La AOD total (bilateral y multilateral) destinada a SSB en el período 2000-2007 se ha multiplicado por dos, pasando de 7.600 a 16.600 millones de USD constantes de 2006. Esta evolución cuantitativa pone de manifiesto la orientación de la ayuda oficial al desarrollo a la cobertura de las necesidades sociales básicas, tal y como se acordó en el consenso de Oslo. La AOD bilateral supone casi el 70% del total destinado a SSB, y aunque ambas (bilateral y multilateral) han experimentado una evolución creciente, la primera ha crecido a un ritmo más acelerado que la segunda, como podemos apreciar en el gráfico 1. Si analizamos la tendencia de la AOD total comprobamos que la tasa de crecimiento no ha sido constante, incluso algunos años ha decrecido (2001 y 2005).

Gráfico 1: Evolución de los compromisos de AOD destinada a SSB: bilateral, multilateral y total (millones de USD constantes de 2006)



Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE

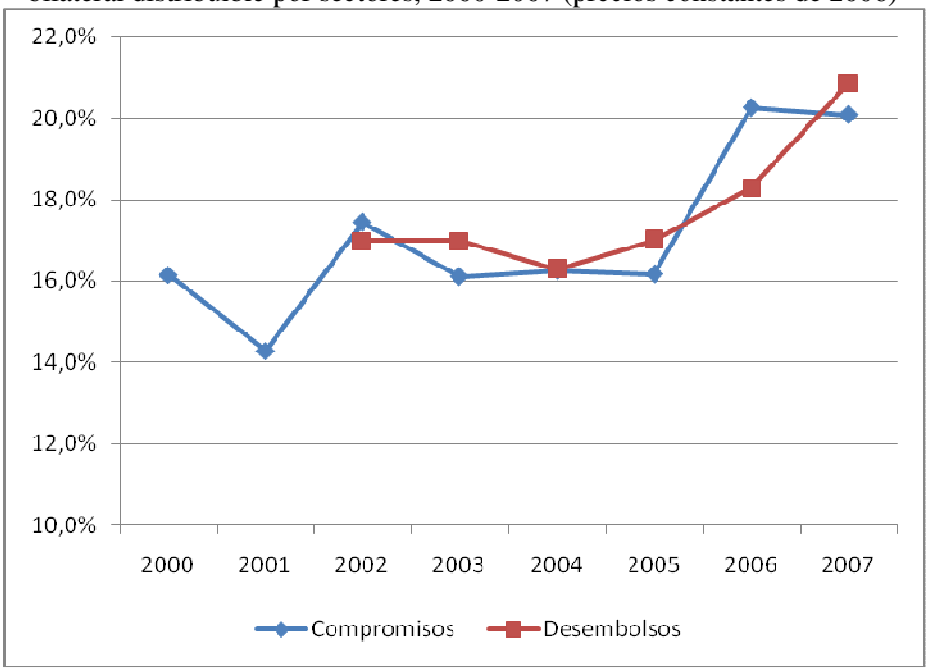
Para analizar el peso relativo que supone la AOD (bilateral o multilateral) para SSB debemos dividir la cuantía total destinada a tal fin entre la AOD distribuible por sectores[2]. Si estudiamos la *AOD bilateral destinada a SSB* (en millones de USD constantes) comprobamos que en los últimos años (2006, 2007) se habría alcanzado el 20% (sobre la AOD total distribuible) que se exige a los países donantes del CAD en la iniciativa 20/20 acordada en el consenso de Oslo[3]. Si tenemos en cuenta los desembolsos de AOD[4] en vez de los compromisos (gráfico 2), los porcentajes serían muy similares, aunque realmente en el año 2006 no se habría alcanzado el 20% si se habría superado en 2007. Si consideramos los desembolsos, podemos confirmar que existe una tendencia creciente en la AOD bilateral destinada a SSB (con la excepción del año 2004).

[2] Según la codificación que utiliza el CAD para clasificar la AOD, determinadas actuaciones se pueden clasificar por el sector de destino de la ayuda (Educación, salud, etc.) pero otras no (Ayuda de emergencia, actividades relacionadas con la deuda, etc). Si deseamos estimar el porcentaje de AOD que se dedica a un sector concreto debemos hacer referencia a la AOD total disribuible tal y como se acordó en el DAC Working Party on Statistics (WP-STAT) en 1999.

[3] Como se observa en el gráfico 1 la tendencia de la AOD bilateral es la de subir un año y bajar al siguiente, esto es común si se utiliza la cantidad comprometida por los donantes como medida de la ayuda concedida al desarrollo, por ello es usual trabajar con promedios bianuales cuando se analiza esta variable. Si se mantiene la tendencia en el año 2008 tocaría subida, por lo que se superaría el 20%.

[4] Los compromisos de AOD que adquieren los países donantes pueden ser llevados a la práctica en varios años consecutivos por ello puede ser que en un determinado período compromisos y desembolsos no coincidan. En el caso de la AOD para SSB, tanto los desembolsos de AOD bilateral distribuible por sectores como la destinada a SSB han supuesto, en promedio del período de análisis, el 77% de las cantidades comprometidas de dicho período.

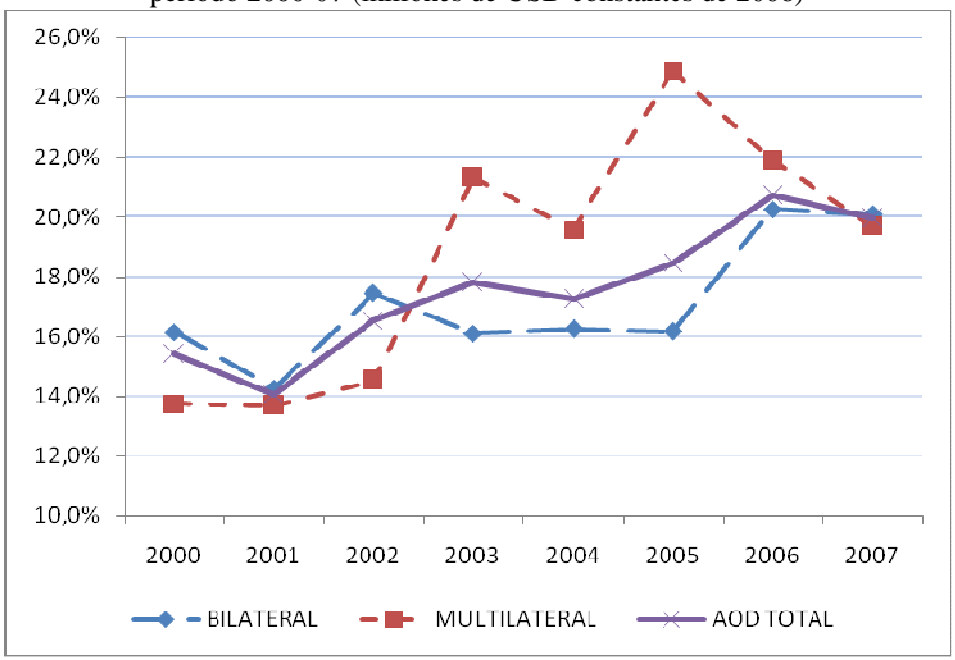
Gráfico 2: Porcentaje de AOD bilateral destinada por los países del CAD a SSB sobre la AOD bilateral distribuible por sectores, 2000-2007 (precios constantes de 2006)



Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE

Si tenemos en cuenta la *AOD destinada a SSB por las Agencias Multilaterales*[5] observamos que a pesar de que el importe total de este tipo de ayuda es inferior al de la ayuda bilateral, la AOD destinada a SSB suele suponer mayor peso relativo en la ayuda multilateral que en la bilateral, superando el 20% desde el año 2003, aunque con altibajos que no han descendido del 20% desde el citado año (véase el gráfico 3).

Gráfico 3: Evolución del peso relativo de los SSB en la AOD total, bilateral y multilateral del período 2000-07 (millones de USD constantes de 2006)



Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE

[5] En este caso trabajamos también sobre compromisos pues existe una importante pérdida de información en la base de datos de desembolsos brutos por sectores de las agencias multilaterales.

3.2. Identificación de los principales donantes internacionales para SSB

3.2.1. AOD bilateral

Al analizar los principales países donantes de AOD para SSB se nos plantean las siguientes cuestiones:

- ¿Coinciden los países que destinan mayor cuantía de AOD a los SSB con los principales donantes bilaterales internacionales en general, sin hacer distinción por sectores?
- ¿Son los países que destinan más ayuda para SSB los que a su vez conceden más importancia a dicha cobertura en el total de su cooperación bilateral?

Para responder a ambas cuestiones hemos elaborado la Tabla 1 en la que se analiza en primer lugar la contribución (en %) a la AOD total bilateral concedida por los países donantes del CAD en el período analizado (desde 2000 a 2007); en segundo lugar la proporción que supone lo que cada país destina a SSB sobre el donado por el conjunto total de los países del CAD para tal fin, y por último hemos medido la proporción de ayuda que cada país a destinado a SSB sobre el total de AOD distribuible por sectores que ha concedido en el período analizado, intentando medir de este modo la importancia que cada país donante le concede en su cooperación bilateral a la cobertura de las necesidades sociales básicas.

Los principales donantes en general (sin distinguir sectores receptores de la ayuda) son: EEUU (27%), Japón (15%), Alemania (9,3%), Francia (9,1%) y Reino Unido (8,5%). España se sitúa en séptima posición, aportando un 3% aproximadamente de la AOD bilateral total durante el período 2000-2007. Si analizamos el porcentaje que cada país ha destinado a SSB sobre el total de AOD bilateral destinado a tal fin comprobamos que EEUU se mantiene en primera posición, sin embargo Japón (segundo donante internacional) bajaría a la sexta posición en cuanto a su aportación a SSB se refiere sobre el conjunto de todos los países. Llama la atención el caso de Francia pues siendo la 4ª donante bajaría a la doceava posición con respecto a su dedicación a SSB. Reino Unido, por el contrario, asciende en el ranking pasando de la 5ª posición como donante internacional a la 2ª en cuanto a SSB se refiere.

Por último, al medir la importancia relativa que cada país donante concede a su cooperación (bilateral) en SSB (última columna de la tabla) comprobamos que Irlanda, que se encuentra en la cola en el ranking de donantes internacionales (18 posición) se sitúa la primera en cuanto a la importancia que concede a los SSB sobre su AOD distribuible, lo cual es muy llamativo. Algo similar ocurre con Luxemburgo, Bélgica y Noruega, países que ocupan posiciones rezagadas en el ranking de donantes (en cuanto a aportación a la AOD total bilateral) y sin embargo están concienciados con la necesidad de priorizar la cobertura de las necesidades sociales básicas en su cooperación. Alemania, sin embargo, a pesar de ser uno de los principales donantes tan solo dedica un 10% de su AOD bilateral a la cobertura de estos servicios. En promedio, los países del CAD habrían destinado un 19,6% de su AOD bilateral distribuible a los SSB.

Según la información anterior solo 7 países habrían destinado en el conjunto del período 2000-2007 más del 20% de su AOD bilateral distribuible a SSB. Sin embargo, como el período analizado comprende 8 años en los que la AOD bilateral de muchos países ha crecido considerablemente, creemos interesante mostrar la importancia relativa de los SSB para los países donantes del CAD en el último bienio con información disponible (06/07). Para ello se ha calculado el importe medio de ayuda comprometida en el bienio 2006/2007 por cada país[6] para SSB y se ha comparado con la media de AOD distribuible por sectores en el mismo bienio por cada uno de dichos países (Tabla 2). En la actualidad, serían 9 los países que han superado el umbral del 20%. En promedio, los países del CAD han comprometido el 23% de su AOD bilateral para SSB. Comparando esta media con la referente al período completo de este análisis

[6] Como ya se mencionó anteriormente es práctica común considerar los importes medios bianuales de los compromisos, pues ofrecen una imagen más real de los mismos.

(2000-07) vemos que ha aumentado el porcentaje de AOD destinado a SSB por el conjunto de donantes internacionales.

Tabla 1: Ranking de donantes internacionales de AOD bilateral (período 2000-2007) para SSB

PAÍS	% AOD total	PAÍS	AOD SSB	% SSB vs total a SSB	PAÍS	% SSB/ AOD DIST SEC PAIS
United States	27,0%	United States	25.857,6	36,4%	Ireland	33,1%
Japan	15,1%	United K.	8.754,7	12,3%	United K.	30,5%
Germany	9,3%	Netherlands	4.982,4	7,0%	Canada	28,3%
France	9,1%	Germany	3.849,7	5,4%	Luxembourg	27,4%
United K.	8,5%	Canada	3.221,2	4,5%	Netherlands	26,3%
Netherlands	6,4%	Japan	2.948,6	4,1%	United States	24,4%
Canada	3,2%	Norway	1.892,7	2,7%	New Zealand	23,8%
Spain	2,9%	Sweden	1.680,8	2,4%	Belgium	19,1%
Sweden	2,8%	Spain	1.618,9	2,3%	Norway	17,3%
Norway	2,7%	Australia	1.409,4	2,0%	Sweden	17,2%
Italy	2,2%	Belgium	1.178,7	1,7%	Greece	16,4%
Australia	1,9%	France	1.173,4	1,7%	Australia	15,6%
Denmark	1,9%	Denmark	1.102,4	1,6%	Denmark	14,7%
Belgium	1,8%	Ireland	795,4	1,1%	Spain	14,6%
Switzerland	1,6%	Italy	636,9	0,9%	Italy	13,4%
Austria	1,0%	Switzerland	613,6	0,9%	Finland	12,8%
Finland	0,7%	Finland	339,5	0,5%	Switzerland	11,4%
Ireland	0,6%	Luxembourg	243,0	0,3%	Germany	10,9%
Portugal	0,6%	NewZealand	193,2	0,3%	Austria	8,9%
Luxembourg	0,2%	Austria	177,4	0,2%	Japan	5,0%
Greece	0,2%	Greece	163,7	0,2%	France	4,4%
New Zealand	0,2%	Portugal	54,3	0,1%	Portugal	3,1%
TOTAL	100%	TOTAL	**71.060,7**	100,0%	TOTAL	19,6%

Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE

Tabla 2: Contribución de la AOD de los países donantes del CAD a los SSB en el bienio 2006-07

PAÍS	(1) AOD Bilateral SSB	(2)AOD bilateral distribuible sect.	% (1) vs (2)
Ireland	171,6	443,6	39%
United K.	1.523,0	4.461,1	34%
Luxembourg	46,7	141,5	33%
Canada	534,4	1.696,8	31%
Netherlands	1.099,6	3.514,7	31%
United States	5.275,5	17.502,8	30%
New Zealand	42,1	174,9	24%
Norway	375,0	1.742,6	22%
Belgium	192,8	911,4	21%
Greece	29,0	166,0	17%
Sweden	265,3	1.540,8	17%
Denmark	128,2	796,2	16%
Spain	303,5	2.034,1	15%
Finland	44,7	371,8	12%
Australia	149,7	1.255,5	12%
Italy	71,5	643,1	11%
Austria	32,0	294,3	11%
Germany	593,3	5.529,7	11%
Switzerland	43,7	721,8	6%
France	236,3	4.409,5	5%
Japan	353,8	8.536,2	4%
Portugal	8,4	217,4	4%
TOTAL	13.025,5	57.105,7	23%

Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE

3.2.2. AOD Multilateral

Si analizamos la información disponible sobre los compromisos de AOD para la cobertura de SSB por parte delos organismos multilaterales podemos destacar que los principales donantes han sido la *Comisión Europea*, (CE), el *Fondo Global contra el SIDA, Tuberculosis y Malaria*, la *Asociación Internacional para el Desarrollo* (ADI), y *UNICEF*.

La CE es el principal donante multilateral, ha aportado el 25% del total destinado por los organismos multilaterales para SSB. La Unión Europea tiene como objetivos prioritarios de su cooperación: garantizar el agua potable y una alimentación mejor, facilitar el acceso a la enseñanza, la sanidad, el empleo, los servicios sociales y territoriales, mejorar las infraestructuras y el medio ambiente, erradicar las enfermedades y permitir el acceso a medicamentos de bajo coste para combatir las epidemias como el sida. Como vemos casi todos están directamente relacionados con los SSB. La contribución de la CE a SSB sufre variaciones importantes a lo largo del período, aunque últimamente su contribución anual parece estabilizarse.

El Fondo Global se creó en el 2002 y desde entonces su contribución sigue una tendencia creciente y progresiva. Como su propio nombre indica sus actuaciones están basadas en la prevención y tratamiento del SIDA, Malaria y Tuberculosis, por lo tanto en el código *Población y salud reproductiva* (60%)y Salud básica (resto). Aporta, en promedio del período el 24,8% del total multilateral a SSB.

La AID fue creada en 1960, su misión es conceder créditos a los 81 países más pobres del mundo. Su ayuda está dirigida a la mejora en el acceso a la educación, salud, abastecimiento de agua y cuidados sanitarios, todos ellos considerados SSB, por ello es uno de los principales donantes internacionales para SSB (24% de media en el período). Las aportaciones anuales de la AID no son constantes sino que sufren variaciones significativas.

El Fondo de las Naciones Unidas para la Infancia (UNICEF): También contribuye a la cobertura de las necesidades sociales básicas de los más desfavorecidos. Entre sus prioridades de actuación figuran la alimentación, atención sanitaria, asistencia materno-infantil; la educación básica e igualdad entre los géneros y la prevención del SIDA (todas ellas consideradas necesidades sociales básicas). Es el tercer donante multilateral para SSB, contribuyendo con un 10,5% del total del período.

Tabla 3: Compromisos de AOD para SSB de los Organismos multilaterales 2000-07 (millones de USD constantes de 2006)

ORGANISMO	2000	2001	2002	2003	2004	2005	2006	2007	TOTAL
Fondo Desarrollo Africano	61,8	29,4	121,2	204,7	136,1	144,9	213,3	17,9	929,4
Fondo Desarrollo Asiático	70,6	259,1	130,6	135,6	125,1	18,8	0,0	56,2	795,9
Comisión Europea	966,4	554,3	483,5	811,5	1.051,8	1.472,9	971,9	1.175,3	7.487,6
Asociación de Desarrollo Internacional	630,3	807,5	696,3	1.077,2	1.369,9	654,6	1.126,1	864,6	7.226,5
Fondo Especial Banco Interamericano Desarrollo	0,0	4,1	39,0	0,0	10,5	68,1	15,0	8,8	145,4
Programa de las Naciones Unidas para el desarrollo (PNUD)	0,0	0,0	0,0	0,0	16,4	19,1	21,5	15,0	72,0
Fondo de las Naciones Unidas para la infancia (UNICEF)	372,5	439,8	377,9	359,1	335,8	408,1	514,6	374,3	3.182,1
Programa de las Naciones Unidas para el SIDA (ONUSIDA)	0,0	115,9	82,7	145,2	142,7	126,0	180,2	122,4	915,2
Fondo de Población de las Naciones Unidas (FPNU)	0,0	271,7	336,5	413,3	205,8	208,4	213,4	200,2	1.849,3
Fondo Global contra el Sida, Tuberculosis y Malaria	0,0	0,0	72,2	1.134,6	861,9	1.461,7	1.716,9	2.235,1	7.482,4
TOTAL	2.101,6	2.481,8	2.339,8	4.281,1	4.256,2	4.582,6	4.973,0	5.069,8	30.085,8

Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE

3.3. Distribución sectorial de la AOD concedida para SSB

En el período 2000-2007 la AOD total (bilateral y multilateral) destinada a SSB ha ido dirigida a *Políticas y programas sobre población y salud reproductiva* en primer lugar (39%), seguida de *Salud básica* (27%), *Educación Básica* (27%) y en las últimas posiciones *Abastecimiento y depuración de agua, sistemas menores* (8%) *y Ayuda multisectorial a SSB* (6%).

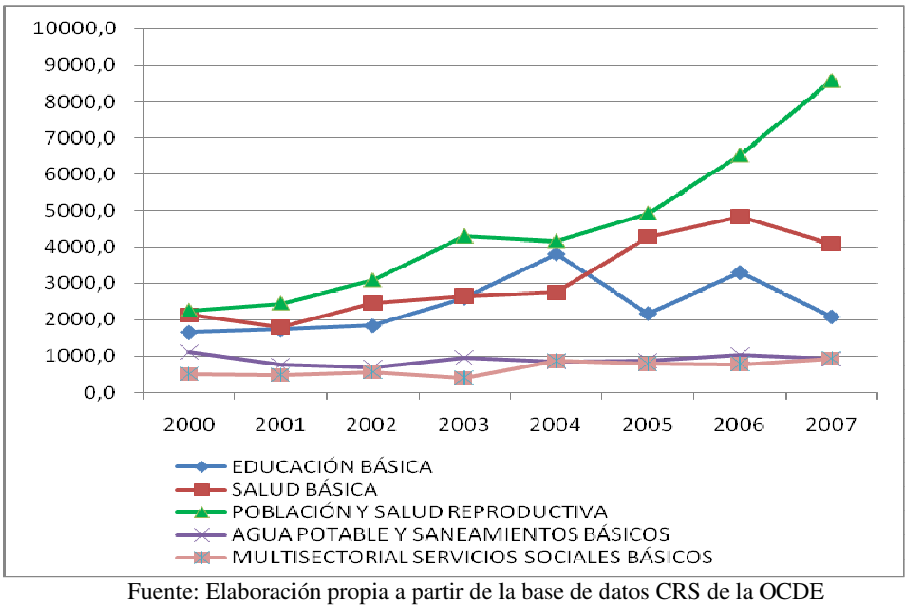
Tabla 4: Distribución sectorial de la AOD total para SSB (millones de USD constantes de 2006)

SECTORES CAD	2000	2001	2002	2003	2004	2005	2006	2007	TOTAL	%
Educación Básica	1663,0	1731,7	1839,7	2611,4	3816,4	2181,7	3300,9	2088,5	19233,3	21%
Salud Básica	2119,9	1778,6	2461,7	2644,6	2740,8	4273,4	4829,5	4080,4	24928,9	27%
Población y Salud reproductiva	2258,7	2458,6	3082,0	4296,9	4164,6	4932,7	6530,8	8576,2	36300,7	39%
Abastecimiento y depuración de agua	1122,8	751,0	680,8	962,1	844,3	882,5	1034,3	916,4	7194,3	8%
Multisectorial	497,8	481,6	559,6	398,8	870,1	788,3	786,3	933,8	5316,2	6%
TOTAL	**7662,3**	**7201,5**	**8623,9**	**10913,8**	**12436,2**	**13058,5**	**16481,8**	**16595,2**	92973,4	100%
Variación interanual		**-6,01%**	**19,75%**	**26,55%**	**13,95%**	**5,00%**	**26,21%**	**0,69%**		

Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE

Si se analiza la evolución temporal de la distribución sectorial comprobamos que existe una clara tendencia creciente en la AOD destinada a *Políticas y programas sobre población y salud reproductiva* y también a *Salud básica* (situándose el primer sector por encima del segundo). Sin embargo la ayuda destinada a *Educación básica* aunque parecía seguir una tendencia creciente desde el año 2000 hasta el 2004, ha sufrido altibajos al final del período analizado (disminuyendo en el año 2007). La partida destinada a *Abastecimiento y depuración de agua, sistemas menores* se mantiene a lo largo del período en valores muy bajos, sin apuntar ninguna tendencia, alternando años de crecimiento con años de descenso. El mismo comportamiento podemos observar en la *Ayuda multisectorial para servicios sociales básicos*.

Gráfico 4 : Evolución de la AOD total destinada a SSB por subsectores CRS



Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE

Si nos centramos en la ***AOD bilateral*** destinada por los países del CAD a SSB comprobamos que se ha duplicado desde el año 2000 al 2007, pasando de 5.560 a 11.525 millones de USD constantes de 2006, manteniendo una tendencia creciente en prácticamente todo el período contemplado.

Como se puede observar en la tabla 5 en la que aparece el conjunto de los países donantes del CAD (ordenados según la cuantía de AOD destinada a SSB), los principales subsectores

receptores de esta ayuda bilateral son los mismos que los de la AOD total: el subsector *Políticas y programas sobre población y salud reproductiva*, así como *Salud básica* (los cuales han recibido en promedio el 35% del total de AOD a SSB), les sigue por orden de importancia cuantitativa *Educación básica* con un 19%; *Abastecimiento y depuración de agua, sistemas menores* (8%) y por último a *Ayuda multisectorial a SSB* (3%).

Dentro del grupo de los países donantes del CAD, EEUU sigue la tónica anteriormente comentada en cuanto a la distribución sectorial de su AOD a SSB; sin embargo, Reino Unido ha destinado más a Educación que a Salud. Algunos países han priorizado la educación sobre otros SSB como es el caso de Holanda, Francia y Nueva Zelanda. Otros países como Japón, le dedican menos cuantía a Población. Los principales donantes de ayuda para abastecimiento de agua son Alemania y Japón. España le dedica la mayor cuantía de su AOD a SSB al sector Salud, Educación y Multisectorial.

Tabla 5: Distribución sectorial de la AOD total bilateral a SSB por países donantes del CAD (millones de USD constantes de 2006) en el período 2000-07.

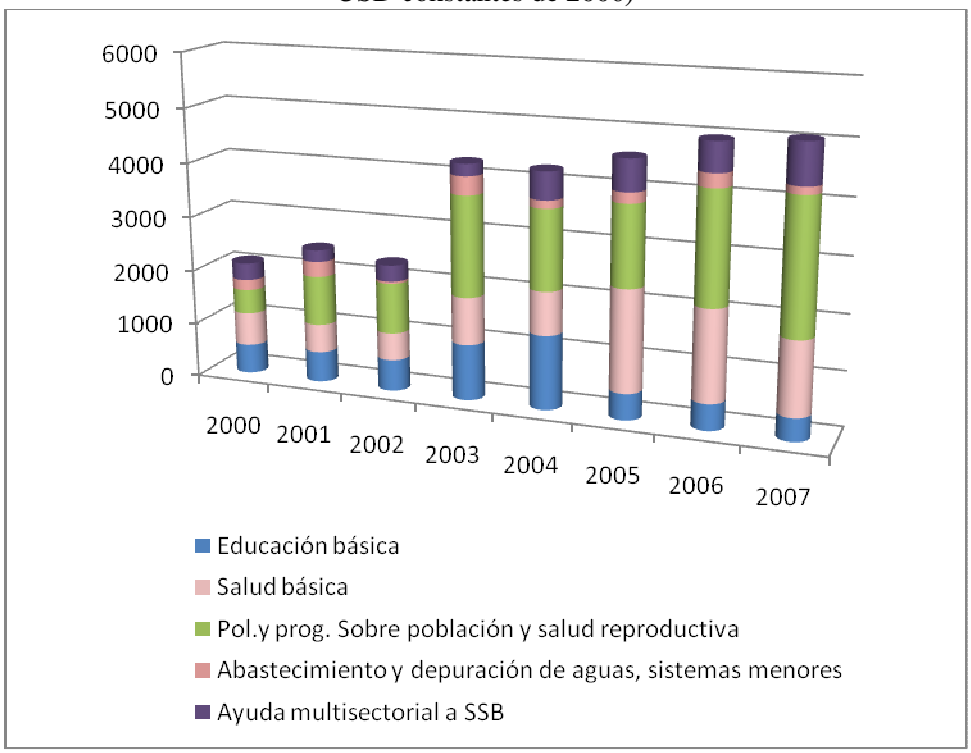
PAÍS	Población	Salud	Educación	Agua	Multisect.	Total SSB	% País vs total SSB
United States	**15.812,5**	**6.658,3**	**2.734,3**	440,9	**211,5**	25.857,6	36,4%
United Kingdom	**4.147,4**	**1.998,0**	**2.277,9**	328,1	3,2	8.754,7	12,3%
Netherlands	797,0	774,9	**2.615,5**	**634,0**	161,0	4.982,4	7,0%
Germany	**857,0**	662,7	830,2	**1.290,4**	**209,4**	3.849,7	5,4%
Canada	735,9	**1.148,0**	1.022,4	167,6	147,3	3.221,2	4,5%
Japan	82,1	851,7	811,8	**1.103,9**	99,1	2.948,6	4,1%
Norway	416,9	648,4	682,3	66,6	78,4	1.892,7	2,7%
Sweden	537,0	518,1	453,8	154,9	17,1	1.680,8	2,4%
Spain	186,4	570,0	370,0	235,6	**257,0**	1.618,9	2,3%
Australia	299,3	572,1	387,7	48,0	102,3	1.409,4	2,0%
Belgium	177,7	495,6	133,6	162,3	209,6	1.178,7	1,7%
France	84,3	194,9	542,7	228,5	123,0	1.173,4	1,7%
Denmark	156,8	325,3	241,8	345,8	32,7	1.102,4	1,6%
Ireland	230,4	283,5	172,3	76,9	32,3	795,4	1,1%
Italy	68,7	363,8	34,7	124,6	45,1	636,9	0,9%
Switzerland	33,5	246,0	122,1	58,4	153,6	613,6	0,9%
Finland	55,7	91,3	80,3	85,6	26,7	339,5	0,5%
Luxembourg	54,3	111,5	50,1	16,2	10,9	243,0	0,3%
New Zealand	25,3	33,2	120,6	7,3	6,9	193,2	0,3%
Austria	12,6	96,9	23,1	39,9	4,9	177,4	0,2%
Greece	15,6	102,8	42,2	2,9	0,2	163,7	0,2%
Portugal	1,0	8,7	41,9	0,9	1,7	54,3	0,1%
TOTAL	**24.787,4**	**24.928,9**	**13.791,3**	**5.619,3**	**1.933,8**	**71.060,7**	**100,0%**
Proporción sector vs total	35%	35%	19%	8%	3%		

Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE

Si analizamos la distribución sectorial de los compromisos de ***AOD de los organismos multilaterales internacionales*** destinados a SSB en el mismo período (2000-07) comprobamos que es muy similar a la de la bilateral, *Población* y *Salud básica* siguen siendo los sectores

prioritarios a los que va destinada la ayuda, aunque en los últimos años ha ido ganando peso la *Ayuda multisectorial a SSB* (que tiene más importancia en la multilateral que en la bilateral) en detrimento de la concedida para *Educación básica*.

Gráfico 5: Distribución sectorial de la AOD multilateral a SSB (compromisos, millones de USD constantes de 2006)



Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE

Si analizamos el sector receptor de la ayuda recibida por los principales donantes multilaterales a SSB comprobamos que:

- *La Comisión Europea (CE)* no presenta una especialización clara en algún sector concreto, sino que hay una gran variabilidad según el año analizado. Así por ejemplo en el año 2001, 2005 y 2006 la salud básica fue el sector prioritario, la educación ha sufrido numerosos altibajos y también la multisectorial, que se ha convertido en la principal receptora de ayuda de la CE en 2007.
- *La Asociación internacional para el Desarrollo (AID)*: ha dirigido su ayuda a Salud, educación y Población. A agua y multisectorial tan solo ha destinado una pequeña cantidad en el año 2007.
- *El Fondo de las Naciones Unidas para la Infancia (UNICEF)*: Sus principales aportaciones van dirigidas a Multisectorial, Salud y Población.

4. PRINCIPALES INDICADORES PARA MEDIR EL DESARROLLO EN SERVICIOS SOCIALES BÁSICOS

Si tratamos de valorar si la ayuda oficial al desarrollo en servicios sociales básicos está llegando a aquellos países realmente más necesitados en este terreno, debemos analizar cuales son los indicadores más importantes relacionados con el nivel de satisfacción en materia de servicios sociales básicos y analizar su valor en los países receptores (socios).

Se van a utilizar dos índices para medir el desarrollo en SSB: El Índice de Desarrollo Humano (IDH) y el Índice de Capacidades Básicas (ICB).

4.1. El Índice de Desarrollo Humano

Se podría decir que el desarrollo humano es el proceso por el que una sociedad mejora las condiciones de vida de sus ciudadanos a través de un incremento de los bienes con los que puede cubrir sus necesidades básicas y complementarias, y de la creación de un entorno en el que se respeten los derechos humanos de todos ellos. Se puede decir que es la cantidad de opciones que tiene un ser humano en su propio medio para ser o hacer lo que él desea ser o hacer.

Antes de adentrarnos en el análisis es importante precisar que "medir" fenómenos como pobreza o desarrollo es sumamente complejo, debido a que se trata de realidades multidimensionales, cruzadas por un amplio rango de vectores; en cualquier caso, es preferible realizar el análisis, con los datos con los existentes, aun a riesgo de utilizar mediciones parciales, que no llegan a abarcar todas las dimensiones.

Cómo se ha comentado en la introducción, los indicadores de pobreza y desarrollo más utilizados internacionalmente son los cálculos del Banco Mundial acerca del número de personas que viven con menos de uno o dos dólares diarios o la clasificación desarrollada por el Programa de las Naciones Unidas para el Desarrollo (PNUD) basada en el Índice de Desarrollo Humano (IDH), que combina cifras de ingreso con indicadores de salud y educación.

El IDH es un indicador social estadístico, es decir, una medida de resumen estadística, referida a la cantidad o magnitud de un conjunto de parámetros o atributos de una sociedad, y permite clasificar las naciones respecto a las variables que se están analizando. El IDH realiza una medición por país y está compuesto por tres parámetros:

- **Vida larga y saludable,** medida según la esperanza de vida al nacer.

  La ***esperanza de vida*** es la media de la cantidad de años que vive una cierta población en un cierto periodo de tiempo. Se ve influenciada por factores como la calidad de la medicina, la higiene, las guerras, etc. (aunque en la actualidad no se tienen en cuentan las muertes violentas)

  La ***esperanza de vida al nacer*** es una estimación del promedio de años que viviría un grupo de personas nacidas el mismo año si los movimientos en la tasa de mortalidad de la región evaluada se mantuvieran constantes.

- **Educación,** medida por la tasa de alfabetización de adultos y la tasa bruta combinada de matriculación en educación primaria, secundaria y superior, así como los años de duración de la educación obligatoria.

  El ***índice o tasa de alfabetización*** es el porcentaje de la población que sabe leer o escribir después de determinada edad. No existe una convención internacional acerca de la edad a tomar en cuenta ni el nivel cualitativo de lectura o escritura.

  La ***educación primaria*** o educación básica es la que asegura la correcta alfabetización, es decir, que enseña a leer, escribir, cálculo básico y algunos conceptos culturales considerados como imprescindibles.

  La ***educación secundaria*** es la que tiene como objetivo capacitar al alumno para proseguir estudios superiores o bien para incorporarse al mundo laboral.

  La ***educación superior*** se refiere al proceso, centro e instituciones educacionales posteriores a la educación secundaria y en la que se obtiene un título superior.

- **Nivel de vida digno**, medido por el PIB per cápita PPA en dólares.

  La ***renta per cápita*** o ***PIB per cápita*** es la relación que hay entre el PIB (Producto Interior Bruto) de un país y su número de habitantes, por lo que hay que dividir el PIB por el dato de población. Es un indicador muy utilizado para medir la riqueza económica de un país, aunque existen muchas limitaciones (ignora las desigualdades de renta, no computa la externalidades negativas, un aumento del PIB no supone necesariamente un aumento del bienestar,…).

El IDH busca medir dichas variables a través de un índice compuesto, por medio de indicadores que se relacionan en los tres aspectos mencionados en forma sinóptica.

Es calculado desde 1990 por el Programa de las Naciones Unidas para el Desarrollo (PNUD). En el informe publicado en 2008 el índice de desarrollo humano fluctuaba entre Islandia con un índice de 0,968 en la primera posición al 0,329 de Sierra Leona en el puesto 179. El PNUD clasifica los países en tres grandes grupos:

- País con desarrollo humano alto (IDH ≥ 0,8): 75 países.
- País con desarrollo humano medio (0,5 ≤ IDH < 0,8): 77 países.
- País con desarrollo humano bajo (IDH < 0,5): 25 países.

4.2. El Índice de Capacidades Básicas (ICB)

Además del IDH, existe otro indicador, desarrollado por Social Watch, denominado Índice de Capacidades Básicas (ICB), como una forma no basada en el ingreso de identificar situaciones de pobreza. El ICB es comparativamente más fácil de construir y puede ser aplicado a nivel subnacional y municipal, sin necesidad de recurrir a costosas encuestas de hogares, como lo requieren los índices basados en el ingreso. Además, al no utilizar el ingreso, el ICB es consistente con las definiciones de pobreza basadas en la privación de capacidades y la negación de los derechos humanos.

El ICB se construye con tres indicadores: porcentaje de niños que llegan a quinto grado, supervivencia hasta los 5 años (en base a la mortalidad de menores de 5 años) y porcentaje de partos atendidos por personal médico especializado. Estos indicadores expresan diferentes dimensiones abordadas por los objetivos de desarrollo acordados internacionalmente (educación, salud infantil y salud reproductiva). Asimismo, las investigaciones señalan que, como medida resumen, el ICB provee un panorama general consistente con la situación sanitaria y el desempeño educacional básico de una población.

Los indicadores del ICB llegan a su máximo nivel posible cuando todas las mujeres reciben asistencia médica durante el parto, cuando ningún niño o niña deja de ir a la escuela hasta completar quinto grado y cuando la mortalidad infantil se reduce a la expresión mínima posible de menos de cinco muertes cada mil niños nacidos vivos. Estos indicadores se asocian estrechamente a capacidades que los integrantes de una sociedad deben tener y que se refuerzan mutuamente para posibilitar un mayor logro de desarrollo individual y colectivo. Refieren, en especial, a aquellas capacidades que logran incorporar sus miembros más jóvenes, potenciando de ese modo el desarrollo futuro de los países.

La utilidad del ICB radica en que ha demostrado estar altamente correlacionado con la medición de otras capacidades humanas relativas al desarrollo social de los países. Mediante este índice se puede asignar un valor a cada país y de este modo compararlo con otros o evaluar su evolución en el tiempo.

Un valor de ICB cercano a 100 no implica necesariamente un nivel elevado de desarrollo social. Apenas significa que el país logró la cobertura universal de los requisitos esenciales mínimos para poder avanzar hacia un bienestar mejor. Es un punto de partida, no de llegada.

En el presente apartado se pretende identificar a los principales países receptores de la AOD para SSB en los últimos años, profundizando en el nivel de desarrollo alcanzado en los mencionados países en determinados indicadores que ponen de manifiesto su desarrollo en esta materia.

4.3. Análisis de la AOD recibida por los países receptores para la cobertura de SSB comparado con el índice de Desarrollo Humano

Los países con IDH más bajo se encuentran en el continente africano, tal y como podemos comprobar en la tabla 6.

Tabla 6: Últimos países de la lista según su IDH

Posición		País	IDH En 2006 (publicado en 2008)
Dato de 2006 (publicado en 2008)	Cambio comparado a datos de 2005 (publicados en 2007)[7]		
169	=	Etiopía	▼0,389
170	=	Chad	▲0,389
171	▲	Guinea-Bissau	▲0,383
172	▼	Burundi	▼0,382
173	▲	Burkina Faso	▲0,372
174	▼	Níger	▼0,370
175	▼	Mozambique	▼0,366
176	▼	Liberia	▲0,364
177	▼	República Democrática del Congo	▼0,361
178	▼	República Centroafricana	▼0,352
179	▼	Sierra Leona	▼0,329

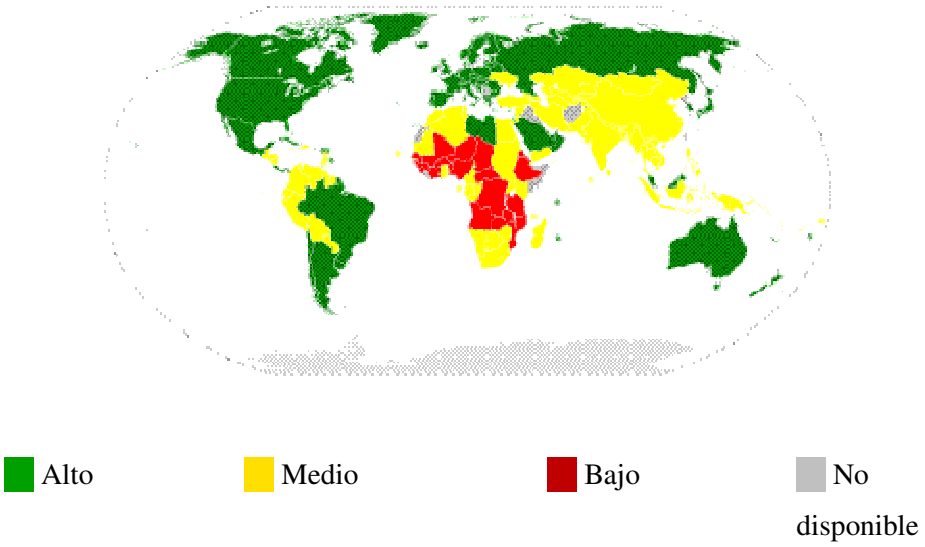
Fuente: Informe de Desarrollo Humano 2008/2009 del PNUD

Cómo se puede observar, todos los países con un índice de desarrollo humano inferior al 0,5 se encuentran en el continente africano y lo que es más alarmante es que en los años correspondientes a los últimos datos disponibles, el indicador, lejos de mejorar, empeora en la mayoría de los países.

El resto del continente africano (con excepción de Argelia) alcanza un IDH medio (entre 05 y 0,8), al igual que la mayor parte de Asia, Centroamérica y la una parte importante del noroeste de América del Sur.

[7] ▲= ascenso comparado a los datos de 2006 (publicados en 2008) =igual a los datos de 2006 (publicados en 2007) ; ▼= descenso comparado a los datos de 2006 (publicados en 2008). Valores similares de IDH en la presente lista no representan igualdad en la posición ya que ésta se determina a partir de valores de IDH con seis decimales.

Gráfico 6: Mapa del mundo para el IDH de los diferentes Estados, según el Informe 2007/2008.



Fuente: Informe de Desarrollo Humano 2008/2009 del PNUD

Si analizamos el destino geográfico de la *AOD total comprometida* durante el *período 2000-2007* para SSB comprobamos que 5 países (el 3%): India, Nigeria, Bangladesh, Tanzania y Etiopía concentran el 20% del total de AOD cuyo destino es conocido, y el 50% de la ayuda se concentra en el 12% de los países.

Si analizamos el importe medio de *AOD comprometida para SSB* en el *bienio 2006/2007* para analizar la situación de los países en el ranking anterior en la actualidad, observamos que los países que concentrarían el 20% de la ayuda serían India, Etiopía, Kenia y Nigeria. Por tanto comprobamos que si bien India se mantiene en primera posición concentrando en 2006/07 un 12% de la ayuda, algunos países como Etiopía, Kenia y Sudáfrica (entre otros) ascenderían de posición en el ranking. A cambio países como Níger, Eritrea o Somalia bajarían su posición. Llama la atención el caso de Hong Kong que ocupando los primeros puestos por la cuantía recibida en el total del período (2000-07) no ha recibido nada para SSB según los datos analizados, en el último bienio.

Si analizamos el IDH de los países que han recibido la mayor parte de AOD para SSB en los últimos años comprobamos que India, con un IDH de 0,619 (medio) casi el doble del IDH de los países peor clasificados, que rondan el 0,35, recibe el mayor porcentaje de la ayuda en SSB, el 7,7% de la ayuda total del periodo, llegando al 12,5% si analizamos la ayuda del bienio.

Observando los países que concentran el 50% de la ayuda del periodo, llegando al 55% de la del bienio, son 21 países, un 11,86% del total de países receptores; podemos comprobar como únicamente 7 de ellos tienen un IDH bajo (inferior al 0,5), mientras que entre los principales receptores podemos encontrar países como Hong Kong (comentado anteriormente), con un altísimo IDH y con un 2,7% del total de la ayuda del periodo, si bien es verdad que en el último bienio no ha recibido ayuda para SSB.

Tabla 7: Distribución geográfica de los compromisos de AOD total (bilateral y multilateral) a SSB por países receptores (millones de USD constantes 2006) en el período 2000-2007.

PAÍS	IDH (2005)	PERÍODO 2000/2007			BIENIO 2006/2007			
		TOTAL	%	% acum.	TOTAL	Media Anual	%	% acum.
India	0.619	7.728,4	7,7%	7,7%	3.232,8	1.616,4	12,5%	12,5%
Nigeria	0.470	3.352,3	3,3%	11,0%	950,1	475,0	3,7%	16,2%
Bangladesh	0.547	3.151,8	3,1%	14,2%	505,6	252,8	2,0%	18,1%
Tanzania	0.467	2.953,5	2,9%	17,1%	852,2	426,1	3,3%	21,4%
Etiopía	0.406	2.798,4	2,8%	19,9%	1.720,7	860,3	6,7%	28,1%
Hong Kong, China	0.937	2.758,1	2,7%	22,6%	0,0	0,0	0,0%	28,1%
Zambia	0.434	2.433,0	2,4%	25,0%	653,8	326,9	2,5%	30,6%
Kenia	0.521	2.424,0	2,4%	27,4%	1.023,1	511,5	4,0%	34,5%
Níger	0.374	2.282,4	2,3%	29,7%	139,6	69,8	0,5%	35,1%
Uganda	0.505	2.115,8	2,1%	31,8%	615,4	307,7	2,4%	37,5%
Sudáfrica	0.674	2.086,8	2,1%	33,9%	835,7	417,8	3,2%	40,7%
Mozambique	0.384	2.073,9	2,1%	36,0%	709,8	354,9	2,7%	43,4%
Indonesia	0.728	1.971,9	2,0%	37,9%	614,2	307,1	2,4%	45,8%
Iraq	..	1.774,3	1,8%	39,7%	491,1	245,5	1,9%	47,7%
Pakistán	0.551	1.744,2	1,7%	41,4%	792,1	396,0	3,1%	50,8%
Eritrea	0.483	1.717,6	1,7%	43,1%	54,8	27,4	0,2%	51,0%
Afganistán	..	1.659,6	1,7%	44,8%	668,6	334,3	2,6%	53,6%
Somalia	..	1.538,2	1,5%	46,3%	114,0	57,0	0,4%	54,0%
Kazajstán	0.794	1.529,9	1,5%	47,8%	17,7	8,9	0,1%	54,1%
Marruecos	0.646	1.382,5	1,4%	49,2%	228,6	114,3	0,9%	55,0%
Vietnam	0.733	1.338,4	1,3%	50,5%	388,7	194,3	1,5%	56,5%
Total especificado		**100.555,9**	**100,0%**	**100,0%**	**25.872,7**	**12.936,3**	**100,0%**	**100,0%**
Total no especificado		**28.529,9**						
TOTAL		**129.085,8**						

Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE y el informe de Desarrollo Humano del PNUD

Analizando el bienio 2006/07, se observa algunas diferencias significativas con los datos agregados para los 8 años. La ayuda está más concentrada que el periodo total, ya que únicamente 18 países, 10% de los países receptores, concentran el 60% de la ayuda recibida en el bienio. De estos países, un 40% poseen un IDH bajo, mientras que el 60% restante alcanza un IDH medio.

Tabla 8: Distribución geográfica de los compromisos de AOD total (bilateral y multilateral) a SSB por países receptores (millones de USD constantes 2006) en el bienio 2006-07

PAÍS	IDH	TOTAL 0607	MEDIA 2006-07	%/total	%ac/total
India	0,619	3232,785686	1616,392843	12,49%	12,49%
Etiopia	0,406	1720,669448	860,3347242	6,65%	19,15%
Kenya	0,521	1023,073749	511,5368747	3,95%	23,10%
Nigeria	0,470	950,0691954	475,0345977	3,67%	26,77%
Tanzania	0,467	852,2034508	426,1017254	3,29%	30,07%
Sudáfrica	0,674	835,6801817	417,8400909	3,23%	33,30%
Pakistan	0,551	792,0574124	396,0287062	3,06%	36,36%
Áreas Palestinas	0,731	781,1708357	390,5854179	3,02%	39,38%
Mozambique	0,384	709,79936	354,89968	2,74%	42,12%
Afganistan	..	668,570547	334,2852735	2,58%	44,70%
Zambia	0,434	653,7786923	326,8893461	2,53%	47,23%
Uganda	0,505	615,4025854	307,7012927	2,38%	49,61%
Indonesia	0,728	614,2320915	307,1160457	2,37%	51,98%
China	0,777	515,6133098	257,8066549	1,99%	53,98%
Bangladesh	0,547	505,6131489	252,8065744	1,95%	55,93%
Iraq	..	491,088726	245,544363	1,90%	57,83%
República Democrática del Congo	0,411	470,4454648	235,2227324	1,82%	59,65%
Mali	0,380	458,2356953	229,1178476	1,77%	61,42%

Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE y el informe de Desarrollo Humano del PNUD

Si comparamos la lista de los países con menor IDH (Tabla 6) con respecto a los principales países receptores de AOD para SSB (tablas 7 y 8) comprobamos como la mayoría de los que están incluidos en la primera no están en la segunda, es decir que la mayoría de los países con peores valores en el IDH están muy alejados de los primeros puestos en cuanto a AOD recibida para SSB, con la única excepción de ***Etiopía***, que con un IDH de 0,406 para 2005, ocupa el puesto 5º respecto a la ayuda total del periodo (2,78%) y la segunda posición si nos centramos en el bienio 2006/07, con el 6,65% de la ayuda. Por su parte, ***Mozambique***, con un IDH de 0,384, se sitúa en el puesto 12 de total del periodo (noveno en el bienio 2006/07) con el 2,06% de la ayuda (2,74% de la ayuda del bienio mencionado). Es digno de mención el caso de ***Níger***, que ocupa la novena posición para el periodo 2000/07 con el 2,27% de la ayuda y analizando el bienio 2006/07 pasa a ocupar la posición 45, con únicamente el 0,54% de la ayuda del periodo.

Sierra Leona, con el menor IDH mundial, el 0,329, ocupa la posición 78 del ranking de países receptores para el total del periodo, recibiendo un 0,3% de la ayuda a SSB. Si analizamos el bienio 2006/07, aunque su posición en el ranking se mantiene en el 78, aunque en porcentaje de ayuda empeora ya que solamente recibe el 0,18%.

La ***República Centrofricana***, con un IDH de 0,352, ocupa el puesto 84 para el total del periodo con el 0,15% de la ayuda en SSB, empeorando al puesto 87 para el bienio 2006/07, si bien mejora en porcentaje de ayuda, obteniendo el 0,24%.

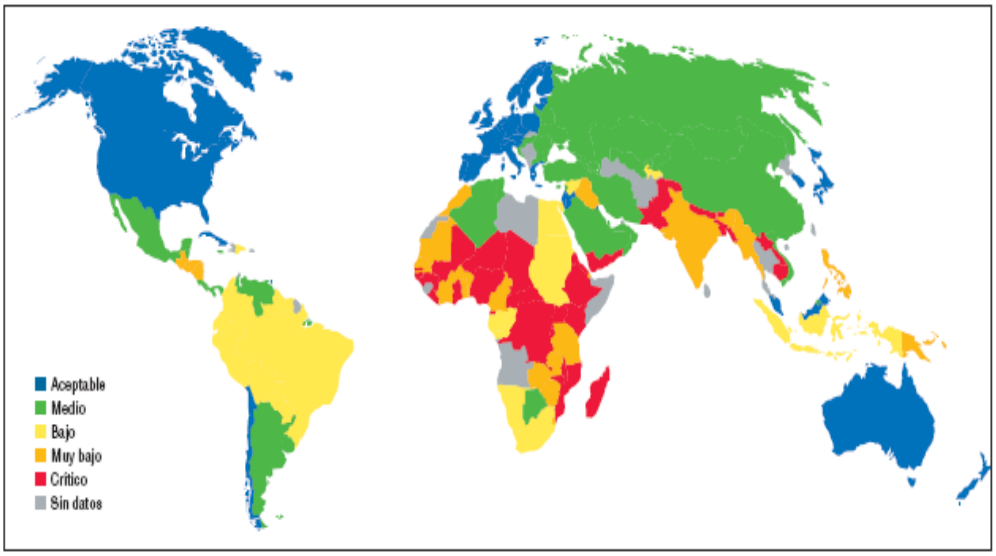
Centrándonos en el total del período, el resto de los países con peor IDH (menos de 0,4) se mueven entre la posición 24, que ocupa la República Democrática del Congo, con un 1,28% de la ayuda en SSB de periodo y la posición 105 de Liberia, con únicamente el 0,17% de la ayuda. Si analizamos las posiciones ocupadas analizando únicamente el bienio 2006/07, no hay diferencias significativas en las posiciones (de la 17 a la 96) y los porcentajes obtenidos por países (del 1,8% al 0,08%)

Este análisis por países permite concluir que los destinatarios prioritarios de la ayuda, en la mayoría de los casos, distan mucho de ser los países con menor desarrollo humano, al menos según el indicador analizado.

4.4. Análisis de la AOD recibida por los países receptores para la cobertura de SSB comparado con el Índice de Capacidades Básicas

Los números del ICB correspondientes a 2007 revelan que la mitad de los países tienen valores de ICB bajo, muy bajo o crítico, y que 25 países están en una situación crítica (Tabla 9). En el mapamundi que se ofrece en el gráfico 7 comprobamos como es el continente africano el que incluye los países con valores más bajos en este índice de desarrollo (crítico y muy bajo).

Gráfico 7: Análisis geográfico del ICB en el mundo



Fuente: Social Watch 2007

Si analizamos el valor del ICB de los países en desarrollo y su evolución (tabla 9) comprobamos que algunos están sufriendo un retroceso importante (Chad, Lesotho, Benin, Suriname, etc.) o leve (Camboya, Guinea Ecuatorial, Zimbawe) en su ICB, lo cual es muy preocupante.

Tabla 9: Valores del ICB por países y su evolución Fuente: Social Watch 2007.

TABLA 1 . ICB y evolución según el nivel de ICB de los países

NIVEL CRÍTICO			NIVEL MUY BAJO			NIVEL BAJO			NIVEL MEDIO			NIVEL ACEPTABLE		
País	ICB	Evolución ICB	País	ICB	Evolución ICB	País	ICB	Evolución ICB	País	ICB	Evolución ICB	País	ICB	Evolución ICB
Gambia	69,7	→	El Salvador	79,8	→	Egipto	89,8	→	Em. Árabes Unidos	97,9	=	Japón	99,9	→
Bhután	69,1	=	Iraq	79,0	→	Brasil	89,1	←	Croacia	97,4	=	Suecia	99,9	→
Lesotho	68,5	←	Djibouti	78,3		Tuvalu	88,7	=	Brunei Darussalam	97,4	=	Alemania	99,8	→
Guinea	68,4	→	Marruecos	78,1	→	Colombia	88,7	→	Lituania	97,4	=	Austria	99,8	→
Kenya	68,3	→	S. Tomé y Príncipe	77,8		Kiribati	88,1	=	Hungría	97,3		Bélgica	99,8	→
Eritrea	66,9	→	Filipinas	77,3	→	Sudáfrica	87,2	=	Belarús	97,2	=	España	99,8	→
Ghana	66,4	=	Swazilandia	76,9	=	Siria	87,0	←	Luxemburgo	97,1	←	Francia	99,8	→
Mali	65,8	→	Côte-d'Ivoire	76,9	→	Maldivas	86,4	→	Omán	97,0	→	Grecia	99,8	→
Yemen	63,8	←	Zimbabwe	76,3	←	Perú	86,0	→	Qatar	97,0	=	Suiza	99,8	→
Madagascar	63,5	→	Honduras	76,3	→	Namibia	85,8	←	Samoa	97,0	=	Dinamarca	99,8	=
Uganda	63,0	→	Mauritania	75,3	→	Tayikistán	85,7	=	Bulgaria	96,9	→	Finlandia	99,8	=
Nigeria	62,8	=	Zambia	74,6	→	Paraguay	85,5	→	Rusia, Fed.	96,8	=	Nueva Zelandia	99,8	→
Malawi	62,8	=	Nicaragua	74,0	→	Suriname	85,4	←	Ucrania	96,8	→	Irlandia	99,8	=
Mozambique	60,8	→	Benin	73,3	←	Vanuatu	85,1	←	Macedonia	96,5	=	Israel	99,8	=
Pakistán	60,4	→	Myanmar	73,1	→	Rep. Dominicana	84,9	←	Rumania	96,3	=	Portugal	99,7	→
Camboya	59,1	←	Papua Nueva Guinea	72,9	→	Indonesia	84,9	→	Costa Rica	96,2	=	Noruega	99,7	=
G. Ecuatorial	58,9	←	Comoras	72,5	=	Gabón	81,9	→	Cisjordania y Gaza	96,0	=	Irlanda	99,7	=
Lao, RDP	58,0	→	Senegal	72,2	→	Guyana	81,2	←	Moldova	96,0	←	Países Bajos	99,7	=
Bangladesh	57,1	=	Guatemala	71,7	→	Sudán	81,0	←	Santa Lucía	95,9	←	Polonia	99,6	=
Burundi	56,4	→	India	71,3	→	Ecuador	80,8	=	Líbano	95,9	→	Malta	99,6	=
Nepal	54,8	→	Burkina Faso	71,1	→	Bolivia	80,2	→	Mongolia	95,8	→	Chipre	99,6	=
Niger	54,6	=	Togo	70,2	=	Congo, Rep.	80,0		Kirguistán	95,7	=	Reino Unido	99,5	→
Rwanda	51,3	→	Camerún	70,2	←				Armenia	95,6	→	Rep. Checa	99,3	=
Etiopía	50,3	→	Tanzania	70,0	→				Uruguay	95,4	=	Chile	99,3	=
Chad	43,0	←							San Vic., y Granadinas	95,3	→	Bahrein	99,2	=
									Albania	95,1	←	Canadá	99,2	→
									Georgia	95,0	=	Estados Unidos	99,2	→
									Kazajstán	94,9	=	Corea, Rep.	99,2	=
									Argelia	94,8	=	Estonia	99,2	=
									Arabia Saudita	94,7	=	Eslovenia	99,1	→
									Túnez	94,6	→	Australia	99,1	→
									Jamaica	94,6	→	Letonia	99,0	→
									Venezuela	94,3	=	Cuba	99,0	=
									Dominica	94,3	=	Kuwait	98,7	→
									Marshall, Islas	94,1		Italia	98,7	=
									México	94,0	→	S. Cristóbal y Nevis	98,7	→
									Tonga	94,0	=	Fiji	98,6	→
									Argentina	93,7	←	Mauricio	98,6	=
									China	93,0	→	Jordania	98,6	→
									Granada	92,3	←	Barbados	98,5	→
									Cabo Verde	92,1	=	Malasia	98,1	=
									Turquía	91,6	→	Trinidad y Tobago	98,0	=
									Panamá	91,5	=			
									Irán	91,3	←			
									Azerbaiyán	91,2	=			
									Botswana	91,1	←			
									Belice	91,0	→			
									Viet Nam	90,0	→			

Referencias: ← Retroceso significativo ← Retroceso leve = Estancamiento → Avance leve → Avance significativo

Si añadimos la información acerca del valor del ICB de estos países al porcentaje de ayuda que reciben para SSB en el periodo, comprobamos que en el grupo de los países que acaparan el 50% de la ayuda (Tabla 10) la mayoría son países con un nivel crítico en el Índice de Capacidades Básicas (Nigeria, Bangladesh, Etiopía, Kenia, Níger, Uganda, Mozambique, Pakistan y Eritrea), o muy bajo (India, Tanzania, Zambia, Irak y Marruecos).

Habría que resaltar que también podemos encontrar algún país cuyo ICB se sitúa en nivel medio, como son Vietnam o Kazajstán.

Tabla 10: Distribución geográfica de los compromisos de AOD total (bilateral y multilateral) a SSB por países receptores (millones de USD constantes 2006) y su ICB

		PERÍODO 2000/2007				BIENIO 2006/2007		
PAÍS	ICB	TOTAL	%	% acum.	TOTAL	Media anual	%	% acum.
India	71,3	7.728,4	7,7%	7,7%	3.232,8	1.616,4	12,5%	12,5%
Nigeria	62,8	3.352,3	3,3%	11,0%	950,1	475,0	3,7%	16,2%
Bangladesh	57,1	3.151,8	3,1%	14,2%	505,6	252,8	2,0%	18,1%
Tanzania	70,0	2.953,5	2,9%	17,1%	852,2	426,1	3,3%	21,4%
Etiopía	50,3	2.798,4	2,8%	19,9%	1.720,7	860,3	6,7%	28,1%
Hong Kong, China	..	2.758,1	2,7%	22,6%	0,0	0,0	0,0%	28,1%
Zambia	74,6	2.433,0	2,4%	25,0%	653,8	326,9	2,5%	30,6%
Kenia	68,3	2.424,0	2,4%	27,4%	1.023,1	511,5	4,0%	34,5%
Níger	54,6	2.282,4	2,3%	29,7%	139,6	69,8	0,5%	35,1%
Uganda	63,0	2.115,8	2,1%	31,8%	615,4	307,7	2,4%	37,5%
Sudáfrica	87,2	2.086,8	2,1%	33,9%	835,7	417,8	3,2%	40,7%
Mozambique	60,8	2.073,9	2,1%	36,0%	709,8	354,9	2,7%	43,4%
Indonesia	84,9	1.971,9	2,0%	37,9%	614,2	307,1	2,4%	45,8%
Iraq	79,0	1.774,3	1,8%	39,7%	491,1	245,5	1,9%	47,7%
Pakistán	60,4	1.744,2	1,7%	41,4%	792,1	396,0	3,1%	50,8%
Eritrea	66,9	1.717,6	1,7%	43,1%	54,8	27,4	0,2%	51,0%
Afganistán	..	1.659,6	1,7%	44,8%	668,6	334,3	2,6%	53,6%
Somalia	..	1.538,2	1,5%	46,3%	114,0	57,0	0,4%	54,0%
Kazakstán	94,9	1.529,9	1,5%	47,8%	17,7	8,9	0,1%	54,1%
Marruecos	78,1	1.382,5	1,4%	49,2%	228,6	114,3	0,9%	55,0%
Vietnam	90,0	1.338,4	1,3%	50,5%	388,7	194,3	1,5%	56,5%
Total especificado		**100.555,9**	**100,0%**	**100,0%**	**25.872,7**	**12.936,3**	**100,0%**	**100,0%**
Total no especificado		**28.529,9**						
TOTAL		**129.085,8**						

Fuente: Elaboración propia a partir de la base de datos CRS de la OCDE y el Informe del PNUD.

A pesar de ello, si analizamos la situación en el ranking anterior de los países con más bajo nivel de ICB (Tabla 9), salvo Bangladesh, que se sitúa en tercera posición del ranking el resto (Chad, Ruanda, Nepal, Burundi, etc.) se posiciona en puestos superiores al 30, no superando en ningún caso el 1% de la ayuda, lo cual indica que persiste un grupo numeroso de países con reducido valor de ICP a los que apenas se dirige la AOD para SSB.

Según el último Informe de Desarrollo Humano sin un aceleramiento sustancial en el actual ritmo de progreso, en 2015 el valor promedio del ICB en los países de Asia Meridional y de África Subsahariana apenas corresponderá al nivel muy bajo y todas las demás regiones, excepto Europa y América del Norte, aún estarán lejos de un nivel aceptable.

5. CONCLUSIONES

Del estudio analítico efectuado sobre la cooperación internacional en materia de SSB se extraen las siguientes conclusiones:

- La AOD total (bilateral y multilateral) destinada a SSB ha mantenido una tendencia creciente en el período 2000-2007, multiplicándose por dos, la cuantía de dicha ayuda, superando los 16.000 millones de USD (constantes de 2006) en 2007. La AOD bilateral la principal donante en esta materia (70%).
- Tanto la AOD bilateral (en promedio), como la multilateral han alcanzado, el compromiso adquirido en el Consenso de Oslo de dedicar al menos el 20% de la ayuda (sobre el total de AOD distribuible por sectores) a SSB.
- Por parte de la AOD bilateral, EEUU y Reino Unido concentran casi el 50% de la AOD concedida para SSB en el periodo 2000-07. Si unimos las aportaciones de Holanda, Alemania, Canadá y Japón alcanzamos el 70% de la ayuda.
- Sin embargo, no necesariamente los principales donantes son los que conceden más ayuda para SSB. Países que ocupan posiciones rezagadas en el ranking de donantes (en cuanto a aportación a la AOD total bilateral) son los más concienciados con la necesidad de priorizar la cobertura de las necesidades sociales básicas en su cooperación (Irlanda, Luxemburgo, Bélgica, Noruega).
- Analizando los organismos multilaterales los principales donantes han sido la *Comisión Europea*, (CE), el *Fondo Global contra el SIDA, Tuberculosis y Malaria*, la *Asociación Internacional para el Desarrollo* (ADI), y *UNICEF*.
- La ayuda ha ido dirigida a *Políticas y programas sobre población y salud reproductiva* en primer lugar (39%), seguida de *Salud básica* (27%), *Educación Básica* (27%) y en las últimas posiciones *Abastecimiento y depuración de agua, sistemas menores* (8%) *y Ayuda multisectorial a SSB* (6%).
- Existe una gran variabilidad en la cuantía de AOD anual recibida para SSB por cada país en desarrollo en el período analizado. Esto dificulta la identificación de los países prioritarios de la cooperación internacional en SSB, así como la realización de planes y políticas para el desarrollo a largo plazo por parte de los países receptores.
- Existe una elevada concentración de la ayuda en cuanto a los países socios (receptores) de la misma. Cinco países socios (de 177): India, Nigeria, Bangladesh, Tanzania y Etiopía concentran el 20% del total de AOD en el período 2000-2007 (cuyo destino es especificado). El 50% de la ayuda se concentra en tan solo el 12% de los países.
- Aunque entre los principales países receptores de AOD para SSB figuran algunos como Tanzania, Nigeria, Etiopía y Mozambique, con valores reducidos en su IDH, existe aún un grupo numeroso de países con IDH muy bajo que apenas reciben ayuda para la cobertura de SSB: Chad, Guinea-Bissau, Burkina Fasso, Sierra Leona, etc.
- La mayoría de los principales países receptores de AOD para SSB tienen un nivel crítico en el Índice de Capacidades Básicas (Nigeria, Bangladesh, Etiopía, Kenia, Níger, Uganda, Mozambique, Pakistan y Eritrea), o muy bajo (India, Tanzania, Zambia, Irak y Marruecos).
- A pesar de lo anterior, persiste un grupo numeroso de países (Chad, Ruanda, Nepal, Burundi, etc.) con reducido valor de ICP que apenas reciben AOD para SSB.
- La evolución de cada país en el ICB desde 2000 revela que casi la mitad de los países ha logrado avances. Sin embargo, 54 países (o 34%) han retrocedido. Los países que más han retrocedido en el ICB son mayoritariamente de África Subsahariana. Algunos países de Asia Oriental y el Pacífico y de América Latina y el Caribe también han sufrido el deterioro de sus capacidades básicas. El retroceso ha sido significativo (más de 5% en el valor del ICB) en siete casos. La situación es preocupante, ya que estos países se encuentran retrocediendo desde un nivel de ICB bajo, muy bajo e incluso crítico. Esto significa que se agrava el rezago de algunos países, aumentando la brecha existente con el resto del mundo.

Por lo tanto, aunque podemos afirmar que se aprecia en los últimos años una orientación de la AOD para la cobertura de los SSB, aún existe un grupo de países en una situación muy desfavorecida, (mayoritariamente del continente africano) que apenas reciben ayuda de la cooperación internacional para lograr un nivel de desarrollo mínimo en la satisfacción de sus

necesidades básicas. Por lo tanto la cooperación internacional debe seguir trabajando para incluir a dicho grupo dentro de sus prioridades geográficas.

BIBLIOGRAFÍA

KOOS N.M. MICHELLE (1998), Hanoi Conference on the 20/20 Initiative.
NACIONES UNIDAS, ASAMBLEA GENERAL (2000), Nuevas iniciativas en pro del desarrollo social.
NACIONES UNIDAS, ASAMBLEA GENERAL (2005), Documento Final de la Cumbre 2005.
NACIONES UNIDAS (2000), Objetivos del Milenio
NACIONES UNIDAS (1995), Informe de la cumbre mundial sobre desarrollo social.
OECD (2006), Analysis of aid for Basic Social Sevices 1995-2004
OCDE (2000), Measuring aid to Basic Social Services, etc.
PNUD (1992), Informe de desarrollo humano, 1992.
PNUD (2009) Human Development Report 2008-09.
WILDEMAN CAROLINA, La iniciativa 20/20 es un trampolín.

PÁGINAS WEB

http://www.cinu.org.mx/eventos/conferencias/johannesburgo/wssd.htm
http://www.cinu.org.mx/temas/desarrollo/desssocial/cumbre/cophenage.htm

EFEITOS DOS PROCESSOS DE INTEGRAÇÃO ECONÓMICA EM DUAS REGIÕES PERIFÉRICAS DA UNIÃO EUROPEIA

MELCHOR FERNANDEZ[*]
E-mail: melchor.fernandez@usc.es
ROBERTO BANDE[*]
E-mail: roberto.bande@usc.es
Universidade de Santiago de Compostela e
Instituto de Estudos e Desenvolvimento de Galiza (IDEGA)
RAQUEL PEREIRA[*]
E-mail: rpereira@iscap.ipp.pt
Instituto Superior de Contabilidade e Administração do Porto e
Universidade de Santiago de Compostela

Resumo

As regiões Norte e da Galiza são duas regiões europeias que, desde a adesão, constituem regiões de objectivo 1 da Política Regional e de Coesão (PRC). Estas regiões, ao partilharem fronteiras, apresentam características geográficas, históricas, culturais e linguísticas muito semelhantes mas também dinâmicas endógenas próprias e diferentes níveis de evolução e desempenho económico, quer na convergência com a média nacional quer com os níveis médios europeus. A análise que nos propomos realizar tem por objectivo analisar e descrever, a evolução económico-social, destas duas regiões, no sentido de compreendermos as suas trajectórias de desenvolvimento e os seus factores de convergência/divergência. A análise indica que o processo de integração económica, ainda que muito elevado, não favoreceu um dinamismo comum destas regiões periféricas mas os espaços centrais das economias nacionais de referência. Isto pode servir de base para questionar a eficácia PRC já que, no interior dos Estados-membros, as disparidades aumentaram.

Palavras-chave: Desempenho económico, Galiza, Norte de Portugal.
Área Temática: Economia Internacional e da Integração.

Abstract

The North and the Galiza regions are two European regions that, since the integration, constitute objective 1 regions to the Cohesion and Structural Policy (CSP). These regions share the same borders and present very similar geographic, historical, cultural and linguistic characteristics. Nevertheless, they have its own endogenous dynamic and different levels of evolution and economic performance, either in terms of convergence with the national or with the European average levels. The aim of this paper is to analyze and to describe the economic and social performance of these regions and, in doing so, to understand the patterns of growth and its factors of convergence/divergence. This analysis indicates that the economic integration, despite its very strong, didn't enhance a common process of dynamism. In contrast, it favored the central regions of the national economies. This argument serves to question the effectiveness of the CSP now that, within the member states, the disparities have increased.

Key words: Economic performance, Galiza, North of Portugal.
Thematic Area: International and Integration Economics.

[*]Os autores agradeçam o apoio financeiro concedido pela Xunta de Galicia no âmbito do projecto INCITE 07TUR005242PR. O presente trabalho enquadra-se no projecto de tese de Doutoramento de Raquel Susana da Costa Pereira. Os autores agradecem os comentários de outros membros do GAME- Research Group (Group of Analysis and Modelling in Economics) e IDEGA (Universidade de Santiago) a uma versão prévia deste trabalho. Naturalmente, os autores são os únicos responsáveis pelas opiniões expressas e pelos eventuais erros cometidos.

INTRODUÇÃO

Desde a sua implementação, a Política Regional e de Coesão (PRC) rege-se pelo princípio fundamental de que a União Europeia (UE) deve promover um desenvolvimento harmonioso de toda a União pela redução das disparidades entre as regiões e pelo reforço da coesão económica e social. As regiões/países beneficiários dos recursos provenientes da PRC, em particular os países da coesão atingiram progressos consideráveis permitindo-lhes a sua aproximação aos níveis médios da UE15. No entanto, e em virtude das suas diferentes condições de atractividade e recursos, também é certo que nem todas as regiões/espaços dentro dos países atingem um mesmo nível de desenvolvimento. Sabemos que há regiões com maior taxa de crescimento, níveis de desemprego mais baixo e níveis de Produto Interno Bruto *per capita* (PIBpc) mais elevados. Nos relatórios sobre a Coesão Económica e Social, (2007, 2004), é referido que, apesar da importante aproximação e convergência dos Estados menos desenvolvidos (sobretudo os da coesão) verificou-se também um aumento das disparidades dentro dos Estados-membros. Portanto a questão fundamental a que pretendemos dar resposta é: será que a região Norte de Portugal e a região da Galiza em Espanha conseguiram reduzir a "distância" que as separava das respectivas médias nacionais? Com efeito, é muito difícil medir adequadamente a convergência real entre regiões. Não obstante, como podemos resolver um problema se não somo capazes de quantificar a sua magnitude e estudar a sua evolução? Portanto, em qualquer análise de convergência, o primeiro passo deve ser quantificar o problema e mostrar qual foi a sua evolução nos últimos anos identificando a evolução de todas as variáveis relevantes no processo de convergência regional.

Porquê o Norte e a Galiza? Apesar de estes territórios estarem divididos por fronteiras políticas, determinando diferentes nacionalidades aos seus habitantes, estas são regiões vizinhas com características geográficas, históricas, culturais e linguísticas muito semelhantes e além disso, constituem duas regiões atrasadas e periféricas no contexto da UE15 apresentando problemas e dificuldades semelhantes no que se refere à competitividade, desenvolvimento e convergência com os níveis europeus. No entanto, e apesar dos respectivos países terem aderido à UE no mesmo ano e de ambas constituírem regiões do objectivo 1, no âmbito da PRC, as duas regiões apresentam uma dinâmica endógena própria determinando, ao longo dos últimos anos, diferentes níveis de desempenho e desenvolvimento. Assim, a análise que nos propomos realizar resume-se a um objectivo simples que é descrever, de forma resumida, a evolução das duas regiões ao nível dos principais indicadores económicos e sociais e tentar compreender as suas trajectórias de desenvolvimento, assim como os seus factores de convergência/divergência. Em particular, interessa-nos mostrar a sensibilidade dos resultados da PRC ao âmbito geográfico de aplicação e verificar até que ponto estes dependem mais da dinâmica nacional do que da eficácia das próprias medidas.

Desta forma, o trabalho organiza-se como se descreve em seguida. Começamos apresentando um conjunto de dados estatísticos através dos quais pretendemos estabelecer o posicionamento e evolução das duas regiões, no contexto nacional e, sempre que possível, no contexto da UE15 como procuramos, tentando, dessa forma, aferir se, nas últimas décadas, as duas regiões conseguiram reduzir a distância que as separava não só dos níveis médios europeus como também dos níveis médios nacionais. Procuraremos efectuar a análise para o período de tempo mais alargado possível, conforme a disponibilidade dos dados. Adicionalmente, e no sentido de compreender o nível de desenvolvimento actual de cada uma das regiões, bem como a sua importância económica dentro do país, recorremos à análise da produtividade, de indicadores de localização e ao método "*shift-share*". Por último apresentamos as conclusões mais pertinentes.

1. DEMOGRAFIA E RECURSOS HUMANOS.

O Norte e a Galiza formam uma euro-região periférica da UE com baixos níveis de desenvolvimento, importantes semelhanças no âmbito sociocultural mas também acentuadas diferenças, sobretudo nos perfis demográficos e produtivos. Ambas as regiões apresentam graves assimetrias intra-regionais a favor das áreas costeiras e litorais, mostrando-se estas áreas

mais dinâmicas em termos de emprego e produção, enquanto as áreas interiores mostram uma tendência à desertificação[1].

Tabela1. Dados sobre a população

Dados sobre a população	UE15		Espanha		Portugal		Galiza		Norte	
Crescimento da população, entre 91-01	n.d.		5,1%		5%		-1,3%		6,2%	
Crescimento da população, entre 01-07	6,9%		8,9%		2,3%		1%		1,5%	
Densidade populacional (em 2006)	n.d.		87,2		114,9		91,8		163,4	
Estrutura da população (em % do total)	**97**	**07**	**97**	**07**	**97**	**07**	**97**	**07**	**97**	**07**
Menos de 15 anos	17,4	n.d.	16,4	14,5	17,6	15,4	14,4	11,4	19,4	16,1
15-65	67	n.d.	68,2	68,8	67,7	67,3	67,5	67,3	68,5	68,8
Mais de 65 anos	15,6	n.d.	15,4	16,7	14,7	17,3	18,1	21,3	14	15,1

Fonte: Elaboração própria a partir dos dados recolhidos em INE e IGE (2005) para dados entre 1991 e 2001. Os dados de 2007 foram obtidos *on-line* nos sítios http://epp.eurostat.cec.eu.int (em Fevereiro de 2009).

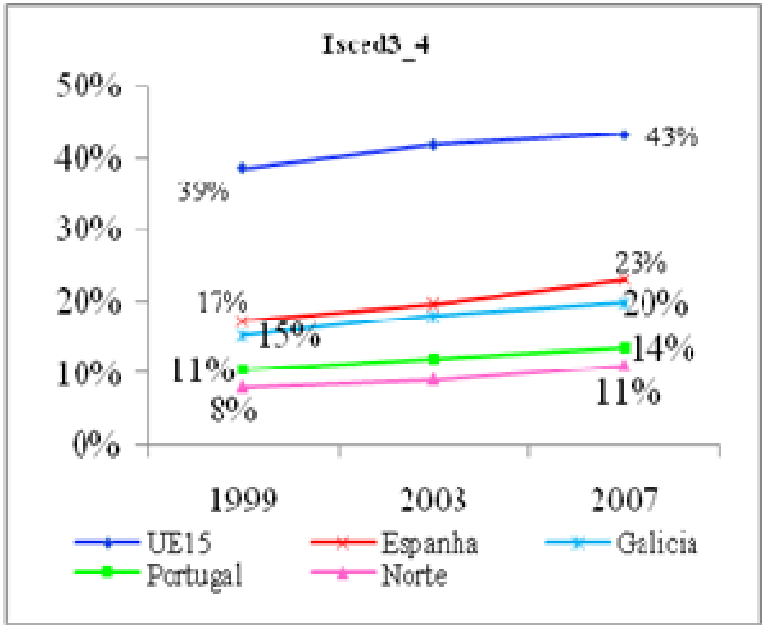
Em relação às tendências demográficas, enquanto no período 1991-2001 se verifica uma divergência entre as duas regiões, no que se refere à evolução da população (crescimento negativo na Galiza e crescimento acima da média nacional no caso da região Norte), nos anos mais recentes observa-se uma alteração da tendência do período anterior. No período 2001-2007, a população da região Norte cresce a um ritmo inferior à média nacional o mesmo acontecendo na região Galega que, apesar de tudo, volta a verificar uma taxa de crescimento positiva. Com efeito, esta tendência ao envelhecimento, que também se pode observar analisando a população por estrutura de idades, é muito mais acentuado na região da Galiza, quer em comparação com a média nacional quer com a da região Norte.

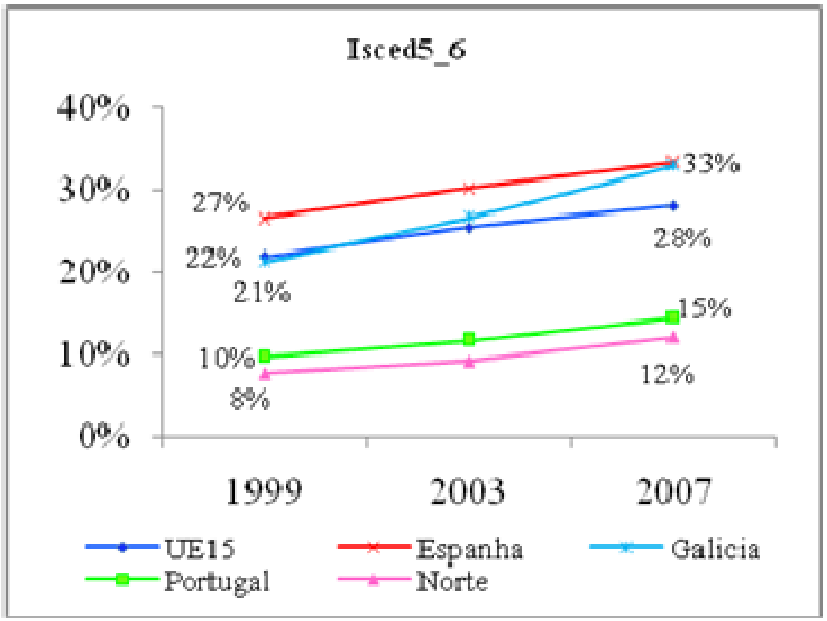
No que se refere aos recursos humanos, a qualificação dos recursos humanos constitui um dos factores chave na análise da competitividade de uma região. Tanto a região Norte como a Galiza mostram uma tendência ao envelhecimento da população o que, a longo prazo, pode afectar negativamente o crescimento, uma vez que o empreendedorismo, a capacidade de inovação e a vitalidade dessa população tende a decrescer. Uma forma de contrariar isso passará pelo aumento da produtividade individual, bem como a acumulação de outros factores produtivos como o capital físico e a tecnologia (Dehesa, 2000, p.20). O aumento da produtividade dos trabalhadores passa pela melhoria da formação e educação dos recursos humanos[2], o que facilita a (re)entrada no mercado de trabalho e possibilita melhores condições de trabalho e rendimentos mais elevados. Assim o considera a CE (2007), que refere que o aumento da qualidade do capital humano explica mais de metade dos ganhos de produtividade da última década.

[1] Tanto no caso da Galiza como na região Norte, são as zonas interiores (Lugo e Ourense na Galiza e Trás-os-Montes e Alto Douro no Norte) onde mais predominam as actividades agrícolas de pequena dimensão, pouco competitivas e, muitas delas, de cariz familiar as quais têm vindo a verificar uma crescente desertificação.

[2] Por exemplo, Blondal e al. (2002), De la Fuente (2003), Engelbrecht (1997) verificam que níveis mais elevados de educação favorecem o crescimento da produtividade. Contudo, a importância do capital humano não se restringe aos aumentos da produtividade do trabalho. O investimento em capital humano tem sido apontado como uma das estratégias fundamentais de política económica na medida em que é determinante para promover a prosperidade económica, o emprego e a coesão social. Para uma análise mais detalhada dos efeitos do capital humano ver, por exemplo, Bassanini e Scarpetta (2001), Benhabib e Spiegel (1994, 2003), OECD (1998), Teixeira (1999), entre outros.

Gráfico 1. Nível educacional da população activa, entre 25-64, em % da população activa com mais de 25 anos e de acordo com classificação internacional ISCED[3]





Fonte: Elaboração própria a partir dos dados recolhidos *on-line* nos sítios http://epp.eurostat.cec.eu.int (em Fevereiro de 2009).

[3] International Standard Classification of Education (ISCED97) – corresponde à classificação Internacional Normalizada da Educação, comummente denominada ISCED, e trata-se de um instrumento de referência da UNESCO que permite a harmonização e comparabilidade das estatísticas educativas. Distingue sete níveis educativos: **ISCED 0**-ensino pré-escolar; **ISCED 1-** ensino básico (1º e 2º ciclo); **ISCED 2** - ensino básico (3º ciclo); **ISCED 3** - ensino secundário; **ISCED4** - ensino pós-secundário; **ISCED5** - ensino superior (bacharelato e licenciatura); **ISCED 6** - ensino superior (mestrado, doutoramento e pós-doutoramento)

A região Norte de Portugal, apesar de ser uma das regiões periféricas da UE15 mais populosas e apesar da sua relativa juventude, apresenta como grande desvantagem a qualificação dos seus recursos humanos. Pela análise do gráfico 1, constata-se que a região Norte acompanha a evolução nacional verificada nos últimos anos mas, tanto os níveis de educação média como superior continuam a situar-se abaixo da média nacional sendo também consideravelmente inferiores aos valores da Galiza. No que respeita à Galiza é de salientar a evolução positiva verificada que, no que respeita à educação superior, apesar de em 1999 ter valores inferiores à média europeia e nacional, eliminou as divergências que tinha.

Que razões justificam estas diferenças tão significativas entre as duas regiões? Será que é um problema de eficiência e elevados índices de abandono escolar? Será que as políticas de ensino e formação profissional estão desajustadas às necessidades estruturais?

Embora não estejam disponíveis os dados desagregados por região, de acordo com os últimos dados disponíveis do Eurostat, Portugal tem um investimento público (em % do PIB) acima da média espanhola (5,37% contra 4,66% em 1995 e 5,40% contra 4,23% em 2005). Adicionalmente, o abandono escolar precoce em Portugal é muito superior ao de Espanha e ao da média europeia (40,1% contra 31,4% 21,6%, respectivamente, em 1996, e 36,3%, contra 31,0% e 16,4% em 2007).

Portanto, o problema poderá não estar relacionado com a falta de investimento mas com o sistema de ensino e a sua própria eficiência. Com efeito, embora seja necessário salvaguardar as dinâmicas positivas verificadas nos últimos anos, tais como o aumento do número de anos de escolaridade e, sobretudo, o aumento da oferta de formação e aumento das taxas de escolarização da população jovem, o que é facto é que a região Norte continua a destacar-se por baixos valores nos indicadores de escolarização. Outros problemas na região Norte que acabam por influenciar os baixos índices de formação e educação são, para além da centralização e insuficiente qualidade de oferta formativa "a baixa procura social, empresarial, individual, e de qualificações" determinada pelo modelo produtivo que domina na região (baseado nos sectores tradicionais, de baixo valor acrescentado e intensivos em mão de obra barata). De facto, embora apenas 15% da população (entre 25-64 anos) possua formação superior ou universitária (em 2007), as oportunidades de emprego nas áreas científicas e tecnológicas oferecidas na região Norte, e que exigem recursos humanos mais qualificados, têm sido insuficientes para absorver esses recursos humanos que acabam, na maior parte dos casos, a emigrar para outras regiões do país ou até para estrangeiro. Assim o refere a Comissão de Coordenação e Desenvolvimento Regional do Norte (CCDRN) (2006, p. 32):

"(….) a região Norte parece estar a aproveitar menos bem o investimento realizado em matéria de qualificação da população indiciando uma menor capacidade da procura para absorver de forma produtiva a força de trabalho mais jovem e escolarizada."

2. EVOLUÇÃO DO MERCADO DE TRABALHO

Observando a tabela 2, e no que se refere à taxa de actividade global, tanto a região Norte como a região da Galiza, em 1987, detinham taxas de actividade superiores à média nacional. No entanto, essa diferença foi-se reduzindo ao longo do tempo. Para tal contribui fortemente a subida da taxa de actividade das mulheres que, ao nível médio nacional, teve um crescimento muito superior ao verificado a nível regional (ver gráfico 2). No caso de Espanha é ainda de referir também os fortes fluxos de imigração verificados nos últimos anos e que não foram tão acentuados na região galega, a qual, pelo contrário, tem verificado emigração líquida com as outras comunidades autónomas espanholas.

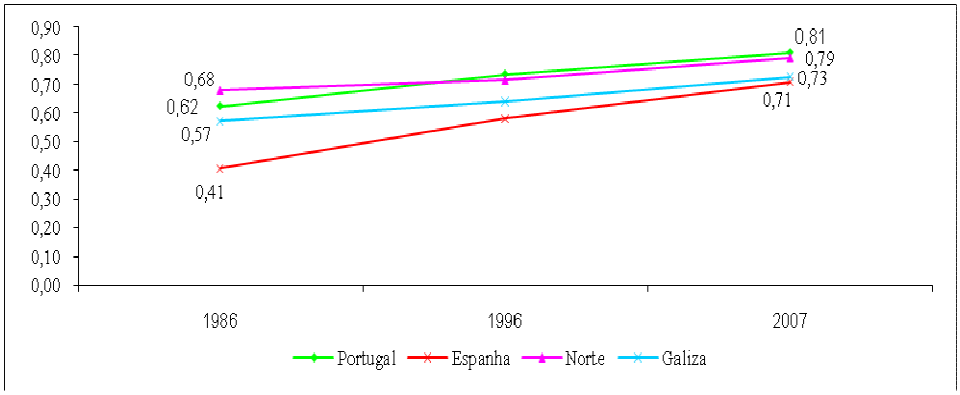
Tabela 2 . Alguns indicadores do mercado de trabalho

	Taxa de emprego		Taxa de actividade		Taxa de desemprego	
	1988	2007	1987	2007	1987	2007
UE15*	61	n.d.	n.d.	n.d.	10,5	7
Portugal	67,2	68,2	57,5	62,6	7	8
Norte	70,1	66,3	61,7	63,1	4,9	9,4
Espanha	41,3	54	46,8	58,2	20,8	8,3
Galiza	56,8	50	51	53,7	13,4	7,6
Diferencial Portugal-Espanha	25,9	14,2	10,7	4,4	-13,8	-0,3
Diferencial Portugal-Norte	-2,9	1,9	-4,2	-0,5	2,1	-1,4
Diferencial Espanha-Galiza	-15,5	4	-4,2	4,5	7,4	0,7
Diferencial Norte-Galiza	13,3	16,3	10,7	9,4	-8,5	1,8

* Em 1987 e 1988 UE12

Fontes: CE (1999), Eurostat: http://epp.eurostat.cec.eu.int , www.ine.pt e www.ine.es.

Gráfico 2 . Taxa de actividade das mulheres vs homens

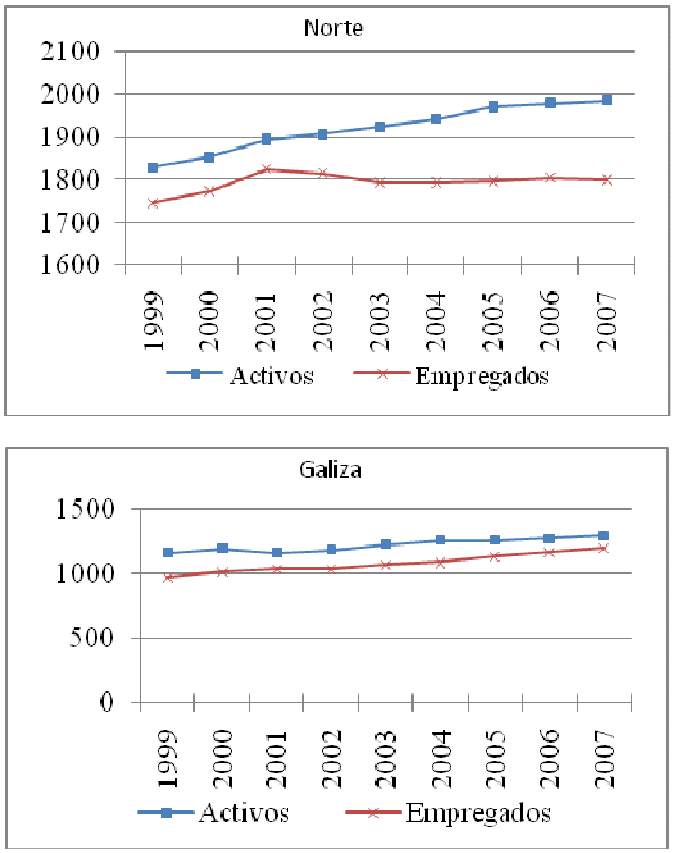


Fonte: Elaboração própria a partir dos dados recolhidos *on-line* em Eurostat: http://epp.eurostat.cec.eu.int em Fevereiro de 2009.

No que respeita à taxa de emprego, no período em causa, Portugal e a região Norte apesar de manterem uma taxa de emprego superior à média espanhola e da Galiza, nos anos mais recentes, verifica-se um maior dinamismo de Espanha relativamente a Portugal. É de salientar a evolução positiva da média espanhola que cresceu 12,7 pontos percentuais contra 1 ponto percentual em Portugal. Há ainda a referir a evolução negativa das duas regiões relativamente às médias nacionais. Em 1988, ambas as regiões detinham um diferencial positivo que passou a negativo em 2007. Com efeito, a criação de emprego nas duas regiões (tanto no período de 1986-1999 como em 1999-2007) é bastante inferior em comparação com as médias nacionais (ver quantificação da evolução do emprego na tabela 3).

Este fraco dinamismo do emprego, quando comparado com o crescimento do número de activos ajuda-nos a perceber o forte crescimento do desemprego na região Norte (gráfico 3). Desde início do século XXI que a evolução tem sido inversa à verificada na Galiza e também à evolução verificada nos períodos anteriores. A taxa de desemprego que, em 1987, era bastante reduzida e inferior à média nacional (4,9 e 7%, respectivamente), em 2007 era já superior à média do país tendo verificado um aumento de 4,5 pontos percentuais. De facto, nos anos mais recentes (no período 1997- 07) a criação de emprego tem sido muito inferior à taxa de crescimento do número de activos (3,1% contra 8,5%). Este facto ajuda a explicar a evolução da taxa de desemprego que tem sido tão desfavorável que passou a situar-se acima da média nacional e também da Galiza. Com efeito, além dos custos salariais terem crescido a ritmo superior ao da produtividade, que favoreceu a deslocalização de multinacionais para mercados mais competitivos, a globalização e liberalização do comércio mundial, e consequente concorrência de outros mercados emergentes como a China, a Índia e alguns países do Magreb, bem como o alargamento da UE15 para UE27 determinaram uma concorrência acrescida sobretudo sobre os sectores tradicionais, os quais predominam na região Norte. Em resultado, a região tem registado uma evolução desfavorável e divergente no que respeita ao emprego, desemprego e riqueza gerada (não só em relação à média nacional como também em comparação com a Galiza e média da UE15).

Gráfico 3. Evolução do número de empregados e activos (em milhares)



Fonte: Elaboração própria a partir dos dados do Eurostat, informação recolhida *on-line* no sitio http://epp.eurostat.cec.eu.int em Fevereiro de 2009.

No que se refere à Galiza, a relação/diferença entre activos e empregados mantém-se estável desde 1999 não se observando aumentos extraordinários do número de activos em relação aos empregados. Pelo contrário, observa-se uma ligeira quebra nos anos mais recentes. Com efeito, no período em causa, a taxa de crescimento do emprego (22,9%) é muito superior à taxa de crescimento no número de activos (11,4%) o que tem ajudado à redução do desemprego que, no período 1987- 07, caiu 5,8 pontos percentuais. Contudo, apesar da evolução positiva, este dinamismo fica aquém do verificado a nível nacional em que o número de empregados aumentou 38,6%, o número de activos 27,3% e a taxa de desemprego caiu 12,5 pontos.

Que factores justificam esta evolução nas duas regiões? Será que as alterações sectoriais e outros factores estruturais ajudam a explicar estas trajectórias divergentes entre as duas regiões?

Tabela 3. Evolução do emprego (em milhares)

	Total		Agricultura		Indústria		Serviços	
	Galiza	**Norte**	**Galiza**	**Norte**	**Galiza**	**Norte**	**Galiza**	**Norte**
1986-91	0,9	96,8	-136	-89	46,1	78,6	92,2	106,7
1991-95	-88,4	-167,2	-37,1	-140,8	-30,6	-88,9	-20,7	64,9
1995-99	13,2	161,2	-93,4	39,2	42,5	155,7	64,1	-33,7
1986-99	-74,3	90,8	-266,5	-190,6	58	145,4	135,6	137,9
Variação 1986-99	**-7,4%**	**5,5%**	**-60,5%**	**-46,3%**	**27,1%**	**21,5%**	**38,5%**	**24,2%**
Variação nacional	**27,1%**	**14,3%**	**-41,8%**	**-32,8%**	**22,4%**	**10,1%**	**52,9%**	**33,9%**
1999-07	222,3	53,4	-69,4	2,8	79,6	-87,5	212,1	138,9
Variação 1999-07	**22,9%**	**3,1%**	**-40,2%**	**1,3%**	**27,9%**	**-11,0%**	**41,3%**	**18,8%**
Variação nacional	**38,6%**	**5,3%**	**-11,7%**	**-3,3%**	**31,5%**	**-6,6%**	**47,9%**	**15,1%**

Fonte: Para o ano de 1986 e 1999, Fernandéz e Polo (2001). Para o período 1999- 2006, elaboração própria partir da informação recolhida *on-line* no sitio http://epp.eurostat.cec.eu.int em Fevereiro de 2009.

A tabela 3 evidencia diferentes sub-períodos da evolução do emprego. Observamos que ambas as regiões conseguem criar emprego líquido nas fases expansivas acontecendo o inverso nas fases recessivas e de forma mais acentuada na região Norte. Adicionalmente, torna-se relevante comparar o período 1986-99 com o período 1999-07 uma vez que surgem mudanças significativas que parecem evidenciar o bom desempenho da economia espanhola relativamente à portuguesa e o da Galiza relativamente ao da região Norte. No período 1986-99, as variações que ocorrem nas duas regiões, em comparação com a média nacional, foram semelhantes nomeadamente as grandes perdas de emprego no sector agrícola superiores à média nacional, o crescimento abaixo da média no sector dos serviços e um crescimento acima da média no sector industrial. Contudo, no período 1999-07, embora o padrão das variações se mantenha no caso da Galiza, o mesmo não acontece no caso da região Norte. Ao contrário do que se verificou nos outros sub-períodos, o emprego líquido gerado foi bastante inferior ao da Galiza. Para além disso, verifica-se uma perda significativa do emprego no sector industrial e um aumento do emprego agrícola. Portugal, no período 1999-07 perdeu 6,6% do emprego industrial enquanto a região Norte perdeu 11% (valor acima da média) o que realça o peso deste sector para a região e, mais uma vez, realça as dificuldades e fragilidades deste sector em se adaptar às novas determinantes da concorrência mundial. A queda do emprego deste sector foi compensada, em parte, pelo crescimento do emprego no sector dos serviços e no sector primário. Contudo, e como já referimos a variação do emprego na região Norte ficou aquém da verificada a nível nacional, o que significa não convergência com a média nacional.

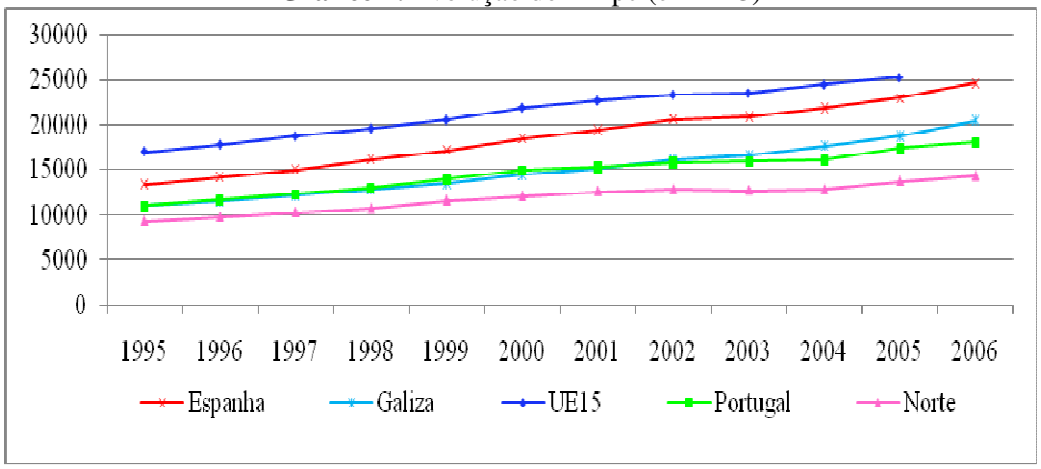
3. ESTRUTURA ECONÓMICA E COMPETITIVIDADE.

Já pudemos constatar que a região Norte e a região da Galiza, apesar dos elementos sócio-culturais e históricos comuns, apresentam também características diferenciadoras no contexto demográfico, qualificação dos seus recursos humanos e mercado de trabalho. Ao longo desta secção iremos também verificar que as diferenças estendem-se às suas estruturas produtivas e dinâmicas de evolução e crescimento.

Com uma estrutura produtiva fortemente exportadora mas baseada nos sectores tradicionais, intensivos em mão de obra, baixos índices de produtividade e uma competitividade ainda muito dependente dos baixos custos salariais, a região Norte que, desde a adesão até finais da década de noventa verificou níveis de crescimento e convergência acentuados, tem registado, nos anos mais recentes, uma estagnação no crescimento da riqueza. Este facto tem afectado, de forma considerável, os tradicionais baixos níveis de desemprego da região, bem como determinou uma dinâmica de afastamento dos níveis médios da UE15 e com a média nacional.

Por sua vez, a região da Galiza que, tal como a região Norte, apresenta uma estrutura produtiva pouco terciarizada, em termos de níveis de emprego, e em comparação com a média nacional e europeia, tem verificado um processo de crescimento sustentado permitindo-lhe uma progressiva aproximação aos níveis médios da UE15. Apesar de ter uma estrutura produtiva historicamente ligada ao sector primário, destacando-se o sector das pescas, após a adesão à UE a região da Galiza conseguiu desenvolver o seu sector industrial através de áreas de negócio que incorporam níveis tecnológicos mais avançados do que aqueles que caracterizam o sector industrial da região Norte, destacando-se a indústria automóvel (localização de unidades produtivas do grupo Citroën-Peugeout) a indústria da moda (grupo Inditex) e o sector do turismo.

Gráfico 4. Evolução do PIBpc (em PPC)



Fonte: Elaboração própria partir da informação recolhida *on-line* no sitio http://epp.eurostat.cec.eu.int em Março de 2009.

Tabela 4. Evolução do PIBpc em PPC (UE15 = 100)

	1986	1996	Diferença 96-86	2005	Diferença 05-96
Portugal	55	65	10	69	4
Espanha	70	80	10	91	11
Norte	51	55	4	54	-1
Galiza	55	65	10	74	9

Fonte: CE (1999) e elaboração própria a partir da informação recolhida *on-line* no sitio http://epp.eurostat.cec.eu.int em Março de 2009.

Tabela5. Diferenciais do PIB *per capita* (em PPC)

	1995	2005
PIBpc Portugal/PIBpc Espanha	0,82	0,76
PIBpc Galiza/PIBpc Espanha	0,82	0,82
PIBpc Norte/PIBpc Portugal	0,85	0,79
PIBpc Norte/PIBpc Galiza	0,85	0,73

Fonte: Elaboração própria a partir da informação recolhida *on-line* no sitio http://epp.eurostat.cec.eu.int em Março de 2009.

Apesar de termos verificado um aumento do PIBpc (em PPC) em todas as regiões, as tabela 4 e 5 evidenciam o melhor desempenho da Espanha e da região da Galiza. Em vinte anos a redução do diferencial relativamente à média da UE15 foi muito superior em Espanha e na Galiza. Em 1996, o PIBpc em Espanha e na Galiza era, respectivamente, 20% e 35% inferior à média da UE15 e em 2005 a diferença era de 9% e 25%, respectivamente. No entanto, se analisarmos o diferencial entre a Galiza e a Espanha verificamos que a região Galega, no período 1995-2005 não reduziu a diferença em relação à média nacional. Com efeito, e analisando a taxa de crescimento real do PIB (ver tabela 6), na Galiza cresceu a um ritmo inferior à média nacional o que reforça a ideia de que parte da convergência verificada refere-se a uma convergência passiva[4].

O desempenho da região Norte tem sido muito fraco reflectindo as dificuldades que a região atravessa. Após um período convergência com a UE15, entre 1986 e 1996, nos anos mais recentes essa tendência inverteu-se. Também em relação à média nacional, e mesmo em comparação com a Galiza, a região Norte viu o seu nível de riqueza reduzir-se uma vez que tem

[4] Como vimos anteriormente, no ponto 1, a Galiza tem perdido população pelo que esta aproximação reflecte, em grande parte, uma convergência passiva, isto é, convergência devido à descida na população e não tanto um grande aumento da produção ou aumento da produção acima da média.

registado taxas de crescimento inferiores à média nacional e europeia (tabela 6). Em consequência, a região Norte voltou a integrar o grupo das 10 regiões mais pobres da UE15.

Tabela 6. Taxa de crescimento real do PIB

	1995-01 (média)	2001	2002	2003	2004	2005	2006
UE15	2,5	1,7	1	n.d.	n.d.	n.d.	n.d.
Portugal	3,5	2	0,8	-0,8	1,5	0,9	n.d.
Norte	2,6	1,3	-1,4	-2,6	0,9	1	n.d.
Espanha	3,7	3,5	2,7	3,1	3,3	3,6	3,9
Galiza	2,8	2,5	2,3	2,6	3,5	3,2	4,1

Fonte: Eurostat, informação recolhida *on-line* no sitio http://epp.eurostat.cec.eu.int em Fevereiro de 2009 e CE (2004) para média 1995-01

Este fraco desempenho no Norte de Portugal, face à média nacional e níveis médios das regiões europeias, prende-se com a estrutura da sua actividade económica e a produtividade, as quais constituem algumas das principais razões que explicam a competitividade de uma região. Como refere Azevedo (2004 p.94):

(…) "*esta situação reflecte as dificuldades conjunturais por que passam as economias europeia e portuguesa e reflecte também dificuldades estruturais inerentes ao modelo produtivo dominante no Norte de Portugal (…)*"

Portanto, as desvantagens estruturais[5] tornaram-se ainda mais acentuadas à medida que a UE15 se alargou a novos membros do Leste, à medida que se verificou a abertura dos mercados asiáticos, sobretudo a China, Índia e países do Magreb que determinou uma concorrência acrescida à qual muitas empresas da região Norte não resistiram acabando por encerrar ou deslocar a produção para outros países mais "*atractivos*" como aconteceu com muitas filiais de multinacionais.

Outro indício de que as regiões da Galiza e do Norte não estão a convergir com as médias nacionais é a perda de importância do Valor Acrescentado Bruto (VAB) e emprego regionais na economia nacional.

Tabela 7. Peso de cada região na economia nacional

		1995	2006
VAB	Norte/Portugal	29,8	28,0
	Galiza /Espanha	5,6	5,1
Emprego	Norte/Portugal	34,9	35,1
	Galiza/Espanha	7,3	6,0

Fonte: Elaboração própria a partir dos dados do Eurostat : Informação recolhida *on-line* no sitio http://epp.eurostat.cec.eu.int em Fevereiro de 2009.

Embora tenham uma estrutura sectorial comparável, o peso económico de cada uma das regiões nas respectivas economias nacionais é muito diferente. Enquanto o Norte contribui com cerca de 28% do VAB e detém 35% do emprego nacional, a importância relativa da Galiza para Espanha é menos significativa com cerca de 5% do VAB e 6% do emprego. No período em causa ambas perderam importância em termos de VAB e a Galiza também em termos de emprego.

Apesar das trajectórias, entre as duas regiões, terem sido divergentes, sobretudo nos últimos anos, a análise que realizamos parece indicar também que o desempenho das duas regiões tem sido mais fraco do que as respectivas economias nacionais, sobretudo ao nível do emprego, desemprego e crescimento do PIB. Em que medida o modelo produtivo das duas regiões e as respectivas dinâmicas de evolução ajudam a compreender a posição relativa das regiões nas economias nacionais?

[5] Além das desvantagens estruturais, não podemos esquecer também a valorização do euro nos últimos anos que em nada favoreceu a já baixa competitividade das empresas exportadoras.

3.1 Aplicação do Quociente de Localização e análise "shift –share"

No sentido de conhecermos a estrutura ou composição interna da economia de cada região face às respectivas economias nacionais apresentamos, de seguida, algumas medidas de análise económica regional. Mais concretamente, e com o objectivo de proceder a uma reflexão sobre a dinâmica da estrutura produtiva das regiões Norte e da Galiza, para o período 1995-2003[6], recorremos ao Quociente de Localização (QL) e ao "método das componentes de variação" *("shift share analysis*"). Os indicadores foram calculados usando duas variáveis: o emprego por sector e o VAB por sector (a preços básicos). O padrão ou a área geográfica tomada como referência foi Portugal para a região Norte e Espanha para a região da Galiza. Os dados usados nos modelos provêm da base de dados do Eurostat (Regio database). As conotações utilizadas são seguintes:

- i – refere os diferentes sectores de actividade;
- R – denota o nível regional, isto é, região Norte e da Galiza;
- N – denota o nível nacional, isto é, Portugal e Espanha;
- X – refere-se à variável em análise, o emprego (E) e o VAB;
- 0 e 1- denotam o momento inicial (1995) e o momento final (2003), respectivamente.

[6] Gostaríamos de ter realizado a análise para um período mais alargado. No entanto, quando fizemos a primeira extracção dos dados, a base de dados (Regiodatabase) só disponibilizava dados para o período 1995-2003. Em Fevereiro de 2009 procuramos actualizar estes dados e o período de análise (nomeadamente o último ano disponível). Contudo, verificamos que a base de dados sofreu alterações. Actualmente, a base está mais alargada em termos regionais (NUT III) mas o nível de desagregação do VAB e emprego por sectores de actividade foi reduzido. Na altura da primeira elaboração desta análise (Junho de 2006) o nível de desagregação sectorial é aquele que está apresentado nas tabelas 8 a 11. Actualmente o nível de desagregação é muito mais reduzido. Passou de um nível de desagregação de 16 sectores de actividade para apenas de 6, isto é, A –B: Agricultura, produção animal, caça e silvicultura; C-E: Indústrias extractivas, transformadoras, electricidade, gás e água; F: Construção; G_H_I: Comércio por grosso e a retalho; reparação de veículos automóveis, motociclos e de bens de uso pessoal e doméstico_ Alojamento e restauração_ Transportes, armazenagem e comunicações; J_K: Actividades financeiras_ Actividades imobiliárias, alugueres e serviços prestados às empresas e L – P: Administração pública, defesa e segurança social obrigatória_ Educação_ Saúde e acção social_ Outras actividades de serviços colectivos, sociais e pessoais_ Famílias com empregados domésticos. Por esta razão, e para não perdermos informação em termos de desagregação de sectores, mantivemos esta análise para o período1995-2003.

3.1.1 Norte vs Portugal

Tabela 8. Quociente de localização e análise shift-share pelo emprego (10^3)

Norte vs Portugal	Quociente localização (Emprego)		1- Variação observada	Componente crescimento nacional			Efeito Estrutural		Vantagem competitiva	
				2 -Taxa crescimento nacional	3 - Variação esperada	4- "Total Shift"	5.1 - Taxa	5.2 - Variação	6.1 - Taxa	6.2 - Variação
	1995	2003	Ei(1)- Ei(0)	N(1)/ N(0) -1	[Ei(0)* N(1)/N(0)-1]	[Ei(1) -Ei(0)] - [Ei(0)*N(1)/ N(0)-1]	Ni(1)/ Ni(0)- N(1)/ N(0)	Ei(0)*[Ni(1)/ Ni(0)-N(1)/ N(0)]	Ei(1)/Ei(0)-Ni(1)/Ni(0)	Ei(0)* [Ei(1)/Ei(0)- Ni(1)/Ni(0)]
A -B Agricultura, produção animal, caça e silvicultura	**1,04**	**1,03**	**-26,4**		**23,34**	**-49,74**	**-0,21**	**-41,85**	**-0,04**	**-7,89**
A - Agricultura, produção animal, caça e silvicultura	1,06	1,05	-25,8		22,79	-48,59	-0,21	-39,89	-0,04	-8,70
B - Pesca	0,58	0,69	-0,5		0,55	-1,05	-0,33	-1,55	0,11	0,50
C-F Indústria	**1,34**	**1,35**	**24,1**		**76,83**	**-52,73**	**-0,05**	**-33,71**	**-0,03**	**-19,02**
C - Indústrias extractivas	0,71	0,90	1		0,42	0,58	-0,07	-0,25	0,23	0,83
D - Indústrias transformadoras	1,48	1,54	3,8		58,22	-54,42	-0,11	-56,26	0,00	1,84
E - Produção e distribuição de electricidade, gás e água	0,77	0,73	-0,8		1,01	-1,81	-0,13	-1,09	-0,08	-0,72
F - Construção	1,08	1,04	20		17,20	2,80	0,11	15,91	-0,09	-13,11
G - P Serviços (excluindo organismos e instituições extraterritoriais)	**0,80**	**0,82**	**124,4**		**83,34**	**41,06**	**0,07**	**52,43**	**-0,02**	**-11,37**
G - Comércio por grosso e a retalho; reparação de veículos automóveis, motociclos e de bens de uso pessoal e doméstico	0,98	0,95	21,8		26,97	-5,17	0,06	13,02	-0,08	-18,19
H - Alojamento e restauração (restaurantes, hóteis e similares)	0,70	0,72	13,9	**0,12**	5,60	8,30	0,19	8,99	-0,01	-0,69
I - Transportes, armazenagem e comunicações	0,71	0,73	1,3		4,37	-3,07	-0,07	-2,70	-0,01	-0,37
J - Actividades financeiras	0,69	0,61	-6,9		3,31	-10,21	-0,23	-6,60	-0,13	-3,61
K - Actividades imobiliárias, alugueres e serviços prestados às empresas	0,69	0,68	18		7,42	10,58	0,23	14,70	-0,07	-4,12
L - Administração pública, defesa e segurança social obrigatória	0,60	0,63	9		8,92	0,08	-0,02	-1,15	0,02	1,23
M - Educação	0,90	0,99	23		9,87	13,13	0,08	6,91	0,07	6,22
N - Saúde e acção social	0,82	0,91	25		7,57	17,43	0,18	11,29	0,10	6,13
O - Outras actividades de serviços colectivos, sociais e pessoais	0,68	0,71	8,5		4,48	4,02	0,11	4,04	0,00	-0,02
P - Famílias com empregados domésticos	0,95	1,00	10,8		4,81	5,99	0,13	5,50	0,01	0,48
Total	**1,00**	**1,00**	**122,1**		**183,51**	**-61,41**	**..**	**..**	**-0,04**	**-61,41**

Fonte: Elaboração própria a partir dos dados do Eurostat.

Os resultados indicam que o sector primário, nos últimos anos, na região Norte perdeu posição competitiva em relação às outras regiões do país mas de forma mais acentuada em termos de VAB. Relativamente ao sector secundário, observando os valores do QL, tanto em termos de emprego como em termos de VAB, verifica-se que o sector industrial tem uma maior importância relativa na região Norte do que a nível nacional. Dentro do sector industrial, ao nível do emprego, essa importância verifica-se para a indústria transformadora e construção e, em termos de VAB, estende-se também à produção e distribuição de electricidade, gás e água. Na verdade, já anteriormente referimos que a região Norte é a mais industrializada do país o que tem como reverso uma menor especialização no sector dos serviços em comparação com as outras regiões do país. Ainda ao nível do sector industrial é de referir a crescente importância da indústria extractiva na região que, apesar de verificar uma divergência estrutural está a beneficiar de um efeito regional positivo, ou seja, a região Norte está a reforçar a sua posição competitiva neste sector. No que respeita ao sector dos serviços, os dados parecem indicar uma evolução positiva para os chamados "*non-market services*", tendo a região reforçado a sua posição competitiva nos serviços de educação, saúde e acção social, sectores que também beneficiaram de uma efeito estrutural positivo, isto é, acompanharam o crescimento médio da economia. Relativamente aos serviços comercializáveis (as actividades G a K), a região Norte está a "perder" posição competitiva comparativamente às outras regiões (tanto em termos de emprego como em termos de VAB) na medida em que o crescimento dessas actividades na região não acompanha o crescimento nacional dessas actividades.

Tabela 9. Quociente de localização e análise shift-share pelo VAB (10^3 Euros)

Norte vs Portugal	Quociente localização (Emprego)		1- Variação observada	Componente crescimento nacional			Efeito Estrutural		Vantagem competitiva	
				2 – Taxa crescimento nacional	3 - Variação esperada	4- "Total Shift"	5.1 - Taxa	5.2 - Variação	6.1 - Taxa	6.2 - Variação
	1995	2003	VABi(1)-VABi(0)	N(1)/N(0) -1	[VABi(0)* N(1)/N(0)-1]	[VABi(1)-VABi(0)] - [VABi(0)* N(1)/N(0)-1]	Ni(1)/Ni(0)- N(1)/N(0)	VABi(0)* [Ni(1)/ Ni(0)-N(1)/N(0)]	VABi(1)/ VABi(0)- Ni(1)/Ni(0)	VABi(0)* [VABi(1)/ VABi(0)- Ni(1)/Ni(0)]
A -B Agricultura, produção animal, caça e silvicultura	**0,87**	**0,74**	**-195,2**		**672,26**	**-867,46**	**-0,45**	**-544,02**	**-0,27**	**-323,44**
A - Agricultura, produção animal, caça e silvicultura	0,89	0,76	-201,2		645,42	-846,62	-0,47	-536,85	-0,27	-309,77
B - Pesca	0,51	0,51	6		26,84	-20,84	-0,28	-13,37	-0,16	-7,47
C-F Indústria	**1,25**	**1,37**	**2658,9**		**4663,18**	**-2004,28**	**-0,19**	**-1617,41**	**-0,05**	**-386,87**
C - Indústrias extractivas	0,46	0,77	39,7		33,72	5,98	-0,43	-26,15	0,53	32,13
D - Indústrias transformadoras	1,35	1,51	1661,1		3274,59	-1613,49	-0,26	-1511,81	-0,02	-101,68
E - Produção e distribuição de electricidade, gás e água	1,20	1,26	169,6		472,18	-302,58	-0,26	-222,99	-0,09	-79,60
F - Construção	1,05	1,11	788,5		882,69	-94,19	0,05	71,04	-0,10	-165,23
G - P Serviços (excluindo organismos e instituições extraterritoriais)	**0,82**	**0,89**	**7961,4**		**7364,93**	**596,47**	**0,12**	**1584,92**	**-0,08**	**-988,45**
G - Comércio por grosso e a retalho; reparação de veículos automóveis, motociclos e de bens de uso pessoal e doméstico	**0,97**	**1,01**	**1122,9**	**0,56**	**1845,96**	**-723,06**	**-0,10**	**-338,21**	**-0,12**	**-384,85**
H - Alojamento e restauração (restaurantes, hóteis e similares)	0,51	0,56	333,2		250,97	82,23	0,23	105,38	-0,05	-23,14
I - Transportes, armazenagem e comunicações	0,70	0,69	402,8		620,26	-217,46	0,02	19,13	-0,21	-236,59
J - Actividades financeiras	0,77	0,62	152,3		659,12	-506,82	0,03	31,51	-0,46	-538,33
K - Actividades imobiliárias, alugueres e serviços prestados às empresas	0,81	0,91	1823,5		1522,97	300,53	0,12	330,16	-0,01	-29,63
L - Administração pública, defesa e segurança social obrigatória	0,69	0,79	1057,8		747,76	310,04	0,20	267,98	0,03	42,06
M - Educação	1,00	1,19	1326,3		837,51	488,79	0,24	357,39	0,09	131,40
N - Saúde e acção social	0,92	1,10	1246,5		600,86	645,64	0,50	533,53	0,10	112,11
O - Outras actividades de serviços colectivos, sociais e pessoais	0,72	0,82	334		178,94	155,06	0,47	148,87	0,02	6,19
P - Famílias com empregados domésticos	1,05	1,20	165,1		100,04	65,06	0,34	60,80	0,02	4,26
Total	**1,00**	**1,00**	**9087,9**		**13448,14**	**-4360,24**	**..**	**..**	**-0,18**	**-4360,24**

Fonte: Elaboração própria a partir dos dados do Eurostat

3.1.2 Galiza vs Espanha

Tabela 10. Quociente de localização e análise shift-share pelo emprego (10^3)

				Componente crescimento nacional			Efeito Estrutural		Vantagem competitiva	
Galiza vs Espanha	Quociente localização (Emprego)		1- Variação observada	2 - Taxa crescimento nacional	3 - Variação esperada	4- "Total Shift"	5.1 - Taxa	5.2 - Variação	6.1 - Taxa	6.2 - Variação
	1995	2003	Ei(1)-Ei(0)	N(1)/N(0)-1	[Ei(0)*N(1)/N(0)-1]	[Ei(1)-Ei(0)] - [Ei(0)*N(1)/N(0)-1]	Ni(1)/Ni(0)-N(1)/N(0)	Ei(0)*[Ni(1)/Ni(0)-N(1)/N(0)]	Ei(1)/Ei(0)-Ni(1)/Ni(0)	Ei(0)*[Ei(1)/Ei(0)-Ni(1)/Ni(0)]
A -B Agricultura, produção animal, caça e silvicultura	**2,85**	**1,59**	**-137,5**		**68,67**	**-206,17**	**-0,36**	**-83,73**	**-0,53**	**-122,43**
A - Agricultura, produção animal, caça e silvicultura	:	1,11	-141,7		60,40	-202,10	:	:	:	:
B - Pesca	:	9,26	4,2		8,27	-4,07	:	:	:	:
C-F Indústria	**0,89**	**1,04**	**70**		**73,35**	**-3,35**	**0,10**	**24,63**	**-0,11**	**-27,97**
C - Indústrias extractivas	:	1,77	-0,3		1,30	-1,60	:	:	:	:
D - Indústrias transformadoras	0,79	1,04	35,4		41,61	-6,21	-0,08	-11,71	0,04	5,50
E - Produção e distribuição de electricidade, gás e água	:	1,28	-1		1,87	-2,87	:	:	:	:
F - Construção	1,06	1,03	35,9		28,57	7,33	0,53	50,67	-0,45	-43,34
G - P Serviços (excluindo organismos e instituições extraterritoriais)	**0,81**	**0,93**	**78,3**		**154,46**	**-76,16**	**0,00**	**1,28**	**-0,15**	**-77,44**
G - Comércio por grosso e a retalho; reparação de veículos automóveis, motociclos e de bens de uso pessoal e doméstico	0,90	1,00	13,9	**0,30**	41,05	-27,15	-0,03	-3,60	-0,17	-23,54
H - Alojamento e restauração (restaurantes, hóteis e similares)	0,89	0,81	-0,1		15,62	-15,72	0,11	5,99	-0,41	-21,71
I - Transportes, armazenagem e comunicações	0,79	0,82	-0,8		13,84	-14,64	-0,08	-3,85	-0,23	-10,79
J - Actividades financeiras	0,73	0,85	-1,1		5,57	-6,67	-0,26	-4,82	-0,10	-1,85
K - Actividades imobiliárias, alugueres e serviços prestados às empresas	0,58	0,75	28,9		10,25	18,65	0,52	17,94	0,02	0,71
L - Administração pública, defesa e segurança social obrigatória	0,76	0,98	7,8		19,41	-11,61	-0,19	-12,35	0,01	0,74
M - Educação	0,90	1,02	4,3		14,64	-10,34	-0,07	-3,43	-0,14	-6,91
N - Saúde e acção social	0,76	0,97	11,2		13,19	-1,99	-0,04	-1,76	-0,01	-0,23
O - Outras actividades de serviços colectivos, sociais e pessoais	0,84	0,96	11,1		8,83	2,27	0,24	7,04	-0,16	-4,77
P - Famílias com empregados domésticos	0,88	1,06	3,1		12,06	-8,96	-0,15	-5,94	-0,07	-3,02
Total	**1,00**	**1,00**	**10,8**		**296,47**	**-285,67**	**..**	**..**	**-0,29**	**-285,67**

Fonte: Elaboração própria a partir dos dados do Eurostat (Informação recolhida *on-line* no sitio http://epp.eurostat.cec.eu.int

Em comparação com a economia nacional, a Galiza é uma região muito especializada no sector primário tanto em termos de emprego como em termos de VAB, sobretudo nas indústrias relacionadas com o sector pesqueiro que, em 2003, apresentava um QL de 9,26 (em termos de emprego) e 8,79 (em termos de VAB). Outro dado que realça a importância e desenvolvimento do sector das pescas Galego é que a vantagem competitiva da Galiza dentro do país aumentou 5%, entre 1995 e 2003, como mostra a análise *shift share* (em termos de VAB). Para além disso, é de referir que o emprego no sector agrícola tem caído acentuadamente como mostra o efeito estrutural o que poderá ser um indício de aumento de produtividade. Ao analisarmos o QL (usando o emprego) as indústrias extractivas e de produção e distribuição de energias apresentam indicadores maiores que a unidade, estendendo-se à indústria da construção se considerarmos o VAB. De facto, este sector foi o que mais cresceu a nível nacional (tanto em termos de emprego como de VAB 53% em ambos os casos) juntamente com a actividade K (K – actividades imobiliárias, alugueres e serviços prestados às empresas) do sector dos serviços. A indústria da produção e distribuição de energia, também assume relevância, sobretudo devido à localização da Repsol na Corunha. Em termos regionais, a Galiza conseguiu reforçar a sua posição competitiva nas indústrias extractivas e transformadoras. Se associarmos o crescimento relativo do emprego, nestas actividades, com o crescimento no VAB, que foi maior, isso poderá ser um indicativo de aumentos de produtividade que é uma condicionante essencial para reforçar as vantagens competitivas numa região. Ao nível do sector dos serviços, tal como acontecia na região norte, embora se tenha verificado uma aproximação aos níveis nacionais, as actividades terciárias ainda têm menor importância relativa na região da Galiza do que a nível nacional com a excepção dos serviços de educação que verificaram um significativo aumento do QL de 1995 para 2003 (em termos de VAB). Tanto em termos de emprego como em termos de VAB, a região da Galiza perdeu posição competitiva em relação às outras regiões de Espanha no sector dos serviços.

Tabela 11. Quociente de localização e análise shift-share pelo VAB (10^3 Euros)

Galiza vs Espanha	Quociente localização (Emprego)		1- Variação observada	Componente crescimento nacional			Efeito Estrutural		Vantagem competitiva	
				2 - Taxa crescimento nacional	3 - Variação esperada	4- "Total Shift"	5.1 - Taxa	5.2 - Variação	6.1 - Taxa	6.2 - Variação
	1995	2003	VABi(1)-VABi(0)	N(1)/N(0) -1	[VABi(0)*N(1)/N(0)-1]	[VABi(1)-VABi(0)] - [VABi(0)*N(1)/N(0)-1]	Ni(1)/Ni(0)-N(1)/N(0)	VAB(0)*[Ni(1)/Ni(0)-N(1)/N(0)]	VABi(1)/VABi(0)-Ni(1)/Ni(0)	VABi(0)*[VABi(1)/VABi(0)-Ni(1)/Ni(0)]
A -B Agricultura, produção animal, caça e silvicultura	**1,75**	**1,54**	**41,2**		**1249,28**	**-1208,08**	**-0,27**	**-538,73**	**-0,34**	**-669,35**
A - Agricultura, produção animal, caça e silvicultura	1,32	1,07	-56,9		875,43	-932,33	-0,25	-350,64	-0,42	-581,69
B - Pesca	7,16	8,79	98,1		373,85	-275,75	-0,52	-306,92	0,05	31,17
C-F Indústria	**0,97**	**1,11**	**4199,1**		**4646,88**	**-447,78**	**-0,03**	**-204,62**	**-0,03**	**-243,16**
C - Indústrias extractivas	1,07	1,58	8,2		100,06	-91,86	-0,80	-126,44	0,22	34,58
D - Indústrias transformadoras	0,81	1,00	2144,8		2445,96	-301,16	-0,15	-585,93	0,07	284,77
E - Produção e distribuição de electricidade, gás e água	1,65	1,74	-27,7		760,08	-787,78	-0,55	-661,43	-0,11	-126,34
F - Construção	1,10	1,17	2073,8		1340,78	733,02	0,53	1130,13	-0,19	-397,11
G - P Serviços (excluindo organismos e instituições extraterritoriais)	**0,88**	**0,92**	**7219,4**		**9454,60**	**-2235,20**	**0,03**	**459,29**	**-0,18**	**-2694,50**
G - Comércio por grosso e a retalho; reparação de veículos automóveis, motociclos e de bens de uso pessoal e doméstico	0,92	0,92	839,6	**0,63**	1705,87	-866,27	-0,09	-247,36	-0,23	-618,91
H - Alojamento e restauração (restaurantes, hóteis e similares)	0,89	0,78	392,9		1073,27	-680,37	0,01	24,95	-0,42	-705,32
I - Transportes, armazenagem e comunicações	0,81	0,81	486,4		1034,75	-548,35	-0,11	-186,26	-0,22	-362,09
J - Actividades financeiras	0,79	0,85	329,3		705,64	-376,34	-0,22	-241,09	-0,12	-135,25
K - Actividades imobiliárias, alugueres e serviços prestados às empresas	0,84	0,85	2107		1776,75	330,25	0,37	1049,99	-0,26	-719,74
L - Administração pública, defesa e segurança social obrigatória	0,84	1,05	853,3		867,94	-14,64	-0,11	-154,04	0,10	139,40
M - Educação	1,06	1,21	794,5		824,16	-29,66	0,03	36,96	-0,05	-66,62
N - Saúde e acção social	1,00	1,08	689,6		871,43	-181,83	-0,01	-11,73	-0,12	-170,10
O - Outras actividades de serviços colectivos, sociais e pessoais	0,87	0,97	598,7		432,10	166,60	0,33	226,50	-0,09	-59,90
P - Famílias com empregados domésticos	0,94	1,17	127,9		162,69	-34,79	-0,22	-57,46	0,09	22,68
Total	**1,00**	**1,00**	**10092**		**16218,71**	**-6126,91**	**..**	**..**	**-0,24**	**-6126,91**

Fonte: Elaboração própria a partir dos dados do Eurostat (Informação recolhida *on-line* no sitio http://epp.eurostat.cec.eu.int.

Que conclusão podemos tirar desta análise?

Em termos gerais, no período de 1995 a 2003, tanto a região Norte como a Galiza têm perdido posição/vantagem competitiva em relação às outras regiões do país. No entanto, ao contrário do que acontece na Galiza em que a perda em termos de VAB é menor do que a observada em termos de emprego, indicando um aumento de produtividade competitiva, na região Norte, verifica-se o inverso, isto é, a perda em termos de VAB é superior à perda em termos de emprego indicando uma perda competitiva em termos de produtividade.

Com efeito, é sabido que num processo de convergência real, o aumento da produtividade constitui um factor chave para o alcance da mesma, juntamente com acumulação de factores produtivos (trabalho, capital físico e tecnológico)[7]. No entanto, a competitividade de uma região e, consequentemente, as decisões de localização de empresas, além de outros factores, dependem da análise da relação entre os custos por trabalhador[8] com os índices de produtividade[9], ou seja, dos custos unitários do trabalho (CTU) [10]. Razão pela qual procedemos a essa análise de seguida.

3.2 Produtividade e custos de trabalho unitários[11]

Portugal e a região Norte apesar de apresentarem custos por trabalhador (em todos os sectores) inferiores a Espanha e à região da Galiza (respectivamente) apresentam também níveis de produtividade inferiores, sendo mesmo dos mais baixos de toda a UE15. A análise dos CTU parece indicar que, a nível agregado de todos os sectores, o CTU é superior em Portugal e na região Norte na comparação com as respectivas regiões. Isto significa que, tanto em 1995 como em 2005, a produtividade média da Galiza e de Espanha, relativamente à região Norte e a Portugal, mais que compensa o maior custo laboral por trabalhador. Outro aspecto a salientar é a perda de competitividade das regiões em relação às respectivas médias nacionais e da região Norte em relação à Galiza (ver gráfico 5). Adicionalmente, e como não foi possível apresentar os CTU para a UE15, devido à indisponibilidade de dados, apresentamos os cálculos e comparação com a França (uma das economias de referência da UE). Pela análise do gráfico 6 constata-se também a perda de competitividade destas regiões, mais acentuada no caso da região Norte e de Portugal.

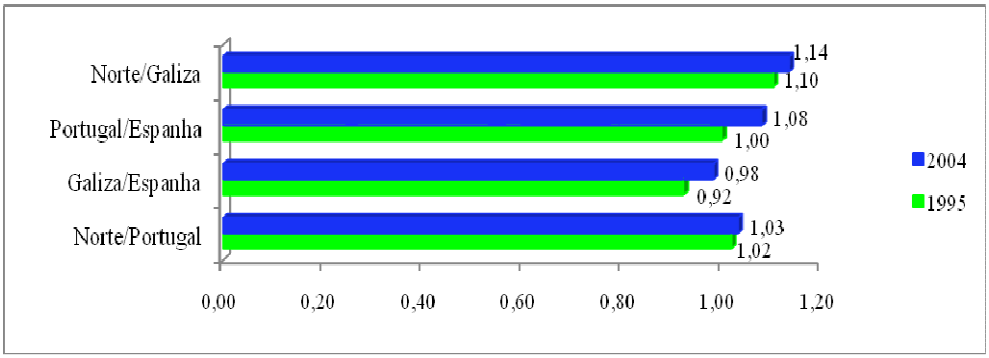
Tabela 12. Evolução do custo de trabalho unitário (CTU)

	1995				2004			
	Portugal	Espanha	Norte	Galiza	Portugal	Espanha	Norte	Galiza
Agricultura, Pecuária, silvicultura e pesca	0,15	0,17	0,14	0,21	0,20	0,19	0,18	0,29
Indústria	0,58	0,57	0,59	0,52	0,56	0,56	0,61	0,54
Construção	0,56	0,61	0,56	0,60	0,65	0,58	0,68	0,53
Serviços	0,53	0,51	0,53	0,48	0,58	0,53	0,59	0,52
Todos os sectores	**0,52**	**0,52**	**0,53**	**0,48**	**0,57**	**0,53**	**0,59**	**0,52**

Fonte: Elaboração própria a partir dos dados do Eurostat. Informação recolhida *on-line* no sitio http://epp.eurostat.cec.eu.int em Fevereiro de 2009).

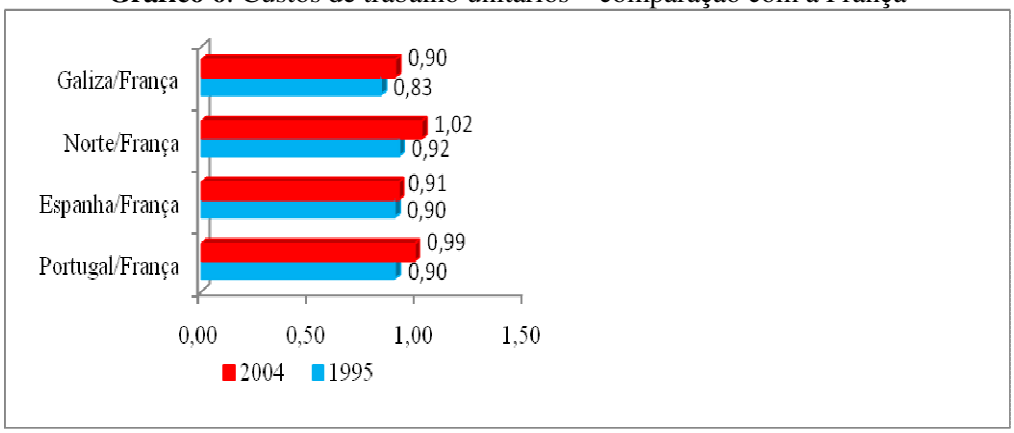
[7] Esteban (1995) num trabalho em que analisa os países da UE, evidencia que a dispersão nos níveis de produtividade explicam mais de 2/3 das desigualdades espaciais de rendimento e a outra terça parte é explicada de forma similar pela taxa de actividade e emprego.
[8] Medimos os custos por trabalhador pelo quociente entre as remunerações e o emprego.
[9] Medimos a produtividade aparente pelo quociente entre VAB (a preços básicos) e o emprego.
[10] Os CTU correspondem ao quociente entre o custo por trabalhador e a produtividade aparente.
[11] Em Fevereiro de 2009 estavam disponíveis os dados do VAB por sector, para 2006, mas não os dados do emprego por sector. Além disso, como o último "Labour Cost Survey" disponível é o de 2004tivemos que restringir a análise ao período 1995-2004.

Gráfico 5. Custos de trabalho unitários relativos



Fonte: Elaboração própria a partir dos dados do Eurostat. Informação recolhida *on-line* no sitio http://epp.eurostat.cec.eu.int em Fevereiro de 2009).

Gráfico 6. Custos de trabalho unitários – comparação com a França



Fonte: Elaboração própria a partir dos dados do Eurostat. Informação recolhida *on-line* no sitio http://epp.eurostat.cec.eu.int em Abril de 2009).

França	**1995**	**2004**
Agricultura, Pecuária, silvicultura e pesca	0,19	0,24
Indústria (incluindo construção)	0,65	0,62
Serviços	0,57	0,58
Todos os sectores	0,58	0,58

CONCLUSÕES

A região Norte e a região da Galiza são duas regiões pertencentes a dois países vizinhos que aderiram no mesmo ano ao mesmo bloco de integração regional. Estes países, Portugal e Espanha, realizaram progressos consideráveis em diferentes níveis e diferentes amplitudes, na redução das disparidades iniciais que os distanciavam dos níveis médios da UE15. No entanto, os benefícios do crescimento económico não se estendem de igual forma a todos os territórios. Autores como Dall'erba (2003), a CE (2004, 2007) reconhecem que apesar dos notáveis progressos verificados nos países da coesão, isso não significou a redução das disparidades no seio dos mesmos estados.

Esta parece ser também uma das principais conclusões que esta análise comparativa parece indicar para a região Norte e Galiza. Ao procurar descrever e analisar o desempenho destas duas regiões no seio das respectivas economias nacionais, os dados parecem indicar que nenhuma das regiões conseguiu reforçar a sua importância e posição competitiva no interior dos respectivos

países. Com efeito, a análise *shift-share* indica que tanto a região Norte como a Galiza têm perdido posição competitiva no interior dos respectivos países.

A qualificação dos recursos humanos também mostra divergências. Os níveis educacionais na Galiza são muito superiores aos da região Norte. Neste indicador, ambas as regiões situavam-se abaixo da média nacional embora, em termos de educação superior, a região da Galiza tenha alcançado a média nacional e superado a média europeia. O baixo nível educacional da região Norte contínua a ser uma grande desvantagem da região Norte não tendo reduzido a divergência nem com a média nacional nem com a média europeia.

Os níveis de desemprego, na região Norte, eram tradicionalmente baixos mas, nos anos mais recentes, têm verificado um forte aumento, estando a divergir com a média nacional. Analogamente, a análise dos dados do desemprego e da actividade parece indicar uma divergência da região da Galiza com a média nacional.

No que se refere à estrutura económica e competitividade, encontram-se também algumas características comuns nas dictomias região-economia nacional. Com efeito, ambas as regiões ainda detêm sectores agrícolas mais extensos que as outras regiões nacionais. Sectores estes que se caracterizam por baixos índices de produtividade na medida em que, nestas regiões, dadas as características geográficas e demográficas, muitas das áreas agrícolas são fragmentadas e de reduzida dimensão não facilitando a mecanização nem a obtenção de economias de escala. Adicionalmente, os dados analisados parecem indicar que tanto na região Norte como na Galiza o sector primário está a perder emprego a favor do sector terciário mas, apesar disso, cada região continua a manter a sua principal característica económica: a especialização industrial na região Norte e a especialização no sector pesqueiro na Galiza.

Por último, quer a região Norte quer a Galiza têm níveis médios de produtividade mais baixos que a média do país em todos os sectores. Podemos ainda referir que, a análise dos CTU, para o período de 1995-2005, parece indicar uma vantagem a favor da Galiza relativamente à região Norte no que se refere à relação custos salariais-produtividade, isto é, os níveis de produtividade superiores da Galiza compensam os mais altos custos salariais por trabalhador.

Assim, apesar das evoluções positivas verificadas em cada uma das regiões, a região Norte e a Galiza continuam a ser duas das regiões mais atrasadas das respectivas economias nacionais e, tal como refere Azevedo (2004, p. 108):

"O Norte de Portugal e a Galiza continuam a apresentar níveis de desenvolvimento dos mais baixos da UE15, revelando fragilidades significativas em matéria de dotação de alguns dos factores dinâmicos de competitividade(…)".

BIBLIOGRAFIA

ANTUNES, M. E SOUKIAZIS, E. (2005). "Two Speed Regional Convergence in Portugal and the Importance of Structural Funds on Growth". *Paper presented at the 4th Annual Meeting of the EEFS, 19-22 May, Coimbra.*

ARMSTRONG, H.W. (2001), "European Union Regional Policy: Reconciling the convergence". *In: Regional Convergence in the European Union, Facts, Prospects and Policies*. Cuadrado-Roura, J.R. e Parellada, M. Editors.

AZEVEDO, RUI (2004). "As regiões do Norte de Portugal e da Galiza: enquadramento no contexto comunitário e perfis de evolução".In: *As eurorrexións e o futuro da Europa: o modelo da eurorrexión Galiza-Norte de Portugal.*

BANDE, ROBERTO E FERNÁNDEZ, MELCHOR (2007). "Integración y mercado de trabajo en la Eurorregión Galicia-Norte de Portugal". *Papeles de Economía Española*, núm.22.

BÄHR, CORNELIUS (2008). "How Does Sub-National Autonomy Affect the Effectiveness of Structural Funds?" *Kyklos,* vol 61-2008, n.º1, 3-18.

BASSANINI, A. E SCARPETTA, S. (2001). "The driving forces of economic growth: panel data evidence for the OECD countries". *OECD Economic Studies*, n.º33.

BENHABIB, JESS; SPIEGEL, MARK M. (2003). "Human Capital and Technology Diffusion". *Development Institute Research Working Papers Series, n.º3, May, 2003.*

BENHABIB, JESS; SPIEGEL, MARK M. (1994). "The role of human capital in economic development. Evidence from aggregate cross-country data". *Journal of Monetary Economics*, 34.

BLÖNDAL, S. ; FIELD, S. E GIROUARD, N. (2002). "L'investissement en capital humain: le rôle de l'enseignement supérieure ". *Révue économique de L'OCDE, n.º 34*

BODENSTEIN, T. AND KEMMERLING, A. (2008). "Ripples in a Rising Tide : Why Some EU Regions Receive More Structural Funds Than Others Do". *Center for European Studies Working Paper Series, 157.*

COMISSÃO EUROPEIA (2007). *Quarto relatório sobre a coesão económica e social na União Europeia.* Comissão Europeia, Serviço das Publicações Oficiais das Comunidades Europeias.

COMISSÃO EUROPEIA (2004). *Terceiro relatório sobre a coesão económica e social na União Europeia.* Comissão Europeia, Serviço das Publicações Oficiais das Comunidades Europeias.

COMISSÃO DE COORDENAÇÃO do Desenvolvimento Regional do Norte (2006). *Norte 2015. Competitividade e Desenvolvimento – Uma visão estratégica.* CCDRN.

DALL'ERBA, SANDY AND LE GALLO, JULIE (2003A). "Competition, Complementary and Increasing Disparities Among the Regions of Spain and Portugal". *Department of Economics, University of Pau* (France).

DEHESA, G. (2000). "Consolidará Galicia a súa converxencia real no século XXI?" *Revista Galega de Economia, vol.9, n.º1, pp7-24.*

DE LA FUENTE, A. (2003a). "El impacto de los Fondos Estructurales: convergencia real y cohesión interna." *Instituto de Análisis Económico – Universidad Autónoma de Barcelona. Marzo de 2003.*

DE LA FUENTE (2003 B), "Human Capital in a global and Knowledge-based economy, Part II: assessment at the EU country level" – *Final Report, Instituto de Análisis Económico (CSIC).*

DE LA FUENTE, A.; AVILÉS, A. Y FERNÁNDEZ, MELCHOR (2002). "Fondos estructurales, inversión en infraestructuras y crecimiento regional". Fundacion Caixa Galicia, CIEF, 2002.

DE MICHELIS, NICOLA (2008). "Regional Convergence: a Relevant Measure of Policy Success?" *CESifo Forum 1/2008*.

EUROSTAT – Regio Database (http://epp.eurostat.cec.eu.int)

FERNÁNDEZ, MELCHOR Y GRELA, M. FERNÁNDEZ (2003). "Comparación de estructuras productivas entre rexións transfronteirizas: Galicia e a rexión Norte de Portugal". *Revista Galega de Economia, vol.12, n.º2*.

FERNÁNDEZ, MELCHOR Y POLO, CLEMENTE (2001). "Reflexiones sobre el mercado laboral de dos rexiones contíguas: el caso de Galicia y la réxion Norte de Portugal". Fundacción Caixa Galicia.

OECD (1998). *Human capital investment. An international comparation.* OECD – Centre for Educational Research and Innovation.

PROENÇA, MARGARIDA (2004). "O impacto dos fundos estruturais e de coesão na perspectiva do alargamento. In: *As eurorrexións e o futuro da Europa: o modelo da eurorrexión Galiza-Norte de Portugal.*

TEIXEIRA, A. (1999). *Capital humano e capacidade de inovação. Contributos para o estudo do crescimento económico português 1960-1991*. Edição do Conselho Económico e Social.

TEMPLE, J. (2001). "*Growth effects of education and social capital in the OECD countries"*. OECD, Economic Studies, n.º 33.

LA NECESIDAD DE UNA INCUBADORA DE EMPRESAS EN EL SUR DEL ESTADO DE MÉXICO

JAVIER JESÚS RAMÍREZ HERNÁNDEZ
E-mail: javjes_uaemex@hotmail.com
Centro Universitario UAEM Tenancingo
Universidad Autónoma del Estado de México, México

Resumen

El presente artículo es parte de un estudio de factibilidad de la creación y operación de una incubadora de empresas en el Centro Universitario UAEM Tenancingo perteneciente a la Universidad Autónoma del Estado de México (UAEM). El diagnóstico de la situación socioeconómica y demográfica indica que la región sureste del Estado de México representó casi tres puntos porcentuales de la economía, con una tasa de crecimiento poblacional casi cero, existe emigración joven masculina al extranjero, aunque la mayoría de los habitantes cuentan con servicios públicos básicos. Por su parte, se utiliza el modelo de incubación de empresas de tecnología intermedia que fue apoyado en su desarrollo por el País Vasco en la década de los noventa. La importancia y los objetivos del proyecto son determinantes en su aceptación en la sociedad y su financiamiento. La creación y funcionamiento de una incubadora de empresas dentro de la UAEM es una forma de contribuir en el crecimiento económico y social de la región, en particular, la generación de empleos. Algunas acciones a realizarse bajo el ámbito de la incubadora de empresas son: consultorías empresarial, servicios especializados y de capacitación.

Palabras claves: Incubadora de empresas, economía mexicana.
Area tematica: Economía Internacional y de la Integración.

Abstract

This article is a part of a biggest study of creating and operating a business incubator in Centro Universitario UAEM Tenancingo at University Autonomous State of Mexico. The social, demographic and economic diagnostic indicates the regions is not important in a global context, there is high migration of young people, almost null growth population. The business incubator model was designed in collaborating with Bask Country years ago. The incubator will contribute whit social and economic growth, specially, new jobs. Its services consist in entrepreneur consulting, specific services and training.

Key words: business incubator, Mexican economy.
Thematic Area: International and Integration Economics.

I. INTRODUCCIÓN

En los países emergentes, como es la economía mexicana, se tiene la necesidad de la generación de empleos dado que la población sigue creciendo y genera problemas de desempelo o de migración. Los gobiernos, además de diversos sectores sociales, económicos y académicos, se han encaminado en dar propuestas para apoyar la generación de nuevas plazas de trabajo.

En este contexto, en el presente artículo se presentan los resultados de un estudio de factibilidad de apertura y operación de una incubadora de empresas de tecnología intermedia,

INCUBASK, en el Centro Universitario UAEM Tenancingo, organismo dependiente de la Universidad Autónoma del Estado de México ubicado al sur del Estado de México[1]. La UAEM cuenta con una red de incubadoras y esta nueva pasa a formar parte de ella.

El estudio sustenta una propuesta de funcionamiento de una incubadora de empresas en dicha zona del Estado de México. Esta es una zona extensa en territorio pero baja densidad poblacional, la actividad económica de la región aporta muy poco al total estatal, con predominación de actividades agropecuarias y de servicios, además de procesos productivos poco tecnificados. Hay problemas socioeconómicos diversos: bajos niveles de escolaridad y deserción escolar, inadecuadas condiciones de trabajo y bajos salarios, alta emigración reciente por razones económicas, en especial de la población joven masculina, entre otros.

De acuerdo con la Secretaría de Economía del gobierno federal, las incubadoras de empresas reconocidas asentadas en el territorio del Estado de México son 9 de negocios tradicionales, 10 de tecnología intermedia y 3 de alta tecnología. En consecuencia, esta estrategia de impulso al crecimiento económico regional y la mejora de las condiciones de vida de la población tiene mucho potencial por desarrollar.

El resto del documento compone por 7 secciones, la segunda aborda el tema de las incubadoras en México, el tercero es el correspondiente a la situación socioeconómica de la región sur del Estado de México, en la cuarta sección el apoyo de los sectores social y económico al proyecto, se prosigue con composición y tamaño de mercado, en la siguiente sección se aborda los servicios posibles demandados por los usuarios, y en la última se tratan los beneficios de la operación de una incubadora de empresas.

2. EL ORIGEN DE LAS INCUBADORAS DE EMPRESAS EN MÉXICO Y EN LA UAEM

En México se registra un bajo ritmo de crecimiento de la cantidad de empresas y la gran mayoría de las generadas cada año cierran sus puertas antes de los dos años de vida, razón por la cual el déficit de nuevas plazas de trabajo sigue creciendo anualmente Así se incrementa el porcentaje de la Población Económicamente Activa (PEA) desocupada o en condiciones de subempleo, convirtiéndose en una situación critica. Ante esto, la generación y consolidación de empresas a nivel nacional se ha vuelto un asunto critico en la agenda política, económica y social de México.

La situación en la economía internacional generada por la presente crisis financiera internacional y desequilibrios macroeconómicos en varias regiones del mundo en un contexto de contracción económica global, el problema de la generación de empleos se enfatiza, porque además de atender el problema estructural de la falta de empleos, también se debe hacer frente al crecimiento de los despidos y mayor tasa de desempleo en general.

En las últimas dos décadas, los diferentes niveles de gobierno han buscado nuevas estrategias dentro de la política económica para promover el crecimiento económico y en consecuencia la generación de empleos. Una de ellas ha sido el impulso a la apertura de nuevos negocios dado que las grandes empresas y mucho menos el propio gobierno pueden ofertar las plazas de trabajo demandadas por la creciente población en edad de laborar.

La incubación de empresas ha recibido por tal motivo, un apoyo y promoción especial de parte de gobierno y las entidades públicas y privadas, a la pequeña y mediana empresa, así como de las universidades, cada uno han comenzado a desarrollar programas y a consolidar los ya existentes para apoyar esta iniciativa nacional. En este sentido se encuentra la Ley para el Desarrollo y Competitividad de las pequeñas y medianas empresas en México, a través de la cual se apoya la creación, desarrollo, capacitación y consultoría de nuevas empresas en el país,

[1] El Estado de México se encuentra en el centro de México, rodea al Distrito Federal (donde se asienta la ciudad de México). El estado presenta grandes contrastes sociales y económicos. La zona conurbada al Distrito Federal es una gran concentración de población y de actividad productiva. Esta zona metropolitana, Distrito Federal y Estado de México, concentra alrededor de 20 millones de habitantes y casi una cuarta parte de la actividad económica del país. Sin embargo, al sur del Estado de México la situación es diferente, es la zona de menor desarrollo relativo y con poca concentración poblacional y económica.

buscando aprovechar la gran oportunidad de que la económica se encuentra al intercambio comercial de bienes y servicios dentro del marco de la apertura comercial[2].

Las denominadas incubadoras de empresas son alternativas que fueron ideadas para crear un ambiente de mayor protección para la creación e implementación de nuevas empresas. Existe otro concepto relacionado y que nace con el cobijo de algunas universidades con fuerte inclinación a la investigación y desarrollo en países desarrollados, este es el de Parques Tecnológicos, ideado como mecanismo de vinculación con la empresa, y además canalizar el espíritu empresarial de sus estudiantes. En el contexto económico, las incubadoras existen para dar soporte a la transformación de potenciales emprendedores en empresarios consolidados y empresas con potencialidades en empresas crecientes y lucrativas. Al reducir los riesgos durante el período inicial de formación de una empresa, las incubadoras pueden contribuir al crecimiento económico y la reactivación regional a través de nuevas empresas y empresarios asociados.

Las incubadoras de empresas son una modalidad de emprendimiento empresarial, que brinda una infraestructura mínima y bajos costos a la persona que tiene la voluntad y el deseo de desarrollar su propio negocio. Es relevante mencionar que las incubadoras de este tipo son entendidas como ambientes especialmente estructurados para apoyar el desarrollo de pequeñas unidades empresariales colocando a su disposición recursos humanos, competencias, instalaciones e infraestructura administrativa/operacional, estimulando el surgimiento de empresas modernas y competitivas, capacitadas tanto en el ámbito tecnológico como gerencial.

En términos generales, las incubadoras deben procurar con el cumplimiento de los siguientes objetivos específicos:

i) Apoyo técnico y gerencial a empresas, a través de identificación y ofrecimiento de consultorías especializadas;

ii) Aceleración de la consolidación de empresas, ayudándolas a superar mas rápidamente las barreras técnicas, gerenciales y de mercado;

iii) Fortalecimiento de la capacitación emprendedora ofreciendo ambiente apropiado y gerencia dinámica;

iv) Desarrollo de acciones asociativas y cooperativas;

v) Reducción de costos para las empresas y para las entidades que apoyan a la incubadora;

vi) Búsqueda de nuevos apoyos (socios estratégicos);

vii) Divulgación de las empresas facilitando la participación en redes y acceso a experiencias en el sector;

viii) Aumento de la interacción entre el sector empresarial y las instituciones académicas;

ix) Contribución a la generación de utilidades y empleo.

Las incubadoras de empresas son un espacio configurado para acoger y amparar a micro y pequeñas empresas que tienen algún perfil emergente. La esencia es la transformación ideas en productos, procesos en servicios, es decir, trata de ser un condominio donde micro y pequeñas empresas, ya sean nacientes o ya estructuradas, reciben condiciones ideales para su consolidación por un período aproximado de 2 años. Posteriormente, la empresa deja este espacio para continuar su proceso de desarrollo y enfrentar desde otra perspectiva la realidad del mercado.

En México, las instituciones de educación superior, institutos tecnológicos y universidades se han insertado casi desde el principio en la operación de incubadoras de empresas. El principal motivo consiste en que las características de una incubadora permiten la vinculación de una universidad con los sectores productivos y con ciertos sectores de la población. Así, es una vía por la cual las instituciones de educación superior participan en el progreso económico y bienestar social.

[2] En México el proceso de creación de incubadoras de empresas o negocios, también así llamadas, se remonta a 1990 con los primeros esfuerzos en Ensenada, Baja California. En esta ciudad donde se creó la primera incubadora formal de Empresas de Base Tecnológica, con la participación del principal organismo de fomento económico del país, Nacional Financiera (NAFINSA), el Consejo Nacional de Ciencia y Tecnología (CONACYT), y el Centro de Investigación Científica y de Educación Superior en Ensenada.

Las incubadoras universitarias, además de vincular la realidad de la universidad a la perteneciente a la empresa, genera espacios que producen una readecuación general en los procesos, se fomenta el desarrollo de interesantes líneas de investigación, cambios en los programas curriculares, etc.; sobre todo la creación de un espíritu que fomenta la actitud emprendedora en estudiantes y académicos. La experiencia de incubadoras universitarias de empresas en Estados Unidos, Canadá y Europa permiten ubicar su importancia. En general, la experiencia ha demostrado que las universidades son por su naturaleza un catalizador importante de las inquietudes de alumnos, académicos y egresados, dada su condición de creadora de conocimiento.

En la Universidad Autónoma del Estado de México, en el marco de colaboración de ésta con el País Vasco, gobierno y empresas, en el año de 1998 se crea el modelo de incubadora de empresas de tecnología intermedia denominado INCUBASK. Dicho modelo es una metodología diseñada para llevar a cabo el proceso de incubación de negocios cuya característica es la innovación el procesos de producción o comercialización, además de innovaciones en el producto o servicio.

En la primera incubadora ubicada en Técamac, zona metropolitana de la ciudad de México, los primeros años no fueron sencillos. Ello debido a que no existe una cultura o instrucción difundida de ser emprendedor y de no hacer planeación. Para la segunda mitad de la presente década se da un nuevo empuje desde la universidad y desde los diferentes niveles de gobierno. Las incubadoras de empresas quedan insertadas en las estrategias de política pública de fomento a la generación de empleos.

En años recientes la UAEM ha registrado un crecimiento acelerado del número de incubadoras y ha constituido una red de ellas, donde la perteneciente a Tenancingo es la décima. De esa forma se cubre casi todo el territorio mexiquense que coincide con la presencia universitaria. Una de dichas incubadoras es de alta tecnología, es decir, se incuban negocios con aplicaciones de innovaciones en el conocimiento. Cada una de las incubadoras de la red emplean el modelo INCUBASK el cual está registrado y reconocido ante la Secretaría de Economía del gobierno federal.

3. LA SITUACIÓN ECONÓMICA Y SOCIODEMOGRÁFICA DE LA REGIÓN SUR DEL ESTADO DE MÉXICO

3.1. Entorno económico

El Estado de México es la segunda economía más grande del país sólo superada por el Distrito Federal. En 2007, la región VI[3] representa sólo el 1.30% del total de la economía mexiquense. A pesar de constituir un pequeño porcentaje en términos económicos, la región tiene una importante extensión territorial, población dispersa con poca densidad poblacional y con necesidades de desarrollo económico. La región sur del Estado se distingue por ser expulsora de población que busca oportunidades de trabajo en las grandes ciudades y en Estados Unidos[4], de ahí la importancia de fomentar el crecimiento económico con la generación de empleos. Por municipios, aquel que destaca es Tenancingo representando un 0.34% del PIB estatal, el resto representa porcentajes muchos más pequeños.

Del total del PIB regional, el municipio de Tenancingo tiene la mayor importancia al representar el 26.30%, seguido de Villa Guerrero con 17.37%, Ixtapan de la Sal y Coatepec Harinas con casi 10%, los cuatro municipio representan dos tercios de la actividad económica regional y el resto se constituye por los otros 11.

La actividad económica es posible analizarla por medio de los sectores productivos, primario secundario y terciario, los cuales constituyen al PIB. Del total regional, el sector

[3] La regionalización realizada por el gobierno estatal indica que los municipios pertenecientes a esta región son: Almoloya de Alquisiras, Coatepec Harinas, Ixtapan de la Sal, Joquicingo, Malinalco, Ocuilan, San Simón Guerrero, Sultepec, Temascaltepec, Tenancingo, Texcaltitlán, Tonatico, Villa Guerrero, Zacualpan y Zumpahuacán. Cabe mencionar que la actual región VI corresponde a la región XI dentro de la regionalización anterior realizada por el gobierno estatal.

[4] Un ejemplo es el municipio de Tonatico, del cual se menciona que casi la mitad de su población nacida originalmente ahí se encuentra radicando fuera de su territorio, en otras entidades y principalmente en Estados Unidos.

agropecuario representa 47.22%, el sector servicios 36.24% y el sector de la transformación 8.79%. Cabe señalar que el municipio de más relevancia, Tenancingo, muestra un distribución diferente: el sector servicios participa con 56.31%, el sector agropecuario con 25.67% y la industria sólo 10.25%. Este municipio se constituye como un polo regional para la prestación de servicios de todo tipo. En general, la región es de perfil de producción primaria seguida de la terciaria, situación totalmente diferente en relación a las regiones conurbadas al Distrito Federal o dentro del Valle de Toluca.

En el 2000 del total de la población de 12 años y más, en el municipio de Tenancingo la población económicamente activa (PEA) representa el 48.90% y la inactiva (PEI) el 51.10%. Del total de la PEA, sólo una pequeña fracción representa a los desocupados, 1.00%, en tanto la gran mayoría se reporta como ocupados. La distribución del empleo por sectores productivos indica se concentra en el sector servicios en el cual se ocupa al 47.40% de la población activa, en segundo término está el sector primario con 26.40% y el secundario sigue de cerca con 24.00%. Cabe señalar que el sector industrial aporta sólo 10.25% a la actividad económica pero hace uso intensivo de trabajo pues emplea a una cuarta parte de la PEA.

3.2. Entorno sociodemográfico

La región VI representó en el año 2000 sólo 3% del total de la población del Estado de México, cerca de 392 000 habitantes, cifra que disminuye en 2005 al mostrar una participación menor la cual llega al 2.8%, cerca de 386 000 habitantes. La tasa de crecimiento poblacional hasta el año 2000 se ubicaba alrededor del 2.2% anual, pero la situación cambia y para el periodo entre 2000 y 2005 la población muestra una decremento anual de 0.2%.

Una primera idea que explica esta reducción del crecimiento poblacional es el índice de masculinidad por rangos de edad (número de hombres por cada 100 mujeres). En los primeros años de vida, hasta los 14 años, se presenta una relación casi proporcional, es decir, al número de hombres corresponde una cifra similar de mujeres. Esta situación cambia en la edad laboral, entre los 15 y los 40 años el índice se ubica en alrededor de 90 hombres por cada 100 mujeres, el caso extremo es en el rango de edad entre 25 y 29 años pues el indicador cae hasta una relación de casi 80 hombres por cada 100 mujeres. Este indicador aporta una idea sobre la reducción del crecimiento poblacional: es debida a la emigración de los habitantes de la región hacia otras dentro y fuera del país, ello motivado por razones de búsqueda de mejores expectativas de trabajo y los ingresos derivados. En 2005, por rangos de edad, la población se concentra entre los 15 y 64 años con un 50.7%, edades que se distinguen por ser las correspondientes a la fase laboral de los individuos. La edad promedio de la población es de 23.2%, esto significa una población joven que conlleva necesidades para ese segmento de la sociedad.

El grado de urbanización es del 31.5% de la población, en otras palabras, sólo una tercera parte de la población habita en poblados grandes o en ciudades. La densidad poblacional es de 111 personas por kilómetro cuadrado, nivel inferior al estatal con 623 habitantes por km^2. Para el año 2000 respectos a las condiciones de la vivienda, el porcentaje de ocupantes de viviendas con drenaje y sanitario es el 67.2%, el porcentaje de ocupantes en viviendas con energía eléctrica es de 94.3%, el porcentaje de ocupantes en viviendas con agua entubada es de 81%, el porcentaje de ocupantes con piso diferente de tierra es de 75.2%.

En cuanto a educación, el grado de escolaridad por municipio es variado, pero en general se encuentra alrededor de los 6 años. El porcentaje de población con alfabetización mayores de 15 años es de 85% y el porcentaje de la población mayor de 15 años con primaria terminada es el 57.2%.

3.3. Necesidad de la incubadora de empresas ante dicho contexto económico y social

En este contexto económico y social se hace necesario impulsar en la región el crecimiento económico y la generación de empleos por parte de diferentes actores. La universidad pública es uno de ellos, dado que tiene como objetivo tener una vinculación con los sectores productivos y sociales para beneficio de éstos.

La creación de una incubadora de empresas dentro de la UAEM es una forma de contribuir en el crecimiento económico de la región, la generación de empleos, el impulso al desarrollo

social y económico, con ello contribuir a frenar fenómenos como los bajos niveles de ingreso y pobreza, la migración de la población joven en busca de mejores empleos, la falta de oportunidades de plazas de trabajo, la falta de arraigo a su lugar de origen, el desempleo de los egresados universitarios, la falta de una cultura y formación emprendedora y empresarial en la población en la edad laboral de la población, además la insuficiente de vinculación de la universidad pública con los sectores productivos y los niveles de gobierno.

Algunas acciones que pueden realizarse bajo el ámbito de la incubadora de empresas es: apoyo técnico y gerencial a empresas, a través de identificación y ofrecimiento de consultorías especializadas; aceleración de la consolidación de empresas, ayudándolas a superar barreras; fortalecimiento de la capacitación emprendedora ofreciendo ambiente apropiado y gerencia dinámica; desarrollo de acciones asociativas y cooperativas; reducción de costos para las empresas y para las entidades que apoyan a la incubadora; búsqueda de nuevos apoyos; divulgación de las empresas facilitando la participación en redes y acceso a experiencias en el sector; aumento de la interacción entre el sector empresarial y las instituciones académicas; contribución a la generación de utilidades y empleo.

La UAEM tiene como una de sus funciones sustantivas la vinculación y extensión con la sociedad y los sectores productivos. En la actualidad está plasmado en el Plan Rector de Desarrollo Institucional[5].

Así, la vinculación con los sectores productivos de la UAEM está planificada y de esa forma puede contribuir desarrollo económico y social de la región sur de la entidad mexiquense, responsabilidad que tiene al ser la máxima casa de estudios del Estado. La comunidad universitaria y los sectores productivos cuentan con la UAEM para impulsar dicha cultura empresarial y emprendedora.

4. SOPORTE DE LOS SECTORES GUBERNAMENTAL, SOCIAL Y ECONÓMICO

Los proyectos que diseñan y llevan a cabo las instituciones educativas y los diferentes niveles de gobierno, incluso las organizaciones sociales y las empresas, requieren del apoyo del entorno social y económico en donde se asentará el proyecto y cuyos efectos beneficiosos serán su aporte a esa parte de la sociedad. Existen problemas en el logro de los objetivos cuando las ideas de los proyectos no cuentan con el financiamiento necesario, no tienen al personal técnico o especializado, no cuentan con el conocimiento suficiente del medio donde operan, no tienen el respaldo o consenso social, entre otras problemáticas.

Los diferentes actores dentro de la sociedad y los sectores productivos consideran que una vía para apoyar el desarrollo regional tanto económico como social es la incubadora de empresas. De manera aislada difícilmente tendría la capacidad financiera y de apoyo social, en consecuencia, interactúa con otras instituciones y sectores sociales y económicos. En concreto, hay varias fuentes de apoyo, las cuales son:

La UAEM a través de diferentes esferas. En primer lugar, la Secretaría de Vinculación y Extensión por medio de la dirección de Vinculación y Desarrollo Empresarial los cuales contribuyen con la orientación y la gestión de trámites y recursos ante los diferentes niveles de gobierno. En segundo término, la transferencia del modelo de incubadora de empresas por parte de INCUBASK, la cual pertenece a la misma Secretaría. Además el Centro Universitario UAEM Tenancingo acoge a la incubadora en sus instalaciones, recursos materiales y aporta el capital humano especializado en la tutoría empresarial.

El gobierno federal a través de la Secretaría de Economía aportará recursos financieros para la puesta en marcha de la incubadora de empresas, los términos de la transferencia de recursos están acordes a las convocatorias y los resultados de los concursos para obtener dicho financiamiento. Además durante la operación de la incubadora, la dependencia federal hará

[5] En específico, en el eje *Vinculación y extensión para una sociedad mejor* la universidad tiene el proyecto de vinculación redituable. Su objetivo es vincular el quehacer de la UAEM y favorecer el desarrollo de una cultura emprendedora y empresarial de los universitarios generando alianzas estratégicas de los diferentes sectores de la sociedad en los ámbitos estatal, nacional e internacional, promoviendo y difundiendo los productos y servicios en la comunidad universitaria.

aportaciones para el pago de las asesorías empresariales de tal forma que los emprendedores y empresarios tengan acceso a estos servicios.

De manera concreta, la inversión inicial sobre la construcción de infraestructura, edificio y su equipamiento, se reparte entre la UAEM y la Secretaría de Economía. El gobierno federal aporta recursos para la construcción del edificio y su equipamiento. La UAEM aporta el diseño del edificio, el cual corresponde al modelo INCUBASK y su equipamiento corresponde a las características del edificio.

La UAEM contribuye en la inversión inicial con la aportación de terreno para la construcción del edificio, el cual estará en las propias instalaciones del Centro Universitario. Además durante la operación de la incubadora, la misma universidad se hará cargo de los salarios de los profesionistas y secretaria que operarán la incubadora, el mantenimiento y la vigilancia del edificio que incluye la limpieza y el pago de servicios públicos.

Existen otros agentes económicos y sociales que han mostrado interés en el proyecto de la incubadora, sin embargo, falta formalizar el tipo de aportaciones y los montos monetarios de éstas.

El gobierno del Estado de México a través del Instituto Mexiquense del Emprendedor, IME, ha apoyado en ocasiones pasadas los proyectos de incubadora de empresas de la UAEM, se sabe del interés del IME por seguir trabajando de forma conjunta con la universidad, además de ser un aportante de recursos importantes para la puesta en marcha del proyecto en Tenancingo. se están generando los acuerdos necesarios para que este nivel de gobierno participe de forma importante.

El Ayuntamiento de Tenancingo está interesado en que opere la incubadora y ha manifestado su disposición de participar en el proyecto. Por ello, la UAEM y el ayuntamiento han firmado recientemente este año un acuerdo de cooperación general entre ambas partes y un acuerdo operativo entre el Centro Universitario como parte de la UAEM y el ayuntamiento. La finalidad de los acuerdos consiste en que la universidad contribuya en el apoyo de ciertos empresarios y emprendedores, en especial micro y pequeños productores, los cuales se han acercado al ayuntamiento para solicitar ayuda. El mecanismo adecuado para este objetivo es la incubadora de empresas. El ayuntamiento aportará las facilidades fiscales de su competencia además de algunos servicios como la excavación y remoción de tierra durante la construcción del edificio. Además ha generado una carta donde expresa su interés por el proyecto de la incubadora de empresas y su respaldo al mismo.

En una serie de reuniones con los integrantes de los sectores productivos, algunos productores o comercializadores ha expresado su interés por el funcionamiento de la incubadora de empresas, correspondientes a diferentes ramos como el artesanal, agropecuario y manufacturero. Se han generado cartas en donde ellos expresan su respaldo al proyecto de la apertura de una incubadora de empresas en la región. Se han mencionado por parte de ellos las diferentes necesidades concretas y otras por definir. De esa forma definir los tipos de apoyo por realizar y trabajar de manera conjunta con la universidad y los diferentes niveles de gobierno.

5. COMPOSICIÓN Y TAMAÑO DE MERCADO

Sin importar el tipo, las incubadoras de empresas tienen un mercado objetivo enorme y con dispersión a nivel nacional y regional. Esta situación se deriva del constante crecimiento de las tasas de desempleo, por tal motivo se observa que cada vez mayor cantidad de personas deciden iniciar su propia empresa ya sea de forma individual o en agrupaciones. Los casos son diversos, uno de ellos es por convencimiento como algunos egresados de universidades que deciden incorporarse al sector productivo participando como empresarios, por otro lado existe un gran número de personas que ven el hecho de poner un negocio propio como una necesidad de sobrevivencia, además de considerar que hay un gran número de empresarios que ya cuentan con una empresa consolidada y desean incursionar en otros negocios, en ocasiones en giros complementarios, ellos también son un nicho potencial de usuarios de la incubadora de empresas.

La demanda de servicios de consultoría empresarial es muy amplia en la región de Tenancingo. Esta se puede clasificar en universitaria y no universitaria, la primera corresponde a la comunidad universitaria en general y la segunda refiere a las personas o unidades económicas que participan o buscan participar en los sectores productivos. Los servicios especializados se pueden ofrecer a través de una incubadora de empresas.

Primeramente se abordan a los potenciales demandantes universitarios. El Centro Universitario en general cuenta con 443 alumnos inscritos, 62 profesores y 27 trabajadores administrativos. De forma desagregada hay 88 alumnos en la licenciatura en Arqueología, en Ingeniero en Floricultura son 97, en licenciatura en Turismo son 157 y en Relaciones Económicas Internacionales 101. Del total de profesores, 7 son de tiempo completo, 1 técnico académico y 54 son de asignatura. Del total de personal administrativos 17 son de confianza, un directivo y 9 son sindicalizado. Ellos pueden tener como opción de emprender o mejorar un negocio propio o en asociación de familiares y amigos.

Los alumnos inscritos y egresados se integrarán al mercado laboral y están en búsqueda de un puesto de trabajo. El Centro Universitario en 2008 tiene sus primeros alumnos egresados de la licenciatura en turismo y de ingeniero agrónomo en floricultura, con 35 y 24 alumnos respectivamente. Los egresados en 2009 en turismo serán 22 y en floricultura 13, en tanto para 2010 en el primer caso serán 31 y en el segundo 20, además la licenciatura de relaciones económicas internacionales en ese año tendrá su primera generación egresada con un total de 28 alumnos. Ellos constituyen la demanda potencial de usuarios universitarios de la incubadora de empresas, pues han manifestado que buscan opciones diferentes para insertarse al mercado laboral y no sólo tienen como objetivo ser empleados.

A partir del primer año de operación, del total de integrantes de la comunidad universitaria se proyecta atender al menos 5 planes de negocios. Dicha cifra se pretende mantener durante los siguientes años: 5 nuevos planes de negocios cada año y 5 planes de negocios consolidados cada año. Es decir, 10 planes se estarán apoyando por año, con excepción del primer año de arranque.

Así, la UAEM queda en posibilidad de brindar a la comunidad universitaria de Tenancingo una opción de desarrollo personal y colectivo, una fuente de empleo y de mejora en las condiciones de vida.

Los no universitarios se refieren principalmente a los participantes en la actividad económica. Las unidades económicas registradas en Tenancingo dentro del más reciente censo económico (2004) son 2724, de las cuales el comercio al por menor representa el 51.36%, seguido de las industrias manufactureras con 15% y en tercer término, servicios de alojamiento temporal y de preparación de alimentos y bebidas con casi 9%, misma cifra corresponde a los servicios de esparcimiento. Además de concentrar una parte importante de las empresas, también la generación de plazas de trabajo. El personal ocupado 7205, se concentra en el comercio al por menor con 41.54%, seguido de las industrias manufactureras con 20% y en tercer lugar el comercio al mayoreo con 8.76%. En el rubro de propietarios, familiares y meritorios cuyo total es 4430, el comercio al por menor representa el 45.73%, seguido de las industrias manufactureras con 22.33% y en tercer término, servicios de alojamiento temporal y de preparación de alimentos y bebidas con 10.32%. Este segmento es de considerarse con énfasis debido a que son los que toman decisiones sobre el futuro de sus negocios. En consecuencia, este segmento debe tomarse como la demanda potencial del sector productivo de dichos servicios de asesoría empresarial.

El sector agropecuario en estos censos económicos casi no aparece registrado, los sectores industrial y de servicios están más representados. Con datos que declara el ayuntamiento de Tenancingo, en el municipio existen alrededor de 1500 productores agrícolas, en especial en el sector florícola. De tomarse en cuenta dicha cifra, la demanda potencial aumenta debido a que la mayoría de ellos son pequeños productores que no operan en condiciones adecuadas que les permitan ser competitivos más allá de los mercados locales.

En general, el municipio de Tenancingo se caracteriza por ser un centro regional para la actividad económica, en especial el comercio, en menor medida la producción florícola[6].

[6] Las actividades del sector terciario que se conforman por todo tipo de servicios, el municipio destaca por prestar servicios, educativos, de salud, financieros, comunales entre otros, además del comercio, el cual está dominado por el comercio de productos

De manera concreta, el trabajo conjunto del Centro Universitario y el ayuntamiento de Tenancingo permite identificar grupos dentro de los sectores productivos que están interesados en los servicios de la incubadora de empresas. Uno de ellos es un grupo de emprendedores que pretende usar desechos orgánicos vegetales para obtener composta y así comercializar fertilizantes; otro ejemplo es un grupo de productores de muebles quienes además de mejorar la comercialización de los productos, desean mejorar el diseño de sus muebles; un caso más son los productores de flores que tratan de darle un mayor valor agregado a sus productos y no sólo comercializarlos sin algún tratamiento adicional.

La atención es importante a este segmento, las unidades económicas o empresas por apoyar son 5 planes de negocios iniciados por año y 5 planes por consolidar cada año. En total se atenderán 10 planes de negocios anualmente, con excepción del primero que sólo serán 5. El giro de las unidades económicas es diverso y con necesidades específicas lo que no se convierte en un obstáculo.

En concreto, con la incubadora de empresas la universidad tiene el propósito de apoyar a su comunidad y a los sectores productivos para que está vinculación contribuya a que la cultura empresarial y emprendedora tenga como resultado la generación de empleos en la zona y mejorar las condiciones de vida.

El segmento de mercado por atender tiene características muy particulares, entre las que se encuentran:

• Alumnos que han cursado una asignatura o asignaturas relacionadas al tema emprendedor, incluso han comenzado a realizar proyectos de planes de negocios o proyectos de inversión, esto con la finalidad de aprobar el curso.

• Alumnos que han mostrado interés por darle continuidad a sus proyectos una vez de haber cursado y aprobado sus asignaturas correspondientes.

• Alumnos que tienen claridad en la forma en como proyectan laborar cuando terminen sus estudios, muestran intenciones de tener su propia empresa y con ello generar sus plazas de trabajo.

• Alumnos que han crecido en un ambiente familiar o en un contexto cercano a la operación de negocios de diferentes tamaños, ya sea porque sus familiares son los propietarios o han laborado en algún establecimiento de este tipo.

• Alumnos que han mostrado características de personalidad pertenecientes a un emprendedor: liderazgo, trabajo en equipo, creatividad, iniciativa, poco temor al riesgo y a lo nuevo o desconocido, actitud positiva, entre otros rasgos.

Cabe hacer énfasis en que dentro de sus planes de estudio, cada carrera contiene unidades de aprendizaje que fomentan su actividad dentro de las empresas y el desarrollo emprendedor así como los planes negocios y proyectos de inversión. Sin embargo, no existe un seguimiento de los planes, proyectos e ideas que desarrollan los alumnos en sus cursos, dado que termina el semestre y ahí se corta el trabajo formal en ellos. Los alumnos manifiestan la necesidad de tener un programa de trabajo permanente que les permita continuar con sus proyectos con la finalidad de consolidarlos. Además consideran que el impulso de sus proyectos no podrían hacerlos de manera individual por no contar con los conocimientos, la experiencia y los recursos materiales y financieros suficientes.

El mercado potencial de los no universitarios se refiere principalmente al universo de participantes en la actividad económica, personas con actividad empresarial o empresas (denominadas unidades económicas éstas últimas). Las unidades económicas registradas en Tenancingo dentro del más reciente censo económico (2004) son 2724, de las cuales el

no alimenticios (ropa, zapatos, muebles); de alimentos (panaderías, carnes y frutas); comercio de bebidas y tabaco (cantinas y bares) y comercio especializado (refaccionarías, ferreterías y tiendas de importación).

En el sector secundario, existen las actividades relacionadas con la industria extractiva, la manufactura, la electricidad y la construcción, hay algunas industrias dentro del municipio que se clasifican en medianas y pequeñas donde se produce artículos de herrería, carnes frías, textiles y materiales para la construcción.

En el sector primario son relevantes actividades como la agricultura con el cultivo de frutos, y es muy significativo el de flores así como el cultivo de granos, además la ganadería en pequeñas explotaciones y de traspatio. Cabe mencionar que la floricultura en el municipio y en la región está ganando importancia y desplazando a otros cultivos. También se realiza la explotación forestal y el cultivo de peces.

comercio al por menor representa el 51.36%, seguido de las industrias manufactureras con 15% y en tercer término, servicios de alojamiento temporal y de preparación de alimentos y bebidas con casi 9%, misma cifra corresponde a los servicios de esparcimiento. La producción bruta total es realizada principalmente por el comercio al por menor ya que representa 44.45%, el comercio al por mayor con 18% y en tercer lugar industrias manufacturas con 17.42%. En el rubro de propietarios, familiares y meritorios cuyo total es 4430, el comercio al por menor representa el 45.73%, seguido de las industrias manufactureras con 22.33% y en tercer término, servicios de alojamiento temporal y de preparación de alimentos y bebidas con 10.32%. Por su parte, el personal ocupado 7205, se concentra en el comercio al por menor con 41.54%, seguido de las industrias manufactureras con 20% y en tercer lugar el comercio al mayoreo con 8.76%.

Se observa que el giro empresarial que tiene más relevancia por unidades económicas, personal ocupado, propietarios e incluso por producción es el comercio al por menor que representa casi la mitad en cada rubro. Es el segmento con mayor presencia en la actividad económica y en consecuencia es al cual se podría cubrir inicialmente. Las autoridades del ayuntamiento de Tenancingo han manifestado que algunos sectores productivos están interesados en mejorar sus condiciones de producción y/o comercialización, ya sea para consolidar o expandirlos. En especial están en esa búsqueda de mejores condiciones los pequeños propietarios de unidades económicas concentradas en el comercio al por menor y en la industria manufacturera.

El segmento de mercado por atender tiene características muy particulares, entre las que se encuentran:

• Empresarios que han detectado necesidades diversas de sus negocios, están interesados y comprometidos en recibir asesoramiento para la mejora de la operación de sus negocios.

• Empresarios que están en búsqueda de expansión de mercados o al menos consolidación en los que tiene presencia en la actualidad.

• Empresarios que tienen la necesidad de mejorar sus condiciones de comercialización (etiquetado, embalaje y empaquetado, transportación y almacenamiento).

• Empresarios que desean hace innovación en el diseño de su productos y ser competitivos con productos diferenciados.

• Empresarios que tienen la necesidad de incorporar mayor valor agregado de sus productos, con ello obtener mayores ingresos, además de llegar a nuevos compradores.

La demanda no ha sido satisfecha, por diversas causas:

i) En esta parte de la región sur del Estado de México, ninguna incubadora de empresas funciona en la actualidad, las más cercanas se encuentran en la ciudad de Toluca, lo cual implica una hora de desplazamiento aproximadamente.

ii) En la región no existe el número suficiente de profesionistas que ofrezcan los servicios de consultoría o asesoría a empresas y emprendedores.

iii) Además, los precios de las consultorías o asesorías son relativamente altos en comparación al poder adquisitivo de la población, en ocasiones se vuelve prohibitivo el hecho de asesorarse.

iv) La información sobre servicios de consultoría y asesoría para empresas y emprendedores que puede ofrecer una incubadora de empresas, en general es escasa o casi nula.

v) Los personas o empresas han recibido una formación cultural sobre los apoyos gubernamentales o de otro tipo de forma errónea, se considera que son transferencia de recursos ya sean subsidios o recursos a fondo perdido. Así esperan primero dinero que apoyo técnico.

vi) Las instituciones educativas de la región, entre ellas la UAEM, no han realizado los vínculos suficientes para apoyar a los sectores productivos en la solución de sus problemáticas.

vii) Las instituciones educativas de la región, entre ellas la UAEM, no han formado estudiantes y egresados con capacidades y habilidades suficientes para que forma de inserción al mercado laboral sea la asesoría y consultoría a los empresarios y emprendedores.

Ambos ámbitos, el universitario y no universitario coinciden en manifestar su intención de tener mejores condiciones en su operación, sin embargo no existe un ente especializado que sirva de guía o asesor. Es por ello que una incubadora de empresas en la región servirá para cubrir la gran demanda insatisfecha.

6. SERVICIOS REQUERIDOS POR LOS USUARIOS

Las necesidades de servicios de los usuarios potenciales de la incubadora de empresas son muy amplias, cada sector productivo tiene requerimientos específicos. De acuerdo con la aportación al PIB del municipio de Tenancingo, es importante considerar la situación en la que se encuentra el sector del comercio y de la industria manufacturera, sin dejar de lado al sector florícola. En general, se tienen problemas en la organización de las actividades dentro de la empresa ya sea producción o comercialización, además en los productos o servicios se carece de calidad, innovación, entre otros.

Con la finalidad de ejemplificar la situación particular de un sector productivo se toma el caso del sector florícola. El Centro Universitario ha generado resultados de investigación sobre la situación de la región, Ramírez Hernández (2008) indica que la producción de flores en México y en el sur del Estado presenta características identificadas, aunque no bien cuantificadas[7]. La comercialización interna de flores, alrededor del 90% de la producción nacional se destina a los mercados nacionales y el resto es enviado a los mercados internacionales[8].

En términos generales, las unidades económicas registradas en Tenancingo dentro del más reciente censo económico (2004) son 2724, de las cuales el comercio al por menor representa el 51.36%, seguido de las industrias manufactureras con 15% y en tercer término, servicios de alojamiento temporal y de preparación de alimentos y bebidas con casi 9%, misma cifra corresponde a los servicios de esparcimiento. Se ejemplifican la situación de otros sectores relevantes.

El sector del comercio al por menor se caracteriza por emplear poca cantidad de personal, regularmente quienes trabajan es el dueño, algún empleado y con frecuencia los familiares laboran sin recibir salario (ejemplo de esto son las misceláneas y tiendas). El número unidades económicas que están en posibilidad de expansión es importante dado los niveles de inversión, sin embargo no lo realizan o no alcanzan a lograrlo. El sector de manufacturas tiene como rasgos distintivos el hecho de que las unidades económicas relativamente emplean más personal en comparación con el comercio al por menor, sin embargo, la mayoría de las unidades económicas son micro o pequeñas empresas. Las posibilidades de expansión son limitadas y muestran problemas en los canales de comercialización tanto por cuestiones de logística como por características propias del producto (estándares de calidad, empaquetado y embalaje, entre otros). Por otra parte, Ramírez Hernández (2007) en un estudio del mercado turístico en el municipio de Tenancingo, los hallazgos señalan que los visitantes buscan mejoras en la infraestructura urbana, señalamientos, pavimentación, estacionamientos, pero además mayor número de actividades recreativas y en consecuencia unidades económicas que los ofrezcan: establecimientos de alimentos , balnearios, algún lugar como una discoteque o cines.

Estos planteamientos señalan todos los posibles rubros que se pueden atender dada la forma en como operan los sectores, las demandas de los consumidores y las posibilidades de reingeniería y expansión planificada, el punto medular es hacerlas realmente con este enfoque.

Dada la situación anterior algunos de los servicios requeridos por los sectores productivos están en la tabla siguiente:

[7] 1. Baja productividad, 2. Baja calidad (para los estándares internacionales), 3. Poca innovación (en cualquier rubro), 4. Falta de inversión de capitales, por tanto, baja tecnificación, 5. Resistencia a asociarse, se prefiere trabajar individualmente, 6. Casi nula investigación y desarrollo, 7. Preferencia por el mercado interno, 8. Uso ilegal de semillas y material de propagación, 9. Desconocimiento de los requisitos para exportar, 10. En general, falta de capacitación a los productores, 11. Infraestructura de transporte y refrigeración obsoleta.

[8] Las características principales son: 1. Inexistencia de índices de calidad, 2. En el manejo post-cosecha las flores son atadas o depositadas en cajas y trasladadas en camiones descubiertos, junto con otros productos, repercutiendo en la vida de anaquel del producto, 3. La comercialización de las flores se realiza de diferentes formas según el tipo de productor y el producto, 4. En general, el pequeño productor acude a los mercados mayoristas como la Central de Abastos del D.F. y mercado de flores de Tenancingo, donde venden el producto a intermediarios y mayoristas de diferentes ciudades de la República, siendo las más importantes: Guadalajara, Monterrey, Culiacán, Cancún, Acapulco.

Tabla 1. Incubadora de empresas UAEM Tenancigo: Servicios a ofertar

Asesoría en comercialización y promoción de productos manufactureros (incluyendo artesanías).
Asesoría en el diseño de nuevos productos o en su caso, un rediseño de los ya existentes.
Asesoría en el diseño de empaques, embalajes y etiquetado de productos manufacturados y no procesados.
Asesoría en la comercialización y promoción de los atractivos turísticos.
Asesoría en el análisis de mercado tanto locales como nacionales e internacionales para los productos florícolas.
Asesoría en el análisis de mercado tanto locales como nacionales de los productos manufacturados.
Asesoría en el análisis de mercado tanto locales como nacionales para los atractivos turísticos, así como el desarrollo de nuevos.
Asesoría en la producción para elevar los niveles de productividad por hectárea cultivada en sector florícola.
Asesoría en la producción florícola con buenas prácticas agrícolas (cuidado y mejoramiento del entorno natural, social y del sector productivo).
Asesoría y orientación sobre los estándares de calidad de los productos florícolas de acuerdo con la normatividad nacional e internacional.
Asesoría y orientación sobre los estándares de calidad de los productos manufactureros (ejemplo muebles, artesanías de diferente tipo) de acuerdo con la normatividad de las empresas dedicadas al comercio (ejemplo centros comerciales o tiendas departamentales).
Asesoría sobre actualización o introducción, según sea el caso, de organización y funcionamiento de las actividades de las unidades económicas.
Asesoría en la solicitud o concurso de recursos de las unidades económicas por recursos de los diferentes niveles de gobierno.
Asesoría y capacitación para el trabajo en equipo y administración de recursos de uso colectiva.
Asesoría técnica en la producción florícola.
Asesoría específica, ya sea para los insumos como para los productos que se distribuyen.
Asesoría para el uso legal de material vegetal.
Asesoría en los trámites de importación y exportación, sea de insumos o de producto terminado.
Asesoría en los trámites y procedimientos para la apertura de empresas.
Asesoría financiera, contable y económica sobre la administración de los recursos de la unidad económica.
Asesoría en todo el proceso de elaboración de un proyecto de inversión.
Asesoría en todo el proceso de elaboración de un plan de exportación.
Capacitación continua en temas específicos de interés para los diferentes sectores productivos.

7. BENEFICIOS DEL FUNCIONAMIENTO DE UNA INCUBADORA DE EMPRESAS

Las incubadoras de empresas son parte de la infraestructura social, que promueven la cultura del autoempleo, la generación de negocios propios y la creación de plazas laborales; compitiendo también por un posicionamiento en el concepto de responsabilidad social empresarial, el cual incluye aspectos del impacto ambiental, social y económico de la actividad del mercado y el reconocimiento de la sociedad. La mayoría de las incubadoras enlazan contactos con diferentes organismos y dependencias que brindan apoyo para su constitución, financiamientos económicos a los emprendedores y promuevan el desarrollo económico a través del fomento a las micro, pequeñas y medianas empresas para su viabilidad y funcionamiento. Estas alianzas son fundamentales para el desarrollo y consolidación de la incubadora ya que de ellas surge apoyo, además logran el reconocimiento y acreditación de la misma.La evaluación de la gestión de la incubadora basada en resultados es central, en particular con el desempeño y el logro de empresas, también en sus efectos e impactos en el resto de la sociedad y no sólo en la actividad económica.

El impacto de las incubadoras se ve reflejado en las condiciones de vida, en las actitudes y prácticas de los participantes/contrapartes, así como también en su entorno al que hayan contribuido con el proyecto. Si bien es complejo cuantificar el impacto social de forma precisa, por ejemplo ¿cómo medir la cultura del autoempleo? Es posible dar algunas ideas de los resultados por esperar.

La creación de empleos es un elemento a mencionar. Hacia el segundo año se espera que existan 5 planes de negocios para concretar, de cada uno de ellos se estima generar al menos de 2 a 3 empleos directos, sin contar todos los indirectos. Así se tendrían hasta 15 empleos en el segundo año. Esa tendencia continuaría los siguientes años, de 10 a 15 empleos por año. Si bien podrían parecer poca cantidad de plazas de trabajo generadas, el impacto podría ser en cascada. Es decir, por imitación de casos de éxito, más estudiantes y personas dentro de los sectores productivos se acercarían a la incubadora de empresas o alguna institución de características similares que les apoye en la realización de sus proyectos. Así, una forma de medir la cultura del autoempleo consistiría en el número de planes de negocios iniciados y por tanto concretados.

Un elemento en ocasiones no cuantificable en el corto plazo son los efectos positivos de una incubadora de empresas en la sociedad a la que está impactando. Algunas de las áreas donde es posible identificar los beneficios son:

i) La oferta de plazas de trabajo atractivas (salario, condiciones laborales) permite el arraigo de la población en edad de trabajar en la región. Los beneficios es que se puede frenar el fenómeno de la emigración y la consecuente desintegración de los vínculos familiares, además se impulsa el crecimiento del mercado (la demanda de bienes y servicios) en la región.

ii) La ampliación de la gama de oportunidades del ejercicio laboral de los egresados de las licenciatura universitarias. Así podrán tener como opción real la apertura de su empresa y generar su propia plaza de trabajo, presionando menos al mercado laboral.

iii) Un efecto cascada de las acciones de las generaciones actuales a las que vienen. Por imitación o ejemplo los jóvenes el día de mañana desearán o al menos lo verán como opción el abrir su propia empresa como sus familiares o amigos que lo han realizado en el presente. De forma implícita esto es el desarrollo de una cultura de emprendedor y empresarial.

A manera de comentario final, la incubadora de empresas en la región se constituye en aquel instrumento de apoyo al desarrollo económico de la región, además de contribuir a la mejora de las condiciones socioeconómicas de la población en la zona. De esa forma cumple la UAEM su función de vinculación con la sociedad y sectores productivos, sin olvidar a los tres niveles de gobierno, cuyo resultado es redituable para todos.

BIBLIOGRAFÍA

AYUNTAMIENTO DE TENANCINGO. (2004). *Plan de Desarrollo Municipal de Tenancingo*, Estado de México. Tenancingo Méx.

AYUNTAMIENTO DE TENANCINGO. (2004). *Plan de Desarrollo Urbano*. Tenancingo Méx.

Gobierno del Estado de México. (2006). *Indicadores sociodemográficos de la región VI, Ixtapan de la Sal 1950-2005*.

INSTITUTO NACIONAL DE ESTADÍSTICA, GEOGRAFÍA E INFORMÁTICA (2001). *Censo de Población y Vivienda,* Aguascalientes.

INSTITUTO NACIONAL DE ESTADÍSTICA, GEOGRAFÍA E INFORMÁTICA (2005). *Sistema Automatizado de Información Censal SAIC 5.0. Censos Económicos 2004*, Aguascalientes.

INSTITUTO NACIONAL DE ESTADÍSTICA, GEOGRAFÍA E INFORMÁTICA (2006). *Conteo Rápido de Población,* Aguascalientes.

MARTÍNEZ VILCHIS, J. M. (2005). *Plan Rector de Desarrollo Institucional 2005-2009*, Universidad Autónoma del Estado de México, Toluca, Méx.

RAMÍREZ HERNÁNDEZ, J. J. (2007). "Características de la demanda turística en el municipio de Tenancingo, Estado de México" en el 4º Foro de Investigación y Reflexión en torno al Turismo, Centro Universitario UAEM Texcoco, Texcoco, Méx. Octubre.

RAMÍREZ HERNÁNDEZ, J. J. (2008). "Desempeño de la floricultura mexiquense en el contexto internacional" en el XI Congreso Nacional y IV Congreso Internacional de Horticultura Ornamental, Universidad Autónoma de Morelos y Asociación Mexicana de Horticultura Ornamental A. C. Cuernavaca Mor. Marzo.

www.contactopyme.gob.mx/snie/PrincipalSNIE. Sistema Nacional de Incubación de Empresas, Secretaría de Economía. Consulta en mayo de 2008.

OS EFEITOS E REPERCUSSÕES DAS PRINCIPAIS OSCILAÇÕES CONTRACIONISTAS FINANCEIRAS INTERNACIONAIS SOBRE A CONJUNTURA ECONÓMICA BRASILEIRA

GILVAN DIAS DE LIMA
E-mail: gilvandl@gmail.com
THIAGO HENRIQUES FERNANDES
E-mail: thiagohenriques@gmail.com
Universidade Federal de Campina Grande

Resumo

A Economia Internacional exige, em seu tratamento e estudo uma sólida compreensão da realidade e dos fatos que lhe permeiam. Dentro desse contexto, esse trabalho se propõe a interpretar de modo descritivo e analítico as principais crises financeiras da economia mundial no século XX. Para isso, serão apresentados os cenários históricos, político-ideológicos e econômicos em que foram constituídos os pilares que sustentaram tais acontecimentos, além do desenvolvimento, conseqüências e repercussões para os mais diversos continentes e países, em destaque para o Brasil.

Palabras Clave: Economia Internacional e Brasileira; Crises Financeiras; Efeitos e Repercussões.
Área Temática: Economía Internacional y de la Integración.

Abstract

The international economy demands, in its treatment and study, a firm understanding about the reality and facts, which constitute it. Within this context, this study is to interpret descriptive and analytical way of the main financial crises of the world economy in the XX century. In order to do so, historical sceneries as well as economical and ideological political ones will be presented, under which they are constituted as pillars for supporting those happenings, in addition to its development, consequences and repercussions for most of countries and continents.

Key words:. Brazilian and International Economy; Financial Crisis; Effects and Repercussions.
Thematic Area: International and Integration Economics.

1. INTRODUÇÃO

Este trabalho apresenta uma visão histórica e analítica das principais oscilações da economia mundial no século XX. Iniciando com a grande depressão de 1929 e encerrando com as turbulências da Ásia e da Rússia nos anos 90.

De cada crise estudada, por sua vez, foi examinado os motivos que lhe originaram, a conjuntura em que se desenvolveu, e os efeitos e repercussões na economia mundial.

2.1 A GRANDE DEPRESSÃO DE 1929

No final da década de 1920, a prosperidade norte-americana afigurava-se em algo extremamente frágil. Os agricultores estavam produzindo mais do que era utilizado pela indústria e adquirido pelos consumidores finais. Tal situação gerava o abarrotamento das lojas, e

repercutia no setor industrial através da redução na aquisição de vários bens manufaturados. Assim, criava-se uma situação viciada em várias linhas produtivas, obrigadas a diminuir a quantidade produzida e conseqüentemente excluindo-se trabalhadores.

A situação da economia parecia não se refletir nas bolsas de valores, especialmente na Bolsa de Nova York, a principal dos Estados Unidos. Baseados na prosperidade dos anos 20, quando a concorrência aos produtos norte-americanos era reduzida, os especuladores faziam com que as ações de diversas empresas atingissem cotações totalmente irreais.

A distorção dos valores encontrou um fim súbito na chamada "quinta-feira negra", dia 24 de outubro de 1929, quando ocorreu de forma brutal a baixa das mais diversas ações, desencadeando-se uma corrida de venda de ações, sem que fossem encontrados compradores. Era o *crack* da economia mundial.

A medida adotada pelos governos das grandes economias capitalistas, para enfrentar a crise foi a aplicação de políticas monetárias restritivas e contracionistas (elevação da taxa de juros; aumento das vendas de títulos ao setor privado; ampliação dos níveis de reservas compulsórias para garantir a liquidez do sistema bancário; a fixação de limites de crédito; e a redução dos prazos de pagamento dos empréstimos, etc.). A opção desses governantes precipitou a transmissão da queda bolsística ao sistema bancário, gerando, finalmente os controles de capitais. Como conseqüência, inúmeros bancos faliram, arrastando em conjunto numerosas empresas, fato esse que aumentava o desemprego e a retração do consumo, num círculo vicioso. Tudo isso, acarretou o fim do padrão-ouro de certa forma espúrio que funcionava no entre-guerras.

Dessa forma, o que tinha começado como uma simples queda dos valores na Bolsa de Nova York foi seguido de uma corrida contra os bancos e, na seqüência, no decreto da inconversibilidade das moedas. Em 1931, como conseqüência de problemas cambiais e do abandono do padrão-ouro pelo Reino Unido, vários países da América Latina também declaram a inconversibilidade de suas moedas (entre eles estavam o Brasil, Argentina, Colômbia e Chile).

As políticas cambiais baseadas na desvalorização agressiva das paridades das moedas geraram perdas para todas as economias, ampliando o clima de protecionismo e de guerra comercial, o que serviu para deprimir ainda mais o ritmo da atividade econômica. Os Estados Unidos foram em grande parte responsável pela depressão mundial que se seguiu. O protecionismo estado-unidense – bastante diferente do praticado na América Latina, já que lá estava articulado a uma política industrial e aqui era o resultado de preocupações com o balanço de pagamento – e os excedentes comerciais periódicos, durante os anos vinte, impediram os países europeus de pagar suas dívidas por meio do comércio; e o fracasso dos EUA em regular os empréstimos privados à Europa fomentou as inversões especulativas e irresponsáveis na Bolsa de Nova York.

Iniciada a crise, originou-se de imediato uma "onda" de protecionismo em todo o mundo. Assistiu-se então a uma depressão severa, com ondas recorrentes de depreciação cambial e um intenso nacionalismo econômico. Os mecanismos empregados foram: quotas e tarifas elevadas, acordos preferenciais de comércio e, em vários países, controles cambiais e comércio de compensação bilateral. Nessas circunstâncias, a recuperação do comércio internacional arrastou-se lentamente atrás da retomada da produção.

O protecionismo teve repercussões mais depreciativas e intensas nos países pobres. Para o Brasil, por exemplo, ocorreu o acumulo de grandes estoques de café, além de uma queda no preço, provocando, nos anos seguintes, uma importante redução nas importações do país, uma vez que o Brasil não tinha como financiar a compra de produtos estrangeiros.

Nessa circunstância, não é de estranhar que o país tenha sido levado, num primeiro momento, a adiar o pagamento de suas obrigações financeiras (1931-1933) e, depois, a declarar uma moratória de fato (1937), assim como tentar poupar o máximo de divisas por meio de acordos de troca de produtos e de *clearing* compensado. Destaca-se que as rupturas de pagamentos sempre foram seguidas de renegociações com credores e nova emissão de títulos, alguns dos quais comportando deságios.

No plano monetário, é nos anos 30 que o Brasil sai da zona *sterlin*g e entra na esfera do dólar, moeda absoluta em suas contas externas. No plano mais geral da política econômica, o

governo brasileiro praticou uma espécie de *Keynesianismo*, ao garantir a demanda agregada pela via da manutenção da renda do setor cafeeiro.

Com respeito ao balanço de pagamentos, entre 1930-1939, os problemas avolumaram-se, não só pela diminuição das receitas da crise internacional, como também pela redução do fluxo de capitais. Como resultado dessa situação, foi contínua a desvalorização da moeda brasileira. Destaca-se ainda que, nesse período, as remessas de lucros das empresas estrangeiras foram reduzidas ao mínimo pelo governo brasileiro.

A desordem econômica dos anos 30 só começou a ser superada a partir da conferência de *Bretton Woods*, em 1944, onde foram criados o Fundo Monetário Internacional e o Banco Mundial, que constituíram-se nos novos pilares da ordem econômica internacional.

2.2 A CRISE MUNDIAL COM O FIM DE *BRETTON WOODS*

A maioria dos estudiosos atribuíram a crise dos anos 70 – estagflação (a coexistência da estagnação econômica com altas taxas de inflação) – ao aumento dos preços do petróleo no período de 1973 e 1979.

Entretanto, a crise também teve outro motivo de origem. A decisão unilateral adotada em 1971 pelos Estados Unidos sem decretar a suspensão da conversibilidade do dólar em ouro, tal como havia sido garantido nas leis impostas a todos os países na conferência de *Bretton Woods* e na convenção do FMI.

Pode-se ainda dizer que foi a desvinculação do dólar ao padrão ouro que causou a alta de preços do petróleo, uma vez que a depreciação monetária imediata da moeda norte-americana, que se seguiu à medida de inconversibilidade do dólar em ouro, representou uma grande redução do poder de compra dos países exportadores de petróleo bruto.

Essa situação se tornou inevitável à medida que a inflação dos Estados Unidos e o acúmulo de imensos estoques dessa moeda em poder de bancos centrais europeus e de outros países tornavam cada vez mais frágil a perspectiva de sua garantia em ouro, tendo em vista os estoques insuficientes de *Fort Knox* e a pouca disposição do governo dos EUA de converter os ativos no exterior em ouro. A conseqüência foi um imenso "calote" internacional aplicado pelo Tesouro Internacional, em contradição flagrante com os compromissos assumidos em *Bretton Woods* (garantia de conversibilidade das demais moedas em ouro, a partir da taxa convencional de 35 dólares por cada onça de ouro).

Logo, em 1971 passa-se a ter de fato um não-sistema monetário internacional, com todas as repercussões em termos de flutuações de moedas e de paridades não-correlacionados. Não se pode, contudo, esquecer os efeitos da crise energética. Como resultado, os dois choques do petróleo, causaram "ondas" depressivas por toda a economia mundial, absorvidas de acordos com a capacidade que cada país tinha de diminuir a necessidade de tal produto (seja promovendo a queda da produção industrial ou produzindo internamente parte do petróleo de que necessitava) ou de reverter em seu benefício os fluxos de divisas carreadas pelos países produtivos (entenda-se, como sendo, a capacidade que o país dispunha de ofertar bens e equipamentos necessários aos países produtores de petróleo, que em troca lhe pagavam com a sua produção).

Com relação ao Brasil, os impactos econômicos de tais mudanças do cenário internacional foram extremamente negativos. Enquanto importador de capitais, o país ficou sujeito a taxas de juros altamente voláteis; enquanto importador de petróleo, o país sofreu bastante pela elevação nos preços e a manutenção de uma dependência cada vez mais custosa.

As duas situações agravaram de maneira dramática o desequilíbrio das contas correntes do país – até então tranqüila e mesmo superavitárias, em vista do grande fluxo de capitais voluntários trazidos com o "milagre brasileiro" de 1967-1973 – e contribuiu para acelerar a inflação que tinha sido controlada no período de 1964-1967. Deve-se salientar que o Brasil já vinha enfrentando um certo processo de auto-alimentação da inflação, com a introdução do sistema de indexação da moeda e das desvalorizações graduais do câmbio. Essas deformações do mecanismo econômico passariam a caracterizar a economia brasileira de forma persistente até uma fase bem recente de nossa história.

Data dessa época, no Brasil, a prática aberta ou velada do protecionismo comercial e o aprofundamento de mecanismos de política comercial ou industrial que isolaram muitas indústrias – para não dizer setores inteiros – da competição internacional e da pressão por ganhos de competitividade.

Destaca-se que o Brasil desenvolveu uma aliança "tática" com os países exportadores de petróleo, uma concentração excessiva de compra desse produto no Oriente Médio, ao lado de uma ativa política de promoção comercial e de expansão das exportações em direção de mercados emergentes, com a ampliação da presença brasileira na África, no Oriente Médio e na Ásia.

2.3 A CRISE DA DÍVIDA EXTERNA DA AMÉRICA LATINA

A grave crise da dívida externa enfrentada pela América Latina nos anos 80, a partir da moratória mexicana de agosto de 1982, seguida logo depois pela do Brasil (em novembro do mesmo ano), foi em grande medida atribuída a políticas "irresponsáveis" de empréstimos comerciais pouco produtivos e de acúmulo excessivo de obrigações externas sem a capacidade correspondente de pagamento em descompasso, portanto, com a progressão das exportações e da obtenção de divisas.

No entanto, uma vez, um estudo do contexto econômico mundial da época mostra que a responsabilidade principal estava em parte fora do alcance e da capacidade de atuação dos países vítimas das crises de inadimplência desses anos. Pois, tal crise resultou, basicamente da decisão adotada em 1979 pela Administração do *Federal Reserva* dos EUA de aumentar brutalmente as taxas de juros, como forma de se garantir um aporte de recursos externos para compensar os formidáveis desequilíbrios comerciais e os sucessivos déficits públicos enfrentados por esse país. Esses juros saltaram de aproximados 6% a 8% ao ano para mais de 12% a 14% - correspondendo na realidade às taxas reais de inflação na zona OCDE – chegando em algumas épocas a 18% ou mesmo a 21% ao ano.

Logo, os empréstimos contraídos pelos países emergentes a taxas de juros flutuantes tornaram-se realmente impagáveis e, como tais, deram origem a moratórias sucessivas.

A conseqüência dessa variação da taxa de juro internacional foi a mais formidável reversão dos fluxos líquidos de capitais ocorridas desde a fase áurea do colonialismo financeiro em princípios do século: os países pobres convertem-se de certa forma em "exportadores de capitais" para países ricos.

As instituições multilaterais (FMI e Banco Mundial) e os países ricos tentaram durante um momento inicial, preservar as aparências de normalidade, transferindo novos recursos para o serviço da dívida, uma modalidade entre outras de socializar os prejuízos dos banqueiros privados e de evitar uma quebra generalizada do sistema bancário nesses países. Mas, para o final da década, reconhecendo a manifesta incapacidade de pagamento dos países mais endividados, os países do G-7, capitaneados pelos Estados Unidos (planos *Baker e Brady)*, chegaram ao fato inevitável da necessidade de uma mudança conceptual na forma de tratamento do problema da dívida: passou-se então a ser aceito por banqueiros e agências públicas dos países credores, a aplicação de algum tipo de desconto do valor nominal ou real dos títulos oficiais da dívida contraída nos anos de euforia financeira.

De fato, pode-se registrar várias fases no tratamento das dívidas dos países latino-americanos, cada uma delas caracterizada pelo aprofundamento gradual e sucessivo de mecanismos informais e formais de abatimento do principal e do serviço das dívidas. Na primeira fase, entre 1983 a 1985 aproximadamente, aplicou-se o tratamento convencional com a concessão de empréstimos pelas instituições financeiras multilaterais, a conclusão de acordos de ajuste e a indução do setor privado a novos empréstimos de curto prazo em condições um pouco mais severas. Com o fracasso dessa estratégia, passou-se à fase de renegociação formal das dívidas oficiais e comerciais, entre 1985 a 1989, com alguns acordos parciais firmados com a banca privada e com os membros do Clube de Paris. Numa terceira fase, entre 1989 a 1993, passa-se a um processo de renegociação com descontos.

O Brasil realizou neste período um acordo com o Clube de Paris (1992), seguido de amplo entendimento – sem aval do FMI – com a comunidade bancária (1994).

Salienta-se, que a deterioração das contas externas do Brasil não esperou a crise da dívida para se manifestar: depois de aproximadamente uma década inteira de saldos comerciais negativos, as contas externas conheceram em 1980, um déficit recorde, equivalente a 6,1% do PIB, resultado que diminuiu ligeiramente em 1981 (para 4,64%), mas que voltou a agravar-se em 1982, com 6,06% do PIB.

Para finalizar, devemos mencionar que o maior problema talvez não fosse a dívida, mas a inflação que chegou a atingir 160% em 1983 e continuava sendo alimentada por múltiplos fatores: os déficits do balanço de pagamentos, as desvalorizações da moeda, os gastos governamentais, a especulação financeira agravada pela falta de investimentos confiáveis.

Tal espiral inflacionária, inédita na história do Brasil, tendia a auto-alimentar-se no medo de todos agentes da economia do país e na corrida entre preços e salários, com nítida vantagem para os primeiros. Aos poucos o Brasil foi mergulhado em uma recessão que tinha o desemprego como principal conseqüência.

2.4 A CRISE DAS BOLSAS DE VALORES DA DÉCADA DE 80

Depois de alguns anos de relativa euforia nos mercados, com a recuperação das principais economias desenvolvidas da estagflação dos nos 70, o mundo voltou a conhecer os sobressaltos típicos das fases de turbulências financeiras do capitalismo. A queda nos títulos chegou a alcançar 22% num único dia, sinalizando o que poderia ser uma repetição da crise de 1929.

As variações extremamente significativas das moedas no decorrer dos anos 80 não corresponderam exatamente a variações nos ciclos econômicos nacionais; mas propriamente ao desenvolvimento extraordinário dos mercados financeiros, com diversos mecanismos de *hedge* e de derivativos que aumentaram o volume de dinheiro – e a fragilidade – circulando diariamente no sistema financeiro.

Ocorreram mudanças relevantes nesse sistema desde a oferta de crédito institucional dos anos 60 – dominada basicamente pelos bancos de desenvolvimento – até os derivativos dos anos 90, passando pelo mercado de *eurodólares* dos anos 60. As crises e as contrações desses mercados financeiros estão associadas às situações de inadimplência temporária de grandes devedores: países emergentes da América Latina e da Ásia, economias socialistas como Polônia, Alemanha e Hungria.

As aplicações em bolsa se desenvolveram bastante nesse período, conduzindo a um primeiro exemplo de "exuberância irracional" que resultaria na crise de outubro de 1987.

Os efeitos para o Brasil foram limitados e variados, uma vez que o país vivia uma situação de descontrole inflacionário de pouco acesso (ou nenhum) ao crédito externo e de ruptura de pagamentos: a chamada "moratória soberana" tinha sido decretada em fevereiro desse ano, seguido-se ao controle e centralização cambiais introduzidos em 1985, no auge da crise das contas externas. As bolsas também sofreram a queda dos títulos e o mercado cambial permaneceu agitado, com pressão para a modificação das paridades, a despeito do Plano Cruzado ter tentado romper com a prática inaugurada em 1967 de minidesvalorizações.

Uma outra conseqüência, de efeitos mais importantes para a política externa brasileira, teria, contudo um impacto mais relevante em termos de sua inserção econômica internacional no resto da década de 1990. O Brasil, cuja política externa esteve basicamente voltada no final dos anos 80 para as relações com os demais países latino-americanos, passou a desempenhar um papel de primeira ordem nos processos simultâneos de formação de espaços econômicos integrados no continente – integração Brasil-Argentina, Mercosul, Iniciativa Amazônica –, de busca de reinserção da região na economia mundial e de um novo papel na política internacional.

2.5 A CRISE DO MEXICO

Integrante tardio da escola liberal e da interdependência, depois de décadas de nacionalismo econômico e de industrialização substitutiva, o México vinha crescendo a taxas razoáveis no começo dos anos 1990, à base de grandes investimentos externos e de uma grande abertura comercial.

Em 1992, no Governo Salinas, o México postulava sua candidatura a OCDE – finalmente em maio de 1994 – para a qual lhe foi precisamente solicitado grande esforço de abertura econômica externa e significativa liberalização financeira.

Acumulando grandes déficits comerciais desde o final dos anos 80, o México apresentou resultados crescentemente negativos nas transações correntes, com mais de três anos de déficits equivalentes a 7% do PIB. O financiamento vinha sendo feito pela emissão de títulos do governo – os "*Tesobonos*", garantidos em dólar – e pela abertura dos mercados financeiros mexicanos, até o momento em que a rigidez cambial começou a lançar dúvidas sobre a capacidade do governo em honrar seus compromissos externos. Uma tentativa de desvalorização controlada, no final de 1994, sofreu o efeito da fuga maciça de capitais – em primeiro lugar dos próprios mexicanos – e do ataque especulativo contra o peso, o que reduziu consideravelmente as reservas de divisas do México.

No começo de 1995, consciente de que os efeitos sobre outros países da região e outros continentes, o governo dos Estados Unidos montou uma gigantesca operação de salvamento, envolvendo um pacote de cerca de 50 bilhões de dólares (dinheiro do FMI, dos bancos multilaterais e outras agências financeiras internacionais e dos próprios países-membros do G-7). A crise foi superada, contudo ela gerou uma nova fase no sistema financeiro internacional: a dos capitais voláteis e da necessidade de melhoria nos mecanismos preventivos e de transparência, monitorados pelas instituições financeiras internacionais.

Os efeitos foram importantes para o Brasil. O primeiro governo de Fernando Henrique Cardoso, eleito em outubro de 1994 fundamentalmente em virtude do sucesso do Plano Real, inicia sua gestão com um desafio externo de grande amplitude: a manutenção da estabilidade cambial, o que foi obtido mediante uma política de deliberada valorização cambial (o que permitiu a manutenção do baixo nível de inflação, mas que depois trouxe graves problemas para o país, uma vez que acarretou déficits constantes na balança comercial, piorando o resultado final do balanço de pagamentos), e através de uma elevada taxa de juros, que tornava atrativo para investidores nacionais e internacionais a aplicação de seus recursos no país (em contra partida, desestimulava a atividade produtiva brasileira).

Em suma, a crise financeira mexicana não custou grandes problemas a economia nacional, mas direcionou a aplicação de medidas econômicas que futuramente viriam a prejudicar o país.

2.6 A CRISE ASIÁTICA

A crise asiática de 1997-1998, que disseminou de forma diferenciada em diversos países da região, não apresentou as mesmas características da crise mexicana. Ela se deu no bojo de um processo de crescimento não especialmente inflacionário e de manutenção de políticas fiscais e macroeconômicas relativamente sólidas e responsáveis. Tratou-se, portanto, de uma verdadeira *"caixa de Pandora"*, pois os países da região eram então louvados como exemplos de boa gestão pública e excelente performance na manutenção de altas taxas de poupança, de investimentos e de crescimento, com grande inserção na economia externa.

Seus efeitos sobre outras regiões foram diferenciados segundo a percepção de riscos implícitos do ponto de vista da solidez das economias, imediatos no caso da América Latina e mais delongados em outros países, provocando ainda assim efeitos devastadores sobre alguns grandes "atores" emergentes como Rússia e Brasil.

A crise inicia-se com a desvalorização sob pressão da moeda da Tailândia em julho de 1997, o que acarretou uma corrida contra as demais moedas da região, provocando um efeito dominó sobre todas elas, com impactos especulativos na Malásia e na Coréia, respectivamente em setembro e em dezembro desse ano, e na Indonésia, em fevereiro de 1998, provocando a queda

da ditadura *Suharto*. Etapa importante no processo de propagação da crise para outras regiões, foi o ataque contra o dólar de Hong-Kong, em outubro de 1997, que quase ameaçou de colapso o sistema de *"currency board"* mantido pela ex-colônia britânica, reincorporada definitivamente à China no primeiro semestre do ano. A crise asiática, portanto, assumiu contorno político com as turbulências na Indonésia, cujos efeitos foram devastadores, uma vez que gerou a inadimplência de diversos devedores (empresas, bancos e mesmo o próprio governo) precipitou a falência de casas bancárias no Japão, em Hong-Kong e na Coréia, países com instituições, mas expostas à crise de liquidez e à desvalorização cambial (em alguns casos superior a 80%) das economias asiáticas mais afetadas pela crise continuada (Indonésia e Malásia, sobretudo).

Destaca-se que a continuada depressão japonesa dificultou de certo modo a retomada do processo de desenvolvimento. Por outro lado, a crise política em vigor em determinados países de governo autoritário – como a Indonésia e Malásia – atuou no sentido de dificultar a correção dos desajustes macroeconômicos e de colocar esses países no caminho da recuperação. Outros países, por sua vez, enfrentam de forma bastante consistente os desafios colocados pelos capitais voláteis, como Cingapura e Taiwan, enquanto Hong-Kong dispunha de uma reserva de segurança adicional, sob a forma do volume de reservas da China.

A gestão da crise asiática pelo FMI foi bastante criticada dentro e fora da região, não apenas no que se refere ao sentido geral das políticas recomendadas – como a contenção fiscal, por exemplo – como também sua dosagem e aplicação inadequada. Deve-se reconhecer, no entanto, que a gravidade da situação da Indonésia, por exemplo, explica-se também por um potencial de conflito propriamente interno. Os pacotes de ajustes montados numa primeira fase pelo FMI seguiram essencialmente o modelo clássico, provocando um certo agravamento das condições macroeconômicas, ao propor um forte ajuste em todos eles, mas de forma geral foi enfatizada a necessidade de correção dos desequilíbrios externos, particularmente graves nos países mais afetados.

Os efeitos da primeira fase da crise asiática se fizeram sentir de imediato na América Latina, com uma repercussão direta nos sistemas financeiros (transferências de ativos e de investimentos de curto prazo para realização de ganhos pelos investidores externos), em especial no Brasil, que começou a sentir as primeiras conseqüências de um retraimento dos capitais voluntários que iria afunilar-se no quadro da crise russa, a partir de agosto de 1998.

A segunda fase da crise asiática, em princípios do mesmo ano, com o aprofundamento da recessão em diversos países da região, continuou a produzir efeitos também para a América Latina, sobretudo no terreno comercial (retração de mercados, alteração de preços relativos, diferenciais de competitividade, etc.). Se o Brasil foi mais afetado na primeira fase, outros países da região (como México, Peru e Chile) tiveram seu comércio com a Ásia bastante diminuído no decorrer da segunda fase. De forma, a adoção de medidas para conter o contágio asiático, no Brasil e em outros países, apresentou um efeito direto sobre a taxa de crescimento econômico, chegando mesmo a provocar recessão em alguns países.

Para evitar maiores riscos, desde o ataque contra Hong-Kong em outubro de 1997, o governo brasileiro aumentou extraordinariamente os juros e adotou uma série de medidas com impacto fiscal. Elas se revelaram momentaneamente eficazes, na medida em que se evitou ataques especulativos contra o Real (desestimulados também por um nível elevado de reservas cambiais). Entretanto, os mercados financeiros julgaram que o caráter dessas medidas não foi forte o suficiente, razão pela qual a política de acumulação de reservas pelas autoridades monetárias foi perseguida com especial denodo desde então. Em todo caso, os juros no Brasil continuaram a ser diminuídos na seqüência da crise asiática, mas tiveram de ser novamente aumentados quando da moratória da Rússia e em função dos desdobramentos mais graves da crise financeira.

2.7 A CRISE DA RÚSSIA

A Rússia mantinha grande parte de sua dívida interna convertida em dólar, atraindo capitais com base em altas taxas de juros e de rendimentos baseados em títulos do governo. O processo

de liberalização financeira conduzida no quadro de uma privatização por si já parecia algo caótico. E foi, pois, levou a Rússia ao acúmulo de obrigações externas sem garantias suficientes de resgate, a não ser a continuidade de um programa de ajuste conduzido por autoridades reformistas, com o apoio do FMI e do G-7. Ao instalar-se a crise política, com a demissão do governo *Kirienko*, seguiu-se uma quebra da confiança na capacidade das novas autoridades em manter a estabilidade da moeda. De fato, após tentativas custosas de se manter a paridade do rublo, submetido a fortes ataques especulativos desde junho de 1998, o governo deixou a moeda flutuar, o que representou níveis espetaculares de desvalorização.

Em 11 de agosto, quase sem reservas em moedas fortes e com as taxas de juros em níveis estratosféricos, o governo russo decretou a moratória sobre a dívida das empresas privadas – cerca de 40 bilhões de dólares – e sobre a dívida interna – aproximadamente 32 bilhões de dólares –, o que representou o equivalente confisco sobre os investimentos estrangeiros, em especial de bancos europeus.

Os efeitos da moratória russa foram devastadores, com a retirada em pânico dos capitais remanescentes, não apenas na Rússia – de toda forma congelados durante 3 meses – mas, sobretudo nos mercados "emergentes" da América e Ásia. Os resultados seqüenciais e a verdadeira interdependência do sistema financeiro internacional ficaram evidentes quando os principais bancos credores da Rússia foram forçados a realizarem fortes provisões de reservas para compensar perdas consideráveis, a essa altura inevitável.

Outro fato de extrema gravidade a evidenciar os efeitos sistêmicos da crise russa foi a "quebra" do fundo de investimento norte americano *Long Term Capital Management*, dirigido por eminentes prêmios *Nobel* de economia, cuja sobrevivência apenas deveu-se à decisiva ajuda financeira do *Federal Reserve*, evitando-se uma quebradeira em cascata.

Em função da gravidade do caso russo, o sistema internacional conviveu durante semanas inteiras, com o risco de um colapso de outros países, o que levou a reuniões de emergência das autoridades financeiras do G-7. A situação da Ásia continuava a ser perigosa e, dentro dela, a irritante inação do governo japonês em combater sua própria recessão interna e a irresponsável atitude da China em relação à paridade externa de sua moeda. Na América Latina, ademais dos problemas de perda de competitividade comercial, continuaram as precauções com a "saúde" financeira da região, sobretudo depois que a Colômbia realizou uma mini-desvalorização de surpresa (felizmente bem sucedida). O México teve de administrar, contra sua própria vontade, um processo moderado de desvalorização do peso, situação igualmente enfrentada pelo Chile. Mas, foi no Brasil que o impacto foi mais dramático, as reservas do país caíram de 72 bilhões de dólares no período anterior a crise para 40 bilhões em meados de novembro de 1998, fazendo com que o país implementasse um novo pacote de ajuste fiscal, paralelamente à implementação de um novo acordo com o FMI e as autoridades financeiras do G-7.

O pacote de ajuda, anunciado finalmente em 13 de novembro de 1998, foi elaborado em plena assembléia anual das instituições de *Bretton Woods,* no mês de outubro, e complementado por contatos diretos das autoridades financeiras do Brasil com seus parceiros das principais economias desenvolvidas e com funcionários do FMI. O Brasil se habilitou a receber cerca de 20 bilhões de dólares no espaço de três meses, a partir de novembro de 1998 e até 37 bilhões no prazo de um ano, dependendo do nível de suas reservas internacionais e do grau de implantação do pacote de ajuste fiscal.

Com respeito ao pacote de ajuste fiscal, o governo brasileiro se comprometeu a manter uma firme disciplina monetária, logrando preservar intacto os fundamentos de sua política cambial, baseada num regime flexível de desvalorizações progressivas internamente a uma banda de flutuação administrada estreitamente pelo Banco Central.

Percebe-se que esse ajuste inaugurou uma nova modalidade de intervenção das instituições financeiras internacionais, já que se tratava de disponibilizar recursos para reforçar as reservas internacionais do Brasil em caráter preventivo, ou seja, antes que se manifestasse uma inadimplência de fato, seguida de eventual decretação de moratória, em um caso imediatamente posterior ao da Rússia.

3 CONCLUSÃO

Ao longo do século XX, no contexto vigente da economia internacional, observou-se o desenrolar de uma série de crises de dimensões e resultados variados em termos de continentes e nações. Contudo, pode-se atribuir a tomada de decisões individuais e unilaterais por parte de alguns líderes de países dominantes e a atuação especulativa de grandes investidores mundiais como fatores responsáveis pela formação de cenários financeiros críticos.

Ao mesmo tempo, as observações feitas ao longo desse trabalho sobre a conjuntura do mercado financeiro internacional nos condicionam a elaboração de medidas e indicações que visam a ampliação e resistência das economias nacionais frente a futuras crises financeiras.

1) A primeira linha de prevenção das crises financeiras deve ser dada pela própria administração de políticas macroeconômicas nacionais, mas particularmente na gestão responsável das políticas monetárias e cambiais, como forma de estabelecer fundamentos sólidos nessas políticas, garantindo a credibilidade e a confiança dos mercados financeiros;

2) Sistemas confiáveis de administração e vigilância bancárias, assim como mecanismos de proteção – multilaterais e nacionais – com suficientes reservas monetárias e cambias para enfrentar diferentes tipos de ataque, também são essenciais para prevenir crises financeiras;

3) No plano internacional, a existência de "agentes credores de última instância" dotado de suficiente capital e a disponibilidade de linhas contigenciais de apoio financeiro – eventualmente preventivas – podem ser essenciais para prevenir e corrigir os aspectos inevitavelmente nocivos da globalização financeira, que são a inversão rápida e brutal dos movimentos de capitais e alavancagem por vezes arriscada dos especuladores cambiais;

4) A utilização de controles temporários de capitais, sobretudo na entrada, assim como o estabelecimento de mecanismos consensuais de renegociação de dívidas de curto prazo pode ser recursos legítimos aos quais os países enfrentando momentaneamente desequilíbrios de suas contas correntes devem poder fazer apelo para administrar eventuais crises financeiras;

5) Deve-se desenvolver o diálogo entre os vários países, em especial as economias "emergentes", e as instituições financeiras internacionais na cooperação internacional (e regional) em prol do estabelecimento de sistemas mais avançados de vigilância financeira e de mecanismos prudenciais na área bancária. O intercâmbio de informações entre nações e essas instituições, assim como, a transparência na divulgação dos principais indicadores macroeconômicos dos países são fundamentais para a formação, consubstanciação e manutenção de um cenário externo diáfano, estável e seguro.

REFERÊNCIAS

ADAS, M. A. América: noções sobre o subdesenvolvimento e o desenvolvimento mundial. V.3, 2° ed.; São Paulo: Ed. Moderna, 1992.

ARRIGHI, G. O longo século XX. São Paulo: UNESP, 1996.

BACHA, E. Os mitos de uma década. São Paulo: Ed. Paz e Terra, 1978.

BATISTA JR., P. N. Da crise internacional à moratória brasileira. Rio de Janeiro: Ed. Paz e Terra, 1998.

BAUMANN, R. Brasil nos anos noventa: uma economia em transição. Revista da Cepal – nº 73. Oficina da CEPAL em Brasília: Mimeo, 2000.

DUPAS, G. Economia global e exclusão social. São Paulo: Ed. Nacional, 1999.

FRITSCH, W.; MODIANO, E. M. A restrição externa ao crescimento econômico: uma perspectiva de longo prazo. Pesquisa e planejamento econômico, 18 (2), agosto de 1988.

FURTADO, C. O capitalismo global. São PAULO: Ed. Paz e Terra, 1998.

GOMES, C. T. F. Captação de recursos no mercado internacional de capitais. Rio de Janeiro: IBMEC, 1982.

IANNI, O. Estado e planejamento econômico no Brasil: 1930-1970. Rio de Janeiro: Ed. Civilização Brasileira, 1971.

INTERNACIONALIZAÇÃO DE EMPRESAS E DESENVOLVIMENTO SUSTENTÁVEL: *O CASO DA NATURA*

THIAGO HENRIQUES FERNANDES
E-mail: thiagohenriques@gmail.com
GILVAN DIAS DE LIMA
E-mail: gilvandl@gmail.com
Universidade Federal de Campina Grande

Resumo

Desde a intensificação do processo de globalização, as empresas direcionam sua internacionalização a partir de estratégias. No caso da empresa brasileira de cosméticos, Natura, a estratégia de internacionalização adotada utilizou-se do modelo de gestão ancorado no desenvolvimento sustentável e no marketing ambiental. A empresa é referencia no setor de cosméticos e vem mudando paradigmas sendo capaz de aliar uma gestão sócio-ambiental inovadora com uma linha de negócios lucrativa e agressiva, num ambiente onde os consumidores, por meio, de *buycott,* vem mudando os padrões de consumo a favor do meio ambiente.

Palavras-chaves: internacionalização, desenvolvimento sustentável, *buycott*, marketing ambiental.
Área Temática: Economía Internacional y de la Integración.

Abstract

Since the intensification of the globalization process, the firms point their internationalization by strategies. In the case of the cosmetics Brazilian company Natura, the internationalization strategy chosen has utilized on model of management structured in the sustainable development and environmental marketing. The firm is reference in the fild of cosmetics and has been changing paradigms being capable of unite an innovative social-environmental management with a profitable and aggressive business line, in one environment where the consumers by boycott are changing the consumption standards in favor of the environment.

Keywords: internationalization, sustainable development, buycott, environmental marketing
Thematic Area: International and Integration Economics.

1 INTRODUÇÃO

O processo de globalização vem mudando a forma que as empresas se organizam e interagem com seus clientes, não só a nível local, mas também a nível internacional. Muitas empresas perceberam que é bem mais lucrativo adotarem uma política mais dinâmica relacionada aos seus negócios internacionais, como a implementação de lojas ou unidades produtivas fora do país sede. Outro fenômeno que vem ganhando importância desde o ano de 1970 são os temas relacionados a preservação do meio ambiente e ao desenvolvimento sustentável, dado a crescente redução do estoque de recursos naturais disponíveis que pode comprometer a produção e consumo das gerações futuras. Muitas consequências já são vistas, como catástrofes naturais de grande escala que só tendem a aumentar devido a degradação e esgotamento ambiental.

Dentro deste contexto a os governos tem criado mecanismos de pressão, para incentivar as empresas a inovarem diminuindo os impactos de sua atividade econômica sobre o meio

ambiente. Entretanto a responsabilidade da preservação da biodiversidade não cabe apenas aos organismos políticos, a sociedade civil organizada e as empresas devem ter igual parcela de comprometimento com as causas ambientais.

Nesta perspectiva vem surgindo empresas que utilizam-se das eco-vantagens como diferencial competitivo, a exemplo disto será abordado o caso da empresa brasileira de cosméticos, Natura.

Este trabalho tem por objetivo apresentar as vantagens advindas da utilização de um modelo de gestão ambiental para sucesso da empresa, não só no mercado doméstico, mas também no mercado internacional, possibilitando a empresa a obtenção de um diferencial competitivo e desenvolvimento da sociedade na qual ela esta inserida. Percebe-se assim que a idéia de lucratividade e eficiência econômica, não esta necessariamente distanciada da idéia de responsabilidade ambiental, uma vez que considera a premissa da utilização ótima dos recursos escassos aliado ao desenvolvimento de uma parcela significativa de consumidores comprometidos com a sustentabilidade e preservação do planeta.

O estudo inicia-se com a abordagem de internacionalização de empresas e as estratégias utilizadas para o sucesso de expansão. Seguido de um resgate histórico da discussão de sustentabilidade ambiental, e posteriormente a sua inclusão no modelo de gestão das empresas. Concluindo-se com a apresentação de um estudo de caso, para exemplificar as idéias anteriormente defendidas.

2 METODOLOGIA

Esse estudo privilegiou duas visões metodológicas: analítico-descritiva e analítico-histórico.

Sendo que a primeira consubstanciou a segunda. A partir de dados estatísticos cedidos gentilmente pela Natura pode-se avaliar a importância do modelo de gestão sustentável desta empresa de cosméticos brasileira.

Para tanto, procedeu-se com a investigação dessa forma apresentada:

- Leitura informativa através de: artigos, jornais, revistas, entrevistas, etc.
- Leitura bibliográfica: Livros, dissertações e teses.

3 INTERNACIONALIZAÇÃO DE EMPRESAS

A internacionalização de empresas é uma decisão de natureza estratégica, que envolve a identificação dos produtos que serão alocados nos mercados internacionais, os mercados alvos e os modos de presença mais adequados.

A internacionalização de uma empresa leva a mudanças significativas na estrutura da organização, principalmente no processo dinâmico de aprendizagem. Ela envolve atividades como: comércio exterior, fluxos de capital, transferência de tecnologia, fluxo de informações e dados, alianças, fusões, aquisições, Investimento Direto Estrangeiro - IDE, o*utsourcing*[1] e outros.

Um dos principais fatores que motivam uma empresa, ou corporação, à buscarem atividades além de suas fronteiras são a saturação do mercado doméstico e a forte competição entre as empresas.

Nessa perspectiva as empresas que obtêm sucesso a nível local, isto é, dominam o mercado doméstico, tendem a buscar novas oportunidades em mercados internacionais.

Na visão de Almeida (2007) as razões pelas quais levam as empresas à investirem em países estrangeiros são diversas e variam de acordo com a realidade de cada empresa; no entanto, o acelerado alargamento dessa prática gera uma pressão sobre as empresas que ainda não se internacionalizaram (exceto em caso de exportações): as vantagens acumuladas pelas empresas

[1] Processo pelo qual um setor da empresa funciona em outro país.

multinacionais ampliam barreiras competitivas para as empresas que restam domésticas e aumentam as pressões competitivas das empresas multinacionais sobre as nacionais, mesmo no mercado doméstico destas últimas.

Fleury (2007), em uma abordagem da internacionalização como estratégia competitiva, considera que a disposição de ingressar em operações internacionais seria justificada pelo jogo competitivo em mercados internacionais.

É a partir desta perspectiva que estratégia, estrutura e comportamento devem ser considerados e avaliados. "Por exemplo, há movimentos que, em princípio, não são justificáveis economicamente, mas que se impõem em função de eventuais movimentos dos demais competidores" Fleury (2007, p. 09).

Também se pode interpretar a internacionalização de empresas como uma forma de diminuir a dependência da empresa por mercados específicos, expandindo seus mercados consumidores e conseqüentemente diminuindo os riscos do ponto de vista comercial.

Kraus (KRAUS apud KIM, 2007) considera que, "internacionaliza-se tem sido uma das estratégias que muitas empresas em torno do planeta têm buscado para fazer frente aos concorrentes internacionais e melhorar a competitividade em seus mercados domésticos".

Muitas empresas se internacionalizam, já que entendem a importância de novos fluxos de negócios para o fortalecimento estratégico de suas organizações, nessa linha de raciocínio, Luis (2006), afirma que além do objetivo de interagir em novos mercados, as empresas que se internacionalizam buscam acesso à novas tecnologias, facilidades e isenções fiscais, redução de custo de mão-de-obra, fortalecimento de sua marca e diversos outros fatores que são relevantes quanto à uma iniciativa desta natureza.

Muitas empresas a nível local e global já perceberam, que na maioria das vezes, é mais simples, efetivo e eficiente iniciar a produção em outro país do que contornar os obstáculos logísticos e legais envolvidos em uma operação de venda para o exterior.

Para Yip (YIP apud Kim, 2007) a maturidade dos mercados domésticos está impulsionando as empresas a procurarem uma expansão internacional. As empresas também buscam a globalização quando integram sua estratégia em diversos países.

3.2 Teorias Comportamentais de Internacionalização de Empresas

Segundo Kim (2007, p. 18), as teorias de internacionalização, de abordagem a partir de uma perspectiva comportamental, explicam o fenômeno de internacionalização das organizações baseado no comportamento, atitude e percepção dos gestores que lideram as empresas.

Isto será aqui trabalhado explorando o Modelo de Uppsala, ou modelo dinâmico de aprendizagem (1977) desenvolvido pelos professores Johanson e Wiedersheim-Paul.

3.2.1 O modelo de UPPSALA ou Modelo Dinâmico de Aprendizagem

Continua o raciocínio de Alem e Cavalcanti (2005) o modelo de Uppsala considera o processo de internacionalização de empresas gradual, com um crescente comprometimento da firma com o mercado externo com base no aprendizado adquirido por meio da experiência e do conhecimento da prática da internacionalização. A empresa passa da exploração de mercados próximos, isto é, passa de uma estratégia regional, e aumenta sua presença em mercados mais longínquos de forma gradual, por etapas e por estágios, de uma maneira evolutiva, formando um *network*[2] que proporcionara para empresa um maior conhecimento sobre os mercados alvos e implicara um maior esforço a cada nova etapa vencida.

O ponto central do modelo seria a distância física, que foi definida como a soma de fatores que interferem no fluxo de informações entre os mercados, como as diferenças que se apresentam entre línguas, cultura gerencial, práticas de negócio, cultura, desenvolvimento industrial, grau de maturidade econômica do país entre outros.

[2] Rede de contatos de trabalho: clientes, fornecedores, etc.

3.3 Internacionalização de Empresas dos *"Late Movers"*

Segundo o relatório da ONU, o *World Investment Report – WIR* (2007), apesar das transnacionais ainda serem dominadas por firmas de países em desenvolvimento, na atualidade este cenário tem mudado.

Segundo o relatório do *Boston Consulting Group - BCG* (2008) existem seis modelos pelos os quais as empresas dos países de rápido crescimento utilizam no seu processo de internacionalização: fazendo suas marcas globais, tornando sua inovação na área de engenharia global, assumir liderança em sua categoria global, monetarizar os recursos naturais implantar novos modelos de negócio para multiplicar os mercados e adquirir recursos naturais.

Segundo Fleury & Fleury (2007) apesar dos "entrantes tardios", ou "*late movers*", serem um assunto recente no que se refere ao estudo de internacionalização de empresas, o mesmo é bastante relevante, já existem estudos como o do *Boston Consultingt Group* (2007), apresentado anteriormente, e o estudo efetuado pelo *Bank of America* que mostram a crescente aquisição de empresas americanas por investidores dos países emergentes.

De acordo com os autores supracitados apesar de existirem muitos estudos sobre os entrantes tardios ainda não existe uma abordagem teórica que trate especificamente estas empresas, isto acontece já que as teorias que tratam do assunto foram criadas em um momento bem diferente do atual, quando o mercado de multinacionais era inteiramente dominado por empresas dos países desenvolvidos.

Todavia existem algumas abordagens que podem ser utilizadas como: a da internacionalização como processo de decisão racional econômico, internacionalização como processo de aprendizagem, internacionalização com estratégia competitiva, internacionalização com abordagem de competências estratégicas, internacionalização centradas nas subsidiárias e por último, internacionalização como abordagem de gestão de operações internacionais.

Segundo Almeida e Cretoiu (2008) existem algumas escolhas que influenciam o tipo de negócio escolhido para as empresas dos entrantes tardios como: adaptação da governança; autonomia das operações internacionais frente à matriz da empresa; os processos de gestão e de padronização a nível administrativo e produtivo; os valores comportamentais e culturais, tanto da empresa quanto do país do seu mercado alvo; e por último, a relação dos recursos humanos da empresa com o seu processo de internacionalização desde o momento do planejamento do mesmo até sua execução.

Ainda na perspectiva de Almeida e Cretoiu (2008), considera-se que as empresas dos *late movers* escolhem três modelos em geral para seu processo de internacionalização.

1. **Modelo de organização por área internacional:** A empresa cria uma área especifica para administrar todos os aspectos diretos e indiretos relacionados às suas estratégias e negócios internacionais.
2. **Modelo geográfico:** A empresa se organiza pela geografia, levando em consideração aspectos relacionados à localização.
3. **Modelo por linhas de negócio:** A empresa tem unidades de negócios, sejam elas, *joint-ventures*[3], subsidiárias, ou representação comercial em outras regiões ou países

Apesar das empresas dos "*late movers*" eram especialmente exportadoras, este cenário tem mudado. Todavia elas estão enfrentando um ambiente muito mais competitivo do que as primeiras transnacionais dos países desenvolvidos enfrentaram, mas a cada ano que passa aumenta a presença de transnacionais de países em desenvolvimento nas maiores corporações do mundo, o que revela uma tendência de uma mudança lenta, porém gradual do cenário econômico mundial.

[3] *Joint venture* ou empreendimento conjunto é uma associação de empresas, que pode ser definitiva ou não,com fins lucrativos, para explorar determinado(s) negócio(s), sem que nenhuma delas perca sua personalidade jurídica.

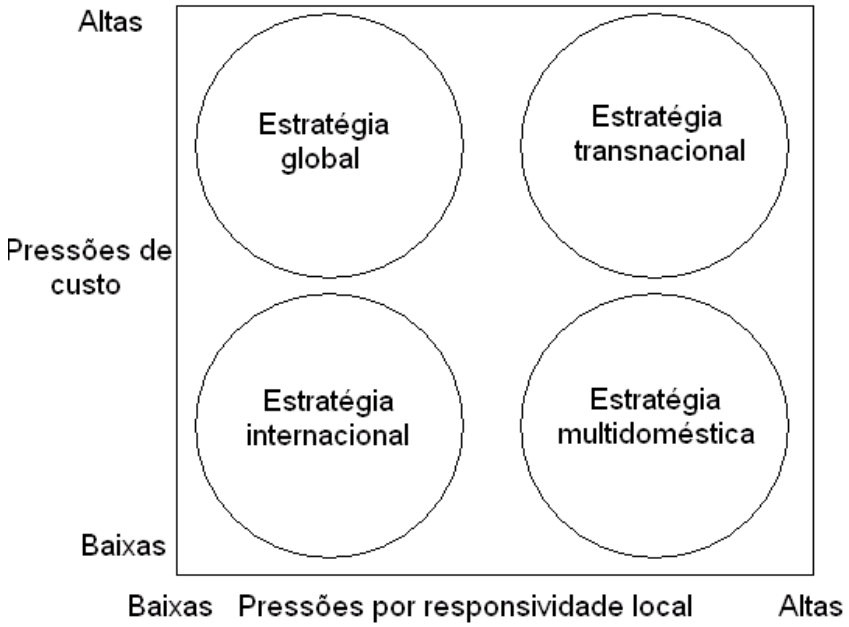
3.4 Estratégias de Internacionalização e Modos de Entreda

Segundo Bennett (2007) existem várias abordagens e modelos que tentam explicar ou orientar empresas e especialistas envolvidos no processo de internacionalização. Para ele uma das mais importantes é a de Hill (2004), que utiliza duas variáveis importantes, isto é, "pressões de custo" e "pressões por responsividade local[4]". Ambas variam para baixo e para cima de acordo com a estratégia escolhida.

Segundo o modelo existem quatro estratégias:

1. **Estratégia Internacional**: Nesta estratégia a pressão que incidem sobre os custos são baixas e as pressões por responsividade local também são baixas. Neste tipo de estratégia não é comum que a empresa crie uma divisão internacional, isto é, que tenha outra unidade fora do país sede, permanecendo etnocêntricas em sua forma de internacionalização.
2. **Estratégia Global**: As pressões relativas aos custos são altas, entretanto, as pressões que dizem respeito à responsividade local ainda podem ser consideradas baixas. Nessa estratégia, as firmas, com freqüência são motivadas a procurar por economias de escala ao criar subsidiárias no exterior, apesar de permanecer majoritariamente etnocêntricas em sua abordagem, tendo a maior parte da produção realizada no país sede.
3. **Estratégia Multidoméstica**: As pressões que incidem sobre os custos são baixas, porém no que diz respeito às pressões por responsividade local são altas. Neste tipo de estratégia as empresas têm maior interesse de tentar uma maior aproximação com o consumidor e mercado local. Adotam uma abordagem policêntrica em suas operações internacionais.
4. **Estratégia Transnacional**: Neste tipo de estratégia tanto a pressão por custos quanto a pressão por responsividade local são altas. Nesta estratégia as empresas, freqüentemente, criam uma estrutura global bastante diversificada procurando tirar vantagem das diferenças entre os países, e tendem a ter uma abordagem geocêntrica

Figura 1: Quatro Estratégias Internacionais básicas



Fonte: (HILL apud FLEURY, 2007,p. 39)

Para Ghemawat (2007) as empresas podem escolher três estratégias para utilizarem em sua estratégia de expansão global. São elas: adaptação para ajustar as diferenças, agregação para ultrapassar as diferenças e arbitragem para explorar as diferenças.

[4] Por responsividade local, entende-se como a influência das questões locais nos aspectos administrativos e financeiros da empresa.

Na visão dele, algum tipo de adaptação é essencial para virtualmente qualquer empresa que deseje se tornar global. Nesta perspectiva é importante analisar as diferenças que dividem os países, ao invés de ignorá-las baseado na falsa evidência que as mesmas não são importantes. O caso da *Wal-mart* mostra um puro caso de sub-adaptação na estratégia de internacionalização, quanto mais longe a empresa fica de sua sede, menor é seu grau de eficiência. Se comparando com o *Carrefour*, o *Wal-mart* ainda tem muito a apreender para obter sucesso na sua estratégia de internacionalização. Segundo o autor para não cair no erro estratégico do *Wal-mart* as empresas devem levar duas estratégias em consideração.

1. Analisar as diferenças que continuam a dividir os países ao invés de ignorá-las.
2. Pensar sobre a completa ordem de ferramentas para se adaptar as diferenças.

Além dessas duas estratégias as empresas também devem se preocupar em outros pontos relevantes como: ter um departamento de Recursos Humanos – RH forte e pronto para os desafios da adaptabilidade, investir na educação formal de seus membros, formação e participação de equipes internacionais de negócios e projetos, utilização de diversos lugares para encontros de grupos e projetos, experiências de imersão em culturais estrangeiras, expatriar atribuições para gerentes de alto potencial, cultivar a diversidade cultural e geográfica no topo da empresa, dispersão de unidades centrais de negócio ou centros de excelência, definir e cultivar o conjunto de valores através da corporação e a abertura através das fronteiras organizacionais.

Com relação à agregação, Ghemawat (2007) considera principalmente o uso de grupos de planos estratégicos para criar maiores e melhores economias de escala do que a adaptação país por país pode fornecer.

Agregação significa inventar e programar mecanismos além das fronteiras que operam a níveis de países individualmente, ou em todo o globo. É geralmente engajada em níveis intermediários dentro de uma empresa, ao invés de apenas nos grandes executivos da empresa na matriz e nas subsidiárias.

Para o autor pensar em agregação, demanda forte atenção tanto nas médias como nas altas posições da organização. O objetivo da estratégia é explorar as similaridades entre os países, mais agressivamente do que a estratégia tradicional de adaptação faz, mas menos agressivamente do que a completa padronização faria. A chave subliminar é a idéia, de diferenças em diferenças. Existe muito a ser ganho pela agrupação, no sentido que, diferenças dentro do grupo sejam minimamente comparados com diferenças entre os grupos.

Existem muitos tipos de estratégias de agregação, todavia a principal estratégia de agregação é por **região geográfica**. A regionalização é particularmente proeminente em um contexto de ultrapassagem de barreiras, a *Toyota* é um exemplo de execução bem feita de estratégia de agregação via regionalização.

É importante notar as diferenças sócio-culturais-econômicas que existem entre as diferentes regiões do globo e dos países para se traçar uma estratégia de agregação por regionalização de forma eficiente. Esta estratégia é mais importante do que nunca devido ao processo de globalização e o aumento das trocas comerciais entre países vizinhos.

A última estratégia, segundo Ghemawat (2007), é a de arbitragem. Para o autor a globalização significa produzir onde é mais eficiente o gasto efetivo, prospectar capital onde é mais barato e vender onde é mais lucrativo. A última das três estratégias Adaptação, Agregação e Arbitragem – AAA lida com as distâncias e ultrapassagem de fronteiras com sucesso. A arbitragem é uma forma de explorar diferenças. Ela implica em procurar "economias absolutas", ao invés de economias de escala ganhas através de padronização. Esta estratégia trata diferenças através dos países como oportunidades e não como barreira.

A estratégia de arbitragem leva em consideração quatro aspectos em potencial: as bases culturais, administrativas, geográficas e econômicas. Para ilustrar a variedade das estratégias de arbitragem, o autor utiliza um caso da indústria farmacêutica, onde tanto as bases administrativas e econômicas da arbitragem são importantes. Existe um cuidado importante de como administrar essa complexa estratégia, para explorar as melhores oportunidades da arbitragem.

Conforme o autor supracitado citado, a arbitragem é, com certeza, a estratégia internacional original. Muitos dos grandes homens de negócio e comerciantes ao longo da história tiveram o comércio de produtos de luxo como fonte de suas receitas, estes produtos eram sujeitos à extremas diferenças em custos absolutos e viabilidade. Todavia, apesar desta retrospectiva histórica, a arbitragem é ilustrada em questões contemporâneas sobre globalização.

Para o autor apesar da total importância da arbitragem para a tomada de decisão nas empresas, existem alguns fatores que levam muitos dirigentes de empresas a não perceberem a importância da arbitragem.

Existe uma crença que os fatores fundamentais da arbitragem, como capital e trabalho, ofereçam apenas oportunidades limitadas, o que também é um erro grave, basta se comparar o custo de mão-de-obra de um chinês e de um americano para ver como existem diferenças a serem exploradas nestes dois fatores produtivos. A noção que o potencial de lucro da arbitragem é bastante limitado também é um erro. Se assim fosse, as empresas americanas não estariam transferindo grande parte de sua produção para China e Índia. Outra razão envolve o assunto do porque as empresas não se esforçam mais em desfrutar das vantagens da arbitragem, mesmo quando elas a reconhecem. Isso se da por fatores organizacionais e falta de estrutura para se internacionalizar e também por fatores políticos como a delicada questão da exportação de empregos.

Abaixo esta um quadro no qual Ghemawat (2007) tenta mostrar de forma didática, como as empresas se movem, planejam e escolhem as três estratégias AAA, e organiza sua produção, baseadas em fatores como a globalização dos mercados, reação local, economias de escala e economia absoluta.

Figura 2: Globalização de Mercados e Produção



Fonte: (GHEMAWAT, 2007, p. 198)

Yongjiang Shi (2007) Johanson e Wierdersheim-Paul (1975) resumem quatro formas de diferentes de modos de entrada em mercados estrangeiros, que são sucessivos e representam maiores graus de envolvimento internacional.

- Estágio 1 – Sem atividades de exportação regulares
- Estágio 2 – Exportação via representantes independentes (agentes independentes de comércio exterior)
- Estágio 3 – Estabelecimento de um subsidiaria de vendas num país estrangeiro.
- Estágio 4 – Unidades de manufatura e produção internacional.

Nesta perspectiva o menor grau de internacionalização se caracteriza quando a empresa nem ao menos tem uma atividade de exportação regular, posteriormente há a possibilidade de representações informais e posteriormente formais, chegando ao grau máximo de internacionalização com a implantação fora do país sede.

4. DESENVOLVIMENTO SUSTENTÁVEL

4.1 Histórico

Segundo Romeiro (2003) um ecossistema em equilíbrio, não é o mesmo que um ecossistema estático, assim considerasse um ecossistema em equilíbrio aquele, que se modifica, embora que de forma lenta, em conseqüência as várias interações existentes entre as diversas espécies nele contida, num processo de co-evolução.

Para o autor, desde dez mil anos atrás, com o desenvolvimento da agricultura o homem deu um grande passo na forma que lidava com a natureza, causando a agricultura um grande impacto sobre os ecossistemas. Todavia a agricultura não necessariamente compromete o meio-ambiente, sendo possível manter um equilíbrio ecológico concomitante a esta atividade econômica.

Com a Revolução Industrial a forma com que a sociedade começou a intervir na natureza passou a se tornar mais predatória e ultrapassar os limites impostos pela mesma. Se não bastasse isto a Revolução trouxe consigo o uso intensivo de combustíveis fósseis, dando possibilidade de produção em escala e pressionando fortemente a base de recursos naturais, antes considerados perfeitamente substituíveis. Já é sabido que a capacidade de carga da terra não poderá ser ultrapassada sem que venham a ocorrer grandes catástrofes ambientais, como tsunamis, terremotos e etc.

Em 1962 a bióloga da marinha americana Rachel Carson, considerada junto com René Dubos um dos precursores do movimento ambientalista mundial. publicou seu livro "Primavera Silenciosa". Este livro causou grande polêmica por fazer criticas ao Dicloro-Difenil-Tricloroetano – DDT, inseticida usado de forma maciça pós década de 40. Carson mostrou que o DDT se acumulava ao longo da cadeia alimentar, e já que era **insolúvel** em água, acabava chegando ao topo da cadeia alimentar, o homem. A partir deste estudo os fabricantes passaram a ter a obrigação de provar que seus produtos eram seguros. Já em 1970 o DDT foi banido em diversos países.

Já René Dubos biólogo francês com cidadania americana, conhecido pela criação da vacina BCG que combate a tuberculose, foi outro precursor de estudos ambientais. O biólogo estudou a interação que existia entre o homem e o planeta terra, este pensador tinha grande preocupação com a forma com que o homem tratava o meio ambiente, todavia era bastante otimista com relação ao futuro e acreditava que os homens conseguiriam resolver os problemas existentes, sendo este movimento conhecido como cornucopiano.

No Clube de Roma, expoentes cientistas e estadistas, banqueiros, industriais de todo o globo organizaram-se preocupados com a situação que se encontrava o planeta, contrataram em 1969 uma equipe de pesquisadores do Massachusetts Institute of Techology – MIT, para realizar um estudo sobre o futuro do planeta, esta previsão com soluções para o futuro foi publicada pela coordenadora de pesquisa Donella Meadows em 1972 com o nome "Limites do Crescimento". Este livro mostrava que com o ritmo de crescimento que a humanidade se encontrava, existiam sérios riscos de sobrevivência, já que os recursos naturais não estavam sendo utilizados de forma racional.

No mesmo ano da publicação deste livro, em 1972 foi convocada a primeira conferencia das Nações Unidas para o Meio Ambiente e Desenvolvimento Humano realizada em 1972 em Estocolmo na Suécia, com a participação de 250 ONGs e 113 países. Houve um grande embate entre os países desenvolvidos e os países em desenvolvimento, já que os países desenvolvidos levantaram a problemática que os países em desenvolvimento deveriam diminuir a velocidade do crescimento econômico, já que o planeta não suportaria tamanha demanda por recursos naturais, por outro lado os países em desenvolvimento defendiam que a pobreza era a maior das

poluições e não aceitavam diminuir seu ritmo de crescimento por causas ambientas, até porque os países desenvolvidos não o tinham o feito. Entretanto, apesar do impasse foi aprovada a Declaração sobre o Ambiente Humano com 110 recomendações que propunham formas de gerir as questões ambientais.

Em 1987 foi publicado pela Comissão Mundial sobre o meio ambiente e Desenvolvimento (CMMAD) mais conhecido como Relatório BRUTLAND, este relatório defendia que:

- O crescimento econômico e proteção ambiental não são excludentes e podem ocorrer simultaneamente
- A pobreza e as questões sociais e não só as econômicas devem ser agrupadas nesta querela ambiental
- Tem que se levar em consideração como as gerações atuais e futuras serão afetadas pelos problemas ambientais criados por nossas ações

No ano de 1992 ocorreu a 2ª Conferência das Nações Unidas para o Meio Ambiente desenvolvimento, tendo o número recorde de 178 países reunidos para debater questões ligadas ao meio ambiente propuseram uma série de convenções e tratados para resolver o problema ocasionado pela exploração desenfreada dos recursos naturais e o aumento da poluição. A conferência tinha como objetivo examinar a situação ambiental, auxiliar a transferência de tecnologias não poluentes principalmente para países em desenvolvimento, e buscar uma cooperação internacional para alcançar as metas estabelecidas. Este conferência teve uma conotação sócio-política bem mais contundente que anterior gerando varias declarações, cartas e convenções como: a declaração de princípios florestais, convenção sobre mudanças climáticas, convenção da biodiversidade, agenda 21 e a carta da terra.

Posteriormente a conferência de 1992, o tema sustentabilidade ganhou bastante força na sociedade, então no ano de 1997 ocorreu no Japão uma conferência sobre mudanças climáticas conhecida como Protocolo de Kyoto, onde ficou estabelecido que os países industrializados deveriam reduzir a emissão de gás carbônico para diminuir o efeito estufa, devendo chegar em 2012 a níveis 5,2% inferiores aos de 1990. Entretanto, o protocolo foi boicotado pelo maior poluidor mundial na época os EUA, e só foi implementado em 2004quando a Rússia assinou o protocolo, em troca de seu ingresso na Organização Mundial do Comércio. Em 2002 ocorreu a Rio +10, em Johansburgo na África do Sul, onde foi apontado que apesar de avanços conseguidos as metas do protocolo de Kyoto não haviam sido alcançadas na primeira década. Já no ano de 2003 entrou em vigor o Protocolo de Cartagena que tratava da questão da biosegurança.

4.2 Conceito

Segundo MAIOLI (2009) em 1973 Maurice Strong e Ignacy Sachs desenvolveram o conceito de ecodesenvolvimento, no qual defendia uma forma alternativa de desenvolvimento diferente da atual. Segundo estes pensadores o desenvolvimento poderia contemplar três fatores: econômico, ambiental e social.

Em 1976 Sachs formulou as dimensões do ecodesenvolvimento: a sustentabilidade social com a satisfação das necessidades básicas, a sustentabilidade econômica com a solidariedade para com as gerações futuras e a busca da participação da população envolvida, a sustentabilidade ecológica que busca a preservação dos recursos naturais e do meio ambiente em geral, a sustentabilidade espacial que busca elaboração de um sistema social garantindo emprego, segurança social e respeito a outras culturas, e a sustentabilidade cultural através do desenvolvimento de programas educacionais.

Esse conceito de Sachs foi aperfeiçoado em 1987 pelo Relatório de Brutland que conceituou desenvolvimento sustentável como "o processo de desenvolvimento que permite às gerações atuais satisfazerem as suas necessidades sem colocar em perigo a satisfação das necessidades das gerações futuras".

Mais adiante a definição de desenvolvimento sustentável foi melhor explicitada pela comissão das nações unidas como "um processo de transformação no qual a exploração dos recursos, a direção dos investimentos, a orientação do desenvolvimento tecnológico e a

mudança institucional se harmonizam e reforçam o potencial presente e futuro a entender às necessidades e aspirações humanas".

4.3 Sustentabilidade e as Empresas

Segundo Leefers e Castillo (1998), as disciplinas de economia e ecologia são geralmente apresentadas como disciplinas opostas. Ambas as disciplinas têm suas virtudes e defeitos. Assim atualmente existe um novo campo de estudo multidisciplinar conhecido como economia ecológica, tendo como objetivo aliar às virtudes de ambas as disciplinas com uma visão de desenvolvimento sustentável de longo prazo.

Para os autores existe um conceito equivocado que economia é sinônimo de finanças (lucro, eficiência, rentabilidade), que é apenas uma das áreas de estudo da economia. Muitos estudiosos não levam em consideração que o estudo econômico e bastante abrangente e engloba várias outras áreas do conhecimento humano.

Nesta perspectiva o estudo da economia ecológica seria uma abordagem multidisciplinar que têm foco no tripé: sustentabilidade, igualdade de distribuição da riqueza e alocação eficiente.

Para Vinha (2003) com o crescente acúmulo de poder das multinacionais existe o questionamento, se estas grandes corporações teriam a responsabilidade de liderar a mudança no padrão de desenvolvimento econômico em direção ao desenvolvimento sustentável e se as mesmas poderiam intervir no rumo das políticas de sustentabilidade ambiental em escala nacional e internacional.

Surge então o grande desafio de conjugar melhoria contínua da qualidade ambiental das instituições com melhores resultados econômicos, em termos de eficiência produtiva e lucratividade. Alguns aspectos importantes devem ser considerados como a economia de recursos naturais e energéticos, o reaproveitamento de resíduos e a reciclagem, a comercialização de resíduos ou o seu tratamento antes do lançamento na natureza, a conquista de novos mercados, a melhoria da comunicação com as comunidades e a instâncias governamentais, inclusive com a redução de custos decorrentes de multas e das indenizações.

Esta visão já é compartilhada por um grande número de empresas que perceberam a consolidação da gestão ambiental e sua projeção a nível internacional. E se utilizam da eco-eficiência como um dos pilares de sucesso de seus negócios. Assim se coloca em questão a responsabilidade não só do governo, mas das empresas e da sociedade civil com a problemática ambiental. Outro questionamento é que a maioria das empresas que estão aderindo a este movimento são do setor secundário, visto que os bancos com raras exceções, ainda estão em uma estágio muito incipiente nesta questão

Para Lustosa (2003) existe uma *trade off* entre a questão ambiental e a eficiência e competitividade econômico, no caso de intervenção governamental via políticas públicas. Todavia alguns economistas teriam flexibilizado este problema, já que algumas empresas encaram a regulação econômica ambiental com inovações, melhorando a competitividade, a chamada "Hipótese de Porter". Para a autora existem vários exemplos de empresas que sofreram pressões para tornarem seus produtos e/ou métodos de produção ambientalmente corretos, fazendo com que as empresas utilizem-se dos fatores de produção de forma mais eficiente, reduzindo assim os custos e compensando os gastos com melhorias ambientais. Desta forma a preservação esta associada ao aumento da produtividade dos recursos utilizados no processo produtivo e concomitante ao aumento da competitividade da empresa.

Este aumento da produtividade dos fatores de produção seria possível, já que muitas vezes a poluição é desperdício econômico. Com as inovações que existem hoje no setor industrial, já é possível a reutilização de vários resíduos indústrias. No que diz respeito ao ciclo do produto também se encontram desperdícios, como o descarte de embalagens e o descarte de produtos que requerem uma disposição final de alto custo. A política de refil resolve parte destes problemas para uma boa gama de produtos. Ambos os desperdícios com resíduos industriais ou com os desperdícios ao longo da vida do produto estão embutidos no preço dos produtos, fazendo os consumidores finais pagarem por este desperdício, sem tomar consciência, desta má utilização dos recursos.

Para a autora supracitada para se alcançar a utilização mais racional dos recursos, é necessário a utilização de inovações tecnológicas, que aumentem a produtividade e tornem a empresa mais competitiva, gerando redução de custos ou melhorias no produto.

Nesta perspectiva as inovações ambientais resultantes de regulamentações podem ser classificadas em duas categorias:

- Ações tecnológicas mais eficientes da empresas no tratamento da poluição que já ocorrera, que também é conhecida como tecnologia *end-of-pipe.*
- Inovações que compensam os custos de implementação, ocorrendo redução do impacto ambiental e melhora do produto ou processo produtivo. Também conhecida como tecnologia de *pollution prevention.*

Segundo Porter e Van der Linde (apud Lustosa 2003) as regulamentações seriam necessárias para:

- Motivar as empresas a inovarem.
- Melhorar a qualidade ambiental dos produtos e processo produtivo.
- Alertar e educar as empresas sobre a atualização ineficiente dos recursos e sobre as áreas que tem potencial de melhorias.
- Aumentar a probabilidade de que as inovações de produto e processo possam ser ambientalmente corretas.
- Criar demanda para as melhorias ambientais até que as empresas e os consumidores sejam capazes de perceber e quantificar melhor as ineficiências no uso dos recursos.
- Nivelar as regras do jogo durante o período de transição para inovações ambientais, para incentivar os investimentos inovativos prol meio ambiente.

O modelo de Porter é muito interessante e sua eficácia já é verificada em vários países que possuem uma legislação ambiental rigorosa como é o caso dos países nórdicos.

Todavia além do processo coercitivo a partir de políticas públicas que protejam o meio ambiente e motivem as empresas a inovarem e melhorarem seu processo produtivo diminuindo a poluição, também existe outro processo, comandando pela sociedade conhecido como "*buycott*".

A palavra *buycott* é o antônimo da palavra *boycott*. O *boycott*, que é o ato político da não aquisição de alguns produtos, já o *buycott* é o ato político da compra de algum bem, baseado no pressuposto que esta ação estará trazendo benefícios para a sociedade. Sendo assim o consumidor que se utiliza do *buycott* é geralmente politizado, informado e responsável. Na Suécia, o percentual de cidadãos que se envolvem em algum tipo de "consumo politizado" chega perto dos 50%. Todavia em países como o Brasil este percentual não chega à casa dos dois dígitos, até por que uma boa parte dos consumidores tem necessidades bem mais primárias para dar prioridade do que questões políticas ou ambientais.

Alguns ambientalistas e estudiosos do assunto defendem que o ato de consumo pode se transformar em ativismo político. A falta de uma regulamentação global transferiu para os consumidores parte da responsabilidade sobre o mercado, fazendo com que esta classe se torne mais consciente e pressionem as empresas à adotarem medidas ecologicamente sustentáveis. Assim pode-se concluir que a sociedade de consumo possui ferramentas sofisticadas e inteligentes para promover ajustes no mercado em favor da sustentabilidade.

Segundo Valeria (2008) com a pressão hoje existente por parte dos governos e órgãos supranacionais para o aprimoramento da legislação ambiental, a pressão de grupos ambientalistas e o surgimento de consumidores que querem qualidade de vida condicionam o mercado através de *boycott* e *buycott,* cria-se um ambiente propício para o desenvolvimento do marketing ambiental. As empresas e outras entidades, entre elas as organizações não-governamentais, reforçam uma imagem positiva frente à sociedade e posicionam seu produto ou serviço no mercado, diferenciando-o dos demais a partir do *marketing* ambiental.

Para a autora supracitada o *marketing* ambiental é uma modalidade que visa enfocar as necessidades de consumidores ecologicamente conscientes e contribuir para a criação de uma sociedade sustentável.

Existem muitas motivações que levam as empresas a adoção do *marketing* verde. A primeira é que este segmento do *marketing* é bom para as empresas porque busca a eficiência no uso dos materiais e também ajuda a reduzir os custos. Segundo, as empresas que não responderem às questões ambientais com produtos mais seguros e ambientalmente mais saudáveis estão se arriscando a perder a sintonia com o consumidor. E por último, ao empregar estratégias de *marketing* ambiental, a entidade pode aumentar sua credibilidade e legitimidade, definir sua personalidade, área de atuação e imagem, além de agregar valor à marca junto aos diversos compradores, agregando valor a sua marca e fidelizando o cliente.

O *marketing* ambiental vai muito além de ações pontuais como a o invento e utilização de refis ou embalagens recicláveis, mas sim em administrar uma cadeia de fatos que envolvem várias etapas, desde a fabricação, relação com fornecedores ou parceiros, clientes, empregados, mídia e comunidade.

3 - O CASO DA NATURA: EFICIÊNCIA ECONOMICA ALIADA À SUSTENTABILIDADE

3.1 - Histórico

Em 1969, abre-se um laboratório e uma pequena loja na Rua Oscar Freire, em São Paulo, por Luiz Seabra, constituindo-se a Natura, empresa de cosméticos. Em 1974, a empresa adota o modelo da venda direta para a distribuição de seus produtos. Essa decisão se baseou na crença da importância das relações e foi determinante para o crescimento da empresa.

Com o tempo, dois novos diretores e sócios passaram a trabalhar ao lado de Seabra: Guilherme Peirão Leal e Pedro Luiz Barreiros Passos, que ingressaram na Natura em 1979 e 1983, respectivamente, ajudando a consolidar os processos de gestão e de produção da empresa dando um novo rumo estratégico a mesma.

A década de 80 foi muito importante para a consolidação da empresa, registrando aumento da receita em mais de 3000%. Em 1983, a Natura foi um dos primeiros fabricantes brasileiros de cosméticos a introduzir o uso de produtos com refil, permitindo economia de custo para o consumidor, redução do desperdício e ampliação da consciência ecológica, ainda hoje marcas registradas da empresa

Entre 1990 e 1992, surgiu a necessidade de expressar de maneira mais clara o jeito Natura de fazer negócios, com a criação de novos produtos que fortaleceram o conceito lançado pela empresa, de preocupação total com todos os períodos da vida de seus consumidores. Em 1994, a empresa começa seu processo de expansão, devido a aproximação cultural o primeiro região escolhida foi a América Latina. Além da expansão internacional a empresa investiu fortemente na responsabilidade corporativa lançando conjuntamente com a fundação Abrinq o programa crer para ver, como a missão da melhora do ensino das escolas públicas e também foi a companhia que apoio a criação do Instituo Ethos.

No ano de 1999 a Natura inova mais uma vez ao adotar o uso sustentável da biodiversidade brasileira como plataforma tecnológica de P&D da corporação, o que guiou a lançar, em 2000, a linha *Natura Ekos,* com fragrâncias e produtos desenvolvidos a partir de ativos obtidos de plantas da flora brasileira, principalmente da região amazônica. Também no final da década de 90, foi aprimorado a governança corporativa da companhia por meio da criação do conselho e comitês, com o objetivo de elaborar estratégias de desenvolvimento para seus colaboradores e administrar a remuneração dos executivos da empresa

Em 2001, foi inaugurado um moderno complexo industrial, de pesquisa e desenvolvimento e de treinamento e logística em Cajamar. Já em de 2004, foi completada outra reorganização societária na qual a Natura Cosméticos incorporou a Natura Empreendimentos e a Natura Participações. Esta reorganização teve como objetivo simplificar a estrutura societária da empresa.

Como parte da estratégia de internacionalização, a Natura inaugurou em 22 de abril de 2005 uma loja em Paris, em ponto privilegiado da capital francesa. O primeiro caso de Investimento Direto Estrangeiro – IDE da natura e neste mesmo ano, a Natura também iniciou operações no México.

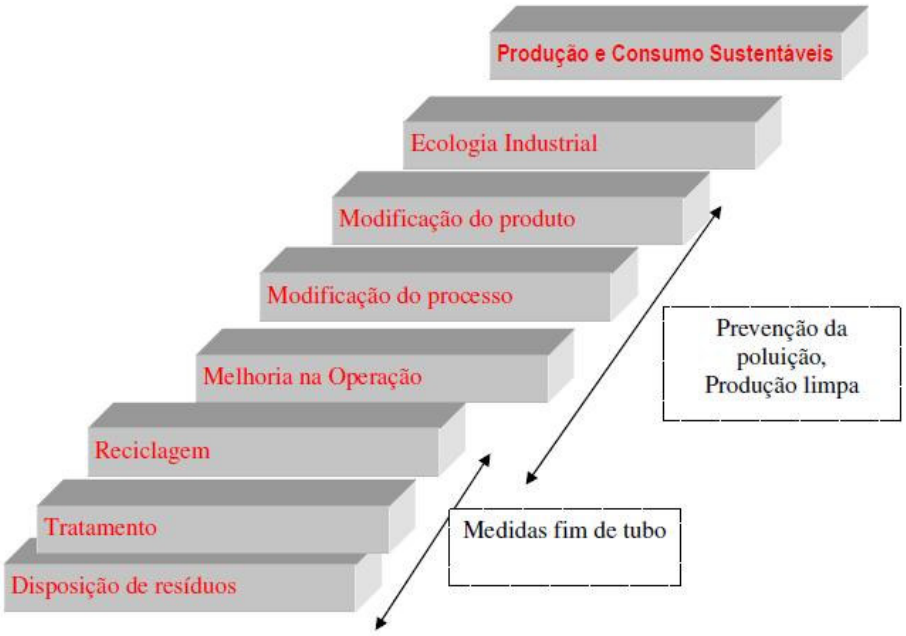
3.2 Modelo de Sustentabilidade da Empresa

A Natura é uma empresa de referência no segmento de cosméticos, não só a nível nacional, mas internacional. Em 2005, a Natura recebeu diversos prêmios. Entre eles destacam-se os seguintes:

Prêmio Top of Mind - 1º lugar na categoria Produtos e Cremes de Beleza – Datafolha, pelo segundo ano consecutivo; Empresa mais Admirada no Brasil, também pelo segundo ano consecutivo – Revista Carta Capital/Intersciense; Melhor Empresa para Mulher Trabalhar, pelo segundo ano consecutivo - Revista Exame – Ranking Você S/A; Quarta Marca Mais Valiosa no Brasil – Revista Isto É/Interbrand; Prêmio Aberje de Comunicação Empresarial; Melhor Relatório Anual - Abrasca; IR Awards Brasil - Revista IR Magazine; Prêmio Balanço Anual - Gazeta Mercantil; Melhor Empresa do Setor Farmacêutico - Revista Isto É Dinheiro; Prêmio L'enseigne d´OR - Revista " Site commerciaux"; Prêmio ECO AMCHAM – Câmara Americana de Comércio; Selo Empresa Cidadã; Prêmio Fiesp de Mérito Ambiental; Classificada entre as 10 empresas-modelo em Responsabilidade Corporativa – Revista Exame; Empresa de Valor e Melhor Empresa do Setor de Higiene e Cosméticos – Jornal Valor Econômico; Prêmio Valor Social – Jornal Valor Econômico; Prêmio DCI empresas do ano – Categoria Cosméticos, Higiene e Limpeza.

Aliado a este prestigio e sucesso a nível empresarial a Natura conseguiu criar uma estrutura de produção que privilegia o meio ambiente. Este aspecto fica melhor vislumbrado na linha *Ekos*, lançada no ano de 2000. Nesta linha são utilizados ativos da biodiversidade brasileira, principalmente da região amazônica, estes ativos são transformados em produtos, respeitando a noção de sustentabilidade social e ambiental.

Figura 3 – A Evolução das Causas Ambientais



FONTE: KIPERSTOK apud DINATO (2006)

A figura acima mostra a evolução das causas ambientais, e a partir dela podemos explicar o modelo de negócios praticado pela Natura principalmente na linha *Ekos*.

Como foi dito anteriormente desde sua fundação a Natura teve compromisso com a causa ambiental, inovando em processo limpos e eficientes. Em 1999 a Natura lançou a política do uso do refil, incentivando a compra dos mesmos por seus clientes e investindo na educação ecológica dos mesmos.

Em 2000 a Natura lança o Programa de Certificação de Ativos, iniciado em julho, com foco na busca pela sustentabilidade dentro da linha *Natura Ekos*. Esta iniciativa teve como objetivo conhecer e garantir que os insumos vindos da flora brasileira sejam extraídos de forma ambientalmente correta e socialmente justa nos locais de origem de cada um dos ativos. Assim o programa funcionou um instrumento que promove o cultivo e o manejo florestal sustentável, por meio da certificação das áreas de manejo das plantações e florestas nativas, envolvendo critérios ambientais, sociais e econômicos.

Todos os ativos dos produtos são extraídos de forma sustentável, permitindo a renovação das fontes de onde provêm. Além disso, vários destes ativos vêm de reservas extrativistas, áreas naturais demarcadas e protegidas pelo Centro nacional das Populações Tradicionais/Instituto Brasileiro de Apoio ao Meio Ambiente – CNPT/Ibama e por seus moradores. Os produtos originários das reservas extrativistas recebem também o certificado de origem do Ibama. Outro ponto bastante importante e a relação existente socioeconômica existente entre estas comunidades extrativistas e a corporação.

A Natura trabalha com a política do preço justo, pagando a estas comunidades um valor digno por seu trabalho e entendendo que essas comunidades são o elo forte para a preservação do meio ambiente. A partir do momento que estas pessoas obtêm uma renda a partir da utilização racional da natureza, não mais necessitam devastar a floresta para sua sobrevivência, pelo contrário começam a se tornar fiscais em prol da manutenção da biodiversidade do ecossistema.

Assim a empresa respeita a declaração dos direitos humanos presentes no artigo XXIII, parágrafo 3 que diz:

- "Toda pessoa que trabalhe tem o direito a uma remuneração justa e satisfatória, que lhe assegure, assim como à sua família, uma existência compatível com a dignidade humana, e a que se acrescentarão, se necessário, outros meios de proteção social."

E o artigo XXV parágrafo 1:

- "Toda pessoa tem o direito a um padrão de vida capaz de assegurar a si e a sua família saúde e bem estar, inclusive alimentação, vestuário, habitação, cuidados médicos e os serviços sociais indispensáveis, e direito à segurança em caso de desemprego, doença, invalidez, viuvez, velhice ou outros caso de perda dos meios de subsistência fora de seu controle."

Com esta linha de negócio, a Natura consegui, de forma exemplar, unir desenvolvimento socioeconômico, eficiência econômica e preservação do meio ambiente, servindo de *benchmark* para outras empresas do setor.

As fórmulas utilizadas pela natura são comprovadamente biodegradáveis, ou seja, se decompõem na natureza em até 28 dias. A linha prevê o uso de embalagens de papel e resinas plásticas recicláveis e entre essas várias são recicladas. Todos os produtos da linha trazem a possibilidade do refil, o que ajuda a preservar o meio ambiente, pelo menor uso de matéria-prima.

A certificação para ativos utilizada pela natura é baseada nos princípios e critérios do Forest Stewardship Council – FSC, organização não governamental internacional que certifica todas as florestas mundiais. No que diz respeito aos ativos que vêm de áreas de cultivo, os princípios são os do Conservation Agriculture Network – CAN. Em síntese, o Programa de Certificação de Ativos é realizado em seis etapas principais:

- levantamento da situação atual do local de origem do ativo,
- preparação do plano de manejo,
- avaliação do impacto social e ambiental,
- implementação do plano de manejo, e
- obtenção do certificado e o monitoramento periódico.

Assim pode-se perceber que a Natura consegue atender toda a evolução das causas ambientais. Inovando na forma de produzir, na forma de relacionar-se com seus fornecedores e

colaboradores e principalmente, mudando o padrão de consumo, oferecendo produtos de qualidade que não agridem o meio ambiente e conseguindo aliar este respeito ao meio ambiente a uma empresa altamente rentável e em pleno processo de expansão internacional.

3.3 Expansão Internacional

O processo de internacionalização da Natura é sintetizado no quadro abaixo.

Tabela 1 – Histórico de Internacionalização da Natura

País	Data	Forma de Atuação	Vendas
CHILE	Começo de 1980*	1° Distribuição local	
ARGENTINA	Começo de 1990*	2° Escritório e Centro de Distribuição	Vendas diretas
PERU			
BOLÍVIA	Final de 1980*	Representante comercial	
MÉXICO	Agosto/2005*	Escritório	
FRANÇA	**Abril/2005***	**Escritório e loja**	**Loja e site**

* Embora não se tenha conseguido precisar o ano de início das operações pela pesquisa em fontes
Fonte: ARAUJO; BLUNDI (2009)

O processo de internacionalização da Natura pode ser interpretado a partir do Modelo de Uppsala, já que a Natura começou a atuar em países de padrão cultural próximos da America Latina. O que corrobora com a visão da Escola Nórdica.

A Natura difere da maioria das empresas que entram em processo de internacionalização. A cultura de internacionalização da Natura é não adotar alianças, mas sim investir diretamente no país, já que desta forma a companhia tem uma maior controle sobre as operações e não necessita compartilhar informações confidencias e estratégicas sobre os meios de produção e as fórmulas de seus produtos. Nesta linha de negócio, Natura adota então uma estratégia de internacionalização via investimento direto e sem parceiros internacionais no novo país.

No que diz respeito a comprometimento dos negócios internacionais, verifica-se que o menor grau de internacionalização é encontrado na Bolívia, uma vez que até hoje a corporação atua neste país por meio de um representante comercial. Isto também se deve ao fato que dos seis países a Bolívia é o de menor poder aquisitivo.

Infere-se também que nos outros países da America do Sul: Chile, Argentina e Peru, a Natura continuou seguindo os passos propostos pela Escola Nórdica. Devido ao fato que a companhia começou seu processo de atuação nestes países por meio de distribuidores, e somente depois de uma maior maturação do negócio, é que houve aumento do comprometimento com a instalação de unidades físicas na forma de escritórios e de centros de distribuição.

No caso mexicano a natura foi mais arrojada em seu processo de internacionalização iniciando suas operações com Investimento Direto Estrangeiro – IDE. Todavia esse grau superior de comprometimento inicial apresentado na entrado no mercado mexicano, pode ser compreendido a partir da experiência obtida anteriormente nos países sul americanos, que são culturalmente e economicamente parecidos com o México.

Até então a Natura só havia se internacionalizado para países culturalmente similares, e em desenvolvimento, sendo assim a entrada no mercado francês foi uma mudança de paradigma no histórico empresarial da empresa.

Em 22 de abril de 2005 a empresa inaugurou a sua Loja em *Saint Germain des Prés*, ponto privilegiado da capital francesa. A escolha da abertura de uma loja se deu devido a importância do novo empreendimento e as diferenças culturais existentes entre Brasil e França, a Natura mudou sua estratégia de venda direta para outro modelo de vendas, o qual se adaptava ao mercado europeu.

Para a entrada no mercado francês a Natura necessitava de um diferencial para obter sucesso, já que a França é considerada por muitos como o principal mercado na área de cosméticos, é já dispunha até de produtos feitos com ingredientes naturais, como os da *Body*

Shop. Entretanto apesar da grande gama de produtos presentes no mercado de cosméticos francês, o mesmo não dispunha de elementos da cultura brasileira e feitos com ingredientes extraídos da biodiversidade brasileira, como os produtos feitos com: andiroba, breu branco, priprioca, buriti, castanha-do-brasil, copaíba, cumaru, guaraná, maracujá, murumuru e pitanga. Para explorar esta idéia de brasilidade a natura escolheu a linha *Ekos* para entrar no mercado francês, uma vez que a mesma utiliza-se de produtos biodegradáveis, com embalagens feitas de material reciclado ou recicláveis.

Desta forma esta preocupação com o desenvolvimento sustentável marcante na linha *Ekos*, unindo modernidade à tradição, foi o diferencial competitivo para o sucesso do IDE feito no mercado francês.

Além da interpretação feita pelo modelo de Uppsala do processo de internacionalização da natura podemos utilizar de três outras visões: A de Ghemawat, a de Benett e a do Boston Consulting Group, e a de Almeida e Cretoiu (2008)

Na perspectiva da interpretação de Almeida e Cretiou (2008), a Natura escolheu dois modelos de estratégia, o geográfico e o pro linhas de negócio.

A empresa se organiza pela geografia, levando em consideração os aspectos relacionados à cultura, economia e costumes locais. O fato da empresa ter sua maior presença na América Latina é um dos aspectos que reforçam esta visão. Entretanto além de se organizar geograficamente a empresa se utiliza de diferentes linhas de negócio, busca se adaptar da maneira mais eficiente em cada país.

Já na visão de Ghemawat (2007) percebe-se que a estratégia utilizada pela Natura é de adaptação. Já que em todos os mercados que a empresa realizou algum tipo de investimento, a mesma sempre tentou se adaptar ao mercado local. Para obter sucesso com esta estratégia a empresa analisou as diferenças que continuam a dividir os países e pensou nas ferramentas que possuía para melhor se adaptar a cada país e mercado.

Na interpretação de Bennett a empresa escolheu a estratégia de internacionalização, já que permanece extremamente etnocêntrica em sua abordagem, possuindo uma modelo de empresarial prioritariamente comercial, e ainda não possuindo nenhuma unidade produtiva fora do país, concentrando assim os aspectos administrativos e produtivos focalizados no país sede que serve de plataforma de exportação.

Já na perspectiva do Boston Consulting Group, a estratégia escolhida pela natura foi a de tornar sua marca global, isso se deu principalmente pelas características do mercado de cosméticos onde o prestigio e o *glamour* são indicadores fortes para o sucesso de um empreendimento. O diferencial competitivo para a consolidação da marca Natura no âmbito nacional e internacional é o modelo de sustentabilidade ambiental e social escolhido pela empresa, que é referência não só no setor de cosméticos, mas também referência de como uma empresa pode aliar lucratividade com sustentabilidade e respeito ao meio ambiente.

Em 2006 a Fundação Dom Cabral conjuntamente com a universidade de Columbia dos Estados Unidos lançaram o *ranking* das vinte empresas mais internacionalizadas do Brasil e a Natura ficou em décimo quinto lugar. Já o *Boston Consulting Group* publicou em 2007 a lista das 100 empresas de países de rápido desenvolvimento econômico, que são ou tornaram lideres globais, e a Natura foi a única empresa de Cosméticos presentes na lista.

Estes estudos mostram o sucesso do processo e estratégia de internacionalização da empresa, que se utilizou de uma estratégia empresarial baseada em princípios de eficiência econômica aliada a sustentabilidade ambiental, inovando não apenas na forma de produzir seus cosméticos, mas também na forma de relacionar-se com fornecedores colaboradores e clientes se tornando sem dúvida em um exemplo a ser seguido.

6. CONCLUSÃO

Desde o começo da década de 80 quando a natura começou seu processo de internacionalização, a companhia vem aliando uma estratégia de negócios agressiva e eficiente com um modelo de gestão sócio-ambiental que respeita e contribui para o desenvolvimento das comunidades que ela esta inserida. Este esforço da empresa vem sendo percebido por seus

clientes, tornando-a referência como empresa que conseguiu aliar os paradigmas econômicos aos ecológicos. A firma é única que representa a indústria de cosméticos na lista do *Boston Consulting Group*, das 100 empresas dos países em desenvolvimento com maior potencial em se tornarem lideres globais, além do que a estratégia utilizada pela empresa e tornar sua marca global, sendo assim o marketing ambiental utilizado pela mesma é o diferencial competitivo para competir com as grandes firmas de cosméticos internacionais.

BIBLIOGRAFIA

ALMEIDA, ANDRÉ (org) (2007), "Internacionalização de Empresas Brasileiras: Perspectivas e Riscos, Editora Campus", 1º edição, Rio de Janeiro.

ALMEIDA, ANDRÉ; CRETOIU, SHERBAN LEONARDO (2008), "Internacionalização de empresas: A experiência brasileira e o contexto latino americano". Disponível em: http://www.fdc.org.br/parcerias/upload/outros/dom/dom5/dom5_campus.pdf.

ARAUJO, LUCIANA VELLOSO DE SOUZA; BLUNDI, MARIA DOMENICA SERPA (2009), "Adaptações Estratégicas no Ambiente Internacional: O Caso da Natura." Disponível em: http://www.coppead.ufrj.br/workshop/docs/artigo_velloso_blundi.pdf

BENNETT, DAVID (2007)," Gestão de operações internacionais: Conceitos Fundamentais", p. 36-52 In: FLEURY, FLEURY (org.) Internacionalização e os Países Emergentes. Editora Atlas , São Paulo.

BOSTON CONSULTANT GROUP (2008), "The New Global Challengers 2007".

Disponível em: http://www.bcg.com/publications/files/New_Global_Challengers_May06.pdf.

BOSTON CONSULTANT GROUP (2008), "The New Global Challengers 2008". Disponível em:http://www.bcg.com/impact_expertise/publications/files/New_Global_Challengers_Feb_200.pdf.

BRASIL, FABIANO LUIS; ORTEGA, LUCIANE MANEGUIN ORTEGA (2009), "Análise de Estratégia de Internacionalização: O Caso Bematech". Disponível em: http://ginebra.incubadora.fapesp.br/portal/referencias/workshop-usp/08.pdf.

CARVALHO, GARDÊNIA MARIA BRAGA DE CARVALHO (2008), "Contabilidade Ambiental: Teoria e Prática". Editora Juruá, 2ª edição. Curitiba.

CYRINO, ÁLVARO BRUNO. BENEFÍCIOS (2007), "Riscos e resultados do processo de internacionalização das empresas brasileiras", p. 79-116 .In: ALMEIDA, André (org). Internacionalização de empresas brasileiras: perspectivas e riscos. Editora Campus, 1ª edição, Rio de Janeiro.

MAIOLI, MARCOS ROGÉRIO (2009) "Desenvolvimento Sustentável: O que é e como chegamos até aqui". Apostila da Aula 1 do MBA em Planejamento e Gestão Estratégica. FACINTER.

MINISTÉRIO DA JUSTIÇA (2009), "Declaração Universal dos Direitos Humanos". Disponível em: http://www.mj.gov.br/sedh/ct/legis_intern/ddh_bib_inter_universal.htm

FLEURY, AFONSO; FLEURY, MARIA TEREZA LEME (org) (2007), "Internacionalização e os Países Emergentes". Editora Atlas, 1ª edição,São Paulo.

FUNDAÇÃO DOM CABRAL; THE COLUMBIA PROGRAM ON INTERNATIONAL INVESTMENT (2008)," A Decolagem das Multinacionais Brasileiras". Disponível em: http://www.fdc.org.br/parcerias/upload/outros/a_decolagem_das_multinacionais_brasileiras.pdf.

GHEMAWAT, PANKAJ (2007)," Redefining Global Strategy: Crossing Borders in a World Where Differences Still", Havard business school publishing corporation, Boston.

JUNIOR, CARLOS HENRIQUE MORA; ZILBER, SILVA NOVAES; LOURENÇO, FRANCISCO (2009),"Estratégias de Internacionalização e Canais de Distribuição no Mercado Externo: O Caso da Natura". Disponível em: http://www.administradores.com.br/producao_academica/estrategia_de_internacionalizacao_e_canais_de_distribuicao_no_mercado_externo_o_caso_natura/1293/

KIM, CRISTIAN (2008), "Estratégias de Internacionalização de Empresas Exportadoras do Estado do Paraná e o Papel de Agentes de Exportação". Disponível em:

dspace.c3sl.ufpr.br:8080/dspace/bitstream/1884/11398/1/Dissertação%20-%20Cristian%20Kim.pdf.

KUAZAQUI, EDMIR (1999), "Marketing Internacional: Como Conquistar Negócios em Mercados Internacionais", MAKRON Books, 1ª edição, São Paulo.

LINHA EKOS NATURA (2009). Disponível em: http://www.natura.net/port/hotsite/ekos/home.asp

LUSTOSA, MARIA CECÍLIA JUNQUEIRA (2003), "Industrialização, Meio Ambiente, Inovação e Competitividade". P. 155-172. In: MAY, Peter H; LUSTOSA Maria Cecília; VINHA Valéria Da (ong) (2003), Economia do meio Ambiente: Teoria e Prática. Editora Campus, 4ª edição, Rio de Janeiro.

MAY, PETER H; LUSTOSA MARIA CECÍLIA; VINHA VALÉRIA DA (ong) (2003)," Economia do meio Ambiente: Teoria e Prática". Editora Campus, 4ª edição, Rio de Janeiro.

MONTEIRO, SOLANGE (2009), " Natura – Mudança de Cor". Disponível em: http://www.global21.com.br/materias/materia.asp?cod=23176&tipo=noticia

SHI, YONGJIANG (2007), "Internacionalização e evolução dos sistemas de manufatura: modelos de processos clássicos, novas questões indústrias e desafios acadêmicos". p. 15-35. In: FLEURY, FLEURY (org.). Internacionalização e os Países Emergentes, Editora Atlas, São Paulo.

SMITH, SANDRA BURLE MARX; RIBAS, JOSÉ ROBERTO (2009), "Gestão com Sustentabilidade: O caso da linha Ekos". Disponível em: http://www.aedb.br/anais-seget/gs.htm.

UNCTAD/ONU (2007), " World Investment Report" . Disponível em: http://www.unctad.org/Templates/WebFlyer.asp?intItemID=4361&lang=1. Acesso em; jul. 2008.

VINHA, VALÉRIA DA (2003), "As Empresas e o Desenvolvimento Sustentável: da Eco-eficiência à Responsabilidade Social Corporativa". p. 173-196. In: MAY, Peter H; LUSTOSA Maria Cecília; VINHA Valéria Da (ong) (2003), Economia do meio Ambiente: Teoria e Prática. Editora Campus, 4ª edição, Rio de Janeiro.

LAS RELACIONES COMERCIALES ENTRE ESPAÑA Y EL MERCOSUR EN LOS ÚLTIMOS DIEZ AÑOS

ALEJANDRO D. JACOBO
E-mail: jacoboa@eco.unc.edu.ar
Departamento de Economía
Universidad Nacional de Córdoba y Pontificia Universidad Católica Argentina

Resumen

El éxito de cualquier asociación económica entre bloques requiere necesariamente el conocimiento y el desarrollo de los flujos comerciales bilaterales entre países que integran uno y otro bloque. Este trabajo describe brevemente las relaciones comerciales entre España y los países del MERCOSUR durante los últimos diez años y sugiere que los flujos bilaterales deberán profundizarse si se desea avanzar en la asociación.

Códigos *JEL*: F13, F15
Palabras clave: Comercio, Integración Económica, Política Comercial
Area tematica: Economía Internacional y de la Integración.

Abstract

The success of any economic association among blocks necessarily requires the knowledge and development of bilateral trade flows between the countries belonging to one and to the other bloc. This paper briefly describes the commercial relationships between Spain and MERCOSUR countries during the last ten years and it suggests that bilateral trade flows must be strengthened to achieve the full association.

Thematic Area: International and Integration Economics.

1. INTRODUCCIÓN

Es un hecho innegable la proliferación de organizaciones regionales con fuerte componente socioeconómico y geopolítico a través de las cuales las naciones han estrechado vínculos y han buscado una cooperación de mutuo beneficio. En este sentido, asociaciones y acuerdos de cooperación han tomado cuerpo a través de bloques, no sólo en países desarrollados, poseedores de sólidas instituciones, sino también en países en vías de desarrollo, caracterizados por instituciones débiles o aún en proceso de formación.

A su vez, algunos de estos bloques han tratado de fortalecer las relaciones entre sí, constituyendo un claro ejemplo de ello el Mercado Común del Sur (MERCOSUR) y la Unión Europea (UE). Este acercamiento tiene características muy particulares, ya que es la primera vez que dos regiones económicamente integradas y distantes geográficamente negocian un posible acuerdo de integración.[1] El nacimiento formal de la iniciativa puede situarse en 1994 a través del Protocolo de Ouro Preto, a pesar de que, como señala Torrent (2005), las conversaciones para plantear un acuerdo de organización (MERCOSUR) a organización (UE) se iniciaron con

[1] Esto último más allá de sus dificultades actuales o los resultados alcanzados hasta el momento.

antelación. En la actualidad, la relación entre ambos bloques está basada en el Acuerdo Marco de Cooperación Interregional MERCOSUR-UE de 1995, firmado en diciembre de ese año en Madrid, y cimentado en tres elementos principales: el diálogo político, la cooperación y el comercio.

Si bien como pilares el diálogo político y la cooperación son importantes, es incuestionable que para lograr fluidez y fortalecer la relación entre los bloques interesa extraordinariamente el tercero de esos elementos: el pilar comercial. E interesa sobremanera debido a que un intercambio profundo significaría la conformación del segundo bloque comercial del planeta, con un producto interno bruto estimado en 9 billones de euros y una población superior a 700 millones de personas.

En este sentido, el patrón de comercio que se genera entre los países de un bloque y de otro puede fortalecer, entorpecer o destruir completamente lo acordado entre los bloques de pertenencia. Un país que no comercie con otro tendrá pocos incentivos a que sus relaciones económicas prosperen y forzará, tarde o temprano, a que los vínculos entre los bloques se deterioren y los acuerdos establecidos sucumban. Es por ello que poco puede avanzar el todo sin el esfuerzo de las partes y que la prosperidad de cualquier asociación, aún entre bloques, requiere necesariamente del conocimiento y fortalecimiento de las relaciones comerciales bilaterales entre cada uno de los países que integran aquéllos

En este marco, el presente trabajo es un intento por aproximarse a caracterizar los intercambios comerciales bilaterales entre España y los países del MERCOSUR en los últimos diez años y observar si tal relación opera en el sentido esperado para una asociación exitosa. Se presume que los intercambios comerciales entre estos países, aunque interesantes, no han tenido mayor dinamismo y ello ha coadyuvado, junto a otros factores y dejando de lado las discusiones políticas y de cooperación de otro nivel, a que las relaciones se debiliten en el tiempo.[2] El seleccionar España por un lado, al MERCOSUR como la otra parte del análisis y los últimos diez años como período de tiempo no resulta un hecho caprichoso.

En efecto, a pesar de lo reducido del comercio entre ambos, España dispone de importantes ventajas competitivas en comparación con sus socios comunitarios, su sector financiero tiene un papel protagonista en las relaciones comerciales con América Latina y, desde principios de los años noventa del siglo que se fue, las empresas españolas han destinado un volumen importante de inversiones en el conjunto de países que integran el bloque sudamericano (Balaguer Franch y Martínez Zarzoso, 2000; p. 121).[3]

Por su parte, el MERCOSUR es el cuarto bloque comercial del mundo, cuenta con un mercado de más de 260 millones de potenciales consumidores y un producto interno bruto (PIB) cercano a 1,3 billones de dólares (80% del PIB sudamericano), una perspectiva de crecimiento favorable, abundantes recursos naturales, minerales y energéticos, y también cuenta con una estructura demográfica joven.

Con relación al período de tiempo, el Acuerdo Marco suscripto en 1995 entró en vigor el 1 de julio de 1999, razón por la cual aparece éste como un año adecuado como punto de partida para explorar, en términos generales, lo sucedido, observando para tal fin la distribución porcentual del comercio exterior como indicador simple de intensidad del comercio.

El trabajo se organiza como sigue. La sección 2 destaca el papel del MERCOSUR como uno de los bloques más importantes en la región. La sección 3 se ocupa de revisar el destino geográfico de los flujos comerciales bilaterales, principalmente del papel que ocupa América en el intercambio comercial con España. La sección 4 recoge alguna evidencia sobre la tendencia de los flujos comerciales específicos con el MERCOSUR. La sección 5 efectúa algunos comentarios finales.

[2] Es probable que haya sucedido algo similar con otros países. Por un análisis más detallado, véase Jacobo (2009).

[3] Esto se ha producido en coincidencia con el fortalecimiento de la estructura empresarial española, y la progresiva y oportuna privatización de los sectores productivos de los países iberoamericanos.

2. ACUERDOS DE INTEGRACIÓN REGIONAL EN AMÉRICA: El MERCOSUR

Existen en América diferentes organizaciones de integración regional que impulsan la creación de espacios económicos y áreas de libre comercio capaces de competir en el escenario internacional y afrontar los retos que impone la globalización. Así, conviven actualmente en el continente americano el Tratado de Libre Comercio de América del Norte (NAFTA, en su sigla inglesa), entre Canadá, Estados Unidos y México; el Mercado Común Centro Americano (MCCA) y la Comunidad del Caribe (CARICOM), que se mantienen activos en América Central[4]; y, en coexistencia con la Asociación Latinoamericana de Integración (ALADI)[5] y la Comunidad Andina, el Mercado Común del Sur (MERCOSUR).[6]

En cuando al MERCOSUR, el Tratado de Asunción de 1991 proporcionó origen formal a un proceso de integración entre Argentina, Brasil, Paraguay y Uruguay, entrando en vigor en 1994.[7] Con la creación del MERCOSUR, los socios buscaron la ampliación de las dimensiones de los respectivos mercados nacionales a través de la integración, como una condición para acelerar sus procesos de desarrollo. El objetivo principal del Tratado es la integración de los Estados parte a través de la libre circulación de bienes, servicios y factores productivos, el establecimiento de un arancel externo común, la coordinación de políticas macroeconómicas y sectoriales y, en general, la armonización de legislaciones en las áreas pertinentes que afiance y fortalezca el proceso de integración.

Desde su nacimiento, el MERCOSUR ha generado una unión aduanera imperfecta, pues no existe plena circulación de factores y de personas entre sus miembros, y tampoco ha avanzado demasiado en la armonización de políticas macroeconómicas y en el fortalecimiento de las instituciones necesarias en toda organización regional para su funcionamiento. De todos modos, es el proceso más joven y exitoso de integración económica en América Latina; aunque en los últimos años se ha visto impedida su profundización por motivos diversos.[8]

Como bloque, el MERCOSUR es la segunda organización regional de mayor envergadura del continente por lo cual, si los países asociados se decidieran involucrar en forma plena en este proyecto de integración regional, se convertiría en una organización poderosa, a la vez que encauzaría el desarrollo de América Latina (Doval Adán, 2008; p. 29).

Las primeras relaciones formales del MERCOSUR con la UE tuvieron lugar con la firma del Acuerdo Marco Interregional de Cooperación Política, Económica y Social. El objetivo fundamental era la creación de una zona de libre comercio a partir del año 2005, lo que no se cumplió. En líneas generales, el contenido del Acuerdo es de cooperación técnica y de avance en el conocimiento mutuo, y cabe mencionar su carácter no preferencial, transitorio y evolutivo. El Acuerdo se sitúa dentro del espíritu de la política comunitaria de cooperación con los países en vías de desarrollo y es conveniente situar dentro suyo cualquier negociación bilateral que se encare.[9]

El Acuerdo firmado supone el primero que la UE suscribe con un bloque económico con el que no comparte frontera física. El acuerdo supone un cambio muy importante en la relación con América Latina, ya que, por un lado, fija como objetivo una gradual liberalización recíproca del comercio y, por otro, establece una relación directa con el MERCOSUR, mejorando las

[4] El Mercado Común Centroamericano (MCCA) fue suscrito en 1960 por Costa Rica, El Salvador, Guatemala, Honduras y Nicaragua. El CARICOM es un acuerdo entre antiguas colonias británicas, hoy convertidas la mayoría en paraísos fiscales, que impulsaron un tímido proyecto de mercado común entre sus miembros: Barbados, Guyana, Jamaica, Trinidad y Tobago, Antigua y Barbuda, Bahamas, Belice, Dominica, Granada, Montserrat, Saint Kitts y Nevis, Santa Lucía, San Vicente y las Granadinas, y Surinam.

[5] El conjunto de los países que hoy integran esta asociación (Argentina, Bolivia, Brasil, Chile, Colombia, Cuba, Ecuador, Venezuela) representan territorialmente una superficie de 20,4 millones de kilómetros cuadrados que aglutina a más de 455 millones de habitantes.

[6] Véase Doval Adán (2008) para una referencia rápida sobre cada uno de estos acuerdos y otros de muy reciente creación.

[7] En los últimos años se han incorporado, en calidad de Estados asociados, Bolivia, Chile, Venezuela, Ecuador y Perú. Véase el sitio *www.mercosur.int* por detalles adicionales sobre la situación actual de cada uno de los países integrantes del acuerdo.

[8] Entre otros posibles factores, la recesión económica que golpeó a la región la mayor parte del período 1999-2002 fue una de las más profundas desde la creación del MERCOSUR y la primera en afectar simultáneamente a los cuatro países miembros. La crisis financiera por la que atravesó desvió su atención del esquema de convergencia macroeconómica e impidió alcanzar algunas metas comunes acordadas.

[9] Entre los aspectos contemplados en el acuerdo se encuentran la mejora del acceso a los mercados, la identificación de productos sensibles y prioritarios, la cooperación en materia de prácticas restrictivas de la competencia, la homogeneización de normas de origen, la cooperación en normas agroalimentarias e industriales, la cooperación aduanera, la cooperación en propiedad intelectual, la cooperación en prestación de servicios, la cooperación empresarial, el fomento de inversiones y la cooperación económica (Balaguer Franch y Martínez Zarzoso, op. cit.; p. 120).

relaciones políticas entre las partes. Con este acuerdo se da un paso fundamental para favorecer el comercio exterior, ya que el MERCOSUR, además de contar con un elevado porcentaje del producto bruto de América Latina, tal cual lo señalado, goza de la mayor concentración tecnológica y financiera del sur de los Estados Unidos.[10]

El Acuerdo debe ser considerado como una vía facilitadora del comercio, en particular, de la exportación latinoamericana a la UE, y deja atrás las opciones autárquicas anteriores. Es, acaso, el primer paso importante en la cooperación entre la UE y el MERCOSUR, convirtiéndose en un factor positivo para el fortalecimiento de las relaciones comerciales.

Este contexto de diálogos entre el MERCOSUR y la UE propicia un incremento del flujo comercial. El MERCOSUR vende principalmente productos básicos y un elevado porcentaje de sus exportaciones están formadas por café, tabaco, carne, pescado, frutas, cuero, cereales y aceites; importando, entre otros, maquinarias y equipos, automóviles y productos químicos.

En este marco, España dispone de importantes ventajas competitivas en comparación con sus socios comunitarios, su sector financiero tiene un rol protagónico en las relaciones comerciales con América Latina (que desempeña un efecto derrame hacia otros sectores), y las empresas españolas han destinado un volumen importante de inversiones en el conjunto de países que integran el continente americano.[11] Por todo, las relaciones España-MERCOSUR son importantes.

3. FLUJO COMERCIAL ESPAÑOL CON DIFERENTES ÁREAS GEOGRÁFICAS: EL CASO DE AMÉRICA

La economía española ha mostrado un grado interesante de dinamismo que se ha visto reflejado en su comercio exterior, más allá del comportamiento neto de sus cuentas externas. Entre 1999 y 2008 los flujos comerciales totales entre España y el resto del mundo pasaron de 243.882,3 millones de euros a 470.435,7 millones de la misma moneda, a la vez que se modificó notablemente el patrón geográfico de su comercio exterior.

La Tabla 1 muestra el flujo comercial bilateral entre España y el resto del mundo por área geográfica, tomando la suma de exportaciones e importaciones dividida por el comercio total entre ambos (España y el área geográfica correspondiente), lo que permite observar el grado de concentración geográfica del comercio exterior.[12]

Tabla 1
Flujo Comercial Bilateral entre España y el Resto del Mundo por Área Geográfica
(en % sobre el flujo bilateral total)

Área	Años									
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
UE	71,2	68,6	69,7	69,9	70,2	68,8	65,7	63,7	64,2	60,5
A. del Norte	5,4	5,5	4,9	4,7	4,3	4,2	4,0	4,1	4,1	4,5
A. Latina	4,5	4,6	4,7	4,4	4,0	3,9	4,4	4,7	4,7	4,9
Asia	9,7	10,3	9,8	10,0	10,0	10,8	11,7	12,4	12,7	13,7
África	4,4	5,6	5,7	5,3	5,4	5,5	6,3	6,5	6,2	7,8
Otras(*)	4,4	5,1	5,0	5,4	5,6	6,4	7,3	7,6	7,8	8,3
Total (**)	243.882,3	293.645,4	302.981,1	308.535,6	323.232,7	355.335,4	387.959,2	433.125,8	470.061,5	470.435,7

Nota: (*) Incluye el resto de Europa y America, Oceania y destinos sin determinar; (**) En millones de euros.
Fuente: Elaboración propia en base a datos del Ministerio de Industria, Turismo y Comercio de España, Secretaría de Estado de Comercio

[10] Claro que para lograr maximizar las posibilidades el MERCOSUR debe afrontar algunos desafíos. Entre ellos la profundización del mercado regional a través de la coordinación macroeconómica, la armonización de las políticas nacionales y la solución de los problemas en los que se encuentran las administraciones aduaneras de los estados parte, ya que esto es fundamental para el desarrollo del comercio exterior.

[11] Esto se ha producido en coincidencia con el fortalecimiento de la estructura empresarial española y la progresiva privatización de los sectores productivos de los países iberoamericanos.

[12] Alternativamente, se hubiera podido tomar un indicador de intensidad relativa del comercio exterior agregado y establecer una clasificación de las principales áreas con los que España comercia teniendo en cuenta el peso de cada uno de los países en el comercio mundial. Se ha optado, en cambio, por una medida más sencilla.

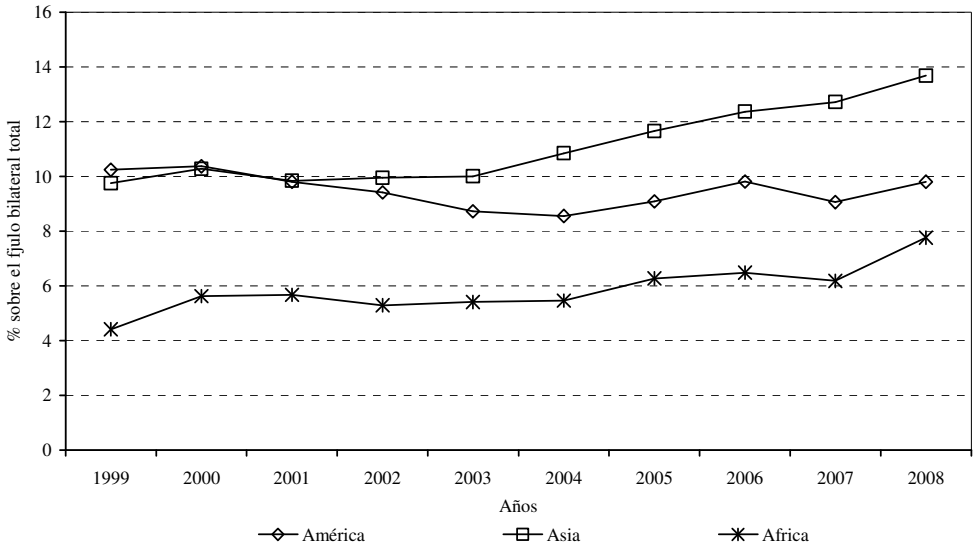
El flujo comercial bilateral entre España y la UE ha descendido casi 10 puntos porcentuales durante el período analizado, pasando de representar 71,2% a 60,5% del flujo total de bienes intercambiados. Como cabría esperar, este descenso ha traído consigo un incremento en la participación de otras áreas geográficas, tal el caso de Asia (que pasó de representar 9,7% a 13,7% del total intercambiado por España), África (de 4,4% a 7,8%) y el de Otras áreas geográficas, cuyo su notable incremento se debe a un flujo comercial muy importante entre España y países europeos que no se encuentran incorporados en la UE.[13]

Con relación a América, el flujo bilateral no ha mostrado un incremento significativo y, afortunadamente, tampoco un descenso, pero pone de manifiesto que si se desean fortalecer las relaciones entre ambos, España y América, los intercambios comerciales deberían robustecerse.

De manera más específica, el Gráfico 1 muestra el flujo comercial bilateral entre España y diferentes áreas con excepción de Europa y Otras. Tal como se aprecia, en 1999 América tenía todavía un mayor comercio bilateral con España en relación con Asia y África, pero a partir del año 2002 el comercio con Asia se intensifica, ganándole a aquél en importancia. Si bien el flujo con América parece recuperarse hacia 2004, la participación de América en el intercambio con España es hoy igual que hace diez años, próxima a 10%. Esta tendencia relativamente estable de los flujos comerciales intercambiados con España para el caso americano contrasta no sólo con la tendencia marcadamente ascendente de Asia sino también con la de África.

En suma, se podría señalar que para España el continente americano importa en sus relaciones comerciales, pero, observando la participación y la tendencia de los intercambios, primero importan otros destinos. Así, dejando de lado el flujo de comercio entre España y Europa, América se encuentra a mitad de camino entre Asia y África.

Gráfico 1
Flujo comercial bilateral entre España y otras áreas geográficas
(excepto Europa)



Fuente: Elaboración propia en base a datos del Ministerio de Industria, Turismo y Comercio de España

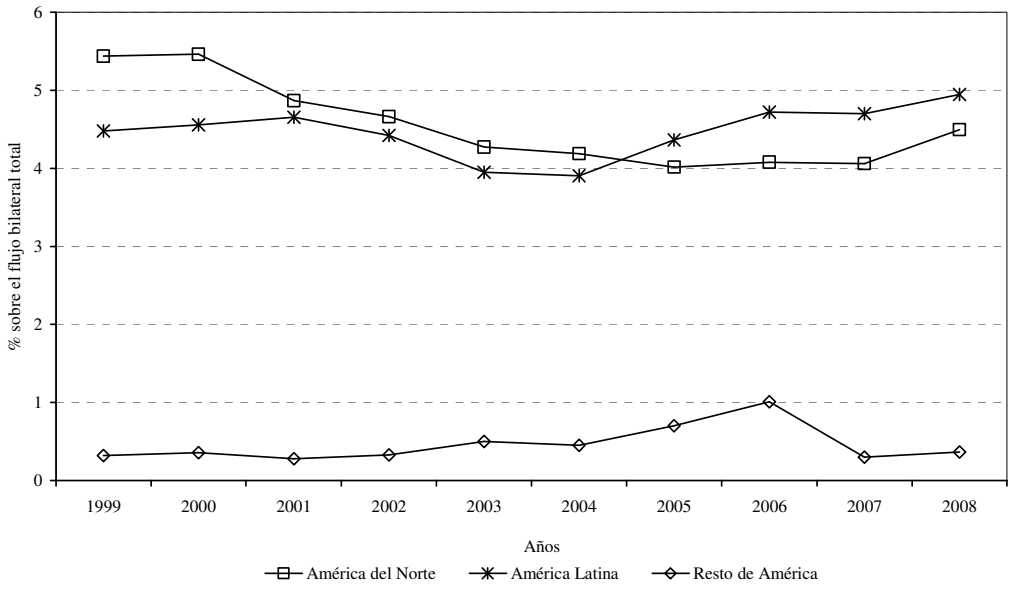
El Gráfico 2 recoge la evolución del flujo comercial entre España y América, diferenciando el que corresponde a América del Norte, América Latina y al resto de América.[14] Como se observa, al comenzar el período, el flujo bilateral de comercio con América del Norte fue superior.

[13] Otras áreas están integradas por el resto de Europa (países que no pertenecen a la zona del euro y a la UE) y Oceanía, junto a países que no se encuentran en América del Norte o América Latina .
[14] Los países que comprenden América del Norte son Estados Unidos, Canadá, Groenlandia y San Pedro y Miquelón.

Los flujos intercambiados con el territorio americano descienden hacia el año 2001 y hasta 2004, mientras que la diferente participación de América del Norte y América Latina tiende a achicarse. En 2005 el intercambio comercial se recupera y, por primera vez en el período, América Latina supera a América del Norte, al menos hasta 2008, último dato disponible. Los flujos intercambiados con España y el resto de América, crecen hasta 2006 (de manera significativa entre 2004 y 2006), para descender notablemente en 2007 y permanecer relativamente estables en 2008. Cabe destacar que este intercambio con el resto involucra a economías más pequeñas, lo que justifica su escasa relevancia sobre el total comercializado con España, llegando a medio punto porcentual del total de intercambios con América.

En suma, a pesar de que los intercambios han permanecido estables durante los últimos 10 años, aquellos correspondientes a América Latina han incrementado ligeramente su participación, aunque no superan el 5% del flujo bilateral total de España con el territorio americano.

Gráfico 2
Flujo comercial bilateral total entre España y América



Fuente: Elaboración propia en base a datos del Ministerio de Industria, Turismo y Comercio de España

4. FLUJOS COMERCIALES BILATERAL ESPAÑA-MERCOSUR

De particular interés resulta el flujo comercial bilateral entre España y el MERCOSUR, especialmente con cada uno de los países que integran este mercado común.[15] La Tabla 2 permite apreciar el flujo comercial bilateral para el período. Mientras en 1999 se intercambiaban 4.821,5 millones de euros, en 2008 esa cifra alcanza 8.274,8 millones.

Como se puede apreciar, Argentina ha ido cediendo participación en el comercio con España, ganándola Brasil de manera significativa, que ha visto incrementado su comercio en el período en una tasa importante, alcanzando en 2008 una participación en el total del comercio España-MERCOSUR próxima al 60%. La participación en el intercambio de los restantes miembros del MERCOSUR (Paraguay y Uruguay) es pequeña, pues, en conjunto, sólo alcanza 5,6% del total intercambiado.

[15] Como la integración de los Estados asociados al MERCOSUR es en grado variable —a diferencia de los que se van incorporando a la UE— se ha preferido en este estudio, para el caso del MERCOSUR, y sin pérdida de generalidad alguna, conservar los cuatro Estados fundacionales para llevar a cabo el análisis.

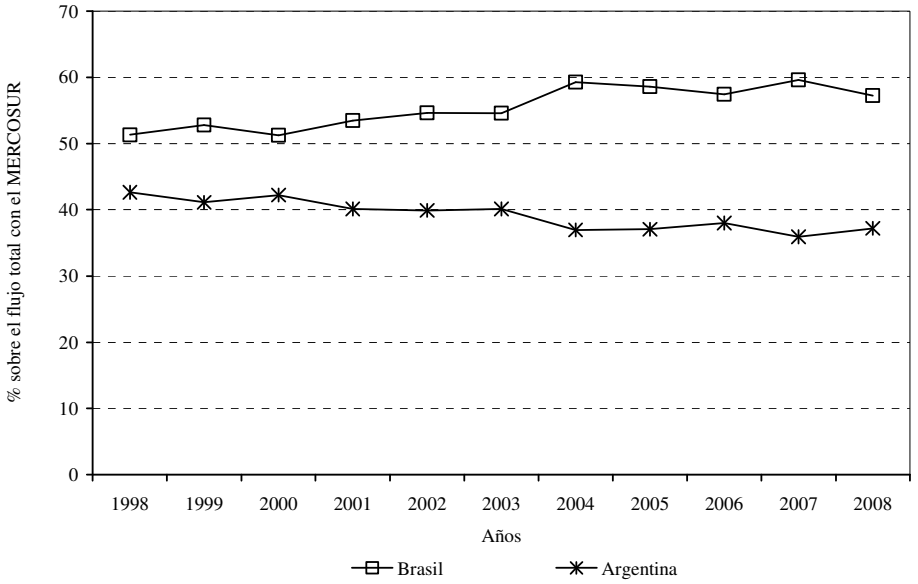
Flujo Comercial Bilateral entre España y los países del MERCOSUR
(en % sobre el flujo bilateral total)

Países	Años										
	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Argentina	42,6	41,1	42,2	40,1	39,9	40,1	37,0	37,1	38,0	35,9	37,2
Brasil	51,3	52,8	51,2	53,5	54,7	54,6	59,3	58,6	57,4	59,6	57,3
Paraguay	1,4	1,5	1,7	1,6	0,9	1,6	0,5	0,5	1,0	1,4	2,2
Uruguay	4,7	4,6	4,8	4,8	4,5	3,6	3,3	3,8	3,5	3,1	3,4
Total (*)	4.650,42	4.821,46	5.036,61	5.210,07	4.400,31	4.678,73	5.057,74	5.277,19	5.723,64	7.312,52	8.274,75

Nota: (*) En millones de euros.
Fuente: Ministerio de Industria, Turismo y Comercio de España, Secretaría de Estado de Comercio

El Gráfico 3 presenta el flujo comercial bilateral entre España y los dos socios principales del MERCOSUR, Argentina y Brasil, pudiéndose apreciar, tal cual lo manifestado, una clara tendencia a la baja en la participación del intercambio con Argentina y un ascenso en el caso de Brasil, ampliándose cada vez más la brecha entre el intercambio de cada uno con España.

Gráfico 3
Flujo comercial bilateral España-MERCOSUR por pricipales países



Fuente: Elabroación propia en base a datos del Minsterio de Industria, Turismo y Comercio de España

Los Gráficos 4 y 5 muestran, respectivamente, las exportaciones e importaciones españolas hacia y desde el MECOSUR por país.[16]

En el caso de las exportaciones españolas, éstas han registrado una tendencia e incremento notables hacia Brasil tomando en consideración el total exportado al conjunto de los cuatro países integrantes del bloque, llegando en la actualidad a representar casi 70% del total exportado a la región sudamericana. No sucede lo mismo en el caso de las exportaciones hacia Argentina, cuya tendencia ha sido claramente descendente en el período de análisis, que condujo a una participación inferior a 30% en 2008, ni tampoco con el caso de las dos economías más pequeñas, Paraguay y Uruguay, cuya participación ha permanecido estable. De modo tal que si hay algún interés de parte de España en comerciar con el MERCOSUR, al menos en lo que a exportaciones se refiere, este interés se exterioriza más claramente con Brasil.

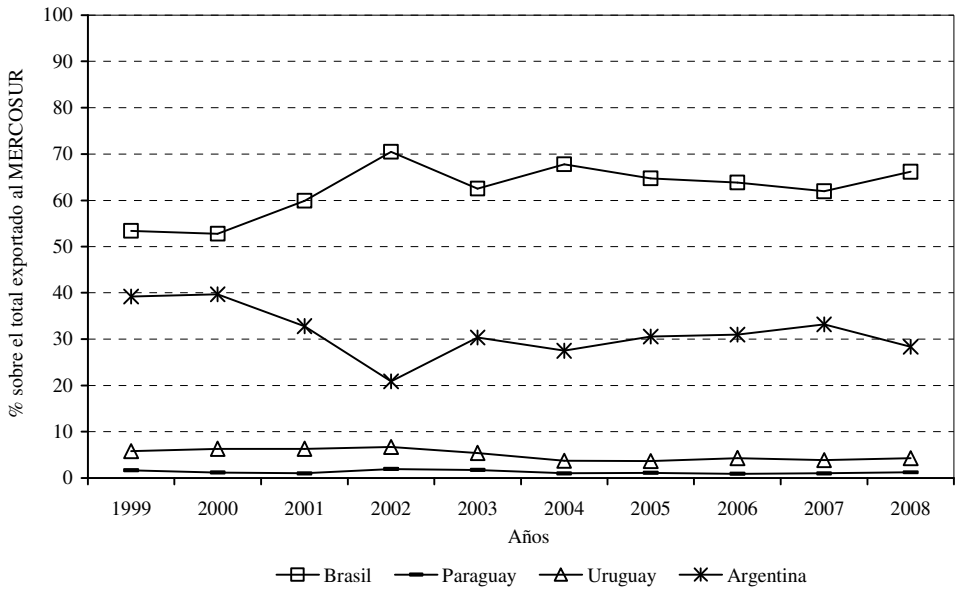
[16] Si bien en este trabajo se ha efectuado una descripción de los flujos totales intercambiados, se diferencia aquí entre exportaciones e importaciones por considerarlo de particular interés. Véase Jacobo (2009) por una distinción entre exportaciones e importaciones españolas hacia y desde otras áreas geográficas diferentes al MERCOSUR.

En el caso de las importaciones, no se manifiesta una tendencia clara ascendente o descendente con ninguno de los países. Antes bien, los números analizados parecen sugerir una situación de relativa estabilidad tanto en tendencia como en participación de los dos socios principales. En el caso de los dos más pequeños, la participación ha permanecido, al igual que el caso de las exportaciones, relativamente estables.

Una observación final es que, en uno y otro caso, es decir, exportaciones e importaciones, un descenso o ascenso en la participación de uno de los dos principales socios del MERCOSUR parece compensarse por un ascenso o descenso respectivamente de la participación del otro socio.

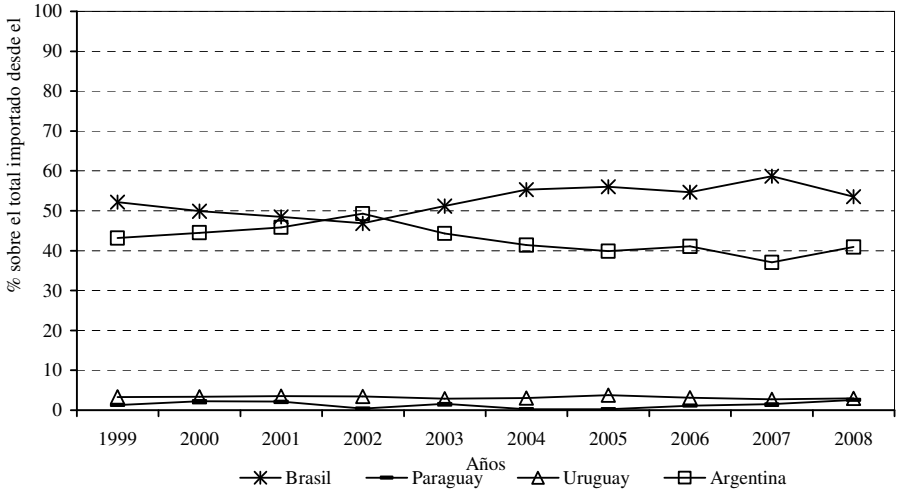
Un punto no menor es que se trata aquí de una descripción sencilla de lo que sucede con los flujos comerciales; descripción que no involucra en modo alguno un análisis de las causas que motivan un cambio en el patrón geográfico de los intercambios comerciales observados.

Gráfico 4
Exportaciones españolas al MERCOSUR por país de destino



Fuente: Elaboración propia en base a datos del Ministerio de Industria, Turismo y Comercio de España

Gráfico 5
Importaciones españolas desde el MERCOSUR por país de origen



Fuente: Elaboración propia en base a datos del Minsterio de Industria, Turismo y Comercio de España

5. COMENTARIOS FINALES

En la actualidad, la relación entre el MERCOSUR y la UE está basada en el Acuerdo Marco firmado en 1995, cimentado en tres pilares, uno de los cuales, el más importante, es el pilar comercial que, de profundizarse, significaría un paso trascendente en la conformación del segundo bloque comercial del planeta, con todas las implicancias del caso.

Sin embargo, la prosperidad y el éxito de cualquier asociación económica de este tipo requiere forzosamente del conocimiento y desarrollo de los flujos comerciales bilaterales entre países que integran uno y otro bloque. Es precisamente en este sentido que el explorar el intercambio entre los países resulta relevante, particularmente en el caso de España, ya que posee ventajas competitivas en comparación con sus socios comunitarios, su sector financiero ha tenido y tiene un papel protagónico en las relaciones comerciales con América Latina, en especial el MERCOSUR, y sus empresas destinaron un importante volumen de inversión en estos países sudamericanos.

Más allá de la buena voluntad por parte de todos los gobiernos para estrechar vínculos, es factible observar que las relaciones comerciales entre España y el MERCOSUR han permanecido muy tranquilas durante los últimos diez años, lo que significa que el comercio entre ellos no ha crecido; no al menos como ha ocurrido en otras regiones y con otros países. Esto sin duda no es auspicioso para lograr los avances que se pretenden y debería revertirse.

A pesar de que esta circunstancia, el escaso comercio mutuo, puede ser una compartida por otros países de la UE, pone de manifiesto una señal de alerta hacia el futuro, desincentivando una relación entre bloques que, no por curiosa coincidencia, hoy no parece prosperar. No por mera casualidad, en el fortalecimiento de los lazos comerciales se encuentra el destino manifiesto que motiva esta singular asociación.

REFERENCIAS

ALADI (2009). *Asociación Latinoamericana de Integración*, http://www.aladi.org.

ALCA (2009). *Área de Libre Comercio de las Américas*, http://www.ftaa-alca.org.

Balaguer Franch, María e Inmaculada Martínez Zarzoso (2000). "Análisis de los flujos comerciales Unión Europea- MERCOSUR", *Información Comercial Española*, número 788, p. 119-132.

CAN (2009). *Comunidad Andina de Naciones*, http://www.can.org.

DOVAL ADÁN, ANTONIO (2008). "Integración regional e inversiones españolas en América", *Boletín Económico del ICE*, número 2934.

IGLESIAS, E. (2005). "Cuatro décadas de Integración Regional en América Latina y el Caribe", conferencia pronunciada en el seminario Los 40 años de la creación del INTAL, Buenos Aires, agosto, *manuscrito*.

MERCOSUR (2009). *Mercado Común del Sur*, http://www.mercosur.int

MINISTERIO DE INDUSTRIA, TURISMO Y COMERCIO DE ESPAÑA (2008). *Informe Trimestral de Comercio Exterior*, varios números.

JACOBO, ALEJANDRO (2009). "Una aproximación a la caracterización de los flujos comerciales MERCOSUR-UE", *manuscrito*.

JACOBO, ALEJANDRO (2008). "Una aproximación a las características de los intercambios comerciales entre el MERCOSUR y la UE", en Jesús Tricás Preckler y Carlos Moslares García (Coordinadores): *Anales de Economía Aplicada*, Barcelona, pág. 3-13.

JACOBO, ALEJANDRO (2007). "Perspectivas de la integración económica en América Latina a comienzos del siglo XXI", en Héctor Gertel, Ernesto Rezk y Alejandro Jacobo (Editores): *Economía Pública de Argentina y España,* Fondo Editorial Consejo, Buenos Aires, p. 95-124.

JACOBO, ALEJANDRO (2004a). "La Argentina frente al Acuerdo de Libre Comercio de las Américas", *Actualidad Económica*, 55, pp. 5-11.

JACOBO, A. (2004b). "La Disyuntiva entre ALCA y MERCOSUR", *Comercio y Justicia*, mayo 31, p. 7.

NAFTA (2009). *Tratado de Libre Comercio de América del Norte*, http://www.nafta-sec-alena.org.

PAIVA, P. y R. GAZEL (2003). "MERCOSUR: Past, present and future", *Nova Economia*, volumen 13, número 2, pp. 115- 139.

TORRENT, RAMÓN (2005). "Las relaciones Unión Europea-América Latina en los últimos diez años: El resultado de la inexistencia de una política. Un análisis empírico y esperanzado", Observatorio de las Relaciones Unión Europea-América Latina, Barcelona, *manuscrito*.

ÁREA X

ECONOMÍA URBANA, REGIONAL Y LOCAL

ECONOMIA URBANA, REGIONAL E LOCAL

HOW EFFECTIVE ARE PUBLIC POLICIES FOR REGIONAL DEVELOPMENT? THE PRIME IN PORTUGAL

PURIFICACIÓN VICENTE GALINDO
E-mail: purivic@yahoo.com
Department of Statistics
University of Salamanca, Spain

TERESA DE NORONHA VAZ
E-mail: mtvaz@ualg.pt
Faculty of Economics and CIEO- Research Center for Spacial and Orga nizational Dynamics
University of Algarve, Portugal

Resumen

The major goal of this work is to evaluate the level of effectiveness of those efforts made by the European Commission towards the socio-economic cohesion in Europe by using the single case-study of the application of PRIME in Portugal. PRIME is one of the best example support system because it has been indirectly financed by European founds while fully structured as a Portuguese policy instrument for support to the modernization of the entrepreneurial tissue.
Further than the goal earlier pointed out, this paper has a dual aim: firstly, to supply an exploratory survey on the patterns of asymmetric use of the financing systems provided by the E.C. and, secondly, to emphasize the north/south endemic dichotomy existing in the country – a path dependence trend that should concern Portuguese policy makers deeper then it does.
Methodologically, we have observed the financing execution of 14910 projects, collected from the official information portal of PRIME in Portugal. Due to limitations related to the information sources of the portal our study deals with a reduced number of variables: project location, year, industrial activity, financial program, NUT, district, investment and incentive. Nevertheless, the application of HOMALS, Cluster and Correspondence Analyses was possible, allowing conclusions about the level of effectiveness of PRIME across, both, location and sectors levels.

Palabras Clave: Public Policy, Regional development, Portugal, Cluster Analyses.
Área Temática: Economía Urbana, regional y local.

1. INTRODUCTION AND THEORETICAL FRAMING

History has made from any spatial context a living organism with memory, action and reactive capacity. We could argue that there is an intelligent attitude in each territory to absorb external inputs or supports in a continuous process of change. Some call it "path dependency", others the historical context. This simple, but very complex, fact may determine the dynamics of local development. Due to this complexity, frequently, the conditions for growth are restrictive and, in spite of the multiple public policies, regions keep gapping. Most of the world regions could serve as an example, but we shall observe the particular case of Portugal. Its peripheral location in southern Europe, its distance from the dominant global trade flows, as well as its lack of strategically located mega-towns should justify a major interest in analysing how the country used the great quantity of European financial supports that resulted from its integration as a full member of the European Union, twenty years ago.

Thus, the major goal of this paper is to evaluate the level of effectiveness of those efforts made by the European Commission towards the socio-economic cohesion in Europe by using

the single case-study of the application of PRIME in Portugal. PRIME is one of the best example support system because it has been indirectly financed by European founds while fully structured as a Portuguese policy instrument for support to the modernization of the entrepreneurial tissue.

Further than the goal earlier pointed out, this paper has a dual aim: firstly, to supply an exploratory survey on the patterns of asymmetric use of the financing systems provided by the E.C. and, secondly, to emphasize the north/south endemic dichotomy existing in the country – a path dependence trend that should concern Portuguese policy makers deeper then it does.

Literature indicates that multiple major driving forces of economic growth have excited economists during the last fifty years. Firstly, Harrod (1939) and Domar (1946) emphasized savings, population and capital as the exogenous determinants responsible for the long-term growth of the economy. Later, Solow (1956) proved that the increases in national savings could, indeed create per capita real income but stressed that continuous growth would only be possible if technological progress would occur. Further contributions to long run growth were made by Arrow (1962) by introducing "learning-by-doing" as a determinant of technological development, Lucas (1988) by including the growth rate of human capital as a factor of technical change and long run growth, and Romer (1986 and 1990) who demonstrated the role of research training in the technological change. But, for a long time the impact of growth as motor of asymmetric or non-asymmetric development was not a source of concern. However, when further studies related to the spill-over effects resulting from growth were discussed (Marshall-Arrow-Romer model by Acs and Audretsch, 1984 and Acs, 2002) local impacts took relevance. Namely, as the output of technological innovation seems to result from knowledge inputs (technology and/or entrepreneurship) then, both, local and regional dimensions are crucial (Fagerberg, 2003, 2004). The most updated growth theories accept the importance of micro and meso-economic levels of activities if growth is supposed to stay sustainable, pulling into entrepreneurship the major drivers of growth (Sogaard, 2009).

In spite of the many authors concerned with growth, development or sustainability, the particular case of those regions tending to a permanent gap in relation to others, more favoured ones, has not attracted much attention. However, political will and democratic equity justify the fairness of symmetric development among different regions and furthering the use of equal opportunities. Therefore, we can accept that there is a global goal that surpasses the historical tendencies and challenges those regions gapping to be supported in finding those corrected choices to push them away from a spiralled movement of atrophy. For such regions target public policy is a must.

In spite of the different directions taken by policy makers, most of the public strategies for development, and corresponding policy proposals, are based in instruments able to promote increases in private investment, employment, regional production and knowledge spillover effects. It is trusted that these combined factors may generate growth and development, if possible sustainable, in gapping regions and there is hope that Europe will find the best solution for socio-economic convergence and regional cohesion.

But cohesion has not been the sole goal of European decision makers: Lisbon Agenda, in March 2000, defined the goal of making EU the strongest and most competitive economy by 2010. So that it can be understood that the need to increase regional capacities for global competition has defined a clear European wish much further from its scientific and technological potential skills. An overwhelming goal, challenging the recent European development trends, may include a major paradox between science and technology and regional and social cohesion: In fact, the application of the R & D policy brings to Europe the preference to those regions able to produce high-tech industries; Also, as the cohesion goals have lead to a considerable aid to scientific research, the relationship between this, the resulting technological advances and the regional development for periphery are still to be launched.

Some quite generic discussions have emerged from research agendas concerned with the notion of geographical space and concluding that its development is a matter of maturity of endogenous systems (Vazquez-Barquero, 2002). The introduction of time as a major factor of the territorial dynamics stresses a whole set of new drivers into the scene of policy making. Concepts such as "learning" and "interacting", give origin to very consistent and original new

approaches in which institutions, public, private or in public-private partnership became regulatory elements of individual and collective practices, the principal assets being therefore of a relational nature (Sierra, 1997). One of the main characteristics inherent to the territory is the proximity among socio-economic agents (Keeble, and Wilkinson (1999). But, whether or not the agents are able to benefit from this attribute depends on their own technological and social capacity to interact and co-operate – the concept of social capital considered as a consistent argument to favour development (Solow, 2000 and Durlauf, 2002).

Public policies should be able to provide the set of factors necessary to speed development, particularly in less favoured regions. Such regions face three types of hindrances to be surmounted: the location disadvantages, the lack of structural adjustments and need for political coherency at national, regional and sector levels (Vaz and Neto, 2005).

Most frequently public policies for regional development, implemented to increase the competitiveness of regional skills, come across targets for technological modernization, at national level, that strongly prejudice policy results. So, how far are the outcomes from the initial targets? The exact answers serve for evaluation of policy effectiveness of regional policy and should be provided by in deep monitoring of implemented programmes.

The next point of this paper follows the previous arguments using the case-study of Portuguese program PRIME.

2. TOOLS FOR DEVELOPMENT IN EUROPE: THE PORTUGUESE PROGRAM PRIME

PRIME[1] - Incentives Programme for the Modernisation of Economic Activities was instituted by the Resolution of the Council of Ministers no. 101/2003, after its formal approval by the European Commission in 2003. The program integrated a set of medium term economic policy instruments for the period 2000-2006 which have been applied to different economic activity sectors in Portugal.

The instrument was developed exclusively for the Industry, Energy, Building Industry, Transport, Tourism, and Trade and Services sectors. The Programme integrated in Area 2 of the Regional Development Plan – *Adjusting the Production Profile towards the Activities of the Future*, and was designed to foster productivity and competitiveness of Portuguese firms, thereby promoting new development capabilities under the following goals: to ensure a generally sound competition environment, which may contribute to revitalise business structure, and to steam in firms research and development on new products and processes; to simplify procedures and stimulate de-regulation; to encourage the qualification of human resources by enhancing technical and scientific training thereby fostering productivity growth, through more motivating and rewarding tasks.

As such, PRIME was promoted along the following priority areas of strategic actions: Area 1 – Enterprise Stimulation; Area 2 – Qualification of Human Resources; Area 3 – Stimulation of the Business Environment. The two first Areas are dedicated to a direct intervention at firm level, and the third, of a more crosscutting nature, aims at stimulating business environment. The 29 support measures and incentives schemes at which the programme is structured are presented in the Annex 1.

The financial supports corresponding to PRIME Programme were non refundable incentive types and each support measure or incentive scheme aimed to generate new economic activities and to introduce a new competitive and innovative context to the Portuguese economy. The financial support approach of this Programme also included a risk capital fund and achievement bonus incentives.

PRIME was a public policy instrument driven to the promotion of economic activity in the

[1] More information related to PRIME can be found in http://www.prime.min-economia.pt/presentationlayer/prime_Home_00.aspx

Portuguese territory specially focused, but not exclusively, on small and medium size enterprises. Some of Programme support measures and incentives included as beneficiaries non profit organisations, municipalities, technological schools, business associative structures and technology-based incubator centres and parks.

As it was supporting a wide range of economic activity sectors the instrument was neither sector nor regional focused. The only exceptions were the URBCOM, MINING ZONES and HISTORIC INNS incentives schemes. As showed in the summary presented on Annex 2, the program did not use, as project selection criteria, the geographical localisation of the project proponent. Just in a few cases localization was a selection criterion: NITEC, SIED, SIPIE, Vocational Training, PITER and TTQ Infrastructures.

3. METHODOLOGY FOR DATA ANALYSIS

3.1. Sampling and choice of variables

Our initial database contained 15307 records reported to 2007 data set. After a filter application based on NUTII clear identification the final database contains 14910 enterprises that received incentives during the period 2000-2006.

Following variables have been defined as active variables: i) Date, ii) Support, iii) Area of strategic action (EIXO), iv) Region (Nuts II) (North, Centre, Lisbon, Alentejo, Algarve, Açores and Madeira), v) District (Aveiro, Baja, Braga, Bragança, Castelo Branco, Coimbra, Évora, Faro, Guarda, Leirea, Lisboa, Portalegre, Porto, Santarém, Setúbal, Viana do Castelo, Viseu, Vila Real), vi) CAE, vii) Investment, viii) Incentive and ix) Ratio (between Invesment and Incentive).

The categorical variables have been recoded to be suitable for the different statistical analyses. The date has been recoded to have the information per year and the variable CAE to capture the information of the first two digits.

Because more than the 50% of the incentives are directed to a few activity branches and in order to simplify the graphical and analytical interpretation of the results, our study is limited to the more frequent categories, grouping the rest under the label “Other”.

The variable Support has been recoded keeping the categories that contain more than 2% of the enterprises and grouping the rest as “other”. For the districts we have kept those that have more than 5% of the Projects.

Investment, Incentive and Ratio were considered as numerical variables, but taking into account the high dispersion, we have categorized them for some statistical analysis using the quartiles of its frequency distribution. See details in Annex 3.

3.2. Statistical methods

For each categorical variable the frequency distribution and an associated chart have been calculated. The association between each pair of categorical variables is tested using the chi-squared test. The association is considered statistically significant when the p-value is lower than 0.05. For each quantitative variable the mean, standard deviation, median and inter-quartile range are calculated. The last two statistics are usually preferred when a high dispersion is present in the data. To compare the means the independent-samples t test has been used and to compare medians the Mann-Whitney test.

3.2.1. Homogeneity analysis (HOMALS)

Homogeneity Analysis is a interdependence statistical technique for graphic interpretation in several axis (Gifi, 1990). It converts categorical data into a graphical display by simultaneously quantifying the categorical variables and the projects; it captures the complex relations between both sets. We use HOMALS because the variables can have non-linear relations and because it is possible to find relations between categories, not only between variables. A three-

dimensional Homals solution was computed for the set of nine variables (Year, Area, Nut, District, System, CAE, Investment, Incentive and Ratio) and 14910 projects, using SPSS.15. The Projects scores and category quantifications are mapped into the same (two or three dimensional) plot.

3.2.2 Cluster Analysis

We have used the numerical scores obtained from HOMALS, to create a K-means cluster analysis with the different projects.

3.2.3. Correspondence Analysis

In a third step, Correspondence Analysis is used to describe the relationships between the three clusters and each nominal variable, representing both in a low-dimensional space, showing the relationships between the categories for each variable and the different clusters. For each variable, the distance between two category points in the plot reflects the similarity between the categories profiles, with categories sharing similar profiles plotted close to each other. The relationship between the variables can be described projecting category points for one of the variables onto the vector from the origin to a category point for the other variable.

4. RESULTS

The following results are based on both univariate and multivariate analysis when adequate.

4.1 Projects typology

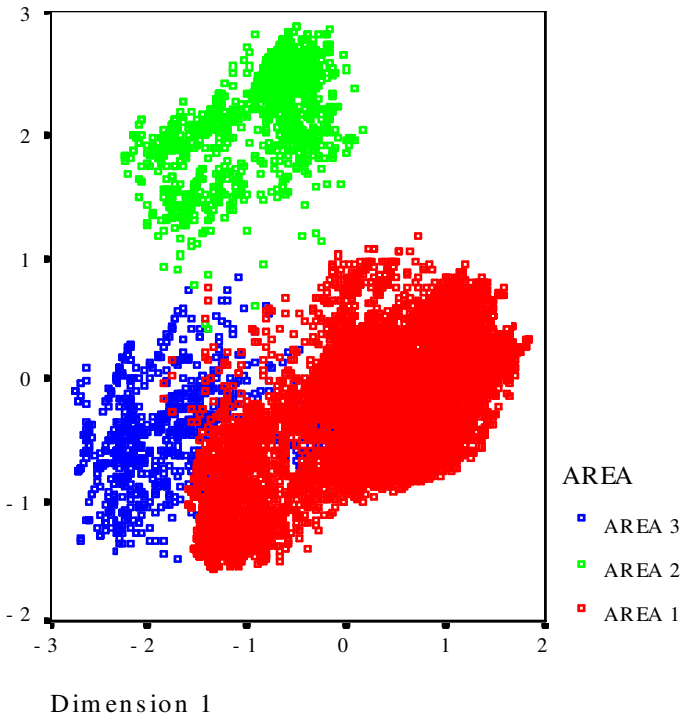
HOMALS´ solution

The HOMALS analysis selected three dimensions. The first dimension can be described as a dimension related to INVESTIMENT, INCENTICV and SUPPORTS. The second dimension was strongly dominated by AREA and the third dimension is related to EAC and SUPPORT, being these variables the most important in the analysis. See Table 1 below. HOMALS solution results in a scatter plot with a point representing each project. Setting the markers by Area clearly identifies three groups, each one related to an Area: the red one with the Area 1; the green one related to Area 2 and the cluster 3, in blue, related to Area 1 and Area 3.

Table 1. Discrimination measures per variable and per dimension

	Dim1	Dim2	Dim3
YEAR			
,402	,425	,275	
CAE			
,483	,004	**,509**	
AREA			
,335	**,723**	,393	
SUPPORT			
,760	**,868**	**,733**	
NUTSII			
,029	,024	,050	
DISTRIT			
,093	,043	,046	
INVEST			
,564	,157	,018	
INCENT			
,582	,081	,080	
RATIO			
,164	,373	,312	

Figure 1. HOMALS map: two first dimensions



A k-means cluster analysis with the Project scores of the first three HOMALS dimensions was performed. We selected K= 3 clusters taking into account the HOMALS map shown above. Table 2 shows project agglomeration into the three clusters.

Table 2. Distribution of the Projects on the three clusters

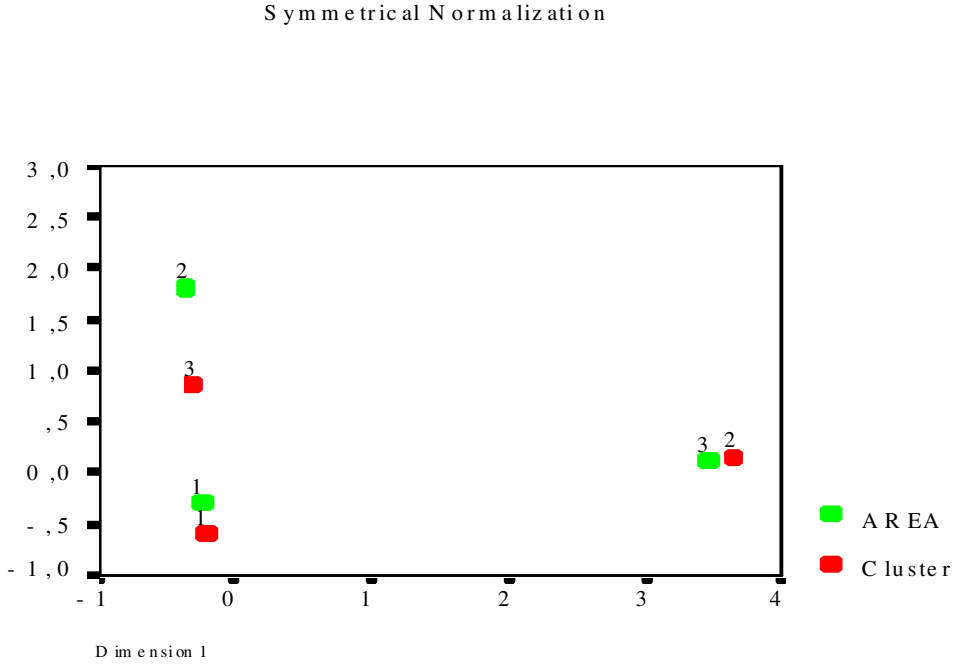
	Frequency	Percent
Cluster 1	8356	57.25
Cluster 2	900	6.03
Cluster 3	5654	37.92

Cluster from the HOMALS scores are closely related to the Areas. See Table 3 below. Most of the projects in the first cluster belong to Area 1, only 1.8% of the points belong to Area 2. Cluster 2 can be identified with Area 3 because only 10% belong to Area 1. Cluster 3 is a mixture of projects in Areas 1 and 2. More details can be seen below.

Table 3. Characterization of the clusters using the Area of strategies action

	CLUSTER1	CLUSTER 2	CLUSTER3	
AREA 1				
Count	8202	90	3759	12051
% within Cluster	98,2%	10,0%	66,5%	80,8%
AREA 2				
Count			1868	1868
% within Cluster			33,0%	12,5%
AREA 3				
Count	154	810	27	991
% within Cluster	1,8%	90,0%	0,5%	6,6%
Total				
Count	8356	900	5654	14910
	100%	100%	100%	100%

Figure 2. CA map for Cluster and Area
Model: Symmetrical normalization and distance Chi square



Correspondence Analysis for the cross-tabulation of clusters and areas, with 90% of the inertia accounted for the first two dimensions, highlights the previous conclusions in a more conclusive way. Cluster 1 can be identified with Area 1, Cluster 2 with Area 3 and Cluster 3 between Areas 1 and 2.

Table 4. Cross Tabulation for Clusters and Regions

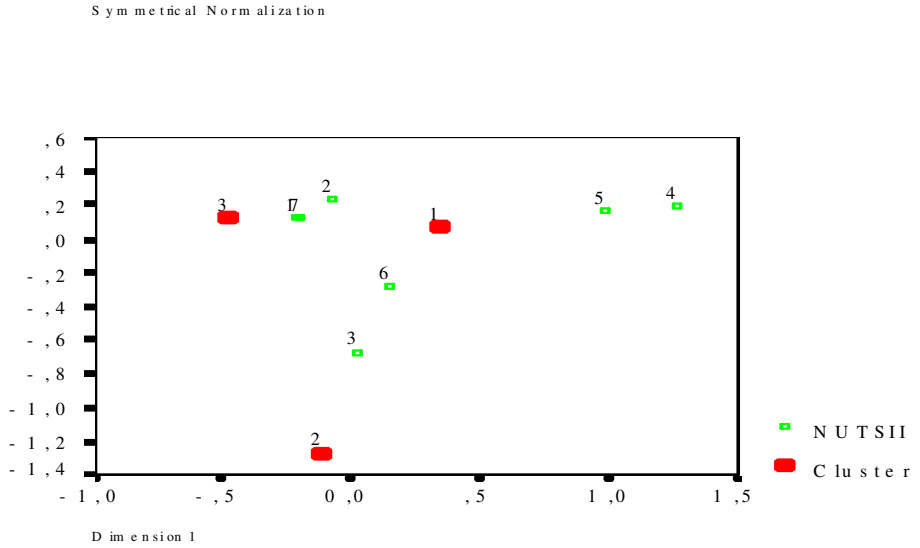
						Total
			CLUSTER 1	**CLUSTER 2**	**CLUSTER 3**	
	North	Count	3459	351	2792	6602
		% within Cluster	**41,4%**	**39,0%**	**49,4%**	44,3%
	Centre	Count	2081	163	1508	3752
		% within Cluster	**24,9%**	18,1%	**26,7%**	25,2%
	Lisbon	Count	1429	296	911	2636
		% within Cluster	17,1%	**32,9%**	16,1%	17,7%
	Alentejo	Count	729	34	147	910
		% within Cluster	8,7%	3,8%	2,6%	6,1%
	Algarve	Count	376	21	105	502
		% within Cluster	4,5%	2,3%	1,9%	**3,4%**
	Acores	Count	163	23	96	282
		% within Cluster	2,0%	2,6%	1,7%	**1,9%**
	Madeira	Count	119	12	95	226
		% within Cluster	1,4%	1,3%	1,7%	**1,5%**
Total		Count	8356	900	5654	14910
			100%	100%	100%	100%

4.2. Factor Analyses by location

CLUSTER- NUTS

The distribution of the clusters is not homogeneous among the regions ($p < 0.001$).

Figure 3. CA map for Cluster and NUTS
Model: Symmetrical Normalization and Distance Chi square.



Cluster 1, identified with Area 1, has more projects from the Centre and North regions.

More than 66% of the projects of this cluster belong to those regions. Lisbon is also important but with a lower percent. See table below. Cluster 2, identified in the previous step with Area 3 (Stimulation of the Business Environment), has a 72% of projects from the North and Lisbon. Cluster 3, identified with Area 2 (Qualification of Human Resources) has a 76% of projects from North and Central Portugal.

See more details on the Table 5, above. The correspondent CA for the cross-tabulation of clusters and regions highlights the previous conclusions.

CLUSTER- DISTRICT

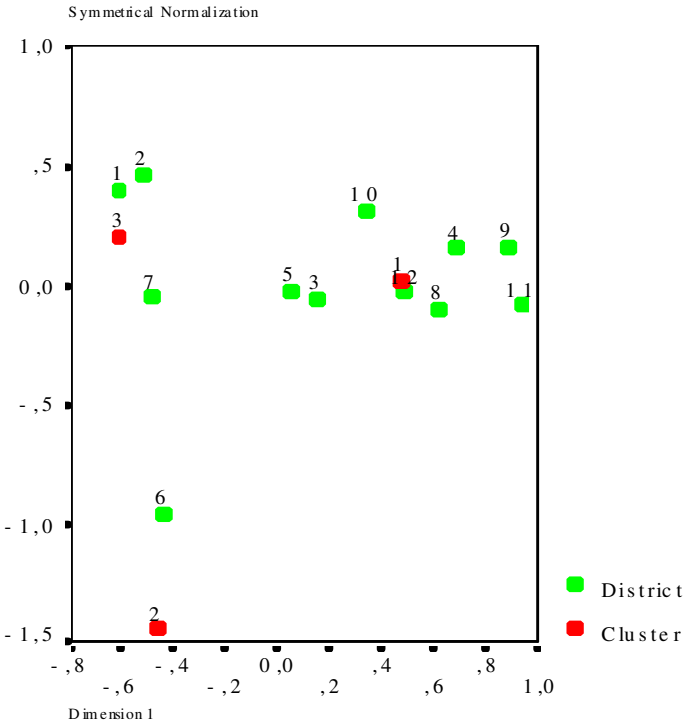
The cross-tabulation of the clusters and districts is shown in the next table. The distribution of the districts is not homogeneous among clusters ($p< 0.001$). From the table, we can state that Cluster 3 has projects from Aveiro, Braga and Porto, the Cluster 2 from Lisbon and Porto, and the Cluster 1 has a similar percentage from all the districts, see Table 5, below.

Table 5. Distribution of frequencies for Cluster and District

DISTRICT		CLUSTER 1	CLUSTER 2	CLUSTER 3	
Aveiro	Count	633	67	864	1564
D1	**% within Cluster**	**7,6%**	**7,4%**	**15,3%**	10,5%
Braga	Count	736	58	913	1707
D2	% within Cluster	**8,8%**	6,4%	**16,1%**	11,4%
Coimbra	Count	401	41	229	671
D3	% within Cluster	4,8%	4,6%	4,1%	4,5%
Faro	Count	362	14	113	489
D4	% within Cluster	4,3%	1,6%	2,0%	3,3%
Leiría	Count	512	54	325	891
D5	% within Cluster	6,1%	6,0%	5,7%	6,0%
Lisbon	Count	659	237	622	1518
D6	% within Cluster	**7,9%**	**26,3%**	**11,0%**	10,2%
Porto	Count	1073	192	1203	2468
D7	% within Cluster	**12,8%**	**21,3%**	**21,3%**	16,6%
Santarém	Count	389	28	123	540
D8	% within Cluster	4,7%	3,1%	2,2%	3,6%
Viana Castelo	Count	570	16	132	718
D9	% within Cluster	6,8%	1,8%	2,3%	4,8%
Viseu	Count	436	16	217	669
D10	% within Cluster	5,2%	1,8%	3,8%	4,5%
Vila Real	Count	426	22	82	530
D11	% within Cluster	5,1%	2,4%	1,5%	3,6%
Otros	Count	2159	155	831	3145
D12	% within Cluster	25,8%	17,2%	14,7%	21,1%
Total					
	Count	8356	900	5654	14910
		100%	100%	100%	100%

The CA map corroborates the previous conclusions: the majority of the points in cluster 3 is from districts D1, D2 and D7; in Cluster 2 from district D6 and Cluster 1 is on the centroid of all the rest.

Figure 4. CA map for Cluster and District.
Model: Symmetrical Normalization and distance Chi square.



4.3. Factor Analyses by time

CLUSTER-YEAR

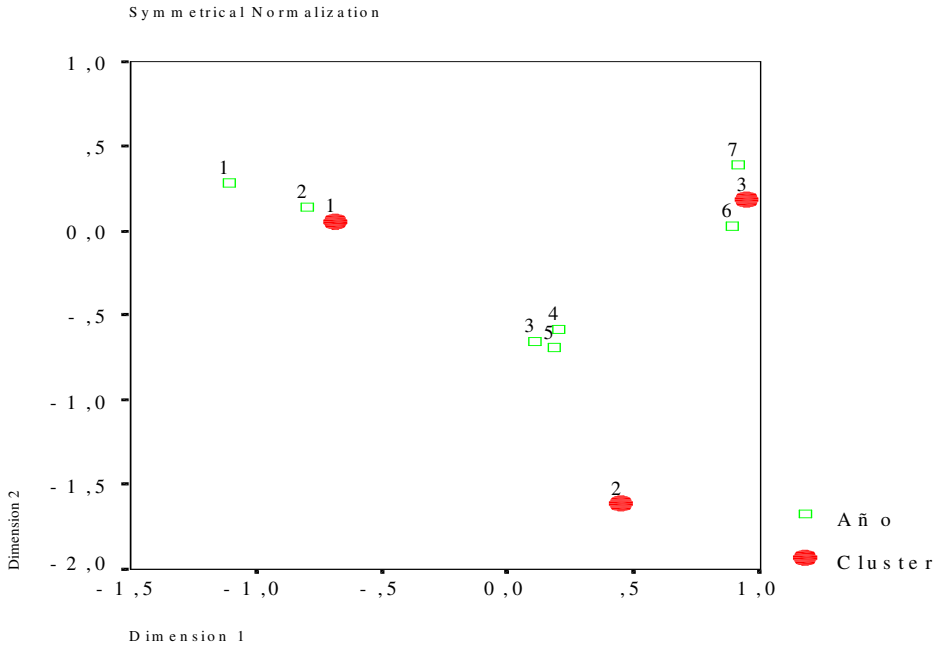
The Incentives in this study correspond to projects presented between the years 2000 and 2006. The distribution of the incentives for the different years is not homogeneous. See table and chart below.

Table 6. Project distribution, by year.

YEAR	Frequency	Percent
2000	2181	14,6
2001	3486	23,4
2002	1934	13,0
2003	1471	9,9
2004	806	5,4
2005	929	6,2
2006	4103	27,5
N		
Total	14910	100,0

The CA for Cluster and year corroborates the conclusions (figure below). The solution accounts for the 92% of inertia. The symmetric normalization model has been chosen.

Figure 5. First principal plane of the CA for Cluster and Year



The factorial plot shows how the first two years stimulate the projects in Cluster 1 (Area 1), while the last years stimulate the projects in Cluster 3 (area 2). The same information is more detailed on the following Table 7.

Table 7. Distribution of frequencies for Clusters and Years.
p-value for Chi square test < 0.01

		CLUSTER 1	CLUSTER 2	CLUSTER 3	
YEAR					Total
2000	Count	2173	7	1	2181
	% within Cluster	26,0%	,8%	,0%	14,6%
2001	Count	3039	89	358	3486
	% within Cluster	36,4%	9,9%	6,3%	23,4%
2002	Count	967	246	721	1934
	% within	11,6%	27,3%	12,8%	13,0%
2003	Count	688	181	602	1471
	% within Cluster	8,2%	20,1%	10,6%	9,9%
2004	Count	379	107	320	806
	% within Cluster	4,5%	11,9%	5,7%	5,4%
2005	Count	204	76	649	929
	% within Cluster	2,4%	8,4%	11,5%	6,2%
2006	Count	906	194	3003	4103
	% within Cluster	10,8%	21,6%	53,1%	27,5%
Total					
	Count	8356	900	5654	14910
		100%	100%	100%	100%

4.4. Factor Analyses by sector

CLUSTER-EAC[2]

80% of the projects of the PRIME Program, during the 7 years studied, are concentrated in the activities shown in the next table and 50% in only 5 (EAC 51, EAC 52, EAC 55, EAC 74 y EAC 91). This is why we have considered only the 5 most representative activities for the multivariate study.

Table 8. Distribution of PRIME funds by activity branch (EAC)

EAC	Frequency	Percent
15		
406	2,7	Manufacture of food products and beverages
17		
411	2,7	Manufacture of textiles and textile products
26		
447	2,9	Manufacture of other non-metallic mineral products
28		
482	3,1	Manufacture of basic metals and fabricated metal products except machinery and equipment
29		
299	2,0	Manufacture of machinery and equipment
36		
300	2,0	Manufacture of furniture; other manufacturing
45		
658	4,3	Construction
50		
540	3,5	Sale, Maintenance and repair of motor Vehicles and Motorcycles retail sale of automotive fuel
51		
1031	6,7	Wholesale Trade and commission trade, except of Motor vehicles and motorcycles
52		
3481	22,7	**Retails Trade, except of motor vehicles and motorcycles; repairs of personal and household goods**
55		
1675	10,9	Hotel and Restaurants
72		
377	2,5	Computer activities
74		
936	6,1	Other service activities (mainly to enterprise)
75		
386	2,5	Public Administration and defence, compulsory social security
91		
832	5,4	**Diverse associative activities**
Total		
12261	80.10	

[2] EAC – corresponding to the Portuguese Economic Activity Classification used in the national accounts, meaning the activity branches and sub-branches

The next table shows the cross-tabulation of the clusters and the activities. The projects in Cluster 1 are mainly in Activity 52; projects in Cluster 2 in Activity 91. The Cluster 3 is associated to the category "Others", that includes several activities representing much lower percents than the ones that are explicitly listed in the table and have much higher frequencies. More detailed information can be shown in the following table.

Table 9. Distribution of the different activities within the clusters
p-value for the Chi square test < 0.01

				Total
	CLUSTER 1	**CLUSTER 2**	**CLUSTER3**	
EAC51				
Count	609		383	992
% within Cluster	7,3%		6,8%	6,7%
EAC52				
Count	3302	1	150	3453
% within Cluster	**39,5%**	,1%	2,7%	23,2%
EAC55				
Count	1380	13	275	1668
% within Cluster	16,5%	1,4%	4,9%	11,2%
EAC74				
Count	464	110	294	868
% within Cluster	5,6%	12,2%	5,2%	5,8%
EAC91				
Count	107	**549**	96	752
% within Cluster	1,3%	**61,0%**	1,7%	5,0%
EAC Others				
Count	2494	**227**	4456	7177
% within Cluster	**29,8%**	**25,2%**	**78,8%**	48,1%
Count	8356	900	5654	14910

Table 10. Distribution of the different Systems within the clusters
p-value for the Chi square test < 0.01

		CLUSTER1	**CLUSTER 2**	**CLUSTER3**	Total
SYSTEM					
Form Prof	Count			1853	1853
	% within Cluster			**32,8%**	12,4%
Inf Assoc	Count		343		343
	% within Cluster		**38,1%**		2,3%
Inf TFQ	Count	6	308	21	335
	% within Cluster	,1%	**34,2%**	0,4%	2,2%
	% within Cluster	1,0%	7,1%	3,5%	2,3%
SIED	Count	80	15	276	371
	% within Cluster	1,0%	1,7%	4,9%	2,5%
SIME	Count	76		1858	1934
	% within Cluster	0,9%		**32,9%**	13,0%
SIME IN	Count	15		426	441
	% within Cluster	,2%		7,5%	3,0%
SIPIE	Count	3941	6	265	4212
	% within Cluster	**47,2%**	,7%	4,7%	28,2%
URCOM	Count	3562	76	89	3727
	% within Cluster	**42,6%**	8,4%	1,6%	25,0%
Others	Count	590	88	670	1348
	% within Cluster	7,1%	9,8%	11,9%	9,0%
Total					
	Count	8356	900	5654	14910
		100%	100%	100%	100%

4.5. Factor Analyses by support system

CLUSTER- SYSTEM

The characterization of the clusters is also very clear in this case: Cluster 1 is associated to the categories SIPIE (Incentive Scheme for Small Business Initiatives) and URCOM (Incentive Scheme for Commercial Urbanism Projects), Cluster 2 to Infrastructures and Associative Infrastructures (Support to the Existing Associative Infrastructures) and Cluster 3 to SIME (Incentive Scheme for Business Modernisation). See details in the above Table 10.

4.6. Cross Effects

CLUSTER – INVESTIMENT and INCENTIVE

The next two tables contain the cross-tabulation of the clusters with the Investment and the Incentive. Both numerical variables have been coded into four categories using the percentiles 25, 50 and 75 as cut points. It can be deduced from the tables that the patterns are very similar for both variables: Cluster 1 is associated to investments/incentives evenly distributed among the first three categories. The projects in Cluster 2 receive higher incentives in categories 3 and 4 but with higher frequencies in category 4. Cluster 3 is characterized by the higher investments/incentives.

Table 11. Distribution of frequencies according to Investment and Incentive

		CLUSTER1	CLUSTER 2	CLUSTER3	
INVESTMENT					
Invt1	Count	2856	14	931	3801
	% within Cluster	34,2%	1,6%	16,5%	25,5%
Invt2	Count	2655	62	1067	3784
	% within Cluster	31,8%	6,9%	18,9%	25,4%
Invt3	Count	2668	254	834	3756
	% within Cluster	31,9%	28,2%	14,8%	25,2%
Invt4	Count	177	570	2822	3569
	% within Cluster	2,1%	63,3%	49,9%	23,9%
Total	Count	8356	900	5654	14910
		100%	100%	100%	100%
INCENTIVE					
Inct1	Count	2872	14	910	3796
	% within Cluster	34,4%	1,6%	16,1%	25,5%
Inct2	Count	2623	41	1124	3788
	% within Cluster	31,4%	4,6%	19,9%	25,4%
Inct3	Count	2679	218	865	3762
	% within Cluster	32,1%	24,2%	15,3%	25,2%
Inct4	Count	182	627	2755	3564
Total	% within Cluster	2,2%	69,7%	48,7%	23,9%
	Count	8356	900	5654	14910
		100%	100%	100%	100%

CLUSTER – INVESTIMENT, INCENTIVE and RATIO

In order to better understand how successfully the programmes have developed, we applied CA to an investment – incentive ratio.

Table 12. Descriptive Statistics

CLUSTER 1	N	Minimum	Maximum	Mean	Std. Deviation
INVEST					
8356	1926,13	22774084,46	130239,70	444696,46	
INCENTIVE					
8356	674,15	4429168,41	54946,80	157597,12	
RATIO					
8356	0,01	1	0,43	0,11	
Valid N					
8356					

CLUSTER 2	N	Minimum	Maximum	Mean	Std. Deviation
INVEST					
900	11237,56	11879900,50	771208,30	1323334,91	
INCENTIVE					
900	4569,17	10401377,00	467527,60	958458,03	
RATIO					
900	0,10	1	0,560214	0,17	
Valid N					
900					

CLUSTER 3	N	Minimum	Maximum	Mean	Std. Deviation
INVEST					
5654	1566,24	750000000,00	2602544,23	17669066,22	
INCENTIVE					
5654	1174,68	60535273,56	605513,54	2486228,94	
RATIO					
5654	0,01	1	0,42	0,20	
Valid N					
5654					

Illustrated by Table 12, we can find that the first cluster, associated previously to Area 1 (Enterprise Stimulation), contains 8539 projects that received a mean incentive of 54946.8€ with a mean investment of 130279.7€. The distribution of the projects within this cluster is highly heterogeneous for both the incentive and the investment, as can be deduced from the high standard deviation. The smallest incentive is 1926.13€ and the highest 4429168€. The minimum ratio was 1%, the maximum was 100% and the mean value is 43%. The second cluster, associated previously to Area 3 (Stimulation of Business Environment), contains 900 projects that receive a mean incentive of 467527.6€ with a mean investment of 771208.3€. The distribution of incentives and investments within cluster 2 is also highly heterogeneous. The minimum incentive is 11237.56€ and the maximum incentive is 771208.30. The average ratio was 56% and the values range between 10% and 100%. The third cluster, associated previously to Areas 1 (Qualification of Human Resources) and 2 (Enterprise Stimulation), contains 5645 enterprises that received a mean incentive of 605513.54€ with an average investment of 2602544.23€. As for the previous clusters, incentive and investment values are highly heterogeneous, the incentive values range between 1174.68 and 60535273€. The average ratio was 42% and the values range from 1% to 100%.

4.7 Cluster TIPOLOGY

To better summarise the results from the application of Factor Analyses, a cluster typology has been developed for which it is possible to clearly understand how the different support types have been used by sectors, by years and by location. This cluster typology is reported in Table 13.

Table 13. Cluster Typology

	CLUSTER1	CLUSTER 2	CLUSTER3
AREA	**AREA 1** Support to firm stimulation	**AREA 3** Stimulation of Business Environment	**AREA 2** Human Resources Qualification
YEAR	2000, 2001	2002, 2003, 2004	2005, 2006
EAC	EAC 52 Retail (Excepting motor vehicles and motorcycles) Repair of household goods Others EAC	EAC 91 Diverse associative activities	EAC others
SYSTEM	SIPIE, URCOM	INFRAEST ASSOC INFRAEST TFQ	FORM PROF SIME
NUT	NORTH	NORTH LISBON	NORTH CENTER LISBON
DISTRICT	D1, D2,D3,D4,D5, D8, D9, D10, D11	D6. D7 Lisbon and Porto	D1, D2, D7 Aveiro, Braga Porto
INVESTIMENT	Inv1, Inv2, Inv3	Inv3, Inv4	Inv4
INCENTICV	Inc1, Inc2, Inc3	Inc3, Inc4	Inc4
RATIO	Rt 2, Rt 3	Rt 4	Rt 1, Rt 4

5. CONCLUDING REMARKS

From the presented results several conclusions can be pointed out. The first one is related to the asymmetric use of the incentives: The use of PRIME (as a public policy instrument for indirect distribution of European funds upon the entrepreneurial tissue), generated a clear geographic concentration of financial help in the northern part of the country. This region accumulated 41.4 per cent of the supports for firm stimulation, 39.0 per cent for stimulation of business environment and 49.4 per cent for human resources qualification. Certainly, the fact that other regions like Algarve, Azores and Madeira could only profit from this program in insignificant amounts (respectively 3.4, 1.9 and 1.5 per cent from the total available funds) is of great concern, suggesting the following question: Is there no entrepreneurial activity in such regions? Or, rather, location determined less use of supports by the existing companies? Is it possible that they could have been unable to follow the requirements for PRIME program calls?

Further observation of results shows that, by district, the concentration of PRIME supports goes clearly to Porto, a metropolitan area that absorbed respectively 12.8, 21.3 and 21.3 per cent of total available amounts. More surprising is the fact that Lisbon, the Portuguese capital, received 26.3 per cent help for stimulation of business environment, mainly driven for associative initiatives and other, not classified categories in the CAE (Classification of Economic Activities). If Lisbon, as cosmopolitan area, needs such a huge support how to classify than those conditions of less favoured Portuguese regions?

From the perspective of the sectors, there are difficulties to analyse a significant part of supports' destiny. Indeed, 48.1 per cent of the financed amounts were canalized to sectors, not specified by the CEA. As far as the remaining is concerned, the associative sector as well as commercial and trade activities had also good shares in the use of PRIME. The activities for

which funds have been mostly used are related to: infrastructure building (for associations) and commercial urbanism, and small business initiatives – not specified, but for improvement of business environment in general. A main question to this issue is: how did PRIME support the industrial and agricultural activities, and by means of which effects for the local, regional or national consumption trends?

A time-series observation can also be quite helpful for a better perception on how PRIME acted to promote activities in general. Rather then starting by shaping human resources and strengthening firm environment followed by firm financing activities, as suggested by the theory of endogenous growth, PRIME used the support system exactly on the inverse order. Our analyses points out the following: Those supports for Area 1, centred in firm stimulation, were attributed at the very start of the program - year 2000 and for a two years period. From 2002 to 2004, the stimulation of business environment was promoted, without integrated re-qualification of human resources; indeed, help for such projects arrived much later as it was the last PRIME target, due in 2005 and 2006. How is it possible to expect that companies can be duly stimulated by a support system without the convenient adequacy of their human resources at the very beginning?

In fact, PRIME system started up with a restricted investment in the area of enterprise stimulation (8539 projects with a average investment of 130279.7€). It was followed by few, but expensive, business environment actions mainly related to building of commercial infrastructures - about 900 projects were associated, now with a much higher average investment, 771208.3€. Finally, in the last execution period, a great amount was allocated to human resources qualification. In this case, 5645 projects have received an average investment of 2602544.23€; and despite these large amounts, the southern part of the country stayed almost away from any action. As proved in this paper, the total amount investments were primarily absorbed to speed the growth process of companies located in the northern and central regions of the country. In such regions, and because investments were concentrating mainly in few commercial activities, the system probably served also to aggravate the Portuguese trade balance. Still, some few indirect positive effects for the primary and secondary sectors should be expected – such results are now very difficult to evaluate in a struggle against the strong international financial crises.

Doubtless, PRIME served to accentuated asymmetries in Portugal. This could be avoidable if policy makers were able to anticipate such a risk and advance an exercise on crossed public policies. Thus, many efforts done by regional policy makers were reduced to ineffectiveness due to a different prioritization established by the Portuguese national innovation policy – whatever it might have been.

REFERENCES

ARROW, K.J. (1962), "The Economic Implication of Learning by Doing", *Review of Economic Studies*, vol. 29, no. 3, pp. 155-173

ACS, Z. and D.B. AUDRETSCH (1984), "Small Business in Industrial Economics: The NewLearning", *Revue d'Economie Industrielle*, no. 7, pp. 21-39

ACS, Z. (2002), *Innovation and the Growth of Citie*s, Edward Elgar, Cheltenham

Conselho de Ministros, Resolução no. 101/2003, Portugal

DOMAR, EVSEY D. (1946), "Capital Expansion, Rate of Growth, and Employment",
Econometrica, vol. 14, no. 2, pp. 137-147

DURLAUF, S. (2002), "On the empirics of social capital". *Economic Journal* 112 (483), pp. 459 – 479

FAGERBERG, J. (2003), "Schumpeter and the Revival of Evolutionary Economics: an Appraisal of the Literature", *Journal of Evolutionary Economics*, vol. 13, pp. 125-59

FAGERBERG, J. (2004), "A Guide to the Literature", in: J. Fagerberg, D. Mowery, M.T.
TERRILL and R. NELSON (eds.), *The Oxford Hand Book of Innovation*, Oxford University
Press, Oxford, pp. 1-26

GIFI A. (1990) Nonlinear Multivariate Analyses. Chichester (England): John Wiley and Sons. 579 p

HARROD, R. F. (1939), "An Essay in Dynamic Theory", *Economic Journal,* vol. 49, no. 193, pp.14-33

KEEBLE, D. and WILKINSON, F. (1999): "Collective Learning and Knowledge Development in the Evolution of Regional Clusters of High Technology SMEs in Europe", *Regional Studies*, vol.33, nº4, pp. 295 – 303.

LUCAS, R.E. (1988), "On the mechanics of Economic Development", *Journal of Monetary Economics*, vol. 22-1988, pp. 3-42

ROMER, P.M. (1986), "Increasing returns and long-run growth", *The Journal of Political Economy*, vol. 94-1986, pp. 1002-1037

ROMER, P.M. (1990), "Endogenous technical change", *The Journal of Political*

Economy, vol. 98, no. 5, pp. 71-102

SIERRA, C. (1997): "Proximité(s), Interactions Technologiques et Territoriales: une revue". Revue d'Économie Industrielle, nº 82, p. 7 – 38.

SOGAARD, V., (2009), "Entrepreneurship and growth: the need to combine micro and macro perspectives", in Galindo, M.-A., Guzman, J. and Ribeiro, D. (Eds), *Entrepreneurship and Business: a regional perspective*. Springer, Heidelberg

SOLOW, R. (2000), "Notes on social capital and economic performances", in Dasgupta P. and Serageldin, I. (Eds), *Social Capital: A multifaceted perspective*. The World Bank, Washigton

SOLOW, R. M. (1956) "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, vol. 70, pp. 65-94

VAZ, M.T.N. and NETO, P. (2005) "Contributos para definição de uma tipologia de inovação em zonas de fronteira". *Papeles de Economia Espanõla*, Economia de las Comunidades Autónomas nº21/2005, Pgs. 133 – 146, FUNCAS Madrid, Espanha

VAZQUEZ-BARQUERO, A. (2002), *Endogenous Development*, Routledge, London.

Annex 1

PRIME Support Measures and Incentives Schemes

AREA 1 - Enterprise Stimulation

DEMTEC – Incentive Scheme for the Implementation of Pilot Project Related to Technologically Inovative Products, Processes and Systems
EXECUTIVES – Executives Programme
IDEIA – Support to Business Applied Research and Development
MAPE – Support Measure for the Maximisation of Energy Potential and for Streamline Consumption
NEST – New Enterprises Based on Technological Support
NITEC – Incentive Scheme for the Establishment of Research and Technological Development Clusters within Business Sector
SICE – Incentive Scheme for Business Co-operation
SIED – Incentive Scheme for Digital Economy
INOV-JOVEM – Young Executives for Innovation in SME's
SIME – Incentive Scheme for Business Modernisation
SIME INTERNATIONAL – Incentive Scheme for the Modernisation of the Economy - International Development
SIME R&TD – Incentive Scheme for Business Modernisation - Research and Technological Development
SIPIE – Incentive Scheme for Small Business Initiatives
SIUPI – Incentive Scheme for the Use of Industrial Proprety
SIVETUR – Incentive Scheme for Tourism Products with a Strategic Dimension
URBCOM – Incentive Scheme for Commercial Urbanism Projects

AREA 2 - Qualification of Human Resources

TECHNOLOGICAL SCHOOLS – Supporting Technological Training
VOCATIONAL TRAINING – Fostering Investments in Human Resources

AREA 3 - Stimulation of the Business Environment

ASSOCIATIVE INFRASTRUCTURES – Support to the Existing Associative Infrastructures
BUSINESS PARTNERSHIPS – Supporting Business Partnerships
ENERGY INFRASTRUCTURES – Support Measure for the Modernisation and Development of Energy Infrastrutures
FSCR – Risk Capital Syndication Fund
GUARANTEE MECHANISMS – Establishment and Strengthening Guarantee Mechanisms
HISTORIC INNS – Support to the Use and Valorisation of the Natural Heritage
INTERNATIONALISATION – Support to the Internationalisation of the Economy
MINING ZONES – Support to the Requalification of Mining Zones
PITER – Support to Integrated Tourism Programmes with Strategic Nature and Regional Basis
TOURISM INFRASTRUCTURES – Support to Tourism Infrastructures
TTQ INFRASTRUCTURES – Support Measure to the Setting-up of New Technological Infrastructures and to the Existing Technological, Training and Quality Infrastructures

Annex 2
PRIME beneficiaries, geographical approach and incentive nature

SUPPORT MEASURES AND INCENTIVE SCHEMES	BENEFICIARIES	NATURE OF THE INCENTIVE	ACTIVITY SECTORS	GEOGRAPHICAL APPROACH
AREA 1 - Enterprise Stimulation				
DEMTEC	Enterprises and NOG 1)	Non refundable incentive	I, BI, C, E, Tr, T, S	No
EXECUTIVES	Small and Medium Enterprises	Non refundable incentive	I, C, T, S	No
IDEIA	Consortia of Enterprises and NTSS	Refundable & Non refundable incentive	I, E, BI, C, T, S	No
MAPE	2)	Refundable & Non refundable incentive	I	No
NEST	3) & 4)	Non refundable incentive	I, E, BI, Tr, C, T, S	No
NITEC	Enterprises	Non refundable incentive	I, E, BI, T, C, S	Yes
SICE	Entreprises cooperation networks	Refundable incentive	I, BI, C, T, S	No
SIED	Micro and SMEnterprises	Non refundable incentive	I, BI, C, T, Tr, S	Yes
INOV-JOVEM	young executives	Non refundable incentive	I, BI, C, E, Tr, T, S	No
SIME	Enterprises	Refundable incentive	I, BI, C, T, Tr, S	No
SIME INTERNATIONAL	Micro and SMEnterprises	Non refundable incentive	I, BI, C, T, Tr, S	No
SIME R&TD	Enterprises	Refundable & Non refundable incentive	I, BI, C, T, Tr, S	No
SIPIE	Micro and SMEnterprises	Non refundable incentive	I, B, C, T, Tr, S	Yes
SIUPI	5)	Non refundable incentive	I, BI, E, C, PI, T, S	No
SIVETUR	Enterprises	Refundable & Non refundable incentive	T	No
URBCOM	6)	Non refundable incentive	C, T, S	No
AREA 2 - Qualification of Human Resources				
TECHNOLOGICAL SCHOOLS	Technological schools	Non refundable incentive	I, BI, C, E, T, S	No
VOCATIONAL TRAINING	7)	Non refundable incentive	I, BI, C, E, Tr, T, S	Yes
AREA 3 - Stimulation of the Business Environment				
ASSOCIATIVE INFRASTRUCTURES	8)	Non refundable incentive	I, BI, C, E, Tr, T, S	No
BUSINESS PARTNERSHIPS	9)	Non refundable incentive	I, BI, C, E, Tr, T, S	No
ENERGY INFRASTRUCTURES	10)	Non refundable incentive	E	No
FSCR	Micro and SMEnterprises	Non refundable incentive	I, BI, C, E, T, Tr, S	No
GUARANTEE MECHANISMS	11)	Non refundable incentive	I,BI, C, E, T, Tr, S	No
HISTORIC INNS	Enatur - Pousadas de Portugal	Non refundable incentive	T	No
INTERNATIONALISATION	12)	Non refundable incentive	I, BI, C, E, T, Tr, S	No
MINING ZONES	13)	Non refundable incentive	I	No
PITER	14)	Refundable & Non refundable incentive	T	Yes
TOURISM INFRASTRUCTURE	Institutional training bodies	Non refundable incentive	T	No
TTQ INFRASTRUCTURES	15)	Non refundable incentive	_	Yes

1) - Non Profit Organisation from the National Scientific and Technological System - NTSS (projects in EAC- Economic Activities Classification - 40)
2) - Firms, municipalities, business and labour associations, schools, health and social bodies, civil protection entities
3) - Persons intending to promote the establishment of new enterprises and recently established technology support enterprises with no significant activities
4) - Entrprises, inventors, independent designers, entrepreneurs at pre-entrepreneur stage and institutions developing research stages
5) - Micro and SMEnterprises, commerce associative structures, municipalities and co-ordinating and supervising units
6) - Enterprises and actors from the business environment and technological schools
7) – National, regional and sector business associative structures classified under EAC (Economic Activities Classification) 91110, federations or confederations of associative structures and regional and local tourism offices
8) - Public organisations and non-profit making private organisations
9) - All concessionaires enterprises of transport and supply of electric power and gas, as well as other enterprises operating those services under public services licensees
10) - Enterprises specialised in the development of SMEs oriented financial innovation, public administration bodies, entities from the National Scientific Technological System, or other representative entities intervening in the development of specific areas of economic activity
11) – Bodies and entities of the public administration, national, regional and sector business associative structures classified under EAC (Economic Activities Classification) 91110, federations or confederations of associative structures
12) Enterprises holding the concession for the rehabilitation works in abandoned mines and bodies from the Ministry of Economy and Innovation
13) - All entities, irrespective of their nature and juridical statute, and namely enterprises, municipal counties, tourism regions and other local or regional tourism bodies
14) - Technology centres, technology transfer centres, technology-based incubator centres and parks, technological schools and interfacing public entities

I - Industry, E - Energy, BI - Building Industry, C - Commerce, Tr - Transport, PI - Pharmaceutical Industry, T - Tourism, S - Services

Annex 3

	Variables	Categories	Number	Labels
AREAS OF STRATEGIC ACTION	AREA	Enterprise Stimulation Enterprise Stimulation Qualification of Human Resources Stimulation of the Business Environment	3	EIXO1 EIXO2 EIXO3
Locational Factors	NUTSII	North Centre Lisbon Alentejo Algarve Acores Madeira	7	North Centre Lisbon Alentejo Algarve Acores Madeira
	DISTRICT	Aveiro Braga Coimbra Faro Leiría Lisbon Porto Santarém Viana do Castelo Viseu Vila Real Others		D1 D2 D3 D4 D5 D6 D7 D8 D9 D10 D11 D12
Temporal Factor	YEAR	2000 2001 2002 2003 2004 2005 2006		2000 2001 2002 2003 2004 2005 2006
Support System	SISTEMCORT	 Profesional Formation **S**upport to the Existing Associative Infrastructures **S**upport Measure to the Setting-up of New **T**echnological Infrastructures and to the Existing Technological, Training and **Q**uality Infrastructures Support Measure for the Maximisation of Energy Potential and for **S**treamline Consumption **S**upport Measure for the Maximisation of Energy Potential and for Streamline **C**onsumption **I**ncentive Scheme for Digital Economy **I**ncentive Scheme for Business Modernisation **I**ncentive Scheme for the Modernisation of the Economy – International Development **I**ncentive Scheme for Small Business Initiatives **I**ncentive Scheme for Commercial Urbanism Projects Others		Form Prof Inf Assoc Inf TFQ MAPE SIED SIME SIME INTERNT SIPIE URBCOM Others
Financial System	INVESTIMENT	INVESTIMENT < 25 Percentil (P25) P25= INVEST MENT< P50 P50= INVESMENT< P75 PINVESMENT≥ P75		Inv 1 Inv 2 Inv 3 nv 4
	INCENTIVE	INCENTIVE < 25 Percentil P25= INCENTIVE < P50 P50= IINCENTIVE < P75 INCENTIVE ≥ P75		Inc 1, Inc 2 Inc 3, Inc 4
	RATIO	RATIO < 25 Percentil P25= RATIO < P50 P50= RATIO < P75 RATIO ≥ P75		Rt 1 Rt 2 Rt 3 Rt 4

REORDENACIÓN Y TRANSFERENCIAS EN LA FINANCIACIÓN ADMINISTRACIONES SUBCENTRALES. UN EJERCICIO COMPARADO ENTRE ESPAÑA Y ALEMANIA[1]

LUIS ÁNGEL HIERRO RECIO
PEDRO ATIENZA MONTERO
E-mail: atienza@us.es
ROSARIO GÓMEZ-ÁLVAREZ DÍAZ
Departamento de Historia e Instituciones Económicas y Economía Aplicada
Facultad Ciencias Económicas y Empresariales
Universidad de Sevilla

Resumen

El vigente Estatuto de Autonomía de Cataluña introduce por primera vez el concepto de ordenación y reordenación en el ámbito de la financiación autonómica, lo cual supone también una novedad en el campo del federalismo fiscal. Dada las implicaciones distributivas que supone la reordenación generada por las transferencias, aquí se intenta medirlas mediante el empleo de índices de concentración, lo cual constituye una profundización respecto a trabajos anteriores de los autores en cuanto a medir el efecto redistributivo de las transferencias. Asimismo se utilizan indicadores basados en matrices de transición y comparación de rangos, para obtener mayor información sobre la magnitud y sentido de la reordenación. Además se realiza un ejercicio comparado con Alemania.

Palabras clave: financiación autonómica, Alemania, reordenación, transferencias intergubernamentales, desigualdad.
Área Temática: Economía Urbana, regional y local.

Abstract

The recent reform of the Statute of Autonomy of Catalonia enters by first time the concept of ranking and reranking in the field of the financing of Autonomous Communities, which supposes also a novelty in the field of the fiscal federalism. Given the distributive implications that supposes the reranking generated by the transfers, here we tries to measure them by means of the employment of indexes of concentration, which constitute a study in depth with regard to previous works of the authors of measurement of the redistributive effect of the transfers. Likewise we use other indicators based in matrices of transition and comparison of ranks, to obtain main information on the magnitude and sense of the reranking. We achieve an compared exercise with Germany.

Key words: financing of Autonomous Communities, Germany, reranking, intergovernmental transfers, inequality
Thematic Area: Local, regional and urban economy.

1. ANTECEDENTES Y OBJETIVOS

Desde hace algunas décadas el concepto de orden y por ende el de reordenación ha merecido una cierta dedicación en el ámbito de la economía. En el sistema de financiación autonómica de régimen común dicho concepto ha adquirido recientemente relevancia como consecuencia de la reciente reforma del Estatuto de Cataluña, plasmada en la Ley Orgánica

[1] Este trabajo forma parte de un proyecto de investigación financiado por el Centro de Estudios Andaluces mediante ayuda concedida al amparo de la Convocatoria de Ayudas a Proyectos de Investigación 2008 bajo la denominación "La incorporación del concepto de reordenación a la financiación autonómica".

6/2006, de 19 de julio, donde se incluye por primera vez la ordenación de las CC.AA. como criterio a considerar en la financiación autonómica. Sin embargo, este ámbito de la financiación federal es muy distante de aquellos en los que tradicionalmente los investigadores han venido preocupándose por el tema de la reordenación, lo que da lugar a la necesidad de un esfuerzo de adaptación de la metodología de medición de la reordenación en esos ámbitos donde se ha analizado a nuestro campo de estudio. En efecto, a lo largo de los últimos cuarenta años la medición de la reordenación ha estado asociada, por un lado, al estudio de la justicia tributaria y, por otro, al de la movilidad económica, resultando que mientras en los trabajos de movilidad económica la mayoría de los instrumentos de medición se obtienen a partir de matrices de transición y correlación de rangos, en el ámbito de la equidad tributaria los instrumentos básicos de análisis y medida se construyen a partir de índices de Gini y de concentración, de manera que ambas líneas de investigación han funcionado como compartimentos estanco.

La primera de las corrientes señaladas, la afincada en materias de carácter fiscal, se ha dedicado fundamentalmente al análisis de la equidad horizontal y vertical, incorporando la reordenación como medida de la pérdida de potencial igualador y por tanto como generadora de inequidad horizontal. A partir de la propuesta pionera de Feldstein (1976), se desarrolla toda una línea de investigación en el ámbito de la teoría de la imposición sobre la reordenación y la equidad horizontal que tiene su origen en los trabajos de Atkinson (1980), Plotnick (1981) y Kakwani (1984). De hecho, el análisis del impuesto sobre la renta se realiza con la agrupación de sujetos o familias por tramos de renta, con lo que la posibilidad del cambio de tramo tras la imposición implica introducir el concepto de reordenación. Esta situación lleva a Aronson, Johnson y Lambert (1994) a plantearse la necesidad de diferenciar entre el tratamiento desigual de los iguales y la reordenación entre los desiguales. Para estos autores la reordenación no tiene su origen en el hecho de que los iguales modifiquen su orden dentro del grupo, sino en que individuos que originalmente pertenecen al mismo grupo pasen a pertenecer a otro diferente tras el impuesto, "saltando" de grupo. Estos autores denominan a este acontecimiento *reordenación por solapamiento*, diferenciándolo de la desigualdad horizontal, que se produce cuando se da un trato desigual a los iguales.

Se abren así dos corrientes en los estudios del impacto distributivo de las políticas impositivas basadas en dos concepciones diferentes de la reordenación y de la forma de medirla. Para la corriente tradicional de Atkinson, Plotnick y Kakwani sólo existe inequidad horizontal si hay reordenación y viceversa, mientras que para quienes siguen la proposición de Aronson, Johnson y Lambert la reordenación no es condición necesaria para la generación de desigualdad horizontal.

Más recientemente, el trabajo de Urban y Lambert (2005) reconcilia ambas visiones. Así, si los agentes están agrupados y se produce reordenación entonces se pueden identificar el cambio de orden dentro de los grupos, entre los grupos y la reordenación por solapamiento, siendo la suma de los tres la reordenación total obtenida según la propuesta tradicional de Atkinson, Plotnick y Kawkani. Gracias a esta aportación se delimitan, se cuantifican y se relacionan las diferentes concepciones de reordenación derivadas de curvas de concentración. Además estos autores clarifican los conceptos de inequidad horizontal y reordenación puesto que la desigualdad horizontal sin reordenación tiene su origen en el agrupamiento de los individuos similares, de forma que es posible que estos tengan un tratamiento desigual sin que ello suponga la reordenación. En el caso de que los individuos de los grupos sean exactamente iguales cualquier tratamiento desigual generaría desigualdad dentro del grupo y por tanto reordenación dentro del grupo.

Por su parte, uno de los estudios pioneros en la corriente de la movilidad económica puede atribuirse a Bartholomew (1973) quien desde el ámbito de la sociología económica, y con objeto de evaluar la movilidad intergeneracional entre clases sociales definidas profesionalmente, planteó la necesidad de medir lo que denominó "*mobility exchange*", es decir la movilidad que se da en una sociedad cuando unos individuos sustituyen a otros en sus posiciones sociales. En esos comienzos, fundamentalmente con el trabajo de Shorrocks (1978), se formaliza la medición de la movilidad económica a través de índices, generándose a continuación una amplia literatura en la que se elaboran múltiples indicadores de movilidad, una parte de los cuales son medidas de reordenación. Fruto de este desarrollo comienzan a forjarse dos ámbitos de trabajo sobre

movilidad económica, el que se concentra en el estudio de las propiedades de los indicadores, aspecto éste donde destacan los trabajos de Shorrocks (1993) y Fiels y Ok (1996), y el que se orienta hacia la medición en términos de bienestar social (King (1983) o Markandya (1982)).

Precisamente siguiendo la metodología propuesta por Aronson, Johnson y Lambert (1994) para el análisis de la progresividad del impuesto sobre la renta, en Hierro, Atienza y Patiño (2007) se incorpora la reordenación en el ámbito de la financiación federal, en este caso como indicador de la pérdida de capacidad igualatoria de las transferencias en relación a los ingresos tributarios obtenidos originariamente por los Estados. En efecto, si el objetivo de los sistemas de transferencias suele ser producir una cierta igualación de los recursos obtenidos por los Estados para cubrir al menos las diferencias de ingresos derivadas de una capacidad tributaria desigualmente distribuida, el hecho de que las transferencias produzcan reordenación implica que el objetivo igualador se traspasa para producir no ya igualación sino la inversión de posiciones. Desde esta perspectiva la reordenación deja de ser una cuestión baladí para constituirse en un tema de gran trascendencia. ¿Permitiría un determinado sistema de transferencia la igualación de los recursos per cápita? ¿Cómo se mide la desigualdad derivada de dotar un sistema de transferencias intergubernamentales que producen reordenación? ¿Es igual que la reordenación se produzca en la parte alta de los ingresos per cápita que en la parte baja? …

La respuesta de estas cuestiones es el objetivo que nos proponemos para el presente trabajo. Para ello realizamos un ejercicio comparado entre la financiación autonómica de régimen común y la financiación de los Länder alemanes, toda vez que el sistema federal alemán en general y, en concreto, en lo que se refiere a la financiación de los gobiernos de nivel intermedio es considerado una referencia a nivel internacional.

En el epígrafe 2 se describe la adaptación de las principales metodologías de medición y valoración de la reordenación a la cuantificación de la reordenación que provocan las transferencias intergubernamentales. En el epígrafe 3 se presentan los resultados de dicha adaptación y aplicación y finalmente el trabajo termina con un breve apartado de conclusiones.

2. METODOLOGÍA

Tal como se ha señalado, la primera y fundamental consecuencia de la reordenación es la merma de las posibilidades de igualación de las transferencias. Para el análisis de este problema vamos a tomar la metodología de índices de Gini y de concentración, aplicada en el campo de estudio de la equidad de los impuestos como anteriormente se ha descrito.

Siguiendo la propuesta de Kakwani (1984)[2] y adaptándola a nuestro objeto de estudio, si denominamos $Lx_a(p)$, a la función representativa de la curva de concentración de los recursos de los Estados antes de las transferencias (recursos tributarios), $Lx_d(p)$ a la de la curva de concentración de los recursos de los Estados después de las transferencias (recursos totales), y por último $\overline{Lxd}(p)$ a la de la curva de concentración de los recursos totales pero ordenados los Estados según el orden inicial, es decir según sus recursos antes de transferencias, podemos definir los respectivos índices de concentración como:

$$C(x_a) = 1 - 2\int_0^1 Lx_a(p)dp$$

$$C(x_d) = 1 - 2\int_0^1 Lx_d(p)dp$$

$$\overline{C}(x_d) = 1 - 2\int_0^1 Lx_d(p)dp$$

El efecto igualador de las transferencias se puede medir mediante el índice del Reynolds-Smolensky, que se define como $RS(x) = C(xa) - C(xd)$ o por el índice Pechman-Okner

[2] Cuya aportación es relevante porque ofrece la base teórica y empírica para evaluar la contribución de la progresividad y la reordenación sobre el impacto igualador final de cualquier política de gasto o impositiva.

$PO(x) = \frac{RS(x)}{C(xa)} \times 100$. Y el impacto de la reordenación de las transferencias, *IR*, vendrá dado por:

$$IR(x) = \overline{C}(x_d) - C(x_d)$$

La contribución por reordenación toma valor negativo[3] y representa la pérdida de impacto igualador derivada de la alteración en las posiciones relativas de las distintas CC.AA. Es por tanto un indicador de la mayor o menor reordenación que produce cada sistema de transferencias.

Una segunda cuestión a considerar son la intensidad y el sentido de la reordenación. No es igual que se produzca reordenación mediante saltos muy grandes pero concentrados en pocos Estados a que sea más generalizada pero de menor intensidad. Igualmente tampoco será igual que la reordenación tenga su origen en que los Estados con menores ingresos escalen posiciones, a que la reordenación tenga su origen en que los Estados originariamente con mayores ingresos sean los perdedores de las posiciones. Estos matices, que en apariencia podrían parecer escasamente relevantes, pueden ser de trascendencia sobre todo cuando existen reticencias a aceptar los resultados distributivos de los sistemas de transferencias, como ocurre en el caso de España.

Para el estudio de estas cuestiones se pueden utilizar las medidas asociadas a matrices de transición, de uso generalizado en el análisis de problemas de movilidad económica. Una matriz de transición se construye identificando los valores cardinales que toman los elementos en dos momentos diferentes y de acuerdo con dichos valores se asignándole un valor ordinal dentro de los *n* posibles. Dicha transformación en órdenes se puede representar mediante una matriz de transición $(A)_{nxn}$ biestocástica[4], de forma que, como es en nuestro objeto de estudio, si la matriz describe hechos pasados y los sujetos son individuos, en nuestro caso sujetos jurídicos como los Estados de cada país, para cada fila *i*, existirá un elemento a_{ij} con un valor igual a 1, que denota la posición *j* final del individuo que se encontraba en la posición inicial *i*, tomando el valor cero el resto de elementos de la referida fila.

La medida más elemental derivada de una matriz de transición es la tasa de cambio, calculada como:

$$C = \frac{n - Traza\ \mathrm{A}}{n}$$

siendo Traza A la suma de los elementos a_{ij} que se sitúan en la diagonal principal de la matriz. La tasa de cambio nos informa del porcentaje de sujetos que no han permanecido en su misma posición pero no informa de la intensidad de los cambios de posición.

Un indicador de reordenación que permite medir la intensidad de la reordenación es la medida de la distancia de Scott y Litchfield (1994). Dicha medida se calcula mediante el cociente de la distancia entre el orden inicial *i* y el final *j,* ponderada por la probabilidad de cambio, y la máxima reordenación posible[5], y viene expresada por la ecuación[6]:

$$I_D = \left[\frac{\sum_{i=1}^{n} \mathrm{aij} |i - j|}{Maxreord} \right]^2$$

$$Maxreord = \sum_{i=1}^{n} \sum_{j=1}^{n} aij(i-j)^2, siendo \quad j = N + 1 - i$$

[3] Esa pérdida ya está descontada al calcular la transformación del índice Reynolds-Smolensky.
[4] La suma de cada columna y de cada fila es uno.
[5] Según Plotnick (1981) la máxima reordenación se produce cuando los sujetos pasan a una posición final inversa a la inicial, de forma que cualquier unidad en el orden *i* pasaría a un orden final igual a *N+1-i*.
[6] Esta fórmula es una adaptación de la propuesta de Plotnick (1982), donde la diferencia en órdenes está elevada a *h,* siendo h un parámetro a elegir por el investigador.

La medida de distancia está normalizada entre 0 y 1, tomando el valor cero en el caso de que no exista ningún cambio en el orden y el valor 1 si se ha producido la máxima reordenación. A su vez este índice se puede descomponer en dos índices para identificar el sentido de la reordenación: ascendente (I_{DA}) y descendente (I_{DD}):

$$I_D = I_{DA} + I_{DD}$$

tal que:

$$I_D = \frac{f_A}{\max} + \frac{f_D}{\max}$$

siendo:

$$f_A = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}(i-j)^2, j \geq i \geq 1$$

$$f_D = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}(i-j)^2, i \geq j \geq 1$$

Otra medida que toma como referencia la inversión del orden, mediante el método de correlacionar rangos, es el coeficiente de Spearman:

$$\rho = 1 - \frac{6\sum_{i=1}^{n}(i-j)^2}{n(n^2-1)}$$

Si el coeficiente toma el valor 1 el orden se mantiene de forma sistemática, y conforme se acerca a 0 mayor es la independencia del orden final de la variable respecto al orden inicial. En el caso de que el coeficiente se aproxime a –1, es una muestra de los órdenes en un momento futuro varían sistemáticamente con respecto a los valores homónimos en el momento inicial pero en sentido contrario, y por tanto, representa la máxima reordenación según la definición de Plotnick ya explicada.

3. LOS RESULTADOS

Para realizar el ejercicio comparativo que nos proponemos se han tomado los datos de la financiación federal alemana de la base de datos online GENESIS de la Oficina Federal Estadística (www-genesis.destatis.de/genesis/online/logon) y del Anuario Estadístico Anual "Statistisches Jahrbuch für die Bundesrepublik Deutschland und für das Ausland 2006" publicado en CD-ROM por el mismo organismo. Por su parte para España la fuente básica han sido los datos suministrados por el Ministerio de Economía y Hacienda en el documento "Financiación de las Comunidades Autónomas por los impuestos cedidos, fondo de suficiencia y garantía de financiación de los servicios de asistencia sanitaria, correspondiente al ejercicio 2006", disponible en la web (www.minhac.es/Portal/Estadistica+e+Informes/Estadisticas+territoriales/Informes+financiacion+comunidades+autonomas2.htm) y los datos del Fondo de Compensación Interterritorial y de los convenios de inversión y contratos programas se han obtenido de la base de datos BADESPE (Base de Datos del Sector Público español), disponible en la web www.estadief.meh.es.

A partir de la cifras de financiación resumidas en los Cuadros 2 y 3, se ha procedido a calcular los indicadores de reordenación, que aparecen en el Cuadro 1. El aspecto más relevante de la reordenación en el ámbito de la financiación subcentral es el hecho de que provoca pérdida de igualación de las transferencias. En consecuencia, comenzando con los indicadores basados en curvas de concentración, se deduce, en primer lugar, que mientras que la desigualdad de la distribución de los recursos tributarios es muy semejante en ambos países (aunque ligeramente superior en Alemania), una vez añadidas las transferencias el impacto igualador de éstas es muy diferente en ambos países, con un índice de Pechman-Okner de 72,4% en España y de 24,2% en Alemania, alcanzando el índice de concentración unos valores de 0,036 y 0,11 respectivamente. De lo cual se deduce que la pérdida de efecto igualador provocado por la reordenación es mucho

mayor en Alemania que en España, pérdida que se produce como consecuencia de que las transferencias no sólo hacen que se igualen los ingresos per cápita sino que en muchos casos hacen que los Estados que reciben menos ingresos originalmente a la postre sobrepasen a los que inicialmente obtenían más. En efecto, dicha pérdida es, en términos absolutos, de -0,13 en Alemania y de -0,055 en España.

Las respectivas curvas de concentración correspondientes a los índices de concentración mencionados se representan en los Gráficos 1 y 2, correspondientes a España y Alemania respectivamente. Se ilustran tres curvas de concentración correspondientes a las ya explicadas *Lxa(p), Lxd(p)* y $\overline{Lxd}(p)$, es decir, a las curvas de concentración de los recursos antes de transferencias, después de transferencias y finalmente después de transferencias pero con el orden inicial. Dichas curvas se han nombrado como las curvas A, B y C respectivamente. Pues bien, el área comprendida entre A y B es el efecto igualador de las transferencias (medido por el índice Reynolds-Smolensky) y entre B y C el efecto de pérdida de igualación derivada de la reordenación.

Cuadro 1. Indicadores de reordenación. Año 2006

		Alemania	**España**
Basados en curvas de concentración			
Índice concentración antes transferencias (C(xa))		0,1460	0,1313
Índice concentración después transferencias (C(xd))		0,1107	0,0362
Pseudo-índice de concentración $\overline{C}(xd)$		-0,0209	-0,0190
Reynolds-Smolensky (RS(x))		0,0353	0,0951
Pechman-Okner (PO(x))		24,16%	72,41%
Reordenación (IR(x))		-0,1316	-0,0552
Basados en matrices de transición y comparación de rangos			
Tasa de cambio		0,8750	0,8667
Medida de distancia (índice de Plotnick con h=2)		0,4721	0,6518
Descomposición medida de distancia	Ascendente (%)	49,53%	43,42%
	Descendente (%)	50,47%	56,58%
Coeficiente de correlación de Spearman		0,0559	-0,3036

Fuente: Elaboración propia.

Gráfico 1.- España. Curvas de concentración.

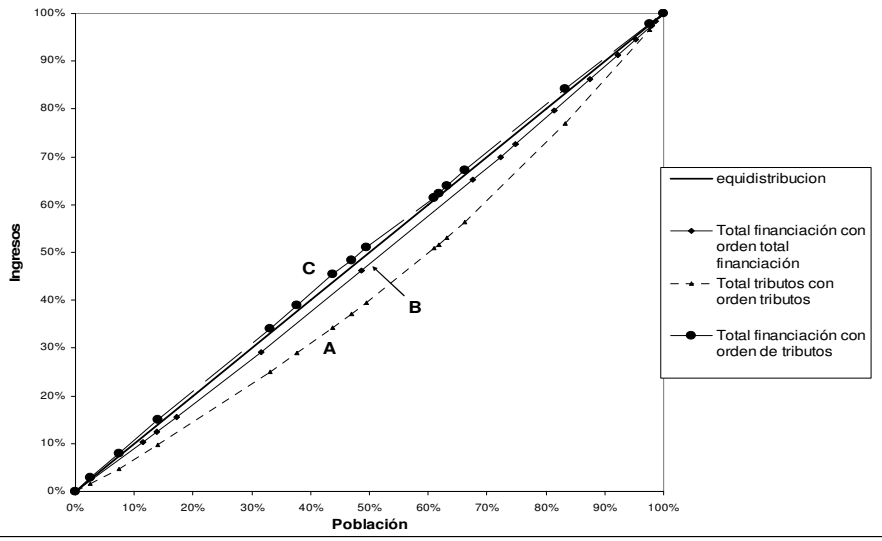
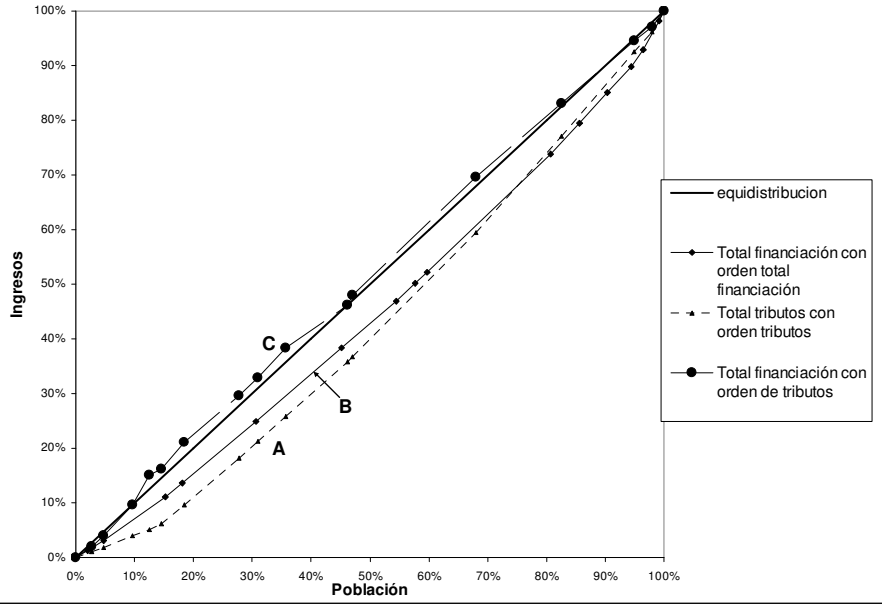


Gráfico 2.- Alemania. Curvas de concentración.



Fuente: Elaboración propia a partir del Cuadro 1 y de datos de población del INE y del Anuario estadístico anual.

Se observa que la pérdida de igualación que provoca la reordenación, el área entre B y C, es mucho mayor en Alemania. Por otra parte, a pesar de que el índice de concentración antes de transferencias es muy semejante en ambos países, de la visualización de las respectivas curvas se desprende que la concentración de la desigualdad antes de transferencias es diferente en ambos países. En efecto, en Alemania se produce una concentración de dicha desigualdad en los *Länder* con menos recursos tributarios per cápita, a saber, Turingia, Meclemburgo-Antepomerania, Sajonia, Sajonia-Anhalt y Sarre, mientras que en España la concentración de la desigualdad es más uniforme.

Información adicional nos puede suministrar otros indicadores de reordenación desligados de la medición de los efectos distributivos y que, en consecuencia, sólo miden la intensidad y el sentido de la reordenación. Tales indicadores son calculados a partir de los Cuadros 2 y 3. Así, un primer indicador es la tasa de cambio, que expresa el porcentaje de *Länder* o CC.AA. que han cambiado de posición al añadir las transferencias. Pues bien, dicha tasa es muy elevada (el 87%) y prácticamente igual en los dos países. En efecto sólo Asturias y Aragón en España no cambian de posición en el *ranking* después de las transferencias y, en el caso de Alemania, son los *Länder* de Schleswig-Holstein y Baja Sajonia los que no cambian de posición. Ahora bien, la tasa de cambio no mide la intensidad de los cambios de posición, que es lo que proporciona la medida de la distancia en términos de la reordenación máxima entendida como inversión del orden. Pues bien, según esta medida, mientras en España se produce una reordenación del 65,2% sobre la máxima, en Alemania es del 47,2%. Esta mayor intensidad de la reordenación en España viene motivada por el gran cambio de posición que experimentan CC.AA. como Baleares, que pierde 13 posiciones, Extremadura, que gana 12, Madrid que pierde 10, Canarias que gana 9 o Cataluña o C. Valenciana que pierden 8 posiciones. En Alemania, por su parte, los Länder que mayor reordenación experimentan son Brandemburgo y Sajonia-Anhalt al perder y ganar respectivamente 11 posiciones, Baden-Württemberg que pierde 9 posiciones o Sajonia que gana 8. Por su parte, el coeficiente de correlación de Spearman nos indica que mientras para Alemania los órdenes inicial y final no están correlacionados, en España el valor negativo indica una tendencia hacia la inversión del orden.

En cuanto al sentido de la reordenación, ascendente versus descendente, en Alemania prácticamente la descomposición es al 50%, mientras que en España hay un cierto predominio de la reordenación descendente (con el 56,6%). En Alemania, los Länder que experimentan mayor cambio de posición compensan el sentido de la misma. Así, la reordenación de Brandemburgo y Sajonia-Anhalt se compensan al perder y ganar respectivamente 11 posiciones; también Baden-Württemberg y Sajonia al perder 9 y ganar 8 posiciones respectivamente; o los cambios de posición de Baviera y Hesse, que pierden 7 posiciones, se compensan con las ganancias de 7 posiciones de Berlín y Meclemburgo-Antepomerania. En España, sin embargo, mientras que las dos CC.AA. que más posiciones ganan y pierden respectivamente compensan en gran medida el sentido de la reordenación, ya la pérdida de 8 posiciones de Cataluña o C. Valenciana no está compensada por reordenación ascendente de la misma intensidad.

Cuadro 2.- España. Año 2006. Ordenación de las CC.AA. de régimen común. Orden de mayor a menor recursos per cápita.

	Financiación per cápita. En euros.		Ordenación		
	Antes transferencias	Después transferencias	Antes transferencias	Después transferencias	Reordenación
I. Baleares	2.635	2.522	1	14	-13
Madrid (Comunidad de)	2.558	2.670	2	12	-10
Cataluña	2.274	2.806	3	11	-8
Aragón	2.031	3.018	4	4	0
Cantabria	1.937	3.394	5	1	4
Rioja (La)	1.888	3.237	6	2	4
Comunidad Valenciana	1.863	2.521	7	15	-8
Asturias (Principado de)	1.687	2.962	8	8	0
Murcia (Región de)	1.675	2.620	9	13	-4
Castilla y León	1.646	3.013	10	6	4
Castilla - La Mancha	1.566	2.947	11	9	2
Andalucía	1.499	2.809	12	10	2
Galicia	1.462	2.984	13	7	6
Canarias	1.205	3.014	14	5	9
Extremadura	1.157	3.192	15	3	12

Fuente: Elaboración propia a partir de Ministerio de Economía y Hacienda "Financiación de las Comunidades Autónomas por los impuestos cedidos, fondo de suficiencia y garantía de financiación de los servicios de asistencia sanitaria, correspondiente al ejercicio 2006", Instituto de Estudios Fiscales, BADESPE (Base de Datos del Sector Público español).

Cuadro 3.- Alemania. Año 2006. Ordenación y reordenación de los Länder. Orden de mayor a menor recursos per cápita.

	Financiación per cápita. En euros.		O		
	Antes transferencias	Después transferencias	Antes transferencias	Después transferencias	Reordenación
Hamburgo	2.344	3.202	1	3	-2
Brandemburgo	1.546	1.882	2	13	-11
Baden-Württemberg	1.533	1.998	3	12	-9
Baviera	1.523	2.012	4	11	-7
Renania del Norte-Westfalia	1.351	2.260	5	7	-2
Bremen	1.328	4.904	6	1	5
Hesse	1.199	1.623	7	14	-7
Renania Palatinado	1.176	2.570	8	5	3
Schleswig-Holstein	1.176	2.101	9	9	0
Baja Sajonia	1.155	2.034	10	10	0
Berlín	1.094	2.663	11	4	7
Sarre	622	1.311	12	16	-4
Sajonia-Anhalt	540	4.166	13	2	11
Sajonia	522	2.480	14	6	8
Meclemburgo-Antepomerania	502	2.167	15	8	7
Turingia	493	1.611	16	15	1

Fuente: Elaboración propia a partir de Oficina estadística federal, "Statistisches Jahrbuch für die Bundesrepublik Deutschland und für das Ausland 2007 CD-ROM"; Oficina federal estadística, base de datos online GENESIS.

4. CONCLUSIONES

La pérdida de igualación de las transferencias que provoca la reordenación es mucho mayor en Alemania, lo que implica que la reducción efectiva de la desigualdad inicial es de mucha mayor magnitud en España que en Alemania, con unas tasas de Pechman-Okner de 72% y 24% respectivamente.

En cuanto a la intensidad de la reordenación, tomando como referencia de máxima reordenación la inversión del orden inicial, llega a situarse en prácticamente 2/3 en España y no llega a 0,5 en Alemania, lo cual implica una cierta tendencia a la inversión del orden en España como corrobora el coeficiente de correlación de Spearman de valor negativo. En efecto, en

España, son las 3 CC.AA. de mayores y menores ingresos tributarios per cápita respectivamente las que experimentan una mayor reordenación descendente y ascendente, de forma que cuanto más cerca se sitúa de los extremos del *ranking* mayor reordenación experimentan. Además la reordenación descendente de las 3 CC.AA. con mayores ingresos tributarios es de mayor intensidad que la ascendente que experimentan las de menores recursos tributarios, lo cual explica básicamente el predominio de la reordenación descendente sobre la ascendente.

En Alemania, por su parte, aunque es cierto que la reordenación descendente /ascendente tiende a concentrarse en los *Länder* con mayores/menores recursos tributarios, no de manera que cuanta más cercanía a los extremos mayor reordenación, lo cual explica el hecho de que el coeficiente de Spearman no registra una tendencia a invertir el orden. Asimismo la intensidad de la reordenación ascendente *versus* descendente es prácticamente igual.

BIBLIOGRAFÍA

AARON, H.J. Y BOSKINS, M.J. (eds) (1980), *The Economics of Taxation* (Washington, DC, Brookings), pp. 3–18.

ARONSON, J.R. Y LAMBERT, P.J. (1994), "Decomposing the Gini Coefficient to Reveal the Vertical, Horizontal and Reranking Effects of Income Taxation", *National Tax Journal*, 47, 2, pp. 273–294.

ARONSON, J.R., JOHNSON, P. Y LAMBERT, P.J. (1994), "Redistributive Effect and Unequal Income Tax Treatment", *Economic Journal*, 104, March, pp. 262–270.

ATKINSON, A.B. (1980), "Horizontal Equity and the Distribution of the Tax Burden", en H.J. Aaron & M.J. Boskins (eds), *The Economics of Taxation* (Washington, DC, Brookings), pp. 3–18.

BARTHOLOMEW, D. (1973), *Stochastic Models for Social Processes.* London: Wiley.

CARSON, M. Y CREEDY, J. (eds.), (1993), *Industrial Concentration and Economic Inequality*, Edward Elgar.

CHAKRAVARTY, S. R., DUTTA, B. AND WEYMARK, J. A. (1985), "Ethical indices of income mobility", *Social Choice and Welfare,* 2 (1), pp. 1–21.

DICKEY H. (2001), "Regional earnings inequality in Great Britain: a decomposition analysis", *Regional. Studies,* 35, pp. 605–612.

DUCLOS, J.Y. (1993), "Progressivity, Redistribution, and Equity, with Application to the British Tax and Benefit System", *Public Finances/Finances Publiques*, 48(3), pp. 350–365.

FELDSTEIN, M.S. (1976), **"**On the Theory of Tax Reform", *Journal of Public Economics 6* (July-August), pp. 77-104.

FIELDS, G. Y OK, E. (1996), "The meaning and measurement of income mobility", *Journal of Economic Theory,* 71, pp. 349–77.

FIELDS, G. Y OK, E.: (1999), "Measuring movement of incomes", *Economica,* 66, pp. 455–72.

GINI, C. (1912), *Variabilità e mutabilità,* Studi Economico-giuridici, Università di Cagliari, III.

HEILM, M. Y WODON, Q.T. (1997), "Inequality in CO2 Emissions Between Poor and Rich Countries", *The Journal of Environment & Development*, Vol. 6, No. 4, pp. 426-452.

HIERRO, L.A., ATIENZA, P., PATIÑO, D. (2007), "Inequality and progressiveness in the distribution of revenues of the states in federal countries. A comparative study", Working Paper 07-03, International Studies Program, Andrew Young School of policy studies, Georgia State University.

KAKWANI, N.C. (1984), "On Measurement of Taxes, Progressivity and Redistributive Effect of Taxes with Applications to Horizontal and Vertical Equity", *Advances in Econometrics*, 3, pp. 149–168.

KING, M. A. (1983), "An index of inequality with applications to horizontal equity and social mobility", *Econometrica,* 51, pp. 99–115.

LERMAN, R. Y YITZHAKI, S. (1995), "Changing Ranks and the Inequality Impacts of Taxes and Transfers", *National Tax Journal*, 48, pp. 45-59.

MARKANDYA, A. (1982), "Intergenerational exchange mobility and economic welfare", *European Economic Review*, 17, pp. 307–24.

MARTIN DAVID AND TIMOTHY SMEEDING (EDS.) (1984), *Horizontal Equity, Uncertainty and Economic Wellbeing*, National Bureau of Economic Research, 1985, pp. 239-263.

PLOTNICK, R. (1981), "A Measure of Horizontal Inequity", *Review of Economics and Statistics*, 63, 2, pp. 283–288.

PLOTNICK, R. (1982), "The concept and measurement of horizontal inequity", *Journal of Public Economics* 17, pp. 373-91.

PLOTNICK, R., (1984), "A comparison of measures of horizontal inequity", in Martin David and Timothy Smeeding (eds.), *Horizontal Equity, Uncertainty and Economic Wellbeing*, National Bureau of Economic Research, 1985, pp. 239-263.

PRAIS, S.J., (1955), "Measuring Social Mobility", *Journal of the Royal Statistical Society. Series A (General)*, Vol. 118, No. 1, pp. 56-66.

SALAS, R., (1999), "Convergencia, movilidad y redistribución interterritorial en España: 1981-1996", *Papeles de Economía Española*, 80, pp. 19-28.

SCHILLER, R., (1977), "Relative Earnings Mobility in the United States", *The American Economic Review*, Vol. 67, No. 5. (Dec., 1977), pp. 926-941.

SCOTT, C. D. Y LITCHFIELD, J. A., (1994), "Inequality, mobility and the determinants of income among the rural poor in Chile, 1968–1986", Discussion Paper 53, Development Economics Research Programme, London School of Economics.

SHORROCKS, A.F., (1978), "The measurement of mobility", *Econometrica* 46, pp. 1013–1024.

SHORROCKS, A.F, (1993), "On the Hart measure of income mobility", en Carson, M. y Creedy, J. (eds.), *Industrial Concentration and Economic Inequality*, Edward Elgar.

URBAN, I. Y LAMBERT, P.J., (2005), "Redistribution, Horizontal Inequity and Reranking: How to Measure them Properly", Economics Discussion Paper, No. 2005-12, University of Oregon.

LOS DETERMINANTES DEL EMPLEO DOMÉSTICO: INMIGRACIÓN Y MERCADO LABORAL. EL CASO DE LA REGIÓN DE MADRID.

JAVIER MORILLAS
E-mail: morgom@ceu.es
Departamento de Economía Aplicada
Universidad San Pablo-CEU

Resumen

Los servicios a los hogares en el ámbito del empleo doméstico han presentado en la última fase del ciclo expansivo de la economía mundial, y hasta 2008 una extraordinaria vitalidad en los países desarrollados. Y especialmente en las denominadas *ciudades globales*, como Roma, Londres, París, o Madrid. Concretamente en la región de Madrid han posibilitado una masiva incorporación de la mujer al mercado de trabajo, y a su vez han retroalimentado una mayor demanda de actividad y empelo en el resto de sectores productivos de más alto valor añadido y contenido tecnológico.
En la ponencia que se presenta se intenta analizar cuáles son los determinantes de la contratación de empleo doméstico por parte de los hogares, intentando una modelización de la misma. Por un lado, basándose en el análisis de la fuerte presión de la oferta, con una abundante y dispar mano de obra inmigrante, en cuanto a procedencia, cualificación, nivel salarial, propensión a la inversión, a remesar, etc. Y por otro lado una alta demanda, condicionada por un conjunto de determinadas variables; no sólo de la renta, que ni siquiera es el más determinante respecto a otras economías regionales en el caso de Madrid, sino de la actividad profesional, el número de hijos en las familias, la ocupación o no de los dos miembros de la pareja, la edad, el sistema regional de transporte, etc y que nos explican una gran parte de la contratación. Una contratación de mano de obra poco cualificada, pero que permite en las fases alcistas del ciclo económico un efecto multiplicador, concretado en una mayor contratación de mano de obra femenina local, y ésta en actividades de más alta cualificación, que favorece una creciente expansión económica y de las actividades de mayor contenido tecnológico.

Palabras Clave: Empleo doméstico. Inmigración. Economía regional. Mercado de trabajo. Determinantes de la contratación. Región de Madrid.
Área Temática: Economía Urbana, regional y local

1. INTRODUCCIÓN

Los servicios a los hogares en el ámbito del empleo doméstico han presentado en la última fase del ciclo expansivo de la economía mundial, y hasta 2008 una extraordinaria vitalidad en los países desarrollados. Y especialmente en las denominadas *ciudades globales*, como Roma, Londres, París, o Madrid. Concretamente en la región de Madrid han posibilitado una masiva incorporación de la mujer al mercado de trabajo, y a su vez han retroalimentado una mayor demanda de actividad y empelo en el resto de sectores productivos de más alto valor añadido y contenido tecnológico.

En la ponencia que se presenta se intenta analizar cuáles son los determinantes de la contratación de empleo doméstico por parte de los hogares, intentando una modelización de la misma. Por un lado, basándose en el análisis de la fuerte presión de la oferta, con una abundante y dispar mano de obra inmigrante, en cuanto a procedencia, cualificación, nivel salarial, propensión a la inversión, a remesar, etc. Y por otro lado una alta demanda, condicionada por un conjunto de determinadas variables; no sólo de la renta, que ni siquiera es el más determinante respecto a otras economías regionales en el caso de Madrid, sino de la actividad profesional, el número de hijos en las familias, la ocupación o no de los dos miembros de la pareja, la edad, el sistema regional de transporte, etc y que nos explican una gran parte de la

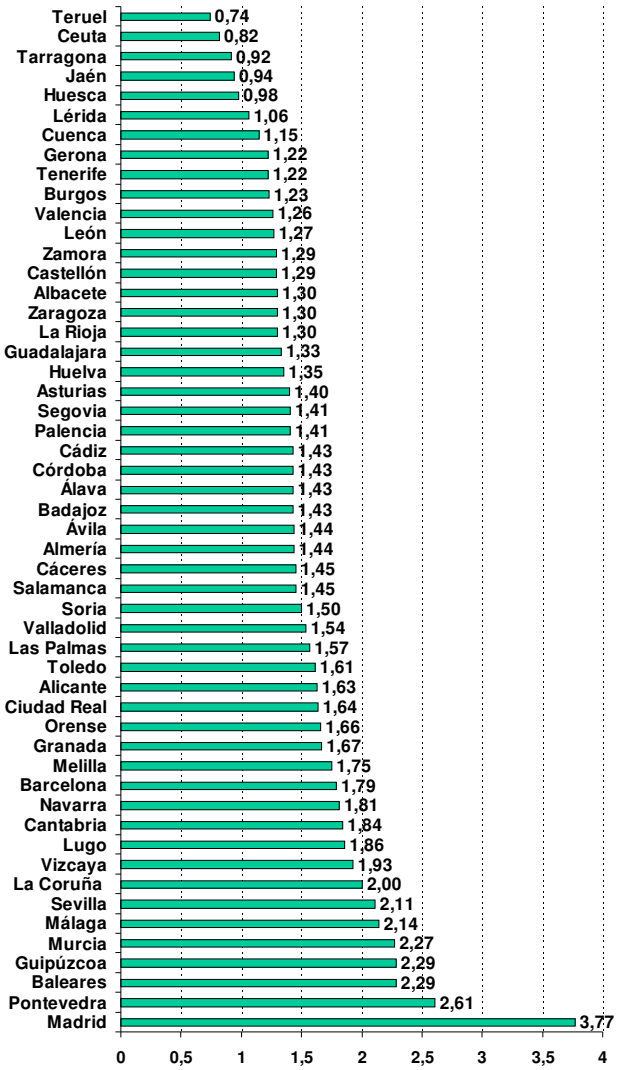
contratación. Una contratación de mano de obra poco cualificada, pero que permite en las fases alcistas del ciclo económico un efecto multiplicador, concretado en una mayor contratación de mano de obra femenina local, y ésta en actividades de más alta cualificación, que favorece una creciente expansión económica y de las actividades de mayor contenido tecnológico.

Para ello hemos recurrido tanto al análisis de las fuentes estadísticas disponibles, principalmente del Instituto Nacional de Estadística y Censo de Población, así como al mayor esfuerzo de encuestación –casi tres mil personas, entre empleados y empleadoresllevado a cabo hasta la fecha para el análisis del mercado de trabajo del empleo doméstico en un país desarrollado.

2. NIVELES DE EMPLEO EN SERVICIO DOMÉSTICO EN ESPAÑA, ¿ES LA RENTA UN DETERMINANTE CONCLUYENTE?

La información del Censo 2001 permite concluir que el número de empleados de servicio doméstico sobre el número total de hogares registra notables variaciones según las provincias españolas, según se muestra en el Gráfico 1. Existen cinco provincias (Teruel, Ceuta, Tarragona, Jaén y Huesca) en las que esta relación toma un valor inferior a 1, mientras que en el extremo opuesto se sitúan ocho cuyo valor es superior a 2. Madrid es un caso extremo, ya que esta relación toma el valor máximo y muy superior al resto (3,8).

Gráfico 1. Número de empleados domésticos por provincias como porcentaje del número de hogares. Censo de Población 2001.



Fuente: Elaboración propia con datos INE

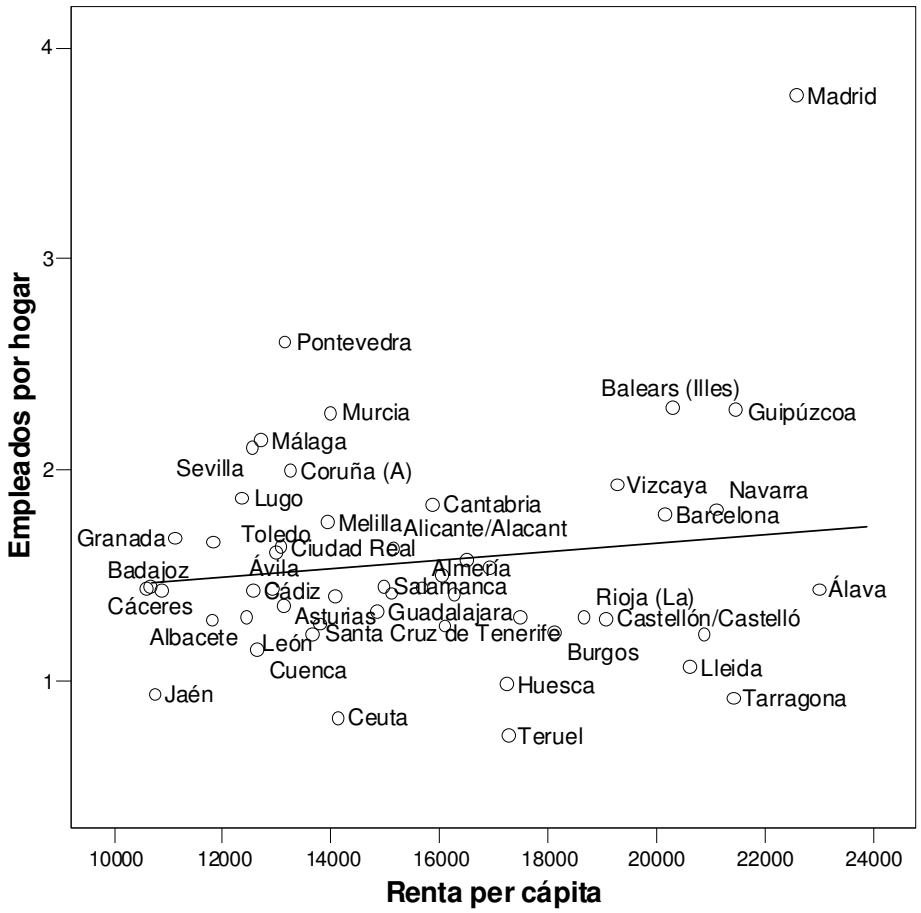
Una explicación posible del distinto número de empleados domésticos entre las regiones españolas es el nivel de renta per cápita, ya que a este tipo de servicios acceden principalmente los hogares con mayor nivel económico. Sorprendentemente, el análisis estadístico realizado, cuyos resultados se muestran en la Tabla 1, permite concluir que la renta per cápita media provincial no tiene prácticamente ninguna capacidad explicativa del número relativo de empleados de servicio doméstico. El Gráfico 2 permite observar la ausencia de relación entre ambas variables.

Tabla 1. Resultados de la regresión: Empleo servicio doméstico-renta per capita provincial.

	Coeficientes no estandarizados	
	B	t
(Constante)	1,206	3,745
Renta	,000023	1,115

R cuadrado = 0,024
Variable dependiente: Empleados por hogar
Programa estadístico utilizado SPSS. Nivel de significación del 95 por ciento
Fuente: Elaboración propia.

Gráfico 2. Relación entre empleados en servicio doméstico y nivel de renta per capita provincial



Fuente: Elaboración propia con datos INE

También se ha analizado la relación existente entre el nivel de empleo doméstico provincial y el grado de implantación de las principales categorías profesionales[1]. En este sentido, un mayor nivel de empleo doméstico se relaciona positivamente con el número relativo de Técnicos profesionales científicos e intelectuales, y en menor medida y de forma negativa con el de personal de las Fuerzas armadas[2]. Otra variable que también tiene capacidad explicativa del nivel de desarrollo del servicio doméstico es el porcentaje de hogares compuestos por personas mayores de 65 años. Los principales estadísticos calculados se muestran en la Tabla 2.

Tabla 2. Posibles variables explicativas del nivel de servicio doméstico en España.

		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ	Beta		
1	(Constante)	-,001	,012		-,104	,918
	Renta	-0,00000008	,000	-,558	-2,284	,028
	0 Fuerzas armadas	-,079	,028	-,486	-2,804	,008
	1 Dirección de las empresas y de las administraciones públicas	,082	,078	,211	1,058	,297
	2 Técnicos y profesionales científicos e intelectuales	,098	,048	,488	2,059	,047
	3 Técnicos y profesionales de apoyo	,054	,060	,278	,903	,372
	4 Empleados de tipo administrativo	,092	,063	,404	1,454	,155
	5 Trabajadores de los servicios de restauración, personales, protección y vendedores de los comercios	,003	,029	,020	,103	,919
	6 Trabajadores cualificados en la agricultura y en la pesca	,010	,019	,090	,547	,588
	7 Artesanos y trabajadores cualificados de las industrias manufactureras, la construcción, y la minería, excepto los operadores de instalaciones y maquinaria	,030	,030	,209	1,004	,322
	8 Operadores de instalaciones y maquinaria, y montadores	-,020	,021	-,178	-,921	,363
	9 Trabajadores no cualificados	,016	,016	,137	1,036	,307
	Tasa actividad femenina	-,017	,021	-,206	-,837	,408
	Tasa actividad masculina	,002	,023	,024	,100	,921
	% Mujer sola más de 65 años	-,141	,070	-,502	-2,032	,050
	% Hombre solo más de 65 años	,207	,133	,387	1,562	,127

Variable dependiente: Empleados por hogar
R cuadrado = 0,60
Fuente: Elaboración propia.

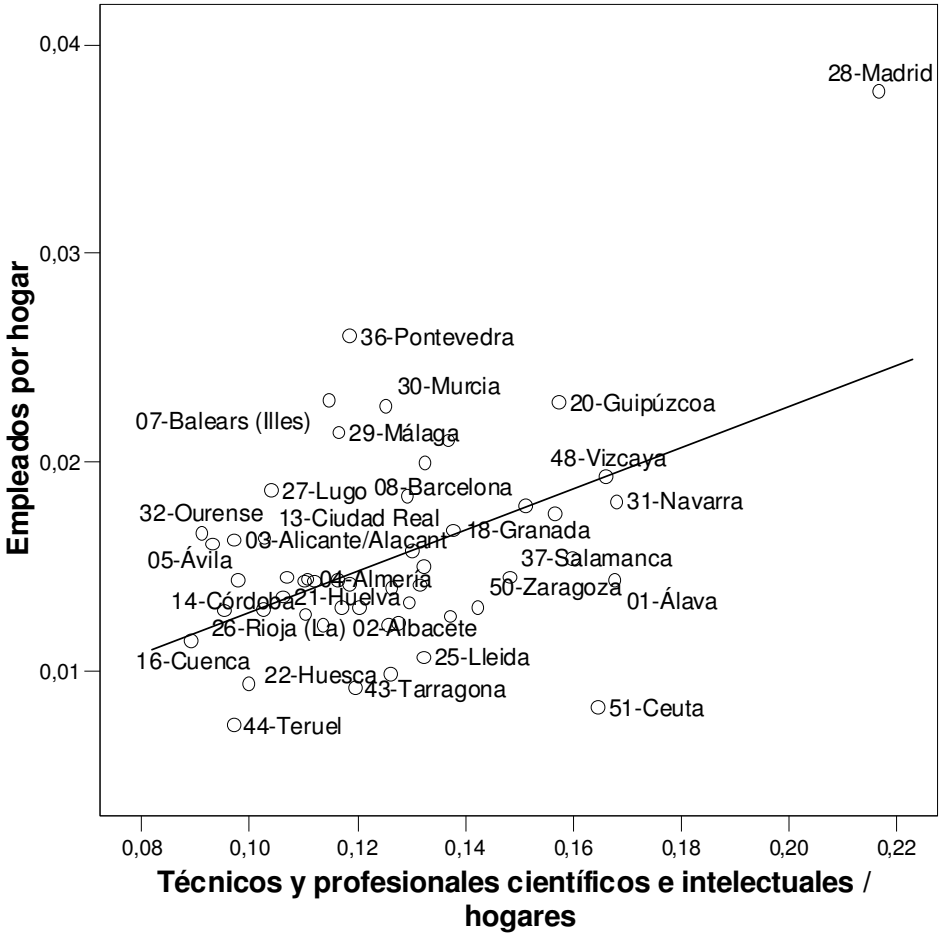
[1] Calculado como: Número de empleados de categoría "i" en provincia "n"/ Número de hogares en provincia "n".
[2] Esta relación se debe en gran parte a la influencia de los valores de Ceuta y Melilla

Tabla 3. Relación entre servicio doméstico y Técnicos profesionales y científicos e intelectuales

	Coeficientes no estandarizados	
	B	t
(Constante)	0,03	1,270
02Técnicos y profesionales científicos e intelectuales	0,091	3,591

R cuadrado = 0,205
Variable dependiente: Empleados por hogar
Programa estadístico utilizado SPSS. Nivel de significación del 95 por ciento
Fuente: Elaboración propia.

Gráfico 3. Relación entre empleados en servicio doméstico y número relativo de Técnicos y profesionales científicos e intelectuales.



Fuente: Elaboración propia con datos INE

3. ANÁLISIS DEL GASTO REALIZADO EN SERVICIO DOMÉSTICO POR LOS HOGARES ESPAÑOLES A PARTIR DE LOS MICRODATOS DE LA ENCUESTA DE PRESUPUESTOS FAMILIARES (EPF)

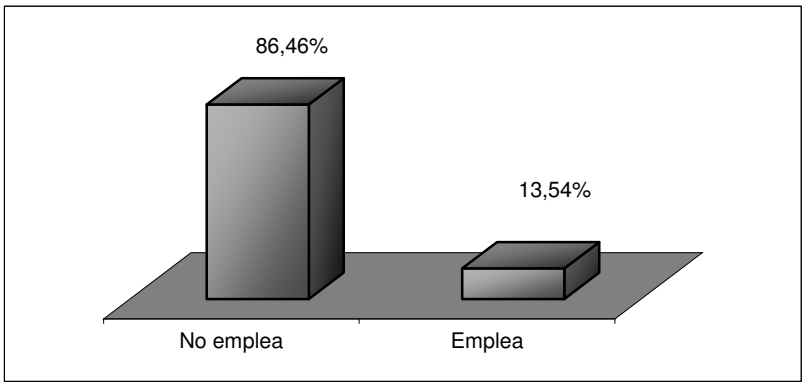
La Encuesta de Presupuestos Familiares nos aporta una información muy exhaustiva a la hora de conocer las características de la población y el gasto que dedican los hogares españoles al empleo de servicio doméstico.

Dicha encuesta nos ofrece los datos anuales de 2002 (microdatos longitudinales) que vamos a valorar a continuación. El análisis trimestral de años más recientes (segundo trimestre de 2004) será desarrollado en una fase más avanzada del proyecto. Por otra parte, también en una segunda fase, será objeto de nuestro estudio la evolución o el análisis dinámico desde años anteriores, con el fin de extraer tendencias.

A continuación, de la lectura de los Gráfico 4 y Gráfico 5, observamos que la CAM es la Comunidad con mayor porcentaje de empleo doméstico, con una tasa de penetración del 22,2 por ciento comparando dicha cifra con la media nacional, que resulta un 13,5 por ciento. Los datos extraídos de la EPF, nos ofrecen la posibilidad de conocer como se distribuye el empleo de este sector, colocando a las regiones del Sur y Canarias después de Madrid; además, nos parece relevante destacar que las regiones del Norte, empleen servicio doméstico en un 25 por ciento menos que las anteriores. No nos cabe la menor duda, que existen otros factores, bien culturales o socioeconómicos, que debemos analizar para conocer esta realidad.

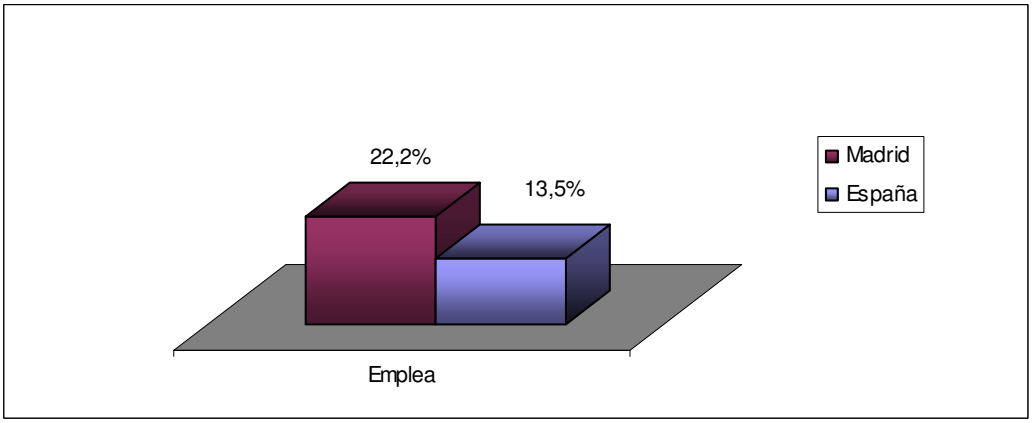
Gráfico 6, a su vez, ofrece el gasto medio anual de los hogares según los NUTS1, indicando de nuevo la diferencia de la partida destinada al servicio doméstico de la CAM con el resto de España, con una diferencia de más del doble entre la zona Noroeste y Madrid.

Gráfico 4. Porcentaje de empleo doméstico en los Hogares Españoles



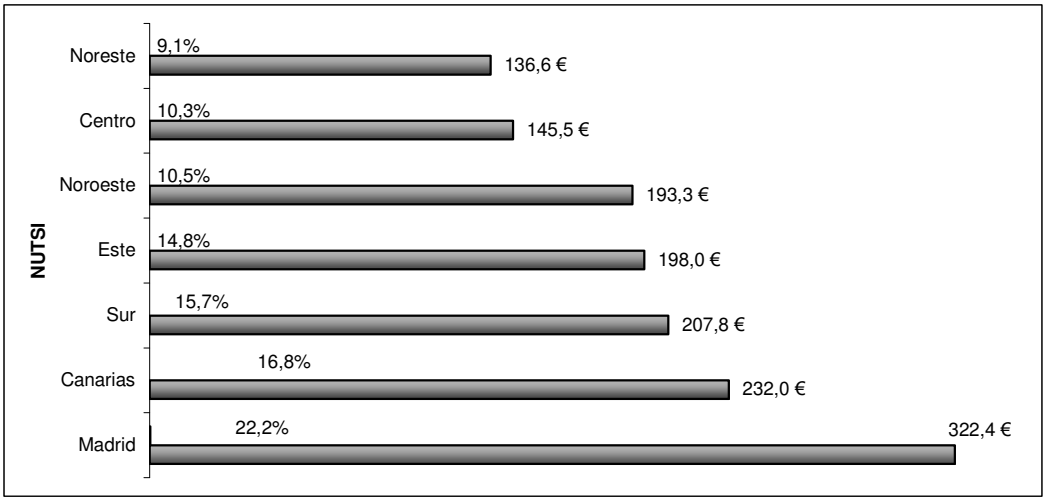
Fuente: Elaboración propia a partir de microdatos ECPF 2002

Gráfico 5. Porcentaje de empleo en servicio doméstico en hogares de España y Madrid.



Fuente: Elaboración propia a partir de microdatos ECPF 2002

Gráfico 6. Empleo de servicio doméstico comparado según NUTS1(a)



Fuente: Elaboración propia a partir de microdatos ECPF 2002

(a) NUTS1:

1 Noroeste (Comunidades de Galicia, Asturias y Cantabria).
2 Noreste (Comunidades de País Vasco, Navarra, La Rioja y Aragón).
3 Madrid (Comunidad de Madrid).
4 Centro (Comunidades de Castilla y León, Castilla-La Mancha y Extremadura).
5 Este (Comunidades de Cataluña, Comunidad Valenciana e Islas Baleares).
6 Sur (Comunidades de Andalucía, Región de Murcia y Ceuta y Melilla).
7 Canarias.

En las tablas siguientes, se detallan los resultados obtenidos de la explotación de los gastos medios anuales de los hogares españoles (Tabla 4), de la CAM (Tabla 5) y de los NUTS1 (Tabla 6), así como la partida de gasto medio en servicio doméstico, en euros.

Tabla 4. Gasto en servicio doméstico en ámbito nacional

ESPAÑA		N	Mínimo	Máximo	Media	Desv. típ.
	GASTO	1.248	301.048	14.779.613	3.102.973,13	1.885.7,71,918
	GASTOSERDOME	1.248	138	1.626.454	142.214,46	159.730,466
	N válido (según lista)	1.248				

Fuente: Elaboración propia a partir de microdatos ECPF 2002

Tabla 5. Gasto en servicio doméstico en Comunidad de Madrid

MADRID		N	Mínimo	Máximo	Media	Desv. típ.
	GASTO	161	644.093	12.681.058	3.756.742,06	2.187.592,986
	GASTOSERDOME	161	3.216	923.943	145.393,71	164.520,508
	N válido (según lista)	161				

Fuente: Elaboración propia a partir de microdatos ECPF 2002

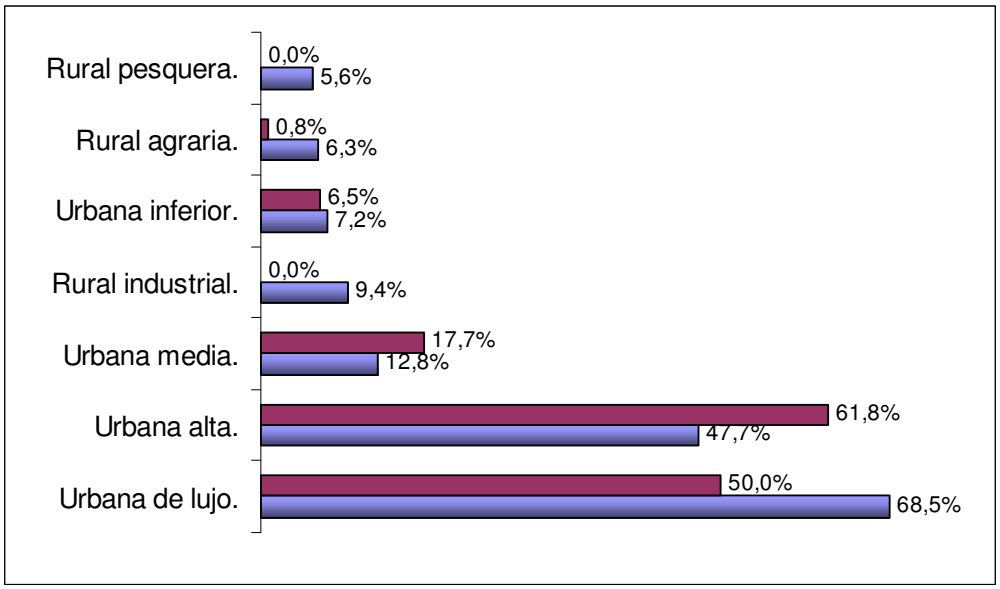
Tabla 6. Gasto en servicio doméstico en áreas geográficas

NUTS1		N	Mínimo	Máximo	Media	Desv. típ.
Noroeste	GASTO	1.375	186.514	9.686.606	2.095.307,82	1.242.972,391
	GASTOSERDOME	1.375	0	1.626.454	19.325,14	90.222,614
	N válido (según lista)	1.375				
Noreste	GASTO	1.401	159.646	7.773.708	2.345.558,96	1.214.689,253
	GASTOSERDOME	1.401	0	804.916	13.657,15	59.107,598
	N válido (según lista)	1.401				
Madrid	GASTO	726	335.236	12.681.058	2.522.774,52	1.559.290,849
	GASTOSERDOME	726	0	923.943	32.242,96	98.116,020
	N válido (según lista)	726				
Centro	GASTO	1.486	163.961	10.335.954	1.786.767,71	1.186.154,541
	GASTOSERDOME	1.486	0	1.009.457	14.552,01	61.641,208
	N válido (según lista)	1.486				
Este	GASTO	2.200	167.513	8.687.559	2.147.584,19	1.216.095,011
	GASTOSERDOME	2.200	0	1.161.703	19.800,30	78.707,613
	N válido (según lista)	2.200				
Sur	GASTO	1.575	188.850	14.779.613	2.097.488,22	1.385.825,289
	GASTOSERDOME	1.575	0	682.494	20.775,54	71.705,703
	N válido (según lista)	1.575				
Canarias	GASTO	451	194.525	14.727.576	1.979.102,81	1.393.889,135
	GASTOSERDOME	451	0	692.543	23.199,78	81.839,648
	N válido (según lista)	451				

Fuente: Elaboración propia a partir de microdatos ECPF 2002

En el caso de estudiar las zonas, tanto urbanas como rurales, en las que divide la EPF el panorama nacional, se observa que el empleo del servicio doméstico se sitúa en zonas urbanas de lujo en el 68, 5 por ciento, podemos ver como en la CAM el mayor porcentaje, el 61,8 por ciento, se obtiene en la zona urbana alta.

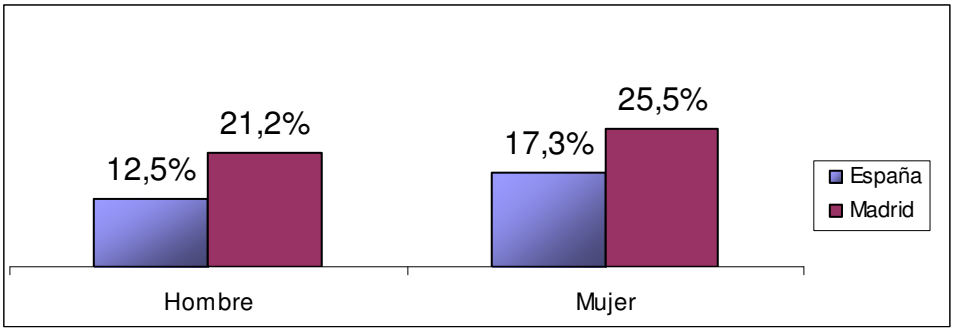
Gráfico 7. Porcentaje de hogares con servicio doméstico según zonas



Fuente: Elaboración propia a partir de microdatos ECPF 2002

En el gráfico siguiente podemos observar que los hogares que emplean y analizan la situación, en función del sexo del cabeza de familia; indicando, que en el caso de ser mujer, aumenta la necesidad de servicio doméstico y por este motivo se eleva el porcentaje de empleo 21.2 por ciento; la situación en el caso de la CAM, es mayor el grado de incidencia de la variable 25,5 por ciento.

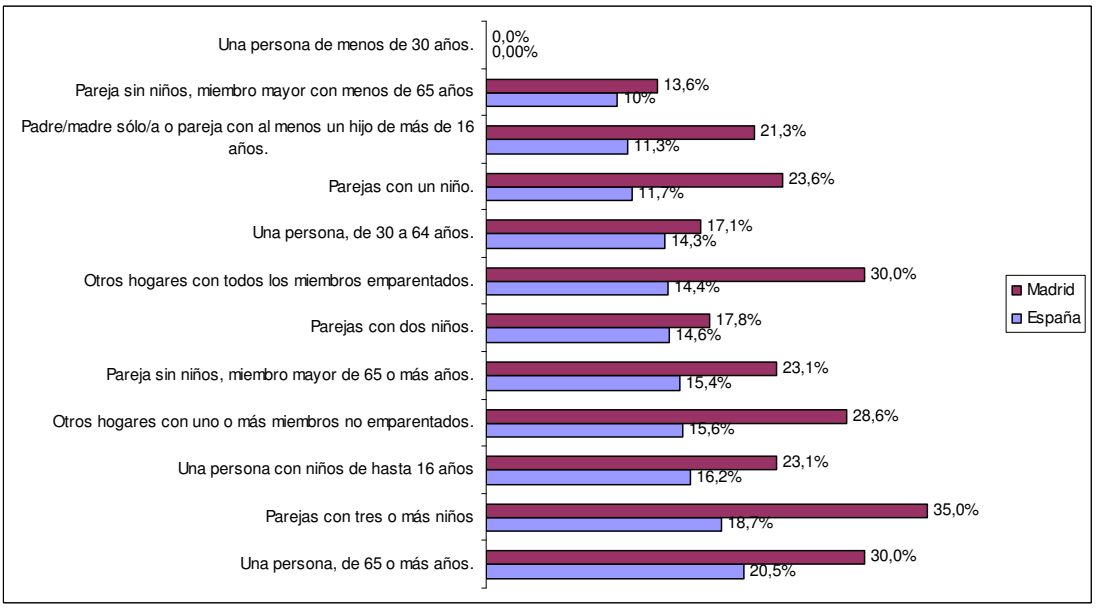
Gráfico 8. Porcentaje de empleadores según el sexo del cabeza de familia



Fuente: Elaboración propia a partir de microdatos ECPF 2002

Resulta interesante conocer los tipos de hogar que emplean servicio doméstico, en este caso, existe una clara diferencia entre los hogares madrileños y el resto de España, a nivel nacional el porcentaje más elevado, el 20,5 por ciento representa hogares de más de 65 años, y en la CAM, el tipo de hogar que más emplea es el de parejas con tres o más niños. Sin descuidar que los hogares de más de 65 años representan el segundo colectivo en números.

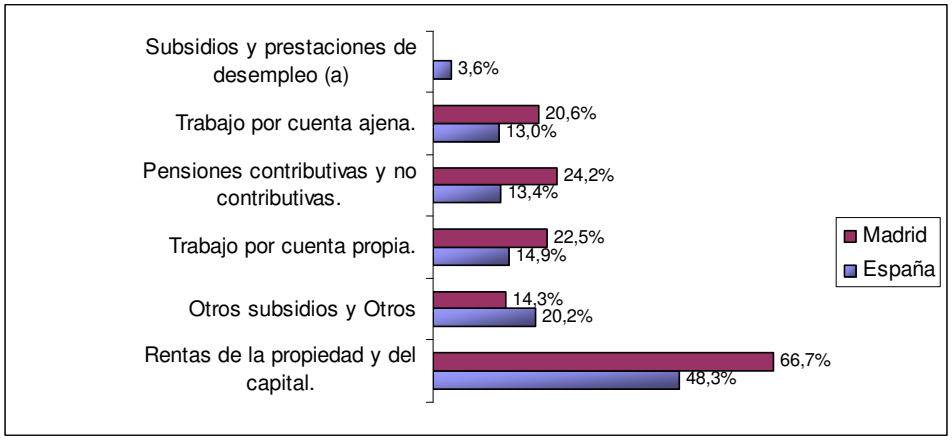
Gráfico 9: Empleadores de servicio doméstico según tipología de hogares



Fuente: Elaboración propia a partir de microdatos ECPF 2002

Otra variable a tener en cuenta sería la fuente de ingresos, en el Gráfico 10 observamos como los empleadores que poseen rentas de la propiedad y del capital contratan en mayor porcentaje que personas que perciben otras fuentes de ingreso.

Gráfico 10. Porcentaje de empleadores de servicio doméstico en función de la fuente de ingresos

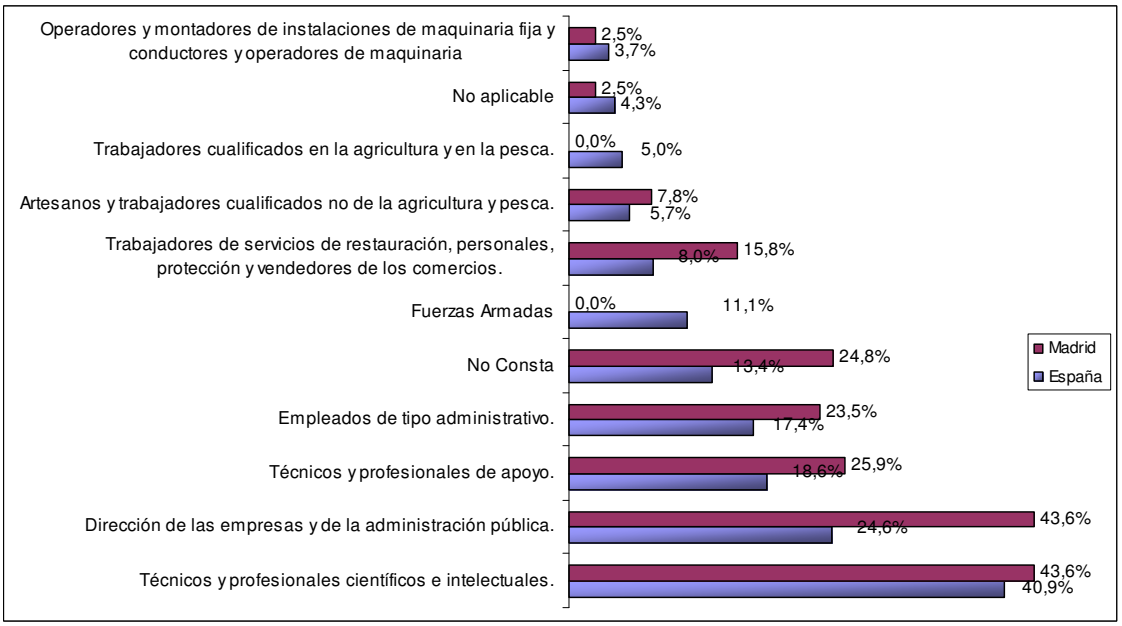


(a) **En el caso de Madrid no es representativo**
Fuente: Elaboración propia a partir de microdatos ECPF 2002

Por último, a tenor de este primer acercamiento a la ECPF, podemos analizar qué tipo de ocupación y qué partida del gasto de los hogares, destinan al empleo doméstico; además podemos abundar en el conocimiento del perfil del empleador según que su actividad se desarrolle en el sector público o privado.

El Gráfico 11 muestra que los directivos de empresas, los técnicos y profesionales son las profesiones que más emplean. Es relevante indicar que en la CAM, el porcentaje se repite, frente al resto de España. Pensamos que a nivel nacional existen pequeñas empresas y el número de directivos es superior. El número de grandes empresas situadas en la CAM, hace que las categorías (1) y (2) se igualen con un 43,6 por ciento. Indicar que la categoría No consta(88), se debe asimilar a las personas jubiladas, mientras que las categorías (00) y (06) no son representativas para el caso de Madrid).

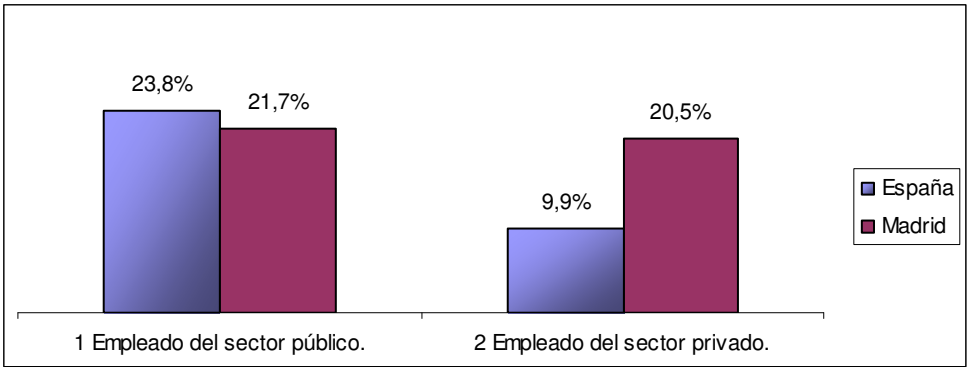
Gráfico 11. Empleadores de servicio doméstico en función de la ocupación



Fuente: Elaboración propia a partir de microdatos ECPF 2002

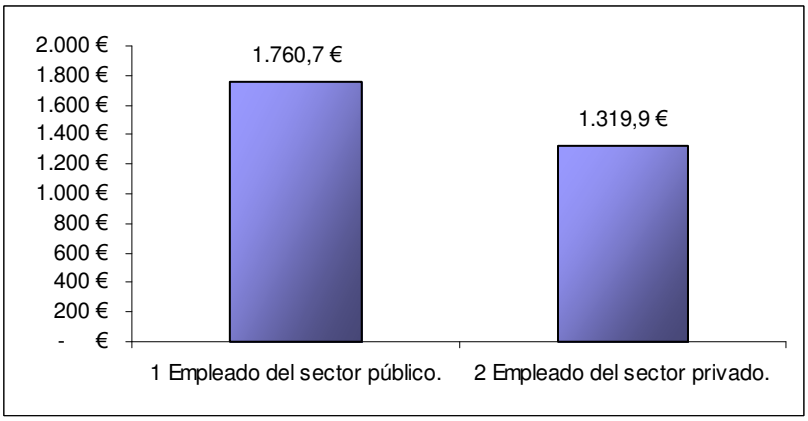
En el Gráfico 12 llama la atención, que las personas ocupadas en el sector público emplean en mayor proporción que en la CAM. A su vez, observamos que la partida de gasto medio de los empleadores, que trabajan en el sector público es de 1.760,7 €uros anuales, frente a los 1.319,9 Euros de la CAM.

Gráfico 12. Porcentaje de empleadores en función de su actividad



Fuente: Elaboración propia a partir de microdatos ECPF 2002

Gráfico 13. Gasto anual en servicio doméstico según su actividad



Fuente: Elaboración propia a partir de microdatos ECPF 2002

4. RED DE INFRAESTRUCTURAS DE TRANSPORTE Y EMPLEO EN SERVICIO DOMÉSTICO.

Al explotar los datos del último Censo de Población en España (el de 2001), se ha utilizado el criterio del municipio de trabajo del empleado doméstico, ya que dentro de la Comunidad es frecuente que la persona trabaje en un municipio distinto del de residencia.

En la Tabla 7 llama la atención cómo el 60 por ciento de los empleados de hogar sólo realiza un desplazamiento diario en la Comunidad de Madrid.

Tabla 7. Número de desplazamientos diarios de los empleados de hogar en la CAM.

Número de viajes diarios	TOTAL	Ninguno	Uno	Dos o más diarios	No es aplicable
Personas	70.726	2.158	42.524	15.417	10.627
%		3,05	60,12	21,80	15,03

Fuente: Elaboración propia con datos del INE 2004. Censos de Población y Viviendas 2001.

En la Tabla 8 resulta significativo que más del 7 por ciento de las empleadas de hogar llegue a su trabajo en menos de 10 minutos; lo que nos estaría indicando una situación en la que encuentran trabajo cerca de su domicilio, sólo posible en determinados distritos donde, a pesar de las previsibles diferencias de renta, se concentran al tiempo empleados y empleadores. Por otra parte más del 60 por ciento de los empleadas de hogar se desplaza a su lugar de trabajo en unos tiempos que oscilan entre los 10 y los 60 minutos. Algo más del 11 por ciento necesitaría entre 60 y 90 minutos, y apenas un 2´31 por ciento de las empleadas de hogar necesitarían más de hora y media para llegar a su puesto de trabajo.

Tabla 8. Tiempo de desplazamiento.

Tiempo de desplazamiento	*Personas*	%
TOTAL	70.726	
Menos de 10 minutos	5.114	7,23
Entre 10 y 20 minutos	9.666	13,67
Entre 20 y 30 minutos	11.247	15,90
Entre 30 y 45 minutos	10.824	15,30
Entre 45 minutos y 1 hora	11.510	16,27
Entre 1 hora y hora y media	7.943	11,23
Más de hora y media	1.637	2,31
No es aplicable	12.785	18,08

Fuente: Elaboración propia con datos del @INE 2004. Censos de Población y Viviendas 2001.

Al mismo tiempo en la Tabla 9 resulta elocuente observar que casi el 63 por ciento de las empleadas de hogar trabajarían -sin contar las internas-, en su mismo distrito de residencia –aunque no se desplacen andando-, que sólo el 25 por ciento de ellas trabajarían en un municipio distinto al de residencia, y que más del 3 por ciento trabajan en varios municipios, incluso en otra comunidad autónoma diferente.

Tabla 9. Lugar de trabajo de los empleados de hogar en Madrid.

Lugar de trabajo	*Personas*	%
TOTAL	70.726	
Domicilio propio (internos)	9.336	13,20
Varios municipios	1.291	1,83
Mismo municipio al de residencia	44.353	62,71
Distinto municipio de la misma provincia	14.769	20,88
Otra comunidad	912	1,29
En otro país	65	0,09

Fuente: Elaboración propia con datos del INE 2004. Censos de Población y Viviendas 2001.

En la Tabla 10 podemos observar que más del 12 por ciento de las empleadas de hogar acuden andando a su trabajo, siendo casi un cinco por ciento de ellas quienes acuden conduciendo su propio coche. Pero quizás lo más significativo es que el autobús y el metro constituirían –en principiolos medios de transporte habituales para más del 50 por ciento de las empleadas de hogar en la comunidad de Madrid, y otro más del 6 por ciento se servirían del tren.

A la espera de los resultados del análisis de campo, podemos avanzar que –en principiola tupida red de transporte existente en la Comunidad de Madrid, puede albergar una densidad de movimientos internos mayor que el que muestran estas cifras. Y que realmente dicha red puede

resultar ser uno de los elementos que contribuyan a explicar el fuerte empuje que viene experimentando el empleo doméstico en la región. A diferencia de lo que hemos comprobado en Estados Unidos, especialmente el norte y medio oeste, donde las redes de transporte público son muy limitadas, la necesidad del automóvil privado se impone para los desplazamientos, y las familias viven más diseminadas por el territorio.

Tabla 10. Medios habituales de deslazamiento de los empleados de hogar en Madrid.

Medio de desplazamiento (combinaciones)	*Personas*	%
TOTAL	70.726	
Sólo coche particular, conduciendo	3.456	4,89
Sólo coche particular, de pasajero	686	0,97
Coche particular, conduciendo y transporte público	391	0,55
Coche particular, de pasajero y transporte público	327	0,46
Sólo autobús	13.967	19,75
Sólo metro	11.410	16,13
Sólo tren	2.517	3,56
Autobús + metro	11.278	15,95
Autobús + tren	2.493	3,52
Tren + metro	2.005	2,83
Sólo moto	125	0,18
Andando	8.665	12,25
Bicicleta	32	0,05
Otros	589	0,83
No es aplicable	12.785	18,08

Fuente: Elaboración propia con datos del INE 2004. Censos de Población y Viviendas 2001.

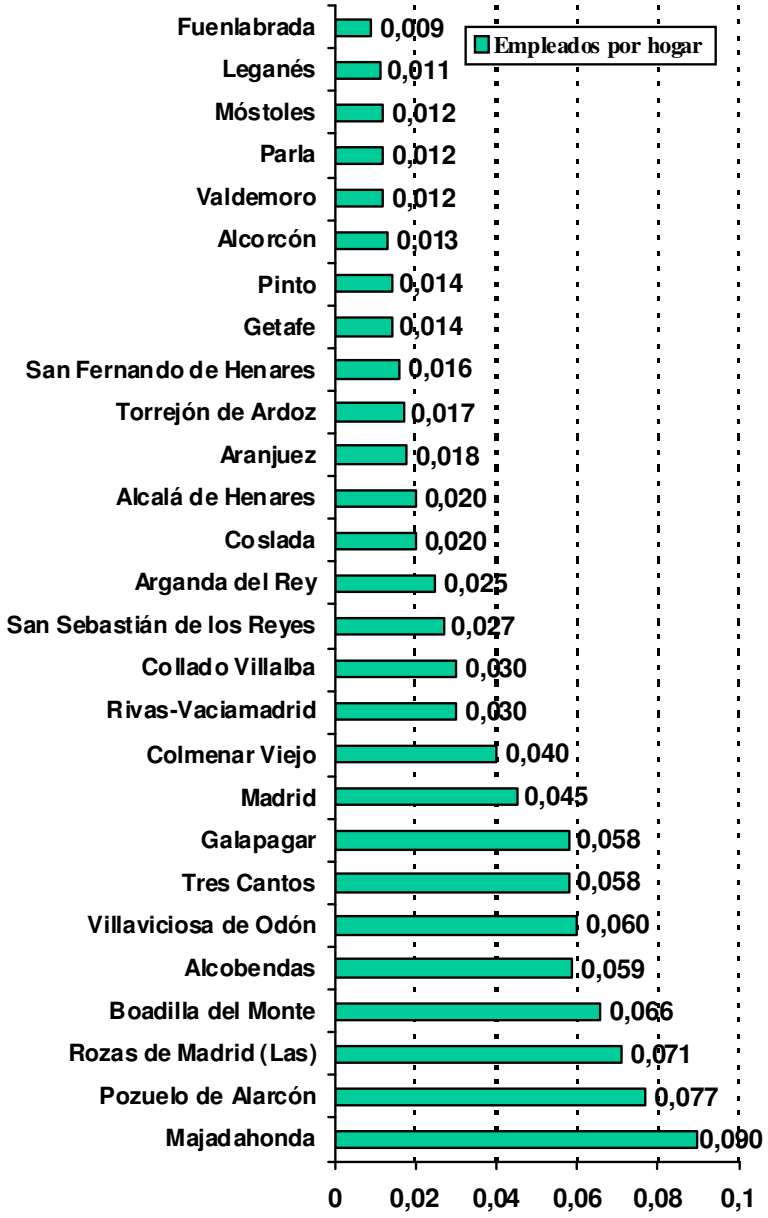
En la Tabla 11 se muestra la distribución de los empleados domésticos por municipios, así como si son españoles o extranjeros, y el número medio de horas de trabajo a la semana. A partir de esa información se han calculado los ratios representativos de la demanda de servicio doméstico en cada municipio: el número de empleados / número de hogares, y el promedio de horas trabajadas a la semana por hogar en cada municipio. Este último indicador es preferible al anterior, ya que se constata que los municipios con mayores valores de Empleados/hogar registran un mayor número medio de horas trabajadas por empleado (ver Gráfico 15.)

Tabla 11. Número de empleados en servicio doméstico, nacionalidades y horas de trabajo a la semana según municipios de la CAM. 2001.

MUNICIPIOS	Núm. hogares	Total empleados servicio doméstico	Nacionalidad Extranjera	Nacionalidad Española	Núm. de empleados por hogar	Horas trabajo a la semana	Horas trabajo por hogar
Alcalá de Henares	56.680	1.118	532	586	0,020	30,3	0,60
Alcobendas	27.670	1.641	1.034	607	0,059	33,8	2,00
Alcorcón	50.850	670	297	373	0,013	30,4	0,40
Aranjuez	13.445	248	81	167	0,018	30,8	0,57
Arganda del Rey	10.024	255	108	147	0,025	31,7	0,81
Boadilla del Monte	8.437	555	394	161	0,066	33,7	2,22
Collado Villalba	15.033	455	310	145	0,030	29,4	0,89
Colmenar Viejo	11.046	446	261	185	0,040	30,6	1,24
Coslada	23.821	478	278	200	0,020	30,3	0,61
Fuenlabrada	56.207	525	149	376	0,009	28,4	0,27
Galapagar	8.286	479	273	206	0,058	31,4	1,82
Getafe	49.879	686	310	376	0,014	30,7	0,42
Leganés	56.712	605	256	349	0,011	30,8	0,33
Madrid	1.080.364	48.863	28.895	19.968	0,045	33,0	1,49
Majadahonda	15.087	1.352	960	392	0,090	33,5	3,00
Móstoles	61.260	722	245	477	0,012	30,1	0,35
Parla	24.329	290	113	177	0,012	30,1	0,36
Pinto	10.592	149	57	92	0,014	29,2	0,41
Pozuelo de Alarcón	19.838	1.524	1.042	482	0,077	35,7	2,74
Rivas-Vaciamadrid	10.942	330	206	124	0,030	31,6	0,95
Rozas de Madrid (Las)	19.944	1.424	1.065	359	0,071	33,2	2,37
San Fernando de Henares	11.828	187	99	88	0,016	30,4	0,48
San Sebastián de los Reyes	19.551	519	292	227	0,027	32,1	0,85
Torrejón de Ardoz	31.057	536	282	254	0,017	29,0	0,50
Tres Cantos	11.566	668	468	200	0,058	33,1	1,91
Valdemoro	10.963	133	73	60	0,012	30,9	0,37
Villaviciosa de Odón	6.664	403	192	211	0,060	31,8	1,92
Total CAM	1.722.075	65.261	38.272	26.989	0,038	31,3	1,11

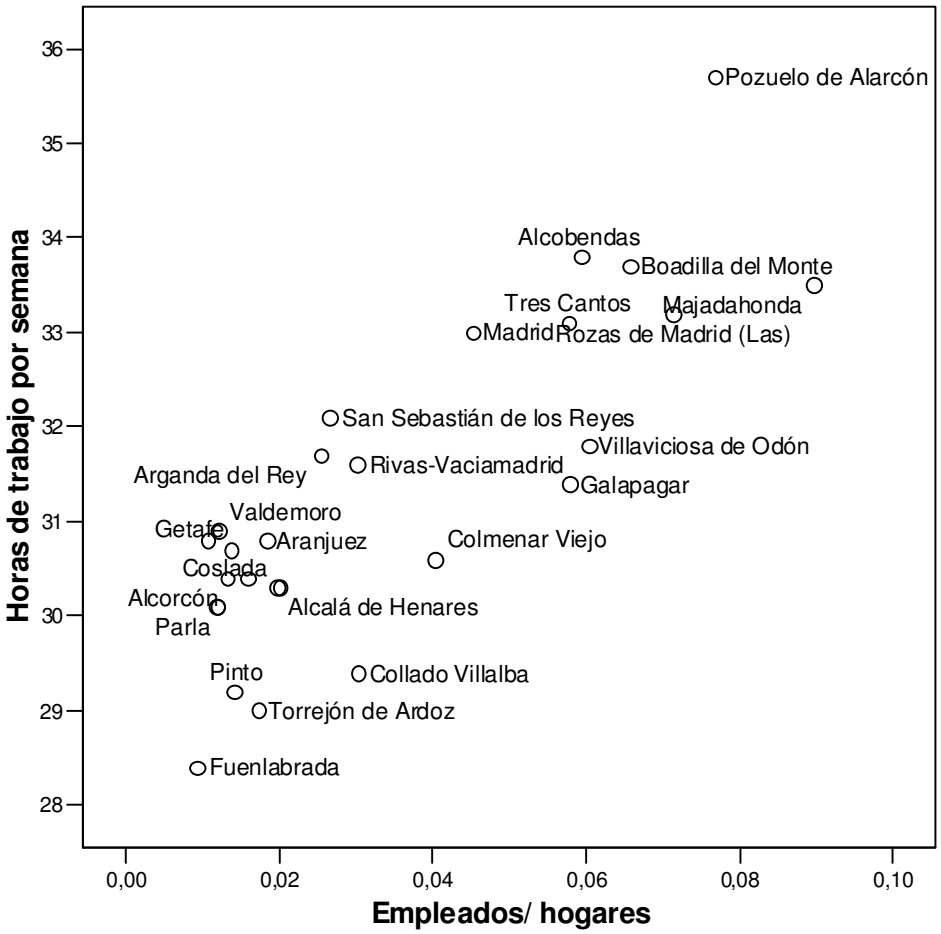
Fuente: Elaboración propia con datos INE.

Gráfico 14. Número de empleados domésticos por hogar según municipios de Madrid.



Fuente: Elaboración propia con datos INE.

Gráfico 15. Relación entre horas medias de trabajo por semana y empleados en servicio doméstico/hogares

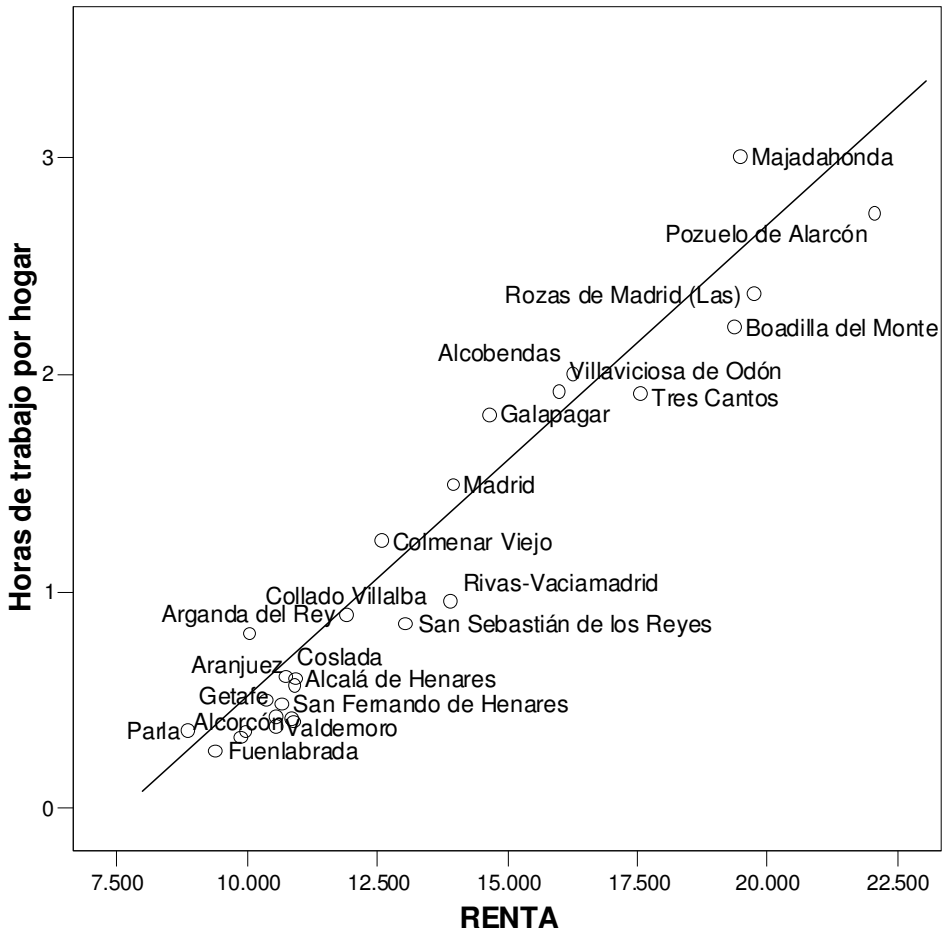


Fuente: Elaboración propia con datos INE.

En el caso de los municipios de Madrid, el número de horas medias de trabajo de servicio doméstico sí está directamente relacionado con la renta per cápita, de tal forma que esta variable permite explicar prácticamente la totalidad (93 por ciento) de las diferencias existentes, como se muestra en la

Tabla 12 y el Gráfico 16. La estimación mejora si se utiliza una variable ficticia, que toma el valor 1 en los municipios con una renta disponible per cápita en 2001 igual o superior al valor de la capital, cifrado según el Instituto de Estadística de la CAM en 13.965 euros, y 0 en caso contrario, según se muestra en la Tabla 13.

Gráfico 16. Relación entre horas medias de trabajo por semana y renta bruta per cápita municipal



Fuente: Elaboración propia con datos del INE.

Tabla 12. Relación entre horas trabajadas por servicio doméstico y renta per cápita.

	Coeficientes no estandarizados	
	B	t
(Constante)	-1,735	-10,6
Renta	0,000216	18,1

R cuadrado = 0,93
Variable dependiente: Horas semanales trabajadas por empleados por hogar / Núm. hogares
Programa estadístico utilizado SPSS. Nivel de significación del 95 por ciento
Fuente: Elaboración propia.

Tabla 13. Modelo utilizado para la estimación de número de horas trabajadas por hogar.

	Coeficientes no estandarizados	
	B	t
(Constante)	-1,137	-4,8
Renta	0,000157	7,4
Nivel de renta (0,1)	0,518	3,2

R cuadrado = 0,95
Variable dependiente: Horas semanales trabajadas por empleados por hogar / Núm. hogares
Programa estadístico utilizado SPSS. Nivel de significación del 95 por ciento
Fuente: Elaboración propia.

5. DETERMINANTES DE LA CONTRATACIÓN DE EMPLEO DOMÉSTICO: UN MODELO PARA SU CONTRASTACIÓN

El análisis de la información contenida en el Censo de Población 2001 y la encuesta de Presupuestos Familiares ha permitido extraer conclusiones sobre los niveles de contratación de empleo doméstico en España y sus factores determinantes en la CAM. De esta forma, la

Tabla **14** muestra los valores estimados de las horas de trabajo en cada hogar y el porcentaje de empleados sobre el total horas para cada uno de los distritos de Madrid., utilizando la ecuación de la Tabla 13..

Tabla 14. Resultados del modelo utilizado para estimar la demanda de servicio doméstico según los distritos de Madrid.

	Reta disponible 2000	Renta 2001	Renta alta	Num hogares	Estimación del número de horas trabajadas por hogar	Horas semana/ trabajador	Estimación del % Empleados/ Núm. Hogares
Madrid	12.768	13.965	1	1.080.364	1,6	32,6	4,85
Centro	12.393	13.555	0	56.424	1,0	32,6	3,06
Arganzuela	13.179	14.415	1	51.954	1,7	32,6	5,06
Retiro	16.832	18.410	1	44.845	2,3	32,6	6,99
Salamanca	17.876	19.552	1	55.055	2,5	32,6	7,54
Chamartín	19.678	21.523	1	49.798	2,8	32,6	8,49
Tetuán	13.161	14.395	1	54.242	1,6	32,6	5,05
Chamberí	17.047	18.645	1	56.820	2,3	32,6	7,10
Fuencarral-El Pardo	14.792	16.179	1	68.083	1,9	32,6	5,91
Moncloa-Aravaca	17.439	19.074	1	38.848	2,4	32,6	7,31
Latina	10.518	11.504	0	89.177	0,7	32,6	2,07
Carabanchel	9.664	10.570	0	78.811	0,5	32,6	1,62
Usera	8.577	9.381	0	41.799	0,3	32,6	1,04
Puente de Vallecas	8.203	8.972	0	81.021	0,3	32,6	0,85
Moratalaz	11.723	12.822	0	37.382	0,9	32,6	2,70
Ciudad Lineal	13.090	14.317	1	81.190	1,6	32,6	5,02
Hortaleza	13.654	14.934	1	49.940	1,7	32,6	5,31
Villaverde	8.674	9.487	0	44.504	0,4	32,6	1,10
Villa de Vallecas	9.854	10.778	0	20.234	0,6	32,6	1,72
Vicálvaro	9.223	10.088	0	18.714	0,5	32,6	1,38
San Blas	9.980	10.916	0	49.496	0,6	32,6	1,78
Barajas	14.579	15.946	1	12.477	1,9	32,6	5,80

Fuente: Elaboración propia.

Para aquilatar los análisis anteriormente realizados se realizaron un total de 3000 encuestas, entre empleados y empleadores en la Comunidad de Madrid, a lo largo del primer semestre de 2006, lo que nos lleva a las siguientes resultados y conclusiones.

6. RESUMEN Y CONCLUSIONES

1) El alto grado de dinamismo del empleo en el ámbito del servicio doméstico ha ayudado a la masiva incorporación de la mujer al mercado de trabajo madrileño en la pasada fase del ciclo expansivo del ciclo económico, retroalimentando, a su vez, la mayor demanda de actividad y empleo en el resto de sectores productivos de más valor añadido y contenido tecnológico alto.

2) Necesidad desde el punto de vista de la economía internacional de conseguir un más alto grado de reconocimiento social para el empleo en el ámbito del servicio doméstico.

3) En tanto no se construyan y desarrollen las costosas infraestructuras públicas previstas en los diferentes Sistemas del Plan Nacional de Dependencia de España y la UE, las familias seguirán de forma creciente cubriendo sus necesidades por sí mismas, y a un coste global menor.

4) Gran flexibilidad y capacidad de adaptación al ciclo económico de este mercado de trabajo pactado en el interior de los hogares, sirviendo de colchón en las fases recesivas, tanto para empleados como para empleadores.

5) La renta per cápita media provincial de España no tiene, por sí sola, capacidad explicativa del número relativo de empleados de servicio doméstico.

6) Una mayor capacidad explicativa de la tenencia de empleo doméstico a nivel provincial se relaciona con la implantación de ciertas categorías profesionales: concretamente "Técnicos profesionales, científicos e intelectuales", "Empleados de tipo administrativo" y "Dirección de las empresas y de las administraciones públicas"; también en el caso de Madrid serían el número de hijos del hogar y el porcentaje de hogares compuestos por personas mayores de 65 años.

7) En el caso de los municipios de la Comunidad de Madrid el número de horas medias de trabajo de empleo doméstico si está directamente relacionado con la renta per cápita de aquellos, de tal forma que esta variable permite explicar casi la totalidad -93 por cientode las diferencias existentes; en este sentido los niveles de renta de los municipios de la región de Madrid sí presentan una clara correlación positiva.

8) De los microdatos de la Encuesta Continua de Presupuestos Familiares, ECPF, deducimos que el porcentaje de los empleadores de servicio doméstico se triplica en los hogares con ingresos provenientes de "Rentas de la propiedad y del capital" sobre los de "Trabajo por cuenta propia", y más que triplica a los de "Trabajo por cuenta ajena".

9) Las personas ocupadas en el sector público contratan más empleo doméstico que los del sector privado, y aquellas emplean en mayor proporción en el conjunto de España que en la Comunidad de Madrid.

10) Una potente red de infraestructuras de transporte es necesaria para el desarrollo del empleo doméstico en los Hogares. En Madrid, más del 60 por ciento de las empleadas de hogar se desplazan a su lugar de trabajo en unos tiempos que oscilan entre los diez y los sesenta minutos; más del 50 por ciento utilizan el metro o autobús para llegar al mismo.

11) En los intentos de modelización de los determinantes de la contratación de empleo doméstico en los Hogares, que hemos realizado, podemos deducir que hay una alto grado de correlación entre el tipo de jornada laboral (interna, completa, media jornada o por horas) y nacionalidad de origen de la empleada de hogar, con un Chi cuadrado de Pearson que tiene un nivel de significatividad próximo a cero.

12) Que las principales variables explicativas del salario/hora en el ámbito del empleo doméstico son: tipo de jornada laboral, si cotiza o no a la Seguridad Social, la edad y la nacionalidad, que explican el 66,4 por ciento de la variabilidad del modelo: con un resultado

del Coeficiente de Student, t, mayor a 2 en todos los casos, hemos comprobado el alto grado de correlación de todas esas variables.

13) Que las variables explicativas de la cotización o no de la empleada de hogar a la Seguridad Social, están en: Su estado civil, si su pareja contribuye o no a la renta familiar, su edad, su nacionalidad, su tipo de domicilio (en propiedad, alquiler u otras variantes), su antigüedad en el hogar de trabajo, entre otras, pronosticando hasta el 84,4 por ciento de las empleadas que son alta en la Seguridad Social.

14) La ocupación previa de las empleadas de hogar está mayoritariamente en el sector servicios: 24,2 por ciento. Pero la mayoría no tenían empleo: 27,9 por ciento, lo que confirma el papel introductorio de esta actividad en el mercado de trabajo especialmente importante para la población inmigrante. Apenas un 2,1 por ciento estaban ocupadas en el sector agrario.

15) Los amigos, los familiares y la Iglesia, son las tres vías principales y mayoritarias a la hora de buscar y encontrar empleo doméstico tanto por parte de la/os empleada/os de hogar y de los empleadores. Tales vías se han contrastado además con la creación del Observatorio Permanente del Empleo Doméstico en Internet en la Comunidad de Madrid. Los servicios públicos de empleo son irrelevantes en este segmento del mercado de trabajo. La Iglesia, por el contrario, aparece como el primer agente integrador del servicio doméstico, a través de las Parroquias, Cáritas y entidades religiosas específicas.

16) En cuanto a las posibilidades actuales de mejora profesional de las empleadas de hogar ligada a eventuales estudios en curso, la mayoría -93 por cientono cursan ningún tipo de estudio. Llama la atención, no obstante las de procedencia magrebí que lo hacen en alguna mayor medida: 12 por ciento -frente al 7 por ciento del restosí estarían asistiendo a algún curso de aprendizaje, a pesar de la oposición marital, lo que muestra el deseo de la mujer musulmana de romper su cerco.

17) La causa principal de no tener empleada/o de hogar es simplemente: "no necesitarlo", en un 51 por ciento; frente a un 34,5 por ciento de: "no poder pagarlo". No tienen relevancia el "haber dejado de trabajar" -3,4 por ciento-, lo que es coherente con la fase expansiva del ciclo económico por la que hasta 2008 ha atravesado la economía madrileña; tampoco la "falta de confianza" o la "mala experiencia" son relevantes: 2,6 y 1,2 por ciento respectivamente. Un 3 por ciento de los hogares no tendrían empleo doméstico por "no gustan convivir con extraños".

18) En las encuestas específicas realizadas a los universitarios, comparando dos universidades de Madrid (Universidad San Pablo-CEU) y Estados Unidos (Carroll Collage y Mount Mary College), hemos constatado que incluso en el ámbito específico de los hogares de los universitarios madrileños se constata el alto porcentaje de empleo doméstico: más del 50 por ciento lo tienen, frente a sus equivalentes de Estados Unidos de universidades también privadas: concretamente sólo el 12´72 por ciento (área de Waukesa, zona tradicional de inmigración en la región de los grandes lagos, por ello seleccionada.)

BIBLIOGRAFÍA

ADDISON, Thomas (2002): "The impact of immigration on the earnings and natives: Evidence from Australian micro data", *Economic Record,* March, vol. 78, issue 240, pp. 68-78.

AGIOMIRGIANAKIS, George, ZERVOYIANNI, Athina (2001): "Macroeconomic Equilibrium with illegal immigration", *Economic Modelling*, April, vol. 18, issue 2, pp. 181202.

ALAÑÓN PARDO, A. Y GÓMEZ DE ANTONIO, M. (2003): "Una evaluación del grado de incumplimiento fiscal para las provincias españolas", *Instituto de Estudios Fiscales*, Papel de Trabajo, nº 9/03.

ANDALL, Jacqueline (2000): *Gender, Migration and Domestic Service. The politics of black women in Italy*, Hampshire-England & Burlington-USA, Ashgate Publishing.

ANDERSON, B. (2000): *Doing the Dirty Work*, London, Zed Press.

ANDERSON, B. (1997a): *Labour Exchange. Patterns of Migration in Asia*, London, Catholic Institute of International Relations.

ANDERSON, B., PHIZACKALEA, A. (1997): *Migrant Domestic Workers: A European Perspective*, Brussels, European Commission, Equal Opportunities Unit (Una version completa en ANDERSON, N., 2000, London, Zed Press).

ANGRIST, Joshua, KUGLER, Adriana (2003): "Protective or counter-productive?. Labour market institutions and the effect of immigration on European Union", *The Economic Journal*, London, June, vol. 113, issue 488, pp. 302-331.

ANGULO, Carlos I. (2004): "Condiciones de vida de la población extranjera en España", *España, un país de inmigración*, núm. 99, Año XXII, pp. 98-107, Colegio de Economistas de Madrid, Madrid.

ANTE, Spenser, MAGNUSSON, Paul (2003): "Too Many VISAS for techies?. U.S. workers gripe that lax rules may cost them their jobs", *Business Week*, New York, August 25, issue 3846, p. 39 ss.

APARICIO GÓMEZ, Rosa (2002): "El impacto económico de la inmigración: costes para el Estado y movimiento de consumos y salarios", *Cuadernos de Trabajo Social*, Vol. 15, pp. 49-52, Madrid.

ARANDA, José (2004): "Los inmigrantes y la delincuencia", *España, un país de inmigración*, núm. 99, Año XXII, pp. 108-115, Colegio de Economistas de Madrid, Madrid.

-ASLUND, Olof (2005): "Now and forever?. Initial and subsequent location choice of immigrants", *Regional Science and Urban Economics*, Sweden, March, vol. 35, issue 2, pp. 141-165.

BALDWIN GROSSMAN, Jean (1984): "Illegal Immigrants and Domestic Employment", *Industrial and Labor Relations Review*, Vol. 37, nº 2, January. Cornell University, School of Industrial & Labor Relations.

-BARBI, A., MIKLAVI-BREZIGAR, I. (1999): "Domestic Work abroad: a necessity and an opportunity for rural women from Slovenia", en Momsen (ed.). *Gender, Migration and Domestic Service*, London, Routledge.

BECKER, Gary S. (2003): "When globalization suffers, the poor take the heat", *Business Week*, New York, April 21, issue 3829, p. 28.

BEAN, Frank (2004): "Immigrant Job Quality and Mobility in the United States", *Work and Occupations*, U.S.A., November, vol. 31, issue 4, pp. 499-518.

BEHR, Michelle, FATA, Edina, KULCSAR, Anita et al (2002): "Who is Hungarian?. Attitudes toward immigration, ethnicity and nationality in rural Hungary", *East European Quarterly*, Fall, vol. 36, issue 3, pp. 281-299.

BÖCKERMAN, Petril (2004): "Does labour market reorganization reduce unemployment?", *Applied Economics Letters*, Finland, October, vol 11, issue 12, pp. 745-748.

BORJAS, George (2003): "The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market", *The Quartely Journal of Economics*, Cambridge, November, vol. 118, issue 4, pp. 1335 ss.

BROWNE, Anthony (2003): "Some truths about immigration", *The Spectator*, London, August 2, vol. 292, issue 9130, p. 18.

BUTCHER, Kristin, DINARDO, John (2002): "The immigrant and native-born wage distributions: Evidence from United States Censuses", *Industrial & Labor Relations Review*, October, vol. 56, Issue 1, pp. 97-121.

CAMACHO BALLESTA, J. A., et al (2006): "Consumo de servicios y ciclo vital: Un análisis de los hogares con datos de corte transversal", *Servilab (Laboratorio de Investigación del Sector Servicios)*, Alcalá de Henares.

CANO, E. Y SANCHEZ, A. (2000): "Empleo Informal y precariedad laboral: las empleadas de hogar de Nazaret". Departamet d´Economía Aplicada, Universitat de Valencia.
http://www.ucm.es/info/ec/jec8/Datos/documentos/comunicaciones/Laboral/Cano%20Ernest.PDF

CARABAÑA, Julio (2004): "La inmigración y la escuela", *España, un país de inmigración*, núm. 99, Año XXII, pp. 62-73, Colegio de Economistas de Madrid, Madrid.

CARD, David (2001): "Immigrants inflows, native outflows, and the local labor market impacts of higher immigration", *Journal of Labor Economics*, January, vol. 19, issue 1, pp. 22-64.

CASARICO, Alessandra, DEVILLANOVA, Carlo, (2003): Social security and migration with endogenous skill upgrading, *Journal of Public Economics*, March, vol. 87, issue ¾, pp. 773-797.

CATANZARITE, Lisa (2003): "Occupational context and wage competition of new immigrants latinos with minorities and whites", *Review of Black Political Economy*, Summer-Fall, vol. 31, Issue ½, PP. 77-94.

CEBRIAN, J.A. et al (2004): "La iniciativa empresarial del inmigrante", *España, un país de inmigración*, núm. 99, Año XXII, pp. 1116-123, Colegio de Economistas de Madrid, Madrid.

CEREIJO, E., VELAZQUEZ, F.J. (2004): "Los determinantes de las inmigraciones en la UE", *España, un país de inmigración*, núm. 99, Año XXII, pp. 38-47, Colegio de Economistas de Madrid, Madrid.

CHANG, Grace (2000): *Disposable Domestics. Immigrant Women Workers in the Global Economy*, South End Press, Cambrigde, Mass. Printed in Canada.

CHATTERJEE, Satyajit (2003): "Agglomeration Economies: The Spark that ignites a city?", *Business Review*, Federal Reserve Bank of Philadelphia, Fourth Quartier, 25 ss.

CHEN, Kong, CHIAN, Shin, LEUNG, Siu (2003): "Migration, Family and Risk Diversification", *Journal of Labor Economics*, April, vol. 21, issue 2, pp. 353-380.

CHERRY, Robert (2003): "Immigration and Race: What we think we know", *Review of Black Political Economy*, New Brunswick, Summer/Fall, vol. 31, pp. 157-184.

CISCEL, David, SMITH, Barbara, MENDOZA, Marcela (2003): "Ghosts in the Global Machine: New immigrants and the redefinition of work", *Journal of Economic Issues*, June, vol. 37, issue 2, pp. 333-341.

COBB-CLARK, Deborah (2000): "Do selection criteria make a difference?: Visa category and the labour market status of Australia" , *Economic Record*, March, vol. 76, issue 232, pp. 15-31.

COLECTIVO IOÉ (2001): "Mujer, inmigración y trabajo". *Instituto de Migraciones y Servicios Sociales*, Ministerio de Trabajo y Asuntos Sociales. Madrid

COMBES, Pierre et al (2005): "The trade-creating effects of business and social networks: evidence from France", *Journal of International Economics*, Amsterdam, May, vol. 66, issue 1, pp. 1-29.

COMISIÓN DE LAS COMUNIDADES EUROPEAS (2005): "Comunicación de la Comisión al Consejo y al Parlamento Europeo de los programas regionales de protección", *Comisión Europea,* COM.(2005) 388 final, pp.1-9, Bruselas.

COMISIÓN EUROPEA (2001): "Undeclared Labour in Europe. Towards an integrated approach of combating undeclared labour". *Regioplan Publication*. Amsterdam.

CONSEJERÍA DE EMPLEO Y MUJER (Marzo, 2006): *Encuesta de Población Activa-INE, IV Trimestre. Balance de datos de 2005*, Madrid. Comunidad de Madrid.

CONSEJO DE LA UNIÓN EUROPEA (2003-a): Decisión del Consejo 2003/578/CE, de 22 de julio de 2003, relativa a la Directrices para las políticas de empleo de los Estados Miembros.

CONSEJO DE LA UNIÓN EUROPEA (2003-b): Recomendación 2003/579/CE, de 22 de julio de 2003, sobre la aplicación de las políticas de empleo de los Estados Miembros.

CONSEJO ECONOMICO Y SOCIAL (1er Trimestre2006): "Panorama sociolaboral de la mujer en España", *Boletín elaborado por el Área de Estudios y Análisis,* núm. 43, pp.1-4, Madrid.

CONSEJO ECONÓMICO Y SOCIAL (1999): "La economía sumergida en relación a la quinta recomendación del Pacto de Toledo".*Informe CES*. Madrid

CONSEJO ECONÓMICO Y SOCIAL DE ANDALUCÍA (2002): "La economía sumergida", *Junta de Andalucía*.

CRUSH, Jonathan (2002): "The global raiders: nationalism, globalization and the South African brain drain", *Journal of International Affairs*, New York, Fall, vol. 56, issue 1, pp. 147-172.

CRUZ ROJA (2003): " Investigación servicio doméstico de cuidado: estrategias de conciliación y condiciones de empleo en diferentes unidades familiares", Cruz Roja, Madrid.

CUADRADO ROURA, Juan (2003): "Madrid, centro nacional e internacional de servicios", *Economía de Madrid*, núm. 95, Año XXI, pp. 65-71, Colegio de Economistas de Madrid, Madrid.

CUANTER S.A. (2001).: "Estudio sociosanitario de las empleadas de hogar". *Instituto de la Mujer*. Documento interno. Madrid

DAVIDOFF, E. , HALL, C. (1987): *Family Fortunes. Men and Women of the English Middle Class*, 1780-1850, London, Hutchinson.

DANIELS, Joseph, VON DER RUHR (2003): "The Determinants of Immigration-Policy. Preferences in Advanced Economies: A Cross-Country Study", *Atlantic Economic Journal*, Vol. 31, Issue 2, pp. 146-159.

DURÁN HERAS, María Ángeles (2001): "La protección de las personas dependientes", *Seminario Aproximación a la dependencia*, Fundación UNIATE, Benidorm.

DE CABO, Gema, et al (2005): " La presencia de las mujeres en el empleo irregular (con especial atención a las empleadas de hogar)", *Centro de Estudios económicos Tomillo S.L.,* Madrid.

DE CASTRO, M.A. (2004): "Fuentes estadísticas sobre la Inmigración", *España, un país de inmigración*, núm. 99, Año XXII, pp. 128-745, Colegio de Economistas de Madrid, Madrid.

DEL BOCA D. (1998): "Labour Policies, Economic Flexibility and Women´s Work", en DREW, E. et al (eds.): *Women, Work and the Family in Europe*, London, Routle, pp. 124-130.

DEL RE, A. (2000): "The paradoxes of Italian Law and Practice", en Hantrais, L. (ed.), *Gendered Policies in Europe*, London, Macmillan, pp. 108-123.

DYMSKI, Gary (2003): "Immigration, finance, and urban evolution: an illustrative model, with a Los Angeles case study", *Review of Black Political Economy*, Spring, vol. 30, issue 4, pp. 27=50.

ECONOMIST.COM (2003): "World on the move", Global Agenda, London, June 11, p. 1.

EDITORIAL (2005): "Q & A: H-1 Bs Helps U.S. Economy", *Computerworl*, Framingham, U.S.A., February 29, vol. 39, issue 9, pp. 4-5.

EDITORIAL (2005): "The H-1B Equation", *Computerworld*, Framingham, U.S.A., February 28, vol. 39, issue 9, pp. 4-5.

EDITORIAL (2005): "How Nations Can Age Gracefully", *Business Week*, January 31, issue 3918, p. 100.

EDITORIAL (2002): "A turning point", *Canada and the World Backgrounder*, Waterloo, Canada, December, vol. 68, issue 3, pp. 4-6.

ESCRIVÁ CHORDÁ, Ángeles (2003): " Inmigrantes peruanas en España. Conquistando el espacio laboral extradoméstico", Revista Internacional de Sociología (RIS), Tercera época, núm. 36, Septiembre-Diciembre, pp. 59-83.

FARKAS, Orsolya, RYMKEVITCH, Olga (2004): "Immigration and the free movement of workers after enlargement: contracting choices", *The International Journal of Comparative Labour Law and Industrial Relations,* vol. 20, issue 3, pp. 369-397.

FOUST, Dean, GROW, Brian, PASCUAL, Aixa (2002): "The changing heartland. An inlux of newcomers

both buoys and burdens small-town America", *Business Week*, New York, September 9, issue 3798, p. 80 ss.

FREEDMAN, Michael (2005): "Welcome to Londongrad", *Forbes*, New York, May 23, vol. 175, issue 11, p. 158 ss.

GARCÍA BALLESTEROS, Aurora (2003): "La inmigración en la población madrileña", *Economía de Madrid*, núm. 95, Año XXI, pp. 40-51, Colegio de Economistas de Madrid, Madrid.

GARRIDO, Luis, et al (2004): "La situación laboral de los españoles y los extranjeros según la Encuesta de Población Activa", *España, un país de inmigración*, núm. 99, Año XXII, pp. 74-87, Colegio de Economistas de Madrid, Madrid.

GASTON, N., NELSON, D. (2000): "Immigration and labour-market outcomes in the United States: a political economy puzzle", *Oxford Review of Economic Policy*, Autumn, vol. 16. issue 3.

GIBEAUT, John (2003): "The paperwork war on Terrorism", *ABA Journal*, Chicago, October, vol. 89, pp. 62-68.

GONZALEZ, Mar (2006): " Mas de la mitad de las empleadas de hogar no tiene contrato", ABC, lunes 13 febrero de 2006, http://www.abc.es/especiales/index.asp?cid=3198

GÓRNY A, RUSPINI P.(eds.) (2004): *Migration in the New Europe*. East-West Revisited, New York, Palgrave Macmillan.

GREGSON, N. ,LOWE,M. (1994): *Servicing the Middle Classes: Class. Gender and Waged Domestic Labour in Contemporary Britain*. London, Routledge.

GROSS, Dominique (2004): "Impact of Immigrant workers on a regional labour market", *Applied Economics Letters*, October, vol. 11, Issue 7, pp. 405-408.

GROSS, Dominique (2002): "Three million foreigners, three million unemployed?. Immigration flows and the labour market in France", *Applied Economics Letters*, London, November 10, vol. 34, issue 16, pp. 196-198

HAGAN, Jacqueline (2004): "Contextualizing Immigrant Labor Market Incorporation: Legal, Demographic and Economic Dimensions"*, Work and Occupations*, U.S.A., November, vol. 31, issue 4, pp. 407-423.

HÄMALAINNEN, Karil, BÖCKERMAN, Petri (2004): "Regional Labor Market Dynamics, Housing, and

Migration", *Journal of Regional Science*, Finland, August, vol. 44, issue 3, pp. 543-568.

HANSON, Gordon, ROBERTSON, Raymond, SPILIMBERGO, Antonio (2002): Does border enforcement protect U.S. workers from illegal immigration?", *Review of Economics & Statistics*, February, vol. 84, Issue 1, pp. 73-92.

HARDILL, Irene, MACDONALD, Sandra (2000): "Skilled International Migration: The experience of Nurses in the United Kingdom", *Regional Studies,* UK, October, vol. 34, issue 7, pp. 681-692.

HENSHALL MOMSEN, Janet (ed.) (1999): *Gender, migration, and domestic service*, London, Routledge.

HERNÁNDEZ-LEÓN, Rubén (2004): "Restructuring at the: Source High-Skilled Industsrial Migration from Mexico to the United States", *Work and Occupations*, U.S.A., November, vol. 31, issue 4, pp. 424-452.

HERNANZ GÓMEZ, Yolanda (1998): "Servicio doméstico y feminización de la inmigración en Madrid", *Ofrim/ Suplementos*, pp. 66-83, Madrid.

HERNANZ GÓMEZ, Yolanda (1992): "Trabajadores latinoamericanos en Madrid", *ES. Economía y Sociedad*, núm. 7, pp. 119-137, Madrid.

HIEBERT, Murray (2002): "The cost of security", *Far Eastern Economic Review*, Hong-Kong, November 28, vol. 165, issue 47, pp. 18-19.

HILL, Charles (2005): *International Business: competing in the global marketplace*, New York, McGraw-Hill.

HONDAGNEU-SOTELO, Pierrette (2001): *Doméstica. Immigrant Workers Cleaning and Clearing in the Shadows of Affluence*, Berkeley, University of California Press.

ICAM (2005): "El servicio doméstico, área productiva donde existe mayor irregularidad laboral femenina", Ilustre Colegio de Abogados,
http://www.icam.es/vernoticia.jsp?id=200503280002&seccion=extranjeria&canal=estanjeria

INSTITUTO ANDALUZ DE LA MUJER (2001): "Servicio Doméstico en Andalucía: Oferta y Demanda". *Instituto Andaluz de la Mujer*, Sevilla.

INSTITUTO NACIONAL DE ESTADÍSTICA (2002): "Encuesta sobre Discapacidades, Deficiencias y Estado de Salud 1999, Resultados detallados", *INE*, Madrid.

IZQUIERDO, Antonio (2001): "La inmigración en España entre dos regularizaciones", *España 2001, un balance*, núm. 91 extra Año XX, pp. 277-282, Colegio de Economistas de Madrid.

JACOBY, Sanford, FINKIN, Matthew (2004): "Labor mobility in a federal system: The United States in comparative perspective", *The International Journal of Comparative Labour Law and Industrial Relations*, September, vol. 20, issue 3, pp. 313-337.

JANIGAN, Mary (2002): Immigrants: "How many is too many?. Who should get in?. Can we tell them where to live and what to do?", *Maclean´s Toronto*, Canada, December 16, vol. 115, issue 50, pp. 20-25.

JASPER, William (2004): "Free Trade Area of the Americas: Forced To Accept Alliens", *The New American*, Appleton, September 6 , vol. 20, issue 18, pp. 14-17.

JOHANNSSON, Hannes, WEILER, Stephan (2004): "Local Labor Market Adjustment to Immigration: The Roles of Participation and the Short Run", *Growth and Change*, Lexington, U.S.A., Winter, vol. 35, issue 1, pp. 61-76.

JONES, Ronald (2005): "Immigration *versus* outsourcing : effects on labor markets", *International Review of Economics and Finances*, U.S.A., vol, 14, issue 2, pp. 105-114.

KOFMAN E., PHIZACKLEA, A., RAGHURAM P., SALES R. (2000): *Gender and International Migration in Europe. Employment, welfare and politics*, New York, Routledge.

KOSSOUDJI, Sherrie, COBB-CLARK, Deborah (2002): "Coming out of the Shadows: Learning about Legal Status and Wages from the Legalized Population", *Journal of Labor Economics*, July, vol. 20, issue 3, pp. 598-628.

KRIKORIAN, Mark (2005): "Borderline Insanity", *The National Interest*, Washington, Spring, issue 79, pp. 70-72.

LIGHT, Ivan (2003): "The New Immigrants in the American Economy", *Journal of American Ethnic History*, U.S.A., Fall, vol. 23, issue 1, pp. 102

LOFSTROM, Magnus, BEAN, Frank, (2002): "Assessing immigrant policy options: labor market conditions and postreform declines in immigrants´ receipt of welfare", *Demography*, U.S.A., November, vol. 39, issue 4, pp. 617-638.

-MACKLIN, A. (1994): "On the outsidelooking in: Foreign domestic workers in Canada", in W. Giles and S. Arat-Koc (eds.), *Maid in the Market*, Halifax, Canadá, Fernwood Publishing.

MAÑAS ALCÓN, Elena, et al (2000): "Los servicios a las familias como motores del crecimiento económico", *Servilab (Laboratorio de Investigación del Sector Servicios),* nº 4/2000, Alcalá de Henares.

MARCELLI, Enrico (2005): "Immigrants and the U.S. Labor Market", *North American Congress on Latin America Report on the Americas*, New York, March/April, vol. 38, issue 5, pp. 47 ss.

MCBRIDE, T. (1976): *The Domestic revolution*, London, Croom Helm

Migration News Sheet (1997): "Migration Policy Group", March, nº 168/97, Brussels.

MEDC (2005): "Datos y cifras. Curso escolar 2003/2004", *Ministerio de Educación y Ciencia*, Madrid, www.mecd

MEIL LANDERLIN, Gerardo (2004): "Continuidad y cambio en la división del trabajo doméstico en la familia urbana", *Ponencia VIII Congreso de Sociología,* Alicante.

MENNITI, A. et al (1997): "Italy: Changing the Family from Whitin", en KAUFMAN F. et al (eds.): "Family Life and Family Policies in Europe", Vol. 1, Oxford, Clarendom Press, pp. 225-52.

MINISTERIO DE TRABAJO Y ASUNTOS SOCIALES (2003): "El servicio del hogar familiar", http://www.mtas.es/infpuntual/HogarFam2003/Indice.htm.

MOLINA IBAÑEZ, Mercedes (2003): "Las inversiones extranjeras", *Economía de Madrid*, núm. 95, Año XXI, pp. 73-79, Colegio de Economistas de Madrid, Madrid.

-MORILLAS, J. (2007): *Mercado de Trabajo e Inmigración: Los servicios a los Hogares en el ámbito del empleo doméstico en la Comunidad de Madrid*, Editorial Universitas, S.A., Madrid.

-MORILLAS, J. (2006): "Evolución del mercado laboral en relación con el fenómeno migratorio", En *IV Jornadas europeas sobre intercambios laborales. Las Migraciones: retos sociales y desafíos culturales*, Madrid-Bruselas, *European Center for Workers´s Questions, EZA*, 2005. Edición bilingüe español-inglés

-MORILLAS, J. (2005): "La inmigración de Europa del Este a España, ¿costes o beneficios?", en *Boletín de Información Comercial Española,* Ministerio de Industria, Comercio y Turismo, 9/15 mayo 2005, nº 2844, pp. 37-48.

MORILLAS, J. (2002): "Inmigración y Cooperación al Desarrollo en España, 2002", en *Boletín de Información Comercial Española, BICE*, Ministerio de Economía, nº 2738, 2/8 septiembre, 2002, pp. 23-28.

MORILLAS, J. (2002): "La modernización de los mercados de trabajo en la Unión Europea. Un estado de bienestar activo. El caso español", en *Boletín de Información Comercial Española, BICE*, Ministerio de Economía, nº 2724, 1/7 abril, 2002, pp.27-32.

MORILLAS, J. (2002): "Movilidad en los nuevos mercados de trabajo de la Unión Europea", en *BICE*, Ibidem, nº 2691.

MORILLAS, J. (2001): "El pleno empleo, un nuevo paradigma en la Unión Europea. Evaluación e Indicadores", en *BICE*, Ibidem, nº 2679.

MOROKVASIC, M. (1995): "In and out of the Labour Market: Immigrant and Minority Women in Europe", *New Community*, 19 (3), pp. 459-83

MTAS (diciembre 2005): "Anteproyecto de Ley de Promoción de la Autonomía Personal y de Atención a las Personas en situación de Dependencia", *Sistema Nacional de Dependencia*, Ministerio de Trabajo y Asuntos Sociales, Madrid.

MÜLLER, Tobias (2003): "Migration, unemployment and discrimination", *European Economic Review*, June, vol. 47, issue 3, pp. 409-427.

MÜLLER, Tobias (2003): "Migration Policy in a Small Open Economy with a Dual Market", *Review of Internacional Economics*, February 2003, Vol. 11, pp. 130-143.

MURPHY, Dean, (2004): "Imagining America Without Illegal Immigrants", *New York Times Upfront*, New York, February 23, vol. 136, issue 9, pp.14-17.

NEWPORT, Frank (2004): "Americans worried about immigration, oppose Bush plan: Majority says immigrants hurt economy", *The Gallup Poll Tuesday Briefing*, Washington, January, pp. 11-12.

OPCIÓN LUNA (2001): "Guía de recursos par alas empleadas de hogar". Asociación de empleadas de hogar "*Opción Luna*" de Córdoba.

OSO, Laura (1998): "La inmigración femenina en Madrid y Lisboa: hacia una etnización del servicio doméstico y de las empresas de limpieza", *Revista de Sociología Papers,,* núm. *60*, pp. 183-207.

PALACIO, Juan Ignacio (2003): "Emigración y crecimiento económico", *Economía de Madrid*, núm. 95, Año XXI, pp. 34-39, Colegio de Economistas de Madrid, Madrid.

PE-PUA, Rogelia (2005): "Vínculos entre Filipinas y España: migraciones y relaciones bilaterales", Revista CIDOB d´Afers Internacionals, núm. 68, pp. 53-65.

PHIZACKALEA, A. (1998): "Migration and Globalization: A feminist perspective", in K. Koser y H. Lutz (eds.), *The New Migration in Europe*, London, Macmillan, pp. 21-38.

PLA JULIÁN, Isabel, et al (2004): "Informalidad del empleo y precariedad laboral de las empleadas de hogar (en la Ciudad de Valencia)", Instituto de la mujer de la Generalitat Valenciana-Universidad de Valencia, Valencia.

REITZ, Jeffery, VERMA, Anil (2004): "Immigration, Race, and Labor: Unionization and Wages in the Canadian Labor Market", *Industrial Relations*, Canada, October, vol. 43, issue 4, pp. 835-854.

RENKOW, Mitch (2003): "Employment Growth, Worker Mobility, and Rural Economic Development", *American Journal of Agricultural Economics*, May, vol. 85, issue 2, pp, 503=513.

REQUES, Pedro y DE COS, Olga (2004): "De lo global a lo local: repercusiones geodemográficas dela inmigración en España", *España, un país de inmigración*, núm. 99, Año XXII, pp. 48-61, Colegio de Economistas de Madrid, Madrid.

RODRIGUEZ, Nestor (2004): "*Workers wanted*: Employer Recruitment of Immigrant Labor", *Work and Occupations*, U.S.A., November, vol. 31, issue 4, pp. 453-473.

ROMERO, Mary (2002): *Maid in the U.S.A*., New YorkLondon, Routledge.

RUHS, Martin, CHANG, Ha, (2004): "The Ethics of Labor Immigration Policy", *International Organization Foundation*, winter, vol. 58, pp. 69-102.

SAIZ, Albert (2003): "The Impact of Immigrants on American Cities: An Introduction to the Issues", *Business Review*, Federal Reserve Bank of Philadelphia, Four Quarter, pp. 14-23.

SANTAMARIA, Arturo (2003): "Politics without borders or postmodern nationality", *Latin America Perspectives*, Mexico-U.S.A., March, vol. 30, issue 2, pp. 66-86.

SARIN, Radhica (2003): "As migration grows, immigrants face greater barriers", *World Watch*,

Washington, May/June, vol. 16, issue 3, p. 8.

SCHEVE, Kenneth, SLAUGHTER, Matthew (2001): "Labor Market competition and individual preferences over immigration policy", *Review of Economics & Statistics*, February, vol. 83, Issue 1, pp. 133-145.

SCHNEIDER, F., Y ENSTE, D. (2000): "Shadow economies: size, causes, and consequences." *The Journal of Economic Literature*. Primavera 2000.

SIPI MAYO, Remei (2004): *Inmigración y género. El caso de Guinea Ecuatorial*, San Sebastián, Tercera Prensa.

SINK, Lisa (2005): "Couple accused of coercing maid face charge of human trafficking", *Milwaukee Journal Sentinel*, Wisconsin, March, 16 th, Journal Sentinel Inc., p. 5. De la misma autora en el mismo medio

SOPEMI (1995): *Trends in International Migration*, Paris, OECD.

SOPEMI (1995): *Trends in International Migration*, Paris, OECD.

STERN, Seth (2005): "An Uneasy Deal With Illegal Workforce", *CQ Weekly*, Washington, March 14, vol. 63, issue 11, pp. 620-626.

STOLL, Michael, MELENDEZ, Edwin, VALENZUELA, Abel (2002): "Spacial Job Search and Job competition among immigrant and native groups in Los Angeles", *Regional Studies*, Los Angeles, California, April, vol. 36, issue 2, pp. 97-112.

SURO, Roberto (2005): "A Growing Minority: Hispanic Americans", *World Almanac and Book of Facts*, *DC*, Primedia Reference Inc. Special Section: Hispanic Americans.

TAYLOR, Peter Shawn (2005): "Help wanted", *Canadian Business*, Toronto, March 14-27, vol. 78, issue 6, pp. 29-34.

THE ECONOMIST (2005): "Finance and Economics: Displacement activity", *The Economist*, London, July 9, vol. 376, issue 8434, pp. 66 ss.

THE ECONOMIST (2002): Survey: "Roll out the welcome mat", *The Economist*, London, October 12, vol. 365, issue 8294, pp. 7-8.

TOBIAS, C. (1997): "Women´s Strategies and the Family-Employment Relationship in Spain", paper presented to research seminar on "Family, Identity and Society in Europe", Bath University, 28-29 November.

TOGNETTI BORDOGNA, M. (1994): "La famiglia che cambia", en Vicarelli, G. (ed.), *Le mani invisibili. La vita e il laboro delle donne immigrate*, Rome: Ediesse, pp. 128-40.

TRAN, Lien, PERLOFF, Jeffrey, (2002) : "Turnover in U.S. Agricultural Labor Markets", *American Journal of Agricultural Economics*, May, vol. 84, issue 2, pp. 427-437.

TREJO, Stephen (2003): "Intergenerational Progress of Mexican-Origin Workers in the U.S. Labor Market", *The Journal of Human Resources*, Summer, vol. 38, issue 3, pp. 467=489.

VENTURINI, Alessandra (2004): "Do Illegal Migrants Compete with National Workers?", *Intereconomics*, January/February 2004.

VICENS, José (2005): "Datos Básicos de la inmigración en la Comunidad Autónoma de Madrid. Resultados de una experiencia DELPHI", *Colección: Documentos de trabajo*, núm. 11, Instituto L.R. Klein-Centro Gauss. U.A.M..DT, Madrid.

WALKER, A., MALTBY, T. (1997): *Ageing Europe*, Buckingham, Open University Press.

WEINSTEIN, Eric (2002): "Migration for the benefit of all: Towards a new paradigm for economic immigration", *International Labour Review*, vol. 141, issue 3, pp. 225-252.

WEISS, Yoram, SAUER, Robert, GOTLIBOVSKI, Menachem 2003 : "Immigration, Search, and Loss of Skill", *Journal of Labor Economics*, July, vol. 21, issue 3, pp. 557-591.

ZAVODNY, Madeline (2003): "The H-1B program and its effects on information technology workers", *Economic Review*, Federal Reserve Bank of Atlanta, Atlanta, Third Quarter, pp. 33-43.

OFERTA LABORAL Y DEMANDA DE VIVIENDA EN ESPAÑA: UN ANÁLISIS MICROECONÓMICO

Mª CONSUELO COLOM ANDRÉS
E-mail: Consuelo.Colom@uv.es
Mª CRUZ MOLÉS MACHÍ
E-mail: Cruz.Moles@uv.es
Departamento de Economía Aplicada
Universitat de València

Resumen

El objetivo de este trabajo es analizar las decisiones que en materia de vivienda realiza un hogar, junto con las decisiones laborales de los cónyuges. En una unidad familiar (hogar) hay una estrecha relación entre las decisiones laborales de sus miembros y las decisiones de consumo, en particular del consumo de vivienda. El conocimiento de este tipo de relaciones ayudará a conocer el comportamiento futuro de los hogares frente a posibles cambios laborales de alguno de los miembros de la pareja o frente a cambios relacionados con la vivienda, y permitirá además realizar mejores evaluaciones de políticas de impuestos y bienestar social.

Palabras clave: demanda de vivienda, oferta laboral familiar.
Área Temática: Economía Urbana, regional y local

Abstract

This paper focuses on the housing decisions that a household realizes together with the labor decisions of the spouses. In a family unit (i.e. a household) there is a strengthened relationship between the labor decisions of partners and consumption decisions, especially housing consumption. Understanding these relationships will help to predict the household's behavior in the face of possible changes. Changes may include both the labor prospect of one member of a couple and changes related to the housing. An understanding of these changes will allow better evaluations of tax and social well-being policies.

Keywords: housing demand, family labor supply.
Thematic Area: Urban, Regional and Local Economics.

1. INTRODUCCIÓN

El mercado de la vivienda y el mercado laboral están fuertemente relacionados. La incidencia del mercado laboral sobre el mercado de la vivienda es a través del efecto que posee la renta permanente o renta a largo plazo sobre las decisiones individuales en materia de vivienda. Por otra parte, el mercado de vivienda tiene importantes efectos sobre el bienestar social, además de implicaciones macroeconómicas notables, como por ejemplo, sobre la movilidad geográfica de los trabajadores.

Un aspecto de gran interés sobre las relaciones de ambos mercados, vivienda y laboral, es el análisis de las decisiones que en materia de vivienda realiza un hogar, junto con las decisiones laborales de los cónyuges, ya que hay una estrecha relación entre las decisiones laborales de sus miembros y las decisiones de consumo, en particular del consumo de vivienda.

En la literatura econométrica podemos encontrar un gran número de trabajos que analizan el comportamiento de los hogares frente a decisiones sobre la vivienda, principalmente la elección del régimen de tenencia y la demanda de servicios vivienda, aunque algunos introducen otros aspectos importantes como movilidad, localización geográfica o se centran en los jóvenes.

Cabría citar, entre otros, a Lee y Trost (1978), Edin and Englud (1991), Rapaport (1997), Goodman (1988, 2002), Colom y Molés (1998) y Colom et al (2002).

Respecto a los trabajos en el mercado laboral, Blundell and MaCurdy (1999) ofrecen una revisión de las aproximaciones existentes a los modelos de oferta laboral. Algunos ejemplos de estudios que utilizan estos modelos son el de Hausman and Ruud (1984), Blundell y Walker (1986), Browning y Meghir (1991), García y Molina (1998), Segura (2002), y Fernández-Val (2003). En estos trabajos se analiza el comportamiento laboral de los hogares considerando la interdependencia entre ambos esposos y la existencia de una estrecha relación entre las decisiones laborales y la demanda de comodidades.

Conjuntamente las decisiones sobre demanda de vivienda y oferta laboral han sido poco analizadas, aunque cabe destacar los trabajos de Kohlhase (1986) con datos de US y Assadian y Ondrich (1993) con datos de la ciudad de Bogotá.

En este trabajo la cuestión planteada es determinar las implicaciones existentes entre las decisiones laborales del hogar y el gasto que dicho hogar realiza en su vivienda. Dada la interrelación entre el mercado de vivienda y el laboral, es de suponer que el consumo en vivienda que realiza un hogar estará influenciado, por las horas dedicadas al trabajo por cada miembro de la pareja y éstas a la vez dependerán de la cantidad de vivienda que el hogar decida consumir; así como del tiempo de trabajo/ocio del otro individuo que forma la pareja. El análisis se centrará en los hogares españoles propietarios de su vivienda en los que ambos cónyuges trabajan.

El conocimiento de las relaciones existentes entre las decisiones laborales del hogar y el consumo de vivienda del mismo ayudará a conocer el comportamiento futuro de los hogares frente a posibles cambios laborales de alguno de los miembros de la pareja o frente a cambios relacionados con la vivienda, y permitirá además realizar mejores evaluaciones de políticas de impuestos y bienestar social.

En el siguiente apartado se presenta el modelo econométrico utilizado. El epígrafe 3 recoge la descripción de los datos y de las variables del análisis y en el epígrafe 4 se comentan los resultados obtenidos en la estimación del modelo. En el último apartado se presentan las principales conclusiones obtenidas en el trabajo.

2. MODELO Y ESTIMACIÓN

En este epígrafe se presenta un modelo que considera que las decisiones laborales de cada miembro de la pareja están condicionadas a las características y decisiones del otro cónyuge (participación o no en el mercado laboral y número de horas trabajadas), así como a las decisiones de consumo de vivienda que tome el hogar. Del mismo modo también las decisiones de vivienda (régimen de tenencia y consumo de servicios de vivienda) estarán determinadas por la oferta laboral de cada uno de los integrantes de la pareja.

El modelo econométrico utilizado para analizar el comportamiento de los hogares españoles frente a las decisiones laborales y de vivienda está formado por cinco ecuaciones simultáneas, las dos primeras representan las decisiones discretas de elección del régimen de tenencia y participación o no de la mujer en el mercado laboral, la tercera recoge la demanda de servicios de vivienda y las dos últimas son las ecuaciones de oferta laboral (número de horas que trabaja el marido y la esposa).

Para modelizar las dos decisiones discretas se considera un modelo Probit Bivariante definido a partir de dos variables latentes: $I_T^* = Z_T'\delta_T + \upsilon_T$ y $I_P^* = Z_P'\delta_P + \upsilon_P$. Estas variables no son observables, pero sí lo es su realización dicotómica definida con las variables I_T que toma el valor 1 si el hogar elige ser propietario de su vivienda y el valor 0 si decide alquilarla, e I_P cuyo valor 1 indica que la mujer participa en el mercado laboral y el 0 indica que no participa:

$$I_T = \begin{cases} 1 & si \quad I_T^* = Z_T'\delta_T + \upsilon_T \geq 0 \\ 0 & en\ otro\ caso \end{cases} \qquad (1)$$

$$I_P = \begin{cases} 1 & si \quad I_P^* = Z_P'\delta_P + \upsilon_P \geq 0 \\ 0 & en\ otro\ caso \end{cases} \qquad (2)$$

siendo Z_T y Z_P vectores de características socio-demográficas y económicas del hogar y (υ_T, υ_P) la componente aleatoria del modelo bivariante de tenencia y participación[1].

Para determinar la demanda de servicios vivienda y la oferta laboral de ambos cónyuges, asumiendo una especificación loglineal en estas ecuaciones se tiene:

$$ln\, D = \alpha_0 + \alpha_1 X_D + \alpha_2\, ln\, h_1 + \alpha_3\, ln\, h_2 + \alpha_4 p_e + \alpha_5 Y + \varepsilon_D \qquad (3)$$

$$ln\, h_1 = \beta_0 + \beta_1 X_1 + \beta_2 D + \beta_3\, ln\, h_2 + \beta_4 w_1 + \beta_5 Y + \varepsilon_1 \qquad (4)$$

$$ln\, h_2 = \gamma_0 + \gamma_1 X_2 + \gamma_2 D + \gamma_3\, ln\, h_1 + \gamma_4 w_2 + \gamma_5 Y + \varepsilon_2 \qquad (5)$$

donde D es la demanda de servicios de vivienda, h_1 y h_2 las horas de trabajo del marido y la esposa, X_D, X_1 y X_2 las características del hogar y sus miembros en las ecuaciones respectivas, p_e el precio de la vivienda, w_1 y w_2 los correspondientes salarios por hora de ambos cónyuges e Y la renta del hogar.

El modelo conjunto viene representado por las anteriores ecuaciones, (1)-(5). Para analizar simultáneamente estas decisiones se admitirá que las perturbaciones aleatorias $(\upsilon_T, \upsilon_P, \varepsilon_{D_v}, \varepsilon_1, \varepsilon_2)$ siguen una distribución Normal Multivariante con vector de medias nulo y una matriz de varianzas-covarianzas que considera correlación entre dichas perturbaciones aleatorias.

Debido a la selección muestral intrínseca del modelo, su estimación debe realizarse mediante un procedimiento en dos etapas. En la primera, se estiman las dos decisiones discretas (tenencia y participación) mediante un modelo Probit Bivariante por máxima-verosimilitud con información completa. En la segunda etapa, para la muestra de hogares que han elegido ser propietarios de su vivienda y la mujer ha decidido participar en el mercado laboral, se estiman las ecuaciones simultáneas de oferta laboral y demanda de vivienda mediante un proceso de mínimos cuadrados en tres etapas, 3SLS, dada la endogeneidad existente entre las variables, e incorporando los factores de corrección del sesgo correspondiente, λ^T y λ^P, calculados desde la estimación del modelo Probit Bivariante[2].

3. DATOS Y VARIABLES

Los datos utilizados en este trabajo provienen del Panel de Hogares de la Unión Europea (PHOGUE) para el año 2000. De todos los hogares disponibles se han seleccionado aquellos que disfrutan de su vivienda principal en régimen de propiedad o alquiler y están formados por una pareja en la que ambos cónyuges están en edad laboral (menos de 65 años) y además el hombre participa activamente en el mercado de trabajo. De estos hogares se han eliminado aquellos con errores de medida o falta de información en alguna de las variables necesarias en el análisis. La muestra resultante está formada por 5 754 hogares.

A continuación se comentan las variables que intervienen en cada una de las ecuaciones del modelo. En la tabla 1 se presenta la descripción de las mismas.

[1] La componente aleatoria se supone que sigue una distribución Normal Bivariante con coeficiente de correlación ρ.

[2] Es una extensión del método de estimación de Heckman de un modelo de regresión con selección muestral (Heckman, 1979), pero considerando un doble criterio de selección, de ahí la existencia de dos factores de corrección. Estas variables de selección muestral se calculan de forma similar a la inversa del ratio de Mills

Tabla 1: Descripción de las Variables

VARIABLES	DEFINICIÓN
Dependientes	
Tenencia	Si el régimen de tenencia elegido es la propiedad = 1; si es el alquiler = 0
Participa_M	Si la mujer trabaja = 1; en otro caso = 0
Demanda	Cantidad demandada de servicios de vivienda (en logaritmos)
Horas_H	Número de horas semanales trabajadas por el hombre (en logaritmos)
Horas_M	Número de horas semanales trabajadas por la mujer (en logaritmos)
Independientes	
Sexo	Si el sustentador principal es varón = 1; si es mujer = 0
*Edad**	Edad en años del individuo (en forma cuadrática)
*Analfabeto**	Individuo analfabeto o sin estudios = 1; en otro caso = 0 (variable de referencia)
*Primarios**	Individuo con estudios primarios = 1; en otro caso = 0
*Secundarios**	Individuo con estudios secundarios o FP2 = 1; en otro caso = 0
*Universitarios**	Individuo con estudios universitarios = 1; en otro caso = 0
Miembros	Número de miembros del hogar (incluyendo al sustentador principal)
*Cuida**	Si el individuo (hombre o mujer) tiene que cuidar niños y/o personas mayores = 1; en otro caso = 0
*Cualificación**	Si el individuo (hombre o mujer) desempeña un trabajo con alta cualificación = 1; en otro caso = 0
Renta	Renta anual disponible del hogar (en logaritmos)
Ahorro	Si el hogar puede ahorrar a final de mes = 1; en otro caso = 0
Ing_salar_H	Ingresos anuales salariales del hombre (en logaritmos)
Ing_nosalar	Ingresos anuales no salariales del hogar (en logaritmos)
Salario_H	Salario por hora del hombre (en logaritmos)
Salario_M	Salario por hora de la mujer (en logaritmos)
Precio	Precio por metro cuadrado de compra de la vivienda (en logaritmos)
Ind_precio	Logaritmo del cociente entre el precio de compra y el del alquiler
Tmuni1	Si el hogar reside en un municipio con menos de 10 000 habitantes = 1; en otro caso = 0
Tmuni2	Si el hogar reside en un municipio entre 10 001 y 50 000 habitantes = 1; en otro caso = 0
Tmuni3	Si el hogar reside en un municipio con más de 50 000 habitantes y no es capital de provincia = 1; en otro caso = 0
Tmuni4	Si el hogar reside en un municipio capital de provincia (variable de referencia)
Murb	Si el hogar reside en una zona urbana (más de 10 000 habitantes) = 1; en otro caso = 0

* El individuo representa al sustentador principal (SP), al hombre (H) o a la mujer (M) según el caso.

Elección del Régimen de Tenencia y Decisión de Participación de la Mujer en el Mercado Laboral

Para la elección del régimen de tenencia se han utilizado como variables explicativas el sexo, la edad (en forma cuadrática) y el nivel de estudios del sustentador principal, esta última introducida en el modelo con cuatro variables ficticias que recogen los siguientes niveles académicos: analfabetos o sin estudios, estudios primarios, estudios secundarios y estudios universitarios. Además para recoger las características demográficas del hogar se ha considerado el número de miembros que lo integran.

Como factores económicos se incluye la renta disponible del hogar y un indicador de la capacidad de ahorro del mismo, así como el precio de las viviendas a través del cociente entre el precio de compra y el precio de alquiler, *ind_precio*. El precio de compra está proporcionado por el Ministerio de Fomento e indica el precio medio por metro cuadrado por Comunidad Autónoma y para el precio de alquiler puesto que no existe ninguna relación oficial del mismo, se ha realizado una aproximación a partir del alquiler anual declarado por los propios hogares de la encuesta, asignando a cada uno el valor medio del alquiler por metro cuadrado de su Comunidad Autónoma. Además se incluyen unas variables ficticias indicadoras del tamaño del municipio en que reside el hogar *tmuni1-tmuni4*, para matizar la variación de precios existente entre municipios con mayor o menor densidad de población.

En lo referente a la decisión de participación de la mujer en el mercado de trabajo, se consideran como variables explicativas la edad (en forma cuadrática) y el nivel de estudios alcanzados por la mujer, el número de miembros del hogar y una variable que indica si la mujer tiene a su cuidado a niños y/o personas mayores o necesitadas. También se consideran dos variables económicas, la renta salarial del marido y el total de ingresos no salariales del hogar. Por último se incluye una variable que recoge si el hogar vive en una zona urbana o rural, para reflejar posibles variaciones en el comportamiento de los hogares asociadas a diferentes culturas y/o hábitos de vida (diferencias sociológicas), así como diferencias en niveles de precios de la vivienda.

Demanda de Servicios de Vivienda y Oferta Laboral

La variable dependiente que recoge la demanda de servicios de vivienda (*demanda*) está definida como el alquiler mensual imputado que declara el hogar. Sus valores son euros y está considerada en logaritmos.

En esta decisión como variables explicativas se consideran la edad (en forma cuadrática) y el nivel de estudios del sustentador principal, el número de miembros del hogar, la renta del hogar y el precio de compra de la vivienda (estas dos últimas variables en logaritmos). También se incluye la variable que indica si el hogar reside en un ámbito urbano o rural.

La oferta laboral se analizará a partir del número de horas de trabajo del hombre y de la mujer. Ambas variables dependientes se definen como las horas de trabajo semanales y se introducen en logaritmos.

Las variables explicativas utilizadas son la edad y el nivel de estudios del individuo, para las que se espera el mismo comportamiento entre hombres y mujeres, el número de miembros que integran el hogar y si el individuo tiene que cuidar de niños o adultos necesitados. Estas dos últimas variables pueden tener efectos distintos en la determinación del número de horas trabajadas por los hombres y las mujeres. También se incluye una variable que recoge si el individuo está o no realizando tareas con alta cualificación[3].

Las características económicas se recogen con el total de ingresos no salariales del hogar y el salario por hora que recibe cada uno de los cónyuges, definido como el cociente entre los ingresos anuales que percibe el individuo por rentas del trabajo y el número de horas anuales trabajadas, tanto para el marido como para la esposa. Para estas variables se espera una influencia negativa. Un elevado salario permitirá al individuo trabajar menos horas para conseguir mayores ingresos salariales, aumentando por lo tanto los ingresos del hogar, y permitiendo al mismo tiempo que el otro cónyuge pueda reducir su jornada laboral manteniendo un determinado nivel de ingresos.

Para realizar la estimación del modelo planteado en la sección anterior, nos encontramos con el problema de la posible endogeneidad de la variable salario por hora[4]. Para evitar este problema se considera que el salario real del individuo es función de características observables propias del individuo y del entorno, y que el salario observado es este salario real con un error de medida. Por lo tanto, una regresión por mínimos cuadrados ordinarios del salario observado nos dará una medida exógena del salario real y será la utilizada en las ecuaciones de oferta laboral. Los resultados de dicha estimación se encuentran en el apéndice A.

4. RESULTADOS

En este apartado se comentan los resultados de la estimación del modelo definido por las ecuaciones (1)-(5). En la tabla 2 se presentan los resultados del modelo Probit Bivariante para las elecciones del régimen de tenencia y la participación o no de la mujer en el mercado laboral. En la tabla 3 se presenta la estimación de las ecuaciones simultáneas de demanda de servicios de

[3] Se considera trabajo de alta cualificación si corresponde al Gran Grupo 1, 2 ó 3 de la Clasificación Nacional de Ocupaciones (directivos, técnicos y profesionales). Son los códigos 11, 12, 13, 21, 22, 23, 24, 31, 32, 33 y 34 de la Clasificación Internacional.

[4] Ver Fernández-Val (2003) para la argumentación de este problema.

vivienda y oferta laboral familiar, realizada únicamente para los hogares que son propietarios de su vivienda y trabajan ambos cónyuges.

La primer conclusión que se desprende del análisis es que es adecuado realizar la doble corrección del sesgo de selección asociado a elección del régimen de tenencia y a la participación o no de la mujer en el mercado de trabajo, puesto que las variables, λ^T y λ^P, que recogen estos efectos han resultado significativas en la estimación conjunta del modelo (ver tabla 3).

Elección del Régimen de Tenencia y Decisión de Participación de la Mujer en el Mercado Laboral

Para la ecuación del régimen de tenencia los resultados obtenidos coinciden con los habitualmente encontrados en los estudios sobre dicha elección. De las variables que recogen las características del sustentador principal y la composición del hogar, afectan de forma significativa el sexo y la edad; mientras que los estudios y el número de miembros del hogar no son determinantes. El *sexo* tiene un coeficiente estimado negativo que indica que las mujeres son más proclives a la compra que los hombres y la edad tiene un impacto no monótono (efecto cóncavo) en la probabilidad de ser propietario de la vivienda. Específicamente, esta probabilidad inicialmente crece con la edad (a un ratio decreciente) y después cae, tal como sugieren los signos positivo y negativo de las variables *edad* y *edad2*, respectivamente.

Tabla 2: Modelo Probit Bivariante: régimen de tenencia y participación de la mujer en el mercado laboral

Régimen de tenencia			**Participación de la mujer**		
Variables	Coeficiente	Estadístico t	**Variables**	Coeficiente	Estadístico t
Constante	-2.8744	-4.724	*Constante*	-0.5554	-1.398
Sexo	-0.1776	-1.721	*Edad**	0.0766	4.767
*Edad**	0.0823	3.783	*Edad2**	-0.0011	-5.947
*Edad2**	-0.0006	-2.536	*Primarios**	0.1713	1.969
*Primarios**	0.0809	0.712	*Secundarios**	0.6450	6.686
*Secundarios**	0.2186	1.723	*Universitarios**	1.3550	13.463
*Universitarios**	-0.0132	-0.101	*Miembros*	-0.0936	-4.974
Miembros	0.0120	0.519	*Cuida**	-0.2488	-6.236
Renta	0.1730	3.599	*Ing_salar_H*	-0.0940	-3.054
Ahorro	0.6624	12.650	*Ing_nosal*	0.0212	2.411
Ind_precio	-0.3979	-2.387	*Murb*	0.0507	1.202
Tmuni1	0.4299	6.080			
Tmuni2	0.3000	4.507			
Tmuni3	0.1829	2.589			
ρ	0.0542	1.684			
Num. observac.	5 754				

* En la ecuación de tenencia de vivienda representa las características del sustentador principal y en la ecuación de participación de la mujer en el mercado laboral representa las características de la mujer.

En cuanto a las variables económicas, la renta disponible del hogar y la capacidad de ahorro del mismo han resultado significativas, siendo esta última la que presenta mayor influencia en esta decisión. Ambas variables presentan el efecto positivo esperado, los hogares con mayor renta tienen mayor probabilidad de disfrutar de la vivienda en propiedad, lo mismo ocurre con los hogares con mayor capacidad de ahorro.

El índice de precios, *ind_precio*, y las variables indicadoras del municipio de residencia son significativas en esta ecuación. El índice de precios presenta un coeficiente estimado de signo negativo, lo que indica que si el precio de compra de las viviendas aumenta más que el precio de alquiler la probabilidad de ser propietario de la vivienda disminuye. Respecto al tamaño del municipio de residencia del hogar sus coeficientes estimados son positivos y decrecientes. A medida que disminuye el tamaño del municipio, los hogares tienen mayor probabilidad de

comprar la vivienda (el coeficiente con mayor valor es el asociado a *tmuni1*), resultado razonable, ya que en las zonas rurales las preferencias por disfrutar de una vivienda en propiedad son mayores que en el ámbito urbano, tanto por motivos culturales como económicos.

Los resultados obtenidos para la ecuación de participación de la mujer en el mercado laboral, en general, son los esperados. Tanto la edad como el nivel de estudios de la mujer afectan a la decisión y su peso es muy importante. Para la edad se ha obtenido una parábola negativa, indicando que la participación de la mujer en el mercado de trabajo aumenta con la edad, hasta llegar a un punto máximo, alrededor de 35 años, y a partir de ahí esta tendencia decrece. Este hecho podría explicarse porque las mujeres jóvenes aún se encuentran estudiando y no se han incorporado al mercado de trabajo y por otra parte, porque a partir de cierta edad las mujeres abandonan el mercado laboral para dedicarse a su familia.

Los coeficientes estimados para el nivel de estudios son positivos y crecientes. A medida que aumenta el nivel educativo de la mujer la probabilidad de participación de ésta en el mercado de trabajo crece. Esto confirma que uno de los principales factores explicativos del aumento en la actividad económica de las mujeres es el cambio del perfil educativo en los últimos años.

El número de miembros del hogar afecta de manera negativa, indicando que cuanto mayor es el número de miembros que forman el hogar menor es la probabilidad de que la mujer trabaje. La razón que avala este hecho es el tradicional rol de la mujer en la familia: si hay más miembros en el hogar la mujer debe dedicarse más tiempo al mantenimiento del hogar y al trabajo doméstico (actividades domésticas no remuneradas).

La variable que indica si la mujer cuida o no a niños y/o ancianos (*cuida*) tiene un efecto significativo y negativo, así la mujer con cargas familiares tiene menor probabilidad de trabajar fuera de casa.

En cuanto a las variables económicas, la renta salarial del esposo y la renta no salarial del hogar son significativas. La primera con un coeficiente estimado negativo, a mayores ingresos salariales del marido menor es la probabilidad de que la mujer esté en el mercado laboral. El salario del hombre desincentivará a la mujer a trabajar. Sin embargo el efecto de la renta no salarial del hogar sorprendentemente es contrario al esperado, incrementos en los ingresos no salariales fomentan la participación de la mujer. Finalmente, la variable *murb* no presenta un efecto significativo.

Demanda de servicios de vivienda y Oferta Laboral

Desde la tabla 3 se observa que existe interrelación entre el gasto en vivienda y la oferta laboral que realizan los hogares; ya que la demanda de vivienda es determinante en las ecuaciones de oferta laboral y recíprocamente. Se observa que la demanda de servicios de vivienda aumenta cuando el hombre y/o la mujer incrementan sus horas de trabajo. Así, los mayores ingresos que tiene el individuo, por aumentar su jornada laboral, el hogar lo destina a incrementar el gasto en servicios de vivienda. A su vez, la cantidad de servicios de vivienda demandados por el hogar afecta positivamente a la oferta laboral del hogar. Si hay más demanda en vivienda se tendrá que ambos cónyuges aumentan sus horas de trabajo para poder contribuir económicamente a cubrir los gastos ocasionados por el incremento en dicha demanda.

El número de horas que trabaja la esposa tiene un efecto importante en el número de horas que trabaja su cónyuge: si la mujer incrementa su jornada laboral el hombre la disminuye. Sin embargo, el número de horas trabajadas por la mujer no se ve afectado por la oferta laboral del hombre.

Los resultados de la estimación de la ecuación de demanda en servicios de vivienda indican que de las variables explicativas, la edad del sustentador principal es la única variable no significativa.

Los coeficientes estimados para el nivel de estudios son positivos y crecientes. Los hogares cuyo cabeza de familia es universitario son los que presentan mayor gasto en servicios de vivienda. El número de miembros del hogar tiene un efecto positivo, el gasto en vivienda aumenta con el número de miembros. Si el hogar crece también crecen las necesidades del

mismo respecto a la dimensión de la vivienda y a otros gastos relacionados con la casa, que hacen aumentar el consumo en vivienda.

Tabla 3: Demanda en servicios de vivienda y número de horas trabajadas por el hombre y la mujer en hogares propietarios de su vivienda y ambos en el mercado laboral

	Demanda de servicios de vivienda		**Número de horas trabajadas por el hombre**		**Número de horas trabajadas por la mujer**	
Variables	Coeficiente	Estadíst. t	Coeficiente	Estadíst. t	Coeficiente	Estadíst. t
Constante	-0.9388	-1.269	4.0357	15.217	2.8655	3.785
*Edad**	-0.0001	-0.003	0.0105	1.890	-0.0232	-1.744
*Edad2**	0.0001	0.434	-0.0001	-1.881	0.0003	1.667
*Primarios**	0.1044	2.204	0.0101	0.346	-0.1076	-1.947
*Secundarios**	0.2119	4.031	0.0125	0.353	-0.2331	-2.529
*Universitarios**	0.3341	5.995	0.0550	1.208	-0.3990	-2.813
Miembros	0.0311	2.948	-0.0073	-1.191	0.0004	0.032
*Cuida**	-------	-------	-0.0359	-3.459	-0.0082	-0.292
*Cualificación**	-------	-------	0.0786	6.761	0.0797	3.645
Renta	0.2279	8.988	-------	-------	-------	-------
Salario_H	-------	-------	-0.3019	-6.781	-0.0722	-1.150
Salario_M	-------	-------	-0.0392	-1.771	0.1042	2.015
Ing_nosalar	-------	-------	-0.0098	-3.969	-0.0117	-2.477
Precio	0.3742	14.873	-------	-------	-------	-------
Murb	0.0789	3.842	-------	-------	-------	-------
Demanda	-------	-------	0.1999	5.968	0.1810	2.566
Horas_H	0.3375	2.081	-------	-------	0.1278	0.536
Horas_M	0.4009	2.898	-0.3084	-4.473	-------	-------
λ^T	0.0694	0.689	-0.0579	-1.015	-0.3006	-2.877
λ^P	-0.1277	-4.155	0.0409	1.621	-0.1885	-1.407
Num. observac	2 243					

* En la ecuación de demanda de vivienda representa las características del sustentador principal y en las ecuaciones de oferta laboral representa las características del hombre o de la mujer según el caso.

Las variables económicas renta del hogar y precio de las viviendas son muy determinantes. Las estimaciones de estos coeficientes corresponden a las elasticidades renta y precio y ambas presentan los signos positivos esperados, aunque su magnitud es inferior a la unidad. Así, un aumento en la renta disponible de los hogares repercute en una mayor demanda de servicios de vivienda. Para el precio de las viviendas, cuando éste crece aumenta el gasto en vivienda, esto puede explicarse ya que, bajo las mismas condiciones, el hogar deberá pagar más dinero por su vivienda.

Para la variable *murb*, que también recoge el efecto del precio de las viviendas, se ha obtenido un coeficiente estimado positivo. Los hogares residentes en zonas urbanas gastan más en servicios de vivienda; ello puede ser debido al hecho de que en las grandes ciudades los precios de las viviendas y todos los gastos que éstas llevan consigo (contribuciones, etc…) son mayores que en los pueblos pequeños.

Para la oferta laboral los resultados obtenidos muestran diferencias de comportamiento entre el hombre y la mujer. La edad presenta un comportamiento de una parábola negativa en el hombre que indica que el número de horas trabajadas por éste aumenta aunque, dado el signo negativo de la forma cuadrática, ese incremento con cada año adicional es decreciente. Mientras que en el caso de las mujeres la edad presenta una parábola positiva que indica que el número de horas trabajadas por la mujer disminuye cuando aumenta la edad, llegando su mínimo alrededor de los 38 años.

El nivel de estudios no afecta al número de horas trabajadas por el hombre y sí que es determinante para la mujer, con coeficientes estimados negativos y decrecientes, mostrando que a medida que crece el nivel de formación de la mujer, ésta tiende a trabajar menos horas.

La variable *cuida* que nos dice si el individuo tiene que cuidar o no a niños y/o ancianos sólo es determinante en el caso de los hombres y presenta un coeficiente estimado negativo. Si el hombre es responsable del cuidado de alguna persona del hogar, se verá reducida su jornada laboral. La variable que recoge la cualificación en el trabajo, *cualificación*, es significativa en ambas submuestras y su influencia es positiva, lo que indica que los individuos con trabajos altamente cualificados trabajan más horas.

Con respecto a los factores económicos vemos que en el caso de los hombres, el salario por hora, tanto el propio como el de la esposa, afecta de manera negativa. Resultado lógico ya que, con salarios más elevados, trabajando menos horas el hogar puede alcanzar el mismo nivel de ingresos. Sin embargo, para las mujeres se ha obtenido que el salario por hora del marido no es determinante en esta decisión, mientras que sí influye su propio salario por hora y éste actúa como un incentivo laboral, de manera que frente a salarios más elevados, la mujer tenderá a trabajar más horas. Hay que notar también que el efecto de esta variable es contrario al que presenta en la ecuación del número de horas trabajadas por el hombre.

Respecto a los ingresos no salariales, en ambos casos, se observa una influencia negativa. Cuando el hogar incrementa su nivel de renta por ingresos que no están asociados al trabajo, el individuo tenderá a disminuir su oferta laboral, ya que no necesitará trabajar más horas para conseguir el nivel económico deseado.

5. CONCLUSIONES

Con el trabajo hemos aportado evidencia empírica que para los hogares españoles las decisiones de consumo de vivienda y oferta laboral familiar son realizadas conjuntamente.

Un resultado importante del estudio es que las decisiones de demanda de vivienda y oferta laboral familiar están interrelacionadas. Kohlhase (1986) para Estados Unidos y Assadian and Ondrich (1993) para la ciudad de Bogotá obtienen resultados coincidentes. Si las horas dedicadas al trabajo por cada miembro de la pareja se incrementan, el hogar aumenta su gasto en vivienda y recíprocamente cuando la cantidad demandada de vivienda aumenta se incrementa la dedicación laboral de ambos cónyuges o de alguno de ellos. El incremento de los gastos en vivienda, principalmente del precio de la misma, lleva a que los miembros del hogar tiendan a alargar sus jornadas laborales por medio de horas extra para obtener ingresos adicionales. Por otra parte, los períodos en que hay un porcentaje de paro elevado la demanda de vivienda cae.

Los factores económicos son fundamentales en las decisiones sobre la vivienda. Los hogares con mayores ingresos tienden a incrementar el gasto en la misma. En la demanda en servicios de vivienda las elasticidades renta y precio son inferiores a la unidad.

Respecto a las decisiones laborales los factores económicos también tienen un efecto importante. La elasticidad del salario propio en la oferta laboral del individuo presenta resultados opuestos entre hombres y mujeres. Mientras que las mujeres presentan una elasticidad positiva, es decir, el salario por hora de la misma actúa como incentivo laboral, para los hombres se ha obtenido un efecto negativo que indica que un mayor salario por hora induce a una disminución de las horas de trabajo. El hombre se plantea conseguir cierto nivel económico y cuando lo alcanza por tener un mayor salario por hora éste disminuye su jornada laboral. Las elasticidades cruzadas presentan valores inferiores a los de las elasticidades propias y su signo es esperado, cuando se asume que incrementos en los salarios se comparten con la pareja y el ocio es un bien normal.

En el trabajo también se revela que la educación tiene efecto en la demanda de servicios de vivienda y en la oferta laboral de la mujer. Cuanto mayor es el nivel educativo del sustentador principal más se gasta el hogar en su vivienda. Las mujeres universitarias son las que menos horas trabajan.

Todos estos resultados nos confirman que el mercado de la vivienda y el mercado laboral están fuertemente relacionados, tanto a nivel macroeconómico como se ha probado en otros trabajos precedentes (Barrios y Rodríguez, 2004), como a nivel individual (hogar) como se

observa con los resultados de este estudio[5]. Ello nos indica que los políticos deben tener en cuenta todas estas implicaciones, considerando distintos escenarios respecto al comportamiento laboral y en materia de vivienda de los hogares españoles para elaborar mejores políticas económicas (vivienda, impuestos, bienestar social,…).

BIBLIOGRAFÍA

ASSADIAN, A. AND ONDRICH, J (1993): "Residential Location, Housing Demand and Labour Supply Decisions of One- and Two-Earner Households: The Case of Bogota, Colombia", Urban Studies, 30, 73-86.

BARRIOS, J. A. AND RODRÍGUEZ, J. E. (2004): "User Cost Changes, Unemployment and Home-Ownership: Evidence from Spain", Urban Studies, 41, 563-578.

BLUNDELL, R. AND MACURDY, T. (1999): "Labor Supply: A Review of Alternative Approaches", Chap. 27 in Handbook of Labor Economics, Vol. 3, Ed. By O. Ashenfelter and D. Card, Elsevier Science B.V.

BLUNDELL, R. AND WALKER, I. (1986): "A Life-Cycle Consistent Empirical Model of Family Labour Supply Using Cross-Section Data", Review of Economic Studies, 53, 539-558.

BROWNING, M. AND MEGHIR, C. (1991): "The Effects of Male and Female Labor Supply on Commodity Demands", Econometrica, 59, 925-951.

COLOM, M.C. AND MOLÉS, M.C. (1998): "Un Análisis sobre el Gasto en Servicios de Vivienda en España", Estadística Española, 143, 147-166.

COLOM, M.C.; MARTÍNEZ, R. AND MOLÉS, M.C. (2002): "Un Análisis de las Decisiones de Formación de Hogar, Tenencia y Demanda de Servicios de Vivienda de los Jóvenes Españoles", Moneda y Crédito, 215, 105-128.

EDIN, P. A. AND ENGLUND. P. (1991): "Moving Costs and Housing Demand. Are Recent Movers Really in Equilibrium?", Journal of Public Economics, 44, 299-320.

FERNÁNDEZ-VAL, I. (2003): "Household Labor Supply: Evidence for Spain", Investigaciones Económicas, 27, 239-275.

GARCÍA, I AND MOLINA, J. A. (1998): "Household Labour Supply with Rationing in Spain", Applied Economics, 30, 1557-1570.

GOODMAN, A.C., (1988): "An Econometric Model of Housing Price, Permanent Income, Tenure Choice, and Housing Demand", Journal of Urban and Economics, 23, 327-353.

GOODMAN, A.C., (2002): "Estimating Equilibrium Housing Demand for Stayers", Journal of Urban Economics, 51, 1-24

HAUSMAN, J. AND RUUD, P. (1984): "Family Labor Supply with Taxes", American Economic Review, 74, 242-248.

HECKMAN, J. J. (1979): "Sample Selection Bias as a Specification error", Econometrica, 47, 153-162.

KOHLHASE, J. E. (1986): "Labor Supply and Housing Demand for One- and Two-Earner Households", The Review of Economics and Statistics, 68, 48-57.

LEE, L-F. AND TROST, R. P. (1978): "Estimation of some Limited Dependent Variable Models with Application to Housing Demand", Journal of Econometrics, 8, 357-382.

RAPAPORT, C. (1997): "Housing Demand and Community Choice: an Empirical Analysis", Journal of Urban Economics, 42, 243-260.

SEGURA, M. (2002): "Comportamiento (condicionado) de las Parejas Casadas en Materia Laboral. Una Evidencia para el Caso Español", Cuadernos Económicos del ICE, 66, 153-181.

STOKER, T.M. (1993) "Empirical Approaches to the Problem of Aggregation Over Individuals", Journal of Economic Literature, 31, 1827-1874.

[5]Stoker (1993) indica que los estudios microeconómicos proporcionan mejor información en demanda de vivienda que los macroeconómicos.

APÉNDICE A

La estimación de los salarios por hora para hombres y mujeres se ha realizado mediante el procedimiento en dos etapas de Heckman que permite corregir el sesgo de selección asociado a la participación del individuo en el mercado laboral[6]. En la tabla A1 están los resultados de dicha estimación.

La variable dependiente, salario por hora, se define como el cociente entre los ingresos anuales que percibe el individuo por rentas del trabajo y el número de horas anuales trabajadas. Como variables explicativas se han considerado las habituales de capital humano (edad y nivel de estudios del individuo y si trabaja en el sector público o privado), junto con variables indicadoras de la región en la que reside el individuo que permiten captar las desigualdades económicas regionales.

Tabla A1: Ecuaciones salariales para hombres y mujeres

	Salario del Hombre		**Salario de la Mujer**	
Variables	Coeficientes	Estadístico t	Coeficientes	Estadístico t
Constante	0.2243	1.565	-0.2072	-0.838
*Edad**	0.0424	6.633	0.0482	4.198
*Edad2**	-0.0004	-5.625	-0.0005	-3.733
*Primarios**	0.1835	5.453	0.2616	3.536
*Secundarios**	0.3845	10.338	0.5471	6.764
*Universitarios**	0.7029	18.526	0.8196	9.487
*Publico**	0.1904	9.744	0.4102	12.577
Murb	0.0799	4.438	0.0892	2.716
Noroeste	0.0095	0.270	-0.2830	-4.425
Noreste	0.1929	5.602	0.0021	0.034
Centro	-0.0257	-0.751	-0.1246	-1.946
Este	0.1491	4.513	0.0782	1.305
Sur	0.0057	0.168	-0.0892	-1.408
Madrid	0.1863	4.831	0.0279	0.403
λ	-0.4871	-7.031	-0.0586	-1.122

* En la primera ecuación representan las características del hombre; mientras que en la segunda son las características de la mujer.

En general, la mayoría de las variables independientes, tanto para la ecuación salarial del hombre como para la de la mujer, son significativas y con el signo esperado. La variable que recoge el sesgo de selección debido a la participación o no en el mercado de trabajo, λ, es significativa para la ecuación salarial de los hombres; mientras que no lo es la asociada a la ecuación salarial de las mujeres. Para las mujeres su participación o no en el mercado laboral no afecta a su salario por hora.

[6] Los resultados de los modelos probit de participación de hombres y mujeres se pueden solicitar a los autores.

THE INCOME INTER-REGIONAL REDISTRIBUTION AND THE INCOME SPATIAL STABILIZATION EFFECTS: AN APPLICATION TO PORTUGAL

CARLA CRISTINA COIMBRA
E-mail: carla.coimbra@ine.pt
Instituto Nacional de Estatística

PEDRO NOGUEIRA RAMOS
E-mail: pnramos@fe.uc.pt
Faculdade de Economia
Universidade de Coimbra

Resumo

A ideia base desta comunicação é que embora o PIB regional, per capita, seja um bom indicador de rendimento segundo o seu local de geração, não tem em conta o processo redistributivo, também relevante no plano regional, que ocorre após a produção. É examinado neste contexto o processo de transformação do produto regional no rendimento disponível ajustado das famílias, assim como os diferentes fluxos redistributivos que o suportam. O artigo tem duas partes. A primeira é uma análise sincrónica (que versa sobre o efeito de redistribuição inter-regional do rendimento) em que a questão é como as desigualdades, patentes nos níveis do PIB, decorrentes das diferentes localizações da produção, vêm atenuadas quando o rendimento é percebido pelas famílias. A segunda parte é uma análise dinâmica (relativa à estabilização espacial do rendimento) onde o propósito é discutir como são absorvidos espacialmente os choques que porventura atinjam o produto de uma particular região.

Palavras chave: Redistribuição inter-regional; estabilização espacial, *risk sharing*, variância *cross-section.*
Área temática: Economia Urbana, Regional e Local.

Abstract

The main idea of this paper is that although the per capita regional GDP is a good indicator of regional income according with its place of generation, it cannot take into account the redistribution process – also significant among regions – that comes after production. The paper focuses on the transformation of the regional product on the adjusted disposable income of the households, and on the different redistributive flows that proceed with it. It has two distinct parts. Firstly we adopted a synchronic analysis (the income inter-regional redistribution effect) where we wonder how the regional income inequalities that are the outcome of the different locations of production are smoothed when the income falls into the hands of the households. In a second part we turn to a dynamic approach (the spatial stabilization effect) where the purpose is to look at the shocks on the regional product, and to discuss how they are absorbed.

Key Words: Inter-regional redistribution; spatial stabilization; risk sharing; cross-sectional variance.
Thematic Area: Urban, Regional and Local Economics.

1. INTRODUCTION

The Gross Domestic Product per capita (GDPpc) is the most common indicator of economic welfare, both at the countries and at the regional level. In fact, the GDP is the source of all the income produced all over the economy, comprising the yields of all its institutional sectors: households of course, but also societies, general government units, and so on. However, in our view, what really matter at the economies is people – that means households. As a consequence,

in the analysis of income distribution we are going to proceed to, our approach will favor the households' income alone, or the personal income as we call it as well, excluding the other kind of institutions (that furthermore, as a rule, do not account significantly as income recipients at last resort).

On the other hand, even when we already look at personal income what is really important for the households welfare is not the so-called Primary Income (the one that is received by virtue of their direct participation in the production process), but the income that actually households benefit. In fact, beyond the compensations by the use of their production factors, households engage as well in a secondary redistribution process of the income, paying taxes, profiting from social benefits, receiving and paying a miscellaneous of other transfers from/to the other institutional sectors of the economy. It is through this secondary redistribution process that the Primary Income is then converted into the Disposable Income of the Households.

Besides, households' economic welfare does not depend exclusively on monetary income, as the one essentially gauged by their Disposable Income. In several countries, governments (and non-profit institutions) provide a multitude of goods and mainly services in kind that households access free of charge or at a trifling price. In other countries, however people must pay by almost everything they consume. Of course, this makes a difference. This is the reason why the distribution of income in kind, as this one, must be taken into account in view of the households' welfare. When these social transfers in kind are added up the Disposable Income transforms into the Adjusted Disposable Income of the Households.

The main idea of this paper is then that it may exist, remarkably at the regional level, a considerable mismatch between the GDPpc and the income that households actually benefit. The Adjusted Gross Disposable Income of the Households per capita (AGDIHpc)[1] is the measure of income in which we focus our attention on. Ramos (1996) gave two main reasons why regional per capita income measures may be so distinct from GDPpc: one is that commuters, that may reach a huge number among regions, bring their income home, which they generated elsewhere, in other regions. The other is that multiregional or multinational firms operating inside one region, may amount to a significant share of its GDP, but distribute the correspondent incomes in a large scale – dividends, interests and so on – to those residing outside the region or even abroad. Of course, both of these arguments may apply to countries analyses as well, but it is plain that they are of great consequence mainly for regions. Behrens (2003) in a statistical note published by the Eurostat made the same point, concluding that GDPpc in an inappropriate measure of regional economic welfare.

As a matter of a fact, this paper looks at the regional mismatch between the GDPpc and our income variable, the AGDIHpc, adopting two different perspectives. In a first step, we deal with the issue of the extent to which observed regional GDPpc disparities are reproduced (or not) at the AGDIH level, on a synchronic basis. We are also seeking at once which channels smooth those GDP disparities – if this is the case – when that product is distributed to the households, or on the opposite – if that happens – how can those disparities be amplified. We are dealing in this stage with what we called the "inter-regional distribution of income".

In a second part of this paper, we adopted instead a dynamic approach, checking how the households succeed in stabilizing their Adjusted Disposable Income, after their region has been hit by a GDP shock. This means that we are then focusing on the growth rates instead of the levels of the aggregates. We called this analysis of "spatial stabilization of income". This effect is also named risk sharing, as the original shocks at production level spread into other regions while product is transformed into income. We are interested, of course, as well, in the risk sharing channels that are the mechanisms that allow that automatic stabilization over space of the adjusted disposable income.

The purpose of this article is to explore these ideas, applying them to the Portuguese NUTS III regions, on the years of 2002 and 2003. Our analysis is mainly founded in the pioneering works of Asdrubali *et al.* (1996) and Sørensen and Yosha (1998), although these authors focused mostly on the inter-regional risk sharing, or as we call it as well on the income spatial

[1] Primary Income, Secondary Distribution, Disposable Income and Adjusted Disposable Income are all National Accounts standard concepts. The official definitions of these aggregates applying to Portugal are provided by the European System of Accounts (Eurostat, 1996a).

stabilization effect. The dichotomy between inter-regional income distribution and spatial stabilization of income, that we have adopted, was proposed by Mélitz and Zumer (1998) and Decressin (1999). After their early works, Asbrubali, Sørensen and Yosha went on searching the same topic (Asdrubali and Kim, 2004; Sørensen *et al*., 2004; Sørensen *et al.*, 2005). Other relevant contributions may be seen as well in von Hagen and Happ (2001), Kim *et al*. (2003), Jüβen (2006) and Andersson (2008)[2].

Figure 1 ahead summarizes our reasoning scheme (based on the sequences of accounts of official National Accounts systems) that beginning with production, which at last resort generates all the income, goes through the income distribution process until reach the AGDIH. This scheme applies both to the "inter-regional distribution of income" and to the "spatial stabilization of income" effects.

FIGURE 1
From Gross Domestic Product until the
Adjusted Gross Disposable Income of Households

GROSS DOMESTIC PRODUCT (GDP)
↓
PRIMARY INCOME OF HOUSEHOLDS (PIH)
↓
GROSS DISPOSABLE INCOME OF HOUSEHOLDS (GDIH)
↓
ADJUSTED GROSS DISPOSABLE INCOME OF HOUSEHOLDS (AGDIH)

It is important to remark that although all the aggregates of Figure 1 are National Accounts concepts, they are not available at NUTS III regional level (except GDP) in the official Portuguese statistics. Portuguese (and as a rule European) Regional Accounts only reach NUTS II for the majority of these indicators. The data by NUTS III that we deal with in this paper were then built by us, taking advantage of our past or present experience as statisticians in that area. The starting point on these estimates were the Portuguese NUTS II Regional Accounts, having then we followed, for stepping down to the NUTS III level, as close as possible, the same rules and proxies used in the official production of the NUTS II statistics[3]. This is the reason (building our own data is hard work) why this paper confines to only two years: 2002 and 2003.

The following two sections of this paper deal with each one of our two types of effects we considered above: section 2 with the "inter-regional distribution of income" and section 3 with the "spatial stabilization of income" effect. Throughout these sections we describe our empirical approach that consists in decomposing the cross-sectional variances of the level of GDPpc (of its log) and of the GDPpc growth rate. This methodology allowed an identification of the channels through which our two effects act, beyond measuring the actual strength of both. Section 4 summarizes our main conclusions.

[2] The majority of these articles do not remain in their analyses at the households' income level, but they step down into the per capita consumption expenditure. In fact, we might argue that the consumption is the very last purpose of everybody, so it should be our focus. However in Portugal we do not have that information, nor can it be estimated in our opinion in a reliable way. On the other hand, even if we had it would be debatable anyway that we should go so far. Although the consumption is really the final aim of the households, that does not mean it is their purpose at every moment. Future consumption grants welfare as well, when households look over their entire life cycle. As savings is the tool for ensuring that future consumption, then it comes that in a given moment saving itself generates welfare. As a result we may choose to look at AGDIHpc – as we are going to do in fact in this paper – even beside the practical point of the non-existence of proper data for dealing with consumption.

[3] These rules are the scope of several methodological documents: Eurostat (1996b, 1996c and 1999)

2. THE INTER-REGIONAL DISTRIBUTION OF INCOME

This section tracks down a double purpose. On the one hand, we wonder if the regional GDP disparities, at NUTS III level in Portugal, are really lessened when the income fall into the hands of the households, forming their Adjusted Disposable Income. On the other, we aim to gauge the weights of the different channels that may produce that smoothing process (if it happens).

Our methodological approach – inspired in the pioneering works of Asbrubali *et al*. (1996) and Sørensen and Yosha (1998) and in the ensuing literature – is based on a cross-sectional variance analysis.

Our starting point is the identity (1) below, that follows the logical scheme of Figure 1:

$$(1) \qquad \frac{GDP_i}{Pop_i} = \frac{GDP_i}{PIH_i} \cdot \frac{PIH_i}{GDIH_i} \cdot \frac{GDIH_i}{AGDIH_i} \cdot \frac{AGDIH_i}{Pop_i}$$

where i refers to the NUTS III region and Pop means the regional population.
Taking logs in equation (1), we can then reckon the cross sectional variance of GDPpc by:

$$(2) \qquad \mathrm{var}(\log \frac{GDP_i}{Pop_i}) = \mathrm{cov}\{\log \frac{GDP_i}{PIH_i}, \log \frac{GDP_i}{Pop_i}\} +$$
$$+ \mathrm{cov}\{\log \frac{PIH_i}{GDIH_i}, \log \frac{GDP_i}{Pop_i}\} + \mathrm{cov}\{\log \frac{GDIH_i}{AGDIH_i}, \log \frac{GDP_i}{Pop_i}\} +$$
$$+ \mathrm{cov}\{\log \frac{AGDIH_i}{Pop_i}, \log \frac{GDP_i}{Pop_i}\}$$

Dividing both sides of the equality (2) by the total variance of GDPpc, we get:

$$(3) \qquad 1 = \beta_1 + \beta_2 + \beta_3 + \beta_4$$

where:

$$\beta_1 = \frac{\mathrm{cov}\{\log(\frac{GDP_i}{PIH_i}), \log(\frac{GDP_i}{Pop_i})\}}{\mathrm{var}(\log \frac{GDP_i}{Pop_i})}$$

and the other βs are the equivalent ratios for the other covariances. Please note that by definition the βs are as well the slopes of the OLS regressions of the logs of the parts in (1) on the log of GDPpc.

Clearly if AGDIHpc disparities among regions are fully independent of GDPpc asymmetries, then $\mathrm{cov}\{\log \frac{AGDIH_i}{Pop_i}, \log \frac{GDP_i}{Pop_i}\} = 0$ and $\beta_4 = 0$ as well. That would mean that the redistribution income process to households lavishly spread over space, insofar that the locals of generation of the income loose any influence on regional AGDIHpc. When, on the contrary $\beta_4 = 1$, regional GDPpc disparities are fully mirrored on AGDIHpc, and any smoothing process among regions did occur. If in the intermediate case we have $0 < \beta_4 < 1$, the regional distribution of the AGDIH is not fully exempted from the influence of the location of the production process, so the inter-regional redistribution process operates partially only. Thus, $1-\beta_4$ is to be regarded as a measure of the degree at which inter-regional distribution of income actually worked. β_1, β_2 and β_3 (that by (3) equal $1-\beta_4$) are of course the different channels of that inter-regional distribution process.

When β_1 is positive and significantly different from zero, we may conclude that the reduction of disparities among NUTS III, felt at the AGDIH level, has been reached at least in part at the expense of the primary distribution of income. The inequalities in regional GDPpc were absorbed by the variability of the income of the other institutional sectors not distributed to the households, and/or are mixed up among the regions by the cross-participations of the residents in ones regions in other regions' production processes.

When we have (instead or additionally) β_2 positive we may then assert that it was secondary redistribution of income that played that role of reducing disparities. That effect was performed by the taxes and social benefits (and other kind of transfers) system. On its turn, if $\beta_3 > 0$, that means that were the social transfers in kind that softened the original income asymmetries.

Table 1 depicts our estimations of these βs.

Table 1. Decomposition of the cross-sectional variance of log GDPpc, 2002 and 2003, Portuguese NUTS III

Anos	β_1	β_2	β_3	β_4
2002	0,349	0,181	- 0,041	0,511
	(4,563)	(4,160)	(-0,762)	(8,077)
2003	0,391	0,161	- 0,024	0,472
	(4,981)	(3,399)	(-0,513)	(7,483)

In this table, as in the others ahead, t-ratios are given in parentheses. As we expected, by (3), the βs exactly added up 1 in each year (although we did not introduce any a priori constraint). The results are rather similar for both years.

The main conclusion is that it does exist an inter-regional redistribution process of income among Portuguese NUTS III regions. As far as β_4 is at the same time significantly different from 0 and 1, we can then say that the regional GDPpc disparities have an influence, but are largely reduced, when after that distribution process the income take the form of AGDIH. Asymmetries may even have been reduced, at least in 2003, by more than 50%.

As for the shares of the different channels on that inter-regional redistribution process we settled that both the primary distribution and the secondary redistribution of income play a relevant role. The channel of the primary distribution seems to be the more robust one, as β_1 clearly exceeds β_2. That primary distribution of income effect (β_1) happens among institutional sectors (the other sectors variance protecting households) and/or among regions. This process is seemingly more important at the spatial level than the smoothing effect of the taxes/transfers system, given by β_2. On the other hand, it looks as though the social transfers in kind have any contribution to the minimization process of regional income disparities, as β_3 took the wrong signal although non-significantly[4].

As the primary distribution and the secondary redistribution of income (caught by β_1 and β_2) are the main routes that bear the inter-regional benign redistribution of income we confirmed to exist in Portugal, we decided go further and decompose these effects in finest and more disaggregated channels that were able to explain this process. We began by the primary distribution of income that led us from the regional GDP to the PIH. The idea was that the primary distribution consists of:

- subtracting the gross operating surpluses (GOS) generated by the production held by the other institutional units (societies, etc.) that is not distributed in that period to the households
- adding up the property incomes (PI) received, less paid, by the households from the rest of the world, comprising in that concept the other regions of the country and the foreigner countries
- adding up the compensation of employees (CE) received, less paid, from the rest of the world

[4] One reason why social transfers in kind seem not to contribute to the reduction process of disparities may be that in regionalizing that flow we considered the place where the services are rendered and not the actual place of residence of the beneficiaries. In many cases, education or health services, or other transferred in kind to the households, concentrate in some regions – sometimes the urban richest ones – but their benefits surpass the borders of these regions reaching the peripheral poorer other ones.

- subtracting the taxes on production and imports, less subsidies (TPLS), that are incorporated in regional GDPs, evaluated at purchasers prices, but are excluded of course from households' incomes.

Having then in mind that:

$$(4) \qquad PIH_i = GDP_i - GOS_i + PI_i + CE_i - TPLS_i$$

It comes that (1) may be transformed into:

$$(5) \qquad \frac{GDP_i}{Pop_i} = \frac{GDP_i}{GDP_i - GOS_i} \cdot \frac{GDP_i - GOS_i}{GDP_i - GOS_i + PI_i} \cdot$$

$$\cdot \frac{GDP_i - GOS_i + PI_i}{GDP_i - GOS_i + PI_i + CE_i} \cdot \frac{PIB_i - GOS_i + PI_i + CE_i}{PIB_i - GOS_i + PI_i + CE_i - TPLS_i} \cdot$$

$$\cdot \frac{PIH_i}{GDIH_i} \cdot \frac{GDIH_i}{AGDIH_i} \cdot \frac{AGDIH_i}{Pop_i}$$

Taking logs and reckoning the cross-sectional variance of GDPpc in the same way than in (2), we then reach:

$$(6) \qquad 1 = \beta_{11} + \beta_{12} + \beta_{13} + \beta_{14} + \beta_2 + \beta_3 + \beta_4$$

where β_{11} until β_{14} are calculated in the same way than in (3), with the covariances of the logs of the parts in (5) with the log of the GDPpc. Note that $\beta_1 = \beta_{11} + \beta_{12} + \beta_{13} + \beta_{14}$, and the remaining βs shall keep the same values than in our previous estimations reported in Table 1.

Table 2 shows then the decomposition of the cross-sectional variance of the GDPpc, extending further the identification of the channels of the inter-regional redistribution of income, taking into account the different types of flows that lead from GDP to PIH. The β_1 channel splits then in β_{11}, β_{12}, β_{13}, and β_{14}.

Table 2. Decomposition of the cross-sectional variance of log GDPpc, with β_1 split, 2002 and 2003, Portuguese NUTS III

Anos	β_{11}	β_{12}	β_{13}	β_{14}	β_2	β_3	β_4
2002	0,194	0,019	0,071	0,065	0,181	- 0,041	0,511
	(4,000)	(1,590)	(2,343)	(4,498)	(4,160)	(-0,762)	(8,077)
2003	0,209	0,034	0,074	0,074	0,161	- 0,024	0,472
	(4,027)	(2,098)	(2,470)	(4,889)	(3,399)	(-0,513)	(7,483)

The conclusion is that all the process of primary distribution of income has an impact on the inter-regional redistribution of income, except perhaps the property income received (less paid) from the rest of the world (β_{12}), mainly in 2002. The most impressive effect is the retained income by the other kind of institutions but households (β_{11}). That non-distributed income seems to mimic the GDP asymmetries, absorbing them before the income fall into the hands of the households. However, both the taxes on production less subsidies (β_{14}), and the net compensations received from the rest of the world (β_{13}), play a significant role as well in the smoothing of the income inequalities among the households residing in different regions.

On the other hand, in a second stage we decided to decompose as well the smoothing impact of the secondary redistribution of income that transforms the PIH into the GDIH (β_2). These calculations were performed as follows:

$$(7) \qquad GDIH_i = PIH_i + SB_i - SC_i - TIWH_i$$

where SB are the "Social benefits other than social transfers in kind received by households", SC the "Social contributions paid by (or on behalf of) households", and TIWH are the "Current taxes on income and wealth due by the households"[5].
Adopting a similar procedure to the one that led to decomposition of β_1 in (6) we then get:

$$(8) \qquad 1 = \beta_1 + \beta_{21} + \beta_{22} + \beta_{23} + \beta_3 + \beta_4$$

As in the case of the primary income we must have $\beta_2 = \beta_{21} + \beta_{22} + \beta_{23}$, and the remaining βs shall keep the same values reported in Table 1.
Table 3 gives then our results of the decomposition of the cross-sectional variance of the GDPpc, splitting β_2 into β_{21}, β_{22} and β_{23}.

Table 3. Decomposition of the cross-sectional variance of log GDPpc, with β_2 split, 2002 and 2003, Portuguese NUTS III

Anos	β_1	β_{21}	β_{22}	β_{23}	β_3	β_4
2002	0,349	0,104	0,005	0,072	- 0,041	0,511
	(4,563)	(3,439)	(0,363)	(4,731)	(-0,762)	(8,077)
2003	0,391	0,088	0,003	0,070	- 0,024	0,472
	(4,981)	(2,826)	(0,175)	(4,460)	(-0,513)	(7,483)

The outcome of this new decomposition of the cross-sectional variance of the GDPpc (of its log) is that social benefits – represented by β_{21} – and the taxes on income and wealth – β_{23} – actually account for inter-regional income redistribution process. Poorest regions benefit from this process because very likely poorer people – that reside there in a relevant percentage – receive more social benefits and pay fewer taxes. On the opposite, richest regions are losers in that redistribution process, as they benefit less from social transfers and have a higher tax burden. On the contrary, the social contributions due by the households (β_{22}) to the social security systems do not seem to play a similar spatial distributive role.

3. THE SPATIAL STABILIZATION OF INCOME ANALYSIS

The spatial stabilization of income analysis is different from the inter-regional redistribution effect, that we dealt with in the precedent section, because the focus now is on the growth – instead of the levels – of the regional incomes. When in our previous analysis the issue was how the richest (poorest) regions become less rich (poor), when we proceeded from the product (that is also the generation of income) to the income that households actually dispose, our point now is how the positive or negative shocks that hit the product are absorbed by the transformation process of that product in income. Note that through this transformation a shock in one region product spreads into the income of other regions too, that being the reason why this phenomenon is known as well by the "risk sharing" analysis.

In this second part of our paper, the starting point is again the same identity (1) of the previous section that emanates from the Figure 1's scheme. The difference now is that we transform each ratio taking the first differences of the logs (which is the same of their growth rates), leading to the following decomposition of the variance:

$$(9) \quad \mathrm{var}\left(\Delta \log \frac{GDP_i}{Pop_i}\right) = \mathrm{cov}\left\{\Delta \log \frac{GDP_i}{PIH_i}, \Delta \log \frac{GDP_i}{Pop_i}\right\} +$$

[5] As a matter of fact we did not take here into account the Other Current Transfers, received less paid, by the households, that amount to a very small value, because we did not find a proper proxy for regionalizing them at NUTS III level.

$$+\operatorname{cov}\{\Delta\log\frac{\mathrm{PIH_i}}{\mathrm{GDIH_i}},\Delta\log\frac{\mathrm{GDP_i}}{\mathrm{Pop_i}}\}+\operatorname{cov}\{\Delta\log\frac{\mathrm{GDIH_i}}{\mathrm{AGDIH_i}},\Delta\log\frac{\mathrm{GDP_i}}{\mathrm{Pop_i}}\}+$$

$$+\operatorname{cov}\{\Delta\log\frac{\mathrm{AGDIH_i}}{\mathrm{Pop_i}},\Delta\log\frac{\mathrm{GDP_i}}{\mathrm{Pop_i}}\}$$

Dividing again both sides of the equality (9) by the total variance of the growth rate of the GDPpc, adopting a similar procedure to what we made in (2), we get:

$$(10) \qquad 1=\gamma_1+\gamma_2+\gamma_3+\gamma_4$$

where the γs (likewise the βs) are the ratios of each covariance in (9) by the left hand side variance.

When γ_4 is less than 1, revealing a small correlation between the growth rates of the AGDIHpc and of the GDPpc, that means that the shocks on regional GDP are partially absorbed, having only a reduced effect on the AGDIHpc of the regions. It exists therefore risk sharing. 1- γ_4 is to be regarded precisely as the risk sharing degree among the regions of the country. Thus we have, as a consequence of this spatial stabilization of income process, by (10), that γ_1, γ_2 or γ_3 or all together must be positive. These γs are the channels through where the spatial stabilization occurs. When γ_1 is positive we conclude that risk sharing happened at the expense of the primary income distribution to the households. When it is γ_2 that is positive – also or instead – it is the secondary redistribution of income (the taxes/transfers system) that plays that role of stabilizing the income over space. In the case of γ_3, the spatial stabilization process is performed through the social transfers in kind that transforms the GDIH into AGDIH.

Table 4 depicts the results we obtained for our γs.

Table 4. Decomposition of the cross-sectional variance of Δlog GDPpc, 2002 – 2003, Portuguese NUTS III

Anos	γ_1	γ_2	γ_3	γ_4
2002-2003	0,530	0,023	0,031	0,415
	(5,398)	(0,525)	(0,380)	(3,932)

These results confirmed that as we expected the shocks on regional product are softened by a spatial stabilization of income process, as γ_4 is significantly different from 1 (and from 0 as well, granting that anyway GDP matters to some extent for explaining the variability of the AGDIH). The risk sharing degree exceeds even 58%, that meaning that less of 42% of a GDPpc shock in one particular region is kept inside that very region. However, very interestingly, it becomes patent that it is the primary distribution of income (γ_1) alone that grasps that process of income stabilizing over regions. As for γ_2 and γ_3, they assume a very low value non-significantly different from zero, preventing the secondary and the in kind distribution from playing the same stabilizing role.

As we concluded that it is the primary distribution of income that hold the full charge for the risk sharing process among regions, we decided to focus on that effect splitting it in the same mode adopted in section 2. γ_1 is then substituted by the sum $\gamma_{11}+\gamma_{12}+\gamma_{13}+\gamma_{14}$.

The γ_{11} accounts to the stabilizing effect of variations of non-distributed gross operating surpluses, protecting the households income, while γ_{12} refers to the impact of the fluctuations of the property income received/paid from the rest of the world (including other regions and other countries), and γ_{13} measures the equivalent effect of the compensation of the employees also received/paid from the rest of the world. γ_{14} finally is the part of the primary income effect concerning the fluctuations of the taxes on production less subsidies, that are excluded when we come up to look at the income of the households instead of GDP. The meaning (and the value) of the other γs is of course the same than in (10).

Table 5 ahead comprises then the decomposition of the cross-sectional variance of the variation of log GDPpc, among the Portuguese NUTS III regions, with the splitting of the primary income effect into its constituent parts.

Table 5. Decomposition of the cross-sectional variance of Δlog GDPpc, with γ_1 split, 2002 - 2003, Portuguese NUTS III

Anos	γ_{11}	γ_{12}	γ_{13}	γ_{14}	γ_2	γ_3	γ_4
2002-2003	0,435	- 0,010	- 0,004	0,109	0,023	0,031	0,415
	(7,819)	(- 0,158)	(- 0,786)	(5,464)	(0,525)	(0,380)	(3,932)

Our conclusion is then that the main channel of the households income stabilizing effect among regions is the non-distribution to the households of part of the gross operating surpluses (γ_{11}). Following a sharp variation, positive or negative, in a regional GDPpc the major impact is felt at the other sectors (societies, government) incomes, that do not propel these shocks (at least in the short run) to the households sector[6]. The fluctuation of taxes on production, less subsidies (γ_{14}), that resembles the GDPpc variations, is also a relevant mechanism – although a weaker one given its lower value – on that process of spatial stabilization of income. On the contrary, the distributions of property income and labor compensations from/to the rest of the word (γ_{12} and γ_{13}) do not seem to have a significant impact on the GDP shocks absorption.

4. MAIN CONCLUSIONS

The purpose of this paper was to look at the regional effects of the income distribution process that comes after the product generation. In fact, the trivial analyses usually based on the regional GDP concept only account for the generation of the income, neglecting its distribution. On the contrary our focus was on the distribution of income ensuing generation, and on its impact at the regional level. We proceeded to our analysis in two steps: the first approach was a synchronic one and consists in wondering if product asymmetries among regions, in a given year (we looked at 2002 and 2003), that arise when it is generated, still remain relevant when the income is distributed to the households. The ultimate households' income concept that we adopted in this study, after the entire distribution process was over, was the Adjusted Gross Disposable Income of the Households, defined in the frame of the official National Accounts systems. We refer to this first approach as the "inter-regional distribution of income". In our second stage we rather adopted a dynamic approach, raising the issue of how the product variations, or shocks, lead to identical income variations at the AGDIH level, if those shocks are buffered by income distribution, or if on the contrary they may be amplified. We looked in this phase at the 2002-2003 growth rates. This latter approach was named of the "spatial stabilization of income" analysis.

Our main conclusions, referring to the Portuguese NUTS III regions, and to the years of 2002 and 2003, were based on a methodology of decomposing cross sectional variances, adopted from the relevant literature. We then asserted that:

- The regional asymmetries, in Portugal, are much lessened as a result of the inter-regional distribution of income, at the AGDIHpc level, than they were previously when we looked at the GDPpc of the Portuguese regions.
- These minor asymmetries result both from a benign distribution of primary income and also from a fairer secondary redistribution of income. The former effect may have been produced by several sub-channels: the income retention by other institutions (societies, government units and so on), the compensation of employees received, less paid, from other regions and countries, and the burden of the taxes on production less subsidies all seem to have played a

[6] This result that the non-distribution of incomes to the households has an outstanding effect on stabilizing their income along GDP shocks is recurrent all over the relevant literature, including the original classic papers of Asdrubali *et al*. (1996) and Sørensen and Yosha (1998)

significant role on the income distribution over space. There is however less evidence that the property incomes received, less paid, from the rest of the world, played an equivalent role in equalizing the regional incomes. As for the secondary distribution of income, its benign effect on regional income was the result of the taxes on income and wealth and of the social benefits received by households, but the social contributions avoided this flattener role. We did not find evidence as well that the social transfers in kind contributed to an equitable income distribution of income over regions.

- Just as regional asymmetries, product shocks are also absorbed when our focus proceed from the product generation, felt at the GDP level, to the income of the households, gauged by the AGDIH. In other words, distribution of income to the households works as a spatial stabilizer in regard to the disturbances that may hit product.
- This spatial stabilizer effect on households' regional income was performed, however, by the primary income distribution alone, the secondary distribution and the distribution in kind having been pulled out from this process. In the scope of that primary income effect, the effective channels were the compression of the non-distributed income to the households when negative shocks happened, and in a small scale the decrease of the taxes bearing on the production and imports, net of subsidies.

REFERENCES

ANDERSSON, LINDA (2008), "Fiscal Flows and Financial Markets: To What Extent Do They Provide Risk Sharing within Sweden", *Regional Studies*, 42, 7, pp. 1003-1011

ASDRUBALI, P. and KIM, S. (2004), "Dynamic Risk Sharing in the United States and Europe", *Journal of Monetary Economics*, 91, pp. 809-836

ASDRUBALI, P.; SØRENSEN, B. and YOSHA, O. (1996), "Channels of interstate risk sharing: Unites States 1963-1990", *The Quarterly Journal of Economics*, 111, pp. 1081-1110

BEHRENS, AXEL (2003), "How rich are Europe´s regions?", Statistics in Focus, 06/2003, Eurostat

DECRESSIN, J. (1999), "Regional income redistribution and risk sharing: how does Italy compare in Europe?", *IMF Working paper*, WP/99/123

EUROSTAT (1996a), "Sistema Europeu de Contas 1995", Serviço de Publicações Oficiais das CE, Luxemburgo

EUROSTAT (1996b), "Métodos das Contas Regionais, Contas das Famílias", Serviço de Publicações Oficiais das CE, Luxemburgo

EUROSTAT (1996c), "Métodos das Contas Regionais, VAB e FBCF por Ramos de Actividade", Serviço de Publicações Oficiais das CE, Luxemburgo

EUROSTAT (1999), "Métodos das Contas Regionais: quadros das Administrações Públicas", Serviço de Publicações Oficiais das CE, Luxemburgo

JÜBEN, F. (2006), "Interregional risk sharing and fiscal redistribution in unified Germany", *Papers in Regional Science*, 85, 2

KIM, S.; KIM, S. and WANG, Y. (2003), "Financial integration and consumption risk sharing in East Asia", East Asian Bureau of Economic Research, Finance Working Paper nº 303

MÉLITZ, J. and ZUMER, F. (1998), "Regional Redistribution and stabilization by the centre in Canada, France, the United Kingdom and the United States: new estimates based on panel data econometrics", CEPR – Centre for Economic Policy Research, Discussion Paper nº 1829

RAMOS, P. (1996), "Contas Regionais das Administrações Públicas: Breves considerações metodológicas", *Revista de Estatística*, 2º Quadrimestre, pp. 123-135

SØRENSEN, B. and YOSHA, O. (1998), "International risk sharing and European monetary unification", *Journal of International Economics*, 45, pp. 211-238

SØRENSEN, B.; YOsha, O. and KALEMNI-OZCAN, S. (2004), "Asymmetric shocks and risk sharing in a monetary union: updated evidence and policy implications for Europe", CEPR – Centre for Economic Policy Research, Discussion Paper nº 4463

SØRENSEN, B.; YOSHA, O., WU, Y. and ZHU, Y. (2005), "Home bias and international risk sharing: twin puzzles separated at birth", CEPR – Centre for Economic Policy Research, Discussion Paper nº 5113

VON HAGEN, J. and HEPP, RALF (2001), "Regional risk sharing and redistribution in the German federation", CEPR – Centre for Economic Policy Research, Discussion Paper nº 2662.

ESTIMACIÓN DEL IMPACTO ECONÓMICO DE LA ACTIVIDAD DEL CENTRO DE ALMACENAMIENTO DE RESIDUOS DE MEDIA Y BAJA ACTIVIDAD DE "EL CABRIL"

Mª CARMEN LÓPEZ MARTÍN
E-mail: mclopez@etea.com
Departamento de Economía, Ciencias Jurídicas y Sociología
ARACELI DE LOS RÍOS BERJILLOS
E-mail: arios@etea.com
Departamento de Gestión Empresarial y Métodos Cuantitativos
Universidad de Córdoba

Resumen

Este trabajo muestra las principales estimaciones realizadas sobre el impacto económico derivado de la actividad del centro de almacenamiento de residuos de media y baja actividad (RMBA) "El Cabril" en la economía andaluza. El período de estudio comprende desde la implantación (1986) hasta el año 2006. El trabajo se estructura en dos grandes partes: en la primera se expone la metodología empleada, insistiendo en la organización de los datos facilitados por ENRESA y su asignación a las distintas ramas de la tabla input-output. Las tablas Input-Output empleadas son las correspondientes a Andalucía para el año 2000 (últimas disponibles). En la segunda parte se muestran los principales resultados obtenidos, distinguiendo el impacto derivado de la demanda efectuada por la instalación del derivado de los gastos del personal que trabaja en ella. A su vez, también se diferencia entre el período de implantación (1986-1996) y el de consolidación del centro (1997-2006). Finalmente, en las conclusiones se valora el impacto total de la instalación sobre la zona de influencia más próxima y sobre el conjunto de la economía andaluza.

Palabras clave: tablas input-output, impacto directo, impacto indirecto, impacto inducido.
Área temática: Economía Urbana, Regional y Local.

Abstract

This text shows the main estimates of the economic impact on Andalusian economy resulting from activities carried out in the "El Cabril" radioactive-waste storage centre (located in Córdoba, Spain) This study covers over 20 years (from 1986 –when the installation began running- to 2006). The work is divided into two main parts: the first outlines the methodology used, emphasizing the organization of the information given by ENRESA and assigning it to different branches in Input-output tables. The input-output matrix used belongs to the Andalusian tables elaborated in 2000 (the most recent). The second section shows the main results obtained, differentiating between the impact of demands coming from the centre itself and the impact of expenses attributable to the employees who work there. Also, two periods are contemplated: the centre's beginnings (1986-1996) and its consolidation period (1997-2006). Finally, the conclusions evaluate the overall impact of the installation on its closest area of influence and on Andalusian economy as a whole.

Key words: input-output tables, direct impact, indirect impact, induced impact.
Thematic Area: Urban, Regional and Local Economics

1. INTRODUCCIÓN

En este trabajo se analiza el impacto económico de la actividad de centro de almacenamiento de RBMA "El Cabril" en el ámbito regional a partir del modelo Input-Output; esta es la metodología más empleada para estimar el impacto económico de determinadas actividades en un ámbito geográfico concreto, ya que permite cuantificar cuál es la contribución

de dicha actividad a la generación de renta y a la creación de empleo en un área de influencia. El impacto sobre la actividad económica andaluza de "El Cabril" tiene un doble origen, por una parte, la demanda propia del centro de almacenamiento de residuos, tanto corriente como de inversión y, por otra, la demanda de consumo derivada de la renta obtenida por los trabajadores del centro.

Las tablas Input-Output empleadas son las correspondientes a la Comunidad Autónoma de Andalucía para el año 2000, por ser las últimas disponibles. Al respecto hay que tener en cuenta dos cuestiones, por una parte, que dada la dificultad de elaboración de estas tablas, su periodicidad suele ser quinquenal, por tanto, un supuesto que se asume en esta investigación es que la estructura productiva de la región permanece constante e idéntica a la existente en el año 2000, durante todo el período de estudio.

Los resultados se exponen distinguiendo entre el impacto directo, indirecto e inducido de la actividad de "El Cabril". Aunque el análisis Input-Output se refiere fundamentalmente al impacto en la comunidad autónoma de Andalucía, también se ha considerado interesante conocer el impacto directo desagregado según distintos niveles territoriales: los cuatro municipios que reciben compensaciones según la OM de 1991, Peñarroya-Pueblonuevo, resto de la provincia de Córdoba, resto de Andalucía y Azuaga. A los efectos del análisis Input-Output, el municipio pacense se ha considerado perteneciente a Andalucía.

2. METODOLOGÍA

2.1. Las tablas Input-Output y el modelo de demanda

A grandes rasgos, una tabla Input-Output es un instrumento que muestra la interdependencia que existe entre las diferentes actividades productivas de una economía[1]. Esta relación de interdependencia se basa en la idea básica de que cada sector, para poder llevar a cabo su producción, necesita bienes y servicios producidos por el resto de los sectores.

La utilidad de las tablas Input-Output estriba en que la información que suministran constituye una base numérica que permite analizar las relaciones que se producen entre las diferentes actividades productivas. Para ello, no obstante, es necesario trabajar con los denominados coeficientes técnicos, que se calculan dividiendo los consumos intermedios de cada sector entre el total de la producción sectorial; de esta forma los coeficientes muestran la necesidad que tiene cada sector de consumos intermedios procedentes de otros sectores y se calculan matemáticamente como sigue:

$a_{ij} = x_{ij} / X_j$,

siendo:

- x_{ij} la cantidad de producto del sector i que utiliza el sector j en su proceso productivo
- X_j la producción total obtenida por el sector j.
- a_{ij} mide en términos relativos las necesidades del sector j de productos procedentes del sector i.

A partir de la tabla Input-Output se estiman los impactos económicos de cualquier actividad en un ámbito geográfico determinado en función del denominado modelo de demanda que, en definitiva, asume que para satisfacer una determinada demanda D es necesario contar con un volumen de producción X. Teniendo esto en cuenta, es posible determinar la cantidad que debe producir cada uno de los sectores cuando se produce un aumento de la demanda final de todos los sectores considerados conjuntamente, o bien cuando dicho aumento se debe a la demanda final de un único sector. El planteamiento en forma matricial del modelo de demanda es el siguiente:

$A \cdot X + D = X$

Siendo:

[1] Para una exposición más detallada sobre el modelo Input-Output véase, por ejemplo, Muñoz Cidad (1989) o Pulido y Fontela (1993).

- A la matriz de coeficientes técnicos
- X el vector columna que contiene las producciones efectivas sectoriales
- D el vector columna de los elementos correspondientes a la demanda de cada sector.

A partir de la expresión anterior, se puede operar en el modelo de forma que se obtenga el valor de la producción en función únicamente de la demanda (teniendo en cuenta que se está operando con matrices y no con valores absolutos). De esta forma se llegaría al siguiente resultado:

$X = (I - A)^{-1} \cdot D$,

Siendo I la matriz identidad y los restantes conceptos, los indicados antes. La matriz $(I - A)^{-1} = M$ se conoce como inversa de Leontief o matriz de multiplicadores. En definitiva, lo que subyace en este modelo es que cuando la demanda de un sector aumenta, su producción debe aumentar también para poder satisfacer la nueva demanda, lo que conlleva un aumento de la demanda de consumos intermedios del resto de sectores. A su vez, estos sectores deberán producir más y, por lo tanto, consumirán mayores cantidades de inputs intermedios y así sucesivamente. Es decir, en la estimación económica, se están recogiendo los efectos multiplicadores (directos e indirectos) que cualquier variación de demanda trae consigo. En definitiva, esto se hace empleando la ecuación anterior en términos de variaciones de la manera siguiente:

$\Delta X = (1- A)^{-1} \cdot \Delta D = M \cdot \Delta D$.

Todo lo expuesto se perfecciona cuando se considera que el consumo privado no es una variable exógena, sino endógena, es decir, está ligado a las variaciones de la producción ya que ésta implica una retribución al factor trabajo, la cual se destina, en su mayor parte, a la adquisición de bienes y servicios. Para reflejar la relación que existe entre el aumento de la renta productiva y el aumento en el consumo final, se puede ampliar el modelo anterior incorporando el consumo a la matriz de coeficientes técnicos y eliminándolo de la demanda final. De esta manera se incorpora el consumo privado como una variable endógena, llegando a la siguiente relación:

$\Delta \overline{X} = (1- \overline{A})^{-1} \cdot \Delta \overline{D} = \overline{M} \cdot \Delta \overline{D}$.

Siendo:

- $\Delta \overline{X}$ el vector columna que contiene la variación en la producción final sectorial y tiene n+l elementos (n actividades de producción y 1 sector de consumo).
- $\overline{A}$ la matriz de coeficientes técnicos, que incorporaría un nuevo vector columna y un nuevo vector fila, el primero correspondería a los coeficientes de consumo sectorial, calculados a partir del cociente del consumo de cada sector respecto al valor añadido total de la economía y el segundo correspondería a los coeficientes de valor añadido, calculados mediante el cociente del valor añadido sectorial respecto al correspondiente valor de la producción final.
- $\Delta \overline{D}$ el vector columna que refleja los cambios en la demando final y tiene n + 1 elementos (n componentes para los sectores y 1 componente para el consumo).
- $\overline{M}$ la expresión contiene los multiplicadores ampliados asociados a la incorporación endógena del consumo.

La anterior ampliación del modelo Input-Output permite incorporar en la estimación del impacto económico de una actividad, no sólo los efectos directos e indirectos de la demanda sobre la actividad sectorial, sino también el efecto inducido sobre el consumo (una mayor producción exige el empleo de una mayor cantidad de factor trabajo, lo que supondrá aumentos en la renta del consumidor que, a su vez, se traducirán en un mayor consumo, volviendo a aumentar nuevamente la producción sectorial).

2.2. Aplicación del modelo de demanda a la estimación del impacto económico del Centro de almacenamiento de RBMA "El Cabril"

A continuación se recoge una explicación de la metodología de cálculo utilizada para evaluar los efectos del Centro de almacenamiento de RBMA "El Cabril" sobre la actividad

económica en Andalucía, distinguiendo los efectos directos, de los indirectos e inducidos, entendidos como se describen a continuación.

- El efecto directo sobre la producción sectorial, recoge la demanda inicial realizada por el Centro de almacenamiento de RBMA "El Cabril" (como consecuencia de su actividad -tanto inversión como funcionamiento-), la cual es atendida por los diferentes sectores productivos. Por otro lado, las remuneraciones pagadas a los trabajadores de la instalación permitirán que éstos realicen unos determinados gastos de consumo, los cuales constituyen también una mayor demanda que deberá ser atendida por los diferentes sectores productivos.
- El efecto indirecto se debe a que los sectores receptores de la demanda directa necesitan, a su vez, inputs y materias primas del resto de sectores, lo que hará también aumentar la producción de dichos sectores, debido a las interdependencias existentes entre las actividades productivas.
- El efecto inducido. Este efecto tiene en cuenta que como consecuencia de la retribución al factor trabajo aumentará la renta familiar disponible que, a su vez, se destinará, en parte, al consumo que es, en sí mismo una nueva demanda que debe ser satisfecha, generando nuevos aumentos de la producción. El efecto inducido recoge precisamente estos aumentos de producción derivados del incremento de renta y consumo.

Para poder llegar a determinar los efectos que se acaban de indicar es necesario, en primer término, determinar cuál ha sido el importe total del gasto directo originado por la actividad del Centro de almacenamiento de RBMA "El Cabril" en la zona de influencia que se quiere analizar. Para ello, debemos tener en cuenta cuáles son las partidas de demanda asociadas a dicha actividad y hacia qué sectores productivos van dirigidas. Puesto que la evaluación del impacto económico exige contemplar de forma exhaustiva todos los gastos asociados, debemos contemplar tanto la inversión necesaria para poner en marcha la instalación como, posteriormente, los gastos relativos a su funcionamiento, así como la demanda de consumo efectuada por los trabajadores directamente vinculados a ambos procesos[2].

Con relación a las partidas de gasto asociadas tanto a la fase instalación del Centro, como a la de su funcionamiento, y su distribución entre las distintas ramas de actividad, hemos partido de la información proporcionada directamente por ENRESA, para posteriormente efectuar su asignación entre las ramas productivas. La información relativa a estas partidas, que abarcan un período de 19 años, procede de dos fuentes diferentes: para el período 1987-1996, se han tomado los datos correspondientes a un estudio de carácter interno elaborado en el propio Centro de almacenamiento de RBMA "El Cabril"; desde 1997 a 2006 se ha contado con la información contable completa del mismo centro. En el primer caso, las partidas contables ya estaban agrupadas en distintos conceptos, por lo que se directamente se realizó la asignación de estos conceptos a las ramas de la tabla Input-Output. Con respecto a los datos del segundo periodo, puesto que no se disponía de una agrupación previa, fue necesario realizar, en primer lugar, una selección de los gastos e inversiones en función del lugar de residencia del proveedor, para contar únicamente con los de Andalucía y Azuaga, que son las zonas objeto de estudio; a continuación, se procedió a agrupar estos conceptos para poder realizar una asignación a las ramas de la tabla Input-Output utilizada de manera similar a como se había hecho en el período anterior.

Todos los datos facilitados por Enresa estaban valorados a los precios corrientes del ejercicio en el que se produjo el gasto o la inversión[3], por ello fueron deflactados tomando como año base 2000[4].

El consumo efectuado por los trabajadores de la instalación, ha sido estimado mediante un procedimiento indirecto. En primer término, se ha tomado como referencia el importe anual de sueldos y salarios pagados a los trabajadores del Centro de almacenamiento de RBMA “El

[2] La determinación del impacto económico derivado del gasto de los trabajadores se analiza más adelante.

[3] Además, los datos correspondientes al periodo 1987-1996 estaban expresados en millones de pesetas, por lo que, previamente, ha habido que expresarlos en euros.

[4] Se ha tomado este año por ser el de realización de las últimas tablas input-output de Andalucía. La valoración a precios del año 2000 se ha llevado a cabo empleando como referencia el índice de precios al consumo de Andalucía (INE, 2007); dado que la forma de cálculo del ipc ha variado en el periodo del estudio, se han homogeneizado las bases utilizadas.

Cabril" durante todo el período considerado[5]; al no contarse con el detalle de dicho gasto por niveles territoriales, se ha estimado proporcionalmente al número de trabajadores en cada municipio. Al igual que con el resto de gastos, se han eliminando las diferencias temporales de precios.

Como es sabido, no toda la renta de los trabajadores se destina directamente a la demanda de consumo, ya que éstos deben pagar los impuestos que recaen sobre la renta y las cotizaciones sociales, en otras palabras, la renta percibida no coincide con el valor de la renta disponible que, a su vez, es aquella susceptible de ser destinada bien al consumo, bien al ahorro. Para estimar el valor de la renta disponible, se ha aplicado al importe anual de los salarios el correspondiente tipo impositivo[6], es decir, reducimos los salarios en la proporción de renta destinada a pagar impuestos directos y cotizaciones. A continuación, puesto que es posible que parte de la renta disponible del consumidor vaya al ahorro, es necesario conocer la proporción de renta disponible que se materializa en consumo (es decir, la propensión a consumir), para estimar cuál es la demanda de consumo.

El tipo impositivo medio se ha calculado contando con los datos de la comunidad autónoma de Andalucía, ya que no se disponía información suficiente para la provincia de Córdoba. Esta estimación se ha hecho partiendo de los datos de la Contabilidad regional de Andalucía[7] correspondientes al total de impuestos directos más cotizaciones efectivas y al saldo de rentas primarias[8] de las familias; el cociente entre ambos conceptos permite conocer cuál sería el tipo impositivo medio, el cual, deducido de los salarios percibidos por los trabajadores, nos permite estimar su renta disponible.

El valor resultante del cálculo anterior equivale a la renta salarial disponible, de la cual sólo una parte se destina al consumo. La propensión marginal al consumo se ha obtenido dividiendo el importe del consumo total anual de los hogares por el valor total de renta disponible bruta de las familias[9]. Así, aplicando la proporción de consumo resultante al valor de renta disponible estimada, hemos determinado el importe destinado a consumo final por parte de los trabajadores de la instalación. Finalmente, se ha repartido el consumo entre los distintos sectores productivos teniendo en cuenta las proporciones sectoriales de consumo contempladas por el Instituto de Estadística de Andalucía para el año 2000[10]. Una vez que se han determinado los importes correspondientes a la demanda de consumo de los trabajadores, estos datos deben ser introducidos en el modelo de demanda expuesto anteriormente con el objeto de calcular los efectos de su actividad.

A continuación se expone el procedimiento seguido para la estimación de los efectos directos, indirectos e inducidos. En primer lugar, hay que indicar que en todos los casos se ha trabajado con la matriz de 30 ramas de actividad; ya que al analizar las matrices se comprobó que un mayor nivel de desagregación no aportaba más información. También hay que indicar que el Marco Input-Output de Andalucía 2000 no proporciona la Matriz de Leontief para 30 ramas de actividad, por lo que ésta ha tenido que ser calculada en todas las ocasiones necesarias.

Una vez obtenidas las matrices con las que se va a trabajar, el siguiente paso es el cálculo del vector (Δi), que contiene los importes correspondientes a los gastos de inversión y funcionamiento, asignados según ramas productivas. También deben tenerse en cuenta los gastos de consumo de los trabajadores de la instalación asignados a las distintas ramas productivas tal y como se ha indicado anteriormente.

Tal y como se exponía más arriba se considera el consumo como una variable endógena del modelo. A partir de la información recogida en el vector Δi, se estima el impacto económico de

[5] Esta información procede de ENRESA.
[6] Para facilitar los cálculos, se ha considerado conjuntamente el pago del impuesto sobre la renta y las cotizaciones sociales.
[7] Esto información procede de los datos publicados por el INE.
[8] Este saldo comprende no sólo la remuneración a los asalariados, sino también otro tipo de rentas, tales como el excedente de explotación y los intereses, etc.
[9] Los valores del tipo impositivo medio, de los años 2005 y 2006 y de la propensión marginal al consumo, de los años 2004, 2005 y 2006, no se ha podido estimar a partir de los datos de la Contabilidad Regional de España, por lo que han tenido que ser estimados aplicando, al último dato disponible, la tasa media de variación acumulada correspondiente a los tres últimos años.
[10] Marco Input-Output de Andalucía 2000 (Instituto de Estadística de Andalucía, 2000).

la actividad de "El Cabril" medido como el aumento de la producción, en el conjunto de sectores, necesaria para satisfacer una demanda determinada. Dicho impacto se obtiene al multiplicar la Matriz inversa de Leontief por el gasto final tanto corriente como de inversión, y el de consumo (Δi).

$\Delta X = \overline{M}^{*} \cdot \Delta i$

Siendo:

- $\overline{M}^{*}$ la matriz que se obtiene al eliminar la fila y la columna correspondientes al sector de consumo de la matriz $\overline{M}$ y, por tanto, la matriz para los n sectores productivos.

Esta expresión se puede descomponer de la siguiente manera:

$\Delta X = 1\ \Delta i + [M - 1]\ \Delta i + [\overline{M}^{*} - M]\ \Delta i$

La expresión anterior divide el efecto total sobre la producción sectorial ocasionado por la demanda en la suma de tres efectos diferenciados:

Siendo:

- $1\ \Delta i$, el efecto directo sobre la producción, que recoge la demanda inicial recibida por los sectores de producción.
- $[M - 1]\ \Delta i$, el efecto indirecto.
- $[\overline{M}^{*} - M]\ \Delta i$, el efecto inducido sobre la producción sectorial.

3. RESULTADOS

3.1. Efectos directos de la actividad de El Cabril: la demanda

Centrándonos en el primero de los subperiodos mencionados, en la tabla 1 se recoge la asignación de las inversiones y gastos de funcionamiento de "El Cabril" a las ramas productivas contempladas en el marco Input-Output elaborado por el Instituto de Estadística de Andalucía. Como se puede apreciar en dicha tabla, los importes se concentran únicamente en cinco ramas de las 30 que contempla el marco Input-Output, lo cual se debe al menor detalle de los datos proporcionados por Enresa. Dentro de ellos, se aprecia que la mayor parte corresponden a los gastos de inversión, lo cual es lógico ya que durante esos años es cuando se produjo la puesta en marcha de la instalación[11], en concreto, casi el 60% de la demanda efectuada se refiere a inversiones, frente a algo más del 40% que corresponde a gastos de explotación.

Tabla 1. Gastos de funcionamiento e inversiones en Andalucía durante el período 1986-96. Asignación a ramas productivas. (euros constantes del año 2000)

Ramas	Gastos	Inversiones	Total
Construcción de maquinaria y equipo mecánico	3.277.149,64	2.205.871,53	5.483.021,17
Construcción	9.498.020,12	52.534.422,79	62.032.442,90
Transporte, almacenamiento y comunicaciones	3.963.881,02	---	3.963.881,02
Actividades inmobiliarias y de alquiler; servicios empresariales	8.197.473,41	111.259,41	8.308.732,82
Administración pública, defensa y seguridad social obligatoria	17.270.069,37	3.707.611,15	20.977.680,53
Total	**42.206.593,56**	**58.559.164,88**	**100.765.758,44**

Fuente: Elaboración propia.

[11] Recordemos que el proyecto de la instalación se inició en 1986 con un estudio sobre la caracterización del emplazamiento, que fue aprobado a finales de 1989 y que la construcción de las instalaciones comenzó en 1990 y concluyó dos años y medio después. La fecha de inicio de su actividad operativa fue octubre de 1992, tras obtener el correspondiente permiso de explotación.

Con respecto a las inversiones, como es lógico, casi el 90% de ellas se adjudican a la rama de construcción ya que se refieren a la edificación de la instalación. Le siguen en importancia, pero ya muy distantes, las inversiones en la administración pública que se refieren a los pagos por licencias, muy vinculadas a la actividad de construcción. Finalmente, con mucha menos importancia relativa encontramos las inversiones adjudicadas a la rama de construcción de maquinaria y equipo mecánico, que están relacionadas con los equipos necesarios para el funcionamiento de la instalación.

Entre los gastos de explotación, la adjudicación de estos a distintas ramas productivas (véase de nuevo tabla 1) no muestra una concentración tan acusada como la comentada en las inversiones. A pesar de ello, la mayor parte de los gastos están de nuevo estrechamente relacionados con la construcción y puesta en marcha de la instalación, lo que explica el alto porcentaje existente en la rama de administración pública (por el pago de tasas e impuestos) y en la rama de construcción. No llegan al 20% los gastos efectuados en la rama de actividades inmobiliarias y de alquiler y servicios empresariales, que comprende fundamentalmente los pagos por estudios de emplazamiento e ingeniería.

Si nos fijamos ahora en la distribución de los gastos de inversión y funcionamiento en función del ámbito geográfico en el que se desarrollan, recogidos en la tabla 2, se aprecia que la mayor parte de los gastos de funcionamiento se llevan a cabo en los cuatro municipios que reciben compensaciones de acuerdo con la Orden Ministerial, de 20 de diciembre, de 1994. En concreto, el 49,38% de estos gastos se llevan a cabo en ellos durante el conjunto del período 1986-96, frente al 29,36% de las inversiones. Esto parece lógico ya que las inversiones suponen un menor número de operaciones pero de mayor importe individual, por lo que probablemente han sido realizadas, al menos parcialmente, por empresas de mayor envergadura que las existentes en estas localidades; por el contrario, los gastos de funcionamiento tienen un importe medio más reducido y habrán sido ejecutados en mayor medida en estas poblaciones.

Tabla 2. Total gastos de funcionamiento e inversiones en Andalucía durante el período 1986-96. Asignación por localidades y áreas de influencia. (Euros constantes del año 2000 y % sobre el total)

Municipios/ámbito territorial	Total Gastos de funcionamiento		Total Inversiones	
	Euros	% sobre total	Euros	% sobre total
HORNACHUELOS	10.298.296,90	22,72%	10.969.957,33	18,72%
FUENTEOBEJUNA	8.959.956,84	19,76%	6.198.023,20	10,58%
ALANIS	1.203.299,94	2,65%	0,00	0,00%
LAS NAVAS DE LA CONCEPCION	1.924.947,46	4,25%	33.017,32	0,06%
Subtotal Orden Ministerial	**22.386.501,14**	**49,38%**	**17.200.997,84**	**29,36%**
AZUAGA	1.734.509,80	3,83%	6.419.717,30	10,96%
PEÑARROYA	4.391.670,23	9,69%	1.618.045,99	2,76%
Subtotal no Orden Ministerial	**6.126.180,03**	**13,51%**	**8.037.763,29**	**13,72%**
TOTAL AREA DE INFLUENCIA	**28.512.681,18**	**62,90%**	**25.238.761,13**	**43,07%**
RESTO CORDOBA	7.317.778,08	16,14%	12.210.561,01	20,84%
Subtotal área de influencia y Córdoba	**35.830.459,25**	**79,04%**	**37.449.322,15**	**63,91%**
Subtotal provincia de Córdoba	**30.967.702,05**	**68,31%**	**30.996.587,53**	**52,90%**
RESTO ANDALUCÍA	9.502.035,81	20,96%	21.143.842,00	36,09%

Fuente: Elaboración propia.

Además, la tabla anterior muestra que durante estos primeros años considerados, la demanda efectuada en la provincia de Córdoba concentra la mayor parte de los efectos directos, ya que más de la mitad de las inversiones (concretamente el 52,9%) y más de dos tercios de los gastos de funcionamiento (68,37%) se efectúan en este ámbito geográfico, lo que evidentemente, tendrá repercusiones también importantes al evaluar el impacto que la actividad del Centro de almacenamiento El Cabril tiene, tanto de forma indirecta, como el efecto inducido a través del consumo.

Respecto a los resultados obtenidos en el segundo de los períodos considerados (1997-2006), puesto que el nivel de desagregación de los datos eran mucho mayor que el del periodo

anterior, la asignación de las distintas partidas de demanda a las ramas contempladas en las tablas Input-Output ha podido realizarse con una mayor exactitud. Ésta es la causa que explica que el efecto directo de los gastos de funcionamiento efectuados por El Cabril no aparezca tan concentrado como en el primer subperiodo analizado, tal y como se muestra en la tabla 3[12]. En esta tabla se han agrupado las partidas de demanda del centro en dos categorías: por un lado las relativas a inversiones (inmovilizado y existencias) y por otro lado, las correspondientes a gastos de funcionamiento, siguiendo un criterio similar al utilizado en el subperiodo anterior.

Tabla 3. Gastos de funcionamiento e inversiones en Andalucía durante el período 1997-2006. Asignación a ramas productivas. (Euros constantes del año 2000)

Ramas	Inversiones	Gastos	Total
Extracción de otros minerales, excepto productos energéticos	0,00	2.397.469,74	2.397.469,74
Industrias de alimentación, bebidas y tabaco	0,00	101.550,72	101.550,72
Industria química	0,00	48.185,38	48.185,38
Industrias de otros productos minerales no metálicos	0,00	703.150,34	703.150,34
Metalurgia y fabricación de productos metálicos	0,00	920.380,17	920.380,17
Construcción de maquinaria y equipo mecánico	496.293,66	759.488,06	1.255.781,72
Industria de material y equipo eléctrico, electrónico y óptico	1.657.208,06	0,00	1.657.208,06
Industrias manufactureras diversas	434.703,48	12.167,30	446.870,77
Producción y distribución de energía eléctrica, gas y agua	0,00	13.570,81	13.570,81
Construcción	6.519.695,55	4.878.552,98	11.398.248,53
Comercio y reparaciones	242.553,90	1.651.689,88	1.894.243,78
Hostelería	0,00	1.963.348,40	1.963.348,40
Transporte, almacenamiento y comunicaciones	0,00	3.458.565,86	3.458.565,86
Actividades inmobiliarias y de alquiler; servicios empresariales	843.349,06	1.873.228,61	2.716.577,67
Administración pública, defensa y seguridad social obligatoria	1.389.341,57	26.226.092,39	27.615.433,96
Educación	0,00	74.077,76	74.077,76
Actividades sanitarias y veterinarias; servicios sociales	0,00	3.523.955,02	3.523.955,02
Otras actividades sociales y servicios prestados a la comunidad; servicios personales	0,00	197.270,21	197.270,21
Total	**11.583.145,27**	**48.802.743,61**	**60.385.888,88**

Fuente: Elaboración propia.

Al contrario de lo que sucedía en el período 1986-1996, en los años comprendidos entre 1997 y 2006, el valor de los gastos de funcionamiento en Andalucía es muy superior al de las inversiones, representando estas últimas aproximadamente una cuarta parte de aquellos. Esta división entre inversiones y gastos de funcionamiento es lógica pues estos son los años en los que el centro ya ha terminado casi totalmente el periodo de puesta en marcha y se encuentra operando con normalidad[13]. Tan sólo hay algunos "picos" que se producen en determinados momentos en los que se ha producido alguna ampliación de las instalaciones.

En cuanto a la distribución de la demanda por ramas de actividad, en la tabla 3 se observa que, al igual que en el periodo anterior, las inversiones se encuentran concentradas en pocas actividades, siendo la más importante la de construcción con el 56,29% del total, que corresponde a la ampliación de la instalación en el trienio 2002-2004. Vinculadas a esta ampliación aparecen las inversiones en material y equipo eléctrico (14,31% del total de las

[12] En el caso de las inversiones, al ser mucho más reducido el número de datos facilitados, las partidas se han asignado a las mismas ramas que las contempladas en el periodo 1986-1996.

[13] En concreto, del total de la demanda efectuada en la región, el 19,18% corresponde a inversiones y existencias y el 80,82% son gastos de funcionamiento.

inversiones), en la administración pública (casi el 12% del total, que corresponden a los pagos vinculados a la construcción) y, ya con menor importancia, las actividades inmobiliarias y servicios empresariales (7,28%). En general, las inversiones se han producido en momentos concretos y guardan una estrecha relación con la ampliación de la instalación, ya indicada.

En relación con los gastos de funcionamiento, ya se ha indicado que constituyen el grueso de la demanda efectuada por El Cabril a lo largo de este periodo, agrupando algo más de 8 de cada 10 euros de gasto total en la región. Como es lógico, al corresponder estos años a la explotación de las instalaciones, la diversificación de estos gastos entre diferentes ramas de actividad es mayor que la que existía en los años comprendidos entre 1986 y 1996[14]; los gastos aparecen asignados a 17 de las 30 ramas de actividad contempladas en el análisis. No obstante, la mayor parte de ellos se encuadran en sector público (rama 26. Administración pública), que supone más de la mitad del total (53,74%). Este elevado porcentaje es debido a que una gran parte de estos gastos corresponden a los pagos realizados a los ayuntamientos de la zona por conceptos tales como impuestos, compensaciones y asignaciones de la Fundación Enresa[15].

Por detrás del sector público, los otros gastos que tienen una importancia relativa más elevada (aunque se sitúan en valores bastante más reducidos que los indicados para dicho sector), son: Construcción (10% del total de gastos de funcionamiento en el período), Actividades sanitarias y veterinarias; servicios sociales (que agrupa al 7,22% de dicho total) y Transporte, almacenamiento y comunicaciones (con el 7,09%). Como se deduce de estos datos, entre las cuatro ramas mencionadas hasta ahora se agrupan aproximadamente las tres cuartas partes de los gastos de funcionamiento en la comunidad andaluza.

El resto de las ramas de actividad a las que se han asignado los gastos de funcionamiento presentan una importancia relativa inferior al 5% del total. Destacan, en este grupo: Extracción de otros minerales, excepto productos energéticos (que suponen el 4,91% de los gastos de funcionamiento de este período), Hostelería (que representa el 4,02%), Actividades inmobiliarias y de alquiler; servicios empresariales (con el 3,84%) y, finalmente, Comercio y reparaciones (3,38% sobre el total).

En cuanto a la evolución experimentada a lo largo del periodo por los gastos de funcionamiento, es posible afirmar, en términos generales, que presentan un comportamiento más estable que el de la inversión, lo cual es lógico dada su naturaleza. No obstante, se pueden señalar algunos datos más destacados. Así, la rama Construcción presenta un valor claramente más elevado en el año 2002 que se debe a los gastos realizados que están asociados a la mejora de carreteras, además en 2004 vuelve a observarse un nuevo dato elevado que se corresponde con las mayores inversiones realizadas en ese ejercicio.

La rama a la que corresponden los mayores importes (administración pública), presenta una tendencia decreciente, aunque este descenso no se produce de forma continuada a lo largo del periodo, sino que se producen disminuciones seguidas de varios años de estabilidad, para luego volver a disminuir. Estos cambios se deben a las variaciones experimentadas por las partidas correspondientes a la Fundación y a las compensaciones (los impuestos locales pagados no presentan grandes oscilaciones).

Centrándonos en la distribución de los gastos de inversión y funcionamiento en función del ámbito geográfico en el que se desarrollan, recogidos en la tabla 4, podemos apreciar que las inversiones son inferiores a las realizadas en el periodo 1986-96 y, además, que están mucho más concentradas en determinados lugares que los gastos de funcionamiento. Así, de entre los cuatro municipios que reciben compensaciones de acuerdo con la Orden Ministerial de 1994, sólo son realizadas por empresas de Hornachuelos; además, al ampliar el ámbito territorial hasta la zona de influencia, encontramos que sólo las llevan a cabo en Peñarroya. De esta forma, entre los dos municipios citados agrupan aproximadamente el 42% de la demanda de inversión y, al ampliar hasta considerar el total de la provincia de Córdoba, se llega a que éstas suponen algo más de la mitad del total de las realizadas en el periodo.

[14] Además, debe tenerse en cuenta que para estos años se ha contado con una información mucho más detallada que la que existía para el periodo anterior, lo que nos ha permitido un mayor grado de precisión a la hora de adjudicar los gastos a las ramas de la tabla Input-Output.

[15] Aunque la información proporcionada por Enresa diferenciaba estos conceptos, a la hora de estimar el impacto tanto regional como provincial, se ha optado por agrupar todos estos gastos e incluirlos en el mencionado epígrafe.

Tabla 4. Total gastos de funcionamiento e inversiones en Andalucía durante el período 1997-06. Asignación por localidades y áreas de influencia. (Euros constantes del año 2000 y % sobre el total)

Municipios/ámbito territorial	Total Gastos de funcionamiento		Total Inversiones	
	Euros	% sobre total	Euros	% sobre total
HORNACHUELOS	18.040.925,79	36,94%	2.651.513,50	22,77%
FUENTEOBEJUNA	6.578.307,03	13,47%	0,00	0,00%
LAS NAVAS DE LA CONCEPCION	4.620.097,74	9,46%	0,00	0,00%
ALANIS	277.163,21	0,57%	0,00	0,00%
Subtotal Orden Ministerial	**29.516.493,77**	**60,44%**	**2.651.513,50**	**22,77%**
AZUAGA	474.155,32	0,97%	0,00	0,00%
PEÑARROYA	4.166.075,67	8,53%	2.248.586,00	19,31%
Subtotal no Orden Ministerial	**4.640.231,00**	**9,50%**	**2.248.586,00**	**19,31%**
TOTAL AREA DE INFLUENCIA	**34.156.724,77**	**69,94%**	**4.900.099,50**	**42,08%**
RESTO CORDOBA	10.316.897,75	21,13%	1.638.619,86	14,07%
Subtotal área de influencia y Córdoba	**44.473.622,52**	**91,07%**	**6.538.719,36**	**56,15%**
Subtotal provincia de Córdoba	**39.102.206,25**	**80,07%**	**6.538.719,36**	**56,15%**
RESTO ANDALUCÍA	3.866.480,08	7,92%	5.105.890,72	43,85%
BAJADOZ	493.817,68	1,01%	0,00	0,00%
Subtotal Resto Andalucía y Badajoz	4.360.297,77	8,93%	5.105.890,72	43,85%

Fuente: Elaboración propia.

Por su parte, los gastos del este período 1997-2006 son superiores a los del período anterior y, además, como ya se ha comentado, están mucho más repartidos entre las localidades consideradas. No obstante, a pesar de ello, los gastos realizados en empresas correspondientes a los cuatro municipios incluidos en la Orden ministerial de 1994 suponen casi el 70% de los realizados, lo que implica que la actividad del El Cabril tiene un impacto mayor en la comarca que el que se producía en el periodo anterior. Algo similar cabría indicar con relación a los datos correspondientes a la provincia de Córdoba, que suponen aproximadamente el 80% del total (para los años 1986-96, en este ámbito no llegaban a representar el 70%).

3.2. Los efectos indirectos e inducidos de la demanda de "El Cabril"

En la siguiente tabla 5 se descompone el efecto total distinguiendo entre el efecto directo, indirecto e inducido derivado de la demanda total de El Cabril. Atendiendo al efecto total y su distribución entre el efecto directo, indirecto e inducido, hay que indicar que el primero representa un 35,96% del efecto total, el indirecto un 20,35% y el inducido un 43,68% del efecto total.

El efecto indirecto es menor que el inducido, lo que se debe a varios factores; por una parte, el menor peso de la industria en Andalucía; recuérdese que el efecto indirecto viene dado por el aumento en la producción de otros sectores productivos para atender a la demanda del El Cabril. Por otra parte, algunos de estos proveedores pueden ser no residentes en Andalucía, incluso del exterior, por lo que el aumento de producción de estos proveedores no queda reflejado en las estimaciones efectuadas. Por el contrario, el efecto inducido es el derivado del gasto en consumo; el suministro de los bienes de consumo doméstico es, probablemente, atendido por un mayor número de empresas de Andalucía, lo que explicaría que el efecto inducido sobre la economía andaluza sea mayor que el indirecto. A esta última razón hay que añadir, que en el efecto multiplicador del impacto inducido es mayor ya que contempla el efecto derivado del consumo de los trabajadores de las distintas empresas que, a su vez, ven aumentar su demanda como consecuencia del gasto realizado por El Cabril.

Tabla 5. Efectos directo, indirecto e inducido derivados del gasto total de El Cabril. (Euros constantes del año 2000)

Ramas	Directo	Indirecto	Inducido	Total
Agricultura, ganadería, caza y silvicultura	0	887.567	7.840.079	8.727.647
Pesca	0	59.130	598.298	657.428
Extracción de productos energéticos	0	33.862	66.656	100.519
Extracción de otros minerales, excepto productos energéticos	2.397.470	2.029.603	235.746	4.662.818
Industrias de la alimentación, bebidas y tabaco	101.551	745.386	11.538.634	12.385.571
Industria textil y de la confección	0	105.241	1.516.485	1.621.726
Industria del cuero y del calzado	0	2.499	117.360	119.859
Industria de la madera y del corcho (excepto muebles)	0	1.322.881	577.283	1.900.163
Industria del papel; edición, artes gráficas y reproducción de soportes grabados	0	701.469	2.028.581	2.730.049
Refino de petróleo	0	4.484.422	5.568.274	10.052.696
Industria química	48.185	1.355.813	1.610.622	3.014.620
Industria de la transformación del caucho y materias plásticas	0	571.320	422.699	994.019
Industrias de otros productos minerales no metálicos	703.150	9.308.445	995.662	11.007.257
Metalurgia y fabricación de productos metálicos	920.380	6.256.444	1.109.297	8.286.121
Construcción de maquinaria y equipo mecánico	6.738.803	1.504.874	572.944	8.816.620
Industria de material y equipo eléctrico, electrónico y óptico	1.657.208	1.056.815	254.050	2.968.073
Fabricación de material de transporte	0	510.158	817.467	1.327.625
Industrias manufactureras diversas	446.871	181.203	2.254.002	2.882.076
Producción y distribución de energía eléctrica, gas y agua	13.571	2.863.241	5.863.599	8.740.411
Construcción	72.234.728	25.566.877	6.628.676	104.430.282
Comercio, reparación de vehículos de motor, motocicletas y ciclomotores y artículos personales y de uso doméstico	1.894.244	2.956.665	35.656.050	40.506.959
Hostelería	1.963.348	1.416.294	27.719.750	31.099.392
Transporte, almacenamiento y comunicaciones	7.422.447	11.013.546	16.956.907	35.392.900
Intermediación financiera	0	3.816.098	9.134.183	12.950.281
Actividades inmobiliarias y de alquiler; servicios empresariales	11.025.310	11.191.446	34.200.533	56.417.289
Administración pública, defensa y seguridad social obligatoria	48.593.114	0	0	48.593.114
Educación	74.078	349.704	3.485.231	3.909.013
Actividades sanitarias y veterinarias; servicios sociales	3.523.955	1.088.005	4.680.846	9.292.806
Otras actividades sociales y servicios prestados a la comunidad; servicios personales	197.270	857.050	10.691.886	11.746.206
Hogares que emplean personal doméstico	0	0	2.455.791	2.455.791
Total	**161.044.588**	**91.147.153**	**195.597.592**	**447.789.333**

Fuente: Elaboración propia.

El efecto directo sobre determinadas ramas es nulo, pero cuanto se analiza el efecto indirecto e inducido en todas las ramas se genera un impacto, excepto en la administración pública y en el caso de efecto indirecto tampoco hay impacto en la rama de hogares que emplean personal doméstico. Atendiendo a las ramas de la economía cabe indicar que el 75% del efecto directo se concentra en dos grandes sectores, la construcción (con un 44,85% del efecto directo total) y el sector público (con un 30,17% del total del efecto directo), seguidos de

otros sectores menos importantes en términos relativos como son: actividades inmobiliarias y servicios empresariales (6,85% del total del efecto directo), transporte, almacenamiento y comunicaciones (4,61%) y construcción de maquinaria y equipo mecánico (4,18%).

Respecto al efecto indirecto, como se ha indicado, se distribuye entre todos los sectores, excepto en el sector público y la rama de hogares que emplean personal doméstico. La construcción es la rama en la que el impacto indirecto es mayor (un 28,05% del impacto indirecto total). Un 12,28% del impacto indirecto total repercute sobre el sector actividades inmobiliarias y servicios empresariales, seguido de la rama que engloba a las actividades de transporte, almacenamiento y comunicaciones (12,08% del efecto indirecto total), de industrias de otros productos minerales no metálicos (10,21%) y de la metalurgia y fabricación de productos metálicos (6,86%).

Al analizar las ramas en las que se concentra el efecto inducido por el consumo, se observa que el principal sector es el correspondiente al comercio (18,23% del efecto inducido total), seguido de las actividades inmobiliarias y servicios empresariales (17,49% del total del efecto inducido), la hostelería (14,17%), las actividades de transporte, almacenamiento y comunicaciones (8,67%) y la industria de la alimentación, bebida y tabaco (5,90%). El efecto inducido sobre la construcción es sólo del 3,39% del total de dicho efecto.

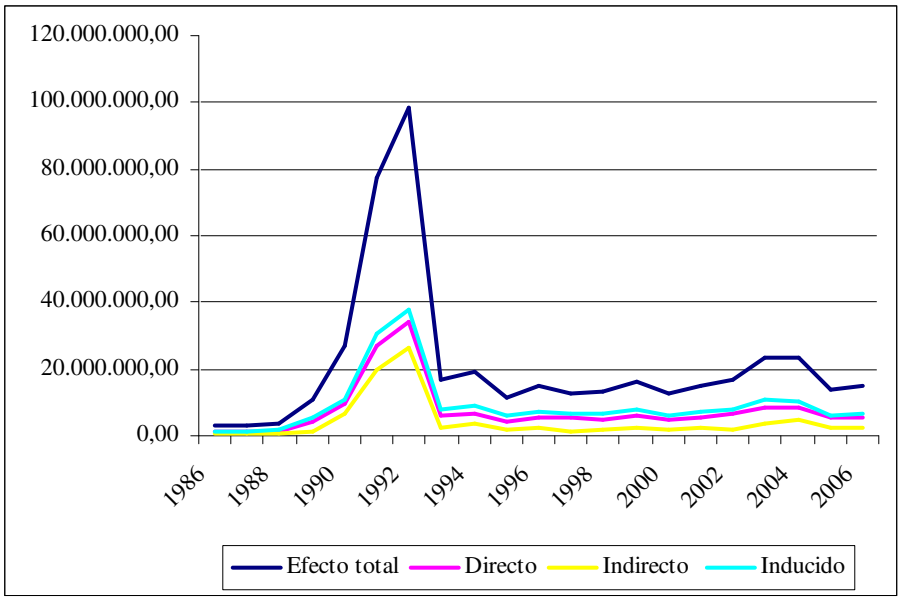
Al comparar las ramas de actividad sobre las que recaen los efectos indirectos e inducidos, observamos que los primeros están más vinculados a sectores productivos y, en concreto, relativos al consumo de capital o bienes de inversión) mientras que los segundos están más relacionados con sectores vinculados al consumo de los hogares, es decir, de bienes de consumo.

Atendiendo a la distribución del efecto indirecto por períodos hay que destacar que el 72,8% del efecto indirecto total se produjo en el período 1986-1996, coincidiendo con el período de implantación, en el que las inversiones fueron, lógicamente, mayores. El 27,2% restante se produce entre 1997 y 2006. Durante 1986-1996, el efecto indirecto medio anual fue de 6.035.436 euros, mientras que entre 1997 y 2006 el efecto indirecto medio anual fue de 2.475.736 euros.

La distribución del efecto inducido por períodos mantiene una proporción diferente aunque similar a la del efecto indirecto, el 60,82% del efecto inducido total se produjo en el período 1986-1996, mientras que el 39,18% restante se produce entre 1997 y 2006. Durante 1986-1996, el efecto inducido medio anual fue de 10.815.402 euros, mientras que entre 1997 y 2006 el efecto inducido medio anual fue de 7.662.817 euros.

A modo de resumen el gráfico 1 muestra la evolución anual del impacto total, directo, indirecto e inducido del gasto corriente y de inversión del El Cabril.

Gráfico 1. Impacto regional de la demanda total de El Cabril (euros constantes del año 2000).



Fuente: Elaboración propia

3.3. Los efectos directos del consumo de los trabajadores

Si analizamos sólo en el efecto que tiene el consumo de los trabajadores de "El Cabril" en el ámbito regional, el primer paso que ha de realizarse, como ya ha indicado en la metodología, es justamente estimar el importe de dicho consumo a lo largo de los años considerados, siguiendo las especificaciones ya indicadas. El resultado de esta estimación es el que se muestra en la tabla 6, que en total asciende a 31.392.257,17 euros constantes del año 2000.

Tabla 6. Estimación del consumo de los trabajadores de El Cabril en Andalucía durante el período 1986-2006. (Euros constantes del año 2000)

	Gastos personal Total (€ corrientes)	Tipo impositivo estimado	Renta disponible estimada	Propensión al consumo	Consumo estimado (€ corrientes)	Consumo estimado (€ constantes)
1986	972.425,62	20,40%	774.010,32	92,08%	712.694,47	1.310.369,85
1987	1.112.957,26	22,88%	858.366,40	96,52%	828.475,66	1.447.335,30
1988	1.207.822,31	22,77%	932.797,20	96,09%	896.356,12	1.501.132,98
1989	1.259.911,10	24,22%	954.793,09	96,24%	918.893,81	1.448.339,48
1990	1.592.680,59	24,54%	1.201.881,82	94,66%	1.137.710,14	1.674.960,59
1991	1.564.896,87	25,36%	1.167.981,49	93,76%	1.095.137,43	1.524.858,41
1992	2.207.454,65	26,87%	1.614.255,98	93,75%	1.513.443,81	1.992.373,10
1993	2.245.834,92	26,52%	1.650.130,48	91,77%	1.514.323,97	1.892.527,59
1994	2.494.103,06	26,88%	1.823.573,17	92,88%	1.693.792,93	2.014.030,60
1995	2.711.333,88	24,98%	2.034.151,52	90,43%	1.839.467,51	2.069.414,70
1996	2.820.616,03	24,37%	2.133.241,78	88,72%	1.892.668,19	2.056.134,35
1997	2.964.839,71	24,38%	2.242.157,33	89,57%	2.008.224,12	2.151.645,35
1998	3.039.482,24	24,37%	2.298.713,12	90,81%	2.087.542,34	2.205.609,33
1999	3.109.527,75	24,44%	2.349.649,05	93,24%	2.190.885,89	2.268.926,84
2000	3.111.227,24	24,53%	2.348.054,00	94,51%	2.219.171,78	2.219.171,78
2001	3.009.236,57	25,04%	2.255.663,97	95,10%	2.145.202,43	2.061.472,01
2002	2.896.602,95	25,35%	2.162.328,09	94,83%	2.050.439,90	1.906.686,74
2003	2.908.393,22	25,25%	2.174.074,81	94,66%	2.057.986,91	1.860.959,85
2004	2.760.037,86	25,38%	2.059.647,89	94,71%	1.950.700,89	1.716.254,81
2005	2.882.585,78	25,49%	2.147.689,10	94,76%	2.035.160,63	1.735.814,92
2006	3.086.459,65	25,61%	2.296.045,70	94,81%	2.176.894,53	1.789.440,16

Fuente: Elaboración propia.

Una vez determinado el importe del consumo anual de los trabajadores, el siguiente paso es repartir este consumo entre los distintos sectores productivos teniendo en cuenta las proporciones sectoriales de consumo contempladas por el Instituto de Estadística de Andalucía. Efectuado este reparto que contempla las mismas ramas de actividad que las empleadas en el análisis Input-Output, se obtiene el consumo efectuado en la región, el cual se muestra en la tabla 7[16].

[16] El total del consumo que aparece en la tabla 1 no coincide con el total del consumo estimado que se recogía en la tabla 2 anterior debido a que no todo el consumo de los trabajadores se realiza dentro de la región y para estimar el impacto en el ámbito regional debe prescindirse de los gastos de consumo efectuados fuera de dicho ámbito.

Tabla 7. Consumo en Andalucía de los trabajadores de "El Cabril" por ramas de actividad durante el período 1986-2006. Datos en euros constantes del año 2000.

Ramas	Consumo (euros constantes)	Consumo (%)
Agricultura, ganadería, caza y silvicultura	565.235,27	1,80%
Pesca	180.800,37	0,58%
Extracción de productos energéticos	893,45	0,00%
Extracción de otros minerales, excepto productos energéticos	4.854,70	0,02%
Industrias de la alimentación, bebidas y tabaco	2.530.361,32	8,06%
Industria textil y de la confección	710.401,50	2,26%
Industria del cuero y del calzado	215.606,03	0,69%
Industria de la madera y del corcho (excepto muebles)	24.020,49	0,08%
Industria del papel; edición, artes gráficas y reproducción de soportes grabados	309.999,09	0,99%
Refino de petróleo	599.750,34	1,91%
Industria química	518.643,30	1,65%
Industria de la transformación del caucho y materias plásticas	33.937,28	0,11%
Industrias de otros productos minerales no metálicos	31.353,81	0,10%
Metalurgia y fabricación de productos metálicos	48.825,02	0,16%
Construcción de maquinaria y equipo mecánico	169.571,89	0,54%
Industria de material y equipo eléctrico, electrónico y óptico	335.268,58	1,07%
Fabricación de material de transporte	1.006.337,92	3,21%
Industrias manufactureras diversas	574.140,68	1,83%
Producción y distribución de energía eléctrica, gas y agua	601.782,81	1,92%
Construcción	279.686,87	0,89%
Comercio, reparación de vehículos de motor, motocicletas y ciclomotores y artículos personales y de uso doméstico	6.118.866,77	19,49%
Hostelería	5.329.195,27	16,98%
Transporte, almacenamiento y comunicaciones	1.871.921,59	5,96%
Intermediación financiera	1.082.884,85	3,45%
Actividades inmobiliarias y de alquiler; servicios empresariales	4.444.671,46	14,16%
Administración pública, defensa y seguridad social obligatoria	0,00	0,00%
Educación	645.834,05	2,06%
Actividades sanitarias y veterinarias; servicios sociales	822.223,60	2,62%
Otras actividades sociales y servicios prestados a la comunidad; servicios personales	1.841.853,24	5,87%
Hogares que emplean personal doméstico	493.335,64	1,57%
Total consumo interior	**31.392.257,17**	**100%**

Fuente: Elaboración propia.

Como se desprende de los datos recogidos en este cuadro, las ramas que reciben los mayores efectos son, en orden decreciente: Comercio (19,49%), Hostelería (16,98%) y Actividades inmobiliarias, alquiler y servicios empresariales (14,16%), que entre las tres suponen poco más de la mitad de los gastos de consumo. Ya con mucha menos importancia (inferior al 10% pero por encima del 5%) se encuentran la industria agroalimentaria (8,06%), transporte (5,96%) y otras actividades sociales (5,87%). Considerando estas tres últimas junto con las anteriores, se alcanza más del 70% del consumo total; el resto de las actividades tienen una importancia relativa mucho más reducida (inferior al 4%).

Siguiendo la misma metodología descrita para la estimación de los efectos sobre el consumo de los sueldos y salarios a nivel regional, es posible determinar los efectos sobre el consumo de los sueldos y salarios en el ámbito provincial y en la zona de influencia de la instalación. Para

ello, se ha tomado de nuevo la información suministrada por Enresa relativa a los sueldos y salarios pagados a sus trabajadores y se ha repartido entre las distintas poblaciones consideradas en función del número de trabajadores residentes en ellas[17]. Los resultados se muestran en la tabla 8, en la que se puede apreciar que la proporción de los gastos de personal en la zona de influencia (en la que se incluyen, además de los cuatro municipios de la Orden Ministerial, Azuaya y Peñarroya-Pueblonuevo) son superiores a los de la provincia de Córdoba debido a la importante presencia de trabajadores de la instalación en Azuaga. Por otra parte, los gastos de personal en la provincia de Córdoba (y, por lo tanto, el consumo realizado por los empleados de la instalación) representan algo más del 86% del total de estos gastos, por lo que el impacto que este concepto tiene sobre la economía andaluza se concentra en su gran mayoría en la provincia cordobesa.

Tabla 8. Estimación del gasto de personal y del consumo de los trabajadores del Centro de almacenamiento de RBMA "El Cabril" según ámbitos territoriales durante el período 1986-2006. Importes en euros constantes del año 2000.

	Total		Provincia		Zona de influencia		Municipios Orden Ministerial	
	Gastos personal	Consumo estimado	Gastos personal	Consumo estimado	Gastos personal	Consumo estimado	Gastos personal	Consumo estimado
1986	972.426	712.694	972.426	712.694	928.224	680.299	77.352	56.692
1987	1.112.957	828.476	1.091.958	812.844	1.070.959	797.212	199.492	148.500
1988	1.207.822	896.356	1.147.431	851.538	1.167.562	866.478	231.499	171.802
1989	1.259.911	918.894	1.174.654	856.713	1.222.019	891.258	284.190	207.269
1990	1.592.681	1.137.710	1.504.198	1.074.504	1.468.805	1.049.222	336.233	240.183
1991	1.564.897	1.095.137	1.338.949	937.016	1.523.055	1.065.856	368.211	257.679
1992	2.207.455	1.513.444	1.938.446	1.329.010	2.025.478	1.388.680	419.337	287.500
1993	2.245.835	1.514.324	2.073.078	1.397.838	2.073.078	1.397.838	540.803	364.653
1994	2.494.103	1.693.793	2.186.807	1.485.102	2.215.392	1.504.515	686.057	465.914
1995	2.711.334	1.839.468	2.407.069	1.633.043	2.400.308	1.628.456	818.133	555.051
1996	2.820.616	1.892.668	2.500.685	1.677.991	2.500.685	1.677.991	829.209	556.409
1997	2.964.840	2.008.224	2.624.612	1.777.772	2.576.008	1.744.850	1.044.984	707.817
1998	3.039.482	2.087.542	2.693.525	1.849.936	2.644.102	1.815.992	1.062.583	729.791
1999	3.109.528	2.190.886	2.808.606	1.978.865	2.607.991	1.837.517	1.053.227	742.074
2000	3.111.227	2.219.172	2.416.986	1.723.985	2.545.550	1.815.686	1.079.930	770.291
2001	3.009.237	2.145.202	2.340.517	1.668.491	2.469.117	1.760.166	1.028.799	733.403
2002	2.896.603	2.050.440	2.241.719	1.586.862	2.367.658	1.676.012	982.326	695.367
2003	2.908.393	2.057.987	2.250.843	1.592.703	2.326.715	1.646.390	1.036.905	733.717
2004	2.760.038	1.950.701	2.141.409	1.513.475	2.212.789	1.563.924	975.531	689.472
2005	2.882.586	2.035.161	2.261.339	1.596.548	2.335.888	1.649.182	1.043.695	736.869
2006	3.086.460	2.176.895	2.380.236	1.678.792	2.458.705	1.734.136	1.098.570	774.827
Total	**29.768.393**	**34.965.173**	**42.495.494**	**29.735.721**	**43.140.090**	**30.191.659**	**15.197.068**	**10.625.281**

Fuente: Elaboración propia

[17] Para realizar la asignación de forma más correcta, debería haberse tenido la información individual relativa a los sueldos percibidos por los trabajadores junto con su lugar de residencia. Dado que no se disponía de este detalle, se ha optado por repartir el total de estos gastos según el número de trabajadores residentes en cada municipio.

Tabla 8. Estimación del gasto de personal y del consumo de los trabajadores del Centro de almacenamiento de RBMA "El Cabril" según ámbitos territoriales durante el período 1986-2006. Importes en euros constantes del año 2000.

	Total		Provincia		Zona de influencia		Municipios Orden Ministerial	
	Gastos personal	Consumo estimado	Gastos personal	Consumo estimado	Gastos personal	Consumo estimado	Gastos personal	Consumo estimado
1986	972.426	712.694	972.426	712.694	928.224	680.299	77.352	56.692
1987	1.112.957	828.476	1.091.958	812.844	1.070.959	797.212	199.492	148.500
1988	1.207.822	896.356	1.147.431	851.538	1.167.562	866.478	231.499	171.802
1989	1.259.911	918.894	1.174.654	856.713	1.222.019	891.258	284.190	207.269
1990	1.592.681	1.137.710	1.504.198	1.074.504	1.468.805	1.049.222	336.233	240.183
1991	1.564.897	1.095.137	1.338.949	937.016	1.523.055	1.065.856	368.211	257.679
1992	2.207.455	1.513.444	1.938.446	1.329.010	2.025.478	1.388.680	419.337	287.500
1993	2.245.835	1.514.324	2.073.078	1.397.838	2.073.078	1.397.838	540.803	364.653
1994	2.494.103	1.693.793	2.186.807	1.485.102	2.215.392	1.504.515	686.057	465.914
1995	2.711.334	1.839.468	2.407.069	1.633.043	2.400.308	1.628.456	818.133	555.051
1996	2.820.616	1.892.668	2.500.685	1.677.991	2.500.685	1.677.991	829.209	556.409
1997	2.964.840	2.008.224	2.624.612	1.777.772	2.576.008	1.744.850	1.044.984	707.817
1998	3.039.482	2.087.542	2.693.525	1.849.936	2.644.102	1.815.992	1.062.583	729.791
1999	3.109.528	2.190.886	2.808.606	1.978.865	2.607.991	1.837.517	1.053.227	742.074
2000	3.111.227	2.219.172	2.416.986	1.723.985	2.545.550	1.815.686	1.079.930	770.291
2001	3.009.237	2.145.202	2.340.517	1.668.491	2.469.117	1.760.166	1.028.799	733.403
2002	2.896.603	2.050.440	2.241.719	1.586.862	2.367.658	1.676.012	982.326	695.367
2003	2.908.393	2.057.987	2.250.843	1.592.703	2.326.715	1.646.390	1.036.905	733.717
2004	2.760.038	1.950.701	2.141.409	1.513.475	2.212.789	1.563.924	975.531	689.472
2005	2.882.586	2.035.161	2.261.339	1.596.548	2.335.888	1.649.182	1.043.695	736.869
2006	3.086.460	2.176.895	2.380.236	1.678.792	2.458.705	1.734.136	1.098.570	774.827
Total	**29.768.393**	**34.965.173**	**42.495.494**	**29.735.721**	**43.140.090**	**30.191.659**	**15.197.068**	**10.625.281**

Fuente: Elaboración propia

3.4. Los efectos indirectos e inducidos del consumo de los trabajadores

En la tabla 9 se muestran los resultados correspondientes al efecto indirecto e inducido derivado del gasto de consumo de los trabajadores. El esquema es el siguiente: en primer lugar se analiza el efecto total comparándose con el efecto directo, indirecto e inducido, para posteriormente centrarnos en la distribución por ramas y períodos del efecto indirecto y del inducido. Recuérdese que, en este caso, el efecto directo corresponde al gasto en consumo de los trabajadores de "El Cabril".

Tabla 9. Efectos directo, indirecto e inducido derivados del gasto de los trabajadores de "El Cabril". Datos en euros constantes del año 2000.

Ramas	Directo	Indirecto	Inducido	Total
Agricultura, ganadería, caza y silvicultura	565.235	1.713.765	1.538.382	3.817.382
Pesca	180.800	53.561	117.398	351.759
Extracción de productos energéticos	893	15.318	13.079	29.291
Extracción de otros minerales, excepto productos energéticos	4.855	62.282	46.258	113.394
Industrias de la alimentación, bebidas y tabaco	2.530.361	1.224.620	2.264.113	6.019.094
Industria textil y de la confección	710.402	93.903	297.565	1.101.869
Industria del cuero y del calzado	215.606	93.335	23.028	331.969
Industria de la madera y del corcho (excepto muebles)	24.020	149.352	113.275	286.647
Industria del papel; edición, artes gráficas y reproducción de soportes grabados	309.999	289.200	398.048	997.248

Ramas	Directo	Indirecto	Inducido	Total
Refino de petróleo	599.750	704.388	1.092.608	2.396.746
Industria química	518.643	317.879	316.037	1.152.559
Industria de la transformación del caucho y materias plásticas	33.937	120.962	82.942	237.842
Industrias de otros productos minerales no metálicos	31.354	220.101	195.369	446.824
Metalurgia y fabricación de productos metálicos	48.825	324.830	217.666	591.322
Construcción de maquinaria y equipo mecánico	169.572	113.921	112.423	395.916
Industria de material y equipo eléctrico, electrónico y óptico	335.269	57.490	49.850	442.608
Fabricación de material de transporte	1.006.338	387.337	160.404	1.554.078
Industrias manufactureras diversas	574.141	80.255	442.281	1.096.676
Producción y distribución de energía eléctrica, gas y agua	601.783	744.802	1.150.557	2.497.141
Construcción	279.687	1.119.436	1.300.680	2.699.803
Comercio, reparación de vehículos de motor, motocicletas y ciclomotores y artículos personales y de uso doméstico	6.118.867	1.165.009	6.996.437	14.280.312
Hostelería	5.329.195	270.055	5.439.174	11.038.424
Transporte, almacenamiento y comunicaciones	1.871.922	1.896.061	3.327.287	7.095.270
Intermediación financiera	1.082.885	883.019	1.792.311	3.758.215
Actividades inmobiliarias y de alquiler; servicios empresariales	4.444.671	2.680.581	6.710.835	13.836.087
Administración pública, defensa y seguridad social obligatoria	0	0	0	0
Educación	645.834	66.452	683.873	1.396.158
Actividades sanitarias y veterinarias; servicios sociales	822.224	121.124	918.476	1.861.824
Otras actividades sociales y servicios prestados a la comunidad; servicios personales	1.841.853	323.477	2.097.964	4.263.294
Hogares que emplean personal doméstico	493.336	0	481.876	975.211
TOTAL	31.392.257	15.292.513	38.380.194	85.064.964

Fuente: Elaboración propia.

El efecto total sobre la economía andaluza derivado del gasto de los trabajadores de "El Cabril" es de 85.064.964,05 euros, mientras que el efecto total derivado de los gastos totales del dicho centro era de 447.789.332,98 euros. El efecto del gasto en consumo de los trabajadores representa un 19% sobre el efecto del gasto total de "El Cabril". La suma de ambos constituye el efecto total sobre la economía andaluza en el período 1986-2006.

Atendiendo al efecto total del gasto de los trabajadores hay que destacar que el efecto directo, que corresponde al consumo de los trabajadores, representa un 36,9% del efecto total, el indirecto un 17,98% y el inducido un 45,12%. De nuevo nos encontramos con que el efecto indirecto es inferior al inducido y al propio efecto directo, lo que como ya se comentaba podía deberse al hecho de que parte de este efecto no se recoja en la economía andaluza, sino que corresponda a importaciones de bienes del exterior y que, por tanto, repercute en el resto de España o incluso en otros países.

Al tratarse del efecto del consumo de las familias que trabajan en "El Cabril", todas las ramas productivas, con la excepción de la rama de administración pública, reciben impactos directos e inducidos, también indirectos, aunque estos últimos tampoco afectan a hogares que emplean personal doméstico.

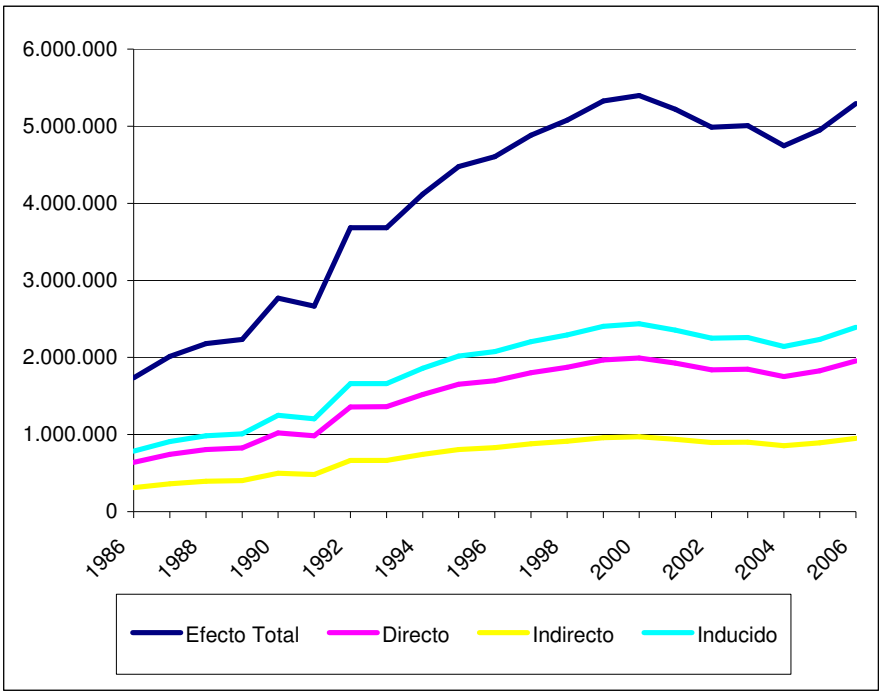
Atendiendo a las ramas de la economía cabe indicar que el efecto directo se concentra en tres grandes sectores, comercio (con un 19,49% del efecto directo total), hostelería (con un 16,98% del total del efecto directo), seguidos de actividades inmobiliarias y servicios empresariales (14,16% del total del efecto directo), la siguiente rama en importancia es la industria de la alimentación, bebida y tabaco (8,06%).

Respecto al efecto indirecto, como se ha indicado se distribuye entre todos los sectores, excepto en el sector público y los hogares que emplean personal doméstico. La rama de actividades inmobiliarias, alquiler y servicios empresariales es la que concentra el mayor impacto indirecto del consumo de los trabajadores de "El Cabril", con un 17,53% del impacto indirecto total, seguido de transporte, almacenamiento y comunicaciones (12,4%), agricultura,

ganadería, caza y silvicultura (11,21%), industria de la alimentación, bebida y calzado (8,01%) y el comercio (7,62%). Al analizar las ramas en las que se concentra el efecto inducido sobre el consumo, se observa que el principal sector es el correspondiente al comercio (18,23% del efecto inducido total), actividades inmobiliarias y servicios empresariales (17,49% del total del efecto inducido), seguido de la hostelería (14,17%) y de la industria de la alimentación, bebida y tabaco (5,90%).

A modo de resumen el gráfico 2 muestra la evolución anual del impacto total, directo, indirecto e inducido del gasto en consumo de los trabajadores de "El Cabril".

Gráfico 2. Impacto regional del gasto de los trabajadores de "El Cabril" (euros constantes del año 2000).



Fuente: Elaboración propia.

4. CONCLUSIONES

Los resultados obtenidos están sujetos a las restricciones propias de la metodología empleada, por lo que a la hora de tomarlos en consideración conviene tener presentes dichas limitaciones. Sin ánimo de ser exhaustivos, cabe señalar los siguientes supuestos o limitaciones inherentes al modelo input-output:

- Es un modelo de corto plazo, pues supone la constancia de los coeficientes estructurales de un año dado y proporciona los efectos que un *shock* de demanda exógena tiene en el año considerado.
- La estructura productiva es constante y no se ve afectada por la inversión realizada. El efecto económico que se recoge de la inversión es el consecuente que tiene que efectuar la economía existente para satisfacer plenamente esa demanda, sin que se contemplen otros beneficios económicos asociados a la mejora productiva de dicha estructura.
- El modelo input-output es lineal y no contempla ni sustitución de factores, ni economías de escala. Este hecho es especialmente relevante a la hora de observar el efecto de creación/mantenimiento de empleo, pues, en última instancia, el modelo acaba asociando de manera lineal una demanda exógena a un nivel de producción y empleo que la satisface. En consecuencia, no se puede hablar, estrictamente, de creación de empleo, puesto que este empleo

dependerá de múltiples factores y, particularmente, de los institucionales del mercado de trabajo.

La cifra total de demanda de la instalación y de consumo de sus trabajadores en la comunidad autónoma andaluza que se desprende de los datos facilitados por Enresa (192.543.904,49 € constantes de 2000) permite suponer, a priori, que la misma tendrá una relativa incidencia, sobre todo en su área de influencia más próxima, como se desprende de la tabla 10. En concreto, este dato implica un efecto total de 532.854.297,02€ constantes de 2000, lo que supone un promedio anual de 25.374.014,14 € que resulta insignificante en comparación con el PIB promedio regional (0,029%). No obstante, cuando se considera el efecto directo provincial (137.381.965,03 € constantes de 2000), éste provoca un efecto total en promedio anual (16.696.936,65 € constantes de 2000) que al confrontarse con las variables representativas del marco territorial más próximo a su zona de influencia (en concreto, .con el PIB promedio de la provincia), da lugar a una significación mayor: 0,211%.

Cuando, en lugar de utilizar como elemento de referencia el PIB, se toman otras variables económicas, por ejemplo, el importe de los presupuestos gestionados por las administraciones públicas de su ámbito geográfico, su importancia relativa aumenta, como también se muestra en la tabla 10. Así, con relación al presupuesto de la Junta de Andalucía, el impacto total medio anual prácticamente se multiplica por cuatro respecto a la calculada con relación al PIB regional. Si se toma el impacto total en promedio anual para en el ámbito provincial y se compara con el total del presupuesto gestionado por la Diputación provincial de Córdoba, el resultado es aún más (12,3%). Algo similar sucede cuando se compara el efecto directo en promedio anual dentro de su zona de influencia[18] lo que da una idea del protagonismo que tiene la instalación en la comarca en la que se ubica y en la provincia cordobesa.

Tabla 10. Actividad del Centro de almacenamiento de RBMA "“El Cabril”" con relación a otras variables económicas de referencia. (Euros constantes de 2000)

	Regional	Provincial	Zona influencia
Inversiones + existencias	70.142.310,15	37.534.324,94	30.138.860,63
Gastos de funcionamiento	91.009.337,17	69.733.996,98	62.669.405,94
Demanda "“El Cabril”"	161.151.647,32	107.268.321,92	92.808.266,57
Consumo trabajadores	34.877.831,11	30.113.643,11	30.191.658,95
Total efecto directo	196.029.478,43	137.381.965,03	126.809.044,17
Efecto total demanda	447.789.332,98	282.066.180,81	
Efecto total consumo trabajadores	85.064.964,05	68.569.488,74	
Efecto total	532.854.297,02	350.635.669,55	
Promedio anual efecto total(1)	25.374.014,14	16.696.936,65	6.038.525,91(a)
PIB Promedio 95-06 - 2000 (2)	87.025.555.521,72	7.916.695.375,94	
Presupuestos (3)	19.391.364.384,30*	135.746.993,28**	
Importancia relativa (1)/(2)	0,029%	0,211%	
Importancia relativa (1)/(3)	0,131%	12,300%	

* Promedio 2000-2007 en euros constantes: ** Promedio 1995-2004 en euros constantes
(a) Corresponde sólo al efecto directo. (b) Importancia tomando el efecto directo promedio.
Fuente: Elaboración propia a partir del INE y del Instituto de Estadística de Andalucía

El impacto sobre la economía andaluza

Como ya se ha indicado, el impacto directo estimado de la actividad sobre la economía regional y provincial, puede interpretarse, en el marco de la metodología input-output, como un verdadero shock de demanda, por la demanda de bienes y servicios derivados de su presencia y por el consumo realizado por sus trabajadores, los cuales generan niveles adicionales de

[18] Recordemos que la zona de influencia está constituida por los cuatro municipios que reciben compensaciones de Enresa, junto con Azuaga y Peñarroya-Pueblonuevo.

producción, renta y ocupación a la economía regional y provincial. Como ya se ha puesto de relieve, para estimar este impacto se ha cuantificado las inversiones y los gastos de funcionamiento realizados por el Centro de almacenamiento de RBMA "El Cabril" durante el periodo 1986-2006; a continuación se ha delimitado con la máxima desagregación posible los sectores en los que incide la demanda de la instalación y de los trabajadores y, finalmente, se ha incorporado a la tabla input-output de la Comunidad de Andalucía esa estructura de gasto sectorial obtenida. Los principales resultados alcanzados se resumen a continuación.

Por una parte, el gasto ejecutado por el Centro de almacenamiento de RBMA "El Cabril" desde 1986 hasta 2006 implica un efecto inicial por el incremento de la demanda final que su actividad genera en distintos sectores productivos. Estos sectores son los mostrados en la tabla 11. Por otro lado, este shock de demanda origina efectos positivos en la economía, al producirse un aumento de la producción que deben realizar los sectores para atender sus necesidades iniciales de demanda. Estos efectos indirectos ascienden a 91.147.153 euros (véase tabla 11). Esto significa que, durante el periodo 1986-2006, la producción adicional generada en Andalucía, para atender la demanda específica realizada por el Centro de almacenamiento de RBMA "El Cabril", representó un 56,60% de la propia demanda.

Como ya se ha indicado, un aumento de la demanda motivado, como en este caso, por la presencia de una empresa, genera efectos no sólo por las relaciones entre las ramas productivas, sino también por el impacto que las rentas generadas provocan en la economía, a través del aumento de gasto en consumo final de los hogares. Así, para atender el incremento final de la demanda y de los efectos indirectos es necesario realizar un aumento de la producción que generará renta en los trabajadores de las empresas, la cual en parte se destinará al consumo de bienes y servicios que, en sí mismo constituye un nuevo efecto del shock de demanda: el efecto inducido sobre el consumo. En el caso que nos ocupa, este efecto inducido asciende a 195.597.592, lo que supone el 121,46% de la demanda. En otras palabras, por cada euro de demanda de la instalación, se realiza una producción adicional de 56,60 céntimos de forma indirecta y 1,21 € adicionales por el consumo realizado por los trabajadores de las empresas que atienden esta demanda.

Tabla 11. Total de efectos de la demanda del Centro de almacenamiento de RBMA '"El Cabril"' sobre el valor de la producción regional (euros) según ramas de actividad (años 1986-2006)

Ramas	Directo	Indirecto	Inducido	Total
Agricultura, ganadería, caza y silvicultura	0	887.567	7.840.079	8.727.647
Pesca	0	59.130	598.298	657.428
Extracción de productos energéticos	0	33.862	66.656	100.519
Extracción de otros minerales, excepto productos energéticos	2.397.470	2.029.603	235.746	4.662.818
Industrias de la alimentación, bebidas y tabaco	101.551	745.386	11.538.634	12.385.571
Industria textil y de la confección	0	105.241	1.516.485	1.621.726
Industria del cuero y del calzado	0	2.499	117.360	119.859
Industria de la madera y del corcho (excepto muebles)	0	1.322.881	577.283	1.900.163
Industria del papel; edición, artes gráficas y reproducción de soportes grabados	0	701.469	2.028.581	2.730.049
Refino de petróleo	0	4.484.422	5.568.274	10.052.696
Industria química	48.185	1.355.813	1.610.622	3.014.620
Industria de la transformación del caucho y materias plásticas	0	571.320	422.699	994.019
Industrias de otros productos minerales no metálicos	703.150	9.308.445	995.662	11.007.257
Metalurgia y fabricación de productos metálicos	920.380	6.256.444	1.109.297	8.286.121
Construcción de maquinaria y equipo mecánico	6.738.803	1.504.874	572.944	8.816.620

Ramas	Directo	Indirecto	Inducido	Total
Industria de material y equipo eléctrico, electrónico y óptico	1.657.208	1.056.815	254.050	2.968.073
Fabricación de material de transporte	0	510.158	817.467	1.327.625
Industrias manufactureras diversas	446.871	181.203	2.254.002	2.882.076
Producción y distribución de energía eléctrica, gas y agua	13.571	2.863.241	5.863.599	8.740.411
Construcción	72.234.728	25.566.877	6.628.676	104.430.282
Comercio, reparación de vehículos de motor, motocicletas y ciclomotores y artículos personales y de uso doméstico	1.894.244	2.956.665	35.656.050	40.506.959
Hostelería	1.963.348	1.416.294	27.719.750	31.099.392
Transporte, almacenamiento y comunicaciones	7.422.447	11.013.546	16.956.907	35.392.900
Intermediación financiera	0	3.816.098	9.134.183	12.950.281
Actividades inmobiliarias y de alquiler; servicios empresariales	11.025.310	11.191.446	34.200.533	56.417.289
Administración pública, defensa y seguridad social obligatoria	48.593.114	0	0	48.593.114
Educación	74.078	349.704	3.485.231	3.909.013
Actividades sanitarias y veterinarias; servicios sociales	3.523.955	1.088.005	4.680.846	9.292.806
Otras actividades sociales y servicios prestados a la comunidad; servicios personales	197.270	857.050	10.691.886	11.746.206
Hogares que emplean personal doméstico	0	0	2.455.791	2.455.791
TOTAL	**161.044.588**	**91.147.153**	**195.597.592**	**447.789.333**

Fuente: Elaboración propia.

Como se aprecia en la tabla 11, el efecto directo que corresponde a este gasto del Centro de almacenamiento de RBMA "El Cabril", se encuentra muy concentrado en algunas ramas de actividad, en concreto, las cinco primeras en orden de importancia, acumulan más del 90% del efecto total en la economía andaluza. Estas ramas son, por orden de importancia decreciente: construcción; administración pública[19]; actividades inmobiliarias y de alquiler; servicios empresariales; transporte; y construcción de maquinaria y equipo mecánico. Cuando se considera el efecto total (por lo tanto, incluyendo además del efecto directo, los efectos indirecto e inducido), las ramas que reciben un mayor efecto son las cuatro primeras ya indicadas y la rama de comercio.

Por otro lado, el impacto ocasionado por la actividad del Centro de almacenamiento de RBMA "El Cabril" no se debe únicamente a la demanda que éste genera, sino también por el pago de rentas (sueldos y salarios) al personal que desempeña en la instalación su trabajo. Si consideramos el efecto derivado del consumo realizado por los trabajadores del Centro de almacenamiento de RBMA "El Cabril", éste se ha estimado en 31.392.257,17 €, lo que representa algo menos del 16% del efecto directo total generado por la instalación. Al igual que sucedía con la demanda del centro, este consumo de los trabajadores provoca, a su vez, un impacto indirecto (15.292.512,93 €), que representa el 48,71% de dicho gasto de consumo, y un efecto inducido (38.380.193,95 €) que supone el 122,26% de dicho dato. En total, el efecto generado por los sueldos pagados a los trabajadores (los cuales dan lugar a su demanda de bienes y servicios) asciende a 85.064.964,05 €, lo que representa el 15,96% del efecto total que provoca la actividad del Centro de almacenamiento de RBMA "El Cabril" (véase tabla 12).

[19] Entre estas dos acumulan el 75% del efecto directo en el periodo 1986-2006, lo cual es lógico ya que una parte muy importante de los gastos ejecutados corresponden a la construcción y acondicionamiento de la instalación y, por otra parte, un elevado porcentaje de los gastos en el ámbito regional tiene como destino la administración pública, no sólo por el pago de impuestos, sino también por los pagos derivados de las compensaciones, y por las cantidades abonadas a éstos por la Fundación.

Desde el punto de vista sectorial, las ramas productivas que se ven afectadas en mayor medida por el consumo de los trabajadores son las mismas que, en términos agregados, recogen la mayor parte del consumo regional[20].

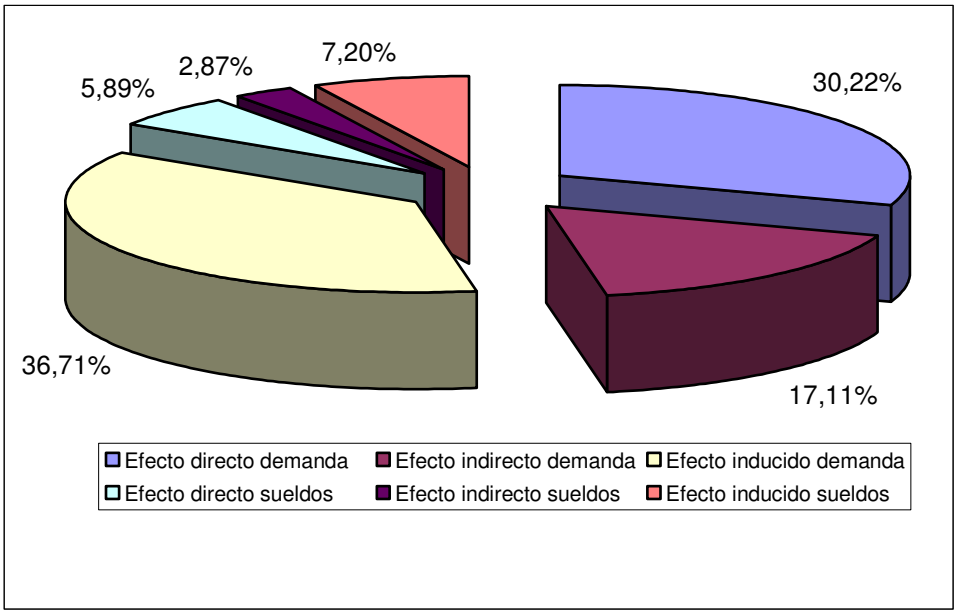
Tabla 12. Total de efectos sobre el valor de la producción regional (euros constantes 2000) (años 1986-2006)

	Efecto directo	Efecto Indirecto	Efecto inducido	Total
Demanda total (1)	161.044.588	91.147.153	195.597.592	447.789.333
Demanda de consumo trabajadores (2)	31.392.257	15.292.513	38.380.194	85.064.964
Impacto demanda sobre total (1)/(3)	83,69%	85,63%	83,60%	84,04%
Impacto salarios sobre total (2)/(3)	16,31%	14,37%	16,40%	15,96%
TOTAL (3)	192.436.845	106.439.666	233.977.786	532.854.297

Fuente: Elaboración propia.

En resumen, el efecto directo de la actividad de la instalación (192.436.845€) da lugar a un efecto global de 532.854.297 €, lo que implica que los efectos indirecto e inducido suponen 340.310.392,53€. Esta cantidad es equivalente a un efecto multiplicador de la producción de 2,77, que indica que cada euro de gasto del centro de almacenamiento de RMBA "El Cabril" provoca un incremento de 1 euro y 77 céntimos de euro más en el conjunto de la economía andaluza. Ésta sería, por lo tanto, una medida aproximada de la creación directa de producción vinculada con la actividad de "El Cabril".

Gráfico 3. Impacto total regional. Distribución porcentual según tipos y demandas (años 1986-2006)



Fuente: Elaboración propia.

[20] En concreto, estas ramas son: comercio (con un 19,49% del efecto directo total), hostelería (con un 16,98% del total del efecto directo), seguidos de actividades inmobiliarias y servicios empresariales (14,16% del total del efecto directo); la siguiente rama en importancia es la industria de la alimentación, bebida y tabaco (8,06%).

BIBLIOGRAFÍA

ALCAIDE INCHAUSTI, J. (2003): "Evolución económica de las regiones y provincias españolas en el siglo XX", Madrid, Ed. Fundación BBVA.

ALCAIDE INCHAUSTI, J. Y ALCAIDE GUINDO, P. (2006): "Balance económico regional (autonomías y provincias) años 2000 a 2005", Madrid, Ed. Fundación de las Cajas de Ahorro (FUNCAS).

APARICIO CASTILLO, S. Y OTROS (2004): "Estudio sobre la incidencia económica y social de la Central Nuclear de Santa María de Garoña en su Zona de Influencia", Burgos, Ed. Fundación Universidad de Burgos y Nuclenor SA.

ENRESA (1997), "El Cabril. Influencia socioeconómica de la instalación. Inversión y Operación", mimeo.

ENRESA (2006), "Impacto económico del desmantelamiento de la central nuclear Vandellós I", Madrid, Ed. Enresa

GARRIDO YSERTE, R. (Dir.), "El impacto de la Universidad de Alcalá sobre la economía local. Un estudio del impacto de la universidad sobre la economía local desde la óptica de la demanda", Alcalá de Henares, Ed. Instituto Universitario de Análisis Económico y Social.

INSTITUTO DE ESTADÍSTICA DE ANDALUCÍA –IEA– (2006), "Marco Input-Output de Andalucía 2000", ed. IEA

MUÑOZ CIDAD, C. (1989): Introducción a la Economía Aplicada: cuentas nacionales, tablas input-output y balanza de pagos, Madrid, Ed. Espasa Calpe.

PULIDO, A. Y FONTELA E. (1993): Análisis input-output. Modelos, datos y aplicaciones, Madrid, Ed. Pirámide.

UTILIDAD CIUDADANA DE LA AGRICULTURA A NIVEL REGIONAL: APLICACIÓN DEL PROCESO ANALÍTICO JERÁRQUICO (AHP)

M. SALAZAR ORDÓÑEZ
E-mail: melania.salazar.ext@juntadeandalucia.es
S. SAYADI
E-mail: samir.sayadi@juntadeandalucia.es
Instituto de Investigación y Formación Agraria. IFAPA
Centro Camino de Purchil
M. J. VÁZQUEZ CUETO
E-mail: pepi@us.es
Facultad de Ciencias Económicas y Empresariales
Universidad de Sevilla

Resumen

En cualquier política económica la mejora del bienestar ciudadano debe ser un precepto para las actuaciones de los gobiernos y sus organizaciones. En la Unión Europea, la Política Agraria Común (PAC), diseñada para el conjunto de los Estados miembros, no es una excepción a este condicionante. Sin embargo, son muchas las críticas sobre su escasa eficacia, particularmente para integrar las prioridades sociales hacia la agricultura.
Así, en esta investigación, a nivel regional de Andalucía (España), se pretende determinar mediante una encuesta a 409 sujetos, las prioridades de los ciudadanos hacia la agricultura y el nivel de satisfacción de las mismas. Posteriormente, se estima un Índice de Utilidad Social (IUS) de dicha actividad. De este modo, se facilitará la contrastación de los objetivos de la nueva PAC con las prioridades regionales identificadas. Para ello, se ha empleado el método multicriterio, Proceso Analítico Jerárquico (AHP), y la estimación de índices.

Palavras chave: Política Agraria Común (PAC); Índice de Utilidad Social; Andalucía; Proceso Analítico Jerárquico (AHP).
Área temática: Economia Urbana, Regional e Local.

Abstract

In any economic policy, to improve the public welfare should be a precept for the actions of governments and their organizations. In the European Union, the Common Agricultural Policy (CAP), designed for the Member States as a whole, is not a exception to this determining fact. However, the criticism about its low efficiency are much, particularly in the integration of the social priorities from agriculture into this policy.
Thus, in this research, at the regional level of Andalusia (Spain) based on a general population survey data (N= 409), we try to determine the citizens' priorities about agriculture and their level of satisfaction. Subsequently, a Social Utility Index (IUS) for this economic activity is estimated. So, the compliance between the objectives of the new CAP and the identified regional priorities are contrasted. With this purpose, the multicriteria method, Analytic Hierarchy Process (AHP), have been applied, as well as index estimation.

Key Words: Common Agricultural Policy (CAP); Social Utility Index (SUI); Andalusia; Analytic Hierarchy Process (AHP).
Thematic Area: Urban, Regional and Local Economics.

1. INTRODUCCIÓN

El bienestar ciudadano debe ser un precepto para las actuaciones de los gobiernos y sus organizaciones, en cualquier política económica (Comisión Europea, 2003). Esta mejora debe ser percibida, a su vez, por los consumidores y contribuyentes, siendo casi imposible que las políticas públicas puedan sustentarse sin su asentimiento. Los decisores políticos de la Unión Europea (UE) aplican dicho precepto, al conceder a la esfera social una gran relevancia, estando la aceptación y justificación de las acciones gubernamentales ante la opinión pública directamente intrincadas con un aspecto fundamental de su marco político.

Dicho aspecto hace referencia al modelo de dirección política establecido, centrado en el objeto político, los ciudadanos. Sin embargo, la Comunidad Europea no posee un proceso de toma de decisiones democrático a escala comunitaria, padeciendo el denominado déficit democrático (Schmitter, 2007), al ser compleja la legitimación[1] de la intervención política, especialmente, si los ciudadanos no se benefician de las acciones de forma igualitaria. Este déficit democrático se ha agudizado recientemente por la crisis económica, conllevando protestas en distintos países, Grecia, Bulgaria, Reino Unido o Francia. De esta forma, "responder a las expectativas políticas normativas de las poblaciones nacionales" (Comisión Europea, 2001) se convierte en un objetivo fundamental y el pilar sobre el que basar la desvalida legitimidad comunitaria. Pese a esta necesidad, se afirma que las distintas políticas, incluyendo la Política Agraria Común (PAC), no han sido capaces de responder de manera satisfactoria a las nuevas demandas ciudadanas, y que cuando lo han intentado "lo han hecho mal con una comunicación aún peor" (García-Azcárate, 2001).

En este sentido, son diversas las opiniones (Givord, 2001; Potter y Burney, 2002; Hall et al., 2004; Arovuori y Kola, 2005, Sayadi et al, 2008, entre otros) tendentes a manifestar que efectivamente la PAC, pese a las reformas[2] ejecutadas y su intento por adaptarse a una realidad llena de incertidumbre, sigue sin responder e integrar adecuadamente los cambios en las prioridades y preferencias de la sociedad hacia la agricultura.

En este contexto, el presente trabajo pretende contribuir a implementar instrumentos analíticos adecuados para la identificación y cuantificación de las preferencias ciudadanas hacia la agricultura y de la satisfacción reportada por dicha actividad, así como a esclarecer si los decisores políticos de la UE han integrado en la nueva PAC las demandas de la sociedad para la que trabajan. La información manejada procede de una encuesta realizada a 409 sujetos a nivel regional de Andalucía (España). Mediante ésta se ha pretendido priorizar las demandas de los ciudadanos hacia la agricultura, y determinar el nivel de satisfacción de las demandas previamente estudiadas. Posteriormente, y fundamentado en los datos anteriores, se ha estimado un Índice de Utilidad Social (IUS) de dicha actividad. Para poder alcanzar los objetivos anteriores, se ha empleado el método multicriterio, Proceso Analítico Jerárquico (AHP), así como la estimación de índices.

El contenido del artículo queda estructurado como sigue. Tras esta introducción, en el segundo apartado se presentan las características e importancia de la agricultura en la región objeto de estudio, Andalucía (España). El tercer apartado muestra la metodología propuesta, el Proceso Analítico Jerárquico y la elaboración del Índice de Utilidad Social (IUS), mientras que en el cuarto se comenta brevemente la estructura del cuestionario y el muestreo realizado. El siguiente recoge los resultados relativos a la priorización de las demandas hacia la agricultura, la determinación del grado de satisfacción de los ciudadanos ante las mismas, y la estimación del IUS. Finalmente, se presentan las conclusiones y recomendaciones.

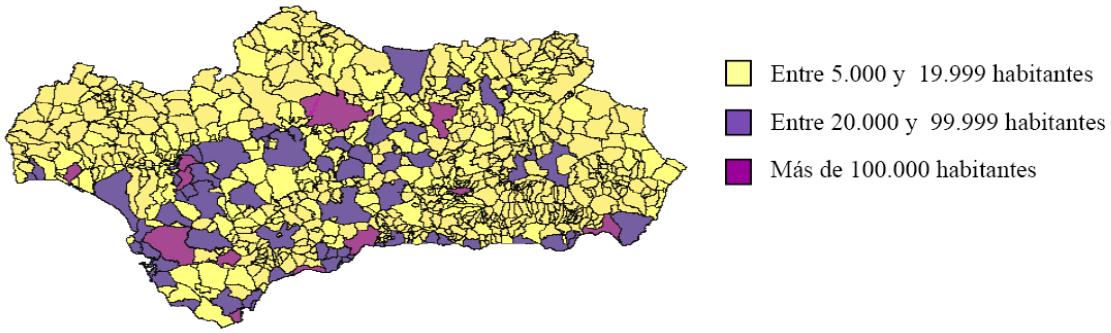
[1] El término legitimidad fue introducido en la Ciencia Política por Max Weber (1864-1920). Cuando se habla de legitimidad se están planteando las razones de la ciudadanía para obedecer al poder, de modo que un Estado es legítimo si existe un consenso entre los ciudadanos para aceptar la autoridad. El proceso mediante el cual una persona obtiene legitimidad se denomina legitimación.
[2] Reforma de 1988, Reforma de 1992, Agenda 2000, Revisión Intermedia de 2003 y Chequeo Médico de 2008.

2. CARACTERIZACIÓN DE LA ZONA DE ESTUDIO: LA AGRICULTURA EN LA REGIÓN DE ANDALUCÍA

La región objeto de esta investigación es Andalucía, una de las 17 comunidades autónomas de España, situada al sur de la península ibérica. Su territorio comprende 87.581 kilómetros cuadrados y cuenta con 8.059.461 habitantes, con un 51% mujeres y un 49% hombres, siendo el 56% de la población menor de 49 años (INE, 2007).

Esta región se organiza administrativamente en ocho provincias y 770 municipios. El 28% de la población se ubica en localidades con menos de 20.000 habitantes; un 34% en ciudades de tamaño medio, entre 20.000 y 100.000 habitantes; y un 38% se sitúa en grandes ciudades, más de 100.000 habitantes (ver Figura 1).

Figura 1. Distribución poblacional de los municipios en la región de Andalucía.



Fuente: Elaboración propia a partir de la Consejería de Economía y Hacienda (2006).

En lo referente a la agricultura en Andalucía, se ha producido una progresiva pérdida del peso relativo de su sector agrario. Actualmente, tiene una aportación al PIB andaluz total a precios corrientes del 5%, alcanzando el 10% si se incluye la aportación de la industria agroalimentaria, y una contribución al empleo del 8,2%. Estas cifras son significativamente superiores a las globales a nivel español, donde la aportación al PIB por parte de la actividad agraria se reduce al 2,6% y del empleo al 4,7%, destacando la inferioridad de los salarios del sector primario en Andalucía respecto al conjunto de España en un 15% (INE, 2007).

Además, la región posee el 18% (4,5 millones de has.) de la superficie agraria utilizada (SAU) española y cuenta con una producción final agraria (PFA) de 10.055 millones de euros (el 25,5% de la producción española). La actividad agraria constituye el principal aprovechamiento del territorio en Andalucía, representando las 4,5 millones de hectáreas (el 68% de las explotaciones son menores de 5 has.) el 55% de la región. En España dicho porcentaje resulta similar al suponer la SAU el 50% del territorio nacional, descendiendo en la UE al 39%.

Los cultivos andaluces más importantes, por su contribución a la PFA, son los hortícolas (35,6%), y el olivar y subproductos (33,1%). Ambos contabilizan más del 68% de la PFA. La Tabla 1 refleja la distribución de la superficie destinada a cada cultivo. De dicha superficie, el 22% es de regadío, aportando un 57% a la PFA y un 60% al empleo agrario.

Tabla 1. Superficie por cultivos en Andalucía año 2005.

Cultivos	Superficie	
	Nº Hectáreas	Porcentaje del Total (%)
Cereales	828.079	25,15
Leguminosas para grano	80.504	2,44
Tubérculos para consumo humano	22.550	0,68
Hortalizas	135.286	4,10
Cultivos industriales	305.605	9,28
Cultivos forrajeros	82.798	2,51
Flores y plantas ornamentales	1.377	0,04
Frutales cítricos	75.074	2,28
Frutales no cítricos	219.342	6,66
Viñedo	41.632	1,26
Olivar	1.499.692	45,56
Total	3.291.939	100

Fuente: Instituto de Estadística de Andalucía (2006).

3. METODOLOGÍA

3.1. El proceso analítico jerárquico

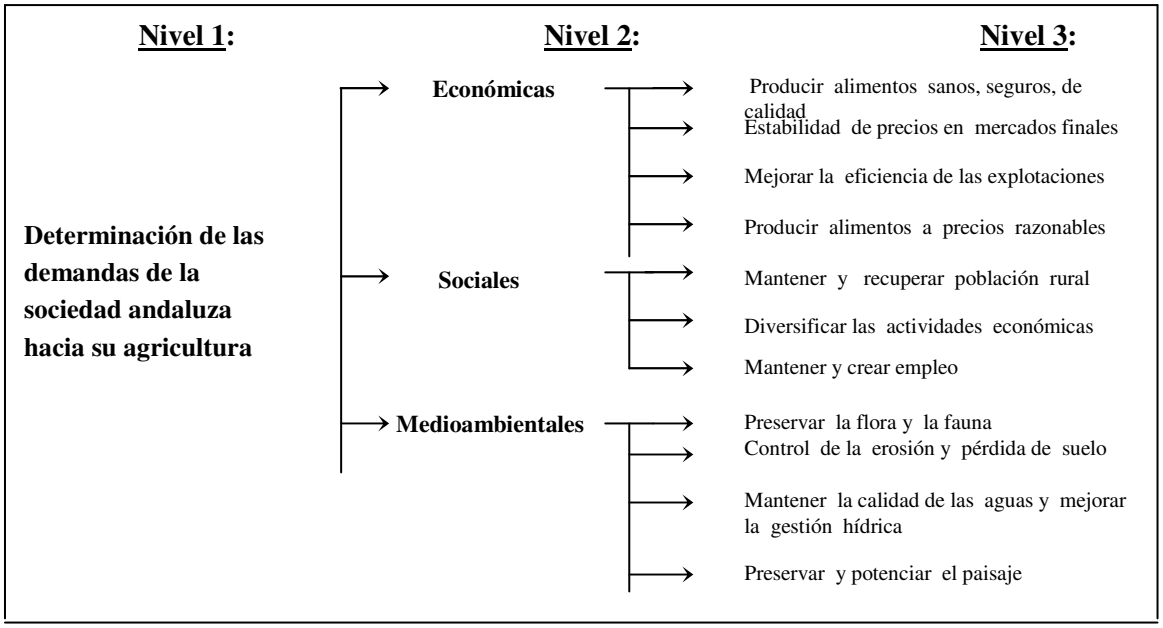
El Proceso Analítico Jerárquico (AHP), creado por Saaty (1977, 1980), consiste en "descomponer una situación en sus partes; ordenar estas partes o variables en un orden jerárquico; asignar valores numéricos a juicios o pensamientos subjetivos sobre la importancia relativa de cada variable, y sintetizar los juicios para determinar qué variables tienen mayor prioridad".

El AHP ha sido empleado, en esta investigación, por la consideración de un modelo complejo de jerarquización con objetivos y suobjetivos, así como por la posibilidad que ofrece de asignar valores a juicios de valor. Aunque en sus inicios dicho método fue concebido como técnica multicriterio para la toma de decisiones, en este caso se ha aplicado a la priorización de las demandas de los andaluces hacia su agricultura. Para ello se han seguido una serie de etapas que se detallan seguidamente.

En primer lugar, se procedió a la identificación y definición del problema, que era la obtención de los pesos de cada una de las demandas hacia la agricultura previamente establecidas. Así, antes de aplicar el AHP, se definieron a priori las demandas, que se iban a incluir en el análisis, mediante la realización de tres Focus Group en distintas provincias andaluzas, a grupos con diferentes características sociales, obteniéndose finalmente 3 demandas genéricas y 11 específicas.

En segundo lugar, se estructuró el modelo de decisión en forma jerárquica, descomponiendo el problema en partes, haciendo grupos (demandas genéricas), y dentro de los mismos subgrupos (demandas específicas). Así, se establecieron tres niveles, como se puede apreciar en la Figura 2, donde se observan las relaciones de los elementos que conforman la estructura mediante la jerarquía.

Figura 2. Estructura jerárquica del modelo AHP de valoración de las Demandas Sociales en relación con la agricultura.



Fuente: Elaboración propia.

En tercer lugar, se estimaron las ponderaciones mediante comparaciones pareadas. El AHP no valora directamente los pesos locales (w_i), sino que su valor se obtiene mediante comparaciones por pares averiguando las razones ($a_{i,j,k}$) entre los elementos de cada grupo de la jerarquía respecto al del nivel inmediatamente superior del que dependen. Para la realización de estas comparaciones, Saaty (1980 y 1994) propone y justifica una escala del 1 al 9 (1: importancia similar entre ambas demandas; 9: importancia absoluta de la primera respecto a la segunda).

Así, una vez que se tienen los juicios de todos los individuos para cada grupo, basados en la información obtenida de las 409 encuestas efectuadas, se construye la matriz teórica de comparaciones pareadas ó matrices de Saaty ($A= a_{ijk}$), representando a_{ijk} el valor de comparación entre el atributo *i* y el atributo *j* del individuo *k*. Esta matriz de comparaciones de pares, *A*, es una matriz recíproca positiva de orden *nxn*, siendo perfectamente consistente o coherente al cumplir: $a_{i,h} * a_{h,j}= a_{i,j}$ $\forall i,j,h \in \{1, 2, 3, \ldots, n\}$. De modo que, los *n* pesos (w_i) podrían determinarse a partir de los *n(n-1)/2* valores de a_{ijk} declarados. Sin embargo, la posibilidad de consistencia perfecta es casi nula, proponiéndose diversas técnicas destinadas a estimar los pesos locales (w_i). El propio Saaty (1980 y 2003) propone como métodos más eficientes la media geométrica y el vector propio principal. Sin embargo, diversos estudios, entre ellos el de Fichtner (1986) ó Zahedi (1986), no han encontrado evidencia estadística de supremacía entre métodos. En este caso, para estimar los pesos locales (w_i), se empleó la media geométrica por filas, aplicando la siguiente expresión algebraica:

$$a_{ij} = \sqrt[n]{\prod_{i=1}^{i=n} a_{ijk}} \qquad \forall i, j \in n$$

Finalmente, se llevó a cabo la síntesis de los pesos locales. La valoración de las demandas, en este estudio, debe ser realizada por la opinión pública, siendo necesaria la agregación de los juicios individuales. Aczél y Saaty (1983), Golden et al. (1989) o Forman y Peniwati (1997) proponen la media geométrica para tal fin, empleando, por tanto, en este caso la agregación de ponderaciones individuales (API). Posteriormente y para poder comparar la importancia relativa de los diferentes subobjetivos (nivel 3 de la jerarquía), se hallaron los pesos globales (w_G). Así, aplicando el modo de síntesis distributiva y multiaditiva se tiene que multiplicando las prioridades locales en cada cluster por las del nodo del que dependen, se obtienen las prioridades globales, es decir:

$w_{G(Sk)} = w_{L(S/i,j)} \times w_{L(Ci)}$

donde $w_{G(Sk)}$ es el peso global del subcriterio *k*; $w_{L(Sk/i,j)}$ es el peso local del subcriterio *k* (nivel 3) respecto al criterio *i, j*; $w_{L(Ci)}$ es el peso local del criterio *i* (nivel 2) respecto a la meta.

De este modo, Zahedi (1986) demuestra que la resolución de problemas mediante AHP es equivalente a la optimización de una función de utilidad multiatributo (MAUF), que ha sido asociada, tradicionalmente, a una función de utilidad aditiva (Kamenetzky, 1982):

$$U(x_j) = \sum_{i=1}^{i=n} w_i \cdot U_i(x_j)$$

donde $U(x_j)$ es la utilidad total proporcionada al decisor por la alternativa *j*, w_i son los pesos de los elementos, y $U_{i(xi)}$ la utilidad parcial proporcionada por el atributo *i* en la alternativa *j*.

Partiendo de esta utilidad aditiva se calcula el Índice de Utilidad Social, al relacionar las utilidades proporcionadas por cada demanda con la satisfacción de las mismas manifestadas por los encuestados.

3.2. El Índice de Utilidad Social (IUS)

Una política económica cuando aumenta el bienestar ciudadano acerca la economía al óptimo social, es decir, al conjunto de ordenaciones ideales de producción de bienes y servicios, asociadas a una determinada distribución de la renta y otros conceptos relevantes, que maximizan el bienestar del conjunto de la sociedad (Nath, 1976). Así, el coste marginal que conlleva a los ciudadanos una inversión pública debe ser igual al beneficio marginal que reciben.

Sin embargo, existe una gran complejidad en la medición de la satisfacción de los individuos, así como en la manifestación de la intensidad de sus preferencias, aceptándose que las estimaciones de utilidad, las del bienestar general, y las comparaciones entre ellas, son juicios éticos sobre los cuales los individuos pueden diferir (Nath, 1976).

Teniendo en cuenta estas dificultades, y para analizar la utilidad que la actividad agraria, como productora de bienes y servicios comerciales y no comerciales[3], reporta a la sociedad de la región de Andalucía, se ha diseñado un índice denominado Índice de Utilidad Social (IUS).

El IUS está basado en tres de las premisas básicas de Pigou (1932), fundador de la economía del bienestar: las unidades en que se mide la utilidad de los individuos es la misma para todos ellos; la utilidad social o utilidad de la sociedad en su conjunto es la suma de las utilidades individuales; y, en la premisa ética de que resulta deseable incrementar la utilidad social. Se asume, igualmente, que la utilidad que un individuo obtiene tanto de bienes y servicios comerciales como no comerciales es medible y cuantificable.

Para su cálculo se ha partido de once demandas que la sociedad de la región analizada realiza a la actividad agraria, estableciéndose, por un lado, la priorización de dichas demandas, es decir su peso para el conjunto de los ciudadanos muestreados; y, por otro, su nivel de satisfacción para cada una de ellas. De este modo, se ha considerado la creación de un índice compuesto basado, en parte, en los modelos de agregación de votos (Cook y Kress, 1990; Puerto et al., 2000) en cuanto al tratamiento del nivel de satisfacción, siendo $\sum_{i=1}^{i} w_i = 1$. De modo que, al ser datos clasificados en categoría ordinales, los conjuntos de información (Φ^i), deben reflejarlo, estableciendo al menos una ordenación de dichos datos de manera que la asignación que se realice a las categorías inferiores no sea más favorable que la estipulada en las superiores (Contreras et al., 2004).

Así, si se considera *N* como el número de unidades a evaluar; D_i, i= 1, …, *N*, como cada una de las once demandas analizadas; y *L* como el número de posiciones de la escala ordinal en la que se miden las valoraciones individuales sobre D_i, se obtendrá la siguiente expresión:

[3] La actividad agraria ejerce una función tradicional de producción de materias primas y alimentos (producción de bienes comerciales), así como todas aquellas funciones de carácter social (por ejemplo, generación de empleo) y ambiental (como preservación de la biodiversidad) (Comisión Europea, 1999) por las cuales el agricultor no obtiene un bien intercambiable en los mercados (producción no comercial).

$$IUS = \sum_{i=1}^{i} \sum_{l=1}^{l} w_i n_{il}/r$$

donde, N_i es el número total de individuos que valoran las demandas i; n_{il} es el porcentaje de votos que la demanda D_i recibe en la categoría l, l= 1, …, L; w_i es la importancia relativa media atribuida por los individuos a cada demanda i; y r es la categoría correspondiente.

Obsérvense que $\sum_{l=1}^{r} n_{il}/r$ representa la suma acumulada de los porcentajes de puntuación de las primeras r categorías (Contreras et al., 2004), pudiendo considerarse, por tanto, la valoración de todas las categorías de votos o sólo las r primeras.

4. DISEÑO DEL CUESTIONARIO Y SELECCIÓN DE LA MUESTRA

4.1. Diseño del Cuestionario

Con el interés situado en conocer las opiniones de la población andaluza acerca de su agricultura y la PAC, así como aplicar la metodología AHP propuesta, se diseñó una encuesta estructurada en cuatro bloques: el primer bloque contenía cuestiones relativas a la opinión de los ciudadanos sobre la actividad agraria en esta región; el segundo, sobre la Política Agraria Común; el tercero incluía el ejercicio multicriterio AHP para la priorización de las demandas de los ciudadanos andaluces en relación con su agricultura; y el último recogía los aspectos socioeconómicos y de estilo de vida de los entrevistados. En el presente artículo se comentan los resultados del tercer bloque. La encuesta se ejecutó entre abril y julio de 2007 mediante entrevistas personales de unos 30 minutos de duración.

4.2. Selección de la muestra

La población objeto de estudio ha sido los ciudadanos de la región de Andalucía mayores de 18 años, por tanto una población objetivo de unos ocho millones de habitantes (INE, 2007). Dada la naturaleza del cuestionario, el tamaño muestral se llevó a cabo para estimación de proporciones, con un error muestral del 5% (nivel de confianza = 95% y p = q = 0,50). Finalmente, se entrevistó a 409 sujetos, y el error muestral resultó del 4,65%.

El muestreo ha sido aleatorio estratificado con afijación proporcional, empleando como variable de estratificación el número de habitantes. Así, se han considerado tres estratos, rural, urbano y metropolitano (Tabla 2), adaptando la clasificación del Instituto Nacional de Estadística (INE, 2007) en función del tamaño poblacional, siendo rural si el número de habitantes está entre 0 y 20.000; urbano, entre 20.001 y 100.000; y metropolitano, más de 100.000. Se ha utilizado como factor de corrección la densidad de población, delimitando los municipios rurales con una densidad menor de 150 hab/km2 (OCDE, 1994).

Tabla 2. Distribución de los entrevistados por estratos.

	Habitantes Andalucía (%)	Número de Entrevistados	% de la Muestra
Rural	27,8	107	26,2
Urbano	34,6	144	35,2
Metropolitano	37,6	158	38,6
Total	100	409	100

Fuente: Cuestionario (2007).

Dentro de cada estrato, con el fin de evitar la infrarepresentación de perfiles, se respetó el porcentaje de representación por provincias, sexos y edades.

5. RESULTADOS

5.1. Priorización de las demandas de los ciudadanos andaluces hacia la agricultura

En la Tabla 3 se presentan los pesos globales para cada uno de los objetivos y niveles propuestos en el modelo jerárquico AHP.

Tabla 3. Pesos AHP globales y parciales de los objetivos y subobjetivos considerados.

Objetivos	Subobjetivos	Pesos globales, $w_{G(Sk)}$	Ranking
Demandas Económicas		0,330	2
	• Producir alimentos sanos seguros y de calidad	0,123*	3
	• Estabilidad de los precios en los mercados finales	0,064	9
	• Mejorar la eficiencia de las explotaciones	0,056	11
	• Producir alimentos a precios razonables para los consumidores	0,086	4
Demandas Sociales		0,313	3
	• Mantener y crear empleo	0,144*	1
	• Mantener y recuperar la población rural	0,085	5
	• Diversificar las actividades económicas	0,083	6
Demandas Medioambientales		0,355	1
	• Preservar la flora y la fauna	0,082	8
	• Control de la erosión y pérdida de suelo	0,083	7
	• Mantener la calidad de las aguas y mejorar la gestión de los recursos hídricos	0,129*	2
	• Preservar y potenciar el paisaje	0,059	10

Fuente Elaboración propia a partir del Cuestionario (2007).
*Diferencias significativas al 5%

Las Demandas Medioambientales son las más valoradas por la sociedad andaluza, con un peso agregado de 0,355, seguidas por las Demandas Económicas y Sociales (0,330 y 0,313, respectivamente). Los pesos obtenidos están estadísticamente muy equilibrados, evidenciando que todas las demandas genéricas (objetivos) son tenidas en cuenta de forma similar por la sociedad a la hora de buscar la plena satisfacción de sus necesidades. La técnica bootstrap ha permitido determinar la no existencia de diferencias significativas entre los pesos ($p \geq 0,05$).

Mantener y crear empleo (MCE) es el subobjetivo más relevante ($w_{G(MCE)} = 0,144$) para la sociedad andaluza. Seguido por *mantener la calidad de las aguas subterráneas y superficiales y mejorar la gestión de los recursos hídricos* ($w_{G(SCAG)} = 0,129$), y *producir alimentos sanos, seguros y de calidad* ($w_{G(SPAS)} = 0,123$). Así, casi el 40%[4] de la preferencia relativa de los entrevistados está representada por dichos subobjetivos. En cuarto lugar aparece *producir alimentos a precios razonables para los consumidores* (PAP) que, con un peso de 0,086, se sitúa a una mayor distancia de los tres anteriores.

La técnica bootstrap ha permitido apreciar que los pesos de mantener y crear empleo (MCE), mantener la calidad de las aguas (CAG) y producir alimentos sanos, seguros y de calidad (PAS) son significativamente diferentes del resto pero no entre sí ($p < 0,05$).

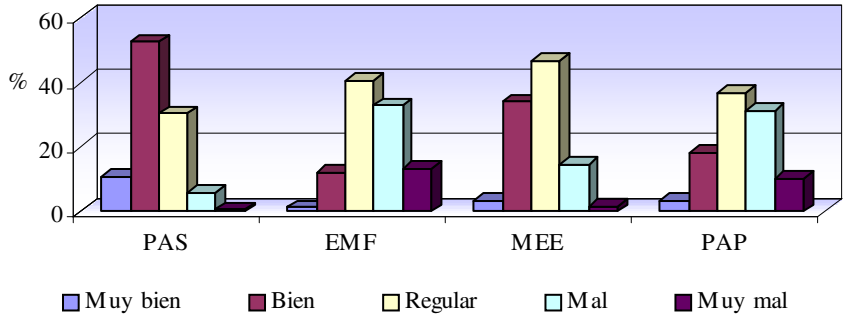
5.2. Demandas sociales y satisfacción de los ciudadanos

La visión que los ciudadanos andaluces tienen sobre cómo satisface la agricultura sus demandas, priorizadas a través del AHP, se cuantificó mediante una escala likert, del 1 (muy mal) al 5 (muy bien).

[4] La suma del peso global de todos los subcriterios debe ser la unidad.

Respecto a las Demandas Económicas (Figura 3), el 63% de los entrevistados afirman que la agricultura satisface al menos bien el *producir alimentos sanos, seguros y de calidad (PAS)*. Contrariamente, el 45,9% y 41,7% estiman una mala y muy mala satisfacción por la *estabilidad de los precios en los mercados finales (EMF)* y por *producir alimentos a precios razonables para los consumidores (PAP)*. Para *mejorar la eficiencia de las explotaciones (MEE)*, el 46,8% de los ciudadanos valoran como "regular" su satisfacción.

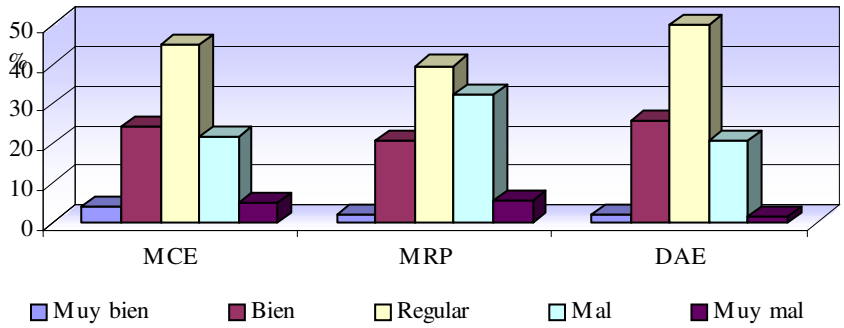
Figura 3. Opinión de los ciudadanos sobre la satisfacción de sus Demandas Económicas (%)



Fuente: Cuestionario (2007).

En el análisis de las Demandas Sociales (Figura 4) se ha detectado que el 44,9% y 49,9% de los encuestados piensan que la agricultura satisface regular el *mantener y crear empleo (MCE)* y *diversificar las actividades económicas (DAE)*, siendo similares los porcentajes de buena y mala satisfacción. Sin embargo, para el 37,8% la agricultura satisface al menos mal el *mantener y recuperar la población rural (MPR)* en Andalucía.

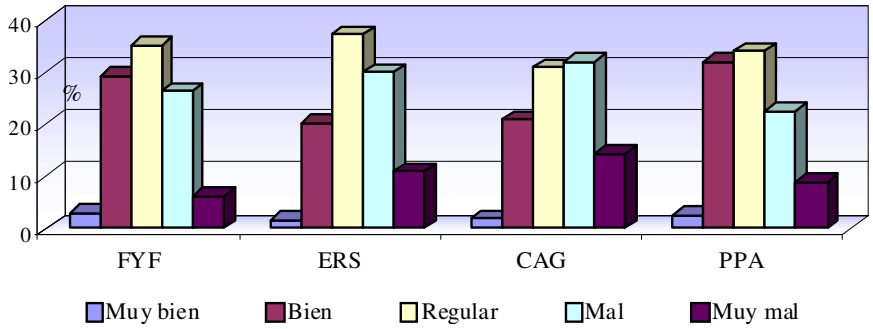
Figura 4. Opinión de los ciudadanos sobre la satisfacción de sus Demandas Sociales (%).



Fuente: Cuestionario (2007).

Entre las Demandas Medioambientales (Figura 5), el 35,1% y 34% de los entrevistados opina que la agricultura satisface "regular" *preservar la flora y la fauna (FYF)* y *preservar y potenciar el paisaje (PPA)*. No obstante, al considerar el *control de la erosión y pérdida de suelo* (ERS), un 40,9% se inclina por una valoración negativa. Algo similar ocurre con *mantener la calidad de las aguas y mejorar la gestión de los recursos hídricos* (CAG), ya que el 46,3% de la opinión ciudadana evalúa su satisfacción como "mal" y "muy mal".

Figura 5. Opinión de los ciudadanos sobre la satisfacción de sus Demandas Medioambientales (%).



Fuente: Cuestionario (2007).

En la Tabla 4 se incluyen las medias de las percepciones sociales del grado de satisfacción reportado por la agricultura andaluza respecto a las funciones evaluadas, así como su ranking y la significación de las diferencias entre dichas medias ($p< 0,05$).

Tabla 4. Percepción social del grado de satisfacción reportado por la agricultura para las demandas establecidas.

Demandas	Percepción Media	Ranking	Coeficiente de Variación	Diferencias Significativas ($p< 0,05$)
Producir alimentos sanos, seguros y de calidad (PAS)	3,68	1	0,206	a
Mejorar la eficiencia de las explotaciones (MEE)	3,23	2	0,244	b
Diversificar las actividades económicas (DAE)	3,05	3	0,257	c
Mantener y crear empleo (MCE)	3,01	4	0,301	c
Preservar y potenciar el paisaje (PPA)	2,97	5	0,338	c
Preservar la flora y la fauna (FYF)	2,96	6	0,326	c
Mantener y recuperar la población rural (MRP)	2,81	7	0,317	d
Producir alimentos a precios razonables para los consumidores (PAP)	2,73	8	0,359	e
Control de la erosión y pérdida de suelo (ERS)	2,72	9	0,355	e
Mantener la calidad de las aguas y mejorar la gestión de los recursos hídricos (CAG)	2,64	10	0,389	e
Estabilidad de los precios en los mercados finales (EMF)	2,56	11	0,355	f
KW= 434,332 ; p= 0,000				

Fuente: Elaboración propia.

De la Tabla 4 se desprende que la mayor satisfacción media es reportada por *producir alimentos sanos, seguros y de calidad* y por *mejorar la eficiencia de las explotaciones*, mientras que la *estabilidad de los precios en los mercados finales* sería la que más se aleja del cumplimiento que la sociedad exige.

5.3. El Índice de Utilidad Social (IUS)

En la Tabla 5 se muestran los resultados de las utilidades parciales calculadas para cada una de las demandas hacia la agricultura (r= 5), así como su priorización y nivel de satisfacción manifestado por los encuestados.

Tabla 5. Utilidades parciales (IUS (D_i)) de las demandas analizadas y ranking (r= 5).

Demandas (D_i)	Utilidades Parciales	Ranking de Utilidades Parciales	Ranking de Priorización según AHP	Ranking de Nivel de Satisfacción
Producir alimentos sanos seguros y de calidad (PAS)	0,060	1	3	1
Estabilidad de los precios en los mercados finales (EMF)	0,020	11	9	11
Mejorar la eficiencia de las explotaciones (MEE)	0,022	9	11	2
Producir alimentos a precios razonables para los consumidores (PAP)	0,030	6	4	8
Mantener y recuperar la población rural (MRP)	0,029	7	5	7
Mantener y crear empleo (MCE)	0,054	2	1	4
Diversificar las actividades económicas (DAE)	0,032	4	6	3
Preservar la flora y la fauna (FYF)	0,031	5	8	6
Control de la erosión y pérdida de suelo (ERS)	0,028	8	7	9
Mantener la calidad de las aguas y mejorar la gestión de los recursos hídricos (CAG)	0,043	3	2	10
Preservar y potenciar el paisaje (PPA)	0,022	10	10	5

Fuente: Elaboración propia.

Como se puede observar, la demanda que reporta una mayor utilidad a la sociedad andaluza es la producción de alimentos sanos, seguros y de calidad (PAP), seguida del mantenimiento y creación de empleo (MCE) y de la calidad de las aguas (CAG), las tres más relevantes según el AHP. Sin embargo, si bien para la primera y la segunda demanda el nivel de satisfacción social parece ser de los más elevados, para la tercera, es de los más bajos. También, se debe señalar que la demanda que menos satisface a los ciudadanos es, a su vez, la que reporta una menor utilidad.

El Índice de Utilidad Social (IUS) de la actividad agraria, ó la medida en qué los ciudadanos andaluces perciben que la agricultura contribuye a su utilidad, y por ende a su bienestar, es 0,371. Por tanto, la sociedad percibe recibir una utilidad tan sólo del 37,1%, bastante reducida, particularmente, si se tiene en cuenta que la agricultura andaluza ha recibido unos 1.500 millones de euros (7% del presupuesto total de la región analizada) en 2008.

Con la intención de estimar un IUS en base a los andaluces que, realmente, perciben alguna utilidad de la agricultura, se han eliminado aquellos sujetos situados en el 1 y 2 de la escala likert (muy mala y mala satisfacción), al suponer que no perciben utilidad (r= 3), siendo los resultados obtenidos los siguientes:

Tabla 6. Utilidades parciales de las demandas analizadas y ranking (r= 3).

Demandas (D_i)	Utilidades Parciales (r= 3)	Ranking de Utilidades Parciales (r= 3)	Ranking de Utilidades Parciales (r= 5)	Tasa de Variación*
Producir alimentos sanos seguros y de calidad (PAS)	0,058	1	1	0,031
Estabilidad de los precios en los mercados finales (EMF)	0,013	11	11	0,341
Mejorar la eficiencia de las explotaciones (MEE)	0,020	8	9	0,098
Producir alimentos a precios razonables para los consumidores (PAP)	0,021	7	6	0,286
Mantener y recuperar la población rural (MRP)	0,022	6	7	0,254
Mantener y crear empleo (MCE)	0,045	2	2	0,170
Diversificar las actividades económicas (DAE)	0,026	4	4	0,149
Preservar la flora y la fauna (FYF)	0,024	5	5	0,201
Control de la erosión y pérdida de suelo (ERS)	0,020	9	8	0,285
Mantener la calidad de las aguas y mejorar la gestión de los recursos hídricos (CAG)	0,029	3	3	0,323
Preservar y potenciar el paisaje (PPA)	0,018	10	10	0,197
IUS= 0,298				

Fuente: Elaboración propia.

*La tasa de variación ha sido calculada como $(D_i\text{-}D_{i\text{-}1}/D_i)$

En la Tabla 6 se reflejan las diferencias obtenidas entre las Utilidades parciales para r= 5 y r= 3, viéndose, en general, reducidas las mismas, siete de ellas entorno a un 20%. Entre las que más varían están, en primer lugar, la estabilidad de los precios en los mercados finales (EMF) y mantener la calidad de las aguas (CAG), y en segundo lugar, producir alimentos a precios razonables (PAP) y el control de la erosión (ERS). Además, se produce una alteración del orden o ranking en cuatro de las demandas estudiadas. El IUS (0,298), también, varía en un 20%.

6. CONCLUSIONES

La función de bienestar de la sociedad andaluza incorpora bienes y servicios comerciales y no comerciales. En este sentido, los resultados confirman la validez del giro experimentado por la PAC desde los años 90, introduciendo paulatinamente para el agricultor el cumplimiento de exigencias medioambientales, así como programas con carácter medioambiental y/o social.

Concretamente, los ciudadanos demandan de la agricultura, en este orden de importancia, que mantenga y cree empleo, mantenga la calidad de las aguas y mejore la gestión de los recursos hídricos, y produzca alimentos sanos, seguros y de calidad. La primera demanda elegida es lógica si se considera que en Andalucía la tasa de desempleo se sitúa en el 14% aproximadamente (INE, 2008), absorbiendo la agricultura el 8% del empleo total. Asimismo, la segunda demanda más preferida también resulta acorde si se tiene en cuenta el grave problema de sequías cíclicas que padece la región, y que incluso ponen el peligro el abastecimiento para uso humano. Igualmente, resulta coherente la preocupación de los andaluces por la producción de alimentos sanos, seguros y de calidad, después de las crisis alimentarias (vacas locas, gripe aviar, etc.) acaecidas en la UE, y su importante difusión en medios de comunicación. Se observa que las demandas priorizadas han sido aquellas que afectan al ciudadano más directamente, bien porque les atañen como consumidores ó bien porque están dentro de sus intereses generales.

Analizando la percepción ciudadana acerca del grado de satisfacción que les aporta la agricultura por el cumplimiento de las demandas establecidas, parece que, en general, esta actividad no satisface dichas demandas, excepto la producción de alimentos sanos, seguros y de calidad.

En relación con el Índice de Utilidad Social (IUS), metodológicamente existen diferencias entre su valor al considerar cinco o tres de las categorías likert, apreciando que desde un punto de vista teórico puede ser más idóneo la inclusión de las tres categorías más valoradas, al reflejar realmente grados de satisfacción del entrevistado. Así, la utilidad reportada por la agricultura a los ciudadanos andaluces resultó tan sólo de un 29%. Para su mejora, y con ello lograr un mayor bienestar social que, a su vez, puede contribuir a la legitimidad de dicha política, a nivel regional se deberían dirigir instrumentos hacia las demandas que reportan una escasa satisfacción, especialmente cuando sean más deseados socialmente, como la creación de empleo en la agricultura y la mejora de la calidad del agua y su adecuada gestión. Así, la PAC, específicamente, para el diseño de medidas agroambientales y sociales debe considerar la componente regional de dichas demandas, siendo recomendable que la UE proponga normativas maleables que posibiliten aplicar medidas enfocadas a aquellos aspectos medioambientales o sociales que el país miembro, e incluso las regiones, aprecie como vitales. Asimismo, debería facilitar un mayor control sobre los precios, sobre todo limitando los crecimientos anormales de márgenes por parte de los intermediarios, para que los productos llegasen a los mercados minoristas con precios ajustados a los niveles de renta de los consumidores, y sin continuas fluctuaciones.

Si se analizan los objetivos por los que se rige la PAC en la actualidad, se observa que están muy relacionados con la competitividad, el sostenimiento de las rentas agrarias, la calidad de los alimentos, el bienestar animal, la sostenibilidad medioambiental, el mantenimiento de los sistemas agrarios tradicionales y la potenciación del desarrollo rural. Sin embargo, estos objetivos no tienen la misma relevancia en términos presupuestarios, reflejo directo de la importancia que se les atribuye por parte de los decidores comunitarios. La competitividad, modernización del sector agrario ó sostenimiento de la renta de los agricultores, demandas no consideradas en esta investigación por los ciudadanos, gozan de una gran relevancia material en la PAC. La seguridad y calidad alimentaria aparece recogida como elemento transversal integrado en todas las políticas comunitarias, fomentándose, dentro de la PAC, con ayudas que provienen básicamente de la Política de Desarrollo Rural (PDR) cuya dotación de fondos no es tan sustanciosa como la Política de mercado. Al bienestar animal, mantener los sistemas agrarios tradicionales, cuidado medioambiental y potenciar el desarrollo rural, prioritarias para la sociedad, se les dirigen medidas más secundarias y menores presupuestariamente que provienen fundamentalmente de la PDR.

Este análisis permite entrever que la Política Agraria Común de la Comunidad Europea, aún, no recoge e integra correctamente las demandas actuales de la sociedad en relación con su agricultura.

Por tanto, es deseable que las autoridades comunitarias realicen diagnósticos de las preferencias sociales ligadas al tema agrario, que permitan identificar, conocer y, posteriormente, cuantificar las demandas de los ciudadanos y sus niveles de satisfacción hacia la agricultura del siglo XXI. Del mismo modo, es importante llevar a cabo un esfuerzo de planificación y diseño de política, que posibilite ajustar adecuadamente la PAC tanto a las necesidades sociales, acercando la producción de bienes y servicios agrarios al óptimo social, como a las del sector.

REFERENCIAS BIBLIOGRÁFICAS

ACZÉL, J., SAATY, T. L. (1983): "Procedures for Synthesizing Ratio Judgements". *Journal of Mathematical Psychology*, nº 27, pp 93-102.

AROVUORI, K., KOLA, J. (2005): "Multifunctionaly policy measures for multifunctional agriculture". XI Congreso de la Asociación Europea de Economía Agraria, *The future of Rural Europe in the global Agri-food system*. Copenhague (Dinamarca), 24-27 de agosto.

COMISIÓN EUROPEA (1999). "Safeguarding the multifunctional role of EU agriculture: which instruments?". *Info-Paper*, octubre. Bruselas.

COMISIÓN EUROPEA (2001): "Libro Blanco de la Gobernanza". COM 2001/428 final. Bruselas.

COMISIÓN EUROPEA (2003): "Consumer Interests in the Common Agricultural Policy efficiency and equity". Directorate General Health and Consumer Protection. Bruselas.

CONSEJERÍA DE ECONOMÍA Y HACIENDA (2006): "Informe Económico de Andalucía 2006", Secretaría General de Economía (Junta de Andalucía), Sevilla (España).

CONTRERAS, I., HINOJOSA, M. A., MÁRMOL, A. M. (2004): "Construcción de índices ponderados multicriterio con información ordinal". *Estadística Española*, nº 155, pp 95-117.

COOK, W. D., KRESS, M. (1990): "A data envelopment model for aggregating preference ranking". *Management Science*, nº 36, pp 1302-1310.

FICHTNER, J. (1986): "On deriving priority vectors from matrices of pairwise comparison". *Socio-economic Planning Science*, nº 20, pp 341-345.

FORMAN, E., PENIWATI, K. (1997): "Aggregating individual judgments and priorities with the Analytic Hierarchy Process". *European Journal of Operational Research*, nº 108, pp 165-169.

GARCÍA-AZCÁRATE, T. (2001): "La Política Agraria Común está en crisis". Disponible en: [http://tga.rippert-online.com/assets/applets/Crisis_de_la_PAC_-_Foro_Agrario.pdf]

GIVORD, D. (2001): "Defensa del modelo rural y agrario europeo en la OMC". *Leader II Magazine*, nº 25.

GOLDEN, B. L., WASIL, E. A., HARKER, P. T. (1989): "Applications of the Analytic Hierarchy Process: a categorized annotated". En Golden, B. L. (ed.): The Analytic Hierarchy Process applications and studies, pp 37-58. Springer Verlag, Berlin (Alemania).

HALL, C., MCVITTIE, A., MORAN, D. (2004): "What does a public want from agriculture and the countryside? A review of evidence and methods". *Journal of Rural Studies*, nº 20, pp 211-225.

INE, Instituto Nacional de Estadística (2007): "Diferentes datos estadísticos". Disponible en: [www.ine.es] Consulta: 2006-2007-2008.

KAMENETZKY, R. (1982): "The relationship between the Analityc Hierarchy Process and the additive value function". *Decision Scienc*e, nº 13, pp 702-713.

NATH, S. K. (1976): "Una perspectiva de la economía del bienestar". Macmillan-Vicnecs-Vives, Barcelona (España).

OCDE (1994): "Territorial Indicators of Employment. Focusing on rural development". Paris.

PIGOU, A. C. (1932): "La economía del bienestar". Macmillan, Londres (Inglaterra).

POTTER, C, BURNEY, J. (2002): "Agricultural multifunctionality in the WTO- legitimate non-trade concern or disguised protectionism?". *Journal of Rural Studies*, nº 18, pp 35-47.

PUERTO, J., MÁRMOL, A. M., MONROY, L., FERNÁNDEZ, F. R. (2000). "Decision criteria with partial information". *International Transactions in Operational Research*, nº 7, pp. 51-65.

SAATY, T. L. (1977): "A scaling method for priorities in hierarchical structures". *Journal of Mathematical Psichology*, nº 15, pp 234-281.

SAATY, T. L. (1980): "The Analytic Hierarchy Process". McGraw Hill, Nueva York (Estados Unidos).

SAATY, T. L. (1994): "Highlights and Critical Points in the Theory and Application of the Analytic Hierarchy Process". *European Journal of Operational Research*, nº 74, pp 426-447.

SAATY, T. L. (2003): "Decision making with the AHP: Why is the principal eigenvector necessary?". *European Journal of Operational Research*, nº 145, pp 85-91.

SAYADI, S.; GONZÁLEZ-ROA, M. C.; CALATRAVA-REQUENA, J. (2008). "Public preferences for landscape features: The case of agricultural landscape in mountainous Mediterranean areas". *Land Use Policy*, nº 26: 334-344.

SCHMITTER P. (2007): "A crisis of representation in European democracies". Instituto Universitario Europeo. 4ª Conferencia, Pisa (Italia), 6-8 septiembre.

ZAHEDI, F. (1986): "A simulation study of estimation methods in the Analytic Hierarchy Process". *Socio-Economic Planning Sciences*, nº 20, pp 347-354.

PERCEPCIONAR DINÂMICAS TERRITORIAIS ATRAVÉS DAS PROXIMIDADES: O CASO DO VINHO VERDE

MANUEL OLIVEIRA
E-mail: moliveira@ubi.pt
ubi_CES – Centro de Estudos Sociais
Universidade da Beira Interior

Resumo

A presente comunicação visa reflectir sobre um conceito introduzido na economia do território em finais do século passado: o de *economia de proximidades*. Trata-se de um arquétipo inovador pois evidencia o papel do território na análise económica regional e atribui um carácter heurístico à análise das relações entre actores económicos. O caso de estudo em análise é um sistema produtivo (do Vinho Verde) com uma identidade própria, uma história e dinâmicas organizacionais e institucionais complexas, ancoradas num território "vivo" (RDVV). Analisam-se as diferentes intensidades relacionais entre actores produtivos e institucionais, o modelo organizacional e os mecanismos de governação construídos e dinamizados no interior daquele sistema produtivo e respectivo território. Conclui-se que, mais do que uma proximidade espacial, importa fundamentalmente uma partilha de redes, regras, espaços de representações, recursos cognitivos e vontades comuns para tornar as externalidades positivas e o jogo *interaccionista* superior à soma das partes.

Palavras chave: Economia de proximidades, dinâmicas territoriais, interacções
Área Temática: Economia Urbana, Regional e Local

Abstract

The present paper aims to reflect on a concept introduced in the regional economy at the end of the last century: The economy of proximities. This is an innovative model that evidences the role of the territory in the regional economy analysis and attributes a heuristically character to the analysis of the relations between the economic actors. The case of study in analysis is a productive system (of the Green Wine) with a proper identity, a history and a complex organization and institutional dynamic, anchored to "an alive" territory (RDVV). The different relation intensities between productive and institutional actors, the organizational model and the governance mechanisms constructed and dinamized inside of that productive system and territory were analyzed. It is concluded that, more than a geographical proximity, basically it matters an allotment of the networks, rules, cognitive resources, spaces of representations, and common will to turn into positive the spillovers and the interactionist game superior than the sum of the parts.

Key Words: Economy of proximities, territorial dynamics, interactions
Thematic Area: Urban, Regional and Local Economics

1. INTRODUÇÃO

Há um longo caminho entre o se que pretende e o que é possível planear e concretizar, e o desenvolvimento dos territórios não foge ao crivo desta premissa transversal. As trajectórias socioeconómicas dos territórios dependem de inúmeros factores relacionados com (i) o tipo de recursos e a sua mobilização/valorização; (ii) a transferibilidade ou não dos recursos e dos saberes e a assimilação/adequação dos recursos e saberes externos e (iii) o tipo de actores envolvidos no processo de co-construção do território e o modo como são coordenados, organizados e se relacionam entre si vertical e horizontalmente. Estas trajectórias, *longe de serem inelutáveis, exprimem relações entre presente, passado e futuro* (Mendez e Mercier,

2006:7), isto é, a história vivida e construtiva, assim como as actuações presentes e estratégias condicionam o ciclo de desenvolvimento dos territórios.

Diversas abordagens explicativas das dinâmicas territoriais baseadas nos sistemas produtivos tiveram início no último quarto do século passado. Estas têm feito emergir linhas de pensamento complementares e, de certa forma, já amadurecidas. Falamos de abordagens como as dos Distritos Industriais (Becattini, 1992 e Antonelli, 2000), dos Meios Inovadores (GREMI, 1986-2004), dos Clusters (Porter, 1990 e 2000) e outros sistemas produtivos como os Distritos Tecnológicos (Zimmermann, 1995) ou Tecnopolos (Courson, 1990). Estas têm contribuído para (i) um maior entendimento das lógicas comportamentais dos agentes as constituem; (ii) o desenvolvimento dos territórios a que pertencem e (iii) o surgimento de dinâmicas de inovação, sob as suas mais diversas formas (Edquist, 1997), e de externalidades positivas (tecnológicas, pecuniárias, de conhecimento) resultantes de uma "Atmosfera Industrial" (Marshall, 1890) própria da aglomeração da actividade económica.

Um entendimento holístico sobre as dinâmicas e efeitos dos sistemas produtivos nos territórios implica pois um tratamento multi-conceptual e dimensional dos comportamentos dos actores e das instituições. Assim, consideramos neste ensaio uma abordagem que permite a congregação de diversos conceitos caracterizadores dos diversos fenómenos territoriais atribuindo ao conceito de e*conomia de proximidades* o papel principal. Neste contexto a presente comunicação visa reflectir sobre uma forma emergente e em construção capaz de percepcionar e analisar as interacções tecidas entre actores económicos e as suas repercussões na construção dos territórios.

Este trabalho divide-se em dois pontos essenciais: (i) num primeiro ponto pretende-se abordar o conceito de economia de proximidades, apresentando as principais contribuições teóricas e introduzindo argumentos teóricos complementares no sentido de alargar o ângulo de análise das dinâmicas territoriais; (ii) o segundo ponto visa operacionalizar, com base num trabalho de investigação já realizado (sobre o sistema produtivo do vinho verde), este conceito, permitindo observar e analisar o mosaico de interacções presente neste sistema produtivo e, assim, entender as dinâmicas de construção do território em causa.

2. ECONOMIA DE PROXIMIDADES: UM CONCEITO EM CONSTRUÇÃO

A introdução do conceito de economia de proximidades na literatura regional tem sido, desde 1993, protagonizada por um conjunto de economistas guiados por uma agenda de investigação denominada de "Dinâmicas de Proximidade", cujo objectivo é *contribuir para a endogeneização da variável espacial na teoria económica e explicar a natureza dos efeitos de proximidade* (Gilly e Torre, 2000:283). De uma forma sucinta, a grande motivação desta nova abordagem conceptual resulta de uma necessidade de acrescentar à análise dos sistemas produtivos (distritos industriais, pólos de competitividade, clusters…) uma interpretação mais abrangente da noção de proximidade, até então entendida numa perspectiva física e determinante na manifestação de efeitos positivos (spillovers pecuniários, tecnológicos e de conhecimento, dinâmicas de inovação) no território. Neste novo entendimento, *a história, as aprendizagens, as dimensões organizacionais, a confiança que se instaura nas relações, contam, por vezes mais que a proximidade geográfica* isoladamente (Rallet e Torre, 2007:11). O facto de nos sentirmos "próximos" não deve ser interpretado como uma condição inequívoca do estabelecimento de relações (Talbot, 2008).

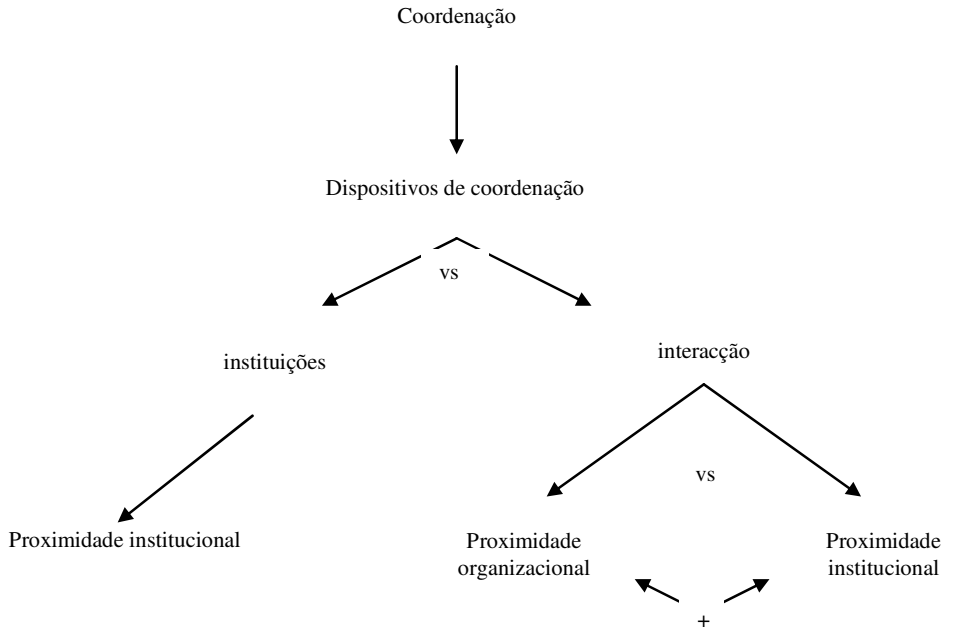
Partindo da ideia de que o espaço não resulta de uma existência pré-determinada, mas de uma construção recíproca dos indivíduos e agentes que dele fazem parte, são compreendidos jogos de interacção e de cooperação proporcionados por uma *percepção identitária*[1] destes com o território e pela capacidade em aproveitar recursos (tangíveis e imateriais) no sentido de criarem oportunidades geradoras de dinâmicas territoriais. Os processos de activação destes recursos levados a cabo pelos agentes institucionais e sectoriais põem a tónica nas lógicas de

[1] Pecqueur e Zimmermann (2002) apontam para uma *singularidade individual territorializada* que estabelece uma ordem e um sentido colectivos alimentados pelo desejo individual de construir a sua própria história num contexto cognitivo colectivo.

mercado tradicionais e na demonstração de jogos de cooperação baseados em factores de proximidade geográficos, institucionais e organizacionais (Pecqueur e Zimmermann, 2002).

De acordo com Bouba-Olga e Carrincazeaux (2008:280) o espaço não fornece por si mesmo as condições necessárias de coordenação territorial e os actores em situação de proximidade são *levados a desenvolver, para resolver problemas de coordenação, dispositivos organizacionais e institucionais*. Pecqueur e Zimmermann (2004) definem duas formas canónicas de proximidade que constroem os modos de coordenação: a proximidade *organizacional* e a proximidade *institucional*. A primeira pode ser considerada como um conjunto de factores que favorece a troca de informação suportada por fluxos de natureza diversa (financeira, comercial, etc.), isto é, estabelece, baseado na reciprocidade de interacções repetidas, um quadro organizacional associado à *similitude* (agentes que possuem o mesmo espaço de referencia e partilham os mesmos saberes) e à *pertença* (actores que pertencem ao mesmo espaço de relações – indústria ou rede – sob as suas diferentes formas) (Gilly e Torre, 2000:12). A proximidade institucional con*strói-se pelo reconhecimento de uma ligação recíproca entre o agente e o colectivo, mediatizada pela estrutura colectiva* (Collectis e Rychen, 2004:220). Por outras palavras, a coordenação entre agentes é definida pelo estabelecimento de espaços de representações, regras e linguagem comuns, quer esta seja estimulada por uma lógica de pertença, ou baseada numa escolha racional.

Figura 1 – Comportamento dos actores sob formas canónicas de proximidade



Fonte: Pecqueur e Zimmermann (2004:31)

Estas formas de proximidade podem implementar dispositivos de coordenação isoladamente, porém não são mutuamente exclusivas. No sentido de construir uma coordenação territorial efectiva e sustentável no tempo, criando mecanismos de travagem a choques externos, esta deverá assentar numa proximidade organizacional sem abdicar do facto de os diferentes actores partilharem um conjunto de representações e regras comportamentais para o seu bom funcionamento (Chalaye e Largeron, 2008).

Apesar de se pretender separar, de um ponto de vista meramente analítico, a proximidade geográfica, esta compagina a estrutura base da constituição dos modos de coordenação territorial pois representa o espaço em que tudo se desenrola. Diversas análises empíricas revelam uma correlação positiva entre a proximidade geográfica e a manifestação de externalidades pecuniárias, tecnológicas e de conhecimento, bem como de lógicas inovadoras (Audretsch e Feldman, 1996, Anselin, Vargas e Acs, 2000, Breschi e Lissoni, 2001, Cooke, 2005). Esta tipologia de proximidade está associada à possibilidade de relacionamento entre os

agentes em função da distância que os separa, bem como aos efeitos de externalidades ligados à aglomeração dos agentes num determinado território. De um ponto de vista normativo e funcional a proximidade geográfica é aqui entendida como a distância que separa duas entidades (indivíduos, organizações, territórios…). De acordo com Torre (2006) esta pode ser de dois tipos: (i) binária – saber se estamos "longe de" ou "perto de" e (ii) duplamente relativa – primeiro a distância é ponderada pelo tempo e pelo custo de transporte; segundo decorre de um *julgamento,* uma representação feita pelos indivíduos sobre a natureza da distância que os separa.

Pecqueur e Zimmermann (2004:36) identificam um conjunto de situações em que a proximidade geográfica influencia positivamente a coordenação na medida em que:

- Pode facilitar o encontro e, dessa forma, estabelecer relações entre agentes que dispõem de um potencial ligado à proximidade institucional;

- Pode, quando a relação foi estabelecida, facilitar a interacção directa através do recurso ao "face a face";

- É susceptível de compensar um defeito ou uma insuficiência de proximidade de natureza não essencialmente geográfica (organizacional ou institucional);

- Produz, em função de algumas condições, uma configuração organizacional particular: o território.

A perspectiva tripartida presente no conceito de economia de proximidades fornece um quadro analítico suficientemente lato para não cingirmos a análise das dinâmicas territoriais a um só dos seus elementos propulsionadores e coerente pois as suas dimensões complementam-se. Não obstante, um maior aprofundamento das suas vertentes é viável dado que apresenta um carácter aberto e flexível.

2.1. Proximidade institucional e governação territorial

O estabelecimento de regras e espaços de representações necessárias à verificação da proximidade institucional pressupõe a participação activa de actores chave do processo de desenvolvimento dos sistemas territoriais. Daí que a esfera da governação deva ser entendida, não só como um modo coordenativo de instituições ligadas ao estado, mas também contemplar o poder dos actores privados através de várias formas com associações empresariais, grandes grupos económicos, etc. através dos quais se procura favorecer uma coordenação partilhada entre vários actores económicos intrinsecamente ligados àqueles. *Isto implica então falar mais de uma mistura dos espaços de coordenação do que de uma estrita hierarquia, sob o controlo de regras jurídicas, entre os níveis pertinentes* (Gilly, Leroux e Wallet, 2004:192-193). O principal desafio da governação consiste pois na *coordenação de diversas ordens relacionais* (Reis 2007:37), sem assentar em *relacionamentos de poder concebidos em função de uma verticalidade entre autoridades ordenadas de forma hierárquica* (Canet, 2004:3).

Assumimos neste texto que a construção dos modos de coordenação depende de duas regras complementares. Do contínuo e sistemático estabelecimento de interacções cristalizam-se nos comportamentos hábitos, códigos e valores (*regras informais*) e da necessidade de tornar mais eficiente o desempenho da economia são impostas *regras formais* (regras políticas e judiciais, contratos, regras de direito económico). Estas regras formam as 'instituições' orientadoras e mediadoras da acção colectiva (Commons, 1951, North, 1990). Aqui as instituições são entendidas como *systems of established and embedded social rules that structure social interactions* (Hodgson 2006:18).

De uma forma sucinta e utilizando argumentos da economia neo-institucionalista (evolucionista), os mecanismos de coordenação deverão minimizar os custos de transacção (informação assimétrica, incerteza, conflitos, comportamentos oportunistas, entre outras *fricções* do mercado) verificados, tendo por base o estabelecimento de convenções, contratos e negociações entre actores constituintes do sistema de governação [territorial] (Coase, 1960, North, 1990, Williamson, 1985). As instituições (no sentido amplo que temos vindo a dar-lhe) serão então o mecanismo de governação apropriado à minimização dos custos de transacção positivos invocados (Grandori, 1997).

Numa perspectiva em que as interacções contam mais do que qualquer pressuposto ou modo de coordenação entendido *ad hoc*, a governação deve traduzir uma interdependência entre os poderes e a acção colectiva, isto é, contribuir para a articulação das variáveis institucionais e as novas formas de acção colectiva (Mayoukou *et al.*, 2003). Este entendimento implica um *feedback* recíproco entre organização e governação, entre indivíduos/actores económicos e instituições.

2.2. Proximidade organizacional e capital social

Como vimos, estão associados à proximidade organizada a similitude (partilha, numa mesma rede de actores individuais e colectivos, dos mesmos códigos interpretativos do conhecimento) e um espaço de referência comum. Ora, a organização colectiva parte, antes de mais, do modo como cada indivíduo, cada organização se posiciona numa rede inter-actuante e para haver interacção é preciso uma ideia, uma iniciativa conjunta, objectivos partilhados, confiança, no fundo, um capital diversificado de sociabilidades.

O conceito de capital social é aqui entendido como um conjunto de recursos (no sentido lato, informação, bens, serviços, estrato social, habilitações, …) geridos pelos indivíduos e organizações ligados por relações sociais e um objectivo comum, portanto é um "stock" que flui e é reciprocamente alimentado pelo estabelecimento de relações numa rede mais ou menos permeável/densa (Bourdieu, 1986, Coleman, 1988, Putnam, 1993 e 2000, Callois, 2004). Os conceitos de norma, confiança e reciprocidade, bem como os de "voice and loyalty" (Cooke e Morgan, 1995) entram enfaticamente no discurso como facilitadores da cooperação e da obtenção de benefícios mútuos. Nesta perspectiva, não devemos considerar a variável confiança exógena ao processo de crescimento económico, como a corrente neo-clássica o faz, mas sim intrínseca às dinâmicas comportamentais dos agentes socioeconómicos. Alguns estudos (Knack e Keefer, 1997 e Alesina e Ferrara, 2002) demonstram que há uma correlação positiva na explicação do crescimento económico pelo nível de confiança das populações em relação ao outro.

Com maior ênfase dada por economistas, as tentativas de medição do capital social, através de métodos recentemente introduzidos na teoria económica, têm sido várias e evoluído positivamente[2]. Os modelos que lhes subjazem não deixam, porém, de assentar em pressupostos utilitaristas e racionais, ocultando determinados comportamentos atípicos ou não, incapazes de serem absorvidos pelos modelos construídos. A complexidade na materialização de comportamentos subjacentes ao conceito de capital social põe a tónica na subjectividade de alguns conceitos como os de confiança, normas, valores e atitudes, que são entendidos como uma forma cognitiva e não estrutural, portanto "a-pressuposta".

Relativamente aos activos estruturais que suportam o capital social as tentativas de medição têm sido diversas e consubstanciam um vasto universo de pesquisa. Uma análise interessante foi introduzida por Landry *et al.* (2001) através de um estudo em que a medição do capital social estrutural assentou em três activos relacionados: (i) activos de rede, (ii) activos de participação e (iii) activos relacionais. Os primeiros dizem respeito à estrutura da rede de relações em que determinado organismo está assente e o tipo de informação partilhada[3] com os diferentes agentes conectados[4], essencialmente, em três áreas fundamentais: mercado; informação e redes de conhecimento. Os segundos pretendem medir o nível de relacionamento formal e informal existente entre organizações, relativamente à intensidade e ao tipo de relações (número de projectos existentes em comum, partilha diversificada de informação, convergência ao nível dos interesses específicos, etc.). Por último, os activos de participação constituem a frequência de comunicação entre os parceiros sectoriais e institucionais a três níveis: local, nacional e internacional.

Ao olharmos para o conceito como um todo é clara a necessidade de apreendermos as dinâmicas cognitivas e estruturais. É natural que as metodologias associadas à interpretação

[2] Um dos modelos mais utilizados assenta na teoria dos jogos, mais especificamente na lógica subjacente ao dilema do prisioneiro. Porém, este baseia-se essencialmente na medição dos comportamentos estratégicos entre indivíduos ou organizações e não nos diferentes comportamentos sociais.

[3] Conhecimento tecnológico, saberes fazer próprios, estrutura de mercados, oportunidades de negócio, bases de dados diversas, disposições regulamentares e legais, etc.

[4] Fornecedores, clientes, concorrentes, centros de racionalidade e competência, instituições regulamentadoras, etc.

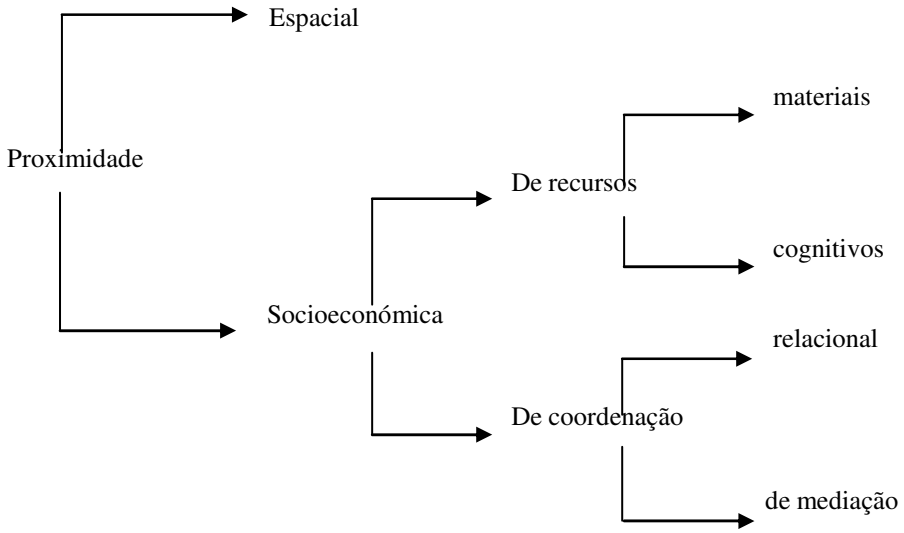
destas dinâmicas deverão assentar em modelos distintos, mas complementares sob pena de revelarmos uma visão apenas parcial do fenómeno em causa.

2.3. Em direcção a uma nova tipologia de proximidades…

A flexibilidade do conceito de economia de proximidade permite a interiorização de diversos avanços conceptuais teorizados e experimentados nas fronteiras entre a economia, a sociologia e a geografia. O carácter "em construção" deste conceito levou recentemente dois investigadores (Bouba-Olga e Grossetti, 2008) a propor uma nova tipologia, considerando todos os avanços adquiridos até então.

Tal como tem vindo a ser enfatizado, esta tipologia apresenta uma separação entre a proximidade geográfica/espacial e a organizada/socioeconómica. A diferença mais marcante entre esta tipologia e a proposta por Pecqueur e Zimmermann reside na centralidade que é dada à proximidade socioeconómica que estrutura as trocas (no sentido lato) e as actividades, em vez de uma coordenação dos indivíduos e das organizações.

Figura 2 – Tipologia das formas de proximidade



Fonte: Bouba-Olga e Grossetti (2008:322)

Nesta tipologia a proximidade socioeconómica é representada por duas tipologias principais (material e de coordenação) que, por sua vez, se subdividem em mais duas categorias.

(i) A proximidade de *recursos* admite um carácter material e outro cognitivo. Por sua vez a proximidade *material* pressupõe a similaridade e complementaridade do património, nível de conhecimento, status social, retorno económico e a proximidade *cognitiva* é entendida como a similaridade e complementaridade de vontades, objectivos, valores, rotinas, referências, etc., exteriorizadas pelos discursos e acções dos indivíduos. A manifestação de proximidades depende da verificação de condições prévias, das quais se destaca a confiança que tem carácter cumulativo na criação de redes sociais sólidas e da partilha de objectivos comuns.

(ii) A proximidade de *coordenação* compreende um carácter relacional e outro de mediação. A proximidade *relacional* é materializada pelo suporte (activos de rede) das interacções e pela construção de regras informais que influenciam as possibilidades coordernativas. No entanto esta estrutura em que tudo flui nem sempre consubstancia uma condição satisfatória de coordenação, sendo necessário recorrer a recursos (proximidades) de *mediação* (regras formais) no sentido de corrigir comportamentos oportunistas, diminuir as incertezas, enfim, minimizar os custos de transacção através das instituições (normas, regras, contratos…).

As relações entre os indivíduos e entre as organizações não têm um comportamento linear, não se podendo considerar as proximidades socioeconómicas de modo isolado, mas sim de forma dialéctica e complementar. A proximidade cognitiva é muitas vezes uma condição prévia à efectivação da proximidade relacional e a causa do recurso à proximidade de mediação.

3. EVIDÊNCIAS EMPÍRICAS: O SECTOR DO VINHO VERDE COMO SISTEMA PRODUTIVO ANCORADO TERRITORIALMENTE[5]

3.1. Contexto histórico e a estrutura da fileira produtiva

O objecto de análise circunscreve-se à região agrícola de Entre Douro e Minho[6] onde se produz, transforma e comercializa o vinho verde. Nele inscreve-se um sistema produtivo que teve uma evolução histórica baseada num recurso intransferível (vinho verde). Desde o século XVI, com intervenção de diversos reinados, este recurso tem sido valorizado e aproveitado economicamente, a princípio com impacto regional, mas que, nas últimas 4 décadas (a partir da entrada de Portugal na CEE), se tem expandido para outras geografias nacionais e internacionais. Com a valorização deste recurso estratégico regional emergiram diversos actores públicos e privados. O forte crescimento da produção (muito influenciado por instrumentos financeiros da PAC e dos Planos de Desenvolvimento Rural das décadas de 80 e 90) e do retorno associado a esta actividade despoletaram a entrada de um número significativo de actores produtivos que concorrem com um modelo pouco eficiente baseado na lógica cooperativa implementada nos anos 50. A entrada na CEE também permitiu a estruturação de uma esfera reguladora e coordenadora do sector vitivinícola nacional (IVV[7]). Com essa entrada e dadas as especificidades vitivinícolas regionais criaram-se, em fins da década de 80, organizações interprofissionais que concorrem para uma afirmação positiva dos vinhos e actores que representam nos palcos nacionais e internacionais (no caso concreto a CVRVV[8]).

A compreensão das dinâmicas territoriais em torno de um sistema produtivo implica primeiro entender a sua estrutura, o posicionamento dos actores que o compõem e as interacções estabelecidas. A estrutura da fileira[9] do Vinho Verde assenta em três eixos fundamentais (viticultura, vinicultura e comercialização).

(i) O primeiro integra a participação de vários agentes, desde os produtores individuais, produtores engarrafadores aos armazenistas vinificadores engarrafadores. Os primeiros são maioritariamente pequenos agricultores que, a título integral ou complementar, produzem a uva e o vinho para ser comercializado por outros agentes. Os segundos produzem o seu próprio vinho e comercializam-no com uma ou várias marcas associadas; para além de uma produção própria e no sentido de esgotarem os stocks existentes, alguns também vendem parte da sua produção a agentes de maior dimensão (armazenistas) e capacidade para colocarem os seus vinhos nos mercados nacional e internacional. Os terceiros também têm explorações próprias de grande dimensão mas, na sua maioria, têm de recorrer aos produtores individuais e engarrafadores para completarem a sua necessidade produtiva.

(ii) Numa fase posterior são empresas, centradas na vinicultura, que acrescentam valor a um produto que encontra a sua diversificação, não apenas nas diferentes castas de vinho produzidas nesta região demarcada, mas também na necessidade de dar resposta a tipos de consumo distintos. Estas sociedades podem integrar grande parte da cadeia produtiva. Por outro lado, encontramos agentes que apenas integram na sua cadeia fases situadas a jusante do processo, tais como as adegas cooperativas que começam a sua actividade produtiva na compra de uva aos seus associados[10] e alguns armazenistas engarrafadores que iniciam o seu processo na compra de vinho a vários produtores (quer eles sejam individuais ou engarrafadores).

(iii) Por fim, a comercialização dos vinhos processa-se por vários canais de distribuição: por um lado, é direccionada uma parte pouco significativa da produção directamente ao consumidor final, através de estabelecimentos próprios para o efeito, por outro, e de forma muito significativa, são comercializados os vinhos a vários agentes, designadamente posicionados em

[5] Com base nos trabalhos empíricos de uma dissertação de mestrado (Oliveira, 2006). A recolha empírica incidiu em duas sub-regiões (Ave e Amarante, que compreendem 7 Concelhos) com elevada representatividade no conjunto da RDVV e pautou-se pela aplicação *in-loco* de questionários diferenciados (aplicados a 15 produtores individuais e 21 empresas) e entrevistas semi-estruturadas (actores institucionais) a actores pertencentes ao sistema produtivo e institucional desse sistema produtivo.

[6] Especificamente a Região Demarcada dos Vinhos Verdes (RDVV) que conta com 47 concelhos agrupados e 9 sub-regiões vitivinícolas.

[7] Instituto da Vinha e do Vinho.

[8] Comissão Vitivinícola da Região dos Vinhos Verdes.

[9] A noção de fileira tem aqui uma interpretação que extravasa o encadeamento lógico e linear do processo produtivo, na medida em que permite interpretar as relações entre as actividades produtivas e os actores relacionados que revelam tanto de interdependência e de complementaridade como de hierarquização.

[10] O número de sócios varia de adega para adega, mas é significativo podendo representar de 600 a 1560 associados (maioritariamente produtores individuais).

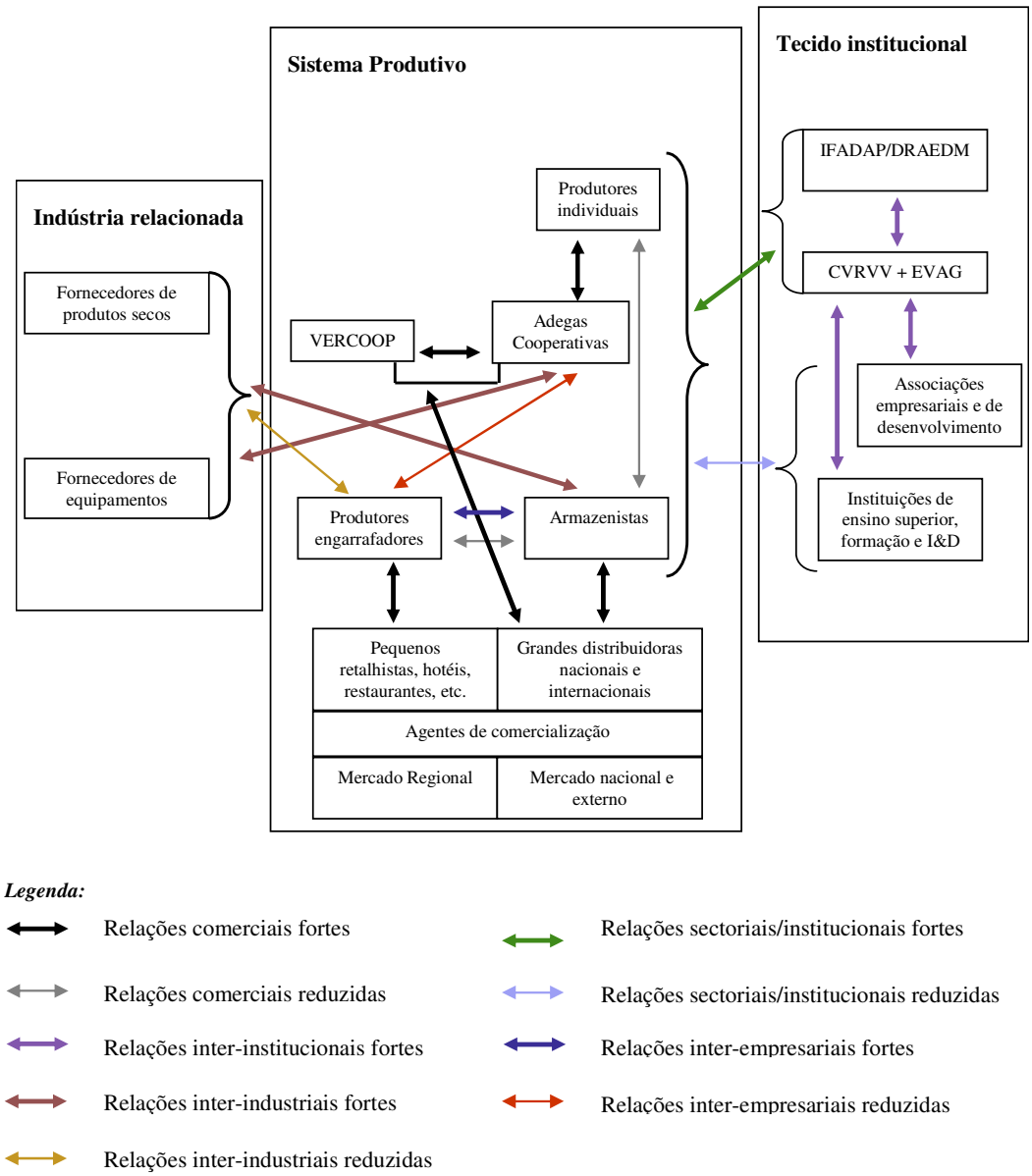
circuitos de distribuição nacionais e internacionais, hiper e super mercados, retalhistas, restaurantes, etc.

Associados a estes agentes principais está um conjunto de fornecedores de matérias-primas específicas (produtos secos: rolhas, garrafas, rótulos, etc.), equipamentos e serviços que entram no processo numa lógica transversal, interceptando, as componentes produtiva (viticultura), de transformação (vinicultura) e comercial.

3.2. Proximidades: recursos, activos relacionais e coordenação

De forma ilustrativa podemos esquematizar (figura 3) a estrutura organizativa do conjunto das esferas (produtiva, institucional e territorial) que compõem o meio envolvente desta actividade económica. Encontramos uma teia de relações com intensidades distintas traduzidas na partilha de competências a diferentes níveis.

Figura 3 – Relações entre os diferentes agentes da RDVV



Fonte: Oliveira (2006:120)

As interacções entre agentes económicos, directa ou indirectamente, relacionados com este sistema produtivo estão imbuídas de particularidades relacionadas com os domínios material, cognitivo, participativo e coordenativo que moldam configurações distintas em função dos tipos de agentes envolvidos. Descrevemos de seguida estas particularidades à luz dos referenciais teóricos propostos.

Relações entre actores do sistema produtivo

A forma mais evidente e simultaneamente menos sustentável no prazo é encontrada na relação entre produtores individuais e cooperativas e/ou alguns armazenistas vinificadores engarrafadores. Se, por um lado, se consegue obter a forma ideal para resolver o escoamento do produto de milhares de produtores, por outro, verificamos o aproveitamento de matéria-prima a custos residuais permitindo uma concorrência muitas vezes predatória (abaixo dos custos médios de produção de muitos agentes produtivos) de agentes que dominam o mercado em detrimento de muitos que vêm o trabalho de anos ser remunerado de forma precária sem saber a causa e os efeitos de tal lógica capitalista ou de ineficiência empresarial. Contribui para alimentar esta tendência uma elevada assimetria de informação entre fornecedores e clientes, assente numa falta de conhecimento das dinâmicas empresariais e numa passividade evidente destes produtores individuais. Regista-se, por um lado, *activos de participação* reduzidos estimulados por recursos materiais (aqui entendidos como níveis de conhecimento e saberes próximos) desiguais e, por outro, um oportunismo gerador de custos de transacção elevados que dificilmente são minimizados pelos mecanismos de coordenação existentes.

No patamar industrial regista-se uma lógica relacional vertical (inter-empresarial) mais do que horizontal (intra-empresarial). Dito de outra forma, as relações de cooperação, baseadas na complementaridade de valências específicas manifestam-se entre tipos de agentes diferenciados. Com efeito, as relações entre agentes do mesmo tipo são pouco frequentes e assentam, essencialmente, na troca de informação sobre fornecedores e dinâmicas da procura, i.e., apesar de se registar uma proximidade cognitiva (mesmos propósitos, anseios, códigos de linguagem, etc.) evidente não verificamos, nem partilha dos recursos materiais, nem interacções eficientes e profícuas, porque complementares, em activos de rede para troca de informação e partilha de objectivos comuns.

Por outro lado, vários são os exemplos de relações profícuas e sinergéticas entre agentes de tipo diferente. As adegas cooperativas encontram nas uniões de cooperativas um aliado estratégico na comercialização dos seus produtos. Neste caso particular, a VERCOOP[11] assume-se como a união de cooperativas com maior poder comercial da RDVV: *a VERCOOP é para nós o que nós somos para os produtores individuais*[12]. A Adega Cooperativa de Amarante permite a vários produtores individuais registarem e engarrafarem a sua própria marca de vinho nas suas instalações. Mas, se por um lado, se cria uma oportunidade, por outro, alimenta-se uma tendência de atomização dos Vinhos Verdes que em nada contribui para a afirmação de um produto regional no mercado global. Outro exemplo notório de cooperação é encontrado na partilha informal de tecnologia (máquinas de vindimar no caso) somente detida por agentes cuja quota de mercado é representativa (armazenistas) e alguns produtores engarrafadores. Essa partilha é precedida da existência de trocas económicas, nomeadamente da venda de uvas com determinadas características pelos segundos que são geralmente associados aos vinhos de quinta, permitindo um duplo benefício gerador de externalidades positivas provocadas, quer pelo aproveitamento tecnológico, quer pelo abastecimento de matéria-prima de elevada qualidade.

Das relações entre os actores produtivos da fileira do vinho verde com os que compõem a indústria relacionada destaca-se primeiro a distância geográfica/espacial que os separa. Com efeito, os fornecedores de produtos secos (rolhas, garrafas, rótulos e caixas de embalagem) e de equipamentos (cubas, prensas pneumáticas, sistema de frio, linhas de engarrafamento) localizam-se, para os primeiros, em diversos Concelhos como Avintes, Figueira da Foz, S^{ta}

[11] A União de Cooperativas da Região dos Vinhos Verdes é constituída pelas seguintes Adegas Cooperativas: Amarante, Braga, Famalicão, Felgueiras, Guimarães, Paredes e Vale da Cambra.
[12] Eng. António Maria da Adega Cooperativa de Amarante.

Maria da Feira, Aveiro e Lisboa e, para os segundos, em países europeus (Itália e França), mas com representações nacionais situadas essencialmente em Gaia e em Aveiro. Esta falta de proximidade espacial é contornada com mecanismos comuns deste tipo de recursos, nomeadamente contratos de assistência, mas também através de relações informais (troca de informações sobre o sector, fidelização, etc.) para as quais a confiança se assume como principal variável explicativa de fortes activos relacionais e de participação.

Relações entre sistema produtivo e tecido institucional

As relações entre os agentes produtivos e o tecido institucional e de serviços sofisticados de apoio à empresa encontram aqui uma forma não raras vezes sustentada pelo apoio e compromisso de vontades partilhadas entre organismos institucionais e sectoriais. A coerência e eficácia das interacções entre os agentes que constituem o território partem, não só do número de vezes que as relações se estabelecem e da quantidade de agentes envolvidos, mas também da pertinência e do valor acrescentado resultante das mesmas.

É necessário encarar distintamente a actuação dos agentes agrícolas dos restantes actores produtivos, pois os produtores individuais são, como temos visto, um caso muito particular. A relação destes com o tecido institucional é muitas vezes despoletada por um sentido de obrigatoriedade; por um lado, a necessidade de obter subsídios agrícolas (Zonas Agrárias – DRAEDM) e de manifestar o vinho para obter possibilidade de comercializar a sua produção (Zonas Agrárias e Adegas Cooperativas – CVRVV), por outro, e de forma manifestamente insuficiente, através dos encontros promovidos pela EVAG[13] com o intermédio de algumas adegas cooperativas. Com efeito, poucos são os que, por iniciativa própria, participam nas iniciativas promovidas por diversos actores no sentido de fazerem o *upgrading* de conhecimentos necessários ao acompanhamento das actuais práticas de viticultura. É errado pensar-se que, por estarem associados a cooperativas e armazenistas, encetam lógicas de verdadeira partilha de saberes, cumprem objectivos comuns, seguem as mesmas regras. De facto, o que encontramos nos seus comportamentos é um síndrome que pode ser resumido pela expressão *pequeno e só*, reduzindo-se os efeitos de uma proximidade geográfica, histórica e cultural, pela falta de partilha dos mesmos códigos de linguagem, activos de rede e confiança recíproca.

Quadro 1 – Tipo de relações entre agentes produtivos e o tecido institucional

Instituições	Nº de actores	Tipo de relações
Vocacionadas exclusivamente para a fileira		
CVRVV	21	Análises laboratoriais, trocas de informações técnicas e regulamentação do sector do Vinho Verde
EVAG	4	Visitas e compra de porta-enxertos
APEVV	2	Defesa dos direitos dos produtores engarrafadores
Associações de Protecção Integrada (nomeadamente AJAP e APUVE)	4	Apoio técnico especializado ao nível da viticultura e produção integrada
Organismos desconcentrados do Estado		
IVV	3	Certificação de vinhos de mesa
DRAEDM/IFADAP (delegações regionais)	15	Candidaturas de projectos ao AGRO (vinicultura) e VITIS (viticultura)
Organizações sectoriais públicas e privadas		
AEP	1	Formação e feiras
ICEP	2	Projectos de internacionalização
ANCEVE	4	Documentação e nova legislação sobre o sector
IAPMEI	1	Projectos de apoio à actividade produtiva
ACIA	2	Prestação de serviços, feiras, projectos ligados à formação e outras iniciativas para micro e PMEs

[13] A Estação Vitivinícola Amândio Galhano tem centrado, desde 1985, o seu trabalho na área do desenvolvimento experimental, e em particular na componente de viticultura através da produção de material vegetativo com qualidade (nomeadamente porta-enxertos), mas também na componente enológica.

Quadro 1 (cont.)

Instituições	Nº de actores	Tipo de relações
Entidades formadoras e prestadoras de serviços de I&D		
IDARN	1	Relações de alto nível técnico e científico; formação especializada
Associação RECIVERDE	1	Formação
CONFAGRI	1	Formação
PROENOL, Indústria Biotecnológica, Lda.	1	Microrganismos, leveduras, bactérias para o vinho
Instituições de ensino profissional e superior		
Católica (Porto e Minho)	2	Investigação direccionada para a vinha e estudo sobre embalagens
Faculdade de Biotecnologia do Porto	5	Protocolos de colaboração, estudos específicos sobre vinho verde tinto, polém da vinha, previsão de stocks, etc.
Escola de Vitivinicultura de Amarante	3	Visitas às quintas pelos alunos da escola
Associações de Desenvolvimento		
Associação de desenvolvimento RIBADOURO	3	Contabilidade e formação
Associação de desenvolvimento DOLMAN	1	Candidaturas a projectos co-financiados pelo LEADER+
Autarquias		
Felgueiras	2	Promoção de feiras locais
Lousada	1	"
Amarante	1	"

Fonte: Inquéritos.

As relações (medidas em termos de frequência e propósitos) entre actores produtivos e o tecido institucional envolvente são sistematizadas no quadro 1. Verifica-se um leque diversificado de actores estratégicos (associativos e institucionais) de apoio a esta actividade produtiva com domínios de intervenção distintos em função das suas valências e dos seus objectivos. Assim, são classificadas 7 categorias de actores que têm um papel distinto, mas muitas vezes complementar no apoio e fornecimento de serviços a estes agentes. As suas formas de actuação podem ser estruturadas em três dimensões:

i) abrangência territorial das suas acções (nacional, regional ou local);

ii) carácter sectorial ou institucional;

iii) âmbito de intervenção (regulação, apoio empresarial, formação, I&D, etc.).

Numa primeira análise evidencia-se a proximidade geográfica/espacial entre estes actores económicos, quer seja através da sua sede, no caso de agentes sectoriais e alguns institucionais, quer seja através dos centros regionais no caso dos organismos de âmbito nacional. Porém, esta proximidade não se manifesta suficiente no sentido de influenciar de forma ideal as relações entre o tecido institucional e o sistema produtivo. Com efeito, como iremos ver, o grau de intensidade e o âmbito das relações, bem como os níveis de confiança subjacentes, compõem um quadro, não raras vezes, insuficientemente capaz de sustentar a partilha de saberes e de informação sistemática e abrangente.

Encontramos na interacção entre entidades com vocação exclusiva para este sector e o tecido produtivo territorial uma situação privilegiada de proximidade. Para além da proximidade geográfica já descrita, existem condições que compaginam uma lógica profícua de coordenação entre agentes assente numa proximidade *institucional*, pois todos se comportam em função de determinados códigos e regras formais e informais estabelecidos quer pela CVRVV, quer por relações bilaterais entre as instituições e as empresas; bem como numa proximidade organizacional assente numa estrutura em rede (espaço colectivo) e na existência de inter-relações baseadas na concretização de objectivos comuns. No entanto, denota-se um predomínio da mobilização pela mediação, pelos contratos, pelas regras e normas, sobre uma verdadeira lógica inter-organizacional, sendo ténue o feedback recíproco entre acção colectiva e os modos de coordenação presentes.

Outras configurações de interacção são verificadas entre estes actores económicos. De uma forma menos evidente do ponto de vista da compreensão das regras do jogo, pela existência, não raras vezes, de assimetria de informação, são estabelecidas relações de cooperação, duradouras no tempo, entre as empresas e o órgão desconcentrado regional do ministério da agricultura (IFADAP/DRAEDM), desencadeado pela necessidade de aproveitar financiamento associado à reconversão e reestruturação das vinhas por um lado (VITIS) e, por outro, ao apetrechamento e reconversão das estruturas de vinificação das empresas (AGRO). O carácter exclusivo (são maioritariamente os actores de elevada dimensão a beneficiar dos financiamentos) das oportunidades criadas, apesar da partilha das mesmas normas e regras do jogo, só pontualmente é 'compensado' por lógicas de partilha de recursos materiais entre os actores produtivos.

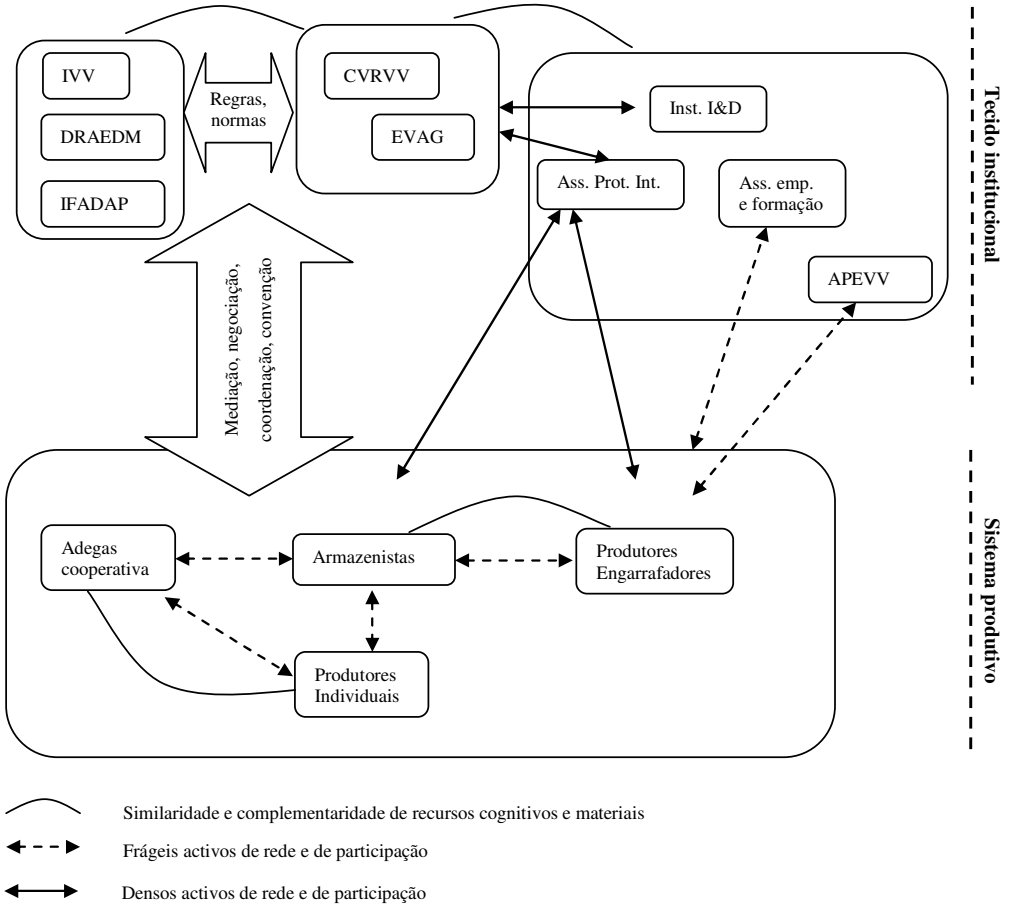
À excepção das associações empresariais, as relações estabelecidas com outros actores institucionais devem essencialmente a sua existência a um critério de proximidade geográfica, fruto de necessidades, muitas vezes esporádicas, de trocas de informação, formação, estudos científicos específicos e apoios ao nível da situação financeira de micro empresas. Não obstante, existe uma consciência (proximidade cognitiva), por parte dos agentes de maior dimensão, da necessidade de recorrer a centros de racionalidade como instituições do ensino superior e à EVAG para tentar solucionar problemas de cariz científico e técnico de elevada complexidade ligada, quer às práticas vitícolas, quer às especificidades enológicas, o que manifesta uma preocupação positiva em busca de um melhor desempenho em termos de I&D. Conjugando estas vontades partilhadas com o estabelecimentos de activos de rede facilitados pela mediação entre as instituições representativas do sector e as instituições de I&D as interacções de retorno positivo serão certamente mais abrangentes e frequentes.

Organização e coordenação territorial

A figura 4 pretende ilustrar as proximidades partilhadas pelos actores económicos enraizados na RDVV que gravitam em torno do recurso vinho verde, demonstrando, nomeadamente, o modo como o sistema de governação medeia, estimula e mobiliza a organização dos actores intervenientes e a função/papel de cada esfera territorial, de cada actor, na construção da sua organização e do seu território.

As dinâmicas de governação estabelecem-se pela coordenação entre as esferas de poder empresarial e institucional. É a este nível que as regras são fixadas e os referenciais são assumidos por todos os que participam neste jogo de construção do território. Não será ao nível das formas institucionais de regulação do sistema que a coordenação poderá não ser eficiente, nem na articulação entre as políticas regionais e externas - os normativos sectoriais (*regras formais*) são negociados e decididos ao nível institucional, a CVRVV formula as regras e as normas juntamente com o IVV e outros órgãos desconcentrados relacionados com a actividade agrícola - mas sim ao nível dos activos organizacionais pois registam-se capacidades diferenciadas de negociação e de poder na tomada de decisões e na construção dos referenciais que estruturam e identificam os caminhos a seguir. A participação dos agentes com menor dimensão crítica apresenta um défice em relação aos que mais influenciam e, consequentemente, aproveitam de forma mais eficiente os benefícios gerados na definição e aplicação das políticas para o sector. Esta assimetria hierarquizada decorre especialmente de défices na similaridade e complementaridade cognitiva e material, assentes na reduzida germinação de confiança, na incerteza na tomada de decisões e em alguns comportamentos oportunistas que travam as relações de cooperação.

Figura 4 – As proximidades presentes entre actores da RDVV



As estruturas de relacionamento entre actores institucionais apresentam um modelo fortalecido pela complementaridade de propósitos e um quadro de referência estratégico comum. É essencialmente entre as entidades reguladoras e os actores pertencentes ao sistema de I&D que se prevêem as acções conjuntas mais vantajosas para o sector. Resta evidenciar os mecanismos formais e informais necessários para uma difusão harmoniosa dos resultados de tais incursões pelos actores da RDVV.

As adegas cooperativas assumem-se, no conjunto dos actores produtivos, como os que mais esforços canalizam para corporizar uma lógica de interacção organizacional e a construção de um espaço cognitivo de aprendizagem colectiva. Por estarem umbilicalmente ligados a diversas estruturas de poder local e de regulação do sector, encetam diversas iniciativas colectivas com o intuito de difundir o conhecimento sobre práticas e tecnologias emergentes para o sector.

Os armazenistas e, principalmente, os produtores engarrafadores apresentam níveis de relacionamento e de cooperação menos evidentes. Os primeiros por levarem a cabo estratégias de desenvolvimento de nível nacional e, portanto, por tecerem relações extra-regionais mais frequentes com alguns organismos prestadores de serviços específicos como centros de I&D e entidades formadoras. Esta característica também acarreta uma relação menos evidente no tocante a algumas instituições locais como associações empresariais ou autarquias[14]. Os segundos por apresentarem um défice de iniciativas de cooperação, quer com actores institucionais indirectamente relacionados com esta actividade, quer com centros de

[14] Esta relação extra-regional é também estimulada pelo facto de alguns armazenistas estarem presentes na região para diversificarem a sua oferta, nestes casos o Vinho Verde aparece como 2ª ou 3ª opção (por exemplo as empresas Borges e Covipor) e, de outras, que têm gradualmente apostado em vinhos maduros localizando unidades de produção noutras regiões com o mesmo objectivo (por exemplo as empresas Quinta da Aveleda e Caves da Cerca). Esta constatação remete-nos para os conceitos de ancoragem vs nomadismo das empresas ao território, mas de forma adaptada pois não se verifica uma deslocalização por parte destas empresas mas sim um relacionamento menos intenso destas a um só território.

racionalidade e competência (centro de I&D, sistema de formação, EVAG), quer mesmo com empresas do mesmo tipo (relacionamento horizontal). Na maioria dos casos, este comportamento está intimamente relacionado com a falta de massa crítica dos empresários (disparidade na proximidade de recursos materiais), da reduzida dimensão e do incipiente desempenho económico destas empresas.

O comportamento dos actores produtivos relativamente ao associativismo não contribui para uma efectiva organização e coordenação territorial. A existência de algumas associações empresariais sectoriais e outras com uma actuação específica como sejam a APEVV[15] e a APUVE ou a AJAP (associações de protecção integrada) não é condição satisfatória de lógicas de cooperação organizacional e institucional eficazes. Por um lado, a APEVV apresenta uma gestão pouco eficiente e uma elevada rotatividade dos seus corpos dirigentes o que contribui para um elevado défice de participação de um universo que unido (institucional e economicamente) poderia contribuir para a construção de uma organização capaz de se afirmar nos mercados mais competitivos. Por outro, ainda que de forma mais eficiente, as associações de protecção integrada poderiam ter um papel mais activo sobretudo no tocante aos produtores individuais que, na sua maioria, não têm conhecimento da sua existência. Quanto às restantes associações com uma actuação sectorial mais abrangente, estas têm essencialmente "servido" os interesses de alguns produtores engarrafadores relativamente aos serviços de contabilidade, formação e informação legislativa sobre o sector, não se evidenciando uma reciprocidades que extravase os objectivos comerciais, nem uma mediação coordenativa das relações.

4. CONCLUSÃO E NOVAS PISTAS DE INVESTIGAÇÃO

Numa perspectiva *interaccionista* a complexidade das dinâmicas territoriais é evidente. Há uma hierarquia de actores e de instituições difusa; há interesses, propósitos e recursos próprios distintos e divergentes; há níveis de participação, envolvimento e apropriação diferentes; há mecanismos de governação territorial mais ou mesmos eficientes; no fundo há incerteza, desequilíbrios, fricções nos comportamentos dos indivíduos, actores e instituições.

Olhar para estas dinâmicas através das interacções engendradas pela acção colectiva permite-nos um entendimento abrangente e integrado da construção dos territórios. O caso analisado apresenta uma miríade de actores que, actuando no mesmo espaço (RDVV), tem distintos (i) recursos materiais e cognitivos, (ii) valores e normas de conduta, (iii) capacidades empreendedoras, e atribuem ao jogo das interacções uma constelação de combinações mais ou menos convergentes.

Assistimos a um jogo de interacções com diferentes propósitos promovidas por actores com posições distintas numa rede, por vezes, frágil (incerteza, oportunismo, reduzido desempenho dos actores), onde fluem informações e se partilham recursos *materiais* e *cognitivos* com base em *activos de participação* mais ou menos intensos. Estes *custos de transacção*, de interacção, ultrapassam em diversos casos as balizas e imposições estabelecidas pelas regras formais sectoriais e territoriais. Os mecanismos de coordenação manifestam-se, portanto, ineficazes face à 'desorganização' alimentada pela falta de *similaridade* e *complementaridade* verificada nalgumas franjas da rede territorial.

Em síntese, confirmamos, à luz do quadro conceptual de partida, que a proximidade geográfica de que partilham os diversos actores que gravitam em torno do vinho verde não é condição *sine qua non* da manifestação de externalidades positivas e de trajectórias em que o resultado do jogo *interaccionista* é superior à soma das partes. Para isso, será então indispensável que os agentes económicos e as instituições se aproximem fundamentalmente através de uma partilha de redes, regras, espaços de pertença, recursos cognitivos e vontades comuns; mediatizadas/coordenadas por estruturas formais e informais de referência.

O conceito de economia de proximidades tem registado uma evolução conceptual relevante, porém ainda não apresenta uma teorização inequívoca. Para além do reforço de reflexão teórica e normativa, este conceito precisa ser "posto à prova" através de uma maior e mais rigorosa

[15] Associação de Produtores Engarrafadores de Vinho Verde.

aplicabilidade empírica, sendo necessário, à luz de referenciais latentes e complementares (nomeadamente teoria dos custos de transacção, teoria dos jogos, teoria das redes), reforçar o acervo de modelos e medidas existentes, passíveis de aplicação a diversos casos de estudo.

Neste contexto, a aplicação do conceito de economia de proximidades para explicar a manifestação de externalidades positivas (pecuniárias, tecnológica e de conhecimento) e dinâmicas de inovação em sistemas produtivos portugueses, nomeadamente de modo a estabelecer uma tipologia, parece-nos da maior pertinência.

REFERÊNCIAS BIBLIOGRÁFICAS

ALESINA A. e FERRARA E. (2002), "Who Trusts Others?", *Journal of Public Economics*, n.º 85, pp. 207-234.

ANSELIN LUC VARGAS ATTILA, ACS ZOLTAN (2000), "Geographic Spillovers and University Research: a spatial econometric perspective, *Growth and Change*, n.º 31, pp. 501-516.

ANTONELLI CRISTIANO (2000), "Collective Knowledge Communication and Innovation: The Evidence of Technological Districts", *Regional Studies*, vol. 34, n.º 6, pp. 535-547.

AUDRETSCH DAVID, FELDMAN MARYANN (1996), "Knowledge Spillovers and the Geography of Innovation and Production", *American Economic Revue*, n.º 86, pp. 630-640.

BECATTINI GIACOMO, (1992), "The Marshallian industrial district as a socio-economic notion", in Becattini, Pyke e Sengenberger (eds.), *Industrial districts and Inter-firm co-operation in Italy*, International Institute for Labour Studies: Geneva, pp. 37-52

BOUBA-OLGA e CARRINCAZEAUX (2008), *« La Proximité, 15 ans déjà ! »*, Revue d'Économie Régionale & Urbaine, n.º 3, ASRDLF.

BOUBA-OLGA OLIVIER e GROSSETTI MICHEL (2008), "Socio-économie de Proximité", *in* Bouba-Olga e Carrincazeaux (2008), *« La Proximité, 15 ans déjà ! »*, Revue d'Économie Régionale & Urbaine, n.º 3, ASRDLF, pp. 311-328.

BOURDIEU, PIERRE (1986), "The Form of Capital" in Richardson, J. (ed.) *Handbook of Theory and Reasearch for the Sociology of Education*, Greenwood Press: New York, pp. 241-258.

BRESCHI STEFANO, LISSONI FRANCESCO (2001), "Localized Knowledge Spillovers Versus Innovative Milieux: knowledge 'tacitness' reconsidered", *Papers in Regional Science*, n.º 80, pp. 255-273.

CALLOIS JEAN-MARC (2004), "Capital social et développement économique local : Pour une application aux espaces ruraux français", *Revue d'économie régionale et urbaine,* n. º4, pp. 551-578.

CANET RAPHAËL (2004), "Qu'est-ce que la Gouvernance ? ", Conférences de la Chaire MCD, disponível em http://www.chaire-cd.ca.

CHALAYE SYLVIE e LARGERON CHRISTINE (2008), "La Veille Scientifique Territorial : messurer les cooperation entre acteurs et territoires", *in* Bouba-Olga e Carrincazeaux (2008), *« La Proximité, 15 ans déjà ! »*, Revue d'Économie Régionale & Urbaine, n.º 3, ASRDLF, pp. 447-466.

COASE RONALD (1960), "The Problem of Social Cost The Problem of Social Cost (this online version excludes some parts)". Journal of Law and Economics n.º 3, pp.1–44

COLEMAN JAMES (1988), "Social Capital in the Creation of Human Capital", *American Journal of Sociology*, n.º 94, pp. S95-S120.

COLLECTIS GABRIEL e RYCHEN FREDERIC (2004), "Entreprises et Territoires: proximités et développement local", *in* Pecqueur Bernard e Zimmermann Jean-Benoît (Dir.) *Economie de Proximités*, Lavoisier: Paris, pp. 207-230.

COMMONS JOHN (1951), *The Economics of Collective Action*, McMillan: London.

COOKE PHILIP (2005), "Regionally Asymmetric Knowledge Capabilities and Open Innovation: exploring 'Globalization 2' – a new model industry organization", *Research Policy*, n.º 34, pp. 1128-1149.

COOKE PHILIP e MORGAN KEVIN (1998), *The Associational Economy*, Oxford University Press: New York.

COURSON JACQUES (1990), "Technopole et projet urbain : l'expérience internationale", *in* Jalabert e Thouzellier (dirs) *Villes et technopoles, nouvelle urbanisation, nouvelle industrialisation*, Les Presses de l'Université du Mirail: Toulouse.

EDQUIST CHARLES (Ed.) (1997), *Systems of Innovation: Technologies, Institutions and Organizations*, Routledge: London.

GILLY JEAN-PIERRE e TORRE ANDRE (Dir.) (2000), *Dynamique de Proximité*, L´Harmattan: Paris.

GILLY JEAN-PIERRE LEROUX ISABELLE e WALLET FREDERIC (2004), "Gouvernance et Proximité", *in* Pecqueur Bernard e Zimmermann Jean-Benoît (Dir.) *Economie de Proximités*, Lavoisier: Paris, pp 187-206.

GRANDORI ANNA (1997), "Governance Structures, Coordination Mechanisms and Cognitive Models", *The Journal of Management and Governance*, vol. 1, pp. 29–47.

HODGSON GEOFFREY (2006), "What Are Institutions?", *Journal of Economic Issue*, Vol. XL, n.º 1, pp. 1-25.

KNACK STEPHEN e KEEFER PHILIP (1997), "Does Social Capital have an Economic Payoff? A Cross-country Investigation", *The Quarterly Journal of Economics*, Vol. 112, n.º 4, pp. 1252-1288.

LANDRY REJEAN *et al.* (2001), "Capital Social, Innovation et Politiques Publique", *ISUMA*, Université du Québec, pp. 64-70.

MARSHALL ALFRED (1890), *Principles of Economics*, Macmillan and Co., Ltd: London.

MAYOUKOU CELESTIN *et al.* (2003), *Gouvernance du Développement Local*, l'Harmattan: Paris.

MENDEZ ARIEL, MERCIER DELPHINE (2006), "Compétences-clés de territoires: Le rôle des relations interorganisationnelles", *Revue Française de Gestion*, pp. 253-276.

NORTH DOUGLAS (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge University Press: Cambrigde.

OLIVEIRA MANUEL (2006), *Dos Agentes aos Sistemas Produtivos Territorializados: o vinho verde como elemento congregador*, Dissertação de Mestrado apresentada na Universidade da Beira Interior: Covilhã.

PECQUEUR BERNARD e ZIMMERMANN JEAN-BENOIT (2002), "Les Fondements d'une Économie de Proximités", *working paper* n.º02A26, GREQAM, Universités d'Aix-Marseille.

PECQUEUR BERNARD e ZIMMERMANN JEAN-BENOIT (Dir.) (2004) *Economie de Proximités*, Lavoisier: Paris.

PORTER MICHAEL (1990), The Competitive Advantage of Nations, Free Press: New York.

PORTER MICHAEL (2000), "Location, Competition, and Economic Development: Local Clusters in a Global Economy", *Economic Development Quarterly*, Vol. 14, n.º 1, pp. 15-34.

PUTNAM ROBERT (2000), *Bowling Alone*, Simon and Schuster: New York.

PUTNAM ROBERT *et al.* (1993), *Making Democracy Work: Civic Traditions in Modern Italy,* Princeton University Press: Princeton.

RALLET ALAIN e TORRE ANDRE (Eds.) (2007), *Quelles Proximités pour Innover?*, L'Harmattan: Paris.

REIS JOSÉ (2007), *Ensaios de Economia Impura*, Almedina: Coimbra.

TALBOT DAMIEN (2008), "Les Institutions Créatrices de Proximités", *in* Bouba-Olga e Carrincazeaux (2008), *« La Proximité, 15 ans déjà ! »*, Revue d'Économie Régionale & Urbaine, n.º 3, ASRDLF, pp. 289-310.

TORRE ANDRE (2006), "Cluster et Système Locaux d'Innovation: retour critique sur les hypothèses naturalistes de transmission des connaissances à l'aide des catégories de l'économie de la proximité", *in* Le Bas (coord.), *Innovation, Région et Connaissance*, L´harmattan: Paris, pp.15-43.

WILLIAMSON OLIVIER (1985), *The Economic Institutions of Capitalism: Market, Hierarchies and Relational contracting*, Free Press: New York.

ZIMMERMANN JEAN-BENOIT (1995), "Le Concept de Grappes Technologiques", *Revue Economique,* Vol. 46, n.º 5, pp. 1263-1294.

INCIDENCIA DE LA CONTRATACION INDEFINIDA SEGUN NACIONALIDAD EN ANDALUCÍA Y EN EL RESTO DE ESPAÑA*

ANTONIO CAPARRÓS RUIZ
e-mail: antonio@uma.es

Mª LUCÍA NAVARRO GÓMEZ
e-mail: L_navarro@uma.es

Departamento de Estadística y Econometría
Universidad de Málaga

Resumen

Desde 1997, las políticas activas implementadas en nuestro mercado de trabajo para mejorar la estabilidad en el empleo se han incrementado notablemente. Así, como resultado de las reformas laborales realizadas desde esa fecha, el conjunto de relaciones laborales indefinidas existentes en nuestro país se puede fragmentar en tres grandes grupos: contratos indefinidos ordinarios, contratos indefinidos bonificados procedentes de la transformación de contratos temporales y contratos indefinidos iniciales bonificados. El objetivo principal de este trabajo es analizar la incidencia de los diversos tipos de contratación indefinida sobre la población inmigrante en comparación con la nacional. Para este fin, la metodología utilizada consiste en la estimación de modelos *logit multinomial* donde la variable dependiente muestra si el trabajador está situado en alguna de las anteriores modalidades de contratación o tiene un contrato temporal, y como variables explicativas se introducen la nacionalidad del trabajador, además del resto de características personales y laborales del individuo. Este análisis se realiza separadamente para España y para Andalucía, y distinguiendo a los trabajadores según su experiencia laboral. La información estadística utilizada para efectuar este trabajo proviene de la Muestra Continua de Vidas Laborales (MTAS 2006).

Palabras clave: Reformas laborales, contratación indefinida bonificada, inmigración
Área temática: Economía Urbana, Regional y Local.

1. INTRODUCCIÓN

El fomento de la contratación indefinida, a través de bonificaciones sobre la cuota empresarial a la Seguridad Social, ha sido uno de los principales instrumentos de política activa de empleo implementado en España desde la reforma laboral de 1997. Este impulso de la contratación indefinida, desde un punto de vista legislativo, ha coincidido temporalmente con el enorme incremento de los flujos migratorios de extranjeros hacia España. La población inmigrante ha pasado de representar el 1,4% sobre el total en 1996 hasta el 8,4% en el 2006, siendo la proyección realizada por el INE para el 2010 superior al 11%.

En este trabajo se pretende modelizar la incidencia de la contratación indefinida (bonificada o no), considerando el país de origen de los trabajadores y diferenciándolos según residan en Andalucía o en el resto de España. Este último tema es de relevante interés económico, ya que los efectos positivos de la inmigración sobre la economía están relacionados directamente con su grado de asimilación en el mercado de trabajo, y en definitiva con la estabilidad y calidad de sus empleos. El hecho de diferenciar geográficamente a Andalucía del resto de Comunidades se debe a que dicha Comunidad es la que presenta una mayor tasa de temporalidad (un 45%).

* Este trabajo se ha realizado en el marco del proyecto SEJ2007-68045-CO2-01/ECON del Ministerio de Educación y Ciencia.

Para llevar a cabo estos objetivos, la metodología utilizada consiste en estimar un modelo *logit multinomial* que determine los factores que influyen sobre la probabilidad de que un trabajador tenga alguno de los siguientes tipos de contrato: temporal, indefinido ordinario (no bonificado), indefinido bonificado procedente de la transformación de un contrato temporal e indefinido bonificado desde el inicio de la relación laboral. Dentro de las variables explicativas se considera la nacionalidad del trabajador y, además, para valorar en mayor medida el proceso de asimilación de la población inmigrante, se realiza el estudio dividiendo a la población asalariada según su experiencia laboral.

El estudio de la asimilación laboral de los inmigrantes a la población nativa del país de destino ha sido un tema tratado extensamente por la literatura económica internacional. En este sentido, cabe señalar los trabajos de Chiswick (1978) y Borjas (1985, 1995) que analizan la asimilación salarial de los extranjeros en EEUU. En el contexto español existen escasos trabajos que hayan abordado la asimilación de la población inmigrante, siendo los más relevantes los realizados por Fernández y Ortega (2006) y Amuedo-Dorantes y De la Rica (2007). En el primero, con datos de la EPA (INE, 1996-2005), se contrastan las diferencias entre nativos y extranjeros con respecto a la probabilidad de participar en el mercado laboral y de encontrar un empleo. Sus resultados muestran que en comparación con los nativos, los trabajadores inmigrantes tienen niveles de participación laboral y de desempleo más altos, aunque la diferencia con los españoles prácticamente se anula a partir de los cinco años. Amuedo-Dorantes y De la Rica (2007), con datos del Censo de Población (INE, 2001) y de la Encuesta de Estructura Salarial (INE, 2002), examinan si las diferencias de probabilidad de encontrar un empleo entre inmigrantes y nativos, y la asimilación ocupacional entre ambos colectivos son funciones crecientes con el tiempo de permanencia en España de los inmigrantes. Sus resultados son, por un lado, que los inmigrantes varones y mujeres, con menos de un año de residencia, experimentan una probabilidad de estar ocupado inferior en 14 y 7 puntos porcentuales, respectivamente, con respecto a los nativos. Estos diferenciales prácticamente desaparecen en un periodo de dos años para los varones y en uno para las mujeres. Por otro lado, también obtienen que los trabajadores españoles realizan tareas más remuneradas que los inmigrantes, incluso cinco años después de la llegada de éstos a España, siendo las diferencias salariales más amplias entre los colectivos de mujeres nacionales y extranjeras.

La aportación que realizamos en este trabajo respecto a los estudios anteriores consiste en abordar de forma más detallada los determinantes de la contratación indefinida en relación con la temporal, en especial, diferenciando la contratación indefinida ordinaria de los dos tipos existentes de contratación indefinida bonificada (desde el inicio de la relación laboral o por transformación de contrato temporal). Esta distinción de contratos permite, además, observar el impacto de las sucesivas reformas efectuadas en el mercado laboral para conseguir la estabilidad en el empleo de la población trabajadora. La única información estadística que permite realizar este análisis en la actualidad, y que se utiliza aquí, es la Muestra Continua de Vidas Laborales (MCVL, MTAS 2006).

La organización del trabajo es la siguiente. En el segundo epígrafe se presentan los datos, ofreciéndose una descripción estadística de la distribución de los trabajadores por tipo de contratos según su nacionalidad y demás variables explicativas usadas en las estimaciones de la probabilidad de tener un determinado contrato. En el siguiente se expone la metodología econométrica utilizada para obtener esa probabilidad, así como los resultados de las estimaciones de los modelos empleados. En la cuarta sección se proporciona un resumen de las principales conclusiones.

2. DATOS Y VARIABLES

Para seleccionar los datos usados en este trabajo, se observan las relaciones laborales iniciadas por los desempleados a partir del año 2000 y que han perdurado hasta el 2006, clasificándose éstas en alguna de las cuatro modalidades contractuales señaladas anteriormente. Hay que indicar que en el año 2006 pueden existir contratos indefinidos ordinarios que procedan

de la transformación de contratos indefinidos bonificados[1]. Ahora bien, como uno de los fines de este trabajo es valorar el éxito de las distintas reformas laborales, al estimar los modelos *logit multinomial*, esos contratos se han considerado dentro de algunas de las dos clases de contratos bonificados. Esta es la razón por la que el periodo de observación se inicia en el año 2000, ya que es en esta fecha cuando la MCVL recoge las variaciones contractuales ocurridas en el empleo.

Las tablas 1 y 2 proporcionan, respectivamente para Andalucía y para el resto de España, el análisis descriptivo de las variables explicativas utilizadas para estimar la probabilidad de tener una determinada modalidad contractual, según el tipo de contrato que tienen los trabajadores y su experiencia previa al empleo. Estas variables explicativas son: nacionalidad[2], edad, sexo, características del empleo (actividad del establecimiento y categoría profesional), número de empleos temporales anteriores como un indicador de la trayectoria laboral de los individuos[3], y Comunidad autónoma de residencia en las muestras correspondientes al resto de España.

En dichas tablas se observa, en primer lugar, que los inmigrantes, al contrario que los trabajadores nacionales, tienen mayoritariamente contratos precarios, independientemente de su región de procedencia y de su experiencia previa. En concreto, los trabajadores originarios de Africa tienen una representación dentro de los temporales que al menos triplica a la observada dentro de los contratos indefinidos ordinarios. Como excepción significativa cabe señalar que, en Andalucía, los trabajadores procedentes de la UE con más de tres años de experiencia laboral previa al empleo tienen una representación dentro del total de contratos indefinidos ordinarios superior a la registrada en los temporales (un 0,8% frente a un 0,5%). Respecto a las otras características personales se constata, en segundo lugar, que la edad media alcanza su valor mínimo para los trabajadores con contratos bonificados por transformación de contratos temporales en indefinidos, seguido por los trabajadores temporales (respectivamente, 29 años frente a 31 para los trabajadores con 3 o menos años de experiencia, y 32 años frente a 35 para el resto, tanto para Andalucía como para España). En tercer lugar, los trabajadores varones están más representados que las mujeres en todas las modalidades laborales, si se exceptúan los contratos indefinidos bonificados iniciales donde ocurre lo contrario. En este último caso, las mayores diferencias se registran para los individuos con 3 o menos años de experiencia donde, tanto para Andalucía como para el resto de España, las mujeres son al menos del 65%.

En referencia a la estructura productiva[4], se observa que en Andalucía la actividad que registra un mayor porcentaje tanto para los contratos indefinidos ordinarios como para los bonificados es el comercio, superando el 30% en el último caso. En cambio para los contratos temporales y para los trabajadores menos experimentados, el sector que predomina es la industria con un 29%, mientras que para los asalariados con más experiencia es la construcción con un 39%. En el resto de España, se constatan diferencias respecto al caso andaluz. Por un lado, ahora el comercio es el sector preponderante sólo para los contratos indefinidos bonificados (aunque no supera el 30%), mientras que la actividad más representada en los contratos indefinidos ordinarios es la industria, con un 25% para los trabajadores con menos experiencia y un 30% para el resto. Por otro lado, el sector que más peso tiene dentro de los empleos temporales es el de la construcción, para los dos grupos de trabajadores.

[1] Esto ocurrirá cuando el periodo de bonificación se extinga y el empresario no despida al trabajador. A partir de ese momento, la relación laboral indefinida no estará bonificada.

[2] La clasificación de los trabajadores según su nacionalidad es la siguiente: españoles (nacidos bien en España, bien en el extranjero o que tienen doble nacionalidad), nacionales de la UE y de EEUU, inmigrantes procedentes de países europeos no integrados en la UE, africanos, sudamericanos y asiáticos.

[3] Esta variable, medida desde el comienzo de la vida laboral de los individuos, es una aproximación del verdadero número de empleos temporales que han tenido, ya que por falta de información en la MCVL, no es posible considerar los contratos temporales anteriores a 1991.

[4] Hay que señalar, por un lado, que en la muestra utilizada no aparecen datos para los trabajadores asalariados en la agricultura. Por otro lado, no se ha incluido el sector de la Administración Pública, ya que las bonificaciones de los contratos indefinidos van dirigidas a los trabajadores del sector privado.

Tabla 1. Análisis descriptivo de los asalariados en Andalucía según tipo de contrato

Variables	3 o menos años de experiencia previa al empleo				Más de 3 años de experiencia previa al empleo			
	Indefinidos ordinarios	Indefinidos bonificados por transf.	Indefinidos bonificados iniciales	Temporales	Indefinidos ordinarios	Indefinidos bonificados por transf.	Indefinidos bonificados iniciales	Temporales
	Media (Dev. St)	Media (Dev. St)	Media (Dev. St)	Media (Dev. St)	Media (Dev. St)	Media (Dev. St)	Media (Dev. St)	Media (Dev. St)
Nacionalidad								
España	0,957 (0,20)	0,940 (0,24)	0,962 (019)	0,889(031)	0,986 (0,11)	0,980 (0,14)	0,978 (0,14)	0,979 (0,14)
UE y EEUU	0,017 (0,12)	0,024 (0,15)	0,023 (0,14)	0,038 (0,19)	0,008 (0,09)	0,010 (0,10)	0,012 (0,11)	0,005 (0,07)
Resto de Europa	0,002 (0,05)	0,002 (0,05)	0,002 (0,05)	0,007 (0,08)	0,001 (0,02)	0,001 (0,04)	0,001 (0,04)	0,001 (0,04)
Sudamérica	0,012 (0,10)	0,023 (0,15)	0,011 (0,13)	0,036 (0,18)	0,001 (0,02)	0,003 (0,06)	0,003 (0,06)	0,005 (0,07)
África	0,006 (0,07)	0,008 (0,09)	0,002 (0,08)	0,027 (0,16)	0,002 (0,04)	0,002 (0,05)	0,002 (0,04)	0,006 (0,08)
Asia	0,006 (0,07)	0,003 (0,05)	0	0,003 (0,05)	0,002 (0,04)	0,001 (0,04)	0,001 (0,03)	0,001 (0,04)
Edad	34,37(9,98)	28,57 (7,57)	33,7 (1,14)	31,42(10,89)	38,36 (9,00)	32,32 (7,65)	39,56(11,3)	35,65 (9,57)
Sexo								
Varón	0,652 (0,47)	0,564(0,49)	0,323 (0,46)	0,600 (0,48)	0,703 (0,45)	0,668 (0,47)	0,414 (0,49)	0,742 (0,43)
Mujer	0,348 (0,47)	0,436 (0,49)	0,677 (0,50)	0,400 (0,40)	0,297 (0,45)	0,332 (0,47)	0,586 (0,49)	0,258 (0,43)
Sector Actividad								
Industria	0,214 (0,40)	0,175 (0,40)	0,139 (0,32)	0,290 (0,45)	0,219 (0,40)	0,159 (0,12)	0,141 (0,33)	0,165 (0,13)
Construcción	0,058 (0,23)	0,080 (0,27)	0,044 (0,47)	0,215 (0,41)	0,053 (0,23)	0,098 (0,29)	0,060 (0,24)	0,390 (0,48)
Comercio	0,266 ((0,44)	0,319 (0,46)	0,34 (0,37)	0,134 (0,34)	0,240 (0,42)	0,312 (0,46)	0,307 (0,46)	0,103(0,30)
Hostelería	0,074 (0,26)	0,069 (0,25)	0,080 (0,37)	0,043 (0,20)	0,069 (0,25)	0,068 (0,25)	0,084 (0,27)	0,030 (0,17)
Transporte	0,105 (0,30)	0,083 (0,27)	0,062 (0,24)	0,036 (0,18)	0,140 (0,34)	0,109 (0,31)	0,081 (0,27)	0,054 (0,22)
Activ. Inmobiliarias	0,123 (0,32)	0,148 (0,35)	0,169 (0,37)	0,137 (0,34)	0,147 (0,35)	0,132 (0,34)	0,152 (0,36)	0,120 (0,32)
Educación	0,039 (0,19)	0,022 (0,15)	0,029 (0,17)	0,040 (0,20)	0,029 (0,16)	0,022 (0,15)	0,035 (0,19)	0,036 (0,18)
Sanidad	0,065 (0,19)	0,043 (0,20)	0,066 (0,24)	0,054 (0,22)	0,041 (0,19)	0,041 (0,20)	0,064 (0,25)	0,061 (0,24)
Otras actividades	0,056 (0,23)	0,061 (0,23)	0,071 (0,25)	0,051 (0,22)	0,062 (0,24)	0,059 (0,23)	0,076 (0,27)	0,041 (0,19)
Categoría profesional								
Ingenieros-Licenciados	0,072 (0,25)	0,051 (0,22)	0,037 (0,19)	0,053 (0,22)	0,064 (0,24)	0,031 (0,17)	0,045 (0,20)	0,043(0,19)
Ingenieros-Técnicos	0,048 (0,21)	0,048 (0,21)	0,025 (0,15)	0,042 (0,20)	0,052 (0,22)	0,034 (0,18)	0,036 (0,18)	0,045 (0,19)
Jefes Admin-Taller	0,058 (0,23)	0,022 (0,15)	0,032 (0,19)	0,013 (0,11)	0,067 (0,25)	0,029 (0,17)	0,036 (0,18)	0,019 (0,13)
Ayudantes no titulados	0,054 (0,22)	0,033 (0,18)	0,027(0,16)	0,015 (0,11)	0,062 (0,24)	0,025 (0,15)	0,023 (0,15)	0,018 (0,13)
Oficiales Administrativ.	0,156 (0,35)	0,138 (0,34)	0,175 (0,38)	0,041 (0,36)	0,178 (0,37)	0,158 (0,36)	0,154 (0,20)	0,051 (0,21)
Subalternos	0,030 (0,17)	0,029 (0,16)	0,030 (0,19)	0,026 (0,15)	0,027 (0,16)	0,039 (0,19)	0,030 (0,18)	0,035 (0,18)
Auxiliares Administrat.	0,138 (0,34)	0,187 (0,39)	0,241 (0,42)	0,115 (0,32)	0,100 (0,31)	0,155 (0,36)	0,231 (0,42)	0,090 (0,28)
Oficiales 1ª y 2ª	0,206 (0,40)	0,202 (0,40)	0,179 (0,38)	0,157 (0,36)	0,249 (0,43)	0,269 (0,44)	0,232 (0,42)	0,350 (0,47)
Oficiales 3ª	0,088 (0,28)	0,134 (0,34)	0,100 (0,30)	0,080 (0,27)	0,086 (0,28)	0,118 (0,32)	0,092 (0,28)	0,081 (0,27)
Peones	0,152 (0,35)	0,152 (0,36)	0,154 (0,35)	0,458 (0,49)	0,115 (0,32)	0,142 (0,34)	0,121 (0,33)	0,268 (0,44)
Nª empleos temporal.	0,72 (2,1)	1,30 (3,48)	0,91 (1,50)	2,42 (3,02)	1,16 (4,14)	3,09 (4,36)	3,04 (7,15)	5,31 (6,17)
Nª Observaciones	5305	3605	1668	6260	6510	2304	1083	7769

Fuente: Elaboración propia a partir de la Muestra Continua de Vidas Laborales (MTAS 2006)

Tabla 2. Análisis descriptivo de los asalariados en resto de España según tipo de contrato

Variables	3 o menos años de experiencia previa al empleo				Más de 3 años de experiencia previa al empleo			
	Indefinidos ordinarios	Indefinidos bonificados por transf.	Indefinidos bonificados iniciales	Temporales	Indefinidos ordinarios	Indefinidos bonificados por transf.	Indefinidos bonificados iniciales	Temporales
	Media (Dev. St)	Media (Dev. St)	Media (Dev. St)	Media (Dev. St)	Media (Dev. St)	Media (Dev. St)	Media (Dev. St)	Media (Dev. St)
Nacionalidad								
España	0,901 (0,29)	0,940 (0,24)	0,936 (0,24)	0,744 (0,44)	0,978 (0,14)	0,980 (0,14)	0,978 (0,14)	0,940 (0,24)
UE y EEUU	0,033 (0,18)	0,024 (0,15)	0,020 (0,14)	0,073 (0,26)	0,009 (0,09)	0,010 (0,10)	0,012 (0,11)	0,013 (0,11)
Resto de Europa	0,003 (0,06)	0,002 (0,05)	0,009 (0,09)	0,009 (0,09)	0,001 (0,02)	0,001 (0,04)	0,001 (0,04)	0,001 (0,04)
Sudamérica	0,040 (0,19)	0,023 (0,15)	0,030 (0,18)	0,114 (0,32)	0,006 (0,08)	0,003 (0,06)	0,003 (0,06)	0,021 (0,14)
África	0,015 (0,10)	0,008 (0,09)	0,004 (0,08)	0,046 (0,21)	0,004 (0,06)	0,002 (0,05)	0,002 (0,04)	0,020 (0,14)
Asia	0,009 (0,09)	0,003 (0,05)	0,001 (0,04)	0,014 (0,10)	0,002 (0,05)	0,001 (0,04)	0,001 (0,03)	0,002 (0,05)
Edad	34,26 (9,76)	28,92 (8,3)	36,4 (12,66)	31,42(10,89)	38,66 (9,24)	32,49 (8,02)	39,56(11,3)	35,47 (9,64)
Sexo								
Varón	0,599 (0,49)	0,566 (0,49)	0,350 (0,47)	0,584 (0,49)	0,645 (0,48)	0,604 (0,49)	0,414 (0,49)	0,645 (0,47)
Mujer	0,401 (0,49)	0,434 (0,49)	0,650 (0,47)	0,416 (0,49)	0,355 (0,48)	0,396 (0,49)	0,586 (0,49)	0,355 (0,47)
Sector Actividad								
Industria	0,251 (0,40)	0,247 (0,40)	0,202 (0,40)	0,165 (0,45)	0,299 (0,45)	0,235 (0,35)	0,199 (0,33)	0,138 (0,30)
Construcción	0,065 (0,24)	0,112 (0,31)	0,064 (0,34)	0,240 (0,42)	0,055 (0,23)	0,107 (0,30)	0,091 (0,28)	0,314 (0,46)
Comercio	0,222 ((0,41)	0,263 (0,44)	0,276 (0,44)	0,123 (0,32)	0,188 (0,39)	0,258 (0,46)	0,256 (0,43)	0,093 (0,29)
Hostelería	0,068 (0,25)	0,070 (0,25)	0,080 (0,27)	0,050 (0,20)	0,043 (0,20)	0,062 (0,24)	0,071 (0,25)	0,032 (0,17)
Transporte	0,122 (0,32)	0,069 (0,25)	0,066 (0,25)	0,037 (0,18)	0,126 (0,33)	0,095 (0,29)	0,078 (0,27)	0,051 (0,22)
Activ. Inmobiliarias	0,152 (0,35)	0,135 (0,34)	0,169 (0,37)	0,196 (0,34)	0,159 (0,36)	0,135 (0,34)	0,162 (0,36)	0,160 (0,37)
Educación	0,036 (0,18)	0,016 (0,12)	0,027 (0,16)	0,060 (0,24)	0,034 (0,18)	0,017 (0,13)	0,028 (0,16)	0,059 (0,23)
Sanidad	0,037 (0,18)	0,039 (0,19)	0,047 (0,21)	0,073 (0,26)	0,049 (0,21)	0,043 (0,20)	0,056 (0,23)	0,108 (0,21)
Otras actividades	0,047 (0,21)	0,049 (0,21)	0,069 (0,25)	0,056 (0,23)	0,047 (0,21)	0,048 (0,21)	0,059 (0,24)	0,045 (0,20)
Categoría profesional								
Ingenieros-Licenciados	0,104 (0,30)	0,051 (0,22)	0,045 (0,20)	0,080 (0,27)	0,095 (0,29)	0,030 (0,17)	0,042 (0,20)	0,059 (0,23)
Ingenieros-Técnicos	0,051 (0,21)	0,049 (0,20)	0,031 (0,17)	0,050 (0,20)	0,058 (0,23)	0,033 (0,18)	0,030 (0,17)	0,068 (0,25)
Jefes Admin-Taller	0,064 (0,24)	0,024 (0,15)	0,037 (0,19)	0,015 (0,12)	0,075 (0,26)	0,030 (0,17)	0,049 (0,21)	0,019 (0,13)
Ayudantes no titulados	0,046 (0,20)	0,028 (0,16)	0,033 (0,17)	0,018 (0,14)	0,055 (0,22)	0,028 (0,15)	0,030 (0,17)	0,022 (0,15)
Oficiales Administrativos	0,144 (0,35)	0,125 (0,33)	0,162 (0,36)	0,055 (0,22)	0,160 (0,36)	0,148 (0,36)	0,165 (0,37)	0,064 (0,25)
Subalternos	0,038 (0,19)	0,028 (0,16)	0,037 (0,19)	0,037 (0,17)	0,039 (0,19)	0,033 (0,17)	0,042 (0,20)	0,037 (0,19)
Auxiliares Administrativos	0,123 (0,33)	0,164 (0,37)	0,203 (0,40)	0,124 (0,32)	0,092 (0,29)	0,159 (0,36)	0,189 (0,39)	0,120 (0,32)
Oficiales 1ª y 2ª	0,184 (0,38)	0,202 (0,40)	0,185 (0,38)	0,166 (0,37)	0,223 (0,41)	0,257 (0,44)	0,236 (0,42)	0,290 (0,45)
Oficiales 3ª	0,117 (0,32)	0,169 (0,37)	0,135 (0,34)	0,133 (0,34)	0,118 (0,28)	0,157 (0,36)	0,112 (0,28)	0,134 (0,34)
Peones	0,124 (0,33)	0,160 (0,36)	0,132 (0,34)	0,322 (0,49)	0,079 (0,27)	0,125 (0,33)	0,102 (0,30)	0,187 (0,38)
Nª de empleos temporales	0,70 (4,95)	1,09 (3,31)	0,84 (3,43)	2,09 (4,83)	1,16 (4,14)	2,50 (5,82)	2,52 (8,52)	4,81 (10,25)
CCAA								
Cataluña	0,241 (0,42)	0,245 (0,43)	0,221 (0,42)	0,167 (0,37)	0,264 (0,44)	0,264 (0,44)	0,246 (0,43)	0,166 (0,37)
País Vasco	0,052 (0,21)	0,057 (0,23)	0,075 (0,25)	0,063 (0,24)	0,058 (0,23)	0,050 (0,23)	0,070 (0,25)	0,067 (0,25)
Castilla-La Mancha	0,041 (0,19)	0,041 (0,19)	0,045 (0,19)	0,058 (0,23)	0,039 (0,19)	0,039 (0,19)	0,029 (0,17)	0,055 (0,22)
Valencia	0,107 (0,30)	0,136 (0,34)	0,135 (0,34)	0,119 (0,32)	0,134 (0,34)	0,136 (0,34)	0,137 (0,34)	0,121 (0,32)
Galicia	0,048 (0,21)	0,065 (0,24)	0,075 (0,25)	0,074 (0,26)	0,055 (0,22)	0,057 (0,23)	0,064 (0,24)	0,075 (0,26)
Castilla-León	0,055 (0,22)	0,050 (0,21)	0,055 (0,22)	0,057 (0,23)	0,051 (0,22)	0,051 (0,22)	0,050 (0,21)	0,063 (0,24)
Extremadura	0,017 (0,13)	0,012 (0,11)	0,013 (0,11)	0,026 (0,16)	0,011 (0,10)	0,011 (0,11)	0,012 (0,11)	0,028 (0,16)
Aragón	0,034 (0,17)	0,037 (0,18)	0,030 (0,18)	0,038 (0,19)	0,036 (0,19)	0,036 (0,19)	0,036 (0,19)	0,034 (0,18)
Canarias	0,054 (0,22)	0,044 (0,20)	0,040 (0,20)	0,049 (0,22)	0,047 (0,21)	0,047 (0,21)	0,036 (0,19)	0,062 (0,24)
Madrid	0,236 (0,42)	0,189 (0,39)	0,200 (0,40)	0,201 (0,40)	0,181 (0,38)	0,176 (0,33)	0,190 (0,39)	0,178 (0,38)
Baleares	0,025 (0,15)	0,023 (0,14)	0,020 (0,14)	0,022 (0,15)	0,025 (0,15)	0,021 (0,14)	0,029 (0,15)	0,026 (0,16)
Murcia	0,028 (0,16)	0,042 (0,19)	0,030 (0,16)	0,046 (0,21)	0,039 (0,19)	0,039 (0,13)	0,027 (0,16)	0,041 (0,20)
Navarra	0,020 (0,14)	0,019 (0,13)	0,018 (0,13)	0,024 (0,13)	0,018 (0,13)	0,018 (0,13)	0,028 (0,14)	0,019 (0,14)
La Rioja	0,009 (0,09)	0,006 (0,07)	0,008 (0,09)	0,013 (0,08)	0,005 (0,06)	0,004 (0,06)	0,010 (0,10)	0,009 (0,09)
Asturias	0,020 (0,14)	0,020 (0,14)	0,023 (0,15)	0,030 (0,17)	0,018 (0,13)	0,018 (0,13)	0,029 (0,16)	0,034 (0,17)
Cantabria	0,013 (0,11)	0,014 (0,12)	0,012 (0,11)	0,014 (0,12)	0,015 (0,10)	0,011 (0,10)	0,019 (0,12)	0,018 (0,13)
Nª Observaciones	37084	28483	12574	23129	55692	20134	8970	32822

Fuente: Elaboración propia a partir de la Muestra Continua de Vidas Laborales (MTAS 2006)

En cuanto a la categoría profesional del trabajador dentro de la empresa, cabe señalar que la ocupación dominante para los trabajadores temporales es la de peón, en el grupo de tres o menos años de experiencia (llegando a alcanzar el 46% de los contratos en Andalucía y el 32% en el resto de España). En cambio para los trabajadores temporales más experimentados, y también para el resto de modalidades contractuales, la categoría profesional predominante es la de oficial de 1ª o de 2ª. Por otra parte, se constata que el porcentaje que representan en términos conjuntos las ocupaciones más cualificadas (ingeniero-licenciado, ingeniero-técnico y jefe administrativo o de taller) es superior en los contratos indefinidos ordinarios que en las modalidades indefinidas bonificadas.

Por lo que respecta al número de contratos temporales previos que tuvo el individuo, como cabía esperar, los valores medios más bajos corresponden a los trabajadores indefinidos ordinarios y los más altos a los temporales. En este último caso, hay que destacar que para los asalariados más experimentados, el número medio de relaciones laborales temporales se sitúa en torno a 5, tanto para Andalucía como para el resto de España.

Por último, en relación a las Comunidades Autónomas que componen el resto de España, son Cataluña y Madrid las más representadas en todas las situaciones laborales, aunque el mayor porcentaje en términos conjuntos lo alcanzan en los contratos indefinidos ordinarios (un 48% para los trabajadores con tres o menos años de experiencia y un 44% para el resto). Como resultado de interés cabe destacar que el País Vasco, Valencia y Galicia son Comunidades con un peso relativo dentro de las modalidades de contratación indefinida bonificada superior al que registran con respecto al total de contratos indefinidos ordinarios.

3. RESULTADOS DE LAS ESTIMACIONES

La especificación econométrica corresponde a un modelo *logit multinomial* con cuatro posibilidades: contratos temporales (Y=0); contratos indefinidos ordinarios (Y=1); contratos indefinidos bonificados por transformación de un contrato temporal (Y=2) y contratos indefinidos bonificados iniciales (Y=3). En este modelo la probabilidad de la alternativa j es:

$$P(Y_i = j) = \frac{\exp(\beta_j' X_i)}{1 + \sum_{j=1}^{3} \exp(\beta_j' X_i)}, \quad (j = 1, 2, 3)$$

β_j es el vector de parámetros asociados con cada alternativa j y X_i es el vector de variables explicativas para cada individuo.

Los coeficientes estimados de este modelo logit multinomial, así como sus efectos marginales, aparecen en las tablas 3 y 4 para Andalucía y el resto de España, respectivamente[5]. Los contrastes estadísticos efectuados sobre dichos coeficientes y efectos marginales muestran que la mayoría son significativos, indicando esto junto al test de la razón de verosimilitudes la idoneidad de los modelos.

En relación a los propios resultados, se observa que los trabajadores nacionales con tres o menos años de experiencia, tanto en Andalucía[6] como en el resto de España, tienen una probabilidad relativa de tener un contrato indefinido (bonificado o no) frente a uno temporal superior que la de los extranjeros. Así, a partir de los *odds ratios* calculados con los coeficientes estimados, se constata que en Andalucía la probabilidad relativa de tener un contrato indefinido ordinario para los españoles cuadriplica a la de los africanos y triplica a la del resto de colectivos de inmigrantes. Estos resultados son similares a los del resto de España, aunque las diferencias de probabilidad relativa son algo menores. La diferenciación dentro de los bonificados muestra la misma pauta, aunque varía en términos cuantitativos. En efecto, para los contratos indefinidos bonificados por transformación de un contrato temporal, las diferencias de

[5] Una de las hipótesis necesarias para la aplicación del modelo multinomial es la independencia de las situaciones consideradas en la variable dependiente, dicha independencia ha sido corroborada en todos los modelos tras la aplicación del test de las alternativas irrelevantes.

[6] En la estimación del modelo *logit multinomial* para Andalucía, no aparece la variable ficticia relativa a los trabajadores asiáticos, ya que no existen observaciones de este colectivo para los contratos indefinidos bonificados iniciales.

los *odds ratios* entre nacionales y extranjeros se estrechan tanto para Andalucía como para resto de España. En particular, la probabilidad relativa de los nacionales sólo duplica a la de los africanos y trabajadores del resto de Europa. Por el contrario, para la probabilidad de tener un contrato indefinido inicial bonificado frente a otro temporal, las diferencias en la mayoría de los casos se amplían entre españoles y extranjeros, especialmente para el resto de España. Así, por ejemplo, esta probabilidad para los primeros es 6 veces superior a la de los africanos en Andalucía, y 14 veces mayor en el resto de España.

Para los trabajadores con más de 3 años de experiencia previa al empleo, los resultados respecto a la variable nacionalidad son diferentes. En Andalucía los *odds ratios* de tener un contrato indefinido ordinario frente a uno temporal de los españoles sólo son superiores a los que ostentan los trabajadores sudamericanos y africanos (en el resto de España, esto también ocurre para los trabajadores de la UE), y además para estos casos las diferencias son inferiores a las comentadas anteriormente para los trabajadores con menos experiencia. Este resultado, que es un indicio de la asimilación de los trabajadores inmigrantes más experimentados pertenecientes a esos grupos étnicos al estatus laboral de los españoles, se acrecienta si se analizan el resto de situaciones. En efecto, en Andalucía, la probabilidad relativa de tener un contrato indefinido bonificado por transformación de uno temporal para los españoles es incluso la mitad de la registrada para los asalariados de la UE, y similar a la constatada para los trabajadores de resto de Europa, Sudamérica y Asia. Por otra parte, es en esta Comunidad no aparecen divergencias significativas en la probabilidad de tener un contrato indefinido inicial bonificado frente a otro temporal según la nacionalidad, y para el resto de España las diferencias se atenúan considerablemente con respecto a los trabajadores menos experimentados. En definitiva, podemos decir que los resultados obtenidos bajo este enfoque están en línea con los procesos de asimilación y convergencia atisbados entre la población inmigrante y la nativa en los trabajos de Fernández y Ortega (2006) y Amuedo-Dorantes y De la Rica (2007) anteriormente mencionados, que incrementan con el aumento del periodo de permanencia de los extranjeros en España.

En cuanto al resto de características personales, cabe señalar, por un lado, que los efectos de la edad sobre la probabilidad de situarse en una determinada modalidad contractual son coherentes con el análisis descriptivo previamente presentado. Así, independientemente de la experiencia y de la zona geográfica donde se desarrolla el empleo, la edad ejerce un efecto positivo sobre la probabilidad de ostentar un contrato indefinido ordinario o un contrato indefinido bonificado inicial. En el primer caso, el resultado es coherente con la teoría que postula que el empleo temporal es más frecuente entre los individuos con menos edad, ya que estos contratos son utilizados como escalones para el logro de un empleo de calidad, que en este caso sería tener un contrato indefinido ordinario (véase, por ejemplo, Booth *et al*. 2005). En el segundo caso, las estimaciones parecen indicar que en el periodo analizado las bonificaciones para contrataciones indefinidas iniciales se han centrado en el paro de larga duración, que está relacionado positivamente con la edad del trabajador (véase, por ejemplo, Bover *et al*. 2002). Por último, se observa que la edad influye negativamente sobre la probabilidad de que un contrato temporal se transforme en indefinido, ya que los empresarios prefieren estabilizar a los trabajadores más jóvenes dentro del colectivo de trabajadores temporales. Una de las posibles explicaciones de este resultado puede ser que los trabajadores con menos edad presentan un horizonte laboral más amplio, lo cual puede implicar una mayor rentabilización de sus inversiones en formación laboral (Caparrós, Navarro y Rueda 2004). Por otro lado, en relación al sexo de los individuos, se corrobora en todas las situaciones que los varones tienen una mayor probabilidad relativa de tener un contrato indefinido ordinario o indefinido bonificado por transformación de uno temporal. En estas situaciones, los empleadores pueden percibir el empleo indefinido femenino como más inseguro que el masculino, dado el mayor riesgo de abandono voluntario asociado a las interrupciones que las mujeres pueden realizar a lo largo de su trayectoria profesional, principalmente por motivos familiares (véase, por ejemplo, Caparrós y Navarro 2007). Sin embargo, para la probabilidad relativa de tener un contrato indefinido bonificado inicial ocurre lo contrario, lo cual es un resultado que se podía anticipar en la medida en que el colectivo femenino está más representado dentro de la población bonificable, valga como ejemplo que a partir de la reforma laboral del 2001 todas las mujeres desempleadas independientemente de su edad podían estar sujetas a la bonificación.

Con respecto a las características del empleo, se observa, en primer lugar, que la actividad ejerce una influencia significativa sobre la modalidad contractual. En segundo lugar, se constata que los trabajadores de la construcción son los que menos probabilidad tienen de ostentar un contrato indefinido ordinario (por ejemplo, 6 y 3 veces inferior que en la industria, según que los trabajadores en Andalucía sean más o menos experimentados, respectivamente) y de beneficiarse de las bonificaciones al empleo indefinido (en este último caso, junto a la educación y la sanidad). Por otro lado, en esta Comunidad e independientemente de la experiencia del trabajador, los sectores con una mayor probabilidad de tener contratos indefinidos bonificados en relación a la industria son el comercio, la hostelería y el transporte, mientras que para el resto de España esto sólo ocurre en el comercio y la hostelería, en la mayoría de los casos.

Tabla 5. Estimación de la probabilidad de situarse en un determinado contrato en Andalucía[a,]

Variables	3 o menos años de experiencia previa al empleo						Más de 3 años de experiencia previa al empleo					
	Indefinidos ordinarios		Indefinidos bonificados por transf.		Indefinidos bonificados iniciales		Indefinidos ordinarios		Indefinidos bonificados por transf.		Indefinidos bonificados iniciales	
	Coef.	Efecto Marg.	Coef.	Efecto Marg.	Coef.	Efecto Marg.	Coef.	Efecto Marg.	Coef.	Efecto Marg.	Coef.	Efecto Marg.
Nacionalidad												
UE y EEUU	-1,09***	-0,15***	-0,38**	0,01	-0,92***	-0,03***	-0,04	-0,05	0,65***	0,09**	0,37	0,01
Resto de Europa	-1,12**	-0,13**	-0,98***	-0,08*	-1,33**	-0,04**	-1,01	-0,18*	0,61	0,14	-0,23	-0,01
Sudamérica	-1,27**	-0,17***	-0,39***	0,01	-1,30**	-0,05***	-1,01**	-0,15**	-0,27	0,01	-0,38	-0,01
África	-1,45***	-0,17***	-0,94**	-0,06**	-1,84***	-0,06***	-1,03***	-0,15**	-0,82*	-0,05	-0,51	-0,08
Asia							0,079	0,04	-0,62	-0,06	-0,31	-0,01
Edad	0,028***	0,01***	-0,03***	-0,01***	0,04***	0,01***	0,03***	0,01***	-0,03***	-0,01**	0,07***	0,01***
Sexo												
Varón	0,41***	0,10***	0,15***	0,02**	-0,86***	-0,09***	0,13**	0,04***	0,19	0,02***	-1,11***	-0,07***
Sector Actividad												
Construcción	-1,26***	-0,17***	-0,75***	-0,04***	-0,88***	-0,03**	-1,82***	-0,25***	-1,35***	-0,08***	-1,73***	-0,04***
Comercio	0,52***	0,01	0,78***	0,07***	1,10***	0,06***	0,29***	0,001	0,74***	0,08***	0,87***	0,04***
Hostelería	0,61***	0,06***	0,57***	0,03*	0,89***	0,05***	0,38***	0,03*	0,64***	0,06***	0,76***	0,03***
Transporte	0,53***	0,04***	0,68***	0,07***	0,59***	0,02	0,44***	0,06***	0,54***	0,04***	0,13	-0,01
Activ. Inmobiliar.	-0,22***	-0,07***	0,21**	0,04**	0,41***	0,04***	-0,07	-0,02	0,06	0,01	-0,07	-0,01
Educación	-0,44***	-0,05**	-0,57***	-0,62**	-0,23	0,01	-0,82***	-0,12***	-0,53***	-0,03*	-0,56***	-0,01
Sanidad	-0,77***	-0,13**	-0,24**	-0,01	0,09	0,03**	-0,79***	-0,12***	-0,40**	-0,01	-0,60***	-0,01*
Otras actividades	0,26***	-0,01	0,45***	0,04*	0,79	0,05***	0,12	0,01	0,36***	0,04*	0,46**	0,02**
Cat. profesional												
Ingenieros	1,26***	0,16***	1,13***	0,08***	0,61***	-0,02*	0,92***	0,17***	0,38**	-0,01	0,85***	0,02**
Técnicos	1,37***	0,14***	1,41***	0,12***	0,78***	-0,16	0,99***	0,18***	0,46***	-0,01	0,86**	0,02*
Jefes de Adm.-taller	2,18***	0,27***	1,52***	-0,01	1,94***	0,04**	1,46***	0,24***	0,88***	0,01	1,48***	0,04**
Ayud. no titulados	2,12***	0,22***	1,80***	0,06**	1,71***	0,01	1,65***	0,30***	0,90***	-0,01	1,22***	0,01
Oficiales Admin.	2,14***	0,19***	1,97***	0,09**	2,17***	0,05***	1,50***	0,20***	1,36***	0,07***	1,61***	0,04***
Subalternos	1,14***	0,11***	1,13***	0,08**	1,07***	0,02	0,28**	0,01	0,61***	0,06***	0,61**	0,02*
Auxiliar Admin.	1,30***	0,11***	1,32***	0,09***	1,55***	0,06***	0,63***	0,06***	0,68***	0,04**	1,51***	0,09***
Oficiales 1ª y 2ª	1,16***	0,08***	1,39***	0,12***	1,48***	0,05***	0,50***	0,06***	0,55***	0,04**	0,98***	0,04***
Oficiales 3ª	0,99***	0,04***	1,37***	0,14***	1,25***	0,03**	0,56**	0,06***	0,68***	0,05***	0,91***	0,04**
Nº Empl. Temp.	-0,39***	-0,06***	-0,16***	0,01**	-0,29***	-0,01***	-0,31***	-0,06***	-0,06***	0,01***	-0,03***	0,01***
Constante	-1,68***		-0,59***		-3,06***		-0,88***		-0,45***		-4,69***	
Probabilidad media predicha		0,31		0,23		0,07		0,30		0,15		0,05
Nº Obsevaciones	16838						17666					
Test RV	6409,61***						8145,8***					

Notas: (a) El individuo de referencia es español, mujer, trabaja en la industria y es peón.
(***) Coeficiente significativo al 1%, (**) al 5%, (*) al 10%.

Tabla 6. Estimación de la probabilidad de situarse en un tipo de contrato en resto de España[a]

Variables	3 o menos años de experiencia previa al empleo						Más de 3 años de experiencia previa al empleo					
	Indef. ordinarios		Indef. bon t ransf.		Indef. bon. inic.		Indef. ordin.		Indef. Bon t ransf.		Indef. bon. inic.	
	Coef.	Ef. M.	Coef.	Ef. M.	Coef.	Ef. M.	Coef.	Ef. M.	Coef.	Ef. M.	Coef.	Ef. M.
Nacionalidad												
UE y EEUU	-1,03***	-0,13***	-0,46***	0,04***	-1,51***	-0,06***	-0,55***	-0,11***	-0,15*	0,03***	-0,44***	-0,01
Resto de Europa	-1,07***	-0,16***	-0,25***	0,09***	-1,55***	-0,07***	-0,19	-0,03	0,06	0,03	-0,71*	-0,03**
Sudamérica	-1,15***	-0,16***	-0,51***	0,04***	-1,40***	-0,06***	-0,88***	-0,16***	-0,39***	0,01	-0,64***	-0,01*
África	-1,25***	-0,16***	-0,69***	0,01	-1,92***	-0,08***	-1,01***	-0,17***	-0,52***	0,01	-1,50***	-0,04***
Asia	-0,83***	-0,04**	-1,08***	-0,08***	-2,63***	-0,09***	-0,23	0,02	-0,70***	-0,06***	-1,54***	-0,05***
Edad	0,03***	0,01***	-0,02***	-0,01***	0,05***	0,01***	0,03***	0,01***	-0,03***	-0,01***	0,08***	0,01***
Sexo												
Varón	0,15***	0,07***	-0,01	0,01***	-0,85***	-0,09***	-0,01	0,04***	0,01	0,01***	-0,99***	-0,07***
Sector Actividad												
Construcción	-1,65***	-0,20***	-1,09***	-0,04***	-1,30***	-0,03***	-2,28***	-0,37***	-1,46***	-0,05***	-1,51***	-0,02***
Comercio	0,03	-0,02***	0,16***	0,02**	0,32***	0,03***	-0,17***	-0,08***	0,30***	0,05***	0,47***	0,03***
Hostelería	0,16***	0,21**	0,04	-0,01**	0,27***	0,02***	-0,33***	-0,11***	0,20***	0,05***	0,29***	0,03***
Transporte	0,34***	0,08***	0,01	-0,03***	-0,01	-0,01***	-0,08**	-0,01***	0,04	0,01**	-0,18**	-0,01**
Activ. Inmobiliarias	-0,88***	-0,09***	-0,82***	-0,05***	-0,70***	-0,01**	-0,94***	-0,15***	-0,74***	-0,03**	-0,71***	-0,09***
Educación	-1,44***	-0,13***	-1,85***	-0,17***	-1,47***	-0,04***	-1,84***	-0,28***	-1,73***	-0,09***	-1,60***	-0,03***
Sanidad	-1,36***	-0,15***	-1,10***	-0,06***	-1,27***	-0,04***	-1,63***	-0,26***	-1,27***	-0,05***	-1,66***	-0,04***
Otras actividades	-0,72***	-0,08***	-0,63***	-0,04***	-0,34***	0,01**	-0,77***	-0,14***	-0,50***	-0,01**	-0,32***	0,01**
Categ. profesional												
Ingenieros	1,13***	0,20***	0,44***	-0,07***	0,50***	-0,02***	1,27***	0,27***	-0,07	-0,11***	0,70***	-0,01**
Técnicos	1,20***	0,13***	0,88***	-0,01	0,85***	-0,01	1,13***	0,23***	0,10**	-0,08***	0,54***	-0,01**
Jefes Admin.-Taller	1,95***	0,23***	1,08***	-0,09***	1,67***	0,02***	1,73***	0,26***	0,62***	-0,09***	1,59***	0,02***
Ayudantes no titulados	1,45***	0,17***	0,92***	-0,04***	1,22***	0,02**	1,38***	0,24***	0,43***	-0,07***	1,01***	0,01
Ofic. Administrativos	1,50***	0,13***	1,15***	-0,01***	1,57***	0,05***	1,31***	0,18***	0,77***	-0,03***	1,33***	0,03***
Subalternos	1,00***	0,10***	0,70***	-0,01***	1,02***	0,03***	0,94***	0,15***	0,34***	-0,05***	1,03***	0,03***
Ayud. Administrativo	0,82***	0,04***	0,77***	0,01***	1,18***	0,06***	0,49***	0,04***	0,34***	-0,01**	1,09***	0,06***
Oficiales 1ª y 2ª	0,82***	0,03***	0,86***	0,03***	1,15***	0,05***	0,44***	0,06***	0,28***	-0,01**	0,75***	0,03***
Oficiales 3ª	0,56***	0,01***	0,65***	0,04***	0,84***	0,04***	0,30***	0,05***	0,16***	-0,01***	0,39***	0,01***
CCAA												
País Vasco	-0,84***	-0,10***	-0,63***	-0,03***	-0,45***	0,01	-0,45***	-0,06***	-0,56***	-0,04***	-0,35***	-0,01
Castilla-La Mancha	-0,72***	-0,07***	-0,74***	-0,05***	-0,60***	-0,01***	-0,73***	-0,10***	-0,72***	-0,04***	-0,76***	-0,01***
Valencia	-0,55***	-0,08***	-0,32***	-0,01***	-0,24***	0,01***	-0,57***	-0,10***	-0,41***	-0,01***	-0,24***	0,01**
Galicia	-1,05***	-0,13***	-0,72***	-0,03***	-0,65***	-0,01***	-0,84***	-0,13***	-0,81***	-0,04***	-0,58***	-0,01
Castilla-León	-0,52***	-0,04***	-0,57***	-0,05***	-0,49***	-0,01***	-0,56***	-0,07***	-0,58***	-0,03***	-0,53***	-0,01**
Extremadura	-0,90***	-0,06***	-1,25***	-0,12***	-1,00***	-0,03***	-1,01***	-0,14***	-1,25***	-0,01***	-1,06***	-0,02***
Aragón	-0,67***	-0,08***	-0,49***	-0,02***	-0,52***	-0,01***	-0,44***	-0,06***	-0,42***	-0,04***	-0,45***	-0,01**
Canarias	-0,40***	-0,01	-0,68***	-0,07***	-0,67***	-0,02***	-0,51***	-0,05***	-0,69***	-0,03***	-0,88***	-0,03***
Madrid	-0,26***	-0,01	-0,45***	-0,05***	-0,26***	-0,01	-0,18***	0,00	-0,42***	-0,08***	-0,32***	-0,01***
Baleares	-0,10*	0,03**	-0,40***	-0,06***	-0,27***	-0,01	-0,42***	-0,04***	-0,57***	-0,02**	-0,46***	-0,01**
Murcia	-0,81***	-0,09***	-0,56***	-0,01***	-0,75***	-0,02***	-0,80***	-0,14***	-0,39***	-0,04***	-0,53***	-0,01
Navarra	-0,33***	-0,01	-0,52***	-0,06***	-0,33***	-0,01	-0,25***	-0,03**	-0,37***	-0,04***	-0,09	0,01
La Rioja	0,01	0,05**	-0,43***	-0,08***	-0,04***	0,01	-0,14	0,01	-0,84***	-0,09***	0,07	0,02**
Asturias	-0,99***	-0,10***	-0,98***	-0,07***	-0,83***	-0,01**	-0,78***	-0,11***	-0,97***	-0,06***	-0,57***	-0,01
Cantabria	-0,55***	-0,05***	-0,51***	-0,03***	-0,56***	-0,01**	-0,63***	-0,09***	-0,74***	-0,05***	-0,37***	-0,01
NºEmp. temp.	-0,17***	-0,02***	-0,08***	0,00***	-0,11***	-0,01**	-0,14***	-0,03**	-0,04***	0,01***	-0,01***	0,01**
Constante	-0,13**		1,41***		-1,88***		-0,01		1,40***		-3,60***	
Prob. media pred.		0,38		0,30		0,11		0,48		0,48		0,07
Nº Observaciones	101333						117618					
Test RV	30009,35***						42599,72***					

Nota: (a) El individuo de referencia es español, mujer, trabaja en industria, es peón y vive en Cataluña.
(***) Coeficiente significativo al 1%, (**) al 5%, (*) al 10%.

En cuanto a la categoría profesional del trabajador en el puesto de trabajo, prácticamente en todas las situaciones los peones muestran una probabilidad relativa de tener un contrato indefinido (bonificado o no) inferior a la del resto de ocupaciones. Así, como rasgo significativo, para los trabajadores con menos de 3 años de experiencia, la probabilidad relativa de tener un contrato indefinido ordinario para los jefes de administración o de taller es 9 y 7 veces superior a la correspondiente para los peones en Andalucía y en el resto de España, respectivamente. Por otra parte, en la Comunidad andaluza el grupo ocupacional que en términos relativos se ha beneficiado en mayor medida de las bonificaciones al empleo indefinido ha sido el de los oficiales administrativos (por ejemplo, para los trabajadores con menos experiencia, es 8 veces superior a la de los peones), seguidos de los ayudantes no titulados y jefes de administración y taller.

Con la variable número de empleos temporales previos, se pretende captar si las trayectorias laborales con varios episodios de precariedad provocan una devaluación de las características productivas del trabajador ante el empresario, es decir, lo que en la literatura económica se denomina "efecto estigma" (véase, por ejemplo, Boyce *et al*. 2007). Los resultados nos muestran que los *odds ratios* son negativos por lo que se concluye que un incremento en el número de empleos precarios provoca una disminución de la probabilidad relativa de tener un contrato indefinido (bonificado o no) frente a otro temporal.

Para finalizar, centrándonos en la influencia de residir en una determinada Comunidad sobre la modalidad contractual en las muestras correspondientes al resto de España, se observa como resultado más destacado que los trabajadores residentes en Cataluña tienen una probabilidad relativa de tener un contrato indefinido (bonificado o no) superior a la existente en cualquier otra Comunidad Autónoma. Concretamente, la mayor divergencia ocurre con Extremadura donde esa probabilidad relativa de poseer un contrato indefinido ordinario es aproximadamente la mitad que la existente en Cataluña.

4. CONCLUSIONES

En este trabajo se ha pretendido arrojar nueva evidencia empírica sobre el proceso de asimilación de los trabajadores inmigrantes, diferenciando a los asalariados entre los residentes en Andalucía frente al resto de España. Concretamente, se ha valorado si la calidad de sus empleos converge hacia la existente para los españoles.

Los resultados obtenidos muestran que para los trabajadores menos experimentados, los españoles tienen una probabilidad de tener un contrato indefinido (bonificado o no) respecto a uno temporal superior a la de los colectivos extranjeros, tanto para Andalucía como para el resto de España. Para los asalariados con más de tres años de experiencia, las conclusiones son diferentes. Así, en primer lugar, en Andalucía el *odd ratio* contrato indefinido/contrato temporal de los españoles sólo es superior al de los trabajadores sudamericanos y africanos. En segundo lugar, la probabilidad relativa de tener un contrato indefinido bonificado por transformación de uno temporal de los españoles es la mitad de la correspondiente a los trabajadores de la UE, siendo similar a la del resto de colectivos si se exceptúan a los africanos. Finalmente, en Andalucía, no hay diferencias según nacionalidad entre la probabilidad relativa de tener un contrato indefinido bonificado inicial, siendo las discrepancias menores en el resto de España. Por lo tanto, los resultados muestran indicios de que el proceso de asimilación de la población inmigrante está correlacionada positivamente con la experiencia del trabajador, y en definitiva con su tiempo de permanencia en España.

REFERENCIAS BIBLIOGRAFICAS

AMUEDO-DORANTES, C. y DE LA RICA, S. (2007): "Labour market assimilation of recent immigrants in Spain". *British Journal of Industrial Relations*, vol. 45, (2).

BLAU, F.D. and KAHn, L. (2005): "Gender assimilation among Mexican Americans". NBER, Working Paper 11512.

BOOTH, A; FRANCESCONI, M; FRANK, J. (2005): Temporary Jobs: Stepping Stones or Dead Ends? IZA Discussion Paper 205.

BORJAS, G. (1985): "Assimilation, changes in cohort quality, and the earnings of immigrants". *Journal of Labor Economics*, 3, pp. 463-489.

BORJAS, G. (1995): "Assimilation and changes in cohort quality revisited: What happened to immigrant earnings in the 1980s?" *Journal of Labour Economics* 13, pp. 201-245.

Bover, O. et al. (2002): "Unemployment duration, benefit duration and the business cycle". *Economic Journal* 112, pp. 223-265.

BOYCE, A. et al. (2007): "Temporary worker, permanent loser? A model of stigmatization of temporary workers". *Journal of Management*, vol. 33, nº 1, pp. 5-29.

CAPARRÓS, A et al (2004). "Efectos de la temporalidad sobre la formación recibida durante el empleo". Cuadernos de Economía, vol. 47, pp. 51-73.

CAPARRÓS, A. y NAVARRO, L (2007): "La naturaleza de la temporalidad en el empleo: nueva evidencia empírica. Estadística Española)" , vol. 49, nº 164, pp. 59-101.

CHISWICK, B. (1978): "The effect of americanization on the earnings of foreing-born men". Journal of Political Economy, 86, pp. 897-921.

FERNÁNDEZ, C y ORTEGA, C (2006): "Labor market assimilation of immigrants in Spain: Employment at the expense of bad job-matches?". FEDEA, documento de trabajo 2006-21.

APPRAISEMENT OF THE ENDOGENOUS RESOURCES OF THE COOPERATIVES FOR THE ECONOMIC INCREASES IN THE RURAL AREAS. CASE OF THE PROVINCE OF TOLEDO

VALENTÍN MOLINA MORENO
Departamento Organización de Empresas
Universidad de Granada
VÍCTOR MANUEL MARTÍN LÓPEZ
Caja Rural de Toledo
IGNACIO RUIZ GUERRA
Ignacio.Ruiz@uclm.es
Universidad de Castilla – La Mancha

Resumen

After the years habits of the Spanish holidays have been changing, modifying the temporality in the tourism, though in a very slightly pronounced way we see that new trends arise to be able to replace the lack of these historical - artistic resources, of these landscapes and of these nature reserves across the endogenous resources of the zone, this is, that the increasing value of all, that has meant a revitalizing element in the life of a social group can be seen from a point of view of communication, exchange and enrichment for everything that one that does not belong to this population.
The patrimony of the rural areas cannot have the same valour that the cultural patrimony of the cities, like culturally after the years so from this article one has wanted to show a new trend in the tourist sector, very in vogue in Europe, which is the Industrial Tourism, and adapted to the rural world, the Cooperative Tourism, as principal representatives of the industrial patrimony of the Social Economy.

Palabras Clave: Cooperative Tourism; Industry; Social Economy; New Trend.
Thematic Area: Urban, Regional and Local Economics.

1.- INTRODUCTION

The importance of the tourism development in the last years has been spectacular up to the point of modifying the habits of life of the persons in his way of enjoying his free time. One must understand this fact in a more global context of conceptual transformation of the rural thing. Of the traditional immobility and isolation, the rural way has met reactivated in the first moment for the external influence of the urban thing and, at present, to turn turned into useful something, indispensable for his material utility and of values.

We are in a process of redefinition of the competitive positioning, in which the tourism of interior assumes an essential protagonism linked to the territory and to the jointly responsible actors of the tourist future that will will to balance revenues, employment and levels of development.

The rural thing has stopped associating to an image of marginality and decadence and begins to be a factor of quality and well-being. This means the putting in value of the ruralidad as social category: a set of spaces, products, practices, images, contents, etc. The city has been the locus of the progress and of the modernization and as the years went by this progress has developed a culture of values completely different from that our forbears apendieron in the rural space, our values now are camouflaged in the processes of marketing of the companies; instead of the manual traditional labors we have taken the route of the consumerism to satisfy our needs

and before the sustainable development of the rural zones that it favors the formal subsistence of certain features diferenciadores in the least industrialized areas, live through a process that eliminates the tradition, the folklore and the sustainability.

The ruralidad, is the form by which the persons are granted and grant to a place, object or idea, an identity attributing to him a few social, cultural, aesthetic and ideological particular characteristics. The cooperatives have assumed this new concept of the rural thing and, beside fulfilling their traditional functions (production, concentration of the offer, supply of inputs, marketing) have seen, in their links to the territory, new opportunities to contribute to the development and to attend to the new demands of the society and the new markets of the rural thing.

The development of the tourist sector and especially the high interterritorial competition provokes the birth of new formulae of leisure and from this point of research we want to show the potential that the province of Toledo and in general the Castilla - La Mancha region they have to be able to increase the annual number of visitors that it has and specially support the quotas of growth of pernoctancy in our land as real economic resource for the society. The generation of the knowledge, the application of the results of research and the permanent debate, they are, cash doubts, instruments that must constitute the transverse foundation of the policies, plannings and processes of the tourist sector.

It is advisable, by virtue of the expressed, the definition of real tourist diverse models as precise but flexible aims, and strategies that allow the attainment of the established purposes and the feedback of the process of planning; it is to say, the tourism, in general, and especially that of interior, must be ordained thanks to alive plannings, capable of assuming the changes and the trends to all the levels.

The territory like fundamental value has to respect and promote his own identity and personality. The revaluation of that one that is considered to be like own, singular or autochthonous is what is going to facilitate the differentiation on the tourist market, a rural environment of interior that offers multiple potentials, still immature, with clear evidences of fragmentation, scanty operative possibilities, as well as difficulties in essential processes as the creation, management, promotion and marketing. It is necessary to answer to unavoidable questions, As the definition of the competitive positioning of the tourist realities of interior, linked to different segments of activity, the design of the model of development of the tourism of interior, the permanent improvement of the structures of management and promotion, the professionalization of the agents and the implied parts and the impulse of an innovative culture in products and processes of improvement of the offer.

The tourist reality of interior, in spite of multiple efforts, suffers still from a lack of real tourist integrated product, mistake in management, ineffective promotion and weak marketing. In occasions, the reality of the tourism in the interior regions diminishes to an inventory of patrimonial resources and incomplete or not strictly tourist services, which are not capable of satisfying to a frustrated demand before unreal promotions.

This us obstacle in order that, at any one time, the tourist development in the above mentioned territory reaches a degree of such economic efficiency that is capable of laro of the time being kept to as a sustainable element inside the rural economy. Well for the flexibility to adapt to new governmental helps of public promotion without cost, for application of new organizational technologies that achieve a major economic efficiency, or for a constant change for reaching a better service and simultaneously a major impact in the opinion of the clientele that helps to increase the level of development.

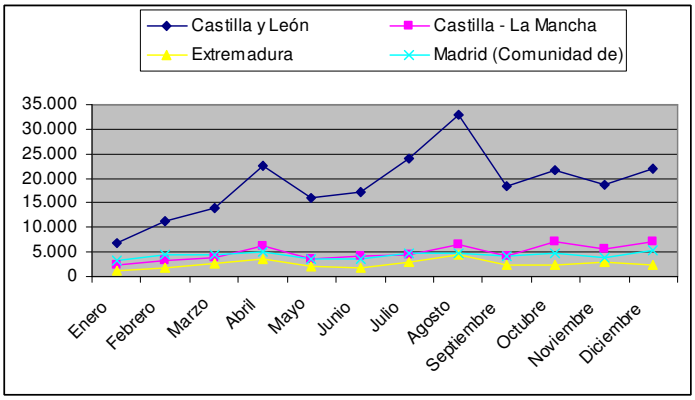
2.- CULTURAL TOURISM GROWING

The Cooperative Tourism, as cultural activity and of leisure, places inside the current trend of the tourism that looks for the differentiation and the specialization demanded by the current tourist. Parallel, the development of the initiative helps to generate and stimulate new activities related to services of housing and restoration, of local guides, of local trade, etc., the appraisement of the resources of the zone and the consistent creation and maintenance of employment in the rural zones.

This way, the offer, complementary of cultural activities to the beach tourism, helps to create a current of interest towards the rural way that makes the consideration of the territory possible as factor of development; to bring the rural thing and the urban thing over.

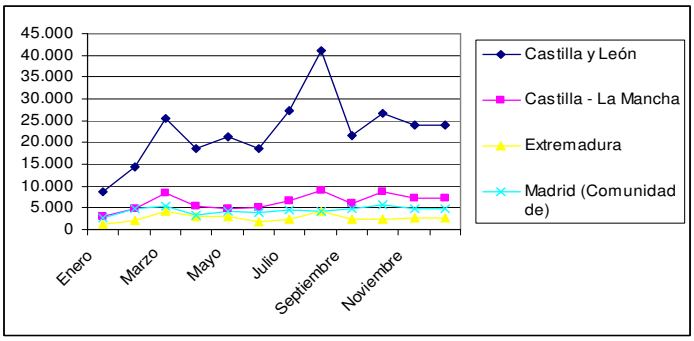
The growth of the tourist option of interior has demonstrated an important evolution in the last years, we can verify it in the tables of number of travelers that we can compare later they mark the trends of last years, from the year 2000 with the change of millenium to which it is not possible to impute a change in the trend of the individuals but in the last years it has grown in 4 autonomic communities in which we recount our study but it is something widespread, because the only autonomous community can be a chance, 2 a coincidence but the proof in four autonomous communities of this growth supported in the time is a proof of a clear push that can suppose losses to the topics of the Sun and beach in our country and that one has to adapt to this new situation, and at the same time as soon as the growth comes to a culminating point it is going to meet a fierce competition of the most ancient sectors of the tourist sector: beach and mountain.

Año 2001 Gráfico 1.-



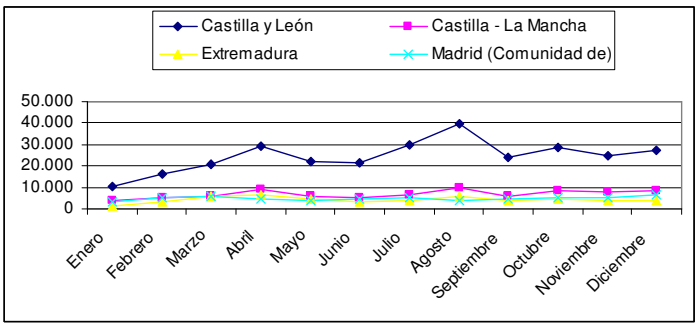
Fuente: INE, 2008.

Año 2002 Gráfico 2.-



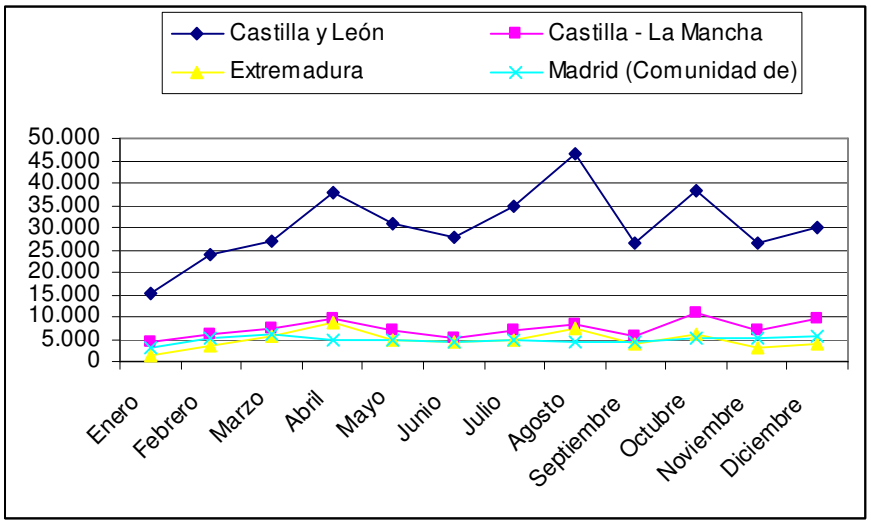
Fuente: INE, 2008.

Año 2003 Gráfico 3.-



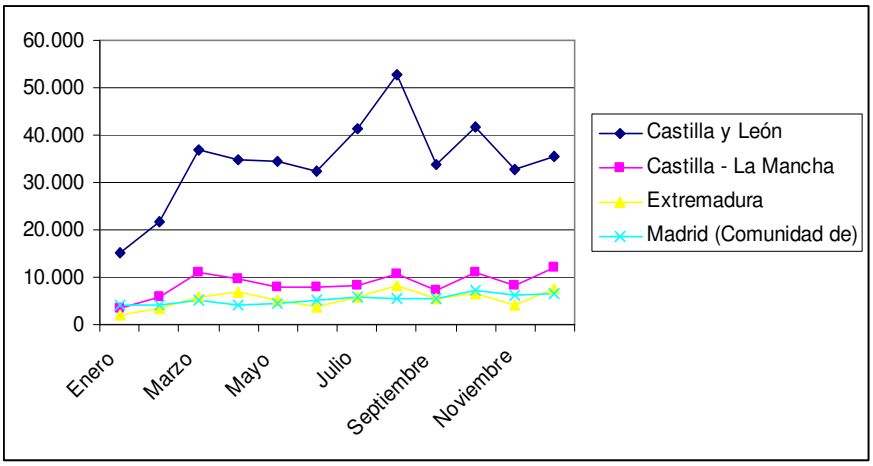
Fuente: INE, 2008.

Año 2004 Gráfico 4.-



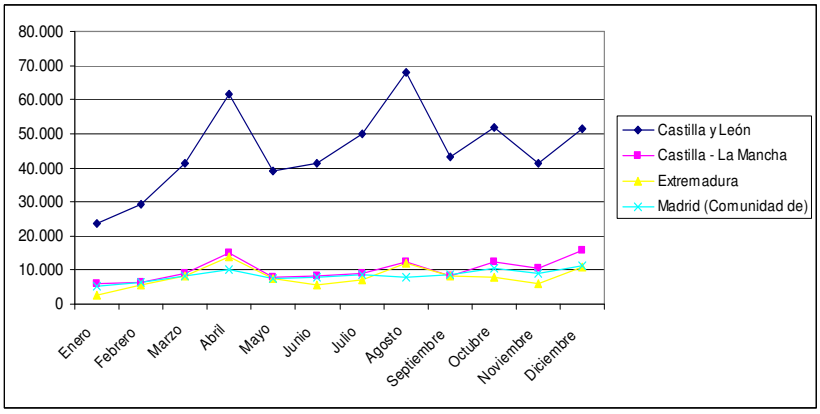
Fuente: INE, 2008.

Año 2005 Gráfico 5.-



Fuente: INE, 2008.

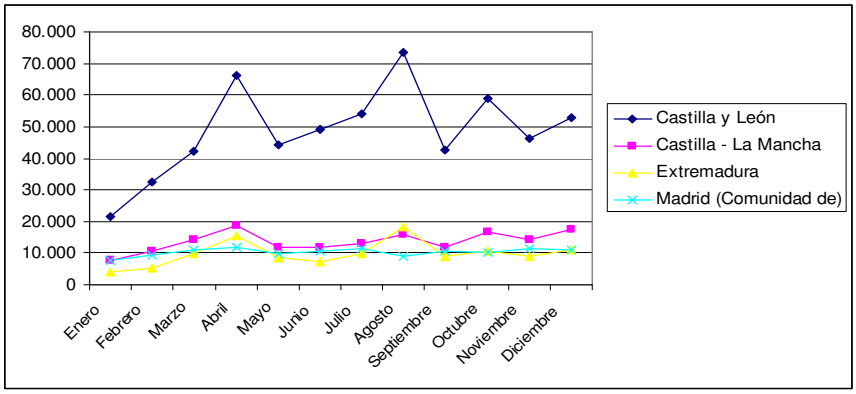
Año 2006 Gráfico 6.-



Fuente: INE, 2008.

Año 2007

Gráfico 7.-



Fuente: INE, 2008.

After viewing the monthly evolution of 4 autonomous communities we can observe and affirm that the termporalidad in the Tourism of Interior is in the same way that the Tourism of the Sun and beach very dependent on the moment of the year in which we are depending on the destination to which us refiramos since they are in the habit of being stays of weekend in that simply the travelers do not need from many kilometres to come. In other destinations a constant linearity is observed in the acceptance of a destination throughout the year, is low in quantity but constant throughout the months.

Later we can find the difference in visits and increase the pernoctancy according to the National Institute of Statistics as for tourists total annual abundance throughout the year and the annual evolution from the year 2001 until already finished 2007. It is very easy to affirm the enormous growth that has lived through the Cultural Tourism in these regions and especially since it has become strong on the tourist market and needs from innovations to support this gradual but constant growth to be able to turn into an economic revitalizing element of depressed zones.

	2001	2002	2003	2004	2005	2006	2007
Castilla y León	225.511	271.674	294.346	366.184	413.962	541.886	583.745
Castilla - La Mancha	58.319	76.399	82.536	87.572	103.649	121.109	163.280
Extremadura	30.192	31.784	50.039	58.683	65.815	95.929	117.980
Comunidad de Madrid	51.920	53.327	57.392	57.794	64.626	101.403	122.943

This one is the evolution in graph of this growth in Castile and León especially and how Castile-La Mancha tries to follow the step very of distant view but in clear competition with the Community of Madrid and Estremadura, which major hándicap is the distance with most of the travelers' broadcasting points of weekend since with a similar trip they can reach coast in any peninsular point what does that it is a disadvantage for the tourism Extremaduran recipient.

3. SOCIAL ECONOMY INNOVATION AND DIFFERENT DELOPMENT

The industrial patrimony has lived through an increase of interest in the public from middle of the 20th century, firstly in Great Britain and The United States, and later in other countries. In Spain the field of the industrial archaeology has been developing even having many obstacles for safeguarding the historical patrimony in all the Autonomous Communities. Regardless there is very much patrimony that has got lost.

As previously we were commenting there are multiple and varied the motives, the facts, the elements or places, definitively the attraction that does that the traveler dresses a certain place, and depending on it it is possible to speak about different types of tourism, business tourism, sports tourism, sun and beach, ecotourism, rural tourism, etc. The fact is that from this study we

try to adapt this trend tourist and care of the industrial patrimony as tourist resource to the technology that has characterized our land during the last years therefore we speak about the cooperative tourism included inside the cultural and scientific tourism. With the expression cultural and scientific tourism, we refer to this tourism which principal attraction is the cultural, scientific, technical or industrial patrimony.

The obstacles to which we are to be able to carry out a conversion of the cooperative patrimony in a tourist resource taken from the point of view of the industrial tourism, according to Dietrich Soyez (Capel, 1996):

- Cognitive barriers: it does not have aesthetic values, is not interesting, is not elegant.
- Economic barriers: it is more profitable to sell the machinery as scrap, to alienate the area, or to dedicate the building to other occupations.
- Legal or administrative barriers: slightly clear competitions on the performances to realize.
- Physical barriers: distance of some facilities with regard to the tourist traditional routes.

The Department of Public Works across the Center of Historical Studies of Public Works is realized a great labor of cataloguing, study and diffusion of this patrimony, translated in numerous exhibitions, catalogues and publications referred to bridges, ports, port equipments, beacons, hydraulic works, channels, fortifications, stations and factories, between others. At the same time the Consejerías of diverse Autonomous Communities they have followed this example and have started worrying for the study and the conservation of the industrial patrimony. The General Law of Historical Patrimony, allows to declare places and landscapes of cultural interest, and relying on it, the Autonomic Communities have protected mining areas, mills, factories or quarries.

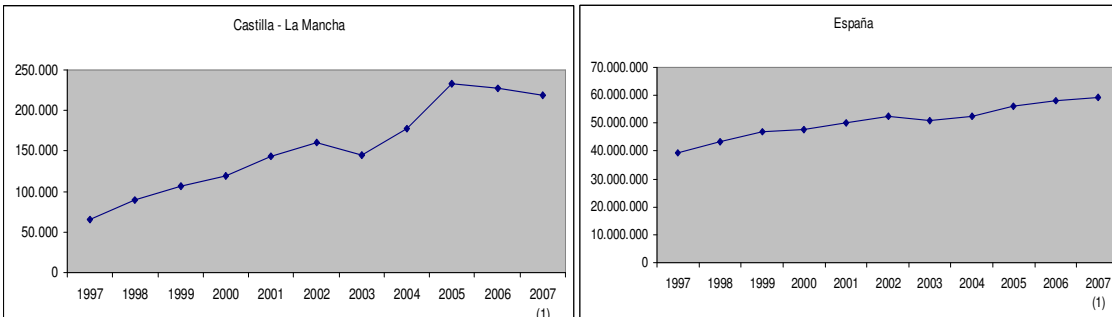
The cooperative tourism helps to support the image of the cooperative as company compromised with his territory, with the maintenance of the culture, of the landscape, of the patrimony and, definitively, it announces how the cooperatives help to create employment, to the rural development and to support the form of life of these zones.

4.- POTENTIALITIES AND ENDOGENOUS RESOURCES

Explanatory model of the population movements during the last years in the regions with a high specialization in the Cultural Tourism. There has been carried out across the information that TURESPAÑA facilitates to us, and the National Institute of Statistics a series of tables in which there is gathered the evolution of a series of "articles" with which we show the increase of the interest in the tourism of interior and the trend in other countries of the Industrial Tourism, and that in our country is in full growth.

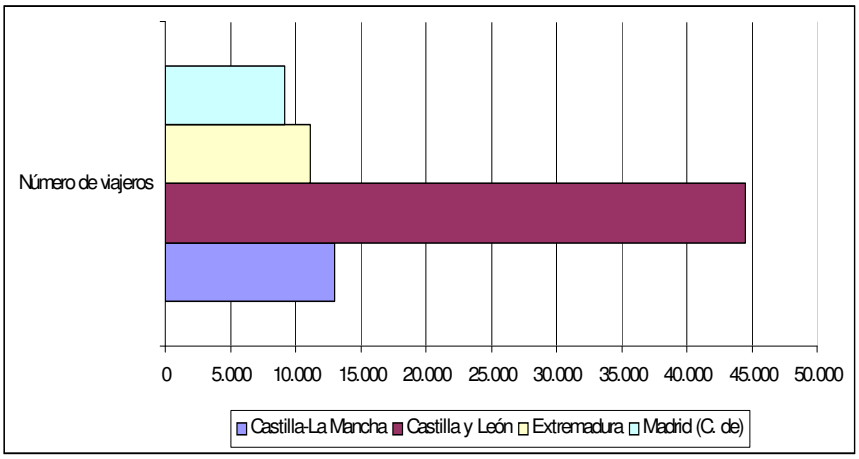
The region of Castile - The Spot is not specially characterized by the massive existence of big industries, and economically it is in the lowest positions of the table of the Spanish regions as for percentage of the national GDP. It is for it that adapting the European trends, and especially knowing the important patrimony of our land and the values that has with we have to be that we do an important labor of putting in value of it that makes to us the others different, the industry that ours major they have created to be able to go out forward without having to emigrate to the city, to survive in the place that saw them to be born and to support the spirit and the characteristics of the territory.

Gráficos 8 y 9.-



The tourist sector differs between two types of persons who realize the process of enjoyment of a trip, the tourists and the trippers, which stay in destination out of his place of origin and those who only realize the trip in the day without using establishments of housing or to happen the night again in his place of origin.

Gráfico 10.-



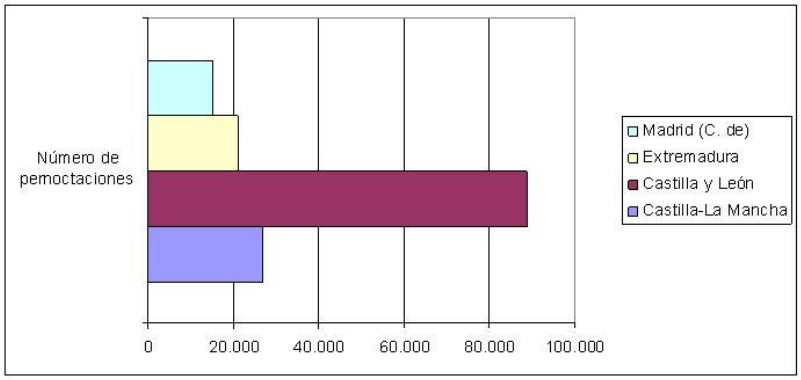
Fuente: Instituto Nacional de Estadística 2008

On the other hand the measurement of the number of pernoctaciones in these 4 autonomous communities on which we have based the study of the Tourism of Interior shows us a constant increase in the number of pernoctancy, that actually is what sustains the theory of the growth of this class of tourism in the rural zones.

The first one in number of nights is the Community of Madrid, it is clear that the type of tourism that monopolizes the capital of the country is completely different but is a clear example of the importance that has this Community for the set of the Tourism of Interior, capitalized in the business Tourism. On the other hand Castile is and León, which principal tourist value is led by his important and varied cultural patrimony and his famous annual campaigns “Las Edades del Hombre”, obtaining an important value added on having stirred the tourist life into action and having chosen one sedates differently every year in the whole Castill y León.

Thirdly Castilla – La Mancha, which still is looking which is the most important value to exploit his tourist value. It is clear that hereby it has certain disadvantage to be able to support a competitive position on the market tourist and more seen(dressed) that the foreign competition bid very loudly with very competitive prices against our important and varied offer of interior.

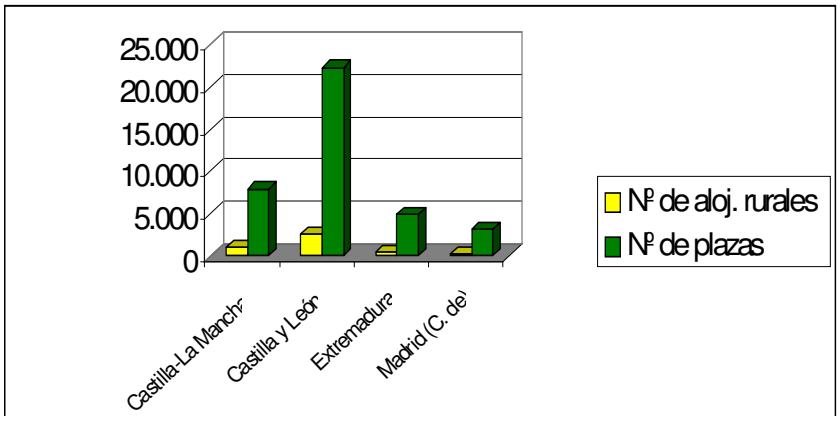
Gráfico 11.-



Fuente: Instituto Nacional de Estadística, 2008

The trend also has favored those territories that have been more dynamic and flexible at the moment of created frames of action and plans of dinamización that have achieved an important growth in the sector, in this case we have in the following table the numerical difference between the housings of rural type between 4 autonomous compared communities and especially the important imbalance remains formed between the capacity of housing of one and other one.

Gráfico 12.-



Fuente: INE 2008

5.- SOCIAL INNOVATION. NEW WAYS IN 21st CENTURY

The Social Economy supposes practically 80 % of the whole of economic developments and of the results of the GDP of our region what supposes an important dependence of the economic sector and of the rural world since the population in Castile - The Spot is in the rural populations of less than 10.000 inhabitants, is for this that in a region in that

In this respect, the Cooperative Tourism is an initiative that answers to these new opportunities of the rural thing and to the need to put in value the small productions of quality. The activity consists of the organization of visits to the cooperative centers of those that the agrarian product sues to transform it into food.

These visits have an eminently cultural character and of leisure from the point of view of the visitor and a promotional character of the cooperative image, of his products and points of sale.

With the development of the Cooperative Tourism there is prosecuted the diversification of the cooperative activity from the development of new lines of business that they contribute to

the appraisement of the cooperative products and of the rural local patrimony. In other words, the promotion of the territory, the culture and the products of the zone of influence, in this case the province of Toledo.

When there are announced the characteristics of a certain product is outlined what it makes singular to this product: the variety, the process of production, the system of culture, etc. Of this form, the cooperative assumes a divulgative paper of the food culture: to differentiate the quality of the products, to emphasize the importance of the product in the diet, etc. Definitively, participant of the production is done to the visitor and approaches the consumer the point of direct sale. It is a question so of a tool of promotion of the product tied to the territory in which the distribution chain is eliminated.

An interesting option to realize a different tourism, to know how the cooperatives work, what products they elaborate and which are his characteristics, simultaneously that we know the cultural, landscape and gastronomic patrimony of the zones where they are located. It is a question of proposing him a series of activities that reach to the leisure and the culture, prepared for groups and that have to be thought to enjoy what is the reality of the rural world of Castilla - La Mancha. An innovative route on the tourist market and a putting in value of the economic resources that have characterized us during the 20th century and that can continue serving as elements of subsistence in the 21st century.

In case of the zone that we place in the province of Toledo we can have to disposition 39 cooperatives of oil with a production that has introduced across the creation of a society comercializadora that moves his production to the general market penetrating the borders of our country to be able to find a place for this production that is of enormous quality but in turn it has many problems to be able to be located on the market for having this character rural and removed from the more important riverbeds of marketing. Therefore this new element ortencial is completely opportune like revitalizing of the economy and especially it is an opportunity to announce this product that lately is having problems to be able to be commercialized in a natural way and without problems.

The territory of Castile - The Spot is a region of constrastes, known by his protected landscapes, by his orographical diversity, by his characteristic plain of the plateau, important producer of oil and wine and an agriculture of dryness in full period of readaptation to the new European situation. From the traditional own productions of the cooperatives, it is a question of constructing a program of every product of agreement with the facilities in those who are elaborated and the environment in which it is placed.

The routes of Cooperative Tourism basing on the trends of the Industrial European Tourism they have to try to obtain several advantages in diverse fields, since it is that of the territorial, patrimonial, cultural promotion, but especially commercial for the products that are manufactured in the cooperatives since are the oil, the wine, the cheese, etc. Also, like singular element of the rural patrimony of the interior, there have to improve themselves the landscapes typical of the long and abundant olive fields as typical landscapes, also the long fields of vineyards of that together with the big warehouses prepared for it they can take advantage to show the quality of the cooperative products taking advantage that the clients come to know first hand, The principal thing is that the offer acquires a sense and attraction when we think that the administration and the private actors interactuan jointly create a series of routes in his set, as collective project, of a group of cooperatives and coordinated by the regional or provincial Administration that there diversifies the cultural predominant offer in our tourism recipient and that can acquire an important role in the future for his enormous potentials.

This type of routes have to have a common structure that allows to the visitor to know the field of culture and the landscape associated with him, the characteristics organolépticas of the product, the process that follows for his production and the facilities, to realize the degustation of a selection of cooperative products and of the traditional gastronomy of the zone. For example, emphasizes the aesthetic value of the vareado of the olive trees, in high season of tourism of interior, with an interesting way of obtaining of the fruit, of a traditional way, the high quality of the olive, also the recovery of the wine tradition that has supposed the production of author's important wines in the private warehouses and in the cooperatives that

have bet for the quality, the differentiation of the oils virgin extra of every region or the singularity of the oil of thousand-year-old olive trees and of the monovarietales.

The tourist activity has before yes an important labor as for the promotion and diffusion of the initiative, tematización of the activity, conciliation and organization of the visits (guided, follow-up, preparation of the degustations, etc.). Though the visit to the facilities at the expense of the own human resources from the cooperative would need from an important commitment of the cooperatives, since they are those who know well the product, the process and are most qualified to transmit his values to the visitor and therefore it is to that there is going to redound the fruit of the success of these initiatives and of that they can take advantage inside the innovative measures for the diversification of income and productivity.

The cultural and divulgative character of this type of activities makes possible the establishment of new alliances between the cooperative world and other entities of the province: cultural associations, educational centers, companies, town halls, etc. Interested parties in knowing and re-discovering the regions of the interior from the point of view of the landscape and the patrimony, the agrarian culture, the local products and the traditional gastronomy. This approximation helps to generate and consolidate the image of the cooperative as entity compromised with his territory: with the maintenance of the culture, of the landscape, of the patrimony and, on the other hand, with the quality and singularity of the food that produces, with the trazability in the processes of production, with the importance of a healthy diet, etc.

Since it has been commented previously, an important function of the activities of Cooperative Tourism is to bring the product over to the final consumer across the direct sale in cooperative shops. This helps to announce in addition the visitor the wide range of qualit products that produce the cooperatives and the points of sale in which these are available.

The innovative character owes to the incorporation, on the part of the cooperatives, new functions related to the development of a new line of business linked to the organization of cultural activities and of leisure. Without stopping fulfilling his traditional functions, the cooperatives have the opportunity to announce his product and his to be able to make, bring the origin of the production over to the final consumer and contribute to the development of his territory putting in value the local resources.

6.- CONCLUSIONS

Undoubtedly, this activity, still incipient, demands to pay attention to aspects also new for the cooperatives as the attention to the visitor, the conditioning of the facilities and of the points of sale that are visited, the own promotion of the activity, the production of materials of interpretation and signposting, that is to say, demands new tools, new approaches to be able to turn the productive activity into a tourist offer. This tematización of the tourist offer places inside the new trends of the tourism in the rural way towards the differentiation and the specialization. Without losing the genuineness and the originality of what already exists (facilities, to be able to do, etc.), it is necessary to provide to the activity of values that make it different and that already are demanded by the "cultural tourist".

From the current situation, with a potential, constructed offer and that had a constant implementation and promotion, the development of this type of tourism would need to bring together other itineraries and tourist cooperative routes that were complementing the already described offer.

In this respect, the intercooperation with cooperatives of other regions that develop cultural activities and of leisure linked to the product and to the territory it would allow to articulate a more diverse offer under the same exposition. Equally, the integration of the offer in other tourist circuits and the integration of other activities they will make possible to come to a more numerous and diverse public.

This type of initiatives are having already an important labor in Castellón and to European level with a project that receives entities of Italy, Holland, England to create networks of exchange that the experiences enrich of some and others for an improvement in the quality of life of the world member of a cooperative.

7.- BIBLIOGRAPHY

ABERNATHY, W., UTTERVACK, J., (1978): "Patterns of industrial innovation", Technology Review nº 80 (7), pp. 40-47.

BALLART, J., TRESERRAS, J., (2001): "Gestión del Patrimonio Cultural", Ariel, Barcelona.

BARRAS, R. (1986): "A Comparison of Embodied Technical Change in Services and manufacturing Industry", Applied Economics nº18 (9), pp. 941-58.

BOULLON, R., (1985): "La planificación del espacio turístico". México: Trillas.

CAPEL, H., (1996): "El Turismo Industrial y el Patrimonio Histórico de la Electricidad", Instituto Andaluz del Patrimonio Histórico, Sevilla.

CASTILLO, J., LÓPEZ GARCÍA, M., (1998): "Arqueología Industrial: Arqueología del trabajo en Madrid", Universidad Politécnica de Madrid, Madrid.

LECOURS, J., (1999): "L`insertion du patrimoine Industriel Dans L`infraestructure Touristique: problemes d`attrait et d`authenticité", Canadá: Université du Québec.

LLURDES I COIT, J., (1995): "El Turismo de Patrimonio Industrial y Minero. Una experiencia de turismo de interior explotada en el Estado español". Barcelona: Universidad Autónoma de Barcelona.

MARCHENA GÓMEZ, M. (coord.), (1999): "Turismo Sostenible y Gestión Municipal". Agenda para Planificadores Locales. Madrid: OMT.

MATHIESON, A., WALL, G., (1990): "Turismo. Repercusiones económicas, físicas y sociales". México: Trillas.

PAZ, C. A., (2001): "Capitalismo, Tecnología e Impacto Ambiental. Las transformaciones socioculturales, estructurales y ambientales del subsistema minero olavarriense". Centro de Investigaciones Ambientales, Universidad Mar del Plata, Argentina.

PAZ, C., VISVEQUI, R., (2001): "Turismo Industrial y Patrimonio Cultural en Olavarría. El pasado de la industria minera como potencial turístico". Royecto GIAAI. NURES. Universidad del Centro de la Provincia de Buenos Aires.

PÉREZ PRADOS, E., (2001): "Turismo cultural: un segmento turístico en expansión", en Congreso Vistual de Turismo Cultural. WEB: www.naya.org.ar.

ROMERO MORAGAS, C., (1994): "Patrimonio, Turismo y Ciudad", Boletín del Instituto Andaluz de Patrimonio. 9: 16-21.

SANTANA, A., (2003): Editorial. "Patrimonios culturales y turistas: unos leen lo que otros miran". PASOS, Revista de Turismo y Patrimonio Cultural. WEB: www.pasosonline.org.

SANTOS, M., (2000): La naturaleza del espacio, Ariel, Barcelona.

VALCARCELT-RESALT, G. (et al), (1993): "Desarrollo local, turismo y medio ambiente", Diputación de Cuenca.

VERA, J. F. (et al), (1997): "Análisis territorial del turismo", Ariel, Barcelona.

ZIPF'S AND GIBRAT'S LAWS FOR MIGRATIONS

JESÚS CLEMENTE
E-mail: clemente@unizar.es
RAFAEL GONZÁLEZ-VAL
E-mail: rafaelg@unizar.es
IRENE OLLOQUI
E-mail: iolloqui@unizar.es
Departamento de Análisis Económico
Universidad de Zaragoza

Abstract

This paper analyses the evolution of the size distribution of the stock of emigrants in the period 1960-2000. Has the distribution of the stock of immigrants changed or has there been some convergence? This is the question discussed in this work. In particular, we are interested in testing the validity of two empirical regularities studied in urban economics: Zipf's law, which postulates that the product between the rank and size of a population is constant; and Gibrat's law, which states that growth rate of a variable is independent of its initial size. We use parametric and non-parametrical methods and apply them to absolute (stock of immigrants) and relative (migration density, defined as the quotient between the stock of immigrants of a country and its total population) measurements.

JEL: J61, R11, R12
Key words: Migration distribution; Zipf´s law; Gibrat´s law.
Thematic Area: Urban, Regional and Local Economics.

1. INTRODUCTION

In the study of the economic landscape, the influence of the location of the productive factors on economic activity is an important element. Some of these factors cannot be moved from one geographical space to another (natural resources, amenities, etc.) but others, such as physical capital, human capital or technology, can. Therefore, the analysis of the distribution of the population in space is an extremely interesting question. From the point of view of urban economics, the attention given to the study of city size is particularly notable and it is focused on the analysis of the concentration of the labor factor in certain areas.

Two laws have been widely considered: Zipf's and Gibrat's. The first refers to the city size distribution and the second to population growth (see Eeckhout, 2004, for a good description of both empirical regularities, although we show the content of both laws below). Recently Rose (2006) analyzed whether other phenomena associated with population size, such as the number of inhabitants of countries, also follow a characteristic distribution, and conclude that the size distributions of cities and of countries are similar[1]. Rose's paper suggests that both laws are associated with different stocks of inhabitants. In this context, in our work, we analyze the distribution of the number of immigrants by countries from the perspective both of stock and of the percentage of this stock over the total population of the country, the migration density.

The recent evolution of migratory flows has led to a considerable growth in the stock of immigrants in some cities and countries. Therefore, given that the total population is the sum of natives and immigrants, it is useful to analyze whether Zipf's and Gibrat's laws hold for both

[1] A new critique of Rose's work is Gonzalez-Val and Sanso-Navarro (2009).

groups. In fact, because the native population grows due to natural causes, it is the evolution of the immigrant population which determines the final size of cities or countries. We can represent the growth of the stock of immigrants (M) in country i by the following function:

$$\frac{M_{it}}{M_{it-1}} = f\left(a_{+}\left(M_{it-1}\right), b_{-}\left(M_{it-1}\right)\right),$$

where $a_{+}\left(M_{it-1}\right)$ and $b_{-}\left(M_{it-1}\right)$ represent, respectively, the positive and negative external effects[2] of the stock of immigrants on growth rate. The positive external effects are associated with the so-called scale effect and the social network effect, while the negative external effect is related to the negative influence that the stock of immigrants has on migratory flows according to traditional theoretical models, for instance through wages. Therefore, the net sign of the combined effect would be a priori undetermined.

In this context, the determinants of migration are of key importance. The traditional model of Harris and Todaro (1970) predicts that migrations will disappear due to the mobility of the factors leading to convergence in expected wages between countries. However, empirical studies do not support this conclusion, as shown by Ghatak and Wheatley (1996). Authors such as Carrington et al. (1996) indicate the importance of the presence of social networks, that is, the existence of a stock of immigrants in the host country prior to the individuals' decisions to emigrate. The stock of immigrants reduces the cost of emigration, increasing the rate of migration. Additionally, authors such as Larramona and Sanso (2006) show that the differences that exist between countries do not always disappear in the long term, so the convergence achieved is limited or conditional because it does not necessarily imply the equalization of per capita income, the capital/labor ratio and of wages. Thus, the final result of the size distribution of the immigrant population is an open question which may be important in order to explain the size countries.

Another useful perspective for this type of analysis is that adopted by Alesina et al. (2000) and Spolaore and Wacziarg (2005). These authors find empirical evidence in favor of the so-called scale effect, that is, that countries with larger populations or GDPs have a larger potential market and their income present greater growth rates. In this context, the migratory stock contributes to increasing the market potential and would have positive effects on the productivity of the country, partly canceling out the tendency to the equalization of wages predicted by the traditional models. This perspective introduces interesting elements related to the effects at the aggregate level of immigration from the point of view of developed economies which present lower birth rates, because of immigration increases the market potential.

Thus, it is interesting to study if there are variations in the distribution of the stock of immigrant population. This is because the mobility of the labor factor is usually associated with differences in socioeconomic characteristics such as wages and the previous stock of immigrants. From this point of view, the flows of the labor factor tend to equalize the labor conditions of the different countries, so there would be some decrease in the stocks of immigrants and a tendency to a lower migration density.

Furthermore, migration is a phenomenon closely related to the job market, because it increases the labor supply in the host country. However, labor mobility has more restrictions than capital mobility. Until the mid 20th century, most of these restrictions were imposed by the transport technology. Since then, the reduction of these costs has been enormous. This decrease has recently been counterbalanced in many countries with the rise of protectionist legislation directed at foreigners. So, the cost is decreasing but the legislation has an influence in the opposite direction, and the question is: has the distribution of the size of the stock of immigrants become more uneven or has there been some convergence? This is the topic discussed in this work.

The result is interesting because it can be considered as evidence in favor of some kind of migration models. That it is to say, a more uneven distribution of migrations stock means that

[2] We call these effects "external" because the present stock of immigrants influences the decisions of other immigrants to emigrate later. But the effects are "internal" in geographical terms, given that the equation shown describes the dynamics in the growth of the stock of immigrants in each country.

the migration rate does not tend to zero and the migration model with permanent migration flows would be more adequate. If the result is the opposite, the traditional migration model must be used again.

To answer it, the article is structured as follows. In Section 2 there is a descriptive analysis. In Section 3, the results relative to Zipf's and of Gibrat's laws are presented. Finally, the last section concludes.

2. DATA AND DESCRIPTIVE ANALYSIS

The data correspond to the total stock of immigrants by country and the source is the Department of Economic and Social Affairs of the Population Division of the United Nations (2004). The Population Division maintains a data bank on international migration statistics covering most countries of the world. The data bank includes information from censuses on the number of foreign-born individuals or, in some cases, the foreign population living in a country. These data provide the basis for estimating the number of international migrants in the world at different points in time.

Four types of data are used by the Population Division to obtain the estimate of the migrant stock. For most countries, the estimate of the migrant stock is derived from data classified by place of birth, so it represents the foreign-born population. Sometimes it is derived from data classified by citizenship, so it represents the foreign population. In some cases, there was no data on either the foreign or the foreign-born population for the country or area concerned, so the estimate is an assigned value (this situation occurs in very few countries and correspond to countries which suffer wars in their territories[3]). And finally, the Population Division deals with the issue of refugees (which are an important percentage of the population of some African and Asian countries) considering the number of refugees as reported by UNHCR or UNRWA. These data were added to the estimates of the international migrant stock for developing countries where they were likely not to have been included in the census data available.

The sample is formed by the 214 United Nations member countries[4]. The period considered is from 1960 to 2000, presenting information by decades. The data on the total population of the countries was taken from the same source.

[3] We have repeated the analysis excluding the African countries, characterised by an unstable political situation, and this has no effect on the qualitative outcomes of our analysis.

[4] Including the former USSR as a single country, because the disintegration of the USSR in 1991 produces a transformation of internal migrants into international migrants generating data discontinuity. The former USSR includes 15 countries: Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Republic of Moldova, Russian Federation, Tajikistan, Turkmenistan, Ukraine, and Uzbekistan.

Table 1. Descriptive Analysis

Panel a: Total Stock

Area	Stock of Immigrants				
	1960	1970	1980	1990	2000
ASIA	29,280,680	28,103,771	32,312,541	41,754,291	43,761,383
EUROPE	14,015,392	18,705,244	22,163,201	26,346,258	32,803,182
NORTHERN AMERICA	12,512,766	12,985,541	18,086,918	27,596,538	40,844,405
AFRICA	8,977,075	9,862,987	14,075,826	16,221,255	16,277,486
LATIN AMERICA AND THE CARIBBEAN	6,038,976	5,749,585	6,138,943	7,013,584	5,943,680
OCEANIA	2,134,122	3,027,537	3,754,597	4,750,591	5,834,976
TOTAL	75,900,698	81,527,177	99,783,096	154,005,048	174,933,814

Panel b: Stock of Immigrants Growth

Area	Stock of Immigrants Growth (%)				
	1960-1970	1970-1980	1980-1990	1990-2000	1960-2000
ASIA	-4.02	14.98	29.22	4.81	49.45
EUROPE	33.46	18.49	18.87	24.51	134.05
NORTHERN AMERICA	3.78	39.29	52.58	48.01	226.42
AFRICA	9.87	42.71	15.24	0.35	81.32
LATIN AMERICA AND THE CARIBBEAN	-4.79	6.77	14.25	-15.25	-1.58
OCEANIA	41.86	24.01	26.53	22.83	173.41
TOTAL	7.41	22.39	54.34	13.59	130.48

Panel c: Migration Density

Area	Migration Density (%)				
	1960	1970	1980	1990	2000
ASIA	1.76	1.34	1.25	1.35	1.21
EUROPE	3.30	4.08	4.59	5.28	6.42
NORTHERN AMERICA	6.13	5.60	7.06	9.73	12.93
AFRICA	3.24	2.76	3.00	2.61	2.05
LATIN AMERICA AND THE CARIBBEAN	2.77	2.02	1.70	1.59	1.14
OCEANIA	13.43	15.57	16.45	17.80	18.80
TOTAL	2.51	2.21	2.25	2.93	2.88

A first analysis consists of describing the evolution of the stocks of immigrants and of the migration density. Panel a of Table 1 shows the total stock by geographical areas and panel b its growth in the period 1960-2000. The most relevant point is the rise in the number of immigrants, which increased by 130.48% during these 40 years. There was a particularly marked increase in North America and Oceania, while Europe is slightly above the mean. In Latin America and the Caribbean the stock has decreased. Thus, the evolution of foreign population by country is not homogeneous.

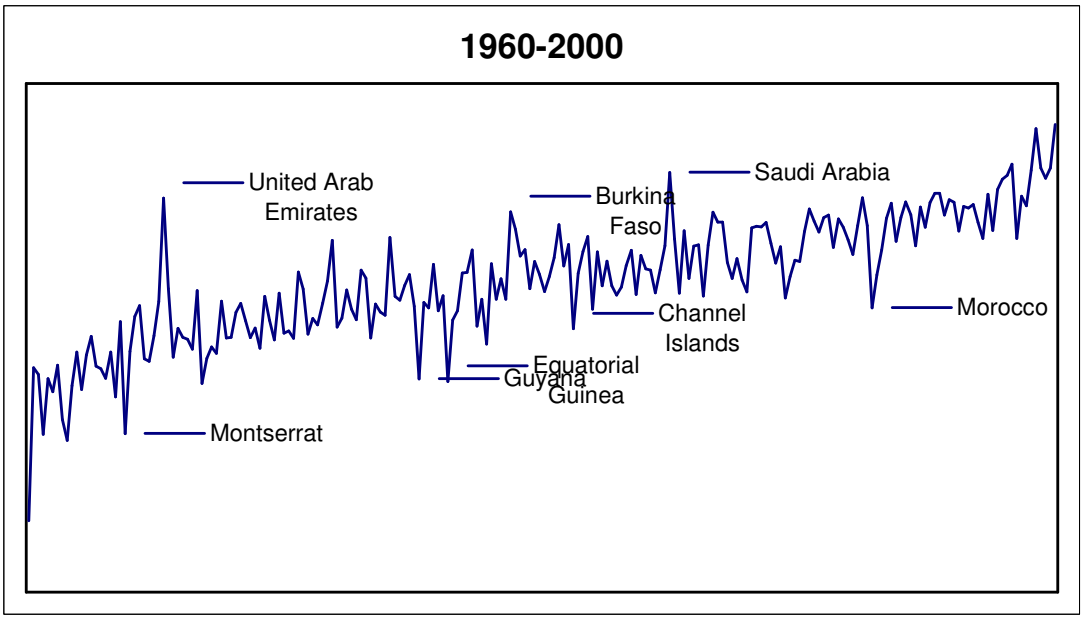
If we look at the migration density, panel c of Table 1, we see that this impression is corroborated. The variable grows in North America, in Oceania and, to a lesser extent, in Europe, while it decreases in Africa and Asia. So, we can conclude that there has been a change in the behavior of international migration.

Returning to the growth rates of the stocks of immigrants, (panel b of Table 1), we can also point out that the rate never reaches the maximum observed in each area in the last decade (1990-2000), so it appears that total immigration has not increased notably. A clear example is Europe, where the rate grows faster in the decade 1960-1970 than in 1990-2000. This fact, together with the increased migration density, which rose 3%, indicates that the birth rate of European countries is responsible for this situation.

Having determined the evolution of the stock worldwide, we may wonder whether its distribution has varied if we consider the countries individually. This is a relevant question

because the countries which present a greater foreign population may not be the same in different sample periods. If the size distribution of the stock of immigrants has not remained stable, there will be countries rising in the ranking while others fall. We follow the non-parametrical approach used by Quah (1997) and Ioannides and Overman (2001) and Figure 1 shows the evolution of the ranking.

Figure 1. Changes in the ranking of Stock of Immigrants: 1960-2000



On the horizontal axis the 214 countries are ordered from smaller to larger according to the ranking of 1960, while the vertical axis shows $\ln\left(M_{i2000} \Big/ \sum_i M_{i2000} \right)$, where M_i is the stock of immigrants in country i in the year 2000. The peaks and troughs show the countries which rise or fall in the ranking in the period being considered, while the changes in the slope inform us of variations in the inequality of distribution. We see that the countries which experience the greatest variations in the ranking are some African countries (Morocco, Equatorial Guinea or Burkina Faso) and some Arab countries (Saudi Arabia or United Arab Emirates), while the most developed countries do not present relatively high variations.

Additionally, the slope tends to be constant, so we can conclude that there have not been significant changes in the unevenness of the distribution. To corroborate this impression, Table 2 presents Spearman's rank correlation coefficients for the data corresponding to each possible pair of decades and for the total period for the stock of immigrants and migration density, both for the total sample and for the top 50 countries with the highest population at the beginning and the end of the period analyzed.

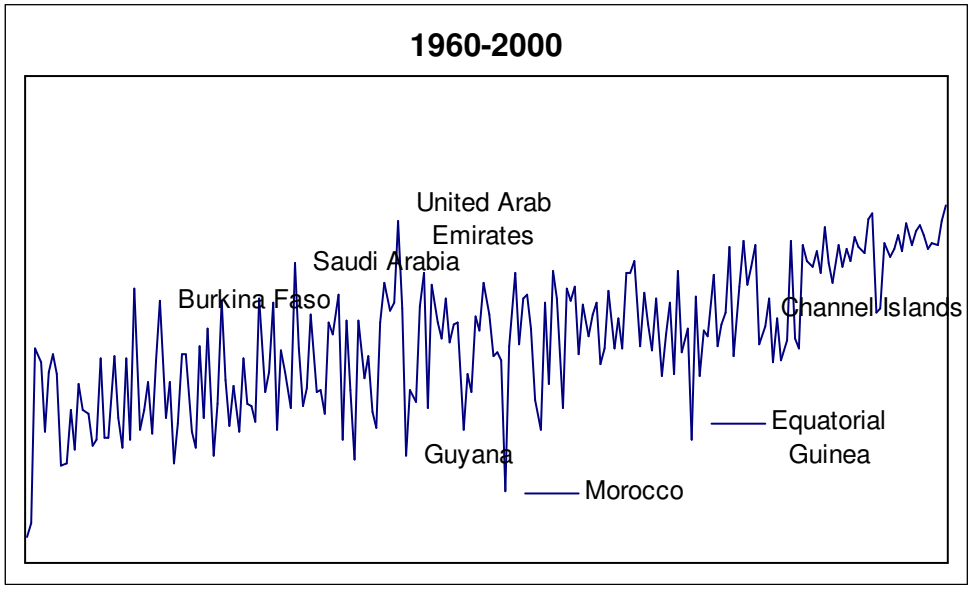
Table 2. Spearman Rank Coefficients

	All countries (214)				Top 50 (population in 1960)				Top 50 (population in 2000)			
Migrations	1970	1980	1990	2000	1970	1980	1990	2000	1970	1980	1990	2000
1960	0.97	0.91	0.89	0.87	0.95	0.87	0.83	0.82	0.93	0.77	0.71	0.54
1970		0.94	0.92	0.91		0.94	0.90	0.89		0.85	0.78	0.63
1980			0.97	0.94			0.98	0.94			0.88	0.73
1990				0.97				0.98				0.81
Migration Density	1970	1980	1990	2000	1970	1980	1990	2000	1970	1980	1990	2000
1960	0.93	0.82	0.77	0.72	0.80	0.75	0.65	0.51	0.83	0.67	0.62	0.51
1970		0.92	0.87	0.83		0.90	0.74	0.62		0.89	0.75	0.64
1980			0.94	0.89			0.91	0.79			0.89	0.77
1990				0.95				0.89				0.88
Growth	All countries (214)				Top 50 (growth in 1960)				Top 50 (growth in 2000)			
	70-80	80-90	90-00		70-80	80-90	90-00		70-80	80-90	90-00	
1960 - 1970	0.47	0.15	0.15		0.28	0.00	-0.11		0.19	-0.29	-0.27	
1970 - 1980		0.32	0.23			0.39	0.35			-0.28	-0.05	
1980 - 1990			0.30				0.12				-0.01	

The Spearman coefficient of the stock of immigrants and migration density is clearly positive in the three sub-samples considered, so we can conclude that the ranking of countries according to criteria of migration has not varied significantly. If we order by rates of variation of stock or growth, although the coefficient continues to be positive if all the countries are considered, its value clearly decreases, so we can affirm that these are more noticeable changes. Furthermore, if the sub-sample selected corresponds to the countries with higher growth in 2000, a negative coefficient appears, which would support a certain degree of convergence between these countries. Thus, in conclusion, we can affirm that, although there have been changes in the order of the countries regarding the growth of stocks, these changes have not been enough to change the order of the total volume of immigrants in the second half of the 20th century.

Regarding the migration density, in Figure 2, the order of 1960 is compared with that of 2000, presenting a profile with greater variability than that of the absolute stock, although the countries with greater variations coincide. In this case, a slight change in the slope appears, which indicates increasing inequality.

Figure 2. Changes in the ranking of Migration Density: 1960-2000



3. ZIPF'S LAW AND GIBRAT'S LAW

The aim of this work is to analyze the structure of the size distribution of the immigrant population. In this section, we present an analysis of two laws traditionally associated with the populations of cities and, recently, with the size of nations. These two laws can be studied from the point of view both of volume and of migration density.

3.1. ZIPF'S law

In this section, we examine the size distribution of stock and of migration density in order to see if there has been convergence or divergence between the different countries of the world. To do this, we use Pareto's distribution (1896), also known as power law. If M denotes the stock of immigrants and R the rank, a power law links both as follows:

$$R(M) = aM^{-b}, \qquad (1)$$

where a and b are parameters. This expression has been used extensively in urban economics to study the size distribution of cities (see, for example, Eeckhout (2004) and Ioannides and Overman (2003)). It has also been used recently to study the size distribution of countries (Rose, 2006), and a theoretical justification as to why the population of cities follows this distribution can be found in Eeckhout (2004) and Duranton (2007).

A particular case which is widely known as Zipf's law (1949), which appears when $b = 1$ and which means that, ordered from larger to smaller, the stock of emigrants of the first country is twice the stock of the second, three times the stock of the third, and so on (the product between rank and size of stock is constant). Another empirical regularity related to Zipf's law and Pareto's distribution is Gibrat's law (1931), which postulates that the growth of the variable is a random variable independent of its initial size. However, both Eeckhout (2004) and Duranton (2007) demonstrate that there is a possibility that only the upper tail adapts to this distribution and that, when the total sample is considered, the distribution which fits best is the lognormal. In this work, we contrast these results for the stock of immigrants.

It is also interesting to test whether Pareto's parameter is more or less than 1, and the evolution of this coefficient in time, that is, for the different cross-sections available. Effectively, the greater the coefficient, the more homogeneous are the stocks of immigrants by countries. Also, a growing evolution would mean a process of convergence in the immigrant stock. And the opposite, the smaller the coefficient the less homogeneous are the stocks of immigrants by countries, and a decreasing evolution would mean a process of divergence in the immigrant stock.

More specifically, the expression (1) of Pareto's distribution is usually estimated in its doubly logarithmic version:

$$\ln R = \ln a - b \cdot \ln M \ . \qquad (2)$$

Different sample sizes have been used, considering the top-50, 100 or 150 countries with the biggest stock of immigrants, and the estimation has also been made with the total countries[5]. Table 3 presents the results of the OLS[6] estimation (the values of the corrected standard error are shown in brackets[7]).

[5] We also consider the possibility of differentiating immigrants by sex. The estimations made show that differentiated behaviours do not exist.

[6] Gabaix and Ioannides (2004) show that the Hill (maximum likelihood estimator) is more efficient if the underlying stochastic process is really a Pareto distribution. As we show below, this is not the distribution that the data follow, and so we use the OLS estimator.

[7] The residues resulting from this regression usually present problems of heteroskedasticity so, to analyze the significance of the parameters, the typical corrected deviation proposed by Gabaix and Ioannides (2004) is used: $\text{GI s.e.} = \hat{b} \cdot (2/N)^{1/2}$, where N is the sample size.

Table 3. Pareto coefficients by decade for the Stock of Immigrants estimated by OLS

	Top 50			Top 100		
Year	b < 0	(GI s.e.)	R^2	b < 0	(GI s.e.)	R^2
1960	0.966	0.193	0.981	0.641	0.091	0.915
1970	0.939	0.188	0.973	0.685	0.097	0.935
1980	1.035	0.207	0.983	0.719	0.102	0.931
1990	0.925	0.185	0.982	0.726	0.103	0.947
2000	0.939	0.188	0.981	0.743	0.105	0.952
	Top 150			All (214)		
Year	b < 0	(GI s.e.)	R^2	b < 0	(GI s.e.)	R^2
1960	0.523	0.060	0.909	0.333	0.032	0.790
1970	0.551	0.063	0.916	0.348	0.033	0.787
1980	0.572	0.066	0.910	0.354	0.034	0.780
1990	0.580	0.067	0.921	0.352	0.034	0.780
2000	0.569	0.066	0.912	0.343	0.033	0.770

(GI s.e.) Gabaix-Ioannides (2004) corrected standard error.

All coefficients are significantly different from zero at the 0.05 level.

The estimation of the parameter is significantly different to one except when we consider only the 50 largest stocks of immigrants. It is very close to one in the upper tail of the distribution, obtaining a good adjustment level (R^2 oscillates between 0.97 and 0.98), while, as the size of the sample increases, the estimated value of b and the degree of adjustment evidently decrease. We also observe that the value of the coefficient increases over time, so some convergence can be detected with a Pareto distribution, especially when we consider the 100 countries with the biggest stock of immigrants. Moreover, the Gini coefficient indicates that the distribution is very uneven and remains almost constant throughout the period examined (see Table A1 for some measures of concentration).

Table A1. Concentration measures for the Stock of Immigrants

Year	Sample Size	Herfindahl	Normalized Herfindahl	Gini Coefficient
2000	All (214)	0.079	0.075	0.846
	Top 150	0.080	0.073	0.787
	Top 100	0.082	0.073	0.712
	Top 50	0.097	0.078	0.610
1990	All (214)	0.074	0.070	0.845
	Top 150	0.074	0.068	0.785
	Top 100	0.077	0.068	0.712
	Top 50	0.091	0.072	0.608
1980	All (214)	0.045	0.041	0.822
	Top 150	0.045	0.039	0.753
	Top 100	0.047	0.038	0.667
	Top 50	0.056	0.036	0.520
1970	All (214)	0.046	0.042	0.836
	Top 150	0.046	0.040	0.772
	Top 100	0.048	0.038	0.691
	Top 50	0.056	0.036	0.550
1960	All (214)	0.051	0.047	0.845
	Top 150	0.052	0.045	0.784
	Top 100	0.053	0.043	0.703
	Top 50	0.061	0.041	0.554

However, the Hill (maximum likelihood) estimator could be more efficient than the OLS estimator in the upper tail, as shown by Gabaix and Ioannides (2004), so Table 4 presents the results of the Hill estimator[8] in the upper tail (Top 50 and Top 100 countries). Although the estimated coefficients are slightly different from the OLS estimations, for the Top 50 they are not significantly different from one at the 5% significance level.

Table 4. Pareto coefficients in the upper tail obtained using Hill's estimator

Stock of Immigrants

	Top 50		Top 100	
Year	Hill	(Hill s.e.)	Hill	(Hill s.e.)
1960	0.830	0.117	0.489	0.049
1970	0.880	0.124	0.548	0.055
1980	0.775	0.110	0.569	0.057
1990	0.891	0.126	0.586	0.059
2000	0.983	0.139	0.554	0.055

Migration Density

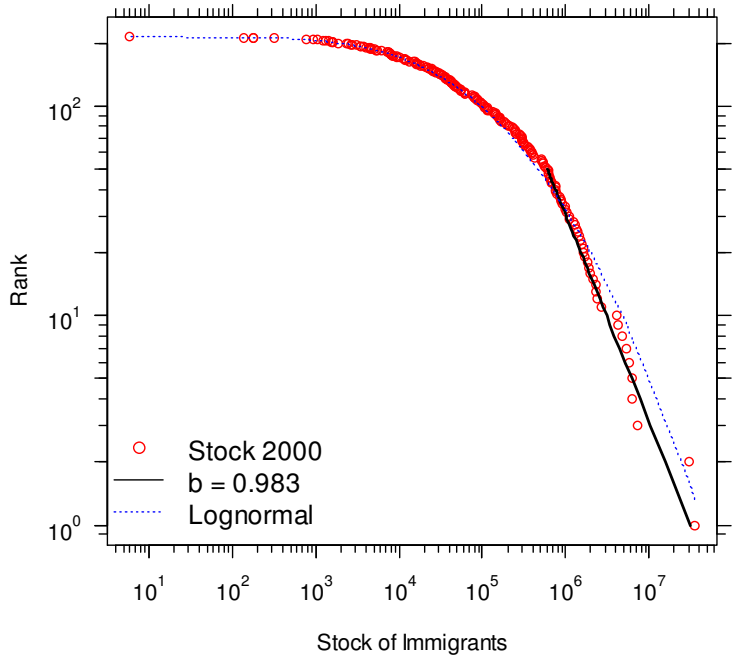
	Top 50		Top 100	
Year	Hill	(Hill s.e.)	Hill	(Hill s.e.)
1960	1.133	0.160	0.961	0.096
1970	1.086	0.154	0.908	0.091
1980	1.188	0.168	0.838	0.084
1990	1.159	0.166	0.852	0.085
2000	1.208	0.171	0.825	0.083

All coefficients are significantly different from zero at the 0.05 level.

The presence of a decreasing Zipf coefficient with the sample size, as shown in Eeckhout (2004), may be because distribution is lognormal, and so, on selecting the countries with higher stock only the upper tail is being considered, which is a good approximation to a Pareto distribution. In fact, for our migration data the Pareto distribution provides a better fit for the upper tail than the lognormal. Figure 3 shows the Rank-Size plot corresponding to the year 2000 (the graphs for the other years are very similar), and we can see that, although the lognormal has a good fit for most of the distribution, for the upper tail, the Pareto fitted by maximum likelihood is closer to the data.

[8] Calculated using the PARETOFIT stata module developed by Jenkins and Van Kerm (2007) available online at http://ideas.repec.org/c/boc/bocode/s456832.html.

Figure 3. Rank-size plot (log scale) for the migrant stock (year 2000)



Note: Top 50 data are fitted by a power-law whose exponent is estimated as b = 0.983 (±0.272) by using Hill's estimator. Also shown is the fit by log-normal distribution for the entire range based on the maximum likelihood estimation.

In order to test whether the distribution is lognormal throughout the sample, in Figure 4, we present the adaptive kernels which represent the estimated distribution (the scale is the same in all the figures in order to make comparisons easier, the y axis represents the estimated density or probability and the x axis shows the logarithm of the stock of immigrants). We observe an approximation to lognormal distribution in Figure 4. Additionally, the Kolgomorov-Smirnov normality test, Table 5, shows that this hypothesis is never rejected, providing evidence against Zipf's law when the entire sample is considered.

Figure 4. Adaptive kernels of the Stock of Immigrants (ln scale)

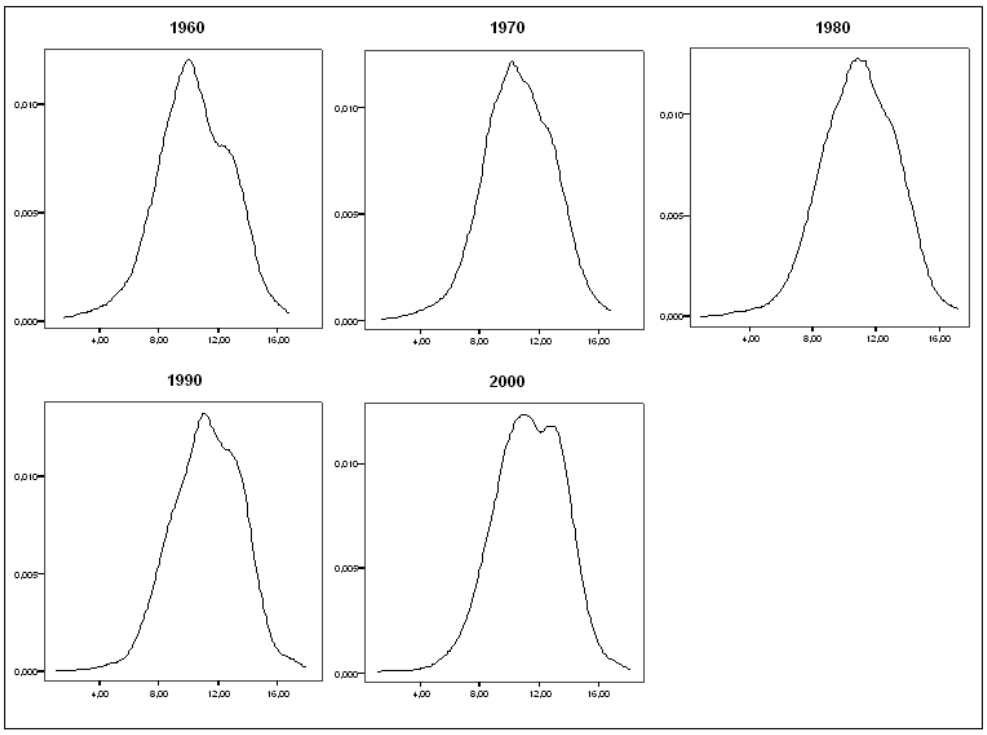


Table 5. Kolmogorov-Smirnov test for the Stock of Immigrants (ln scale)

	1960	1970	1980	1990	2000
Kolmogorov-Smirnov Z	0.617	0.464	0.544	0.696	0.597
Asymp. Sig. (2-tailed)	0.841*	0.982*	0.929*	0.718*	0.868*

* Normality null hypothesis not rejected at the 0.05 level

Finally, we analyze the distribution of migration density. If we consider the earlier results referring to the stock of immigrants, as well as those obtained by Rose (2006) for the total population, it is to be expected that the difference between the logarithms of the two variables follows a lognormal distribution. The results of the OLS estimation of the Pareto exponent are shown in Table 6, while Table 4 displays the results for the upper tail using the Hill estimator.

Table 6. Pareto coefficients by decade for Migration Density estimated by OLS

	Top 50			Top 100		
Year	b < 0	(GI s.e.)	R^2	b < 0	(GI s.e.)	R^2
1960	1.327	0.265	0.916	1.118	0.158	0.944
1970	1.363	0.273	0.909	1.071	0.152	0.931
1980	1.449	0.290	0.877	1.047	0.148	0.898
1990	1.654	0.331	0.904	1.106	0.156	0.893
2000	1.713	0.343	0.890	1.071	0.151	0.874
	Top 150			All (214)		
Year	b < 0	(GI s.e.)	R^2	b < 0	(GI s.e.)	R^2
1960	0.941	0.109	0.935	0.558	0.054	0.775
1970	0.874	0.101	0.915	0.535	0.052	0.773
1980	0.843	0.097	0.895	0.541	0.052	0.794
1990	0.838	0.097	0.872	0.521	0.050	0.776
2000	0.804	0.093	0.862	0.499	0.048	0.771

(GI s.e.) Gabaix-Ioannides (2004) corrected standard error.
All coefficients are significantly different from zero at the 0.05 level.

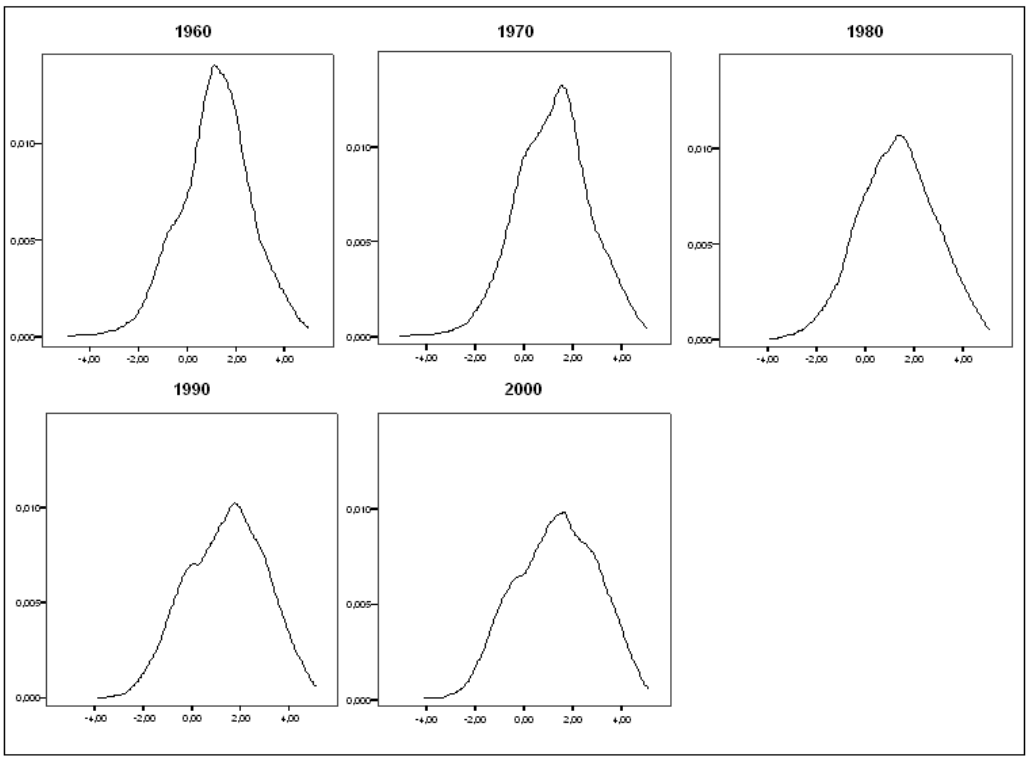
Again, if we take the whole sample, the distribution is uneven, although not as much as in the case of the stock of immigrants. We reject Zipf's law, and the concentration indexes (see Table A2) indicate a slight tendency to divergence between countries the estimated parameter tends to decrease, which means that inequality has increased.

Table A2. Concentration measures for Migration Density

Year	Sample Size	Herfindahl	Normalized Herfindahl	Gini Coefficient
2000	All (214)	0.015	0.010	0.667
	Top 150	0.016	0.009	0.555
	Top 100	0.018	0.008	0.440
	Top 50	0.025	0.005	0.275
1990	All (214)	0.015	0.010	0.659
	Top 150	0.016	0.009	0.547
	Top 100	0.018	0.008	0.440
	Top 50	0.026	0.006	0.290
1980	All (214)	0.016	0.011	0.669
	Top 150	0.017	0.010	0.565
	Top 100	0.019	0.009	0.466
	Top 50	0.027	0.007	0.319
1970	All (214)	0.017	0.013	0.668
	Top 150	0.018	0.011	0.569
	Top 100	0.021	0.011	0.481
	Top 50	0.029	0.009	0.350
1960	All (214)	0.017	0.012	0.654
	Top 150	0.018	0.011	0.554
	Top 100	0.021	0.011	0.474
	Top 50	0.030	0.011	0.364

The graphic representation of the adaptive kernel of migration density, Figure 5, also shows an evolution towards a lognormal distribution, starting from a very leptokurtic distribution in 1960 (again the scale is the same in all the figures in order to permit comparisons). The centre of distribution has lost importance compared to the tails, which indicates that growth was not convergent.

Figure 5. Adaptive kernels of Migration Density (ln scale)



However, the Kolgomorov-Smirnov test, Table 7, shows that the hypothesis of normality is not rejected any decade, providing evidence against Zipf's law when considering all the countries of the sample.

Table 7. Kolmogorov-Smirnov test for Migration Density (ln scale)

	1960	1970	1980	1990	2000
Kolmogorov-Smirnov Z	0.689	0.567	0.469	0.611	0.679
Asymp. Sig. (2-tailed)	0.730*	0.905*	0.980*	0.849*	0.745*

* Normality null hypothesis not rejected at the 0.05 level

Therefore, we can conclude that both the stock of immigrants and migration density follow similar distributions to those found in Eeckhout (2004) and Rose (2006) for the size of US cities and of countries, respectively, confirming the presence of an empirical stylized fact when the spatial distribution of the population is considered. Moreover, given the value of the concentration indexes of the estimated coefficient of the Pareto distribution and that Spearman's coefficient is high, we can affirm that no sign of convergence appears in the stock of immigrants and in migration density. This may be considered as evidence in favor of the theoretical models which find long term migration rates different to zero, recognizing the presence of factors such as technology or social costs, which compensate the effects of traditional factors such as income or wages. However, the question of convergence must be analyzed in the framework of Gibrat's law, which we do in next section.

3.2. GIBRAT'S law

The above section has shown that there is a certain stability in the distribution of the immigrant population size and in migration density, although a certain tendency towards divergence was observed in the latter. However, for a more rigorous dynamic analysis we have to use growth rates. In particular, we were interested in verifying whether Gibrat's law holds or not, meaning that the growth of a population variable does not depend on the initial situation[9]. In parametric terms, this relationship between growth and size is usually estimated as:

$$\frac{Y_{t+1}}{Y_t} = K + a \cdot Y_t, \qquad (3)$$

where Y_t is the corresponding variable (in our case, stock of immigrants or migration density) and K is a constant. If the parameter a is not significant, we can conclude that the growth is independent of the initial level of the variable.

[9] Gibrat (1931) observed that the distribution of size (measured by sales or the number of employees) of firms tends to be lognormal, and his explanation was that the growth process of firms can be multiplicative and independent of the size of the firm.

Table 8. Results of estimations of parametrical growth regressions

Stock of Immigrants parametrical growth regressions

		All countries (214)			Top 100		
Initial Year	Final Year	a < 0	(s.e.)	R^2	a < 0	(s.e.)	R^2
1960	1970	-1.45E-07	1.27E-07	0.006122	-4.96E-08	4.34E-08	0.013158
1960	1980	-1.04E-06	1.53E-06	0.00217	-2.09E-07	2.08E-07	0.01022
1960	1990	-1.65E-06	3.00E-06	0.001434	-3.27E-07	4.45E-07	0.005471
1960	2000	-2.00E-06	3.70E-06	0.001388	-3.69E-07	5.42E-07	0.004705
1970	1980	-3.31E-07	5.17E-07	0.001926	-9.23E-08	9.23E-07	0.010095
1970	1990	-2.64E-07	2.08E-07	0.005981	-1.02E-07	1.91E-07	0.002884
1970	2000	-4.40E-07	4.07E-07	0.005473	-1.78E-07	2.49E-07	0.00516
1980	1990	-2.37E-08	7.41E-08	0.00048	-2.72E-08	5.45E-08	0.002527
1980	2000	-9.56E-08	1.91E-07	0.001211	-2.58E-09	6.70E-08	0.000015
1990	2000	-9.40E-09	1.57E-08	0.001699	-9.61E-09	1.54E-03	0.003975

None of the coefficients is significantly different from zero at the 0.05 level.

Migration Density parametrical growth regressions

		All countries (214)			Top 100		
Initial Year	Final Year	a < 0	(s.e.)	R^2	a < 0	(s.e.)	R^2
1960	1970	-1.46E-02	1.28E-02	0.006122	-4.99E-03	4.37E-03	0.013158
1960	1980	-1.04E-01	1.54E-01	0.00217	-2.10E-02	2.09E-02	0.010122
1960	1990	-1.66E-01	3.01E-01	0.001434	-3.28E-02	4.47E-02	0.005471
1960	2000	-2.01E-01	3.71E-01	0.001388	-3.71E-02	5.45E-02	0.004705
1970	1980	-3.32E-02	5.21E-02	0.001926	-9.28E-03	9.28E-03	0.010095
1970	1990	-3.66E-02	3.25E-02	0.005981	-1.02E-02	1.92E-02	0.002884
1970	2000	-4.42E-02	4.10E-02	0.005473	-1.79E-02	2.50E-02	0.00516
1980	1990	-2.38E-03	7.48E-03	0.00048	2.73E-03	5.48E-03	0.002527
1980	2000	-9.61E-03	1.90E-02	0.001211	-2.59E-04	6.73E-03	0.000015
1990	2000	-9.45E-04	1.58E-03	0.001699	-9.66E-04	1.54E-03	0.003975

None of the coefficients is significantly different from zero at the 0.05 level.

Table 8 shows the results of the OLS estimation of (3) for all possible combinations of sample periods considered, both for the overall sample and for the top-100 countries with the greatest stock of immigrants or migration density. The first conclusion is that all the coefficients are negative, which would imply an inverse relationship between size and growth. Thus, countries with larger initial stocks present less growth, producing convergence. But the second conclusion is that all the estimated coefficients are non-significant, meaning that the null hypothesis can not be rejected and Gibrat's law is accepted.

Quah (1993) indicates that this type of parametrical analysis so habitually used to study economic growth is too simple, because the problems of regressions towards the mean, and proposes the use of non-parametrical methods, precisely, transition matrices. We will use the methodology followed by Ioannides and Overman (2003) and Eeckhout (2004). It consists of taking the following specification:

$$g_i = m(S_i) + \varepsilon_i , \qquad (4)$$

where g_i is the normalized growth rate (subtracting the mean and dividing by the standard deviation) and S_i is the logarithm of the stock of immigrants, and instead of making suppositions about the functional relationship of m and supposing a linear relationship, as in equation (3), $\hat{m}(s)$ is estimated as a local average around point s and is smoothed using a kernel, which is a symmetrical, weighted and continuous function around s.

In order to analyze the period 1960-2000, the Nadaraya-Watson method is used, exactly as it appears in Härdle (1990), based on the following expression[10]:

[10] The calculation was done with the KERNREG2 stata module, developed by Cox et al. (1999), available online at http://ideas.repec.org/c/boc/bocode/s372601.html. This program is based on the algorithm described by Härdle (1990) in Chapter 5.

$$\hat{m}(s) = \frac{n^{-1}\sum_{i=1}^{n} K_h(s - S_i) g_i}{n^{-1}\sum_{i=1}^{n} K_h(s - S_i)}, \qquad (5)$$

where K_h denotes the dependence of the kernel K (in this case an Epanechnikov kernel) on the bandwidth h (0.5). Starting from this calculated mean $\hat{m}(s)$, the variance of the growth rate g_i is also estimated, again applying the Nadaraya-Watson estimator starting from:

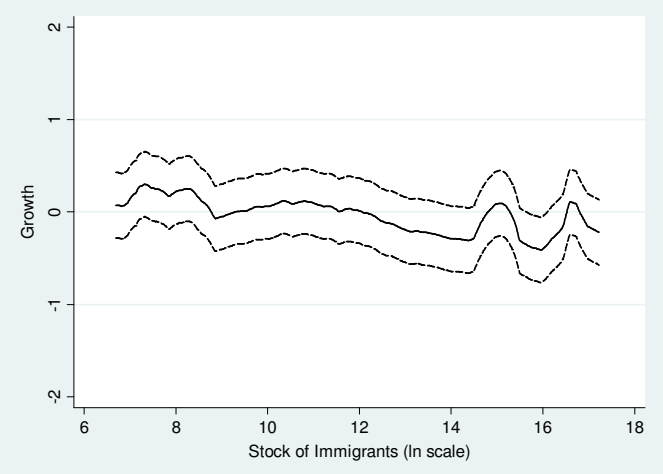
$$\hat{\sigma}(s) = \frac{n^{-1}\sum_{i=1}^{n} K_h(s - S_i)(g_i - \hat{m}(s))^2}{n^{-1}\sum_{i=1}^{n} K_h(s - S_i)}. \qquad (6)$$

The estimator is very sensitive, both in mean and in variance, to atypical values. Thus, we eliminate 5% of the lowest observations of the distribution, both for the stock of emigrants and for migration density, as these observations are characterized by a very high dispersion both in mean and in variance[11]. For the case of stock, four more very atypical values are eliminated[12].

If growth were independent of size, the estimated kernel would be a straight line on the zero value. Values different to zero imply deviations with respect to the mean. And variance would also be a straight line, supposing that variance does not depend on the size of the variable analyzed.

For each size of the stock of immigrants, Figure 6 represents the estimation of the mean growth and Figure 7 that of the variance of that growth. Bootstrapped 95% confidence bands are also displayed. They have been calculated using 500 random samples with replacement. For the calculation, the 809 available observations covering the entire sample period have been taken into account. Some conclusions can be highlighted from the results obtained. First, the estimated kernel of the mean is close to zero. However, the trend is slightly decreasing; the bigger the size, the smaller the growth rate[13]. But the null hypothesis of this mean being equal to zero can be rejected at a 5 % significance level only for some values in the upper tail, so, except for these values of the distribution, Gibrat's law holds in the period examined. Thus, we find evidence of slight convergence in the stock of immigrants because the biggest countries have had less mean growth.

Figure 6. Kernel estimate of Stock of Immigrants Growth 1960-2000 (Bandwidth 0.5)

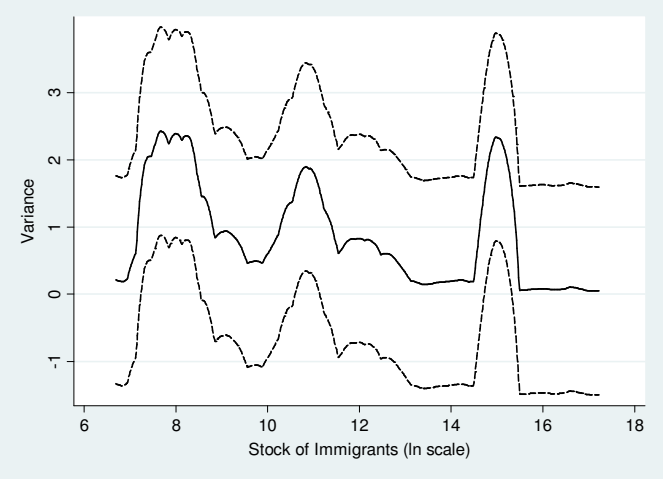


[11] The majority of these 43 excluded observations for the stock of immigrants correspond to African countries and islands which constitute independent states. In the case of migration density, Asian countries also appear (China, Vietnam, for example), which, due to their high population, have low migration density.

[12] These four observations correspond to United Arab Emirates (1960-1970), Djibouti (1970-1980), Mozambique (1970-1980) and Somalia (1970-1980).

[13] Although clusters of countries are detected that differ from the trend, both in mean and in variance (but we cannot reject to be equal to zero for the mean growth or equal to one for variance).

Figure 7. Kernel estimate of the Variance of Stock of Immigrants Growth 1960-2000 (Bandwidth 0.5)



For migration density, we have 813 observations available. Figures 8 and 9 present the kernels corresponding to the mean and the variance of its growth. Although the estimated kernel is close to zero and we cannot reject the null hypothesis of this mean being equal to zero at a 5% significance level, we observe two clearly differentiated behaviors: the countries with a lower rate have grown more slowly than those that began with a higher rate. Therefore we find divergence. And it is also observed that variance is independent of size, except for some upper tail distribution values, so evidence against Gibrat's law in countries' migration density is not found.

Figure 8. Kernel estimate of Migration Density Growth 1960-2000 (Bandwidth 0.5)

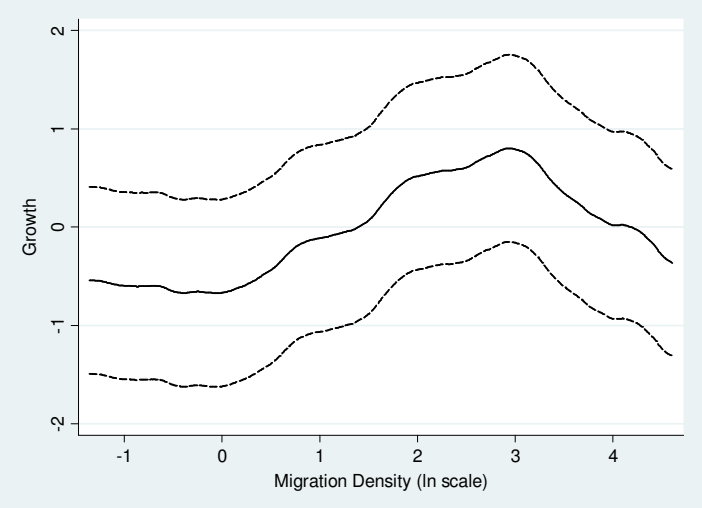
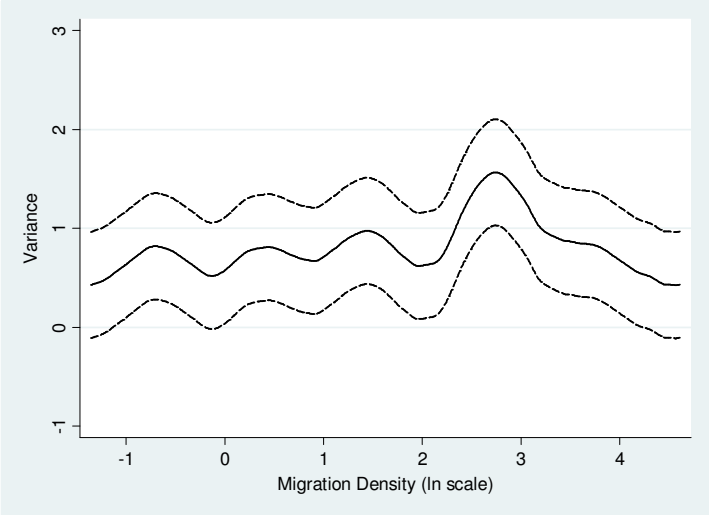


Figure 9. Kernel estimate of the Variance of Migration Density Growth 1960-2000 (Bandwidth 0.5)



If we interpret the two variables together, we can say that, although the growth of the stock of immigrants does not appear to have been especially relevant when establishing the size distribution, migration density is. This may be because the host countries have lower birth rates than the origin countries so that, while the stock of immigrants grows at the same or a similar rate, migration density increases.

This may be important for several reasons. The first is that less population growth in the host countries may be generating a scarcity in the labor supply, which creates wage differences that encourage migration. As long as these differences are maintained, migration will continue. On the other hand, the stock of immigrants reduces immigration costs and, as long as the importance of this fact is greater than the rate at which wages converge, migration will continue. Finally, in a context where a scale effect of population on economic growth is important, as discussed in Alesina (2000) and Spolaore and Wacziarg (2005), this result may be useful for understanding the existence of persistent differences in wages. Furthermore, it may help us to understand why migratory patterns are maintained and to give new perspectives on the mechanisms by which migration affects the economic growth and welfare of the host countries.

4. CONCLUSIONS

This paper studies the evolution of the worldwide distribution of the stock of immigrants, focusing on two well-known empirical regularities in urban economics, Zipf's law and Gibrat's law. We use parametric and non-parametrical methods and obtain the following results.

First, for the stock of immigrants, the estimated Pareto exponent is about 1/3 if we consider the whole sample very close to one in the upper tail of the distribution while, as the size of the sample increases, the estimated value decrease., which indicates that the distribution is very uneven. The estimation of the parameter remains almost constant throughout the period examined. Moreover, the Gini coefficient indicates that the distribution is very uneven and remains almost constant throughout the period examined.

Also, the estimated kernels show that the distribution that fits best is lognormal, while the upper tail is represented by a Pareto, a statistical regularity already shown in Eeckhout (2004) for the case of the North American cities. We show that growth is independent of size, and the non-parametric estimation corroborates this result, although we found a weak convergence in the stock of immigrants.

We have repeated the analysis for migration density, defined as the percentage of immigrants over the total population of the country. In this case, if we take the whole sample, the distribution is uneven, although not as much as in the case of the stock of immigrants, The

estimated Pareto exponent for the entire sample in this case is about ½, but presents a tendency to decrease, which would and the concentration indexes indicate a slight tendency to divergence between countries. This coincides with the results offered by the estimated kernels, which show a loss of kurtosis in the center of the distribution. Although the estimated kernel of the growth of migration density is close to zero, we observe two clearly differentiated behaviors: the countries with a lower rate have grown more slowly than those that began with a higher rate. Consequently, in the period examined, we observe a divergent behaviour. We also observe that variance is independent of size, except for some upper tail distribution values, so evidence against Gibrat's law in countries' migration density is not found.

REFERENCES

ALESINA, A., E. SPOLAORE and R. WACZIARG, (2000). Economic Integration and Political Disintegration. American Economic Review 90: 1276-1296.

CARRINGTON, W. J., E. DETRAGIACHE and F. VISHWANATH, (1996). Migration with Endogenous Moving Cost. American Economic Review, 86, 909-930.

CoX, N, I. SALGADO-UGARTE, M. SHIMIZU and T. TORU TANIUCHI, (1999). KERNREG2: Stata module to compute kernel regression (Nadaraya-Watson). http://fmwww.bc.edu/repec/bocode/k/kernreg2.ado

DURANTON, G., (2007). Urban Evolutions: The Fast, the Slow, and the Still. American Economic Review 97(1), 197-221.

EECKHOUT, J., (2004). Gibrat's Law for (All) Cities. American Economic Review, vol. 94(5), 1429-1451.

GABAIX, X., and Y. M. IAONNIDES, (2004). The evolution of city size distributions. Handbook of urban and regional economics, Vol. 4, J. V. Henderson and J. F. Thisse, eds. Amsterdam: Elsevier Science, North-Holland, 2341-2378.

GHATAK, L., and P. WHEATLEY, (1996). Migration Theories and Evidence: an Assessment, Journal of Economic Surveys, 10, 159-198.

GIBRAT, R., (1931). Les inégalités économiques, Paris: Librairie du Recueil Sirey.

GONZALEZ-VAL, R., and M. SANSO-NAVARRO, (2009). Gibrat's Law for Countries. Journal of Population Economics, forthcoming.

HÄRDLE, W., (1990). Applied nonparametric regression. Econometric Society Monographs. Cambridge, New York and Melbourne: Cambridge University Press.

HARRIS, J. R., and M. P. TODARO, (1970). Migration, Unemployment and Development: A Two-Sector Analysis. American Economic Review, 60, 126-42.

IOANNIDES, Y. M., and H. G. OVERMAN, (2001). Cross-Sectional evolution of the U.S. city size distribution. Journal of Urban Economics 49, pages 543-566.

IOANNIDES Y. M. and H. G. OVERMAN, (2003). Zipf's law for cities: an empirical examination. Regional Science and Urban Economics 33, 127-137.

JENKINS, S. P., and P. VAN KERM, (2007). PARETOFIT: Stata module to fit a Type 1 Pareto distribution. http://ideas.repec.org/c/boc/bocode/s456832.html.

LARRAMONA, G., and M. SANSO, (2006). Migration dynamics, growth and convergence. Journal of Economic Dynamic and Control, 30(11), 2261-2279.

PARETO, V. (1896). Cours d'Economie Politique. Geneva: Droz.

QUAH, D. T., (1993). Galton's Fallacy and Tests of the Convergence Hypothesis. The Scandinavian Journal of Economics, Vol. 95, No. 4, Endogenous Growth (Dec., 1993), 427-443.

QUAH, D. T., (1997). Regional cohesion from local isolated actions: Historical outcomes. Discussion paper 378, Centre for Economic Policy Research, London.

ROSE, A. K., (2006). Cities and countries. Journal of Money, Credit, and Banking, 38(8), 2225-2246.

SPOLAORE, E., and R. WACZIARG, (2005) Borders and Growth. Journal of Economic Growth 10: 331-386.

UNITED NATIONS, (2004). Trends in total migrant stock: The 2003 revision. Database maintained by The Population Division of the Department of Economic and Social Affairs.

ZIPF, G. (1949). Human Behaviour and the Principle of Least Effort. Cambridge, MA: Addison-Wesley.

LA FINANCIACIÓN DESCENTRALIZADA EN ESPAÑA: NECESIDAD DE UNA ORDENACIÓN POLÍTICA PACTADA

ALFREDO IGLESIAS SUÁREZ
E-mail: Alfredo.Iglesias@uclm.es
CRISTINA GARCÍA NICOLÁS
E-mail: Cristina.Garcia@uclm.es
Universidad de Castilla-La Mancha
Facultad de Ciencias Jurídicas y Sociales

Resumen

Con el fin de estabilizar los sistemas de financiación de los tres niveles de Gobierno plantearemos qué posibilidades existen en la España actual para elaborar algo parecido a los Pactos de la Moncloa. Además de la importante merma de ingresos locales provocada por la crisis del sector inmobiliario, consideramos imprescindible introducir en el debate la creación de las áreas metropolitanas como institución intermedia entre Municipio y Comunidad Autónoma, que mejore la cobertura y la eficiencia de los servicios públicos. En cuanto a las CCAA de Régimen Común, deberían cumplir dos objetivos fundamentales: estabilizar el sistema de financiación, aportando un fuerte grado de solvencia que lograra un nivel elevado en la cobertura de sus competencias; y dotar al modelo de un nivel de suficiencia en estrecha relación con las competencias asumidas. Al mismo tiempo nos parece que el Estado central debería ser realista a la hora de otorgar ingresos a los gobiernos subcentrales, tal como requiere un proceso de Federalismo Fiscal.

Palabras clave: Federalismo Fiscal, descentralización, Pacto Político
Área temática: Economía urbana, regional y local

Abstract

In order to stabilize the systems of funding of three levels of Government we will raise what possibilities exist in the current Spain to elaborate something similar to the Moncloa Agreements. Besides the important decrease of local income caused by the crisis of the real-estate sector, we consider indispensably to introduce in the debate the creation of the metropolitan areas as intermediate institution between Municipality and Autonomous Community to improve the coverage and the efficiency of the public services. As for the CCAA of Common Regime, they should comply with two fundamental aims: to stabilize the system of funding, contributing a strong degree of solvency that was achieving a high level in the coverage of its powers; and to endow the model with a level of sufficiency in close relation with the assumed powers. At the same time it seems to us that the central State should be realistic at the moment of granting income to the subcentral governments, as there needs a process of Fiscal Federalism.

JEL classification: H70, H71, H74, H77
Key words: Fiscal federalism, decentralization, Political Agreement
Thematic Area: Local, regional and urban economy.

1. INTRODUCCIÓN Y FINALIDAD

Con este trabajo queremos plantear la necesidad que tiene España de consolidar y dar estabilidad al proceso de Descentralización Política, Territorial y, por supuesto, Fiscal. Pensamos que el momento es el adecuado, pues las Comunidades Autónomas (CCAA) de régimen común están preocupadas y pendientes del acuerdo que propone el Gobierno Central, vía el Consejo de Política Fiscal y Financiera (CPFF). Nosotros proponemos un Pacto Político que resucite en cierta medida el antiguo "consenso" de los años 70 y 80, y que se produzca entre los partidos parlamentarios, tanto para la dotación suficiente de recursos como para la asunción de competencias por los niveles de Gobierno Autonómico y Local.

Esta circunstancia puede derivar en la posible existencia de una cierta multiplicidad de normativas autonómicas que produzcan en los agentes cierto rechazo. No obstante, si la capacidad se jerarquiza entre los tres niveles de Gobierno, se pueden evitar las distorsiones en la actividad económica española. A mayor abundamiento, una cierta competencia fiscal no daña la eficiencia, ni del Sector Público ni del Sector Privado.

Este "Pacto" que proponemos debe agrupar las diversas fuerzas políticas, sociales y, por supuesto, empresariales. Con ello pretendemos pactar un sistema de financiación que traiga estabilidad, suficiencia, autonomía, solidaridad y, muy especialmente, la creación de un órgano de Coordinación, con poder decisorio, entre los tres niveles de Gobierno (Central, Autonómico y Municipal). Posteriormente, vamos a tratar de esbozar cuáles pueden ser las líneas básicas del Pacto, pero esto es el desarrollo del trabajo.

2. CARACTERÍSTICAS DEL MODELO

El sistema de financiación vigente se deriva de la Ley 21/2001 de 27 de diciembre, por la que se regulan las medidas fiscales y administrativas del nuevo sistema de financiación de las Comunidades Autónomas de régimen común y Ciudades con Estatuto de Autonomía.

Los principios y, para nosotros, características son los siguientes:

1. ***Generalidad***. Tal como se indica en la Ley 21/2001 este principio o característica se manifiesta de dos formas. La primera de ellas hace referencia al carácter integrador del nuevo sistema de financiación, porque comprende la financiación de todos los servicios que es factible traspasar a las CCAA. Hasta ahora existían tres maneras de financiación diferenciadas:
   - i. Para los servicios públicos comunes.
   - ii. Por los servicios de asistencia sanitaria de la Seguridad Social.
   - iii. Para los servicios sociales de la Seguridad Social

   El segundo aspecto de la "generalidad" señala que el sistema de financiación es sólo aplicable a las CCAA que cumplan las condiciones del mismo.

2. ***Estabilidad***. Significa que el sistema es estable y obviamente más permanente. Ello permite que las CCAA puedan planificar sus actividades a largo plazo con perspectivas de más seguridad y certidumbre.

3. ***Suficiencia***. El nuevo sistema conecta este principio por medio de tres elementos: el cálculo de las necesidades de recursos, la regulación de los ingresos y las posibilidades de evolución en el tiempo. También garantiza que todas las CCAA de régimen común dispongan de recursos (suficientes) necesarios para atender sus competencias transferidas y aquellas que puedan serlo a corto, medio y largo plazo. El siguiente cuadro recoge los impuestos del nuevo sistema consensuado para todas las CCAA, contribuyendo al avance en la corresponsabilidad fiscal al aumentar considerablemente la cesta de impuestos cedidos (el modelo 1997-2001 incluía Patrimonio, Sucesiones y donaciones, Transmisiones patrimoniales, Actos jurídicos documentados, Tasa sobre el juego, y el 30% del IRPF), y la capacidad normativa otorgada en algunos de ellos:

Cuadro 1. Impuestos del sistema LOFCA, 2008

Impuestos	Capacidad Normativa	Gestión
IRPF	Fijar tipos para un 33% de la base del impuesto. Podrán establecer deducciones y modificar la deducción estatal por inversión en la vivienda habitual.	Agencia Tributaria
IVA	Sin capacidad normativa. Reciben el 35% de lo recaudado en su territorio.	Agencia Tributaria
Impuesto sobre patrimonio	Reciben la totalidad de la recaudación en su territorio. Capacidad normativa.	Gobiernos Regionales
Impuesto sobre sucesiones y donaciones	Reciben el 100% de la recaudación. Capacidad normativa.	Gobiernos Regionales.
Impuesto de transmisiones patrimoniales	Reciben el 100% de la recaudación. Capacidad normativa.	Gobiernos Regionales.
Tributos sobre el juego	Reciben el 100% de la recaudación. Capacidad normativa.	Gobiernos Regionales.
Impuestos especiales sobre tabaco, alcohol e hidrocarburos	Reciben el 40% de la recaudación. Sin capacidad Normativa	Agencia Tributaria
Impuesto sobre electricidad	Reciben el 100% de la recaudación. Sin capacidad Normativa.	Agencia Tributaria
Impuesto especial sobre determinados medios de transporte	Reciben el 100% de la recaudación. Sin capacidad Normativa.	Agencia Tributaria
Impuesto sobre ventas minoristas de determinados hidrocarburos	Impuesto nuevo. Propio de las CCAA. Tramos complementarios al estatal con limitaciones (LOFCA)	Gobiernos Regionales.

Fuente: Ministerio de Economía y Hacienda. Elaboración propia.

4. ***Autonomía***. El nuevo sistema amplía la Corresponsabilidad Fiscal en el ámbito de los impuestos y otros ingresos de carácter tributario, tanto en lo referente a la forma de obtención de recursos, como a las condiciones de prestación de sus servicios.

5. ***Solidaridad***. Su finalidad es garantizar que los servicios que provean los Gobiernos Autonómicos sean en condiciones equivalentes. Los artículos 156.1 y 158.2[1] de la vigente Constitución hablan de la existencia de esta característica o principio para garantizar que las CCAA presten los servicios básicos en niveles similares. Incluso hablan de que el Gobierno Central debe garantizar un nivel mínimo de prestación de los servicios fundamentales. Esto debe hacerse a través del Fondo de Compensación Interterritorial (FCI), Fondo de Nivelación que tal como apunta el artículo 15 de la LOFCA debe de garantizar un nivel mínimo de prestación de los servicios, que se determina en función de "media aritmética" de la mencionada prestación en las 17 CCAA y Ceuta y Melilla, teniendo en cuenta el Fondo de Suficiencia. Además, como novedad se crea el Fondo Complementario del FCI. El punto 5 del artículo 15 de la Ley 21/2001 de 27 de diciembre con respecto al Fondo Complementario del FCI señala: el Fondo Complementario se dotará anualmente para cada CA y Ciudad con Estatuto de Autonomía propio, con una cantidad equivalente al 33,33 de su respectivo FCI. Dicha cuantía se destinará a gastos de inversión. No obstante, a solicitud de los territorios beneficiarios del mismo, podrá destinarse a financiar, durante el período que determine la Ley, gastos de funcionamiento asociados a las inversiones financiadas con cargo al FCI o al Fondo Complementario del FCI.

[1] Artículo 156
1. Las Comunidades Autónomas gozarán de autonomía financiera para el desarrollo y ejecución de sus competencias con arreglo a los principios de coordinación con la Hacienda estatal y de solidaridad entre todos los españoles.
Artículo 158
2. Con el fin de corregir desequilibrios económicos interterritoriales y hacer efectivo el principio de solidaridad, se constituirá un Fondo de Compensación con destino a gastos de inversión, cuyos recursos serán distribuidos por las Cortes Generales entre las Comunidades Autónomas y provincias, en su caso.

También se crea el Fondo de Suficiencia, tal como nos dice la Ley 21/2001, en su artículo 13. Las CCAA y las ciudades de Ceuta y Melilla participarán, por medio del Fondo de Suficiencia, en los ingresos del Estado. Dicho Fondo cubrirá la diferencia entre las necesidades de gasto de cada CA y Ciudades con estatuto propio, y su capacidad fiscal. Según el punto 4 del art. 13, únicamente podrá ser objeto de revisión en los supuestos siguientes casos:

a) cuando se produzca el traspaso de nuevas competencias (servicios) o se amplíen o revisen valoraciones de traspasos anteriores.
b) Cuando cobre efectividad la cesión de nuevos ingresos (tributos u otros).

El Fondo de Suficiencia será la diferencia, positiva o negativa, entre las necesidades totales de financiación de cada CA y el valor de los recursos establecidos con anterioridad, siendo 1999 el año de referencia (base).

No obstante, debería plantearse una mayor dotación, pensando en que a partir de 2014 desaparecerán los Fondos de Cohesión, salvo para Andalucía y Extremadura. Este Fondo de Suficiencia debería colaborar a financiar los déficits del Presupuesto de cada CA. Algunos autores apuntan que la "solidaridad" se puede lograr con dos tipos de instrumentos diferentes que son el FCI y los Fondos Estructurales europeos, por un lado, y por otro, aquellos instrumentos destinados a garantizar que en todos los territorios exista un nivel de servicios públicos básicos mínimo y que al mismo tiempo esto provenga de las transferencias de nivelación, cuyo objetivo es equiparar al nivel medio existente en el país en la cobertura de dichos servicios (Herrero Alcalde, Díaz de Sarralde, Loscos Fernández, Antiqueira, Tránchez Martín, 2007). Nosotros, en principio, entendemos la solidaridad como un todo y hacer esta distinción puede ser interesante pero creemos que poco operativa esencialmente por la multiplicidad de "fondos" que conlleva.

6. ***Participación de las Comunidades Autónomas en la Agencia Estatal de la Administración Tributaria***. Tal como apunta la Ley 21/2001 en los arts. 63, 64 y 65, la colaboración con la Agencia Estatal Tributaria de las CCAA se producirá en la adopción de decisiones que afecten a la gestión tributaria. Para ello se crean tres nuevas instituciones: el Consejo Superior de Dirección de la Agencia Estatal Tributaria, la Comisión Mixta de Coordinación de la Gestión Tributaria, y los Consejos Territoriales de Dirección para la Gestión Tributaria. Sin embargo, dado su escaso o casi ningún poder decisorio, se convierten en la práctica en instituciones de asesoramiento o de nuevo debate. Consideramos que lo más sencillo hubiera sido o bien la creación de diecinueve Agencias Tributarias, o bien colocar la Administración Tributaria en el ámbito de los Gobiernos de las CCAA. Modelo, este último, que se aplica en países como Alemania, Bélgica o Canadá, entre otros países federales, y en el que la Administración Central se quedaría con la función de control y vigilancia de la gestión. Es obvio que cualquier modificación en los espacios fiscales condicionaría o podría incluso provocar rechazo de alguno de los niveles de gobierno. Por tanto, pensamos que quizás la existencia de la Agencia Tributaria sea un obstáculo para buscar una distribución de la Administración Fiscal que recayese en el nivel intermedio pero con una superestructura en el Estado federal (central) y evidentemente en el ámbito local. Esto conlleva una modificación importante en la administración y gestión tributaria que desde el punto de vista real no sería dificultosa, pues en la actualidad en cada espacio existen: administración central, administración autonómica o regional, administración local. Entendemos que esta estructura no va en contra de cuál es el nivel de gobierno que debe realizar, manejar y organizar la gestión. En los diferentes países con estructura política federal se utilizan distintas alternativas, y la que proponemos resulta adecuada para España toda vez que no altera demasiado la estructura actual, salvo la supresión de la Agencia Tributaria.

7. ***Participación en los Órganos de Control Jurisdiccional*** (Tribunales Económico-Administrativos de ámbito regional y central) nos parece interesante y oportuno, pero es preciso adaptarlo a las circunstancias que se producen en el momento presente. Pensamos que es necesario debatirlo y plantearlo con un grado importante de realismo, especialmente en el caso de España por no existir un modelo de financiación homogéneo, pues la existencia de dos

modelos (común o LOFCA y Foral), dificulta su funcionamiento, salvo si se extendiera lo aplicado en Navarra y País Vasco a las restantes CCAA.

8. ***La coordinación entre los diferentes niveles de Gobierno en materia fiscal y financiera.*** El artículo 3 de la Ley 21/2001 establece, para la adecuada y necesaria coordinación de la actividad financiera de las CCAA, la creación del Consejo de Política Fiscal y Financiera de las Comunidades Autónomas, que estará formado por el Ministro de Economía y Hacienda, el Ministro de Administraciones Públicas y los Consejeros de Economía y Hacienda de las 17 CCAA, o sólo por estos últimos (art.3.1). Sigue la Ley (art. 3.2) señalando las funciones del CPFF en materia fiscal y financiera, que –sin tener poder decisorio, pues es un órgano de mero debate- debe tratar los asuntos siguientes:

a. La coordinación de la política presupuestaria de las CCAA con el Estado.
b. La emisión de informes y la adopción de los acuerdos previstos en la Ley Orgánica 18/2001, complementaria de la Ley General de Estabilidad Presupuestaria.
c. El estudio y valoración de los criterios de distribución de los recursos del FCI.
d. El estudio, la elaboración, en su caso, y la revisión de los métodos utilizados para el cálculo de los costes de los servicios transferidos a las CCAA.
e. La apreciación de las razones que justifiquen, en cada caso, la percepción por parte de las CCAA de las "asignaciones presupuestarias", así como los criterios de equidad seguidos para su afectación.
f. La coordinación de la política de endeudamiento.
g. La coordinación de la política de inversiones públicas.
h. En general, todo aspecto de la actividad financiera de las CCAA y de la Hacienda del Estado que, dada su naturaleza, precise de una actuación coordinada.

El principio de coordinación entre niveles de Gobierno para hacer frente al mismo Sistema de Financiación, incorpora las normas concretas de localización y atribución de los rendimientos de los tributos cedidos a las CCAA, así como de colaboración entre las Administraciones en la gestión, liquidación, recaudación e inspección, y también en la revisión de los actos dictados en vía de gestión tributaria, con la finalidad de dar mayor seguridad y estabilidad al ejercicio de las potestades fiscales de las CCAA.

No obstante, este principio o característica de coordinación adolece de alguna insuficiencia o falta de concreción. Nos referimos a la inclusión de los tres niveles de Gobierno, pues no hay referencia alguna al ámbito local.

III. DEFECTOS Y/O LAGUNAS DEL MODELO DE 2001

El modelo de financiación de las CCAA en la actualidad, es el que entró en vigor el primero de enero del año 2002. Éste no afronta las insuficiencias y problemas existentes en ese momento y por supuesto en la actualidad. Por tanto, es necesario en el momento actual (abril 2009) buscar un "Pacto Político" entre las diferentes fuerzas parlamentarias para llegar a un consenso que posibilite afrontar las insuficiencias que presenta nuestra Constitución y además adecuar las competencias –gasto autonómico- a una distribución adecuada entre los tres niveles de gobierno: Central, Autonómico y Local. Además, el sistema actual presenta importantes deficiencias en orden a los siguientes aspectos:

1) Falta de homogeneización al contemplar la Carta Magna dos modelos de financiación totalmente diferenciados, pues uno de ellos –el modelo común- se generaliza en quince CCAA y dos Ciudades Autónomas, y otro con importantes ventajas respecto al anterior se circunscribe a sólo dos CCAA –Navarra y País Vasco-. Por tanto, es el momento de plantearse si no sería necesario unificar este criterio quizás por la vía de extender el modelo foral a aquellas regiones que lo soliciten. No creo que en el año 2009 vayamos a considerarnos excesivamente partidarios de que esto provocaría la falta de unidad de España.

2) No existe en la Constitución un criterio válido para la atribución de competencias a los gobiernos autonómicos y por ello me remito a la experiencia que se ha derivado del contenido del artículo 148.2 de la vigente Constitución[2]. Esto se deriva de que a veces en los distintos gobiernos de la democracia las mayorías que los apoyaban no eran suficientes y era necesario pactar con otros grupos políticos, esencialmente con los partidos nacionalistas (PNV, CIU y otros) para aprobar leyes importantes como, por ejemplo, la de Presupuestos Generales del Estado.
3) Es preciso crear –y si es posible a nivel constitucional- una institución de coordinación entre los tres niveles de gobierno, que coordine las políticas -tanto en orden económico y social como estrictamente político-, las opiniones y los planteamientos de los distintos niveles de gobierno. Es decir, el Consejo de Política Fiscal y Financiera habría que suprimirlo, puesto que ni sirve ni servirá para nada, porque sus acuerdos no son vinculantes y además no dispone de poder decisorio –respeto por los votos de la mayoría de sus integrantes-.
4) Otro aspecto que hay que tener en cuenta es que a partir de 2014 no sabemos si vamos a contar con fondos europeos suficientes, toda vez que si esto se produce evidentemente las grandes receptoras van a ser Andalucía y Extremadura. ¿Qué va a pasar con otras Comunidades hoy en día beneficiarias de dichos Fondos? La solución al problema no es fácil, pues no debemos olvidar que si comparamos el presupuesto de la UE para 2009 con el anterior el incremento no ha sido superior al 2%. Esto nos lleva a pensar que si la posición de los países contribuyentes netos sigue siendo cicatera la tendencia se mantendrá.
5) Con respecto a los Fondos de Cohesión quizás sea el momento de configurarlos como ayudas a aquellas regiones que bien por razones económicas o de otro tipo se ven perjudicadas en un periodo determinado. En la práctica sería combinar la finalidad de aplicar la cohesión económica, social y territorial a las regiones, y complementarlo con un apoyo a lo que se podría denominar "fondo de contingencias", que se dedicaría fundamentalmente a aquellos sucesos que se producen de forma aleatoria y además imprevisible.
6) Otro aspecto a tener en cuenta es el significado económico y político de la Cooperación Transnacional y Transfronteriza. Esta debe potenciarse y especialmente en la denominación de los territorios de forma homogénea con el fin de diseñar actuaciones concretas ajenas totalmente a la circunstancia de que una parte de un territorio pertenezca a un país y otra a otro. Además, no sólo debe ser importante esta cooperación a nivel de infraestructuras y a nivel de energía y telecomunicaciones, sino que debe ampliarse de manera significativa a diferentes sectores productivos, que evidentemente son complementarios entre sí.
7) Por lo que respecta a las competencias atribuidas a los niveles de gobierno pensamos que es preciso ajustar aquéllas teniendo en cuenta la eficacia y eficiencia en la cobertura y prestación de lo que podemos denominar "servicios públicos esenciales". Esta situación no es una cuestión fácil porque se trata de buscar criterios lo más neutrales posibles entre los territorios a la hora de distribuir las ayudas del gobierno superior (Central) con respecto a los gobiernos autonómicos y los gobiernos locales. Esta posibilidad requiere que la cooperación en materia de cobertura de servicios de los niveles autonómico y local deba establecerse prácticamente *ex novo*, puesto que exceptuando Cataluña, Navarra y País Vasco principalmente, y en menor medida las islas Canarias a través de los Cabildos, es preciso tener un planteamiento homogéneo entre las 17 CCAA y los 8.108 municipios. No obstante, no hay que olvidar que dada la pluralidad de municipios sea necesario para una mayor eficacia en la gestión crear de *iure* las áreas metropolitanas como una institución de gestión de servicios públicos que se prestan de manera conjunta a varios municipios. Para ello podríamos pensar comparativamente en el caso de Francia con alrededor de 36.700 municipios y una gestión bastante eficaz de los mismos con respecto

[2] 148.2. Transcurridos cinco años, y mediante la reforma de sus Estatutos, las Comunidades Autónomas podrán ampliar sucesivamente sus competencias dentro del marco establecido en el artículo 149.

a los españoles. Esta circunstancia comparativamente con España es reveladora, y su aplicción colaboraría a una mayor eficiencia de la ordenación del territorio de España.

8) También deberían considerarse otros aspectos que es preciso tener en cuenta: el comportamiento del gasto sanitario, es decir, el incremento que el gasto corriente de los presupuestos de las CCAA puede suponer tras la asunción del 1 de enero de 2002 por parte de todas las CCAA y Ciudades autónomas de la competencia de sanidad. Es decir, ¿van a ser las CCAA las que soporten el déficit sanitario debido a la cicatería de la Administración Central? Otro aspecto a considerar es el relativo a los impuestos cedidos pues la CCAA han hecho poca utilización de la normativa propia salvo para la supresión de los impuestos de patrimonio y sucesiones en casi todas las CCAA, aunque las donaciones se mantienen. Es más, posiblemente en dos o tres años salga una normativa europea con respecto a este impuesto. Cumple además con un objetivo importante que es someter a tributación las donaciones inter vivos, dado que son un incremento de patrimonio. En este contexto nos podemos plantear un interrogante: ¿deben las CCAA asumir nuevas competencias tributarias? (hechos imponibles y nuevos impuestos). Esta idea no ha sido hasta ahora considerada por el gobierno central.
9) Un diseño real de los instrumentos del Fondo de Nivelación exige conocer la dinámica del sistema de financiación y una actualización periódica de los criterios de distribución. Un correcto funcionamiento de estos mecanismos llevaría a que los territorios con mayores necesidades de gasto deberían obtener más recursos y aquellos con menor necesidad de gasto obtuvieran menos recursos.

IV. REFLEXIONES FINALES

Pensamos que con todo lo expuesto se tiene una idea clara de la urgencia de que el sistema de financiación de las CCAA de régimen común acometa su reforma, a pesar de que conlleva dificultades de orden político y que quizás con la crisis económica que padecemos sea poco oportuno, pero también pensamos que a veces esta situación puede propiciar el acuerdo y el consenso. De todo lo mencionado podríamos resaltar como deficiencias y desajustes más importantes los siguientes:

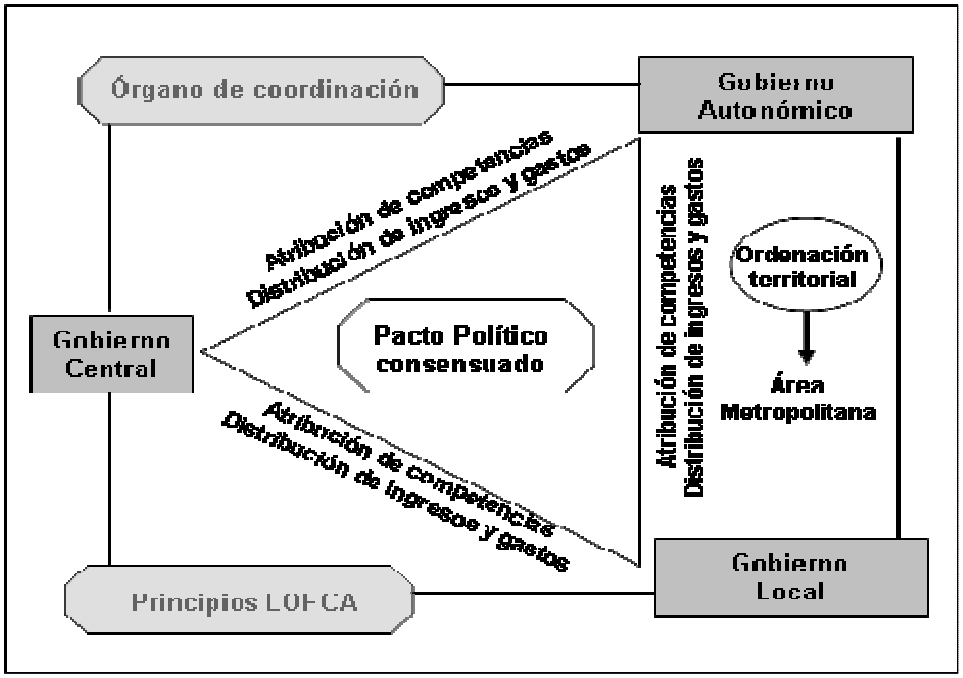
1. No se incluye el nivel de gobierno municipal, lo que evidentemente plantea cuestiones que sería imposible tratar en un trabajo de esta amplitud. Requiere un nuevo trabajo que previsiblemente afrontaremos en el futuro.

2. En segundo lugar, la falta de una institución de coordinación que diseñe el papel de cada nivel de gobierno en un asunto concreto. Además debe contar con poder decisorio porque sin él carece de eficacia.
3. Es necesario que pensemos en la posibilidad de que a partir de 2014 las ayudas procedentes de los Fondos europeos se reduzcan casi en su totalidad, si exceptuamos los territorios de Andalucía y Extremadura.
4. Es preciso que los Gobiernos autonómicos tengan dos preocupaciones importantes: utilizar la capacidad normativa de que disponen para introducir modificaciones en los impuestos cedidos o compartidos -en este último caso se encuentra el IRPF-, y que esta posibilidad se pueda extender con un planteamiento realista a los municipios si se cree necesario.

Por tanto, todo esto debe converger en la firma de un Pacto político entre las diferentes fuerzas parlamentarias, e inclusive las no parlamentarias, para tratar de consensuar un sistema de financiación estable y que sus modificaciones no sean tan continuas como ahora, teniendo lugar sólo y exclusivamente en aquellos periodos necesarios. Este tipo de Pacto no es ajeno a los españoles, ya que en 1977 se firmaron -a iniciativa del Presidente Adolfo Suárez y del Vicepresidente Económico Fuentes Quintana- los Pactos de la Moncloa por los que se diseñó la que –con algunas rectificaciones puntuales- ha sido la política económica de España desde esa fecha.

Los elementos fundamentales del citado Pacto Político que debe tener un marcado carácter de consenso, son recogidos en el gráfico siguiente y responden a las deficiencias y desajustes mencionados con anterioridad:

Gráfico 1. Financiación autonómica y Pacto Político



Fuente: Elaboración propia, 2009.

El núcleo, en torno al cual gira el Pacto que aquí proponemos, es una estructura marcada por los tres niveles de Gobierno –Central, Autonómico y Local- y por la atribución de competencias, así como evidentemente por la distribución de ingresos y gastos entre cada uno de dichos niveles. Se pone especial énfasis en el nivel de Gobierno Local, cuya relación con el nivel intermedio (Gobierno Autonómico) favorecería una ordenación territorial más lógica y eficaz, más atenta a la cobertura las necesidades y a las economías de escala a través del diseño de áreas metropolitanas.

Esta estructura triangular formada por los niveles de gobierno se apoyaría evidentemente en los principios recogidos tanto en la Constitución vigente –solidaridad en sus tres dimensiones (solidaridad de las regiones con el Estado, solidaridad interregional y solidaridad intrarregional) y equidad- como en LOFCA (estabilidad, generalidad, suficiencia, autonomía, corresponsabilidad fiscal, coordinación, además de la solidaridad).

Tendría un elemento nuevo que es la creación de una institución de coordinación a la que ya hemos hecho referencia, y que tendría poder decisorio en el diseño de la política de financiación autonómica. Entre sus funciones estaría la regulación de las competencias compartidas con el fin de profundizar en el principio de Corresponsabilidad Fiscal, y esencialmente la coordinación de los tres niveles de gobierno, estableciendo un sistema de control de arriba abajo, que garantizara, especialmente, la inclusión del nivel local como un elemento fundamental en el modelo de financiación.

En conclusión, se trata de un proyecto de Pacto basado en el consenso multilateral que supone evidentemente un esfuerzo político para mantener los principios fundamentales recogidos en la Constitución y en la LOFCA, y para establecer el difícil equilibrio entre generalidad y especificidad, siendo el principal objetivo la ciudadanía y un nivel mayor en la prestación de los servicios públicos esenciales. Se trataría de un sistema que actualizado –no necesariamente renovado-, fuera lo suficientemente flexible para adaptarse a la evolución económica del entorno, es decir, la Unión Europea.

V. BIBLIOGRAFÍA

CARAMÉS VIEITEZ, L., Y VÁZQUEZ NAVARRETE, J. A. (1997), *Los problemas de la financiación autonómica. Un perspectiva desde Galicia*, La Coruña.

DE LA FUENTE, A.; GUNDÍN, M. (2007), "La financiación autonómica en los nuevos estatutos regionales", *Hacienda Pública Española/ Revista de Economía Pública*, 182-3, 163-199.

ELAZAR, D. (1994) (ed.), *Federal Systems of the World: A handbook of Federal, Confederal and Autonomy Arrangements*, Longman Group Limited, Harlow.

ELAZAR, D. (1995), *Federalism: An Overview*, HSRC, Pretoria.

HERRERO ALCALDE, A., DÍAZ DE SARRALDE, S., LOSCOS FERNÁNDEZ, J., ANTIQUEIRA, M., TRÁNCHEZ MARTÍN, J. M. (2007), "La reforma de la Financiación Autonómica: reflexiones y escenarios", *ICE, Política Económica en España*, nº 837 (Julio-Agosto).

HIERRO RECIO, L.; PATIÑO RODRÍGUEZ, D.; ATIENZA MONTERO, P. (2008), "Un análisis de las diferencias territoriales en los saldos de las balanzas fiscales y de sus posibilidades de corrección", *Revista de Estudios Regionales*, nº 83 (septiembre-diciembre), Ed Universidades Públicas de Andalucía, pp. 79-109.

LAGO PEÑA, S. (2007), *La Financiación de las Autonomías: Perspectivas de Futuro*, Instituto de Estudios Fiscales, Madrid.

LEY 21/2001, de 27 de diciembre, *por la que se regulan las medidas fiscales y administrativas del nuevo sistema de financiación de las Comunidades Autónomas de régimen común y Ciudades con Estatuto de Autonomía*. (B.O.E. núm. 313 de 31/12/2001).

LÓPEZ CASASNOVAS, G. Y PONS NOVELL, J. (2005), *Análisis de la redistribución entre las Comunidades Autónomas. Criterios de evaluación y propuesta de nuevos principios básicos*. Ed. Universitat de Barcelona, Barcelona.

MITXELENA CAMIRUAGA, C. (2006), *Financiación autonómica y cohesión territorial*, Thomson Civitas, Cizur Menor (Navarra).

MONASTERIO ESCUDERO, C. Y FERNÁNDEZ LLERA, R. (2008), "Hacienda autonómica, reglas fiscales y modernización presupuestaria", *Revista de Presupuesto y Gasto Público*, nº 51, Madrid, pp. 83-106.

RÍOS BERJILLOS, A. DE LOS (2004), *Financiación Autonómica. Cuestión de principios*, Ed. Desclée de Brower, S. A. Bilbao.

WATTS, R. L. (2006), *Sistemas federales comparados*, Marcial Pons, Madrid.

IMPACT OF THE COMMON AGRICULTURAL POLICY (CAP) ON THE ECONOMIC INDICATORS OF FARMS IN THE MIDDLE-EUROPEAN POST-COMMUNISTIC COUNTRIES

BOHUMILA ZIZKOVA
E-mail: zizkova@fame.utb.cz
Tomas Bata University in Zlin
Faculty of Management and Economics

Abstract

The Common Agricultural Policy (CAP) is a system of European Union (EU) agricultural subsidies and programmes. The issue of EU subsidies for farmers is linked to the historical traditions of the CAP, the specific circumstances of the small farms, a political support for maintaining rural agricultural communities and naturally influences the economic standard of farms in the European Union. The aim of this paper is the analysis of the Common Agricultural Policy and its impact on the agriculture of the four middle-European post-communistic countries - the Czech Republic, the Slovakia, the Poland and the Hungary. The paper also includes case study of the less-favoured areas payment system.

Key-Words: Agricultural subsidies, The Common Agricultural Policy (CAP), LFA payments
Thematic Area: Local, regional and urban economy.

1. THE COMMON AGRICULTURAL POLICY (CAP)

The Common Agricultural Policy has been the basic pillar of the European integration already from the beginning of the European Economic Community (EEC). Already in the Treaty of Rome about establishment of the EEC there are included general targets of the Common Agricultural Policy and are also outlined the instruments, by which the existing targets are to be reached.

„*Article 33*

1 The objectives of the common agricultural policy shall be:

a) to increase agricultural productivity by promoting technical progress and by ensuring the rational development of agricultural production and the optimum utilisation of the factors of production, in particular labour;

b) thus to ensure a fair standard of living for the agricultural community, in particular by increasing the individual earnings of persons engaged in agriculture;

c) to stabilise markets

d) to assure the availability of supplies

e) to ensure that supplies reach consumers at reasonable prices

Article 34

…In order to attain the objectives set out in Article 33, a common organisation of agricultural markets shall be established….

…In order to enable the common organisation referred to in paragraph 1 to attain its objectives, one or more agricultural guidance and guarantee funds may be set up….“ [10]

The Common Agricultural Policy (CAP) is a system of European Union agricultural subsidies. Europe's agricultural policy is determined at EU level by the governments of Member States and operated by the Member States. It is aimed at supporting farmers' incomes while also encouraging them to produce high quality products demanded by the market.

The efforts to improve the Common Agricultural Policy and to upgrade its targets proceeded already from the very beginning. The first attempt at the reform is from the end of the sixties of the 20th century. It concerned the Mansholt's plan, but it failed. The reform from the year 1992 focused on decrease of support prices on cereals, beef and milk, the support of early retirements, the support of alternative incomes at countryside and so on. The reform from the year 2003 focused especially on the support of the quality against previous quantity approach and in particular in the area of the environment, the consumers' health and the living conditions of animals.

Along with the analysis of the recent courses of the CAP I keep to the Communication from the Commission on Simplification and Better Regulation for the Common Agricultural Policy and to the Council Regulation (EC) n° 1234/2007 establishing a common organisation of agricultural markets and on specific provisions for certain agricultural products (Single CMO Regulation) and from the so-called 'Health Check' of the CAP plan by the European Commission for streamlining and further modernising the European Union's Common Agricultural Policy. The target of the simplification is the increasing of the transparency and the comprehensibility of regulations, liberalization of their demands and cost reduction for companies. The main task of the 'Health Check' of the CAP is then more thorough review of regulations and supports within the frame of CAP and their adaptation for new tasks and opportunities in the EU formed by already 27 member states. With the Council Regulation (EC) n° 1234/2007 was created the integrated set of harmonised rules for all common market organizations.

2. IMPACT OF THE COMMON AGRICULTURAL POLICY (CAP) ON THE AGRICULTURE OF THE MIDDLE-EUROPEAN POST-COMMUNISTIC COUNTRIES

Each country has many specifics which determine its agriculture and agricultural subsidies. Therefore I have divided this part of the paper into four sections – each section for one country. Firstly I describe the most important specifics for each country and than the development of subsidies. At the end of this part I compare the countries and show the impact of subsidies of CAP to economic indicators of farms.

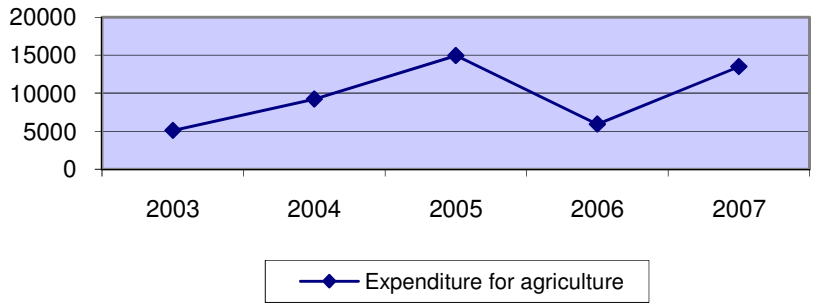
2.1 Poland

The total area of Poland is 312 700 km^2 and it covers 7.1 % of of the total EU area. Agricultural land create 16,3 million ha, which constitutes 52,2% of the total area of the country. Poland is a lowland country. It has a moderate climate with changeable weather conditions. The soil and climatic conditions as well as traditions in the respective regions determine production specialisation. Cultivation of intensive cereals, mainly wheat and sugar beet and rape, predominates and where the conditions are worse cultivations of more resistant plants, i.e. rye and potatoes, predominate. Agricultural production conditions differ in Poland, especially as compared to EU agriculture. In majority, agricultural holdings have small areas, too high number of employees; apply relatively little chemical production means. Considerable part of farmers applies traditional production methods, where production is only for farmer's own family. Animal production is usually carried out with small intensity of breeding, which contributes to the protection of the natural environment.

Agricultural subsides in Poland are divided into three parts in the year 2007. First part are subsidies to agriculture for example biological advancement in plant production, biological advancement in animal production, subsidies to calcium fertilisers, plant protection, organic farming, etc. Second part create subsidies to rural areas development include Agency for

Restructuring and Modernisation of Agriculture, Restructuring and modernization of food sector and rural development, Rural Areas Development Plan, complementary payments for arable land, etc. Third part is to the agricultural markets due to Agricultural Market Agency. Development of expenditure for agriculture, agricultural markets and rural development in Poland is illustrated on next chart.

Fig. 1 Expenditures for agriculture, agricultural markets and rural development in Poland (in mil. PLN)



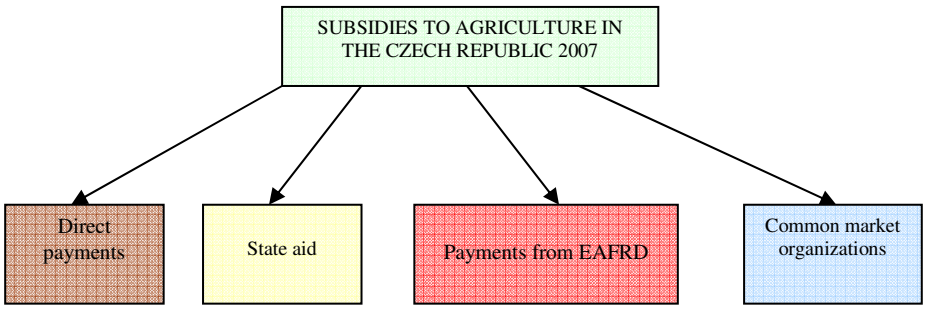
Expenditures for agriculture, agricultural markets and rural development in Poland increase in the years 2003 – 2005. This growth was due to Poland accession to the European Union. As compared to 2003, i.e. the year before the accession, expenditure on agriculture, rural development and agricultural markets tripled in 2005. Within the same period of time, national funds increased by 59%, whereas EU funds increased eleven times. The peak of subsidies was reached in the year 2005 because of many of subsidies from the year 2004 were paid in the year 2005. The fall in the year 2006 was due to dramatic reducing of expenditure to Common Agricultural and Fisheries. The growth causes the beginning of the new programming period 2007 -2013 in the year 2007. The programming period is focused on rural development and expenditure to rural development rise from 5000 mil. PLN in 2006 to 11 000 mil PLN in 2007. In 2004-2006, domestic and European Union funds totalled PLN 89,800 million, including those without payments to farmers for national insurance amounting to PLN 45,200 million.

2.2 Czech Republic

The Czech Republic is a Central European country with a total area of 7 886 000 ha. 67% of total area is situated at an elevation of not more than 500 m above sea level, 32% at an elevation between 500 and 1000 m above sea level and 1% of the territory lies at an elevation above 1000 m above sea level. Mutual pervasion and mixing of oceanic and continental effects are characteristic for the climate in the Czech Republic. Agricultural land covers about a half (54%) of the total area of the Czech Republic. Present agricultural enterprises use more than 90 % of the rented land. Physical entities farm 29.8 % of agricultural land, legal entities farm a total of 70.2 % of agricultural land. For one enterprise of private farmers there are 23.6 ha of farmed agricultural land, for one company limited liability company there are ca 475.7 ha of agricultural land, for one cooperative there are 1 456.60 ha of agricultural land and for one joint stock company there are ca 1 404 ha.

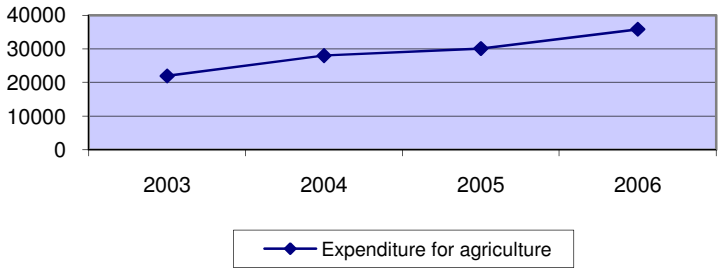
In the year 2007, subsidies in Czech Republic were divided into four main groups: Direct payments, the payments from the European agricultural fond for rural development (EAFRD), state aid and the common market organizations, which are illustrated in scheme.

FIG. 2 : Subsidies to agriculture in the Czech Republic 2007-2013



Direct payments are Single Area Payment Scheme (SAPS) and Top – Up State Aid. These supports are for farmers basic European financial source. The National Strategic Rural Development Plan of the Czech Republic is co-financed by the national state budget of the Czech Republic and European Agricultural Fund for Rural Development and create the second group of subsidies. Strategy of NSRD is created by 4 axes. . Axis 1 is focused on innovation transfers and knowledge of agri-food sector and the modernising of agricultural holdings. The main priority of axis 2 is the protection quality of the environment and water sources. Axis 3 supports the development of living conditions in rural areas and diversification of economic activities. The LEADER (Axis 4) integrates all three axes and follows the local strategies and partnership [local action groups or micro-regions] which are supported by Leader. State aid is divided into two main groups - subsidies from the Ministry of Agriculture and the Support and Guarantee Fund for Farmers and Forestry (SGFFF) subsidies. Expenditure for agriculture is shown on the chart. Communal market organizations, within them, agricultural products are organized by harmonized rules. In terms of simplifying the CAP, a single Communal Market Organization for all agricultural products was established in October 2007, after a proposal of the European Commission in December 2006.

Fig. 3 Expenditures for agriculture and rural development in Czech Republic (in mil. CZK)



Agricultural subsidies increase in Czech Republic in years 2003 – 2006. Level of subsidies and aids to agriculture increased namely the direct payments. The entrance Czech Republic to the European Union brought a significant increase of income from all agriculture activities of Czech Republic.

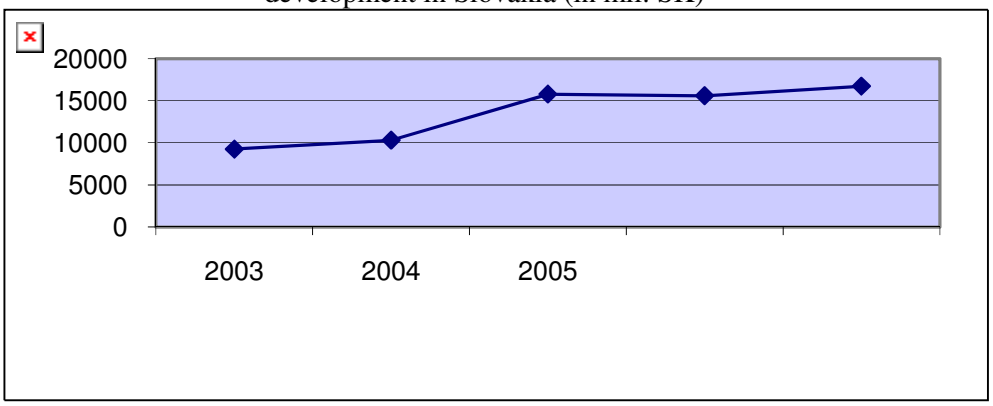
2.3 Slovakia

From the total area of Slovakia 49 035 km^2 is the area of agricultural land in use represented 1,941,380 ha. The greatest share in that area was represented by arable land (69.9%) and by permanent meadows and pastures (27%). The highest share in the structure of utilised arable land was represented by cereals (58.9%), arable fodders (17.8%) and industrial crops (16.1%) More than 90% of utilised agricultural land was rented. The major share in the group of corporate organizations is held by farms with more than 1000 ha, which make up for 33.8% of

farms and which farm at 75.4% of the area of agricultural land. Most of the farms held by natural persons (26.9%) still fall into the area bracket from 1 to 5 ha, although the farms with area of 100-500 ha hold most of the land area (41.1%).

Structure of agricultural subsidies is analogous as in Czech Republic. The subsidies were also divided into four main groups: Direct payments, the payments from EAFRD, state aid and the common market organizations in the year 2007. However the structure of the 4 main groups is different. Direct payments consisted of SAPS, complementary payments (CNDP), and separate payment for sugar. The fund EAFRD funds the Rural Development Programme of the Slovak Republic for programming period 2007-2013. Programme consists of four axes: Axis 1 - Improving the competitiveness of the agricultural and forestry sector, Axis 2 - Improving the environment and the countryside, Axis 3 - Quality of life in rural areas and diversification of rural economy and Axis 4 - Implementation of LEADER approach. Since 2004, the state aid in the sector of agriculture was provided under the Decree No 806/2004-100 of the Ministry of Agriculture of the Slovak Republic. The Common market organizations or Market oriented expenditure (MOE) represents the expenditure allocated to market measures. This expenditure was based on the binding regulations which stand at the core of the EU Common Market Organisation. The support to market in agrarian commodities included export and product subsidies, and production subsidies.

Fig. 4 Expenditures for agriculture and rural development in Slovakia (in mil. SK)



Expenditures for agriculture increase in Slovakia in years 2003 – 2007. The most significant growth was in the year 2005 when was paid the most part of first payments of subsidies of CAP. In other years there was a slight increase. The biggest share in total expenditure in the year 2007 was that of the Rural Development (46.6%). Direct payments in 2007 totalled to SKK 7,291.6 million and represented 38.2% of the total volume of subsidies in agriculture. MOE expenses are reduced each year within the EU budget. A total of SKK 232.5 million was paid out in 2007 in the Slovak Republic, which is 24.9% of the level in 2006 when the amount of MOE subsidies stood at SKK 933.2 million. The total amount of state aid represented SKK 293.7 million, which was 40.5% less than in the previous year. The largest portion (37.45%) was represented by subsidies to compensate for losses caused by adverse weather conditions and their impact on agricultural crops.

2.4 Hungary

Hungary has a total area of 9.3 million hectares. The agricultural area is 5,807,000 hectares, which represents an outstandingly high proportion in Europe. 78% of this agricultural area is made up of arable land and 17% is grassland, while kitchen gardens, orchards and vineyards accounted for 5%. According to certain calculations, the share of agribusiness in GDP production is 12–13%. the number of employees in the national economy was 3.93 million, of which agriculture, forestry and fishery employed 182,900 people, 4.7% of all employees in the national economy. In Hungary we can find all the different types of entrepreneurship. A unique

feature the bi-polar economic structure is that it consists mainly of large farms and small units. There is a large number of cooperatives (13855) but, more importantly, there are 619,000 private household farms involved in agricultural production. The high number of individual farms is explained by the EU's statistical methods. Farmers are continuously involved in statistical monitoring until the range monitored covers 99% of production. For instance, a household rearing one pig per year qualifies as an individual farm. This explains the fact that individual farms are substantively different in the extent to which farming is their actual purpose:

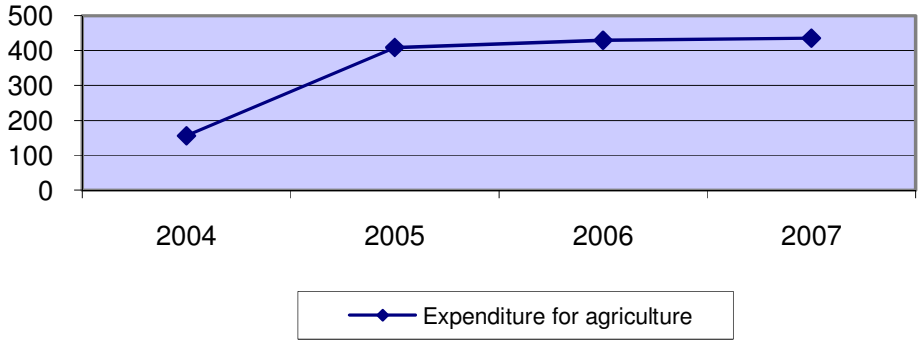
- Enterprises producing exclusively for own consumption (363,000 enterprises, 51.3%);
- Enterprising selling the surplus over own consumption (234,000 enterprises, 33.1%);
- Enterprises producing mainly for sale (109,000 enterprises, 15.5%);
- Agricultural service providers (707 enterprises, 0.1%).

The typical size of individual enterprises exceeded 300 hectares, whereas 73% of private farms ranged between 10 and 300 hectares. In 2007 almost one third (29%) of enterprises and over half (55%) of private farms were involved in animal husbandry. 90% of the stock was accounted for by cattle and pigs in enterprises and cattle, pigs, sheep and horses in private farms. This situation influence agriculture subsidies.

The structure of agricultural subsidies is quite same as in the Czech Republic or Slovakia with the nationality differences for example LFA payments in mountain areas are not in the New Hungary Rural Development Program because mountain areas are not in the Hungary. You can find Expenditure for agriculture and rural development in Hungary on chart no .

The development of agricultural subsidies in Hungary is dramatic after the entrance of Hungary in European Union. The growth of subsidies was from 155,9 billion HUF in 2004 to 409 billion HUF in 2005. The entrance was in the year 2004 but many of subsidies from the year 2004 were paid in the year 2005 therefore the growth is in the year 2005. The share of subsidies from national budget to EU budget is about 1.3x more from EU budget than from national budget.

Fig. 5 Expenditures for agriculture and rural development in Hungary (in mil. HUF)



2.6 Overall impact

Development of subsidies in the surveyed countries indicates that implementation of CAP when enter the European Union has brought rapid increase of subsidies to agriculture at all the surveyed countries. For discover the impact of CAP on economic returns of farms is needed to choose the suitable indicators for comparison and to find credible data source. For the whole comparative analysis I have used Farm Accountancy Data Network of the European Union (FADN). From indicators I have chosen Total output, Farm net value added and Family farm income. Then I show count of hectares over average farm in the surveyed countries because the investigated returns are presented in EUR over average farm and the area of farmland is very different according to the each single country. First I am comparing total output, farm net added

and operating subsidies. This comparison is accomplished to two versions - in total count and recounts over one hectare of farmland, because of better predicative capability.

Chart 1 Selected economic indicators in 2006

Country	Total output		Farm Net Value Added		Operating subsidies		Total Utilised Agricultural Area
	EUR	EUR/ha	EUR	EUR/ha	EUR	EUR/ha	ha
(CZE) Czech Republic	272673	1120,91	94026	386,52	62780	258,08	243,26
(HUN) Hungary	49648	1035,20	18076	376,90	10367	216,16	47,96
(POL) Poland	23282	1344,23	10706	618,13	4921	284,12	17,32
(SVK) Slovakia	368455	676,52	10241	18,80	104380	191,65	544,63

In the chart you can see the greatest output and amount of subsidies achieves the Slovakia and the Czech Republic, but if we recount this sum over one hectare we would give the lowest sum at the Slovakia and the highest at the Poland. Net added value is the biggest in the Czech Republic, but the Poland achieves the double by recount over hectare. Slovak farms gains only 18 EUR over hectare of net added value, compared with 618 EUR/hectare in the Poland. This difference is very huge. Current subsidies have the Slovakia the highest, but after the recount the Czech Republic have. All the amounts of overall indicators of the Slovakia are high but after recount are low. It is given by large area of farmland of farm here.

Comparisons indicators of income, operating subsidies and income minus operating subsidies of farms in the years 2004 – 2006 are in the chart no 2.

Chart 2 Selected economic indicators in 2004-2006

Country	Year	Family Farm Income	Operating subsidies	Family Farm Income - Operating subsidies
(CZE) Czech Republic	2006	23826	62780	-38954
	2005	18522	47501	-28979
	2004	26609	37526	-10917
(SVK) Slovakia	2006	-101158	104380	-205538
	2005	-8879	91248	-100127
	2004	-9558	52889	-62447
(POL) Poland	2006	9087	4921	4166
	2005	7293	3314	3979
	2004	6398	2066	4332
(HUN) Hungary	2006	7930	10367	-2437
	2005	7062	10511	-3449
	2004	6713	9594	-2881

The chart shows that the lowest level of income reached farms in Slovakia. Family farm income is in Slovakia in minus. In the year 2006 the family farm income was in the highest level in comparison with 2004 and 2005 in all countries except Slovakia. Indicator of income minus operating subsidies was in Czech Republic, Slovakia and Hungary in all years in minus. It means that farms would be in loss without subsidies. Farms in Poland reached more Family Farm income than got operating subsidies so only these farms would be in profit without subsidies.

3. IMPACT OF LFA PAYMENTS ON THE ECONOMIC INDICATORS OF FARMS

Payments for less-favoured areas (LFA) are one of the most important subsidies from the CAP. LFA payments are paid as a compensation of disadvantages which affect farming in the less-favoured areas. Firstly I want to analyze basic economic indicators of farms in different types of LFA and than evaluate the impact of LFA payments on the economic indicators of farms.

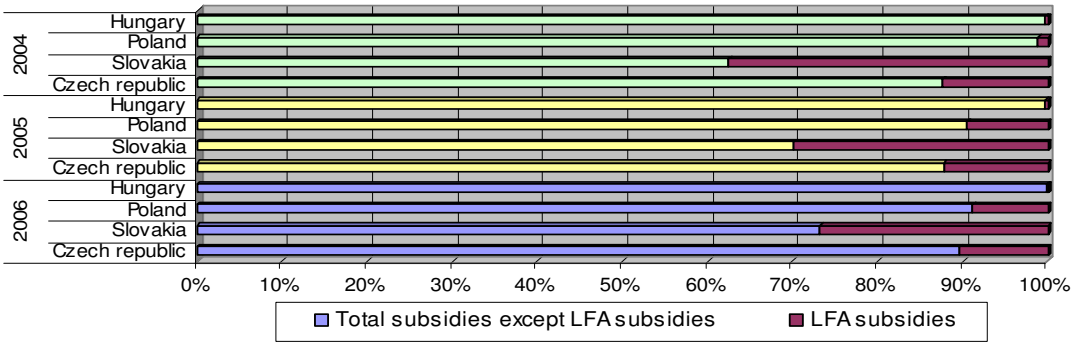
Chart 3 Basic economic indicators of farms in different types of LFA

		Annual work unit	Total Utilised Agricultural Area	Total Inputs	Total Output	Difference between total output and inputs
Czech Republic	**Total (LFA)**	**8,55**	**249,26**	**309543**	**272673**	**-36870**
	LFA - mountain areas	6,98	238,48	213879	158842	-55037
	LFA - not mountain areas	6,69	222,87	222393	184661	-37732
	not in LFA	*9,68*	*253,5*	*365917*	*332487*	*-33430*
Slovakia	**Total (LFA)**	**17,63**	**544,63**	**575538**	**368455**	**-207083**
	LFA - mountain areas	24,59	801,51	711724	377470	-334254
	LFA - not mountain areas	18,37	629,12	642622	408720	-233902
	not in LFA	*12,86*	*314,3*	*432129*	*324725*	*-107404*
Poland	**Total (LFA)**	**1,78**	**17,32**	**18544**	**23282**	**4738**
	LFA - mountain areas	1,74	11,5	11477	14599	3122
	LFA - not mountain areas	1,73	18,39	17319	21186	3867
	not in LFA	*1,83*	*16,6*	*20243*	*25973*	*5730*
Hungary	**Total (LFA)**	**1,82**	**47,96**	**51931**	**49648**	**-2283**
	LFA - mountain areas					0
	LFA - not mountain areas	1,88	44,47	46038	43043	-2995
	not in LFA	*1,78*	*50,22*	*55736*	*53912*	*-1824*

If we do compare the manpower level exploited at the farms, we would not manifest any strong boost of the manpower level at the farms which are farming in LFA. But there are significant differences between the countries. It is inspectional the Slovakia and the Czech Republic need the manpower many times over the Poland and the Hungary. If we do compare areas of farms and give the manpower over one hectare of farmland, we would receive value around 0.03 AWU/ha for the Czech Republic, the Slovakia and the Hungary, but 0.1 AWU/ha for the Poland which is approximately three-times more than for the three previous mentioned countries. Through the comparison of total amount of input and total amount of output we can say farms in LFA shows longer interval between outputs and inputs than farms which are farming out of LFA. The biggest deprivation of farms and difference between farming in LFA and out of LFA is in the Slovakia. Only one of the surveyed countries the Poland gains profit at the whole branches. At the surveyed inputs and outputs there are not included subsidies.

The portion of LFA supports in current subsidies gives us level of exploitation of these subsidies in the single country. On the next figure there is comparison of the surveyed countries in the three subsequent years since the integration of these countries to the EU.

Fig. 6 Share LFA payments in operating subsidies



There is evident the Hungary exploited LFA subsidies just in very limited supply as you can see on the figure (Fig. 6). This situation is given by the geographical specification of the country and portion of LFA areas on farmland in amount of 14 %. In the Poland LFA subsidies frame c. 10 % of all the subsidies since the year 2005. In 2004 the Poland almost did not exploit them. This state was because of legislation. In the Slovakia and the Czech Republic has been going down the portion of LFA payments off direct payments. The portion of LFA payments in total amount of current subsidies was in the Czech Republic c. 11 % in the year 2006.

If we give the LFA subsidies over one hectare of farmland (for more transparent comparison because of serious differences between areas of the farms) we would receive that the highest sums of subsidies are paid in the Czech Republic (from 47 up to 106 EUR/ha), next in the Slovakia (from 42 up to 98 EUR/ha), in the Poland (from 25 up to 46 EUR/ha). The lowest sums of subsidies are paid in the Hungary, c. 0,7 EUR/ha. For investigation the impacts of LFA on economic outcomes of farms I have compared Gross Farm Income, Farm Net Value Added, Family Net Income, total amount of current subsidies and LFA payments. This comparison is illustrated on the next figure (Fig. 7).

If we do compare the chosen economic ratings we would see at first sight there is negative Family Net Income in the Slovakia, namely at LFA areas and also at the areas out of LFA. This data tells us that farms in the Slovakia are despite of subsidies receipt at loss. , Farm Net Value Added is as low as it contains just half of LFA subsidies. Rough added value overtops total amount of current subsidies by 38 thousand EUR. This data says that in the Slovakia there is unfavourable situation of farms and their strong dependence on LFA subsidies.

Fig. 7 Comparison of the chosen economic ratings of farms in the year 2006 (EUR) (self-made)

		Gross Farm Income	Farm Net Value Added	Family Net Income	Total subsidies	LFA subsidies
Czech Republic	**Total (LFA)**	**120504**	**94026**	**23826**	**62780**	**6557**
	LFA - mountain areas	99735	81261	29265	85627	25459
	LFA - not mountain areas	93129	72925	23779	62523	10597
	not in LFA	**136658**	**105935**	**22937**	**59075**	**1522**
Slovakia	**Total (LFA)**	**142752**	**10241**	**-101158**	**104380**	**28013**
	LFA - mountain areas	224186	15658	-130433	197486	78635
	LFA - not mountain areas	142901	3454	-112939	115620	26482
	not in LFA	**95193**	**13573**	**-72853**	**39425**	**0**
Poland	**Total (LFA)**	**14239**	**10706**	**9087**	**4921**	**438**
	LFA - mountain areas	10572	8039	7066	4177	527
	LFA - not mountain areas	13670	10473	9067	5779	861
	not in LFA	**15052**	**11113**	**9235**	**4090**	**0**
Hungary	**Total (LFA)**	**24409**	**18076**	**7930**	**10367**	**13**
	LFA - mountain areas					
	LFA - not mountain areas	20510	15143	6091	9448	33
	not in LFA	**26927**	**19970**	**9117**	**10960**	**0**

In the Czech Republic there is better situation. We gain positive Family Net Income and LFA subsidies contain c. ¼ of this income. Contrariwise current subsidies are, compared to Family Net Income, double, it means in case of not received subsidies Czech farms would be at loss. There is also at the first sight reasonless data in LFA subsidies for farms out of LFA in the Czech Republic. This distortion is given by continuous split-off of certain areas, so called S^x, out of LFA and compensating their payments because of the split-off.

The portion of LFA subsidies in the Poland is only 10 % of total subsidies and they reach Family Net Income neither, it means in case of not received subsidies farms could gain profit. There is low difference between Gross Farm Income, Farm Net Value Added and Family Net Income compared to the Czech Republic and the Slovakia. This situation is caused by low rebates and extreme factors such as low labour costs, ground rent, or interests.

The portion of LFA subsidies on total subsidies in the Hungary is imponderable. Although if we do compare Family Net Income and subsidies we would see that in case of not received subsidies farms would not gain any profit.

4. CONCLUSION

The accession to the European Union has resulted in a considerable increase in the amount of financial support allocated for agriculture. The EU budget subsidies for payments made within the scope of the Common Agricultural Policy, as well as investment subsidies from the EU funds, have influenced significantly the favourable trend. Following the EU accession, there was a change in funding of agriculture, which consisted in a reduced share of indirect subsidies for subsidies obtained directly by agricultural producers. The CAP brings also the rules which must farmers keep to obtain subsidies. These rules are focused on protection of environment and water, welfare of animals, extensive farming etc. and determinate new direction of agriculture in EU.

Subsidies from CAP are very important for farms in all focused countries. The Slovakia has the worst economic output. The best output gains the Polish farms. Indicator of income minus operating subsidies was in Czech Republic, Slovakia and Hungary in all years in minus. It means that farms would be in loss without subsidies. Farms in Poland reached more Family Farm income than got operating subsidies so only these farms would be in profit without subsidies.

LFA subsidy is given by the geographical position of the country, therefore e.g. in the Hungary is mountain area tariff missing. And thanks to favourable conditions for agriculture contains this subsidy only marginal account of all the subsidies to the agriculture sector. In the Czech Republic, the Slovakia, and the Poland there is c. 50 % of farmland included in less-favoured areas, therefore this subsidy has more important role here. This geographical indicator also affects the impact of this subsidy on economic outcomes of farms. If we do compare the difference between the incomes and outcomes we would see significant difference between farms which are farming in less-favoured areas and out of them. Then there are important differences between all the countries. There farms gain negative Family Net Income and lower added value than the amount of LFA subsidies is paid (these subsidies amount at most c. ¼ of total subsidies in LFA areas). If we subtract the total amount of current subsidies from the net added value in the other countries, we would receive negative Family Net Income. The lowest effect of LFA subsidies is, of course, on the Hungarian farms.

Acknowledgement

Author is thankful to the Research Support Fund of Faculty of Management and Economics in Tomas Bata University in Zlin for financial support to carry out this research.

BIBLIOGRAPHY

AGRICULTURE AND FOOD ECONOMY IN POLAND. Warsaw: The Ministry of Agriculture and rural development of the Poland, 2008.

AGRICULTURE: HEALTH CHECK TO STREAMLINE COMMON AGRICULTURAL POLICY AND ADDRESS NEW CHALLENGES (on line). 20.11.2007. [cit. 25.4.2008]. Available from: http://europa.eu/rapid/pressReleasesAction.do?reference=IP/07/1720&format=HTML&aged=1&language=CS&guiLanguage=enFADN. Farm accountancy data network CR (on line). 2008. [cit. 25.2.2008]. Available from: http://www.vsbox.cz/fadn/

FADN. FARM ACCOUNTANCY DATA NETWORK EU (on line). 2008. [cit. 25.2.2008]. Available from: <http://ec.europa.eu/agriculture/rica/index_en.cfm>

NEW HUNGARY RURAL DEVELOPMENT PROGRAMME. Ministry of Agriculture and Rural Development of Hungary. (on line). September 2007. [cit. 29.10.2008 Available from: <http://www.fvm.hu/main.php?folderID=945>

PROGRAM ROZVOJA VIDIEKA SR 2007 – 2013. Ministerstvo podohospodarstva Slovenskej republiky. (on line). November 2007. [cit. 29.10.2008]. Available from: <http://www.land.gov.sk/sk/index.php?navID=2&navID2=2&sID=26&id=506>

REPORT ON AGRICULTURE AND FOOD SECTOR IN THE SLOVAK REPUBLIC 2008. Bratislava: 2008. ISBN 978-80-89088-73-7.

RURAL DEVELOPMENT PROGRAMME FOR 2007-2013. Ministry of Agriculture and Rural Development of Poland. (on line). July 2007. [cit. 29.10.2008]. Available from: <http://www.minrol.gov.pl/DesktopDefault.aspx?TabOrgId=1210&LangId=1>

THE HUNGARIAN AGRICULTURE AND FOOD INDUSTRY IN FIGURES. Budapest: The Ministry of Agriculture of the Hungary, 2008.

THE NATIONAL STRATEGIC RURAL DEVELOPMENT PLAN OF THE CZECH REPUBLIC for the period 2007 – 2013. The Ministry of Agriculture of CR (on line). January 2008. [cit. 16.1.2008]. Available from: http://www.mze.cz/UserFiles/File/EAFRD/PRV_oficiln_schvlen.pdf

THE TREATY ESTABLISHING THE EUROPEAN COMMUNITY (on line). 2008. [cit. 25.4.2008]. Available from:
< http://eur-lex.europa.eu/cs/treaties/dat/12002E/htm/C_2002325EN.003301.html>

THE STATE AGRICULTURAL INTERVENTION FUND (on line). 2008. [cit. 12.3.2008]. Available from: <http://www.szif.cz>

2004 SUMMARY REPORT. The Ministry of Agriculture of CR. May 2008. [cit. 11.5.2008]. Available from: < http://www.mze.cz/UserFiles/File/kOTTOVA/Green%20report%202004/agriculture.pdf>

ESTIMACIÓN DE LA MATRIZ DE PRODUCCIÓN DE LOS CUADROS DE OFERTA DE LA CUENTA SATÉLITE DEL TURISMO DE ESPAÑA DE 1999 (BASE 1995)

EVA Mª BUITRAGO ESQUINAS
E-mail: esquinas@us.es
Departamento de Economía Aplicada III
Universidad de Sevilla

Resumen

La Cuenta Satélite del Turismo de España (CSTE) no calculó los cuadros de oferta de 1999 puesto que para ello era necesario disponer de las tablas input-output (TIO) de dicho año y estas se publicaron con posterioridad a las de 2000. Tras la publicación de las TIO de 1999, el Instituto Nacional de Estadística no ha considerado entre sus prioridades volver atrás para completar la serie. Por ello, en este trabajo nos planteamos la estimación de la matriz de producción de las ramas de actividad características del turismo desagregada por productos característicos y no característicos del turismo: matriz 22x15. Para ello, hemos empleado una metodología de interpolación específica a parir de los datos de oferta de la CSTE de 1998 y 2000 y de las TIO de la economía española de 1998, 1999 y 2000.

Palabras clave: Estadísticas de Turismo, Cuenta Satélite del Turismo de España, Análisis input-output.
Área Temática: Economía Urbana, regional y local

Abstract

The Tourism Satellite Account of Spain (CSTE) did not calculate the supply tables for 1999. To do that it was necessary to have obtained the input-output tables (TIO) for that year and those tables were published after the ones for the year 2000. After divulging the TIO for 1999, the National Institute of Statistic of Spain has never considered to complete the before quoted series as one of its priorities. That is the reason why, in this work, we try to establish the estimation of the Production Matrix for the activity branches characteristic of tourism disaggregating products characteristics and non-characteristics of tourism (matrix products-branches 22x15). To that ain we have used an specific methodology of interpolation starting from the data supplied by CSTE in 1998 and 2000 and the ones by TIO related to the Spanish economy in 1998, 1999 and 2000.

Thematic Area: Urban, Regional and Local Economics.

1. PLANTEAMIENTO Y OBJETIVOS.

La CSTE es una fuente estadística que permite la medición y el análisis de los impactos económicos del turismo en España desde 1995 teniendo en cuenta para su delimitación el enfoque de demanda. Asimismo, es un satélite de la Contabilidad Nacional; por lo que emplea metodologías coherentes con ella. Por otro lado, para su elaboración se han seguido las recomendaciones de la OMT (OFICINA DE ESTADÍSTICA DE LAS NACIONES UNIDAS, 2000); por lo que sus resultados son comparables a nivel internacional.

La CSTE parte de la estimación de la *demanda turística* y, a partir de ella, calcula la producción de las ramas que la abastecen. La OMT clasifica las ramas que abastecen la demanda turística en características y no características del turismo. Las ramas características son aquellas que dejarían de existir o su consumo se vería sensiblemente reducido en ausencia

de turismo[1]. Al conjunto de ramas características se les denomina *industria turística*. Esta industria se caracteriza por presentar una demanda dual; por un lado, turistas y, por otro, residentes. De este modo, la demanda turística adquiere tanto productos elaborados por la industria turística como por otras industrias (enfoque de demanda); y, a su vez, la industria turística vende su producción tanto a turistas como a residentes (enfoque de oferta). La información recogida en la CSTE permite integrar ambos enfoques (de demanda y oferta) al estimar la *producción turística*: aquella que satisface la demanda turística con independencia de que haya sido o no producida por la industria turística.

Por todo ello, podemos afirmar que la CSTE es, en la actualidad, la fuente más completa para la medición y el análisis de la dimensión económica del turismo en España. No obstante, la información de oferta que ofrece está incompleta al producirse un salto en los datos de 1999.

El INE no publicó los cuadros de oferta de la CSTE 1999 (base 1995) dado que para su cálculo es imprescindible disponer de tablas input-output (TIO) y las de ese año se publicaron con posterioridad a las de 2000. Una vez publicada las TIO de 1999 el INE no ha considerado ni considera entre sus prioridades volver atrás en sus estimaciones y completar la serie. Este hecho explica el salto que existe en la información de oferta de la CSTE.

Por ello, en este trabajo, nos hemos planteado *estimar la matriz de producción por ramas y productos de 1999* con el objetivo de poder disponer de unas series completas de información de oferta y comparables con las series de demanda[2].

2. METODOLOGÍA GENERAL Y FUENTES.

Tras una consulta realizada al INE, este nos recomendó una interpolación a nivel agregado entre los detalles de 1998 y 2000 tomando como referencia las TIO de 1999. Nosotros hemos optado por una *interpolación más desagregada por ramas y productos* utilizando, no solo las TIO de 1999, sino también las de 1998 y 2000.

Con el objetivo de obtener la información de de la matriz de producción por ramas y productos de la CSTE de 1999, partimos de los datos reales y definitivos existentes recogidos en las TIO de 1999 hoy publicadas y, a partir de ellos, elaboramos nuestras estimaciones únicamente cuando sea necesario. En este sentido, nos apoyamos en la información recogida en las siguientes fuentes:

a. *Tablas input-output de la economía española de 1998, 1999 y 2000* (base 1995) con información desagregada por ramas de actividad y productos (INE, varios años, a).

b. *Cuadros de oferta de la CSTE* (base 1995) *de 1998 y 2000* (año anterior y posterior al que pretendemos estimar) con información desagregada por ramas de actividad y productos (INE, varios años, b).

La base contable en la que están elaboradas ambas fuentes es la misma: base 1995, y las dos fuentes forman parte de la Contabilidad Nacional de España (CNE); por lo que sus metodologías son coherentes y, por lo tanto, comparables. En ambas fuentes los datos de producción se desagregan tanto por ramas de actividad como por productos; el principal problema radica en que estas desagregaciones no son las mismas. La CSTE desagrega a un mayor nivel que la CNE base 1995 (CNAE-93 y CNPA-96) las actividades y productos característicos del turismo y, por el contrario, agrega más las actividades y productos no característicos[3].

[1] En España, las ACTIVIDADES CARACTERÍSTICAS DEL TURISMO se han concretado en la siguiente lista: hoteles y similares, actividades de alquiler inmobiliario, restaurantes y similares, transporte de viajeros por carretera, transporte por ferrocarril, transporte marítimo de pasajeros, transporte aéreo, agencias de viajes, anexos al transporte, alquiler de bienes de equipo de transporte, actividades culturales, recreativas y deportivas de mercado, actividades culturales, recreativas y deportivas de no mercado. (INE, 2002).

[2] Los cuadros de oferta de la CSTE incluyen dos grandes grupos de información: *producción de cada rama característica del turismo por productos* característicos y no característicos y *distribución de la producción* de las diversas ramas en consumos intermedios (a su vez desagregados por grupos de productos), impuestos y VAB. En este trabajo únicamente nos planteamos realizar la estimación de la información del primer grupo.

[3] La CSTE desagrega los productos no característicos únicamente en: *bienes*, *márgenes de distribución* y *otros servicios*. Los *bienes* se corresponden con los grupos 1 a 59 menos 3 de las TOD base 1995 (es decir, todos los productos); los márgenes de distribución se identifican con los márgenes comerciales y de transporte de las TOD (salvo los márgenes de transporte que ya se han incluido en

ESQUEMA 1. Equivalencias de las clasificaciones por ramas de actividad utilizadas en la CSTE y en la CNE (base 1995)

CSTE *(Ramas de actividad características del turismo)*	**CNE base 95 (CNAE 93)** **Marco I-O** *(A.70 ramas)* (1)
Ramas de alojamiento	**44 (P) + 54**
Hoteles y similares	44 (P). Hostelería (P)
Actividades de alquiler inmobiliario	54. Alquileres inmobiliarios
Restaurantes y similares	**44 (P).** Hostelería (P)
Transporte de pasajeros	**46 (P) + 45 + 47 (P) + 48**
Transporte de viajeros por carretera	46 (P). Transporte terrestre y transporte por tubería (P)
Transporte por ferrocarril	45. Transporte por ferrocarril
Transporte marítimo de pasajeros	47 (P). Transporte marítimo (P)
Transporte aéreo	48. Transporte aéreo y espacial.
Agencias de viaje	**49 (P).** Actividades anexas a los transportes (P).
Anexos al transporte	**49 (P).** Actividades anexas a los transportes (P).
Alquiler de vehículos.	55 (P). Alquiler de maquinaria y enseres domésticos (P).
Activ. culturales, recreativas y deportivas	**63 + 69.2**
Activ. culturales, recreativas y deportivas de mercado	63. Actividades recreativas, culturales y deportivas.
Activ. culturales, recreativas y deportivas de no mercado	69.2 Actividades recreativas y culturales de las AAPP.

NOTA: (P) Únicamente recoge parte de la rama característica del turismo. (1) En la base 1995, las ramas en las que quedan incluidas las actividades características son idénticas en el núcleo central de la CNE y en el marco I-O.
FUENTE: Elaboración propia a partir de: INE, *CSTE y CNE (base 1995)*, www.ine.es (consultado en enero 2008).

las actividades de transporte características del turismo) y los otros servicios con aquellos servicios que no se incluyen en las actividades características del turismo.

ESQUEMA 2. Equivalencias de las clasificaciones por productos de la CSTE y la CNE (base 1995).

CSTE *(Productos característicos del turismo)*	**CNPA-96**	**CNE (base 1995)** **Marco I-O** *(110 grupos de productos)*
Servicios de alojamiento		
- Hoteles y similares	*55.1 y 55.2*	68. Servicios de hostelería
- Servicios de alquiler inmobiliario	*70*	81. Servicios inmobiliarios
Restaurantes y similares	*55.3, 55.4 y 55.5*	69. Servicios de restauración
Transporte de pasajeros		
- Transporte de viajeros por carretera	*60.2 (P) y 60.3 (P)*	71 (P). Otros tipos de transporte terrestre
- Transporte de viajeros por ferrocarril	*60.1 (P)*	70 (P). Servicios de transporte por ferrocarril
- Transporte marítimo de viajeros	*61 (P)*	72 (P). Servicios de transporte marítimo
- Transporte aéreo de viajeros	*62 (P)*	73 (P). Servicios de transporte aéreo y espacial
Servicios anexos al transporte	*63.1 (P), 63.2 (P) y 63.4 (P)*	75 (P). Otros servicios anexos al transporte
Alquiler de bienes de equipo de transporte	*71.1*	82. Servicios de alquiler de automóviles
Servicios de agencias de viajes	*63.3*	74. Servicios de agencias de viajes
Servicios culturales, recreativos y deportivos		
-Servicios culturales, recreativos y deportivos de mercado	*92.1-92.6 (P), 92.7*	105 (P), 106 y 107. Servicios artísticos y agencias de noticias. Servicios culturales y deportivos de mercado. Otros servicios recreativos.
-Servicios culturales, recreativos y deportivos de no mercado	*92.1-92.6 (P)*	105 (P) y 108 Servicios artísticos y agencias de noticias. Servicios culturales y deportivos de no mercado.
Servicios turísticos no mercado de las AAPP	*75 (P)*	93 (P). Administración pública.

NOTA: (P) Únicamente recoge parte de la rama característica del turismo.
FUENTE: Elaboración propia a partir de: INE, *CSTE y CNE (base 1995),* www.ine.es (consultado en enero 2008).

Por ello, el primer paso que nos planteamos para la estimación matriz de producción por ramas y productos de la CSTE de 1999 es establecer la correspondencia entre las clasificaciones de ramas de actividad y productos utilizadas por la CSTE y la CNE. En este sentido, tras realizar una exhaustiva comparación, en los ESQUEMAS 1 y 2 hemos recogido las principales equivalencias en las clasificaciones por ramas y por productos respectivamente.

En aquellas ramas y productos en los que la equivalencia sea exacta, la información de oferta de 1999 se obtiene directamente de las TOD de ese año, para el resto debemos hacer algunos supuestos restrictivos y, a partir de ellos, estimar los datos.

Para la estimación de la producción por ramas y productos de 1999 siguiendo el mismo formato que los cuadros de oferta de la CSTE de otros años, necesitamos obtener una matriz (productos x ramas) que incluya: la producción de las distintas ramas características del turismo (12 ramas), de las ramas no características (1 rama agregada) y del total de ramas, desagregadas por productos característicos (13 grupos de productos), productos no característicos (3 grupos de productos) y productos totales. Por lo tanto, debemos estimar una *matriz 22x15*.

A partir de las TIO de 1998, 1999 y 2000 y los cuadros de oferta de la CSTE de 1998 y 2000 es posible conocer directamente el dato real de 1999 para un número importante de celdas: aquellas en las que las clasificaciones de la CNE y la CSTE coinciden de algún modo[4].

Para las celdas en las que esto no es posible, se han estimado los valores a partir de las metodologías que hemos considerado como más adecuadas teniendo en cuenta: las características de cada rama y grupo de productos y la información existente para cada uno de ellos.

De este modo, en el CUADRO 1, en el que se recogen los resultados de las estimaciones de la matriz de producción por ramas y productos (22X15); se presentan una serie de datos de producción de 1999 que pueden tener la consideración de reales y definitivos y otros que son de estimación propia, con las limitaciones que ello puede suponer.

En general, para la elaboración de estas *estimaciones* hemos calculado la estructura (% que representan las desagregaciones de la CSTE respecto de la rama más agregada de la CNE) de las diferentes ramas para 1998 y 2000 y hemos supuesto que la estructura de 1999 es el punto medio entre la de estos dos años. Por ello, se aplican dichos porcentajes a los datos de la tabla de origen de 1999 para la obtención de los cuadros de oferta de 1999. De esta forma recogemos tanto la evolución que han experimentado las distintas variables (al menos, la evolución de las ramas y productos considerados en la CNE) como la dinámica interna de su estructura.

Por otro lado, contamos con una información adicional: dado que queremos obtener una matriz por productos y por ramas, podemos comenzar estimando las columnas (producción de cada rama por productos) y, una vez hecho esto, utilizar las filas (producción de cada grupo de productos que ha producido cada rama) para contrastar las estimaciones realizadas y completar la información si se ha producido alguna carencia. Estos dos procedimientos son complementarios y es posible establecer un feed-back entre ellos.

Dada la enorme heterogeneidad existente entre las ramas de actividad características del turismo (tanto conceptual como en la información disponible) hemos considerado oportuno aplicar diversos *procedimientos específicos* para la estimación de la producción de cada rama y cada grupo de productos. En todo caso, hemos tratado de respetar las características de cada rama / grupo de productos y las peculiaridades de la información disponible, tomando siempre como punto de partida los datos reales existentes para 1999 y estimando únicamente los imprescindibles.

3. METODOLOGÍA ESPECÍFICA PARA LA ESTIMACIÓN DE LA MATRIZ DE PRODUCCIÓN POR RAMAS DE ACTIVIDAD Y GRUPOS DE PRODUCTOS DE LA CSTE 1999.

1. Datos de las ramas de hostelería y restauración:

A partir de la tabla de origen de 1999 se conoce con exactitud la producción de la rama de *hostelería* (HH) desagregada por productos; aunque desconocemos como se distribuye entre sus dos componentes principales: *hoteles y similares* y *restaurantes y similares*. Así, en este caso tenemos información de una rama característica del turismo pero a un nivel más agregado del que necesitamos, por lo tanto, el objetivo es repartir dicha producción de hostelería desagregada por productos en las dos ramas que considera la CSTE. Dado que esta rama no tiene producción en los grupos de productos problemáticos (los que no coinciden las clasificaciones), no es necesario realizar ningún ajuste en este sentido.

A partir de la observación y comparación de las tablas de origen (TO) de 1998 y 2000 y los cuadros de oferta de estos dos años, se sabe que la producción de servicios de hostelería y de otros servicios anexos al transporte se asignan completas a la rama de hoteles y similares; por lo que estos valores para 1999 se pueden obtener directamente de la TO 1999 (datos reales).

[4] Como se recoge en el ESQUEMA 1, por ramas de actividad coinciden: la suma de hoteles y similares y restaurantes y similares de la CSTE con la rama de hostelería de la CNE; la rama de alquiler inmobiliario, las ramas de transporte aéreo y por ferrocarril; las actividades culturales, recreativas y deportivas de mercado y no mercado. Como se recoge en el ESQUEMA 2, Por productos, coinciden: los hoteles y similares, los servicios de alquiler inmobiliario, los restaurantes y similares, los servicios de agencias de viaje, los servicios anexos a los transportes y los servicios culturales, recreativos y deportivos. Por ello, aquella celda fruto del cruce de estas ramas y estos productos se pueden obtener directamente de la TO de 1999.

Dado que para 1998 y 2000 existe información tanto de las TIO como de los cuadros de oferta de la CSTE, puedo conocer los porcentajes que representan para esos dos años la producción de la rama de hoteles y similares (CSTE) respecto del total de la hostelería (TO) y la producción de restaurantes y similares (CSTE) respecto del total (TO), y todo ello desagregado para los distintos grupos de productos. Supongo que estos porcentajes para 1999 son el punto medio entre los calculados para 1998 y 1999 (no obstante, para los productos característicos estos porcentajes se mantienen constantes en los años considerados, por lo que el porcentaje de 1999 es prácticamente idéntico al de estos años). Aplico dichos porcentajes medios a los datos globales de hostelería de 1999 (TO 1999) y obtengo los datos de producción de las ramas de hoteles y similares y restaurantes y similares (CSTE) desagregados por productos[5]. Los datos del total de productos característicos y no característicos para cada una de las dos subramas se obtienen por agregación de los cálculos anteriores.

Es importante señalar que con la metodología seguida se han respetado todas las identidades contables ya que el total de la suma por filas y por columnas de las estimaciones para los cuadros de oferta de 1999 coinciden con la suma por filas y por columnas de la tabla de origen de 1999 (producción total de los distintos grupos de productos y producción total de la rama de hostelería). Se han podido producir algunas pequeñas desviaciones irrelevantes por el redondeo de decimales.

2. Rama de alquiler inmobiliario:

La rama de alquiler inmobiliario de la CSTE coincide con la utilizada en las TO y su producción desagregada por productos únicamente incluye su producción principal: producción de servicios de alquiler inmobiliario, y producción de otros servicios (otras construcciones y servicios de informática). Estos grupos de productos son idénticos en ambas fuentes; por lo que los datos de esta rama para completar la información de oferta de la CSTE de 1999 se pueden obtener directamente de la TO de 1999: datos reales, no estimados.

3. Ramas de transporte de viajeros por carretera.

Las *ramas* de transporte por ferrocarril y transporte aéreo y espacial de la CSTE corresponden con las de la CNE[6]; sin embargo, las ramas de transporte marítimo y por carretera de la CSTE únicamente consideran el transporte de pasajeros por esa vía mientras que en la CNE se considera el transporte total. Por lo tanto, para las ramas de transporte por ferrocarril y aéreo podemos obtener los datos directamente de las TO y para las ramas de transporte marítimo y por carretera deberemos realizar un trabajo de estimación para calcular la parte que se corresponde con transporte de pasajeros.

Por otro lado, la desagregación por *productos* de la CSTE únicamente recoge los servicios de transporte de pasajeros mientras que la desagregación por productos de las TO considera los servicios de transporte de pasajeros y mercancías conjuntamente. Por ello, será necesario estimar la parte que corresponde al servicio de transporte de viajeros en cada una de las producciones de estos servicios de cada rama.

Teniendo en cuenta los dos aspectos anteriores y tras realizar una comparación exhaustiva de las TO y los cuadros de oferta de 1998 y 2000, podemos concluir que: salvo el caso del transporte aéreo y espacial, el resto de ramas de transporte características del turismo desagregadas por productos (CSTE) no coinciden en todos sus valores con las ramas de transporte de la CNE; lo que nos ha obligado a estimar la mayor parte de los datos de 1999 para dichas ramas y grupos de productos.

Partimos de la *rama de transporte aéreo y espacial* que, como coincide conceptualmente en las dos fuentes consideradas, es posible obtener los datos reales de producción de 1999 a partir de la TO de dicho año (no tienen producciones de los grupos de productos problemáticos).

[5] Se ha aplicado cada porcentaje a la producción de los diferentes grupos de productos y, posteriormente, por agregación de los datos obtenidos se calculan los totales para cada una de las dos ramas (total producción de productos característicos, total producción de productos no característicos y producción total de cada rama).

[6] En estos casos, la CSTE considera de forma conjunta el transporte de pasajeros y el de mercancías. En el caso del transporte aéreo, porque el transporte de mercancías que se realiza por esta vía es irrelevante y, en el caso del transporte por ferrocarril, porque en España tanto el transporte de mercancías como el de pasajeros lo realiza la misma empresa en régimen de monopolio: RENFE.

Para la *rama de transporte por ferrocarril* el único dato que no coincide en las TO y en los cuadros de oferta de la CSTE es la producción de servicios de transporte de pasajeros por ferrocarril ya que en las TO se incluye el conjunto de servicios de transporte por ferrocarril. En este caso, para la estimación de dicho dato se calculan los porcentajes que han representado los servicios de transporte de viajeros por ferrocarril sobre los servicios totales de trasporte por ferrocarril para 1998 y 2000 (a partir de las TO y cuadros de oferta de las CSTE de dichos años) y se aplica la media a los datos de servicios de transporte por ferrocarril de 1999 (TO). El resto de datos se obtienen directamente de la TO de 1999[7].

Como se ha señalado, la *rama de transporte por carretera* considerada en la CSTE no coincide con la de la CNE ya que como actividad característica del turismo únicamente incluye la rama de trasporte de viajeros por carretera. Por ello, ha sido necesario estimar todas las producciones de la rama desagregadas en los distintos productos. Para ello, hemos seguido la misma metodología empleada hasta el momento: calcular los porcentajes que representa la producción de transporte de viajeros por carretera para cada grupo de productos sobre la producción total del transporte por carretera para los años 1998 y 2000 y aplicar la media de dichos porcentajes al dato de la TO de 1999.

Finalmente, para estimar los datos de la *rama de transporte marítimo* de pasajeros hemos utilizado el mismo procedimiento.

4. Rama de agencias de viaje y anexos a los transportes

Estas dos ramas de la CSTE se encuentran agregadas en *la rama de actividades anexas a los transportes* de la TO (rama 49); por lo que observando la distribución de su producción por productos (columna de la TO) se pueden obtener algunas conclusiones importantes: la producción de la rama de AAVV de la CSTE desagregada por productos se corresponderá con la producción de servicios de AAVV y de servicios auxiliares a la intermediación financiera de la rama conjunta de actividades anexas a los transportes (TO 49) y, por lo tanto, el resto de la producción de esta rama se corresponde con la *rama de anexos a los transportes* de la CSTE. De este modo, dado que conocemos la TO de 1999, podemos obtener directamente de ella los datos reales (no estimados) de la rama de anexos a los transportes de la CSTE.

Por su parte, la *rama de AAVV* de la CSTE presenta una peculiaridad importante respecto a los valores de la TO: en la CSTE no se considera el valor completo de los productos de AAVV porque estos incluyen los paquetes turísticos y la CSTE ha tratado de repartir la producción de dichos paquetes en sus ramas correspondientes. Por ello, el valor de los servicios de AAVV de la TO de 1998 y 2000 es superior al de los cuadros de oferta de la CSTE de estos años. Esto nos obliga a estimar el dato de 1999, ya que no es posible obtenerlo directamente a partir de estas fuentes. En este caso, no utilizamos la metodología explicada en los casos anteriores porque disponemos de información adicional que nos permite conocer el valor real. Dado que la producción que la CSTE imputa a las AAVV durante todos los años para los que se dispone de información es exactamente la misma que el *consumo interior total de servicios de agencias de viaje* y los cuadros de demanda de la CSTE de 1999 se conocen, se tiene el dato real de la producción de servicios de agencias de viaje por parte de la rama de AAVV (3.773,9 millones de €). Para la producción de productos no característicos por parte de la rama de AAVV de 1999, tenemos el dato real ya que se corresponde únicamente con los servicios auxiliares a la intermediación financiera de la rama conjunta de actividades anexas a los transportes.

5. Alquiler de vehículos.

En el marco input -utput esta rama queda incluida en la de alquiler de maquinaria y enseres domésticos (rama 71). Esta rama de las TIO incluye dos grandes producciones principales (la de servicios de alquiler de automóviles y la de alquiler de maquinaria y enseres domésticos) y un importante número de producciones secundarias vinculada a cada una de ellas. Para la estimación de la producción de la rama de alquiler de vehículos de las CSTE únicamente nos

[7] Los *bienes* se obtienen como suma de toda la producción de todos los productos (grupos 1 a 59 menos 3 de las TOD base 1995); los *márgenes de distribución* se identifican con los márgenes comerciales y de transporte de las TOD (salvo los márgenes de transporte que ya se han incluido en las actividades de transporte características del turismo) y los *otros servicios* corresponden a aquellos servicios que no se incluyen en las actividades características del turismo, es decir, se calculan por diferencia entre los productos no característicos y la suma de los dos grupos anteriores.

interesa la producción de servicios de alquiler de automóviles y las producciones secundarias asociadas a esta. A partir de la TO de 1999 conocemos el dato real de producción de servicios de alquiler de automóviles de la rama de alquiler de maquinaria y enseres domésticos, este dato se corresponde con el de la producción de alquiler de bienes de equipo de transporte de la CSTE (lo cual se ha comprobado con las TIO y los cuadros de oferta de 1998 y 2000). Por lo tanto, para completar la información de los cuadros de oferta de 1999 para la rama de alquiler de vehículos necesitamos estimar los servicios de alquiler inmobiliario y los otros servicios que produce esta rama. Para ello, calculamos los porcentajes que han representado en 1998 y 2000 estos grupos de productos respecto de la producción principal de la rama (servicios de alquiler de automóviles) y aplicamos los puntos medios al valor principal del dato de 1999 (que conocemos a partir de la TO 1999). Con esta estimación estamos considerando tanto la evolución en las cantidades globales como las diferencias en la estructura de la rama.

6. Ramas de actividades culturales, recreativas y deportivas.

La rama de actividades culturales, recreativas y deportivas de mercado se corresponde en su totalidad con la rama 63 de la TO y la rama de actividades culturales, recreativas y deportivas de no mercado con la 69.2; por lo que únicamente ha habido que ajustar algunos elementos de la clasificación por productos para calcular sus producciones[8] (por lo tanto, obtenemos datos reales, no estimados). Es de destacar que rama de actividades culturales y recreativas de las ISFLSH se incluye como rama no característica.

7. Conjunto de ramas características del turismo.

Siguiendo los procesos descritos anteriormente, hemos obtenido la matriz de producciones de las ramas de actividad características (columnas) del turismo desagregada por productos (filas). Por lo tanto, por agregación de las columnas, se obtiene la producción de estas ramas (tanto la de productos característicos como de no característicos). Por agregación de las filas, obtenemos la producción de productos característicos y no característicos que realizan las ramas de actividad características.

8. Ramas no características del turismo y producción total de los diferentes grupos de productos.

Una vez completa la matriz de producciones de las ramas características del turismo desagregada por productos de 1999, nos planteamos completar la información de la producción de las ramas no características. Para ello, nos hemos centrado en el análisis por filas: *producción de los distintos grupos de productos*.

En este sentido, para aquellos *grupos de productos en los que las clasificaciones de la CSTE y la CNE coincide*n (hoteles y similares, servicios de alquiler inmobiliario, restaurantes y similares, servicios de transporte aéreo, otros servicios anexos al transporte, alquiler de bienes de equipo de transporte, servicios culturales recreativos y deportivos de mercado y de no mercado) el procedimiento es el siguiente: dado que conocemos la producción total para cada uno esos grupos de productos (TO 1999) y la producción de las ramas características (estimada anteriormente); por diferencia entre ellas, se obtiene la producción de las ramas no características.

Para los *servicios de transporte*, en los que hay importantes diferencias entre las desagregaciones realizadas por la CSTE y la TO, hemos empleado la siguiente metodología: se ha observado la estructura de la producción reflejada en las TO de 1998 y 2000 y se ha comparado con los cuadros de oferta de la CSTE para esos años y se han extrapolado las conclusiones para la elaboración de los datos de 1999. Según los datos de las TO de 1998 y 2000, la producción de servicios de transporte por carretera es realizada principalmente por la rama de transporte por carretera, aunque también hay implicadas otras ramas. Tras la comparación con los datos de oferta de las CSTE de esos años llegamos a la conclusión de que

[8] El grupo de servicios culturales, recreativos y deportivos de mercado utilizado en la desagregación de la CSTE se corresponde con la suma de tres grupos de la TO: servicios artísticos y agencias de noticias, servicios culturales y deportivos de mercado y otros servicios recreativos.

únicamente se considera que realicen producción de servicios de transporte de viajeros la rama de educación de mercado; por ello, únicamente asignaremos el valor de esta producción de servicios de transporte de viajeros por carretera a las ramas no características. Por agregación con la producción de las ramas características obtenidas anteriormente, se calcula la producción total. Tanto la producción de servicios de transporte por ferrocarril como marítima es íntegramente realizada por las ramas de transporte por ferrocarril y transporte marítimo; por ello, las ramas no características no producen este grupo de productos.

La producción de *servicios de agencias de viaje* es realizada en su totalidad por la rama de agencias de viaje; por lo que tampoco hay producción de este grupo de productos por parte de las ramas no características.

La producción de *servicios turísticos de las AAPPP* no es producida por ninguna rama característica del turismo (esta producción se corresponde con la rama de administraciones públicas) y, por lo tanto, queda incluida en las ramas no características. Para calcular la producción de servicios turísticos de las AAPP hemos recurrido a los cuadros de demanda de la CSTE de 1999: consumo turístico interior desagregado por productos. En ellos se incluye el dato de *consumo turístico interior de servicios turísticos de las AAPP* que, como hemos podido comprobar tras la comparación de los años para los que hay información de oferta y demanda, coincide plenamente con el dato de producción de dichos servicios. De este modo obtenemos el dato real.

Por agregación obtenemos la producción de todos los productos característicos total y desagregada por *ramas características* y *no características*.

El total de la producción para el conjunto de las ramas características y no características y para el conjunto de productos característicos y no característicos de la CSTE no coincide con el de las TO puesto que se le ha deducido la parte de los paquetes turísticos que se habían contabilizado doblemente (en agencias de viaje y en sus componentes); por ello, la producción total en la CSTE se diferencia de la producción total de la CNE en la misma cantidad que se diferencian los servicios de AAVV en ambas fuentes.

Si a la producción total de la economía española (restándole la doble contabilidad de los paquetes turísticos) le deducimos la producción de productos característicos obtenemos la producción de productos no característicos del turismo.

Para la estimación de los datos de la desagregación de la producción de los productos no característicos (bienes, márgenes y otros servicios) seguimos los siguientes procedimientos. Los *bienes* se calculan por agregación de los grupos de productos de las TO. El total de la fila de los *márgenes de distribución* de la CSTE se obtiene de los márgenes de comercio y transporte incluidos en la TO (columnas para transformar los datos a precios básicos en precios de adquisición). Los márgenes de comercio se consideran completos, pero a los márgenes de transporte hay que deducirles la parte que ya queda incluida en la producción de las ramas características de transporte. Dado que ese dato no nos es posible obtenerlo directamente de la TO de 1999, lo estimamos. Para ello, se calcula el porcentaje que supone los márgenes totales de transporte sobre la producción total de servicios de transportes y aplicamos dicho porcentaje a la producción de servicios de transporte de viajeros. Este margen ya está incluido en la producción de servicios de viajeros por las distintas vías, por ello tenemos que deducirlo del total de márgenes.

La fila de *otros servicios* se calcula por diferencia entre el total de productos no característicos (que a su vez se calculó como la diferencia entre el total de la producción y la producción de productos característicos) y los bienes y márgenes de distribución estimados siguiendo la metodología especificada anteriormente. De este modo se completa la última columna de la matriz: CUADRO 1.

4. RESULTADOS: MATRIZ DE PRODUCCIÓN POR RAMAS Y PRODUCTOS DE LA CSTE 1999.

En el CUADRO 1 hemos recogido los resultados de los cálculos anteriores; de modo que sintetiza los ***cuadros de producción estimados para 1999***.

CUADRO 1. Matriz de producción por productos y ramas de actividad. 1999 (E). Base 1995. (Millones €).

	Hoteles y similares	Alquiler inmobiliario	Restaurantes y similares	Tte. de viajeros por carretera	Tte. por ferrocarril	Tte. marítimo de pasajeros	Tte. aéreo	AAVV	Anexos al transporte	Alquiler de vehículos	Activ. cultur., recreat. y deport. de mercado	Activ. cultur., recreat. y dep. de no mercado	TOTAL RAMAS CARACTER. DEL TURISMO	TOTAL RAMAS NO CARÁCTER.	TOTAL PRODUCCIÓN
Sº. ALOJAMIENTO	10.700,6	53.684,9	7,5	7,6	39,3	1,3	23,6	0,0	9,2	30,7	101,3	0,0	64.605,9	2.246,2	66.852,1
- Hoteles y similares	10.671,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	10.671,2	86,4	10.757,6
- Sº. alquiler inmobiliario	29,4	53.684,9	7,5	7,6	39,3	1,3	23,6	0,0	9,2	30,7	101,3	0,0	53.934,7	2.159,8	56.094,5
RESTAURANTES Y SIMILARES	1.949,0	0,0	55.719,3	0,0	0,0	28,7	0,0	0,0	0,0	0,0	234,9	0,0	57.931,9	325,6	58.257,5
TRANSPORTE DE PASAJEROS	0,0	0,0	0,0	5.523,4	2.020,2	260,2	5.263,7	0,0	0,0	0,0	0,0	0,0	13.067,5	0,0	13.067,5
- Tte. viajeros por carretera	0,0	0,0	0,0	5.523,4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	5.523,4	0,0	5.523,4
- Tte viajeros por ferrocarril	0,0	0,0	0,0	0,0	2.020,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	2.020,2	0,0	2.020,2
- Tte maritimo de viajeros	0,0	0,0	0,0	0,0	0,0	260,2	0,0	0,0	0,0	0,0	0,0	0,0	260,2	0,0	260,2
- Tte aéreo de viajeros	0,0	0,0	0,0	0,0	0,0	0,0	5.263,7	0,0	0,0	0,0	0,0	0,0	5.263,7	0,0	5.263,7
Sº. AAVV	0,0	0,0	0,0	0,0	0,0	0,0	0,0	3.773,9	0,0	0,0	0,0	0,0	3.773,9	0,0	3.773,9
Sº. ANEXOS AL TRANSPORTE	3,2	0,0	0,0	85,1	4,1	10,8	447,6	0,0	9.883,3	0,0	0,0	0,0	10.434,1	461,6	10.895,7
ALQUILER BIENES EQUIPO DE TTE.	0,0	0,0	0,0	28,7	0,0	0,0	0,0	0,0	0,0	948,6	0,0	0,0	977,3	97,1	1.074,4
Sº. CULT., RECRE. Y DEPORT.	29,2	0,0	945,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	12.567,0	2.637,0	16.178,2	1.586,4	17.764,6
Sº. cult., recrea. y deport. de mercado	29,2	0,0	945,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	12.567,0	30,0	13.571,2	1.524,8	15.096,0
Sº. cult., recrea. y dep. de no mercado	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	2.607,0	2.607,0	61,6	2.668,6
Sº. turísticos no mercado de las AAPP	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	331,5	331,5
PRODUCTOS CARÁCTER.	**12.682,0**	**53.684,9**	**56.671,8**	**5.644,8**	**2.063,6**	**301,0**	**5.734,9**	**3.773,9**	**9.892,5**	**979,3**	**12.903,2**	**2.637,0**	**166.968,8**	**5.048,4**	**172.017,2**
Bienes	0,0	0,0	0,0	0,1	36,6	0,0	154,0	0,0	0,0	0,0	64,1	0,0	254,8	379.523,1	379.777,9
Márgenes de distribución de bienes	13,0	0,0	400,4	0,0	3,4	0,0	44,0	0,0	0,0	0,0	10,1	0,0	470,9	83.120,8	83.591,7
Otros servicios	26,4	6,1	172,4	0,0	227,4	124,8	113,1	3,7	23,5	106,3	2.325,6	0,0	3.129,3	379.169,5	382.298,8
PRODUCTOS NO CARÁCTER.	**39,4**	**6,1**	**572,8**	**0,1**	**267,4**	**124,8**	**311,1**	**3,7**	**23,5**	**106,3**	**2.399,8**	**0,0**	**3.855,0**	**841.813,4**	**845.668,4**
TOTAL PRODUCCIÓN	**12.721,4**	**53.691,0**	**57.244,6**	**5.644,9**	**2.331,0**	**425,8**	**6.046,0**	**3.777,6**	**9.916,0**	**1.085,6**	**15.303,0**	**2.637,0**	**170.823,8**	**846.861,8**	**1.017.685,6**

NOTA: Los datos son de estimación propia a partir de las TOD 1998, 1999 y 2000 (base 1995) y los cuadros de oferta de la CSTE de 1998 y 2000 (base 1995).
FUENTE: INE, *CSTE (base 1995)*, *CNE (base 1995), TOD (base 1995),* www.ine.es (consultado en mayo de 2008) y elaboración propia.

BIBLIOGRAFÍA

BUITRAGO ESQUINAS E. M., (2009): "Medición de los impactos macroeconómicos del turismo en España: Explotación de la Cuenta Satélite (CSTE) 1995-2006". *Tesis Doctoral.* Universidad de Sevilla, Sevilla.

CAÑADA MARTÍNEZ, A. (2002): "La CN según el SEC95 como marco para la estimación de la CSTE", *Revista de Estudios Turísticos*, nº 151, Instituto de Estudios Turísticos, Madrid, pags. 5-37.

CAÑADA MARTÍNEZ, A. (2004): "Instrumentos de medida del turismo: la Cuenta Satélite del Turismo", en *Papeles de Economía* nº 102, Madrid, págs. 2-28.

CARRASCO, F. (1999): *Fundamentos del Sistema Europeo de Cuentas Nacionales y Regionales. (SEC 1995)*, Ediciones Pirámide, Madrid.

EUROSTAT (1996): *Sistema Europeo de Cuentas Nacionales y Regionales 1995*, Oficina de Publicaciones de las Comunidades Europeas, Luxemburgo.

EUROSTAT (2002): *Manual Europeo de Implantación de la Cuenta Satélite de Turismo*, Oficina de Publicaciones de la Unión Europea, Luxemburgo.

INSTITUTO NACIONAL DE ESTADÍSTICA (2002): *La Cuenta Satélite del Turismo: Metodología y primeras estimaciones (1996-1999)*, INE, Madrid.

INSTITUTO NACIONAL DE ESTADÍSTICA (varios años): *La Contabilidad Nacional de España. Marco Input-Output*, Base de Datos Tempus, http://www.ine.es/tempus/.....

INSTITUTO NACIONAL DE ESTADÍSTICA (varios años): *La Cuenta Satélite del Turismo*, Base de Datos Tempus, http://www.ine.es/tempus/.....

LEONTIEF, W. (1970): *Análisis Económico Input-Output*, Ed. Gustavo Gili, S.A., Colección Ciencia Económica, Barcelona.

OFICINA DE ESTADÍSTICAS DE LAS NACIONES UNIDAS (2000): *Cuentas Satélites del Turismo: Referencias Metodológicas Comunes*, Naciones Unidas, Nueva York.

OFICINA DE ESTADÍSTICAS DE LAS NACIONES UNIDAS (2008): 2008 *Tourism Satellite Account: Recommended Methodological Framework (TSA: RMF 2008)*, http://unstats.un.org/unsd/statcom/doc08/BG-TSA.pdf.

PULIDO, A. y FONTELA, E. (1997): *Análisis input-output. Modelos, datos y aplicaciones*, Pirámide, Madrid.

IMPACTO ECONÓMICO DE LA DEPENDENCIA Y EL ENVEJECIMIENTO EN LA CIUDAD AUTÓNOMA DE CEUTA

FRANCISCO JAVIER BLANCO ENCOMIENDA
RAFAEL HERRERÍAS PLEGUEZUELO
JOSÉ CALLEJÓN CÉSPEDES
FRANCISCO JAVIER BLANCO ENCOMIENDA
E-mail: jble@ugr.es
Universidad de Granada

Resumen

El aumento de la esperanza de vida y el envejecimiento progresivo de la población están haciendo que cada vez más personas necesiten ayuda para realizar las actividades básicas de la vida diaria. En esta línea, la estimación del coste de los cuidados de larga duración es uno de los aspectos que cobra especial relevancia en torno a las personas dependientes.
En el presente estudio se aporta una metodología para estimar el número de personas de edad avanzada y dependientes en la Ciudad Autónoma de Ceuta, empleando para ello el modelo de Markov. Esto permitirá aproximar su impacto en la economía de la ciudad, la cual va a tener que seguir destinando fondos para satisfacer la demanda de cuidados de aquellos ceutíes que se encuentran en situaciones de especial vulnerabilidad.

Palabras clave: envejecimiento, dependencia, coste, cuidados de larga duración.
Área temática: Economía Urbana, Regional y Local.

Abstract

The increase of the life expectancy and the progressive aging of the population are provoking that increasingly more people need help to perform the basic activities of daily living. In this line, the estimation of the cost of long-term care is one of the aspects that receives a special relevancy around the dependent people.
In the present study a methodology is shown to estimate the number of older and dependent people in the Autonomous City of Ceuta, using for it the Markov model. This will make possible to bring its impact on the economy of the city, which is going to have to continue destining funds to satisfy the demand of care of those persons who are in situations of special vulnerability.

Key words: aging, dependency, cost, long-term care.
Thematic Area: Local, regional and urban economy.

1. PRIMERA APROXIMACIÓN

Recientemente, cuestiones como la dependencia y el envejecimiento se han convertido en un importante campo de estudio, en gran parte, debido a los cambios demográficos y sociales que están teniendo lugar desde hace algunos años.

Si al aumento de la esperanza de vida se le añaden los accidentes de tráfico y las consecuencias de enfermedades varias, el incremento en el número de personas con restricciones para desempeñar las actividades básicas de la vida diaria (ABVD) es inevitable. De ahí que los organismos públicos deban actuar y prever los servicios sociales y sanitarios (socio-sanitarios) que cubran las necesidades de este colectivo de población.

El trabajo propuesto se centra en la Ciudad Autónoma de Ceuta y pretende abordar la interrelación entre la dependencia y el envejecimiento, por una parte, y la demanda futura de

cuidados de larga duración (CLD)[1], por otra. Todo ello, bajo la pretensión de que sirva como base para la planificación económica, a corto plazo, del gasto en la atención y asistencia requeridas.

2. LA DEPENDENCIA

Aunque el término *dependencia* no es nuevo por lo que respecta a su definición, sí lo es en relación con la frecuencia en la que últimamente está apareciendo y con la problemática social y económica que está provocando.

Según la *Ley de Promoción de la Autonomía Personal y Atención a las personas en situación de Dependencia* (LPAPAD, 2006) la dependencia es el estado de carácter permanente en que se encuentran las personas que, por razones derivadas de la edad, la enfermedad o la discapacidad, y ligadas a la falta o a la pérdida de autonomía física, mental, intelectual o sensorial, requieren la atención de otra u otras personas o ayudas importantes para realizar ABVD o, en el caso de las personas con discapacidad intelectual o enfermedad mental, de otros apoyos para su autonomía personal. En este sentido, se entiende por ABVD las tareas más elementales de la persona, que le permiten desenvolverse con un mínimo de autonomía e independencia, como pueden ser: el cuidado personal, las actividades domésticas básicas, la movilidad esencial, reconocer personas y objetos, orientarse, entender y ejecutar órdenes o tareas sencillas.

De la definición anterior se desprenden tres factores básicos que delimitan la situación de una persona dependiente. En primer lugar, la existencia de una limitación física, psíquica o intelectual; en segundo lugar, la incapacidad para realizar de manera autónoma las actividades básicas de la vida diaria; en tercer lugar, la necesidad de asistencia o ayuda por parte de un tercero.

Como bien se apunta en el *Libro Blanco de la Dependencia*, editado por el IMSERSO en 2005, la dependencia puede estar presente desde el nacimiento, desencadenarse a consecuencia de un accidente o de una enfermedad aguda en la infancia, la juventud o la vida adulta o, más frecuentemente, ir apareciendo a medida que las personas envejecen, a partir de enfermedades crónicas (alzheimer, artritis, osteoporosis, etc.) o como reflejo de una pérdida general en las funciones fisiológicas, atribuible al proceso global de senescencia.

La necesidad de asistencia y/o cuidados derivada de las situaciones de dependencia no es una cuestión reciente. En todas las épocas ha habido una parte de la población que por razones de edad, enfermedad o deficiencia ha necesitado, de forma más o menos intensa, la atención de terceras personas en el desarrollo de su vida cotidiana. Lo que ha cambiado es su dimensión (debido, sobre todo, al crecimiento del número y de la proporción de personas mayores), su importancia social (ha dejado de verse como un problema exclusivamente individual o familiar para pasar a percibirse como un problema que afecta a la sociedad en su conjunto) y su naturaleza (en tanto que supone una redelimitación de los objetivos y funciones del Estado de Bienestar e implica nuevos compromisos de protección y financiación). Todo ello hace de la dependencia una cuestión tanto social como familiar y, al mismo tiempo, abre un nuevo campo de intervención que pone a prueba la capacidad de la sociedad y de sus instituciones para adaptarse a las nuevas realidades de fragilidad social (IMSERSO, 2005).

3. PROGRESIVO ENVEJECIMIENTO DE LA POBLACIÓN

El desarrollo social y la mejora en la calidad de vida que están teniendo lugar en nuestra sociedad están provocando un incremento en la longevidad y, por ende, la existencia de una gran masa de población envejecida.

Existe una fuerte vinculación entre el envejecimiento y la dependencia. Conforme aumenta la edad de los individuos, éstos presentan más limitaciones funcionales y, por lo tanto, son más

[1] Los CLD consisten en una amplia gama de servicios sanitarios ÿÿsociales, que generalmente son requeridoÿÿante el desarrollo de una discapacidad crónica, la cual puede deberse a un accidente, a una enfermedad o simplemente a una edad avanzada.

proclives a entrar en el colectivo de personas dependientes. Así que, aunque el fenómeno de la dependencia afecta a todos los grupos de edades, es en las personas mayores donde con más intensidad puede preverse que lo haga.

A pesar de que siempre haya habido personas en situación de dependencia (aunque no se haya considerado como tal), es cierto que el proceso acelerado de envejecimiento que actualmente está experimentando la población, así como las altas tasas de supervivencia (debidas, en gran parte, a los avances médicos) y las modificaciones en la estructura familiar están provocando una mayor conciencia social por la protección de los mayores que se encuentran en dichas condiciones.

No obstante, este proceso de envejecimiento no debe ser considerado únicamente como un problema, sino también como una oportunidad social que es necesario aprovechar. El envejecimiento de la población es la expresión de un logro humano: vivir más y vivir mejor. Constituye una de las transformaciones sociales más importantes producidas en el último tercio del pasado siglo. Pero es, a la vez, un reto al que se tiene que dar respuesta. Se trata de un reto complejo, pues el incremento del número de personas mayores y las necesidades crecientes de atención derivadas de ese crecimiento, coincide en el tiempo con una crisis de los sistemas de apoyo informal que han venido dando respuesta a esas necesidades, motivada por otras dos grandes transformaciones sociales de este final de siglo: el cambio en el modelo de familia y la creciente incorporación de las mujeres al mundo laboral (IMSERSO, 2005).

4. METODOLOGÍA

Con el fin de estimar el número de personas dependientes, se parte de las tasas de prevalencia de la dependencia (apartado 4.1) para calcular las probabilidades de transición entre estados (activo, dependiente -con tres grados de severidad- y fallecido).

Lo ideal sería que se pudiesen obtener las probabilidades de transición directamente de nuestros datos, habiendo realizado un seguimiento de un número de personas y observado si a lo largo de un periodo hubiesen cambiado de estado; se trataría, entonces, de información longitudinal. No obstante, los datos más recientes -los que nos aporta la EDAD (2008)- permiten conocer las tasas de prevalencia pero no cuándo se produce una transición entre estados.

Tal y como plantea Leung (2006), una posibilidad podría ser comparar las tasas de prevalencia en dos o más encuestas consecutivas (en nuestro caso, la EDDES de 1999 y la EDAD de 2008) y calcular las estimaciones de máxima verosimilitud de la probabilidad de que, trascurridos t años, una persona de edad x haya experimentado una transición del estado i al estado j, a través de la siguiente expresión:

$$ {}_t p_x^{ij} = \frac{n_{x,x+t}^{ij}}{\sum_k n_{x,x+t}^{ik}} $$

donde $n_{x,x+t}^{ij}$ es el número de personas en el estado i, de x años de edad en 1999 y en el estado j, de x+9 años en 2008.

El mayor inconveniente que puede aparecer al aplicar este método es que las encuestas que se quieran comparar se hayan diseñado de distinta manera. Aunque en nuestro caso, ambas presentan -en la medida de lo posible- una serie de datos común, existen ciertas diferencias conceptuales, ya que mientras la EDDES se basa en la Clasificación Internacional de Deficiencias, Discapacidades y Minusvalías (CIDDM), la EDAD se impregna de la filosofía de la Clasificación Internacional del Funcionamiento, de la Discapacidad y de la Salud (CIF), la cual difiere sustancialmente de la anterior.

Por este motivo, la estimación de máxima verosimilitud no resulta útil en la presente investigación, de manera que en el apartado 4.2 se plantea una alternativa para calcular las probabilidades de transición: el modelo de Markov.

4.1. Estimación de las tasas de prevalencia de la dependencia

Se necesita contar, por un lado, con el número de personas mayores de 65 años en situación de dependencia (aquellas que presentan alguna discapacidad para realizar las ABVD) y, por otro, la población total para cada una de las edades más avanzadas. El primer valor se obtiene de la EDAD (2008), mientras que el segundo se extrae de las estimaciones del INE de la población actual por Comunidades Autónomas, calculadas a partir del censo de 2001.

En cuanto a la primera cifra requerida, la EDAD (2008) aporta el número de personas con alguna discapacidad o limitación en general, así como el de aquellas a las que, en particular, dichas discapacidades les impiden realizar ABVD. Dado que en este trabajo se pretende estimar el coste de los CLD, se opta por tomar como eje el número de personas con alguna discapacidad para las ABVD.

Uno de los inconvenientes de la EDAD (2008) es que, de momento, sólo se ha publicado un avance de resultados. Por ello, el no contar aún con los microdatos hace que se deba trabajar con valores no especificados al detalle. Así, en varias de las tablas que presenta el INE las cifras no se encuentran desagregadas año por año, sino que lo están en rangos de 4, 5, 7, 9 o 10 años en unas y de hasta más de 30 años en otras. Esta cuestión hace que no sea posible realizar cálculos directos, obligando a plantear métodos más ingeniosos para obtener las tasas de prevalencia.

En el caso que nos ocupa, en la mencionada encuesta aparece el número de personas de 65 y más años con alguna discapacidad para las ABVD según el máximo grado de severidad por sexo y por los siguientes grupos de edad: de 65 a 79 y de 80 y más años. También se dispone del número de personas mayores con alguna discapacidad por sexo y por los grupos de edad que se detallan a continuación: de 65 a 69, de 70 a 74, de 75 a 79, de 80 a 84, de 85 a 89 y de 90 y más años.

Inicialmente, en la tabla que presenta mayor especificación de datos se considera la ponderación de cada uno de los rangos de edades más concretos respecto a los coincidentes con la tabla de discapacidades para las ABVD. Después, se pondera del mismo modo el número de personas de 65 y más años con alguna discapacidad para las ABVD según el grado de severidad y, por último, se distribuye equitativamente la ponderación entre los distintos años que componen cada uno de los grupos de edades.

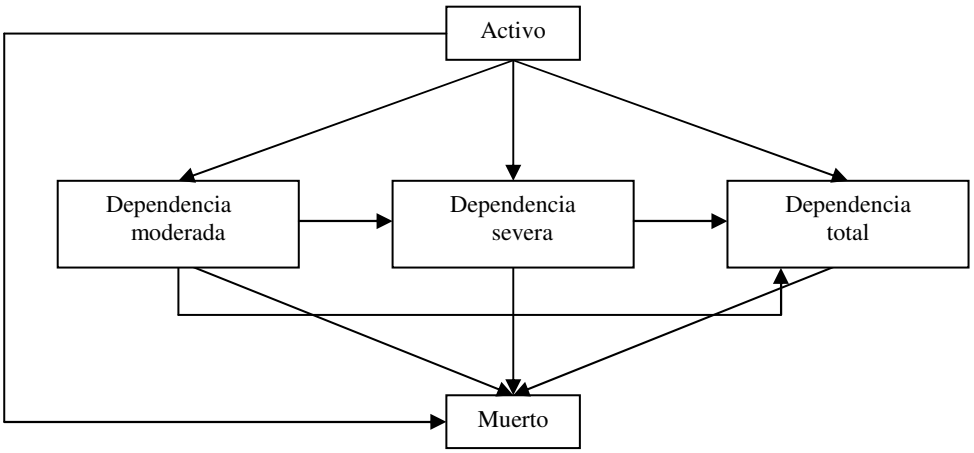
4.2. Modelo Markoviano de transiciones discreto en el tiempo

Son abundantes los trabajos, nacionales e internacionales, en los que se analiza el modelo markoviano de transiciones en torno al campo de la dependencia (Alegre, Pociello, Pons, Sarrasí y Varea, 2004; Artís, Ayuso, Guillén y Monteverde, 2007; Pociello, 2000). Especial mención merece el realizado por Haberman y Pitacco (1999), el cual sienta las bases a aplicar en numerosos estudios a este respecto, incluidos algunos de los citados anteriormente.

Alegre, Pociello, Pons, Sarrasí y Varea (2004) determinan las transiciones anuales entre estados de dependencia, asumiendo la hipótesis de estacionariedad propuesta por Pitacco (1995); es decir, se considera la dependencia como un estado irreversible en el que la persona ni puede volver a ser autónoma ni pasar a un nivel de menor dependencia. Por la idoneidad de dicha propuesta, para la presentación de este apartado se toma como referencia la misma.

Las transiciones entre estados que pueden darse son las que se reflejan a continuación:

Figura 1: *Transiciones en el modelo de estados múltiples*



A partir de estas posibles transiciones se puede definir el conjunto:

$$\wp = \{a \to d_1, a \to d_2, a \to d_3, a \to m, d_1 \to d_2, d_1 \to d_3, d_1 \to m, d_2 \to d_3, d_2 \to m, d_3 \to m\}$$

donde la correspondencia entre los distintos términos y estados es la siguiente:

$a \to$ Activo.

$d_1 \to$ Dependencia moderada.

$d_2 \to$ Dependencia severa.

$d_3 \to$ Dependencia total.

$m \to$ Muerto o fallecido (estado absorbente).

Así pues, las probabilidades de transición de un estado a otro entre las edades x y $x+1$ se pueden recoger en la siguiente matriz:

Tabla 1: *Matriz de transición* M_x

	a	d_1	d_2	d_3	m
a	p_x^{aa}	$p_x^{ad_1}$	$p_x^{ad_2}$	$p_x^{ad_3}$	p_x^{am}
d_1	0	$p_x^{d_1d_1}$	$p_x^{d_1d_2}$	$p_x^{d_1d_3}$	$p_x^{d_1m}$
d_2	0	0	$p_x^{d_2d_2}$	$p_x^{d_2d_3}$	$p_x^{d_2m}$
d_3	0	0	0	$p_x^{d_3d_3}$	$p_x^{d_3m}$
m	0	0	0	0	1

donde p_x^{ij} es la probabilidad de que una persona de edad x pase del estado i al estado j en el transcurso de un año, siendo $i, j =$ activo, dependencia moderada, dependencia severa, dependencia total o muerto/fallecido.

Siguiendo la evolución de estado de un individuo de edad inicial x a través de un proceso discreto en el tiempo $\{S(x),\ x \in N\}$, sus probabilidades satisfacen la propiedad de Markov:

$$\text{Prob}\left\{S(x+1)=s_{x+1}/S(x)=s_x \wedge S(x-1)=s_{x-1} \wedge ... \wedge S(0)=s_0\right\}=$$
$$=\text{Prob}\left\{S(x+1)=s_{x+1}/S(x)=s_x\right\}$$

Es decir, que la probabilidad de que el individuo alcance la edad *x*+1 en el estado s_{x+1}, depende únicamente del estado que tenía en la edad *x*.

Con el fin de cuantificar las probabilidades de transición, Alegre, Pociello, Pons, Sarrasí y Varea (2004) utilizan el método actuarial propuesto por Pittacco (1995), que permite obtenerlas a partir de la ley de supervivencia de la población general, l_x, y de las tasas de prevalencia, λ_x, para los distintos estados de dependencia. Para ello, se supone conocida la probabilidad de fallecimiento de la población general, q_x, así como dichas tasas de prevalencia:

$\lambda_x^{d_1} \rightarrow$ Tasa de prevalencia de una persona con dependencia moderada de edad *x*

$\lambda_x^{d_2} \rightarrow$ Tasa de prevalencia de una persona con dependencia severa de edad *x*

$\lambda_x^{d_3} \rightarrow$ Tasa de prevalencia de una persona con dependencia total de edad *x*

A partir de éstas, establecen que:

$$l_x^{d_1}=\lambda_x^{d_1}\cdot l_x$$
$$l_x^{d_2}=\lambda_x^{d_2}\cdot l_x$$
$$l_x^{d_3}=\lambda_x^{d_3}\cdot l_x$$
$$l_x^{a}=\left(1-\lambda_x^{d_1}-\lambda_x^{d_2}-\lambda_x^{d_3}\right)\cdot l_x$$
$$l_x=l_x^{a}+l_x^{d_1}+l_x^{d_2}+l_x^{d_3}$$

donde,

$l_x^a \rightarrow$ Número de supervivientes activos de edad *x*

$l_x^{d_1} \rightarrow$ Número de supervivientes con dependencia moderada de edad *x*

$l_x^{d_2} \rightarrow$ Número de supervivientes con dependencia severa de edad *x*

$l_x^{d_3} \rightarrow$ Número de supervivientes con dependencia total de edad *x*

El número de supervivientes de edad $x+1$ a partir de la edad del periodo anterior, para los tres grados de dependencia, puede obtenerse a través de las siguientes ecuaciones:

$$l_{x+1}^{d_1}=l_x^{d_1}+l_x^{a}\cdot p_x^{ad_1}-l_x^{d_1}\cdot p_x^{d_1d_2}-l_x^{d_1}\cdot p_x^{d_1d_3}-l_x^{d_1}\cdot p_x^{d_1m}$$
$$l_{x+1}^{d_2}=l_x^{d_2}+l_x^{a}\cdot p_x^{ad_2}+l_x^{d_1}\cdot p_x^{d_1d_2}-l_x^{d_2}\cdot p_x^{d_2d_3}-l_x^{d_2}\cdot p_x^{d_2m}$$
$$l_{x+1}^{d_3}=l_x^{d_3}+l_x^{a}\cdot p_x^{ad_3}+l_x^{d_1}\cdot p_x^{d_1d_3}+l_x^{d_2}\cdot p_x^{d_2d_3}-l_x^{d_3}\cdot p_x^{d_3m}$$

Al estar ante un sistema de tres ecuaciones y nueve incógnitas se incorporan diversas hipótesis para así resolver la indeterminación. Estas hipótesis se establecen respecto a:

- Las probabilidades de fallecimiento para cada uno de los grados de dependencia, que se determinan recargando la probabilidad de fallecimiento de la población general, q_x:

$$p_x^{d_1 m} = \left(1 + \delta_x^{d_1 m}\right) \cdot q_x$$

$$p_x^{d_2 m} = \left(1 + \delta_x^{d_2 m}\right) \cdot q_x$$

$$p_x^{d_3 m} = \left(1 + \delta_x^{d_3 m}\right) \cdot q_x$$

siendo $\delta_x^{d_1 m}$, $\delta_x^{d_2 m}$ y $\delta_x^{d_3 m}$ los recargos por fallecimiento asociados a los distintos niveles de severidad.

Las probabilidades de transición entre los diferentes grados de dependencia, que se obtienen recargando las probabilidades de transición que tienen los activos de pasar a cualquiera de los tres grados de dependencia:

$$p_x^{d_1 d_2} = \left(1 + \delta_x^{d_1 d_2}\right) \cdot p_x^{a d_2}$$

$$p_x^{d_1 d_3} = \left(1 + \delta_x^{d_1 d_3}\right) \cdot p_x^{a d_3}$$

$$p_x^{d_2 d_3} = \left(1 + \delta_x^{d_2 d_3}\right) \cdot p_x^{a d_3}$$

donde $\delta_x^{d_1 d_2}$, $\delta_x^{d_1 d_3}$ y $\delta_x^{d_2 d_3}$ son los recargos asociados a las distintas transiciones entre los grados de dependencia y las probabilidades de transición que tiene una persona activa de pasar a cualquiera de los tres grados de dependencia se calculan sustituyendo las tres últimas expresiones en el sistema de ecuaciones.

5. EL COSTE DE LOS CUIDADOS DE LARGA DURACIÓN EN CEUTA

El objetivo principal de este trabajo es estimar el coste de los CLD requeridos por los ceutíes mayores de 65 años en el 2013, para lo cual es preciso conocer cuántos de ellos se encuentran en situación de dependencia.

A partir de las tasas de prevalencia y de la ley de supervivencia de la población general se puede obtener el número de supervivientes en cada uno de los estados para cada edad *x*, así como las probabilidades de transición, en base a distintas hipótesis. No obstante, al contar con los resultados (por Comunidades Autónomas) de las proyecciones de la población a corto plazo realizadas por el INE para 2013 y dado que se desea estimar el coste de la dependencia para dicho año, en esta investigación no presentamos las tasas anuales de transición resultantes. Así, a partir de las tasas de prevalencia de la dependencia y de la población mayor de 65 años proyectada por el INE, se estima el número de ceutíes de edad avanzada en situación de dependencia, determinando el coste total de los CLD a través del coste individual asignado a cada grado de severidad.

En cuanto a las tasas de prevalencia, se podrían establecer diferentes supuestos en función de que se mantuviesen constantes, se incrementasen o disminuyesen en los próximos años. A este respecto, en Monteverde (2004) se pone de manifiesto la tendencia decreciente registrada por las tasas de prevalencia de las personas mayores dependientes; sin embargo, esta tendencia resulta tan poco significativa que sería bastante realista llevar a cabo proyecciones considerando tasas de prevalencia constantes en el tiempo, supuesto que se tiene en cuenta en este estudio.

Por lo que respecta al coste de la dependencia, se toman como punto de partida los datos aportados en el Informe 2006 sobre las Personas Mayores en España (IMSERSO, 2007), en el que, teniendo como fuente a la Dirección Territorial del IMSERSO en Ceuta, se recogen los costes individuales de los servicios sociales que ofrece la ciudad (ayuda a domicilio, teleasistencia, centros de día y residencias). En este sentido, se pueden plantear multitud de alternativas de asignación de los cuidados a cada grado de severidad. En nuestro caso, según las características que presenta cada uno de los niveles, se ha optado por considerar para el

dependiente moderado ayuda a domicilio, para el severo una plaza en un centro de día y para el total en una residencia.

En la siguiente tabla se recoge el coste estimado de los CLD precisados por los ceutíes mayores de 65 años para 2013, suponiendo que para los próximos años la tasa interanual del IPC (base 2006) se sitúe en el 2% (hipótesis 1), en el 3% (hipótesis 2) o en el 4% (hipótesis 3).

Tabla 2: Costes totales de los CLD. Personas mayores en situación de dependencia en Ceuta, 2013

HIPÓTESIS 1	**Dependientes de 65 años y más en 2013**	**Coste individual (en euros)**	**Coste total (en euros)**
Dependencia moderada	220	3.171,63	697.198
Dependencia severa	440	8.989,88	3.952.372
Dependencia total	2.386	17.889,68	42.690.480
Total Mayores Dependientes	3.046		**47.340.051**
HIPÓTESIS 2	**Dependientes de 65 años y más en 2013**	**Coste individual (en euros)**	**Coste total (en euros)**
Dependencia moderada	220	3.297,85	724.944
Dependencia severa	440	9.347,64	4.109.661
Dependencia total	2.386	18.601,62	44.389.398
Total Mayores Dependientes	3.046		**49.224.003**
HIPÓTESIS 3	**Dependientes de 65 años y más en 2013**	**Coste individual (en euros)**	**Coste total (en euros)**
Dependencia moderada	220	3.427,80	753.510
Dependencia severa	440	9.715,98	4.271.599
Dependencia total	2.386	19.334,60	46.138.526
Total Mayores Dependientes	3.046		**51.163.635**

Fuente: Elaboración propia

Para terminar, cabe decir que las estimaciones realizadas se han hecho sin considerar la intervención privada ni los cuidados informales (aquellos que son dispensados fundamentalmente por familiares y allegados), lo cual reduciría el coste anterior. Así pues, en cualquiera de los tres casos el gasto en cuidados a los mayores dependientes de Ceuta supondría menos del 5% del PIB de la ciudad.

6. BIBLIOGRAFÍA

ALEGRE, A., POCIELLO, E., PONS, M. A., SARRASÍ, F. J. Y VAREA, J. (2004). Modelo discreto de transiciones entre estados de dependencia. *Anales del Instituto de Actuarios Españoles*, *10*, 91-113.

ARTÍS, M., AYUSO, M., GUILLÉN, M. Y MONTEVERDE, M. (2007). Una estimación actuarial del coste individual de la dependencia en la población de mayor edad en España. *Estadística Española*, *49* (165), 373-402.

HABERMAN, S. & PITACCO, E. (1999). *Actuarial Models for Disability Insurance*. London: Chapman & Hall/CRC.

IMSERSO (2005). *Atención a las personas en situación de dependencia en España. Libro Blanco*. Madrid: Ministerio de Trabajo y Asuntos Sociales, Secretaría de Estado de Servicios Sociales, Familias y Discapacidad, IMSERSO.

IMSERSO (2007). *Las Personas Mayores en España. Informe 2006*. Madrid: Ministerio de Trabajo y Asuntos Sociales, Secretaría de Estado de Servicios Sociales, Familias y Discapacidad, IMSERSO.

INE (1999). *Encuesta sobre Discapacidades, Deficiencias y Estado de Salud*. Madrid: INE, IMSERSO y Fundación ONCE.

INE (2008). *Encuesta sobre Discapacidad, Autonomía personal y situaciones de Dependencia*. Madrid: INE, Ministerio de Educación, Política Social y Deporte, Fundación ONCE, CERMI y FEAPS.

Leung, E. (2006). A Multiple State Model for Pricing and Reserving Private Long Term Care Insurance Contracts in Australia. *Australian Actuarial Journal*, *12* (2), 187-247.

Monteverde, M. (2004). *Discapacidades de las personas mayores en España: prevalencia, duraciones e impacto sobre los costes de cuidados de larga duración*. Tesis doctoral inédita. Barcelona: Universidad de Barcelona.

PITACCO, E. (1995). *Modelli attuariali per le assicurazioni sulla salute*. Milano: EGEA.

POCIELLO, E. (2000). *Modelización y cobertura de operaciones actuariales en colectivos con múltiples estados*. Tesis doctoral inédita. Barcelona: Universidad de Barcelona.

UNIDADE DE PROMOÇÃO E DESENVOLVIMENTO E OBSERVATÓRIO TRANSFROTEIRIÇO DO VALE DO MINHO

FERNANDO AUGUSTO SEGADÃES REBELO
E-mail: fernandorebelo@frebelo.mail.pt
FERNANDO MIRANDA TORRADO
E-mail: fernando.miranda@usc.es
Universidade de Santiago de Compostela

Resumo

Com este trabalho procuramos tornar evidente a necessidade de realizar uma observação dinâmica e permanente do espaço transfronteiriço, em especial nas áreas do emprego e das infra-estruturas. Assim, vamos desenvolver o conceito de Observatório num projecto mais concreto e sobretudo o meio para o desenvolver, que será uma Unidade de promoção e desenvolvimento transfronteiriço para o Vale do Minho.
O desenvolvimento deste projecto poderá ser feito como uma acção inovadora de cooperação no desenvolvimento dos programas operativos do FSE dos dois países. No caso português, também poderia servir como experiência piloto para analisar a viabilidade de um programa destas características.

Palavras-Chave: Observatório, desenvolvimento transfronteiriço, promoção.
Área Temática: Economía Urbana, regional y local.

Abstract

With this work we look for a way to evidence the need to realize a dynamic and lasting observation of the cross-border space, especially in the employment and infrastructures areas. So, we are going to develop the concept of an observatory in a more concrete project and, mainly, the way to develop it, which will be a promotion and development cross-border unit for the Minho´s Valley.
The development of this project could be done as an innovative action of cooperation in the development of the operative programs of the FSE of both countries. In the Portuguese case, it could also be used as a pilot experience to analyse the viability of a program with these characteristics.

Key Words: Observatory, cross-border development, promotion.
Thematic Area: Local, regional and urban economy.

1. QUADRO GERAL DO PROJECTO

1.1. Introdução

Através do diagnóstico realizado na primeira fase, ficou evidente a necessidade de realizar uma observação dinâmica e permanente do espaço transfronteiriço, em especial nas áreas do emprego e das infra-estruturas. Com o projecto que se desenvolve na presente ficha vamos desenvolver este conceito de Observatório num projecto mais concreto e, sobretudo, o meio para o desenvolver, que será uma Unidade de promoção e desenvolvimento transfronteiriça.

As unidades de promoção e desenvolvimento são instrumentos de planeamento e desenvolvimento local, já existentes na Galiza e relativamente aos quais não existe uma correspondência no Norte de Portugal. São financiados através do Fundo social Europeu e, na Galiza e em Espanha em geral, têm-se convertido num dos principais instrumentos de desenvolvimento local de uma perspectiva local de carácter supra municipal. A sua função

básica centra-se na promoção de programas de emprego e desenvolvimento local, em especial de escolas oficina e casas de ofícios, como programas de emprego – formação.

Ao longo da presente ficha de projecto vamos partir da ideia das Unidades de promoção e desenvolvimento da Galiza para desenvolver um projecto de características similares sobre o espaço transfronteiriço do Vale do Minho. Sem desenvolver ainda o projecto em toda a sua amplitude, queremos aproveitar esta breve introdução para destacar a importância deste projecto de duas ópticas diferentes:

- A unidade de promoção e desenvolvimento como equipa técnica encarregada de conceber e realizar a observação permanente do espaço transfronteiriço.
- A unidade de promoção e desenvolvimento como equipa técnica multidisciplinar de referência no espaço transfronteiriço, encarregada de conceber e planificar as acções de desenvolvimento local executadas pelo Vale do Minho, assim como apoio técnico às acções de desenvolvimento local dos municípios integrantes.

De destacar, por último nesta introdução que ainda que os projectos de Unidades de promoção e desenvolvimento sejam co-financiadas pelo FSE na Galiza, as características do projecto que se descrevem a seguir e o carácter transfronteiriço do projecto fazem com que seja também possível a sua inclusão em linhas de financiamento de cooperação transfronteiriça, similares à actual iniciativa comunitária Interreg IIIA. Além disso, o desenvolvimento do projecto poderá ser feito como uma acção inovadora de cooperação no desenvolvimento dos programas operativos do FSE da Galiza e do Norte de Portugal. No caso de Portugal, o projecto também poderia servir como experiência piloto para analisar a viabilidade de um programa destas características neste país.

1.2. Referências: linhas estratégicas nas quais se enquadra o projecto

Como elemento de referência para o desenvolvimento do projecto dever-se-á ter em conta:

1.2.1. Projecto de regulamento do FEDER

No seu art. 6, Cooperação territorial europeia, fala-se especialmente nas intervenções que tenham por objectivo "a realização de actividades económicas, sociais e ambientais transfronteiriças, através de estratégias comuns de desenvolvimento territorial sustentável". Neste sentido, é fundamental dispor de informação homogénea e conjunta sobre o espaço transfronteiriço, uma informação que seja capaz de desagregar os dados disponíveis ao nível NUT III e que permita articular verdadeiras estratégias de desenvolvimento local à escala municipal.

Tal como se indica no próprio projecto de regulamento, o FEDER também pode contribuir para promover a cooperação judicial e administrativa, a integração dos mercados de trabalho transfronteiriços, as iniciativas locais em matéria de emprego, a igualdade entre os sexos e a igualdade de oportunidades, a formação e a inclusão social, assim como a utilização partilhada dos recursos humanos e dos meios destinados à I+DT.

1.2.2. Projecto de regulamento do FSE

Relativamente ao possível financiamento do equipamento técnico da unidade de promoção e desenvolvimento através do FSE, temos de ter em conta que estes projectos já são financiados na actualidade pelo FSE e que o novo projecto de regulamento, no seu art. 3 estabelece que o "FSE também apoiará as actuações a nível transnacional e inter-regional, em particular, através da partilha de informação, experiências, resultados e boas práticas e do desenvolvimento de planeamentos complementares e acções coordenadas ou conjuntas". Actualmente, tudo parece indicar que as unidades de promoção e de desenvolvimento vão continuar a existir no próximo período de programação comunitária, pelo que para o desenho do presente projecto seguiremos as indicações estabelecidas na Ordem de 24 de Março de 2006 da Xunta de Galiza, para poder

ser mais fácil a colaboração através de programas do FSE, tal como se estabelece no art. 8 do projecto de regulamento do FSE relativo à cooperação transnacional e inter-regional.

1.2.3. Orientações do futuro Objectivo 3: programa operativo de cooperação transfronteiriça

Dentro da diversa documentação que está actualmente ao dispor dos estados-membros e do quadro do programa operativo de cooperação transfronteiriça para Espanha e Portugal, é necessário destacar que as mudanças não são especialmente significativas no que toca à programação da iniciativa comunitária Interreg IIIA 2000-2006. No que respeita aos eixos de actuação nos quais se pode enquadrar este projecto, poderá fazer sentido tanto nos programas de apoio à competitividade e promoção de emprego (antiga prioridade 3 do Interreg IIIA) como no fomento da integração económica e social (antiga prioridade 4 do Interreg IIIA).

Em modo de conclusão, podemos observar que o projecto, tanto na sua vertente de serviço transfronteiriço de desenvolvimento local (observatório transfronteiriço) como na sua vertente de emprego e equipa técnica multidisciplinar (Unidade de promoção e de desenvolvimento) pode e deve ser co-financiado pelos fundos comunitários, tanto na sua vertente FEDER como FSE, assim como através do futuro Objectivo 3 de cooperação transfronteiriça. Por este motivo procurar-se-á em futuras secções estabelecer um desenvolvimento modular do projecto, que permita apresentar subprojectos à medida, para que possam aparecer janelas de financiamento adequadas, sem perder a possibilidade de o apresentar como um único projecto, no caso de isso ser possível, no quadro de um único instrumento de financiamento (como poderia ser actualmente o Interreg III).

1.3. Referências Transversais: relação com outras linhas estratégicas do Vale do Minho

O presente projecto enquadra-se no terceiro objectivo estratégico do Vale do Minho: Potenciar as iniciativas empresariais, a formação e a criação de emprego. Não obstante, e tal como se mencionou anteriormente, o conceito de observatório transfronteiriço que está no núcleo do Vale do Minho como instituição, é um dos pilares básicos da sua actuação futura e aquilo que lhe permitirá posicionar-se como referência na cooperação transfronteiriça. Por isso, o presente projecto incide sobre todas e cada uma das linhas de actuação propostas para o Vale do Minho.

1.4. Quadro Normativo do Projecto

Respeitando o sistema de concessão de projectos de unidade de promoção e de desenvolvimento (Ordem de 26 de Março), os projectos UPD deverão desenvolver a sua memória de pedido, entre outros, os seguintes elementos:

- Identificação da entidade promotora;
- Conteúdo do projecto (finalidade, objectivos, metodologia, estratégias previsíveis de fomento de emprego no território, colectivos afectados e calendário de actuações);
- Data de inicio de projecto (duração máxima de 2 anos);
- Orçamento de gastos e rendimentos, incluindo os contributos da entidade promotora.
- Os aspectos a valorizar num projecto de unidade de promoção e desenvolvimento para a sua concessão são:
- Número de pessoas desempregadas no âmbito de actuação: 5 pontos;
- Carácter inovador do projecto: 10 pontos;
- Qualidade do projecto atendendo à sua finalidade, objectivos, estratégias previstas de fomento do emprego e inserção laboral: 15 pontos;
- Contributos da entidade promotora: 5 pontos;
- Resultados de anteriores projectos: 5 pontos;

- Participação de agentes do território quando esteja vinculado ou surja na raiz de um pacto territorial: 5 pontos.

Nas secções seguintes vamos desenvolver cada um dos aspectos comentados.

2. DADOS BÁSICOS DO PROJECTO

2.1. Denominação do projecto

Para identificar o projecto e permitir a união dos dois conceitos subjacentes ao mesmo, pensamos no acrónimo **UVM-t** (Unidade de observação e promoção do desenvolvimento local e o emprego transfronteiriço).

Através deste nome pretende-se evitar possíveis confusões com as UPD da Galiza e tentar reflectir as duas principais vertentes do projecto: a observação transfronteiriça e a promoção das politicas de desenvolvimento local.

2.2. Objectivos do projecto

Mais uma vez, temos de falar de uma dupla perspectiva. Por um lado, a necessidade de dispor de informação relevante sobre a região transfronteiriça para articular os projectos de desenvolvimento e a criação de um equipamento multidisciplinar que seja capaz de analisar a informação e valorizá-la através da articulação de estratégias e projectos concretos. Contudo, o objectivo último do projecto é contribuir para o desenvolvimento local do espaço transfronteiriço, com a preservação do ambiente e o aproveitamento de recursos endógenos, a utilização intensiva das TIC's e a luta contra o desemprego, com uma perspectiva de género, como eixos centrais de actuação.

2.2.1. Objectivos gerais

- Melhoria dos sistemas de informação sobre o espaço transfronteiriço do Minho, aplicados ao planeamento estratégico, que melhorem a tomada de decisões e possibilitem uma gestão integrada de equipamentos locais a nível transfronteiriço e o desenvolvimento sustentável deste espaço territorial;
- Fomentar o desenvolvimento local e a sustentabilidade das estratégias de desenvolvimento através da identificação e valorização dos recursos próprios do espaço transfronteiriço;
- Promover a activação de politicas de emprego, em especial, de acções de formação e o fomento de politicas activas de emprego;
- A captação de recursos externos para o desenvolvimento de projectos de emprego, em especial no âmbito de projectos de emprego – formação como as escolas oficinas e os centros de emprego.

2.2.2. Objectivos específicos

- Desenvolvimento de um sistema de indicadores e bases de dados sobre o espaço transfronteiriço, que permita efectuar comparações intermunicipais, acompanhar a sua evolução e analisar o impacto das políticas e actuações;
- Desenvolvimento de um sistema de informação para a análise e a gestão de equipamentos locais de ambos os lados da fronteira, que possibilite a utilização conjunta dos mesmos;
- Desenvolvimento de um sistema de indicadores e bases de dados cartográficas de carácter ambiental;
- Estabelecer um sistema de informação conjunto que facilite a circulação de informação orientada para a gestão territorial;

- O desenvolvimento de sistemas de informação ao cidadão e empresas com a informação gerada;
- Apoiar as politicas locais de emprego e as acções concretas realizadas pelas câmaras municipais do Vale do Minho;
- Fomentar o trabalho em rede dos técnicos/as do território que trabalham no âmbito do desenvolvimento local: Técnicos/as de emprego e agentes de desenvolvimento local;
- Fomentar e apoiar as políticas de criação de empresas e empreendedores que se desenvolvem no território;
- Assessorar os municípios do Vale do Minho na implementação de projectos de emprego e formação, tais como escolas oficina;
- Promover a utilização de metodologias de intervenção próprias, adaptadas às necessidades do território, contemplando a igualdade de oportunidades como eixo transversal das mesmas;
- Facilitar informação sobre políticas europeias e programas de actuação da Comissão Europeia aos agentes públicos e privados do território;
- A detecção de boas práticas em matéria de desenvolvimento local, sistematizando a informação recolha;
- Estimular a utilização das novas tecnologias da informação e comunicação nos projectos de desenvolvimento local promovidos pelas câmaras municipais do Vale do Minho.

2.3. Descrição do projecto

Em conformidade com a dupla vertente do projecto descrito, comentaremos os principais aspectos do projecto, na perspectiva da constituição de um observatório transfronteiriço, bem como na perspectiva da criação da equipa que se encarregará da realização dos trabalhos.

2.3.1. Estrutura do Observatório

Relativamente ao observatório, este será executado de acordo com os seguintes critérios gerais:

- Os trabalhos de observação serão concebidos de modo assegurar a comparação dos resultados com os dados existentes de trabalhos anteriores, permitindo assim a apresentação da evolução das diferentes variáveis observadas, coisa que proporciona um maior valor, do ponto de vista da observação. Trata-se pois de aproveitar ao máximo a informação disponível em fontes secundárias;
- Os âmbitos de informação e análise terão como referência a estruturação dos dados realizados no plano estratégico de Pontevedra 2010, para dispor já de um primeiro elemento de referência. A classificação funcional da observação será a seguinte:
- Localidade, território e habitação;
- Desenvolvimento social e qualidade de vida;
- Economia;
- Análise empresarial;
- Meio ambiente;
- Mobilidade;
- Inovação e sociedade do conhecimento;
- Turismo;
- Gestão pública.
- Será necessário estabelecer um processo de homologação dos dados disponíveis na Galiza e Portugal, já que não se utilizam sempre os mesmos critérios na construção de indicadores. Nesse sentido, trabalhar-se-á preferencialmente com as fontes estatísticas oficiais dos institutos de estatísticas nacionais e regionais;

- Pretende-se conseguir uma análise comparada com as diferentes regiões e áreas territoriais nacionais e europeias, pelo que, na medida do possível, se seguirão os critérios e indicadores de observação propostos pela eurostat;
- Relativamente à observação própria, realizada pela equipa técnica de unidade de promoção, procurar-se-á disponibilizar informação quantitativa e qualitativa;
- A informação quantitativa será articulada num inquérito anual de conjuntura económica e empresarial e num inquérito de satisfação ao cidadão. Estes dois inquéritos poderão realizar-se telefonicamente, sobre uma amostra representativa da Galiza e de Portugal. Em alguns casos, e quando for possível, poderão realizar-se pequenos inquéritos através de correio electrónico para colectivos específicos ou para a análise da administração local;
- A informação quantitativa centrar-se-á na realização de painéis de peritos e mesas de trabalho. Estas mesas e grupos deverão ter um carácter permanente, funcionando como uma verdadeira mesa de concertação do espaço transfronteiriço;
- Como resultado da observação própria, editar-se-á um relatório anual, em suporte de papel e em formato electrónico através do site do Vale do Minho. Existirá também um resumo executivo da informação que será enviado aos autarcas dos municípios integrantes do projecto;
- É importante destacar a necessidade de manter as variáveis de observação no tempo, já que num breve espaço de tempo a informação disponível poderá ter um carácter dinâmico, possibilitando o estudo da evolução histórica do território;
- Em colaboração com as universidades de Vigo e do Minho, propõe-se também a realização de um estudo prospectivo das principais variáveis sociais e económicas, para prever potenciais situações não desejadas ou tomar decisões correctoras com a devida antecedência.

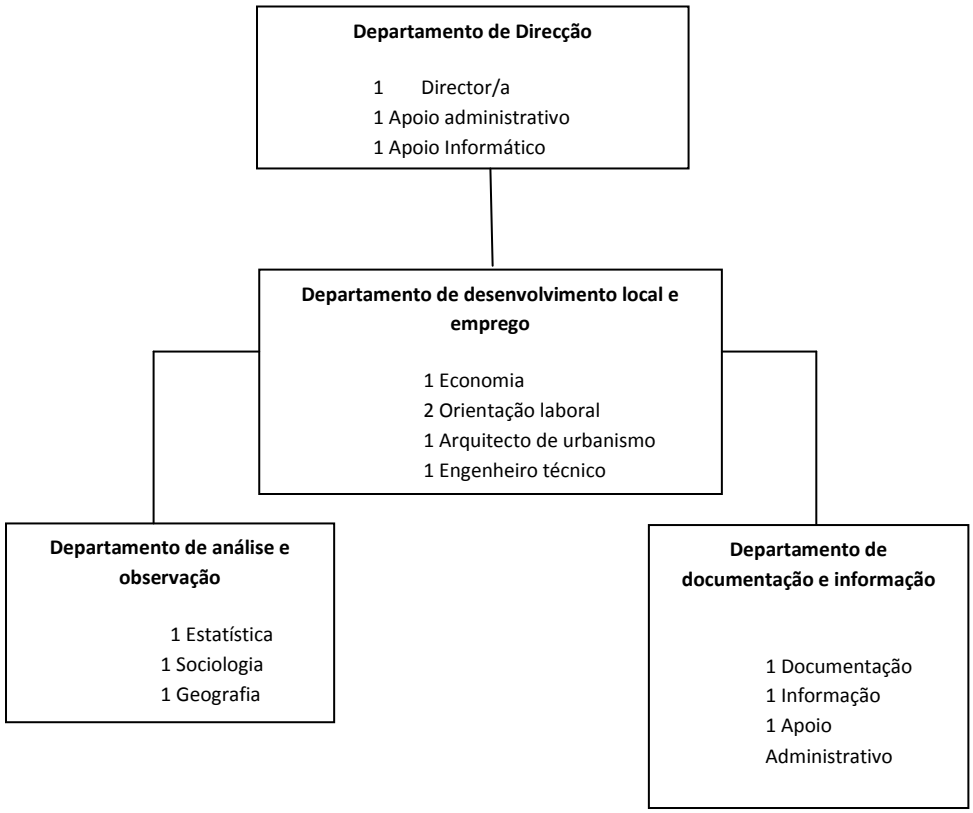
2.3.2. Estrutura da equipa da Unidade Técnica

Para facilitar a compreensão das actividades específicas da unidade de promoção e de desenvolvimento local, apresenta-se em primeiro lugar, um organigrama da equipa em função das áreas de trabalho:

Departamento de Direcção
- 1 Director/a
  - Perfil: Diploma superior e experiência acreditada na direcção de equipas (mínimo 5 anos).
- 1 Pessoa para dar apoio às tarefas administrativas.
  - Perfil: Formação profissional no ramo administrativo. Não é necessário ter experiência.
- 1 Pessoa para dar apoio nas tarefas informáticas.
  - Perfil: Formação profissional no ramo informático. Preferencialmente com experiência em tarefas de assistência.
- Departamento de análise e observação
- 1 Especialista em estatística
  - Perfil: Formação superior em matemáticas ou economia. Experiência profissional prévia.
- 1 Especialista em Sociologia.
  - Perfil: Diploma superior, preferencialmente formação superior em Sociologia ou Politicas. Experiência profissional prévia.
- 1 Especialista em Geografia e análise territorial.
  - Perfil: Formação superior em Geografia. É necessário ter experiência profissional prévia.
- Departamento de desenvolvimento local e emprego
- 1 Especialista em território

- Perfil: Preferencialmente arquitecto especialista em urbanismo. Não é necessário ter experiência prévia.
- 1 Especialista em desenvolvimento de projectos técnicos de desenvolvimento rural.
- Engenheiro agrónomo ou florestal, no mínimo ao nível da engenharia técnica. Preferencialmente com experiência profissional.
- 4 Titulados/as superiores
- Perfil: os perfis serão, em 50% duas pessoas ligadas à actividade económica e empresarial (licenciados/as em economia) e duas pessoas ligadas aos recursos humanos (Licenciados/as em psicologia, pedagogia ou assistentes sociais). Recomenda-se experiência profissional prévia em pelo menos, 50% dos postos de trabalho.
- Departamento de documentação e informação
- 1 Especialista em documentação
- Perfil: Diploma superior com experiência em arquivismo. É necessário ter experiência profissional prévia.
- 1 Informático/a
- Perfil: Diploma médio ou superior. Não é necessário ter experiência prévia, mas sim conhecimentos de desenvolvimento de bases de dados e enquadramento WEB.
- 1 Pessoa de apoio
- Perfil: Formação profissional de grau médio ou superior, preferencialmente no ramo administrativo.

Imagem nº. 1: Organigrama



Como podemos observar, a partir de uma Direcção do projecto, a unidade estrutura-se em 3 grandes departamentos, que correspondem a cada uma das áreas de intervenção no território:

- Departamento de Direcção: Encarregado pela Direcção estratégica do projecto, pela redacção de primeiro nível com outros projectos e departamentos do Vale do Minho, pela gestão dos RH, pela coordenação operativa dos 3 departamentos e seu planeamento, pela captação de financiamento, acções de negociação e defesa dos projectos, apresentações dos produtos resultantes e em geral, pelas acções de acompanhamento e controlo.

- Departamento de Análise e Observação: é o departamento responsável pela concepção do observatório, pela realização dos inquéritos (quantitativo) e pela coordenação de mesas ou painéis de peritos (qualitativas). Será responsável também pela utilização da informação, com a elaboração de relatórios e publicações, assim como pela exploração dos dados nos diferentes projectos executados pelo Vale do Minho.
- Departamento de desenvolvimento local e emprego: é o departamento mais numeroso, já que se encarregará do desenho e execução dos projectos, incluindo projectos de emprego, escolas oficina e casas de ofícios e pela dinamização da rede de agentes de emprego.
- Departamento de Documentação e Informação: este departamento está encarregado de pôr em marcha um centro de documentação sobre o espaço transfronteiriço, oferecer informação e assessoria aos municípios integrantes do Vale do Minho e, em especial, lançar na sua fase inicial um centro de informação Europeia, seguindo o modelo da Europa Direct (antigos Info Point e Carrefour).

É de comentar, por último que se procurará que as nacionalidades dos técnicos e técnicas participantes no projecto contem com representantes de ambos os países, numa relação de 50/50.

2.4. Activadades a desenvolver pela equipa técnica: projecto de unidade de promoção e de desenvolvimento

Antes de desenvolver as actividades específicas desta equipa técnica, é necessário salientar que, ainda que a maior parte das actividades seja efectuado directamente pelas pessoas contratadas, em alguns casos pontuais, será necessário contar com a assistência técnica ou colaborações externas, tal como se verá posteriormente.

Nesta secção, vamos descrever as principais actividades que a equipa técnica da unidade de promoção desenvolverá, assim como as estratégias previstas, metodologia aplicada, etc., respeitando as condições de subsídio (26 de Março 2006).

2.4.1. Metodologia de trabalho

A intervenção no território não é possível sem uma abordagem ascendente que possibilite a participação dos agentes locais na concepção das actividades, garantindo o seu envolvimento posterior. Para isso, e desde o primeiro momento, através do observatório transfronteiriço criar-se-ão mesas de trabalho, painéis e grupos de discussão que facilitarão a integração dos agentes socioeconómico e peritos do território.

Em casos especiais, deve ser mencionada a oportunidade de integrar as duas Universidades (Vigo e Minho) no desenvolvimento dos sistemas de informação e consulta. O critério de igualdade de oportunidades será aplicado transversalmente a todas as actividades, começando pelo próprio Unidade, que será constituída por, pelo menos 50% de mulheres.

2.2.2.Estratégias de intervenção e acções

O projecto de unidade de promoção articula-se em torno das seguintes estratégias, sobre as quais se coordenarão as diferentes actividades que desenvolve directamente a Unidade.

- Estratégia 1: Observação. Através desta estratégia, é desenvolvido o observatório permanente do espaço transfronteiriço, com uma análise de informação secundária e com informação primária através de inquéritos e mesas de trabalho com peritos e agentes locais. Merecerão especial atenção as acções de difusão e consulta de resultados.
- Estratégia 2: Aproveitamento de recursos locais. A partir dos resultados da observação, desenvolver-se-ão acções para:
  - Dinamizar o tecido social, e em especial a rede de técnicos/as de emprego de território, para o desenvolvimento de projectos de desenvolvimento local.
  - Criar um inventário de recursos, acessível também em base de dados.
  - Identificar boas práticas no território.
  - Desenvolver iniciativas de fomento ao auto-emprego e à cultura empreendedora.

- Desenvolver programas de sensibilização empresarial para a contratação de grupos com especiais dificuldades.
- Acções de ápio ao tecido empresarial local, incluindo o sector comercial e o artesanato.

- Estratégia 3: Apoio ao programa de escolas oficina e centros de emprego. Em consonância com o planeamento do financiamento deste projecto, esta unidade encarregar-se-á da promoção de projectos de escolas oficinas no âmbito de intervenção do Vale do Minho. Da mesma forma, serão realizados trabalhos de assessoria aos projectos em curso e serão elaborados materiais e ferramentas de intervenção para os técnicos/as do programa. No caso de Portugal, procurar-se-á sensibilizar as instituições para o desenvolvimento deste tipo de projectos.
- Estratégia 4: Olhar a Europa**.** Esta estratégia tem duas áreas de actuação. A primeira, já comentada, a criação de um centro de informação europeia (que possa ser a génese de um centro Europa Direct). A segunda, a captação de financiamento comunitário através da concepção de projectos e da apresentação das respectivas candidaturas.
- Estratégia 5: Difusão de resultados e dinamização. Através desta estratégia são coordenadas as actuações de difusão dos produtos e resultados que a unidade vai produzindo. De destacar a celebração de jornadas de carácter anual sobre o desenvolvimento transfronteiriço, com a presença de peritos de nível internacional, e a celebração de jornadas bimensais de curta duração (uma tarde) que tratem de aspectos monográficos do desenvolvimento sobre os quais a Unidade técnica esteja a trabalhar.
- Estratégia 6: Transversalidade. Além da inclusão de critérios transversais em todas as actuações que se venham a desenvolver, a unidade de promoção não pode ser um projecto independente dos demais que o Vale do Minho promova. Além disso, esta unidade técnica deve ser eixo central das acções de cooperação e desenvolver funções de coordenação e apoio aos projectos executados de ambos os lados da fronteira. Em especial, procurar-se-á que haja colaboração transversal nos âmbitos de:
  - Igualdade de oportunidades;
  - Ambiente;
  - Desenvolvimento das tecnologias da informação e comunicação;
  - Turismo sustentável.

2.5. Estruturação do projecto, fases e actividades para a execução do projecto

As principais fases pelas quais este projecto vai passar, pormenorizando os aspectos ou as actividades que necessitem de dotação económica de forma independente, isto é, que não possam ser desenvolvidos integralmente com os meios humanos e materiais da unidade, são descritas a seguir:

Fase 1: Lançamento do projecto

Actividade 1. Dotações e recursos materiais. Durante esta fase realizam-se os investimentos necessários para dispor de um espaço de trabalho para a unidade, desde o aluguer ou reabilitação de espaços, às dotações informáticas (incluindo servidores e sistemas de segurança, assim como estações de trabalho para sistemas de informação geográfica), passando pelo imobiliário.

Actividade 2. Selecção da equipa técnica. A selecção realizar-se-á através da combinação de processos de oferta genérica nos centros de emprego e através de anúncios na imprensa e outros meios de comunicação. Trata-se de captar os perfis mais adequados para um projecto com estas dimensões e, em muitos casos, será imprescindível contar com as pessoas com alguma experiência profissional. Na estrutura orçamental prevista, inclui-se o custo do pessoal e do equipamento técnico durante 2 anos de funcionamento.

Fase 2: Análise Territorial

Actividade 4. Criação de base de dados, sistemas de consulta. Nesta fase, adquirem-se e criam-se os sistemas de bases de dados e mapas, incluindo-se também os gastos com os inquéritos, os gastos de edição de um inventário de recursos e uma dotação económica para a assinatura de convénios com as duas universidades existentes no território para a organização e manutenção desta informação.

Fase 3: Trabalhos técnicos

Actividade 5. Desenvolvimento das actividades da unidade temática. Nesta fase, estão incluídos os gastos necessários para o funcionamento operativo da unidade, excluindo os gastos de pessoal, nomeadamente: gastos gerais, telefone, viagens e deslocações, etc.

Fase 4: Comunicação

Actividade 6. Desenvolvimento do centro de documentação transfronteiriço e do centro de informação europeia. Incluem-se os custos de aquisição de fundos bibliográficos, dipticos de informação específica e apoio de assistências técnicas externas.

Actividade 7. Publicações. Incluem-se os custos de preparação e edição dos produtos, inventário de recursos e a publicação de um boletim ou revista de carácter periódico.

Actividade 8. Dinamização de agentes locais. Organização de 2 encontros anuais de dimensão internacional sobre a cooperação e o desenvolvimento transfronteiriço, com a participação da Comissão Europeia. Esta actividade tem por objectivo tornar o Vale do Minho uma referência a nível europeu, em matéria de cooperação transfronteiriça. É complementada com a organização de pequenas jornadas e mesas temáticas de peritos, com uma periodicidade determinada: turismo, emprego, cultura, etc.

2.6. Orçamento económico por actividades

Fase 1: Lançamento de projecto

Actividade 1: Dotações e recursos materiais

	Orçamento
3. Outros investimentos materiais	150.000,00
8. Despesas gerais	48.000,00
TOTAL	€ 198.000,00

Actividade 2: Selecção da equipa técnica

	Orçamento
5. Despesas com o pessoal	924.000,00
TOTAL	€ 924.000,00

Actividade 3: Comunicação

	Orçamento
4. Prestações de serviços	30.000,00
7. Promoção e divulgação	32.000,00
TOTAL	€ 62.000,00

Fase 2: Análise territorial

Actividade 4: Criação de base de dados e sistemas de consulta

	Orçamento
3. Outros investimentos materiais	90.000,00
4. Prestações de serviços	60.000,00
9. Outros	100.000,00
TOTAL	€ 250.000,00

Fase 3: Trabalhos técnicos

Actividade 5: desenvolvimento de actividades de unidade técnica

	Orçamento
5. Despesas com o pessoal	48.000,00
8. Despesas gerais	48.000,00
9. Outros	30.000,00
TOTAL	€ 126.000,00

Fase 4: Comunicação e divulgação

Actividade 6: desenvolvimento do centro de documentação transfronteiriço e centro de informação europeia

	Orçamento
2. Aquisição de materiais	60.000,00
4. Prestações de serviços	12.000,00
7. Promoção e divulgação	6.000,00
TOTAL	€ 78.000,00

Actividade 7: Publicações

	Orçamento
7. Promoção e divulgação	60.000,00
TOTAL	€ 60.000,00

Actividade 8: Dinamização agentes locais

	Orçamento
5. Despesas com o pessoal	24.000,00
7. Promoção e divulgação	84.000,00
TOTAL	€ 108.000,00

Resumo global

	Orçamento
1. Obras públicas	0,00
2. Aquisição de materiais	60.000,00
3. Outros investimentos materiais	240.000,00
4. Prestações de serviços	102.000,00
5. Despesas com o pessoal	996.000,00
6. Formação	0,00
7. Promoção e divulgação	182.000,00
8. Despesas gerais	96.000,00
9. Outros	130.000,00
TOTAL	€ 1.806.000,00

3. INFORMAÇÃO COMPLEMENTAR

3.1. Localização do projecto

A área geográfica de intervenção será as câmaras municipais a sul da província de Pontevedra, que compreende as comarcas de A Paradanta, O Condado e Baixo Minho, assim como as seis câmaras municipais do Vale do Minho: Caminha, Melgaço, Monção, Paredes de Coura, Valença e Vila Nova de Cerveira.

Ocupa uma superfície total de 1.907.04 Km2, ou seja 3.8% da superfície da Galiza – Norte de Portugal. Nesta zona residiam em 2001, 185.819 habitantes, isto é cerca de 3% da população residente na Galiza – Norte de Portugal.

3.2.1 Apresentação da problemática que se pretende abordar

No que respeita ao mercado de trabalho, o Vale do Minho apresenta a seguinte problemática, sobre a qual deverá actuar a unidade de promoção:

- Fortes diferenças dos indicadores de emprego, que reflectem diferentes problemáticas de ambos os lados da fronteira, com elevadas taxas de desemprego no lado Galego e um crescimento da mesma taxa do lado português;
- Persistem as diferenças entre Homens e Mulheres, de forma especialmente negativa na região Galega;

- Escassa qualificação da mão-de-obra e problemas de reconhecimento e homologação de terminados;
- Escassa coordenação relativamente à gestão do mercado de trabalho;
- Fortes expectativas de desenvolvimento devido à PLISAN e à proximidade da área de Vigo;
- Sectores com amplo potencial de crescimento e geração de emprego: ambiente, turismo, florestal, etc., que aproveitam os recursos locais;
- Diferentes custos laborais, que dificultam a integração num único mercado de trabalho;
- As carências quanto à mão-de-obra qualificada podem retrair a implantação de iniciativas empresariais;
- População fortemente envelhecida, com um índice de renovação de activos muito inferior às médias da Galiza e de Portugal.

3.3. Situação actual do projecto

Actualmente não existe uma informação homogénea de carácter permanente sobre este espaço transfronteiriço, uma referência para o futuro observatório do Vale do Minho. Existem experiências de catalogação de recursos que podem ser aproveitadas dentro da Interreg IIIA (Projecto SIGN), com um esforço muito importante quanto á estruturação da informação através dês sistemas de informação geográfica. Não obstante, este projecto centra-se sobretudo na homologação e gestão cartográfica, sem dispor de uma vertente dinâmica do projecto que permita uma observação no tempo.

No âmbito do emprego há que destacar a existência do Eures transfronteiriço de Tui – Valença, como referencia na observação do mercado de trabalho. Também se tem desenvolvido projectos de formação ocupacional no quadro da Interreg III, os projectos FPtrans e Foriet, este último actualmente em execução. Ambos os projectos incidem principalmente no âmbito da gestão conjunta de acções de formação e no reconhecimento da formação, em ambos os lados da fronteira.

Relativamente às unidades de promoção e de desenvolvimento existentes na Galiza, a UPD da Associação de municípios de Vigo e a Unidade de Diputación de Pontevedra, actuam no âmbito territorial do projecto, pelo que seria conveniente limitar o âmbito territorial do projecto ao espaço delimitado do Vale do Minho.

3.4. Resultados esperados e objectivos quantificados

Como resultados finais do projecto, espera-se chegar a:

1 Observatório transfronteiriço sobre a realidade económica e social do Vale do Minho;

2 Sistema de informação geográfica;

3 Bases de dados sobre o Vale do Minho, agrupados sobre os seguintes âmbitos de intervenção:

1. Localidade, território e habitação;
2. Desenvolvimento social e qualidade de vida;
3. Economia;
4. Análise empresarial;
5. Meio ambiente;
6. Mobilidade;
7. Inovação e sociedade de conhecimento;
8. Turismo;
9. Gestão pública;

- 2 Jornadas internacionais;
- 10 Jornadas técnicas (bimensais);
- 9 Mesas de trabalho sobre desenvolvimento local (peritos) constituídas;
- 1 Inventário de recursos do espaço transfronteiriço (publicação);

- 1 Base de dados de boas práticas;
- 1 Boletim de informação/Revista;
- 2 Inquéritos a cidadãos e a empresas;
- 1 Site do projecto;
- 5 Projectos de escolas oficina realizados;
- 5 Projectos de desenvolvimento local elaborados;
- 100 Consultas efectuadas (desenvolvimento local e emprego);
- 300 Consultas dadas no centro de informação europeia;
- 10 Projectos empresariais assessorados;
- 100 Pessoas em acções de sensibilização empresarial/local.

3.5. Inovação e valor acrescentado do projecto

Ao longo da descrição do projecto aparecem 3 tipos de inovação de carácter relevante:

Inovação temática do produto: Embora existam referências prévias, não há, actualmente um observatório com estas características, que disponha de informação homogénea de ambos os países e que tenha escala local. Também não existe um projecto de trabalho multidisciplinar que actue sobre os dois países em simultâneo, e que integre pessoal qualificado de ambas as nacionalidades de carácter permanente, dentro de uma mesma equipa de trabalho. Esta experiência poderá servir além disso para o arranque de equipas de Promoção e desenvolvimento no Norte de Portugal.

Inovação do processo: A participação activa de todos os agentes socioeconómicos, com a aplicação do princípio de uma abordagem ascendente, é uma inovação neste tipo de processos técnicos. A articulação de mesas de peritos/as permanentes contribui para a estabilidade e para o compromisso do Vale do Minho no sentido de que esta participação seja permanente e desenvolvida ao longo de todo o processo.

O valor acrescentado deste projecto centra-se fundamentalmente na cooperação transfronteiriça. Tanto o observatório como a unidade técnica têm o seu eixo de actuação exclusivamente sobre o Vale do Minho como unidade de actuação. Não se pretende o apoio de projectos de desenvolvimento local que actuem apenas numa das margens da fronteira, mas antes com os projectos desenvolvidos que sejam verdadeiros projectos de cooperação, com impacto em Pontevedra e na área de Minho-Lima.

Como conclusão, há que destacar a capacidade de transferência deste projecto para outros territórios da União Europeia, e a sua inclusão como projecto de referência ou linha orientadora no âmbito do projecto objectivo 3 de cooperação. Durante a execução do projecto, a própria Unidade de desenvolvimento procurará a transferência de iniciativa com apresentações às administrações nacionais de Espanha e Portugal e também na União Europeia, através da Semana Europeia do Emprego.

3.6. Calendário de trabalho

A duração do projecto é de 24 meses. Atítulo meramente indicativo, estima-se que o projecto se inicie em Janeiro de 2010 e que termine a 31 de Dezembro de 2011.

Fase 1: Lançamento do projecto. Duração: 4 Meses
- Actividade 1. Dotações e recursos materiais: 1 Mês
- Actividade 2. Selecção da equipa técnica: 1 Mês
- Actividade 3. Comunicação às administrações públicas e agentes do território: 1 Mês

Fase 2: Análise territorial. Duração: 15 Meses
- Actividade 4. Criação de bases de dados, sistemas de consulta: 15 Meses.

Fase 3: Trabalhos técnicos. Duração: 20 Meses
- Actividade 5. Desenvolvimento das actividades da unidade técnica: 20 Meses.

Fase 4: Comunicação: 18 Meses
- Actividade 6. Desenvolvimento do centro der documentação transfronteiriço e do centro de informação europeia: 6 Meses
- Actividade 7. Publicações: 12 Meses
- Actividade 8. Dinamização de agentes locais: 18 Meses

	2010				2011			
	T1	T2	T3	T4	T1	T2	T3	T4
FASE 1: LANÇAMENTO DO PROJECTO	x							
Actividade 1. Dotações e recursos materiais	x							
Actividade 2. Selecção de equipa técnica	x							
Actividade 3. Comunicação	x							
FASE 2: ANÁLISE TERRITORIAL		x	x	x	x	x		
Actividade 4. Criação de bases de dados, sistemas de consulta		x	x	x	x	x		
FASE 3: TRABALHOS TÉCNICOS		x	x	x	x	x	X	x
Actividade 5. Desenvolvimento actividades da unidade técnica		x	x	x	x	x	X	x
FASE 4: COMUNICAÇÃO			x	x	x	x	X	x
Actividade 6. Centro de documentação			x	x				
Actividade 7. Publicações					x	x	X	x
Actividade 8. Dinamização de agentes locais			x	x	x	x	X	x

3.7. Sistema de acompanhamento e indicadores

O sistema de acompanhamento será reunido num manual de gestão interno e num quadro de comando específico, através de um programa de acompanhamento e avaliação do projecto, que será desenvolvido pelos técnicos dos mesmos, no qual será exposto de forma detalhada o processo sistemático de identificação e obtenção de informação válida acerca da concepção, planeamento, realização, resultados e impacto do projecto. O sistema de acompanhamento permitirá monitorizar a evolução dos indicadores de realização, o resultado e o impacto seleccionados de acordo com critérios de relevância, de quantificação, de fiabilidade e de disponibilidade. Estes indicadores irão abranger os seguintes aspectos:

1. De acompanhamento físico e financeiro;
2. De tempos de execução do Projecto;
3. Acompanhamento relativo e deficiências de implantação dos serviços (incidências).

Para recolher a evolução destes indicadores será concebida uma ferramenta informática que permita obter e tratar a informação necessária referente à gestão, ao acompanhamento e à avaliação do projecto com os seguintes objectivos:

- Garantir a correcta gestão dos fluxos financeiros e obter informação quantitativa e qualitativa sobre o conteúdo, desenvolvimento e resultados;
- Facilitar a tomada de decisões em relação ao desenvolvimento e execução das acções;
- Contribuir para a adequada canalização da informação para as bases de dados nacionais;
- Facilitar a visibilidade, difusão e transferência de resultados.

A sistematização na recolha e no tratamento da informação permitirá realizar num processo de avaliação contínua, que se materializará através da realização de estudos de avaliação parciais e de elaboração de relatórios globais por períodos semestrais de gestão.

Relativamente à realização de indicadores específicos, tornar-se-ão, como referência os resultados esperados do projecto e objectivos quantificados referidos anteriormente. A titulo de proposta inicial, e sem querermos ser exaustivos, propomos a utilização dos seguintes indicadores que indicamos.

3.7.1. Indicadores de realização

Bases de dados criadas;
Jornadas realizadas;
Documentos/relatórios realizados;
Projectos de escola oficina e de desenvolvimento local apresentados e/ou aprovados;
Inquéritos realizados/pessoas entrevistadas;
Apresentações públicas realizadas (transferência);

Consultas dadas;
Financiamento conseguido para projectos de desenvolvimento;
Mesas de peritos/as constituídas;
Projectos de auto-emprego assessorados.

3.7.2. Indicadores de resultados

Indicadores disponíveis sobre o Vale do Minho;
Assistentes das jornadas;
Distribuição de revistas e boletins (exemplares);
Reuniões de mesas de peritos/as (periodicidade);
Entidades colaboradoras nos projectos de desenvolvimento.

3.7.3. Indicadores de impacto

Projectos de cooperação transfronteiriça implementados sobre o espaço territorial do Vale do Minho;
Redução da taxa de desemprego (2 anos);
Pessoas beneficiárias directas dos projectos de emprego concebidos;
Aumento do investimento público por habitante (global ou projectos de cooperação transfronteiriça).

3.8. Identificação de fontes de financiamento

Tal como se referiu na introdução, o projecto é susceptível de receber financiamento directo através de projectos de cooperação no quadro do FSE. Não obstante, parece razoável candidatar também o projecto ao programa operativo de objectivo 3. Relativamente a este ultimo e aos eixos de actuação em que se pode enquadrar este projecto, podia fazer sentido tanto nos programas de apoio da competitividade e promoção de emprego como do fomento da integração económica e social. A inclusão nesta janela de financiamento permitir-nos-ia dispor de financiamento tanto para a gestão RH (FSE) como para os investimentos e equipamentos necessários para o correcto funcionamento da equipa (FEDER).

Previsivelmente, através do objectivo 3 (antigos Interreg IIIB e C) também poderiam existir mecanismos de financiamento, se bem que com uma maior componente transnacional.

3.9. Identificação de possíveis parceiros e agentes envolvidos

O projecto deve ser promovido directamente a partir do Vale do Minho, como entidade marca da cooperação transfronteiriça e integrado na estrutura desta instituição. Não obstante, e para candidatar o projecto adequadamente, contará inicialmente com o apoio da Deputación de Pontevedra e da Comunidade Intermunicipal do Vale do Minho.

Para conseguir o financiamento adequado, o projecto deveria envolver desde o primeiro momento as entidades regionais com competências na matéria, CCDRN e Xunta de Galiza através da Secretaria de Trabalho.

Por outro lado, através das mesas de peritos/as e das jornadas de trabalho, o projecto deverá contar, entre outros, com o envolvimento de:

- Todos os municípios do Vale do Minho;
- Associações de empresários/as e câmaras de comercio;
- Associações de turismo rural, hotéis e restaurantes;
- Associações do carácter cultural;
- ONG e Sociedade civil, em especial nas áreas do ambiente e da assistência social;
- Associações de mulheres;
- Responsáveis técnicos dos principais equipamentos;
- Universidades;
- Comunidades de montes e movimentos comunitários.

4. VALORIZAÇÃO E CONTEXTO

4.1. Impacto transfronteiriço das acções do projecto

A cooperação transfronteiriça não só está presente nas diferentes actividades propostas no projecto, como está na base do projecto, na sua própria concepção estratégica e, portanto é a razão de ser do mesmo.

Com efeito, o projecto de criação do observatório e da unidade de promoção e desenvolvimento ocuparão um espaço preferencial na gestão publica conjunta entre Galiza e o Norte de Portugal com:

- Um sistema de informação único e conjunto do espaço transfronteiriço no âmbito local;
- Com uma equipa interdisciplinar à disposição, constituída por pessoas de ambas as nacionalidades;
- Nasce com a vocação de possibilitar a gestão conjunta de equipamentos e serviços públicos transfronteiriços de âmbito local;
- Conceberá projectos de desenvolvimento local que incidam simultaneamente em ambas as fronteiras, e não só Galiza ou em Portugal;
- Apoia a criação de empresas com vocação para operar na Galiza e Norte de Portugal;
- Definirá um conjunto de indicadores e informação de carácter ambiental que possibilite o desenvolvimento sustentável do espaço transfronteiriço.

O projecto surge como resposta a uma problemática comum, identificada claramente em projectos de cooperação anteriores e é fruto da experiencia e reflexão conjunta e continua das entidades publicas participantes. É também uma resposta às necessidades de uma população transfronteiriça que vive cada vez mais de forma conjunta e procura respostas concretas. Que impacto terá a futura plataforma logística de Salvaterra – As Neves nos municípios transfronteiriços? Como é que os cidadãos do outro lado da fronteira podem aceder a equipamentos públicos locais? Como é que se podem racionalizar os equipamentos sanitários de um ponto vista transfronteiriço? Que parte da população pode aceder a um determinado equipamento e como é que isso afecta a sua sustentabilidade? As respostas a estas perguntas requerem sistemas de informação de grande capacidade, que permitam conjugar um elevado número de variáveis dentro de um enquadramento local que muda cada vez mais rapidamente.

4.2 Coerência com as políticas comunitárias nacionais e regionais

A nível europeu, tornou-se como referencia e procurou-se a coerência do projecto com prioridades do novo quadro comunitário para o período 2007-2013. Definido pelo União Europeia através de objectivos de Desenvolvimento sustentável de Gotemburgo e da Estratégia de Lisboa renovada, que se centra no crescimento e na criação de emprego através de objectivos específicos para conseguir este fim. Estes são mais concretamente:

- Fazer da Europa um lugar mais atractivo para investir e trabalhar;
- O conhecimento e a criação;
- Criar mais empregos e de melhor qualidade.

O projecto proposto também se enquadra nas seguintes directrizes e referencias das políticas europeias e nacionais:

- ETE, Estratégia territorial europeia. Para um desenvolvimento equilibrado e sustentável do território da EU;
- NORTE 2015: Uma região. Um futuro. Uma estratégia;
- As directrizes estratégicas de desenvolvimento rural para o período 2007-2013, publicadas a 25 de Fevereiro de 2006;
- O programa operativo de cooperação transfronteiriça Espanha-Portugal para o período 2007-2013;
- Quadro estratégico nacional de referência para o período 2007-2013;
- Plano estratégico nacional de desenvolvimento rural 2007-2013 (Espanha);
- Quadro de referência estratégico nacional (QREN 2007-2013) em Portugal;
- Plano nacional de desenvolvimento rural 2007-2013 para Portugal.

BIBLIOGRAFIA

CAPELA, JOSÉ VIRIATO (1987), "Entre-Douro e Minho 1750-1830. Finanças, Administração e Bloqueamentos estruturais no Portugal Moderno", 3 vol., Doutorado, Universidade do Minho.

CASTILLO, ANA MARIA (1988), "Apróximacion metodológico al contenido de la responsabilida social de la empresa", *Cuadernos de Ciencias Económicas y Empresariales*, Junio.

CASTRO, J. B. (1745), "Mappa de Portugal pelo Padre João de Castro", Officina de Miguel Menescal da Costa, Impressor do Santo Officio, 5 Vols., Lisboa.

FEIJÓ, RUI GRAÇA (1985), "Panorâmica sobre as gentes e os Campos do ano 1870. Notas de uma pesquisa de História regional em curso", *Les Campagnes Portugaises de 1870 a 1930: Image et Realité*, Foundation Callouste Gulbenkian / Centre Culturel Portugais, Paris.

FERRO, JOÃO PEDRO (1995), "A população portuguesa na final do Antigo Regime (1750-1815)", Editorial Presença, Lisboa.

GALEGO, JÚLIA, E SUZANE DAVEAU (1986), "O Numeramento de 1527-1532. Tratamento Cartográfico", Centro de Estudos Geográficos, Lisboa.

GARCIA ERQUIAGA, E. (1990), "Las características del empresário en la creación de empresas" Tesis doctoral. Universidad de Navarra.

GARCIA ERQUIAGA, E. (1996), "La incidência del entorno en la competitividad de la PYME: Pequeñas empresas de alto rendimiento*", Revista Alta Dirección*, nº. 185, Enero-febrero.

GODINHO, VITORINO MAGALHÃES (1977), "Estrutura da Antiga Sociadade Portuguesa", Arcádia, Lisboa.

LANDABASO, M. Y DÍEZ, M. A., "La gestión social de la empresa*", Economia e Empresa*, nº. 3.

LIVI BACCI, MASSIMO (1971), "A Century of Portuguese Fertility", Princeton University Press.

MAGALHÃES, JOAQUIM ROMERO (1988), "O Algarve Económico (1600-1773)", Editorial Estampa, Lisboa.

OCDE (1997), "La compétitivité industriele: comparaison internationale de l'énvironnement des entreprises", OCDE, Paris.

OLIVEIRA, AURÉLIO DE (1974),"A Abadia de Tibães e o seu domínio". Doutorado, Faculdade de Letras, Universidade do Porto, Porto.

PORTER, M. (1998),"On competition", *Harvard Buisiness School Press,* Boston.

PORTER, M. (1998),"*Clusters and the new economics of competition", Harvard Business Review*.

PRECEDO LEDO, A. (1994),"Desenvolvemento territorial e planificación comarcal", Colección Monografias, nº. 2, Xunta de Galicia.

RIBEIRO, ORLANDO (1955),"Portugal", *Geografia de España y Portugal*, Org. M. Teran, Barcelona.

RIBEIRO, ORLANDO (1967),"Portugal o Mediterrâneo e o Atlântico", Livraria Sá da Costa Editora, Lisboa.

RIBEIRO, ORLANDO (1987),"Introdução ao estudo da geografia regional", Edições Sá da Costa, Lisboa.

RIBEIRO, ORLANDO (1991),"Opúsculos Geográficos. O mundo Rural", Fundação Calouste Gulbenkian, Lisboa.

RIBEIRO, ORLANDO (1995),"Opúsculos Geográficos. Estudos Regionais", Fundação Calouste Gulbenkian, Lisboa.

RODRIGUES, MANUEL AUGUSTO (1993),"As Estruturas populacionais", *História de Portugal*, Org. J. Mattoso, Editorial Estampa, Lisboa.

ROWLAND, ROBERT (1997),"The population of Portugal", *Histoire de la population européenne*, Org. J. Dupâquier e J. P. Bardet, Fayard, Paris.

SAEZ DE VITERI, D. Y CABANELAS OMIL, J. (1997),"Cooperar para competir com êxito", Pirâmide, Madrid.

SANTOS, GERALDA MARIA MARQUES FERREIRA DOS (1979),"A freguesia de São Martinho da Arrifana de Sousa 1700-1829", Publicações do Centro de Estudos Demográficos, Lisboa.

SILBERT, ALBERT (1978),"Le Portugal Méditerranéen a la fin de l'Ancien Régime", INIC, Lisboa.

STOREY, D. Y JOHNSON, G. (1987),"Job reation in small and médium sized entreprises: sumary report", Comisión de las Comunidades Europeas, Oficina de Publicaciones, Luxemburgo.

VAZQUEZ BARQUERO, A. (1996),"Desarrollo económico local: Cursos de verano", Cord. Ourense: Copytemo Editores.

VAZQUEZ BARQUERO, A. (1996),"Infraestruturas, Território y Desarrollo local", *ICE*, nº. 757, Outubre.

LA INTEGRACIÓN DE LAS REGIONES PERIFÉRICAS EN EL ESPACIO COMUNITARIO: EL CASO DE CEUTA

JOSÉ AURELIANO MARTÍN SEGURA
jmartin1@ugr.es
Mª JOSÉ AZNAR UNZURRUNZAGA
mjaznar@ugr.es
CECILIO GÓMEZ CABRERA
GABRIEL GARCÍA-PARADA ARIAS
MANUEL HERNÁNDEZ PEINADO
Mª CARMEN MORÓN PÉREZ
ALEJANDRO RAMÍREZ HURTADO
Facultad de Educación y Humanidades de Ceuta
Universidad de Granada

Resumen

Partiendo del carácter de Ceuta de no ser parte integrante del territorio aduanero común y de poseer un régimen económico-fiscal especial, en la presente comunicación se aborda el papel de los flujos comerciales ceutíes como una vía para acelerar su proceso de integración en el esquema comunitario.
Basándonos en un modelo que relaciona las importaciones con los precios relativos de los productos adquiridos y el PIB, se trata de estimar el impacto que la reducción gradual de la protección exterior marroquí (en virtud del acuerdo de asociación UE-Marruecos) va a ejercer sobre los tráficos comerciales locales, y por ende sobre la economía ceutí.
Los resultados del panel de datos del modelo econométrico elaborado con el citado fin, marcarán la reorientación de que debe ser objeto la política comercial exterior de Ceuta con vistas a no perder oportunidades de integración en el contexto europeo.

Palabras Clave: Ciudad Autónoma de Ceuta / Marruecos / Acuerdo de libre comercio / Unión Aduanera.
Área temática: Economía Urbana, Regional y Local.

Abstract

Ceuta outside the customs union of the European Community and has a special economic and fiscal system.
This paper examines the foreign trade of Ceuta, variable essential to accelerate its process of integration into the European Union. We rely on a model that relates the relative price of imports and Gross Domestic Product (GDP). We must calculate the impact of eliminating tariffs on Morocco (EU-Morocco) on the foreign trade of Ceuta and its economy.

Thematic Area: Local, regional and urban economy.

1. INTRODUCCIÓN

A través de los siglos, la ciudad de Ceuta ha sido sucesivamente objeto de dominación por parte de distintas civilizaciones: fenicia, cartaginesa, romana, visigoda, bizantina e islámica, llegando a formar parte de los reinos ibéricos desde el siglo XV. Primero de Portugal, y ya en el siglo XVII de España.

Emplazada en el noroeste de África, se encuentra ubicada en pleno estrecho de Gibraltar. Limitando al norte, este y sur con el mar Mediterráneo y al oeste con Marruecos, posee un escarpado territorio, densamente poblado, que apenas abarca 20 km2.

La economía ceutí se encuentra sumamente condicionada por su emplazamiento geográfico, que le imprime el carácter de área periférica de la UE, separada por una frontera marítima, el estrecho de Gibraltar, que en ocasiones se convierte en una auténtica barrera para el correcto desenvolvimiento de las comunicaciones. Tal es el caso de los fuertes temporales de viento que afectan a la ciudad, que la dejan prácticamente aislada vía marítima del resto del país, con los consiguientes contratiempos sobre el tráfico de personas y mercancías.

Al mismo tiempo, su situación geográfica también es fuente de potencialidades, al configurarse como el punto de conexión de Europa con el continente africano, pues no en vano la frontera de Ceuta con Marruecos es parte integrante de las fronteras exteriores de la Comunidad.

Esas peculiaridades derivadas de sus condicionantes geográficos han propiciado por otra parte el establecimiento de un especial estatus económico-fiscal para la ciudad, que se justifica en la extraterritorialidad aduanera comunitaria, lo que no es óbice para que Ceuta sea, como lo es, territorio integrado de pleno derecho de la UE y perceptora de sus ayudas.

Por otra parte, para su futuro socioeconómico resulta vital tomar en consideración el esquema de relaciones que está operando en su entorno, que no es otro que el configurado por el Proceso Euromediterráneo (o de Barcelona). Dicho proceso, al consistir en buena medida en la creación de una zona de libre comercio entre la propia Unión y el reino de Marruecos (entre otros estados del Magreb y del Mashreq, como se sabe), está llamado a afectar de forma indirecta a los flujos comerciales ceutíes. Esta aseveración es perfectamente admisible si se tiene en cuenta la distribución geográfica de las exportaciones, dirigidas al país vecino casi exclusivamente. A este dato hay que añadir también que se trata de flujos que circulan extraoficialmente (aunque con aquiescencia de las autoridades de uno y otro lado de la frontera), gracias a la inexistencia de una aduana comercial entre ambos.

En este trabajo planteamos qué orientaciones debe seguir la economía ceutí según sea el impacto del inminente libre comercio UE-Marruecos promovido por el Proceso Euromediterráneo. Y ello con miras a no entorpecer su proceso de convergencia real y por ende evitar la pérdida de posibilidades de una favorable integración de Ceuta en el espacio comunitario.

2. EL PROCESO EUROMEDITERRÁNEO: IMPLICACIÓN PARA CEUTA

Si bien el objetivo prioritario establecido en la Conferencia Euromediterránea de Barcelona en 1995 era garantizar la paz y la estabilidad en el área mediante el diálogo en materia de política y seguridad, la cooperación económica y financiera, y la cooperación social, cultural y en educación, el contenido del acuerdo denotaba un marcado carácter comercial. De manera que el citado objetivo general habría de lograrse mediante el establecimiento de una zona de libre comercio euromediterránea para el año 2010. Ello debía conseguirse mediante la suscripción de los respectivos acuerdos de asociación con Argelia, Egipto, Jordania, Israel, Líbano, Marruecos, la Autoridad Palestina, Siria, Túnez y Turquía, que incluían acuerdos de liberalización comercial.

En definitiva, el proyecto perseguía, y persigue, promover la libertad de circulación de mercancías entre la Comunidad y los países mediterráneos, suprimiendo progresivamente los derechos de aduana, así como las restricciones cuantitativas a las exportaciones e importaciones y las medidas de efecto equivalente entre las partes. Estas normas se aplican a la importación de

productos manufacturados y a los intercambios de productos agrícolas y de la pesca, si bien en estos dos últimos casos con ciertas limitaciones. No obstante, pueden adoptarse medidas de salvaguardia por motivos de interés público o con objeto de proteger los sectores económicos sensibles.

Como es lógico, Ceuta ha de tomar muy en serio este nuevo modelo de relaciones comerciales que se está fraguando entre las economías de su entorno. Y ello porque corre el riesgo de desmoronarse el flujo de comercio bidireccional pseudo-oficial del que se ha venido beneficiando ampliamente la economía local.

En efecto, la ciudad ha sabido aprovechar el diferencial de aranceles existente entre Marruecos y Ceuta: el considerable mayor grado de protección exterior que otorga el país magrebí a las importaciones de la UE, en comparación con los tipos notablemente más bajos del Impuesto sobre la Producción, los Servicios y la Importación (IPSI), hacen que numerosos marroquíes efectúen sus adquisiciones en la ciudad, aprovechando unos precios más competitivos.

Ahora bien, cabe plantearse la cuestión de qué va a suceder en el ámbito local toda vez que este esquema se vea alterado como consecuencia del proceso de liberalización comercial que está operando entre la UE y el reino alauí. En otras palabras, cuál va a ser el impacto de esa zona de libre cambio que está gestándose en derredor. Sin duda alguna, la ventaja que actualmente detenta la ciudad gracias al aprovechamiento de ese diferencial de protección, se verá erosionada al albur del desarme arancelario de Marruecos.

La cuestión es hasta qué punto va a quedar afectada la economía local como consecuencia del progresivo desmantelamiento de la protección marroquí. El planteamiento de esta pregunta es vital para ir allanando el terreno de la previsión, de manera que Ceuta no se aparte de las regiones más favorecidas de la Unión ni retroceda en materia de convergencia real, quizás por una debacle económica que afecte a la localidad a tenor de la transfiguración del mapa comercial.

3. EL IMPACTO ECONÓMICO DEL DESARME ARANCELARIO MARRUECOS-UE SOBRE LA ECONOMÍA DE CEUTA: PLANTEAMIENTO

3.1 El marco de referencia e hipótesis

Ceuta, como todas las zonas transfronterizas, es un lugar de constante trasiego de personas y mercancías, que dispone de un régimen fiscal especial, como es sabido, y ya ha sido mencionado, orientado a suplir sus carencias de infraestructuras y de materias primas y a fomentar su crecimiento económico. No es por tanto un régimen privilegiado, sino una necesidad.

El crecimiento de la economía ceutí entre 1955 y 1998 estuvo condicionado por las diferencias en la dotación de factores productivos. Así, mientras que su tasa de actividad creció por encima de la media nacional (1,34% frente al 0,89%), la dotación de capital en relación a la población creció en ese período el 2,70 %, inferior al 3,58% de promedio nacional. Asimismo, las tasas de desempleo, aunque crecieron a una tasa inferior al promedio nacional, sin embargo en 1998 estaban en el 26,2%, por encima del 18,8% nacional. En la actualidad, el diferencial con la tasa nacional supera el 55%.

Además de lo anterior, el modelo económico de Ceuta está desequilibrado y es excesivamente dependiente de variables externas que no puede controlar. Así, el valor añadido de su economía se genera en el sector público casi en un 50%, y en un 30% en el comercio, excesivamente dependiente de los vaivenes del vecino país. Su población es muy joven (22% de 0 a 15 años, frente al 15% nacional), y el porcentaje de alumnos que inicia estudios universitarios es muy bajo (15 alumnos por 1.000 habitantes frente a 36 a nivel nacional). Y las industrias que podrían crearse al amparo de las reglas de origen, dando cierto dinamismo e independencia a la situación, son prácticamente inexistentes.

No obstante, en los últimos años la renta *per cápita* en Ceuta ha experimentado una evolución muy positiva. Su ritmo de crecimiento en el período le ha permitido converger en PIB *per cápita* con la UE-25. Así, el índice de convergencia de la renta *per cápita* en paridad del poder adquisitivo ha pasado del 79 % de la media comunitaria en el año 2000 al 87,8% en 2004.

Por su parte, es un hecho indiscutible que Marruecos mantiene importantes diferenciales con Europa y con España respecto a sus niveles de renta y bienestar, como lo ilustra el Fondo Monetario Internacional (FMI): la renta por habitante en España supera más de trece veces la marroquí en términos nominales (32.066 USD frente a 2.389), siendo casi cinco veces superior en términos de paridad de poder adquisitivo (27.914 dólares PPA frente a 5.765).

Sobre la base configurada por el escenario descrito, la hipótesis de la que se parte en nuestro estudio reviste los siguientes aspectos:

-La economía de Ceuta se encuentra inmersa en un proceso de cambio y adaptación a las nuevas circunstancias.

-La presión migratoria y de trabajadores transfronterizos, es creciente.

-El flujo no registrado de mercancías vendidas a Marruecos, alcanza cifras importantes.

-El desarme arancelario de Marruecos con la UE, junto con las circunstancias internas de la Ciudad, podrían generar un impacto negativo en la economía local, que es el que se trata de medir.

3.2. Metodología

Los estudios sobre comercio internacional tienen una larga tradición en economía desde Adam Smith y David Ricardo. En 1950 los trabajos de Maurice Byé, Herbert Giersch y Jacob Viner desarrollaron la problemática de las uniones aduaneras. La obra de Viner *The Customs Unions Issue (1950)* constituyó el punto teórico de partida de la investigación posterior.

Para analizar el impacto económico se debe partir de modelos económicos apropiados. Una vez localizados dichos modelos, habrá que definir las variables que puedan ser capaces de influir en los mismos. A partir de aquí se dará una especificación econométrica al modelo.

Al objeto de examinar en qué medida la reducción de barreras arancelarias puede afectar a los distintos sectores comerciales, serán útiles los modelos de equilibrio parcial, o general. De lo que se trata es de modelizar el comportamiento de las importaciones, conforme al enfoque clásico, que parte de que un país importa por las necesidades de dicho bien, o también al aumentar su actividad económica y su nivel de desarrollo.

Para estimar el *efecto* de la reducción de aranceles sobre el volumen de comercio se puede partir de una ecuación simple que relaciona las importaciones (variable dependiente) con los precios relativos de los productos importados entre países y con el PIB del país importador, o algún otro indicador de la actividad económica interna.

Lo anterior se puede especificar en una ecuación lineal sencilla del tipo:

$$Log(M_{ijk}) = \beta_{1ij} + \beta_{2ij} Log(PR_{jk}) + \beta_{3ij} Log(PIB_j) + \cdots + \varepsilon_{ij}$$

En donde *i* se refiere a los productos importados a los países *j k*

M= Importaciones

PR= Precios Relativos (aranceles y otros impuestos de importación)

PIB= Producto Interior Bruto del país importador

La ecuación econométrica lineal, en escala logarítmica, permite calcular las elasticidades precio de la demanda de importaciones, desagregando por sectores industriales. Se harían tantas estimaciones como sectores. O también se podría construir un modelo de panel de datos, dado que la serie temporal disponible no es muy amplia. De esta forma la estimación que se realizara obtendría unos resultados más consistentes. En nuestro caso, al depender de la reacción de Marruecos ante el desmantelamiento de sus aranceles, se puede incluir también otra variable con el PIB de dicho país.

Una vez conocidas las elasticidades demanda precio de importaciones, se puede predecir lo que ocurrirá ante una bajada de las mismas. Con los resultados se estará en condiciones de calcular el impacto de dichas cantidades (positivas o negativas) sobre los distintos aspectos de la

economía (gasto público, comercio, empleo...). Para ello se emplearían distintas estrategias de simulación, dependiendo del caso.

Ahora bien, el problema con el que se tropezaría en el análisis es que Ceuta no es un país. Los modelos consultados estudian el impacto del desarme arancelario entre países. Para ello analizan los impactos sobre la producción sectorial interna de cada producto. En Ceuta el sector mayoritario es el comercio. No hay producción sectorial interna. En estas circunstancias: ¿cómo podría construirse un modelo de análisis adaptado al caso?

3.3 Especificación del modelo para Ceuta

Los modelos consultados estudian el impacto del desarme arancelario entre países y después analizan los impactos sobre la producción sectorial interna de cada producto. Sin embargo, en Ceuta el sector mayoritario es el comercio. No hay casi producción sectorial interna. En ese caso se ha tenido en cuenta que parte de las importaciones desde la península se convierte en exportaciones a Marruecos. Por tanto, la disminución de aranceles en Marruecos en esos productos podría influir en un incremento de las exportaciones desde la península hacia ese país, lo que podría conllevar una disminución de las exportaciones desde Ceuta. Pero también podría dar lugar a un incremento del PIB interno en Marruecos y, por tanto, a un incremento del consumo de estos productos o a un incremento de la producción interior de los mismos, con lo cual se compensaría dicho impacto negativo sobre nuestra economía, ya que se continuaría comprando por ambos cauces.

Por lo tanto, la primera variable a utilizar es el total de importaciones que se realizan desde la península a Ceuta. De esta cifra, se estima (de la forma que se especifica más abajo) la cantidad que se exporta a Marruecos, aunque no quede registrada. Una vez localizados estos productos, se debe tener constancia del desarme arancelario que tendrán, según el acuerdo de libre comercio. A partir de aquí estaremos en condiciones de saber qué productos se exportan a Marruecos, sus precios relativos, su importancia en la economía local y, mediante los cálculos econométricos detallados anteriormente, intentar verificar el impacto sobre los distintos sectores económicos.

Partiendo de las cifras de importación facilitadas por los servicios aduaneros y tributarios de Ceuta, si se comparan las mismas con las cifras medias de consumo nacional de dichos productos, por diferencia se deduce, en una primera aproximación, la cantidad de productos que se dedica al comercio con Marruecos. Pero ha sido necesario afrontar un problema metodológico, puesto que la clasificación de importaciones en la aduana difiere de la clasificación de productos para el consumo que utiliza el INE de cara a realizar la encuesta de presupuestos familiares. En el primer caso se trabaja con la clasificación internacional de importaciones; en el segundo, con la *Classification of Individual Consumption by Purpose* (COICOP), que divide los productos de consumo en 12 grandes grupos, pudiéndose desagregar hasta los cuatro dígitos, e incluso hasta cinco en el caso de los dos primeros grupos. El problema se ha resuelto estableciendo una equivalencia entre productos atendiendo a criterios de coherencia.

Ahora bien, cabe formular las siguientes precisiones. Por una parte, algunos bienes de consumo (como el tabaco o las bebidas alcohólicas) podrían reexportarse también a la península. Otra partida importante son los hidrocarburos, que suponen casi el 21% del total de importaciones y que no se exportan al citado país, aunque podría ser un elemento competitivo importante en el nuevo puerto de Tánger-Med, por lo que debe mantenerse en los cálculos. Por otra parte, las cifras de importaciones se han comparado con el consumo medio nacional. Sin embargo, las pautas de consumo son distintas en las grandes ciudades y por tramos de renta. Por tanto, en una fase posterior habrá que aplicar factores correctores para tener en cuenta dichas circunstancias.

También se han analizado las series de aranceles de los productos seleccionados del vecino país. Tras una comparación de dichos aranceles, por productos y años, con el impuesto local de Ceuta (IPSI), se ha obtenido una variable que hemos denominado PR, y que representa los precios relativos entre Marruecos y Ceuta.

Por último, se han incluido las variables representativas del Valor añadido de Ceuta y del Producto Interior Bruto de Marruecos.

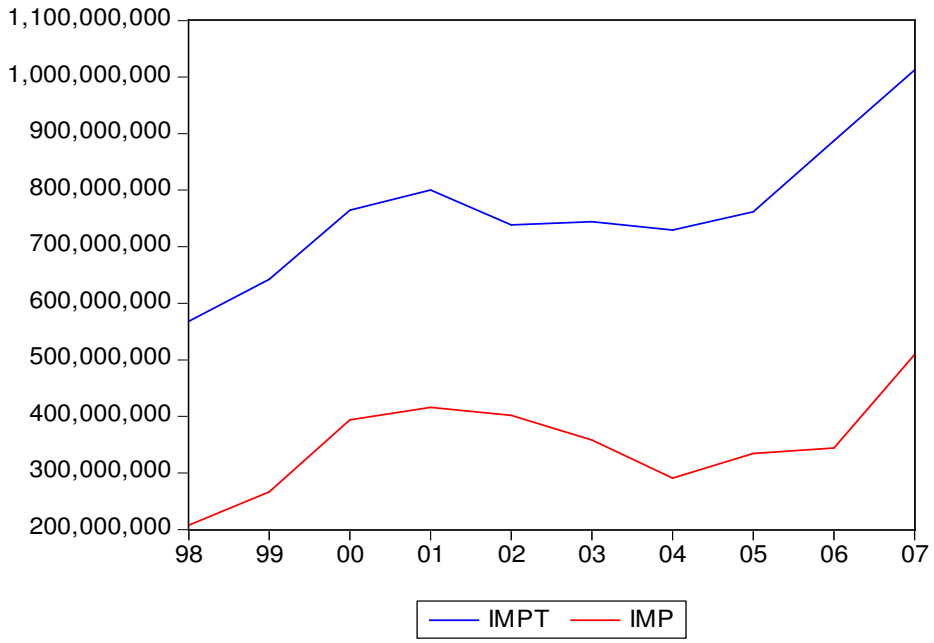
Con todas estas variables se ha formado un panel de datos con el que se han realizado diversas estimaciones que se muestran en el siguiente apartado.

4. RESULTADOS

4.1 Análisis descriptivo de los datos

Los primeros resultados obtenidos arrojan las siguientes cifras globales de importaciones a Ceuta (IMPT), y por diferencia, el saldo neto estimado de exportación a Marruecos (IMP):

AÑO	IMPT (en €)	IMP (en €)	%IMP/IMPT
1998	568,459,256.92	207,634,749.43	36.52
1999	642,078,003.01	266,611,070.75	41.52
2000	764,953,139.15	393,486,795.43	51.44
2001	800,547,221.73	416,103,192.83	51.97
2002	738,637,864.41	401,920,075.40	54.41
2003	744,672,669.73	358,401,273.52	48.13
2004	729,358,381.12	290,786,782.72	39.87
2005	761,647,502.64	334,611,175.99	43.93
2006	887,330,645.21	344,222,572.56	38.79
2007	1,012,426,467.07	509,475,030.18	50.32



La gráfica anual de ambas series nos incida que la evolución es similar, lo que sería un indicativo de que la estimación realizada por diferencias es bastante ajustada a la realidad, y que la demanda de productos desde Marruecos es correspondida con las importaciones que se realizan a Ceuta, aumentando o disminuyendo en razón de dicha demanda.

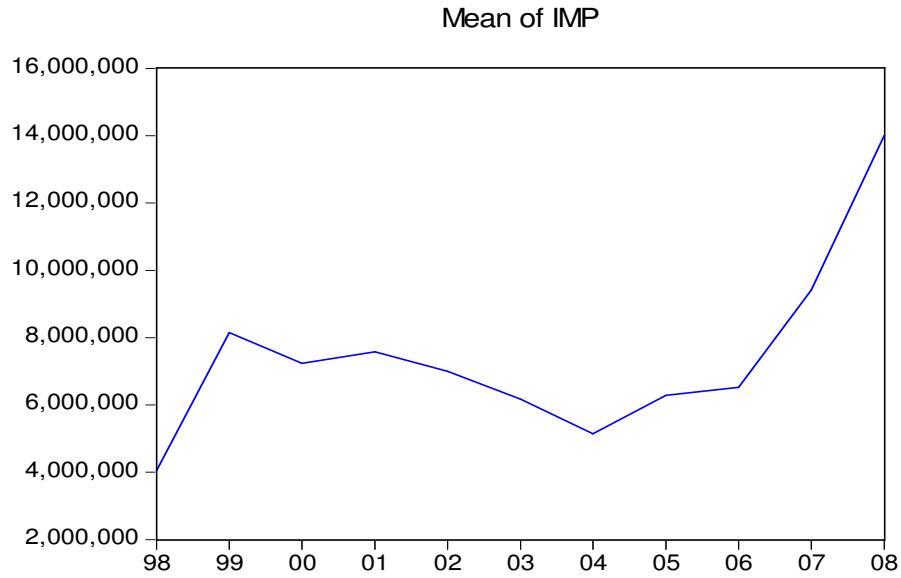
Si analizamos la estructura de esas importaciones, por ejemplo, en el año 2006, veremos que los **887,330,645.21 €** de importaciones representa un 68% respecto al PIB a precios corrientes,

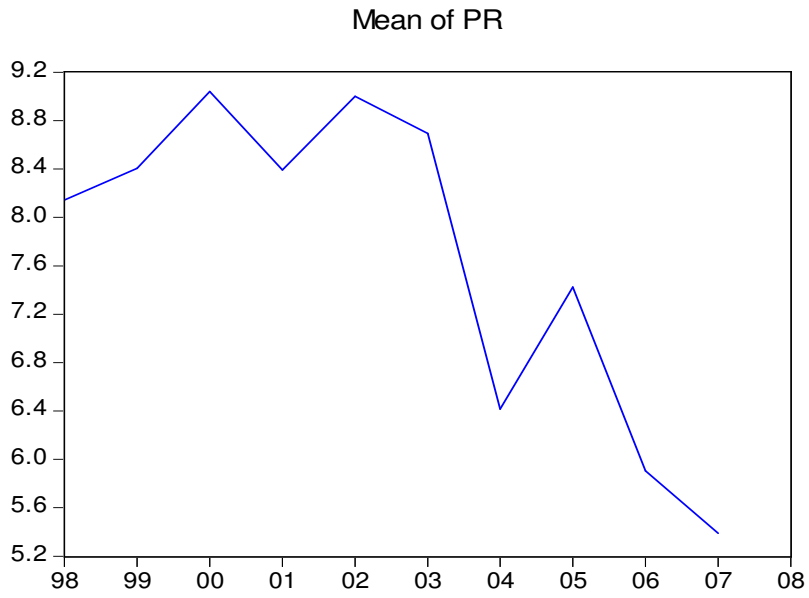
mientras que a nivel nacional las importaciones suponen un 26% del PIB en los mismos términos. Estas cifras muestran la importancia de las importaciones para la economía ceutí.

Respecto a las diferencias, los cálculos ofrecen un montante de posibles exportaciones a Marruecos equivalente a **405,285,933.04 €**, lo que significa el 45.6% del total. Por su parte, la cifra a la que asciende el consumo por debajo de la media nacional se eleva a **-61,063,360.48 €**, lo cual supone aproximadamente un 6.9% del total y representa aquellas partidas de productos que se adquieren tanto Marruecos como en la Península, pero que no se registran como importaciones.

Por su parte, en el ejercicio 2007 se mantienen porcentajes muy similares.

A continuación vamos a mostrar las dos gráficas de la evolución media anual de la variable representativa de las exportaciones a Marruecos desde Ceuta (IMP) y la de los precios relativos (PR), al objeto de ver si su evolución es similar.





Como se puede observar, la evolución de los precios relativos de las mercancías entre Marruecos y Ceuta, que como hemos dicho compara los aranceles de Marruecos con los impuestos locales en Ceuta, es claramente descendente desde el año 2003. Sin embargo, desde ese año, la gráfica de las exportaciones desde Ceuta a Marruecos muestra una tendencia

creciente, es decir, que aparentemente no estaría influyendo dicho desarme arancelario en la actividad que se realiza desde Ceuta, quizás porque la estructura de distribución comercial de Marruecos aún no ha sido adaptada a la nueva situación, o porque los productos que más se estén demandante vengan de países de fuera de la Unión Europea. Este análisis se hará en una fase posterior.

4.2 Estimaciones Econométricas

Como se ha dicho antes, se ha construye un panel de datos con 56 grupos de productos de consumo, en las que se incluyen las diferencias con la media nacional respecto a las importaciones a Ceuta, y una serie temporal de 1998 a 2007. Estas cantidades se ponen en relación con los precios relativos (aranceles/IPSI), el VA de Ceuta y el PIB de Marruecos. Posteriormente se pone en relación el VA de Ceuta con dichas importaciones y con el PIB de Marruecos.

Los resultados del panel de datos sobre importaciones, calculado por Mínimos Cuadrados en dos etapas (MC2E) y por efectos fijos en las secciones transversales (productos), son los siguientes:

Dependent Variable: IMP
Method: Panel Two-Stage Least Squares
Date: 03/13/09 Time: 21:20
Sample (adjusted): 2000 2006
Periods included: 7
Cross-sections included: 56
Total panel (balanced) observations: 392
Instrument list: C PR VA VA(-1) PIBMADHPCR

	Coefficien	Std. Error	t-Statistic	Prob.
C	9122626.	2342632.	3.894178	0.0001
PR	86100.01	42169.41	2.041765	0.0420
VA	47.43790	46.75142	1.014684	0.3110
VA(-1)	-64.69124	50.67891	-1.276492	0.2027
PIBMADHPCR	14.73949	11.36647	1.296752	0.1956

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.920676	Mean dependent var	6562528.
Adjusted R-squared	0.906580	S.D. dependent var	16363894
S.E. of regression	5001581.	Sum squared resid	8.31E+15
F-statistic	65.31163	Durbin-Watson stat	1.439266
Prob(F-statistic)	0.000000	Second-Stage SSR	8.31E+15
Instrument rank	60.00000		

Como se observa, la única variable que resultó con coeficiente positivo y estadísticamente significativo fue la correspondiente a los precios relativos, lo que significa que conforme disminuyan los precios relativos, las exportaciones desde Ceuta también tenderán a disminuir, y ello a pesar de que gráficamente la tendencia parecía contraria.

Si ahora mostramos la estimación del panel de datos sobre el Valor añadido local, tanto en su versión normal, como con las correcciones efectuadas por el método Arellano y Bon, que en general consigue mayor significancia en los coeficientes individuales:

Dependent Variable: VA
Method: Panel Two-Stage Least Squares
Date: 03/13/09 Time: 21:22
Sample (adjusted): 2000 2006
Periods included: 7
Cross-sections included: 56
Total panel (balanced) observations: 392
Instrument list: C IMP VA VA(-1) PIBMADHPCR

	Coefficient	Std. Error	t-Statistic	Prob.
C	-17117.83	2528.695	-6.769433	0.0000
IMP	7.08E-05	6.38E-05	1.110747	0.2675
VA(-1)	1.075863	0.007721	139.3480	0.0000
PIBMADHPCR	0.022514	0.013256	1.698453	0.0904

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.999740	Mean dependent var	783240.6
Adjusted R-squared	0.999694	S.D. dependent var	335127.9
S.E. of regression	5858.802	Sum squared resid	1.14E+10
F-statistic	22051.56	Durbin-Watson stat	2.888890
Prob(F-statistic)	0.000000	Second-Stage SSR	1.14E+10
Instrument rank	60.000000		

Dependent Variable: VA
Method: Panel Generalized Method of Moments
Transformation: First Differences
Date: 03/13/09 Time: 21:28
Sample (adjusted): 2001 2006
Periods included: 6
Cross-sections included: 56
Total panel (balanced) observations: 336
White period instrument weighting matrix
White period standard errors & covariance (d.f. corrected)
Instrument list: @DYN(VA,-2) IMP PIBMADHPCR

	Coefficien	Std. Error	t-Statistic	Prob.
VA(-1)	1.077110	0.000873	1233.224	0.0000
IMP	0.000128	2.18E-05	5.853456	0.0000
PIBMADHPCR	0.020735	0.001585	13.08409	0.0000

Effects Specification

Cross-section fixed (first differences)

Mean dependent var	51076.20	S.D. dependent var	23242.26
S.E. of regression	9234.060	Sum squared resid	2.84E+10
J-statistic	55.63534	Instrument rank	11.00000

Como se dijo antes, la variación de los aranceles de Marruecos ejerce una influencia positiva y significativa sobre el intercambio comercial que se realiza desde Ceuta. Es decir, la evolución de las exportaciones desde Ceuta van a tener la misma evolución futura que los precios relativos, disminuyendo si lo hacen estos.

Sin embargo, sobre la evolución del VA de Ceuta sólo ejerce una influencia significativa el VA del año anterior, y también la evolución del PIB de Marruecos, en sentido positivo. Con la estimación más fina para panel dinámico (MGM, Arellano y Bond) resultó también significativa la variable de intercambio comercial IMP.

Esto nos estaría indicando que, si bien la disminución de las importaciones que podría producir el desarme arancelario ejercería una influencia negativa en la evolución de la economía local, ésta podría quedar neutralizada por el crecimiento económico de Marruecos, tanto si sigue avanzando al ritmo actual, como si dicho desarme lo hace incrementarse aún más.

5. CONCLUSIONES

Las primeras estimaciones indican que el desarme arancelario de Marruecos ejercerá una influencia significativa sobre la evolución de la economía local a través de la disminución de las importaciones**,** aunque esto puede quedar compensado vía crecimiento económico de dicho país.

Ahora bien, el hecho de que los vínculos comerciales entre nuestros vecinos y la Unión lleguen a su culmen mediante la realización plena del área de libre cambio, puede dejar a Ceuta en una situación de aislamiento. Esto sería así en el caso de que se intensificara el comercio Marruecos-UE a expensas de la reducción de la intensidad comercial entre Ceuta y la Unión, por una parte, y con el país vecino, por otra. Ante esta hipótesis, no cabría duda de que la ciudad quedaría marginada de los tráficos comerciales dentro del triángulo UE-Marruecos-Ceuta, en cuyo caso lo más conveniente será integrarla en unión aduanera comunitaria, por razones ante todo estratégicas.

Y ante tal situación, sería además necesario dotarla de las ventajas de región ultraperiférica, como única forma de frenar el continuo deterioro de su modelo económico, por un lado, y de reconocer su peculiaridad de aparecer ligada al desarrollo económico de un país emergente, a pesar de pertenecer a la UE.

En definitiva, el aislamiento de Ceuta de los centros económicos nacionales y comunitarios, que ya se ha encargado la geografía de marcar, no debe verse agudizado por el nuevo esquema de relaciones comerciales que involucra a las economías de su entorno. La política económica local debe aprovechar la sinergia que desencadene el nuevo proceso de integración comercial, evitando toda posibilidad de un mayor aislamiento económico de la ciudad, y situándola en una posición favorable de cara a su convergencia real en el conjunto de regiones de la UE.

BIBLIOGRAFÍA

Acuerdo de Asociación UE-Marruecos. *Diario Oficial número L 70 de 18/3/2000.*

ARCE BORDÁ, R. DE Y MAHÍA CASADO, R. (2004): "Estimación analítica de los efectos de la creación de un área de libre comercio agrícola entre la UE. y Marruecos", *ICE*, nº 819, pp. 191-210.

BATALLER MARTÍN, F. Y JORDÁN GALDUF, J.M. (2000): "El área euromediterránea: esperanzas, logros y frustraciones del proceso de Barcelona", *ICE*, nº 788, pp. 161-177.

BONET MADURGA, A. (2002): "Métodos casuísticos de evaluación de impacto para negociaciones comerciales internacionales", Banco Interamericano de Desarrollo.

CONSEJO ECONÓMICO Y SOCIAL DE CEUTA: *MEMORIA 2005 sobre la situación socioeconómica y laboral*.

DECLARACIÓN FINAL Y PROGRAMA DE TRABAJO DE LA CONFERENCIA EUROMEDITERRÁNEA de 27 y 28 de noviembre de 1995.

GRACIA ANDÍA, A.B. Y SANZ VILLARROYA, I. (2006): "Cambios estructurales en las series de comercio exterior español 1960-1997", *Ekonomiaz*; nº 62.

HOOPER P. Y KOHLHAGEN S.W. (1978): "The effects of Exchange rate uncertainly on the Price and volumen of international trade", *Journal of International Economic*, vol. 8, pp. 483-511.

PROCESA, Sociedad de Fomento de Ceuta: *Anuario Estadístico Ciudad Autónoma de Ceuta 2004*.

PROCESA, Sociedad de Fomento de Ceuta: *Anuario Estadístico Ciudad Autónoma de Ceuta 2006*.

PROCESA, Sociedad de Fomento de Ceuta: *Anuario Estadístico Ciudad Autónoma de Ceuta 2007*.

PROCESA, Sociedad de Fomento de Ceuta: *Guía para invertir en Ceuta 2004*.

RAMIL DÍAZ, M. (2001): "Las importaciones de mercancías en la economía española", *Estudios de Economía Aplicada*, vol.19, nº3, pp. 123-128.

LA PARTICIPACIÓN LABORAL Y LA JUBILACIÓN ENTRE LAS PERSONAS MAYORES EN CASTILLA Y LEÓN.

JESÚS MARÍA GÓMEZ GARCÍA
E-mail: jmgomez@eco.uva.es
YOLANDA GONZÁLEZ GONZÁLEZ
E-mail: yolanda@eco.uva.es
Departamento de Economía Aplicada
Facultad de Ciencias Económicas y Empresariales
Universidad de Valladolid (España)

Resumen

Este trabajo analiza la relación de las personas mayores que residen en la Comunidad Autónoma de Castilla y León con el mercado laboral, teniendo en cuenta el progresivo envejecimiento que viene experimentando la población regional. El objetivo último es perfilar los rasgos que caracterizan al colectivo de trabajadores de más edad y de las personas mayores en relación con la actividad laboral y su experiencia en la jubilación. Para ello, se analizarán indicadores básicos del mercado laboral en España y en Castilla y León tomando como referencia el año 2007. La principal fuente estadística utilizada será la "Encuesta de Población Activa" (EPA), junto con datos de la "Encuesta de Condiciones de Vida de las Personas Mayores" (ECVM) y del "Informe 2006. Las personas mayores en España". También se han utilizado resultados del estudio "Situación económica y social de las personas mayores de Castilla y León, Año 2007" para el caso específico de la Comunidad Autónoma de Castilla y León.

Palabras clave: Envejecimiento demográfico, Participación laboral, Jubilación, Personas mayores.
Área Temática: Economía Urbana, regional y local

Abstract

This paper analyzes the relationship to the labour market of older persons living in the Autonomous Community of Castilla y León, taking into account the aging which has experienced the regional population. The objective of this paper is to outline the features that characterize a group of older workers and the elderly in relation to their work and experience in retirement. Accordingly, it will analyze key indicators of the labour market in Spain and Castilla y León taking 2007 as the base year. The main statistical source used is the "Labour Force Survey" (LFS), together with other data from the "Survey of Living Conditions of Older Persons" (ECVM) and the "2006 Report Older people in Spain". Data from the study "Economic and social situation of older people of Castile and Leon, Year 2007" have also been used, but only for the specific case of the Autonomous Community of Castilla y León.

Keywords: Aging population, Labour force participation, Retirement, Elderly people.
Thematic Area: Urban, Regional and Local Economics.

1. INTRODUCCIÓN.

Una de las dimensiones básicas de la vida de las personas mayores es su relación con el mercado de trabajo, ya sea ésta resultado de la participación de forma activa en el mismo o, al contrario, por su inactividad laboral. A continuación se analiza la relación de las personas mayores de la Comunidad Autónoma de Castilla y León con el mercado laboral. El objetivo es perfilar los rasgos que caracterizan al colectivo de trabajadores de más edad y de las personas mayores en relación con la actividad laboral y su experiencia en la jubilación. Para ello, se analizarán indicadores básicos del mercado laboral en España y en Castilla y León tomando como referencia el año 2007. La principal fuente estadística utilizada será la "Encuesta de Población Activa" (EPA), junto con datos de la "Encuesta de Condiciones de Vida de las

Personas Mayores" (ECVM) y del "Informe 2006. Las personas mayores en España". También se han utilizado resultados del estudio "Situación económica y social de las personas mayores de Castilla y León, Año 2007" para el caso específico de la Comunidad Autónoma de Castilla y León.

2. LA ACTIVIDAD, LA OCUPACIÓN Y EL PARO.

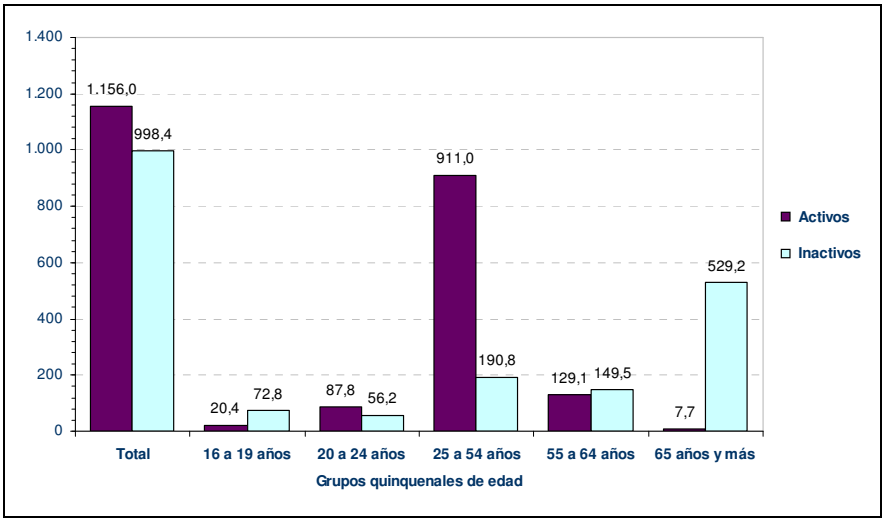
Los datos que proporciona la Encuesta de Población Activa señalan que en la Comunidad Autónoma de Castilla y León había en promedio para el año 2007 un total de 2.154.400 personas en edad de trabajar (16 y más años), de las que 1.156.000 se encontraban en situación laboralmente activa y el resto eran inactivos (Tabla 1). Dentro del conjunto de activos, el 16,8% tenían 55 y más años y sólo el 1,4% tenían 65 años o más de edad. Esta realidad pone de relieve que las personas mayores mayoritariamente llevan a cabo su vida al margen del mercado de trabajo, siendo muy pocos los que aún se mantienen laboralmente activos una vez que alcanzan la edad de jubilación (Gráfico 1). Como se aprecia, la participación de las personas mayores de Castilla y León en el mercado de trabajo y la actividad económica esconde amplias diferencias debido a que se trata de un colectivo social muy heterogéneo. Este hecho se pone de manifiesto cuando se realiza un análisis del mercado laboral detallando por sexo y grupos de edad, al igual que si se realiza la comparación de la situación de Castilla y León con la de otras regiones y con el contexto nacional.

Tabla 1. Población de 16 y más años según relación con la actividad, según sexo y grupo de edad, 2007 (Media anual, en miles de personas)

	España				Castilla y León			
Ambos sexos	**Pob. 16+**	**Activos**	**Ocupados**	**Parados**	**Pob. 16+**	**Activos**	**Ocupados**	**Parados**
Total	37.662,9	22.189,9	20.356,0	1.833,9	2.154,4	1.156,0	1.073,0	83,0
16 a 19 años	1.844,7	548,2	390,8	157,4	93,2	20,4	14,3	6,1
20 a 24 años	2.797,6	1.885,9	1.600,5	285,3	144,0	87,8	75,0	12,7
25 a 54 años	20.946,9	17.349,5	16.093,2	1.256,4	1.101,8	911,0	852,5	58,5
55 años y más	12.073,7	2.406,4	2.271,6	134,8	815,6	136,9	131,2	5,7
Varones								
Total	18.480,7	12.802,5	11.987,3	815,2	1.062,2	680,1	647,8	32,4
16 a 19 años	948,3	329,0	250,4	78,6	47,9	13,1	9,6	3,5
20 a 24 años	1.431,2	1.032,6	903,8	128,9	73,9	51,2	45,3	5,8
25 a 54 años	10.665,6	9.879,2	9.343,8	535,4	565,9	524,0	503,6	20,4
55 años y más	5.435,5	1.561,7	1.489,3	72,4	374,6	91,9	89,2	2,7
Mujeres								
Total	19.182,2	9.387,4	8.368,8	1.018,7	1.092,2	475,9	425,3	50,6
16 a 19 años	896,4	219,2	140,4	78,8	45,3	7,3	4,7	2,6
20 a 24 años	1.366,4	853,3	696,8	156,5	70,1	36,6	29,7	6,9
25 a 54 años	10.281,3	7.470,4	6.749,4	721,0	535,9	387,0	348,9	38,1
55 años y más	6.638,2	844,7	782,2	62,4	441,0	45,0	42,0	3,0

Fuente: Elaboración propia a partir de datos de EPA (INE).

Gráfico 1. Población de 16 y más según grupos de edad en Castilla y León, 2007. (Miles de personas)



Fuente: Elaboración propia a partir de datos de EPA (INE).

2.1. La actividad.

La población activa de Castilla y León, que alcanzó la cifra de 1.156.000 personas en el año 2007, aumentó el 1,6% en promedio anual con respecto al año anterior. La tasa de actividad de la región creció en igual cuantía que la media nacional (0,6 puntos porcentuales) a lo largo de 2007 hasta situarse en el 53,7% de la población en edad de trabajar (de 16 y más). Pese a ello, dicha tasa refleja que Castilla y León sigue siendo una de las regiones españolas con menor actividad ya que sólo es superada por Extremadura, Asturias y las ciudades autónomas de Ceuta y Melilla, manteniendo una posición lejos de la tasa de actividad media para España (del 58,9% en ese mismo año) (Tabla 2).

Tabla 2. Tasas de actividad, empleo y paro de la población de 16 años y más y de la población de 55 años y más por CC.AA., 2007 (Media anual)

	Población de 16 años y más			Población de 55 años y más		
	Actividad	Empleo	Paro	Actividad	Empleo	Paro
Andalucía	56,2%	49,1%	12,8%	17,1%	15,3%	10,2%
Aragón	58,8%	55,7%	5,2%	19,1%	18,3%	4,1%
Asturias (Principado de)	50,5%	46,3%	8,5%	15,9%	15,3%	3,4%
Baleares	64,1%	59,6%	7,0%	23,9%	23,0%	3,7%
Canarias	61,0%	54,7%	10,4%	22,6%	20,9%	7,7%
Cantabria	56,2%	52,9%	5,9%	18,9%	18,4%	2,4%
Castilla y León	53,7%	49,8%	7,2%	16,8%	16,1%	4,1%
Castilla - La Mancha	56,0%	51,8%	7,6%	16,6%	15,8%	5,0%
Cataluña	62,5%	58,4%	6,5%	23,3%	22,1%	4,9%
Comunidad Valenciana	60,1%	54,9%	8,8%	20,2%	19,2%	5,1%
Extremadura	52,9%	45,9%	13,1%	16,4%	14,8%	9,8%
Galicia	54,3%	50,1%	7,6%	18,0%	17,1%	5,1%
Madrid (Comunidad de)	63,9%	59,9%	6,3%	23,8%	22,7%	4,7%
Murcia (Región de)	60,5%	55,9%	7,6%	19,3%	18,2%	5,7%
Navarra	60,8%	57,9%	4,8%	21,6%	21,0%	2,7%
País Vasco	57,6%	54,1%	6,1%	19,5%	18,8%	3,8%
Rioja (La)	58,9%	55,6%	5,7%	19,2%	18,7%	2,8%
Ceuta y Melilla	52,8%	42,6%	19,3%	18,5%	16,5%	10,8%
Total Nacional	**58,9%**	**54,0%**	**8,3%**	**19,9%**	**18,8%**	**5,6%**

Fuente: Elaboración propia a partir de datos de EPA (INE).

La tasa de actividad según sexo para Castilla y León presenta también una peor situación que el caso nacional. El 64,0% de los varones castellano-leoneses están en activo frÿÿte al 69,3ÿÿnacional, siendo el 43,6% de las mujeres de la región las activas frente al 48,9% del país (Tabla 3). El estudio de la tasa de actividad dentro del colectivo de trabajadores de 55 años y más, muestra un comportamiento de crecimiento en Castilla y León al igual que en la media nacional, pero unos valores para la primera inferiores. Concretamente, del 0,6% desde el año 2005 al 2007 para la región y de algo menos de un punto en España (0,8 puntos). La Comunidad Autónoma de Castilla y León, situada por debajo de la media nacional, ha tenido también un crecimiento, pero en una proporción algo menor (el 0,6%). Por otra parte, la región muestra una de las menores tasas de actividad en este tramo de edad (el 16,8% en 2007), situándose por debajo solamente las tasas correspondientes a las Comunidades Autónomas de Castilla-La Mancha (16,6), Extremadura (16,4) y Asturias (15,9), exceptuando Ceuta y Melilla.

Tabla 3. Tasas de actividad, empleo y paro según sexo y edad, 2007
(Media anual)

	España			Castilla y León		
Ambos sexos	**Actividad**	**Empleo**	**Paro**	**Actividad**	**Empleo**	**Paro**
Total	58,9%	54,0%	8,3%	53,7%	49,8%	7,2%
16 a 19 años	29,7%	21,2%	28,7%	21,9%	15,3%	30,0%
20 a 24 años	67,4%	57,2%	15,1%	61,0%	52,1%	14,5%
25 a 54 años	82,8%	76,8%	7,2%	82,7%	77,4%	6,4%
55 años y más	19,9%	18,8%	5,6%	16,8%	16,1%	4,1%
Varones						
Total	69,3%	64,9%	6,4%	64,0%	61,0%	4,8%
16 a 19 años	34,7%	26,4%	23,9%	27,4%	20,0%	26,9%
20 a 24 años	72,1%	63,1%	12,5%	69,2%	61,3%	11,4%
25 a 54 años	92,6%	87,6%	5,4%	92,6%	89,0%	3,9%
55 años y más	28,7%	27,4%	4,6%	24,5%	23,8%	2,9%
Mujeres						
Total	48,9%	43,6%	10,9%	43,6%	38,9%	10,6%
16 a 19 años	24,4%	15,7%	35,9%	16,1%	10,3%	35,6%
20 a 24 años	62,4%	51,0%	18,3%	52,2%	42,4%	18,8%
25 a 54 años	72,7%	65,6%	9,7%	72,2%	65,1%	9,8%
55 años y más	12,7%	11,8%	7,4%	10,2%	9,5%	6,7%

Fuente: Elaboración propia a partir de datos de EPA (INE).

Adicionalmente, si se diferencia, dentro de los trabajadores de Castilla y León de 55 años y más, a aquellos que tienen menos de 65 y a aquellos de 65 años y más años, se constata un brusco descenso de las tasas de actividad en estos tramos de edad, puesto que, por un lado, solamente menos de la mitad de las personas de entre 55 y 64 años siguen siendo laboralmente activas (en torno al 47% para España y para Castilla y León) y, por otro, son una escasa minoría las personas activas mayores de 64 años (el 2% y el 1,4%, respectivamente). Esto indica que la retirada del mercado de trabajo se ha adelantado temporalmente a la edad de los 65 años que marca el momento de la jubilación (Tabla 4).

Tabla 1.4. Tasas de actividad de las personas mayores según sexo y edad, 2007 (Media anual)

	España	Castilla y León
Ambos sexos		
55 a 64 años	47,4%	46,3%
65 años y más	2,0%	1,4%
Varones		
55 a 64 años	63,1%	61,5%
65 años y más	3,2%	2,3%
Mujeres		
55 a 64 años	32,5%	30,9%
65 años y más	1,1%	0,7%

Fuente: Elaboración propia a partir de datos de EPA (INE).

El resultado anterior se corrobora también si distinguimos la población por género, observando unas tasas de actividad femeninas muy inferiores a las masculinas. En el caso de los varones de entre 55 a 64 años, en torno al 60% permanecen laboralmente activos, mientras que tal situación solo la presenta algo menos de un tercio de las mujeres de este grupo de edad. Es muy posible que buena parte de estas mujeres nunca hayan realizado una actividad laboral formal y remunerada. Estas diferencias se manifiestan tanto en Castilla y León como dentro del contexto nacional, pero con unas tasas de actividad de la región, para ambos sexos, notablemente más bajas que el promedio nacional.

Hay que destacar también, que la evolución del mercado de trabajo en los últimos años en España ha sido positiva en el sentido del aumento producido en las tasas de actividad para ambos sexos. Particularmente, sobresale el crecimiento de las tasas para las mujeres y para todos los grupos de edad, con mayor magnitud en los jóvenes de 20 a 24 años. Este incremento de las tasas de actividad ha sido menos importante para los trabajadores de más edad, sin embargo, constituye un cambio de tendencia respecto de la última década[1]. En Castilla y León la evolución ha sido algo diferente, pues la tasa de actividad de los varones de 55 años y más sufre un leve descenso entre los años 2005 a 2007, mientras que la tasa correspondiente a las mujeres ha experimentado un ascenso más acusado que el que se experimenta en el contexto nacional (CES-CyL, 2009).

2.2. La ocupación.

La población ocupada de Castilla y León se situó en 1.073.000 personas, en promedio para el año 2007. Esta cifra representa un aumento del 2,6% respecto del año anterior, frente al 3,1% que experimenta la ocupación a escala nacional. La representación del empleo en la región fue, en el citado año, del 49,8%, superando en un punto porcentual a la del año anterior. En España la tasa de empleo alcanza el 54% (un 0,7 puntos porcentuales superior a la registrada en 2006). Estos datos sitúan a Castilla y León entre las Comunidades Autónomas con menor tasa de empleo, solo por encima de los casos de Andalucía, Extremadura y las ciudades autónomas de Ceuta y Melilla (Tabla 1).

El estudio de esta variable atendiendo al género, muestra que la tasa de empleo de los varones de Castilla y León se situó en el 61,0% y la de las mujeres en el 38,9%, frente a unas tasas del 64,9% y del 43,6% en España. Por lo que respecta al colectivo laboral de mayor edad (55 y más años), la tasa de empleo regional fue del 16,1% en el año 2007, frente al 18,8% que se registra a escala nacional en ese año (Tabla 3). Esta cifra, al igual que ocurría con las tasas de actividad, sitúa a Castilla y León entre las regiones españolas con menor proporción de población

1 En España, la tasa de actividad de los varones de 65 a 69 años ha aumentando entre 2004 y 2005 en algo más de 1 punto porcentual y la de las mujeres en algo menos de medio punto. Igualmente, varones y mujeres de 60 a 64 años han aumentado su vinculación con el mercado en unos 3 puntos en dicho período (IMSERSO: "Las personas mayores en España. Informe 2006", pág.205).

ocupada de mayor edad, sólo superando las correspondientes a Extremadura (14,8%), Andalucía (15,3%), Asturias (15,3%) y Castilla La Mancha (15,8%) (Tabla 1).

La evolución de la tasa de empleo para los trabajadores de más edad (55 años y más) muestra una tendencia positiva en los últimos tres años, tanto en España y como en Castilla y León, con incrementos de 0,8 y 0,6 puntos porcentuales, respectivamente, aunque, evidentemente, lejos todavía de los casi 2 puntos porcentuales en que aumenta dicha tasa para el total de la población ocupada en igual período temporal. Dicha evolución muestra también con claridad la peor situación que, en general, presentan las mujeres frente a lo que ocurre para los varones en términos de empleo. Concretamente, en 2007 y para el colectivo de trabajadores de 55 y más años, la tasa de empleo de Castilla y León fue del 23,8% para el caso de los varones, frente al 9,5% para el de las mujeres. Ambas tasas regionales se encuentran muy por debajo de los valores medios que se alcanzan dentro del contexto nacional, que son del 27,4% y del 11,8%, respectivamente. Se aprecia, asimismo, un estancamiento del ritmo de crecimiento de empleo entre los varones de mayor edad de Castilla y León, hecho que contrasta con el creciente dinamismo en el empleo que se observa entre las mujeres de la región integradas en dicho grupo de edad. Estas tendencias se manifiestan también a escala nacional, si bien se encuentran agudizadas en el contexto laboral de la región, particularmente en lo que respecta a las mujeres trabajadoras de mayor edad, entre las que la tasa de empleo creció 1,7 puntos porcentuales entre 2005 y 2007, frente a los 1,1 puntos de crecimiento promedio resultantes para España (CES-CyL, 2009).

Un recorrido por la distribución de la ocupación sectorial de los trabajadores de más edad en España refleja que, en 2007, dicha ocupación se concentró en la industria y, fundamentalmente, en el sector de los servicios (Tabla 5).

Tabla 5. Población ocupada de 55 años según sexo y sector de actividad, 2007.
(Media anual, en miles de personas)

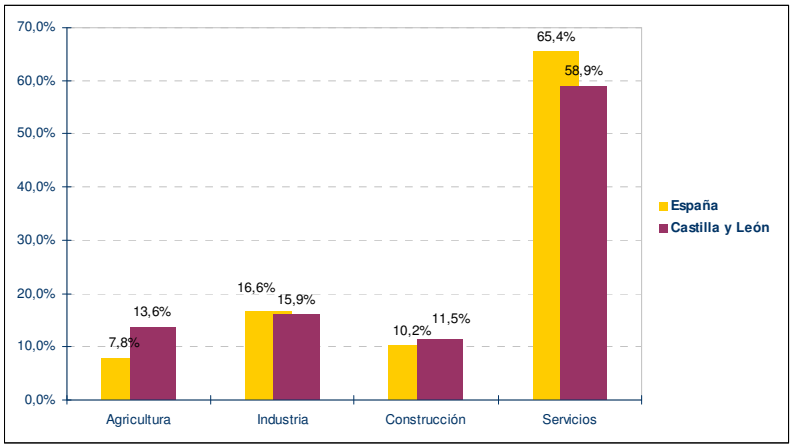
		Agricultura	Industria	Construcción	Servicios	Total sectores
Castilla y León	Total	17,9	20,9	15,1	77,3	131,2
	Varones	14,2	18,8	14,7	41,6	89,3
	Mujeres	3,7	2,1	0,5	35,7	42
España	Total	178	377,1	230,8	1.485,8	2.271,7
	Varones	132,2	319,4	225,4	812,3	1.489,3
	Mujeres	45,8	57,6	5,3	673,5	782,2

Fuente: Elaboración propia a partir de datos de EPA (INE).

La misma pauta distributiva anterior se mantiene en Castilla y León, siendo las ramas del sector terciario (el 58,9%) y las industriales (el 15,9%) las que aglutinan a la mayoría de los ocupados de este grupo de edad. No obstante, destaca en el ámbito regional el mayor peso proporcional que representan los empleos en el sector agrario, frente al contexto nacional (13,6% y 7,8%, respectivamente), lo que refleja la relevancia que aún hoy ostenta la actividad agraria entre los trabajadores de más edad de la región (Gráfico 2).

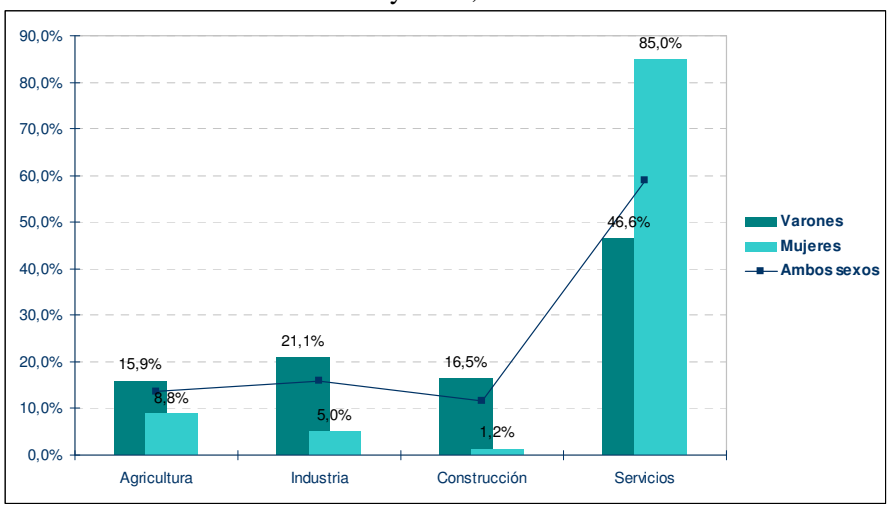
La distribución sectorial de la población ocupada de 55 años y más de Castilla y León presenta amplias diferencias cuando se distingue por género. Destaca la abrumadora presencia de las mujeres con empleos en el sector servicios (el 85%), la cual, casi duplica la de los varones. En el resto de sectores, en cambio, la relevancia de las mujeres ocupadas de 55 años y más es muy inferior a la de los varones. Las diferencias con respecto al contexto nacional se centran particularmente en el sector agrario regional, en el que encuentran ocupación un porcentaje mucho mayor de varones de más edad que el promedio para España, con el 15,9% y 8,9% del total, respectivamente (Gráfico 3).

Gráfico 2. Distribución sectorial de la población ocupada de 55 años y más en España y en Castilla y León, 2007.



Fuente: Elaboración propia a partir de datos de EPA (INE).

Gráfico 3. Distribución sectorial de la población ocupada de 55 años y más según sexo en Castilla y León, 2007.



Fuente: Elaboración propia a partir de datos de EPA (INE).

Otra de las características del colectivo de trabajadores ocupados de 55 años y más es la baja proporción de empleo asalariado. Concretamente, dicha proporción es del 9,8% del total de este grupo de edad en Castilla y León, frente al 9,2% en España (Tabla 6). Este dato sugiere que las personas que siguen trabajando en el tramo de edad más alto son, en buena parte, trabajadores por cuenta propia y, por tanto, que tienen mayor capacidad para decidir cuándo abandonar la actividad laboral que desempeñan.

Tabla 6. Asalariados según sexo y edad en España y Castilla y León, 2007.(Media anual)

		16 a 19 años	20 a 24 años	25 a 34 años	35 a 44 años	45 a 54 años	55 años y más
Castilla y León	Total	1,4%	8,2%	30,7%	27,4%	22,4%	9,8%
	Varones	1,6%	8,5%	28,6%	27,2%	22,9%	11,1%
	Mujeres	1,2%	7,9%	33,6%	27,8%	21,6%	8,0%
España	Total	2,1%	8,9%	31,8%	27,8%	20,1%	9,2%
	Varones	2,4%	8,7%	30,7%	27,7%	20,2%	10,3%
	Mujeres	1,8%	9,1%	33,3%	28,0%	20,0%	7,8%

Fuente: Elaboración propia a partir de datos de EPA (INE).

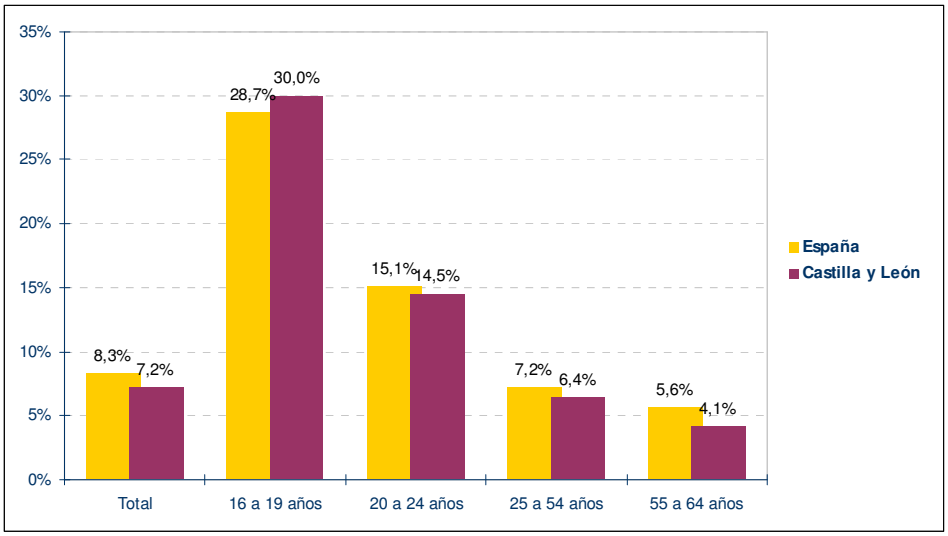
Si dentro de los asalariados tenemos en cuenta el género, se aprecia que la proporción de varones asalariados entre los trabajadores de mayor edad supera ampliamente la correspondiente a las mujeres y, manteniendo este comportamiento tanto en España como en Castilla y León. No obstante, en el caso concreto de los varones, la proporción regional de asalariados es mayor que el promedio para España.

2.3. El paro.

El total de población activa desempleada en Castilla y León fue, en promedio para el año 2007, de 83.000 personas. Dicha cifra representa una disminución del 10,2% (9.400 personas) en relación con el año anterior, mientras que en España el descenso sólo fue del 0,2% en igual período (Tabla 1.1). Los valores de la tasa de paro en Castilla y León son inferiores a la nacional y muestran una evolución similar a la del conjunto español, tanto para el conjunto de la población, como para el grupo de personas entre 55 y más años. Así, en promedio para el año 2007, la tasa de paro fue del 7,2% en Castilla y León y del 8,3% en el conjunto nacional, lo que supone una reducción de 0,9 y 0,2 puntos porcentuales sobre el año anterior en la región y en España, respectivamente. De acuerdo con estos resultados, Castilla y León, ocuparía el octavo puesto de las regiones españolas (incluidas Ceuta y Melilla) con menor tasa de paro en el año 2007 (Tabla 1).

La tradicional menor tasa de paro de los trabajadores de Castilla y León respecto de la nacional en todos los grupos de edad se rompe para los trabajadores más jóvenes (16 a 19 años) en cuyo caso la tasa de la región supera la nacional. El grupo de 55 años y más presenta una tasa de paro del 4,1% en Castilla y León, frente a un promedio del 5,6% en España (Gráfico 4), lo que la sitúa entre las regiones españolas con menor nivel de paro en este grupo de edad. La evolución del desempleo entre estos trabajadores ha sido, en general, positiva en los últimos tres años (CES-CyL, 2009).

Gráfico 4. Tasa de paro por grupos de edad en España y Castilla y León, 2007.

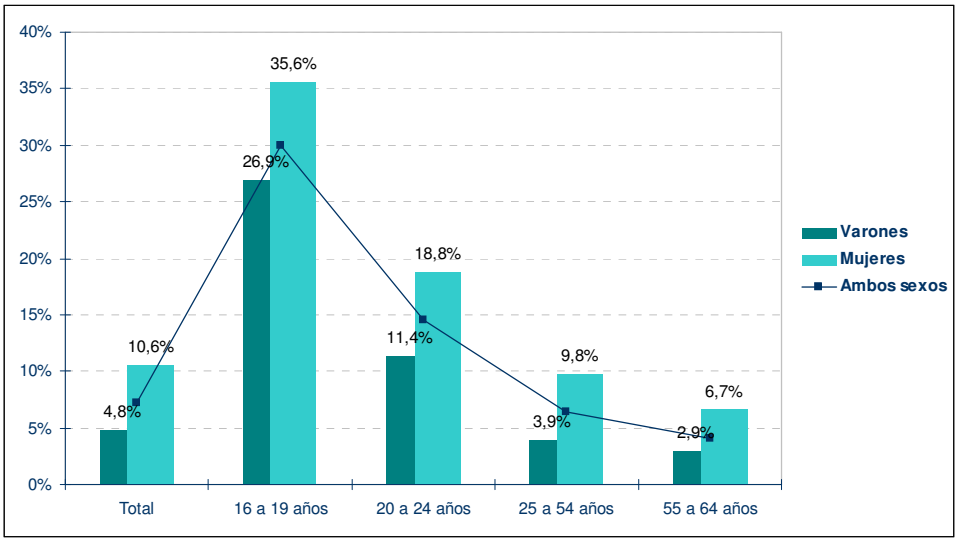


Fuente: Elaboración propia a partir de datos de EPA (INE).

La tasa de paro para los trabajadores varones de mayor edad de Castilla y León es del 2,9% en 2007. Este valor representa un retroceso de 1,4 puntos porcentuales sobre el año anterior, mientras que para España se observa un estancamiento de dichas tasas en torno al 4,6% en el último año. Las tasas de paro femeninas de la región, aunque superan siempre a las de los varones en todas las edades, han evolucionado positivamente, reduciéndose en casi 3 puntos porcentuales durante los tres últimos años para el total de mujeres y en 0,6 puntos porcentuales para las de más edad, siendo en 2007 del 10,6% y del 6,7%, respectivamente (Gráfico 5). En

España, la tasa de paro femenina también se redujo hasta el 10,9%, pero creció 0,3 puntos porcentuales entre las trabajadoras de más edad (7,4%). Las tasas de paro regional y nacional son muy similares en el caso de las mujeres, pero más bajas entre los varones de la región (Tabla 1).

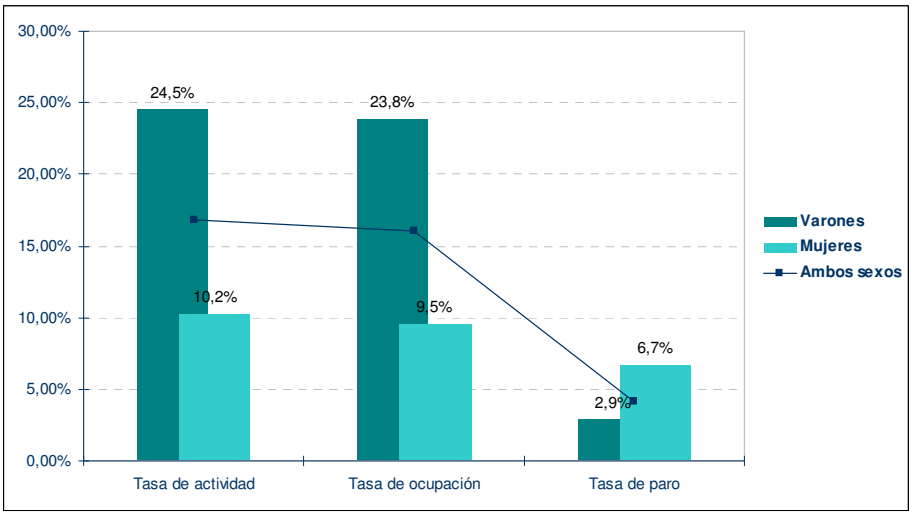
Gráfico 5. Tasa de paro por grupos de edad y sexo en Castilla y León, 2007.



Fuente: Elaboración propia a partir de datos de EPA (INE).

Considerando las tasas de actividad, de ocupación y de paro entre los trabajadores de 55 años y más de Castilla y León se observa que existe una amplia disparidad entre sexos pues las tasas de actividad y de ocupación de los varones más que duplican las de las mujeres, mientras que ocurre lo contrario en cuanto a las tasas de paro (Gráfico 6).

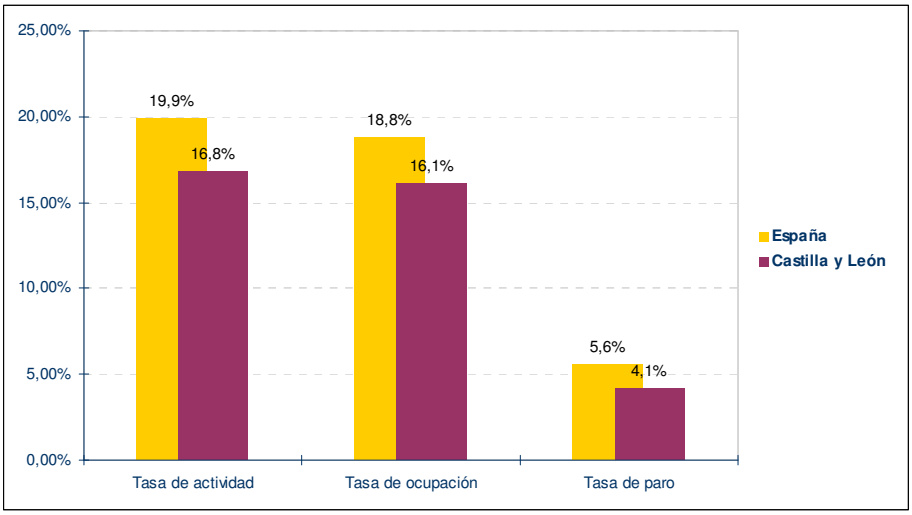
Gráfico 6. Actividad, ocupación y paro de los trabajadores de 55 años y más según sexo en Castilla y León, 2007.



Fuente: Elaboración propia a partir de datos de EPA (INE).

Por otra parte, la comparación de los valores regionales y nacionales de las tres tasas analizadas, pone de manifiesto diferencias entre la situación de los trabajadores de 55 años y más en Castilla y León y en España. Específicamente, las disparidad muestra que las tasas de actividad, de ocupación y de paro regionales son inferiores a las nacionales, aunque la distancia se estrecha en el caso de la tasa de desempleo (Gráfico 7).

Gráfico 7. Actividad, ocupación y paro de los trabajadores de 55 años y más en España y en Castilla y León, 2007.



Fuente: Elaboración propia a partir de datos de EPA (INE).

3. LA INACTIVIDAD Y LA JUBILACIÓN.

3.1. Población inactiva.

La población inactiva total de Castilla y León se estima en 998.400 personas, según fuentes de la EPA para 2007. Dicha población se ha reducido en un 0,7% respecto del año anterior. Este dato positivo no se mantiene, sin embargo, en el grupo de edad de 55 a 64 años, ni tampoco en el grupo de las personas de 65 años y más. Los valores de inactivos en estos grupos apenas han variado en los últimos tres años, totalizando 149.500 y 529.200 personas en 2007, respectivamente. En España, se produce un crecimiento en el total de la población en 2007 igual al 0,3%, aumentando igualmente el número de inactivos en los dos grupos de edad citados (Tabla 7).

La inactividad laboral en el grupo de población de 65 años y más es una tendencia en la que se encuentran, en 2007, un 98,6% de dicho grupo en Castilla y León y que alcanza a más de la mitad (el 53,7%) de las personas entre 55 a 64 años en nuestra región. Estos resultados se observan también en España, aunque la inactividad entre las personas mayores es menos frecuente que en Castilla y León.

Tabla 7. Población inactiva según edad en España y Castilla y León, 2007 (Media anual)

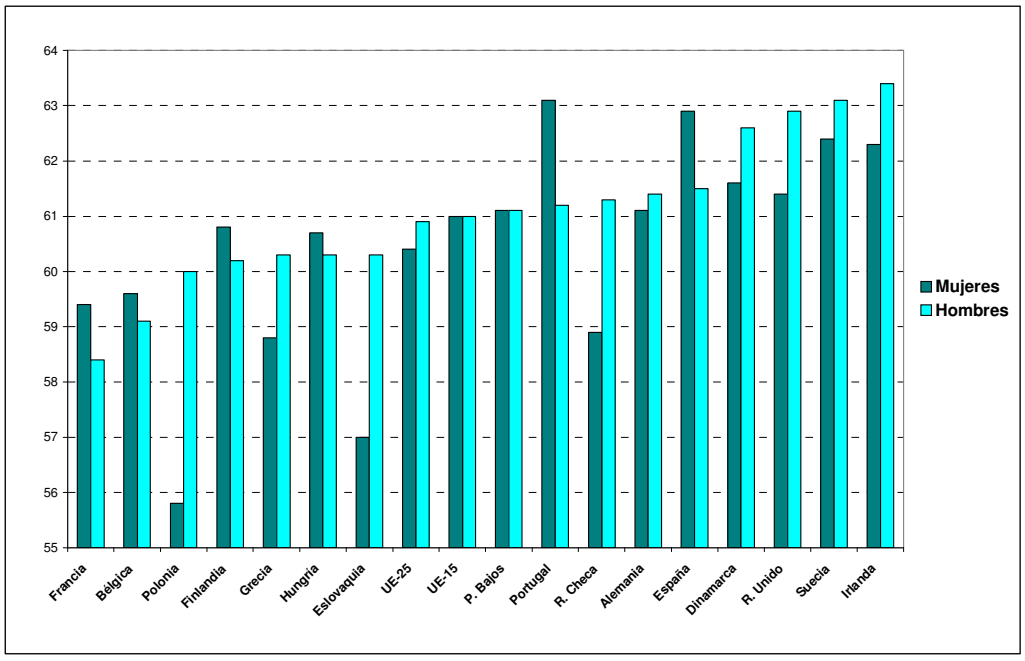
	España		Castilla y León	
	Población (miles)	Porcentaje	Población (miles)	Porcentaje
Ambos sexos				
55 a 64 años	2.510,7	52,6%	149,5	53,7%
65 años y más	7.156,7	98,0%	529,2	98,6%
Varones				
55 a 64 años	855,9	36,9%	54,1	38,5%
65 años y más	3.017,9	96,8%	228,6	97,7%
Mujeres				
55 a 64 años	1.654,7	67,5%	95,4	69,1%
65 años y más	4138	98,9%	300,6	99,3%

Fuente: Elaboración propia a partir de datos de EPA (INE).

Si nos situamos en el la población castellano-leonesa masculina, podemos observar que la de más edad presenta un brusco incremento de la inactividad al llegar a los 65 años (el 97,7%, frente al 38,5% del grupo de edad de 55 a 64 años). Así, los datos muestran que el total de varones inactivos de la región ha aumentado durante el período 2005-2007, afectando particularmente al grupo de edad de 55 a 64 años. Atendiendo a la población femenina, las inactivas representan la práctica totalidad de las que tienen 65 años y más años (el 99,3%) en 2007, si bien su número tiende a disminuir en Castilla y León, al contrario que en España (CES-CyL, 2009).

Aunque la población inactiva puede ser también numerosa en otros colectivos de trabajadores más jóvenes (por ejemplo, los de 16 a 24 años), la situación de inactividad difiere cualitativamente de la que presentan las personas de más edad. Entre estos últimos, la retirada de la actividad laboral se explica fundamentalmente por la jubilación. En este sentido, Eurostat proporciona información sobre el momento de la retirada del mercado laboral, indicando que la edad media de jubilación de la Unión Europea (UE-25) es de 60,7 años, siendo un poco más baja para las mujeres (60,4 años) que para los varones (60,9 años). España es uno de los países con una jubilación más tardía (62,2 años), siendo ésta más baja en el caso de los varones (61,5 años) que en el de las mujeres (62,9 años). Globalmente, entre 2001 y 2004, la edad media de jubilación ha aumentado en casi todos los países europeos. En España, en igual período, la edad media aumentó en más de un año, aunque debido, fundamentalmente, al retraso en casi tres de la edad media de jubilación femenina (Gráfico 8).

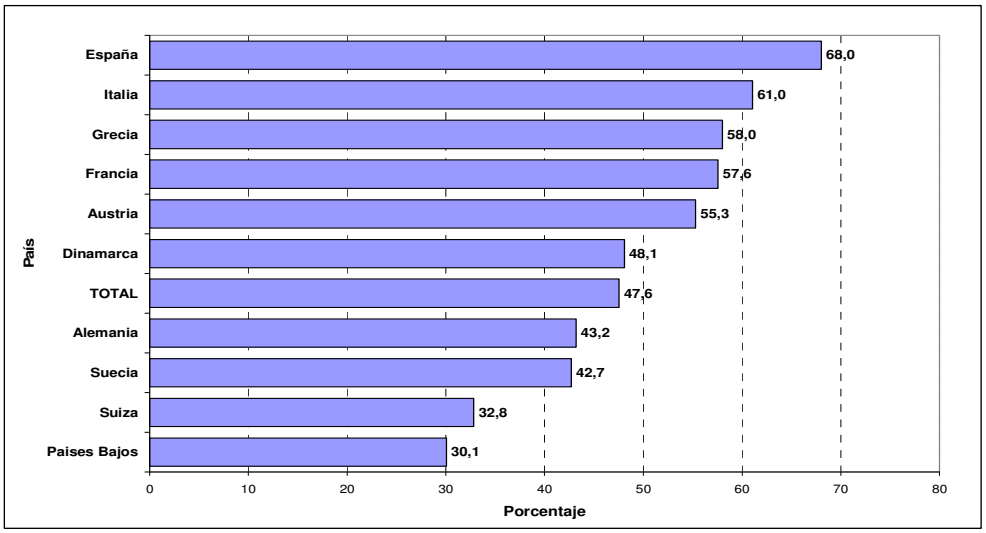
Gráfico 8. Edad media de retirada del mercado laboral según sexo en los países de la Unión Europea, 2004



Fuente: Elaboración propia a partir de datos de Sustainable Development Indicators (Eurostat).

Los datos procedentes de Eurostat parecen indicar que en Europa, y también en España, se está produciendo una inversión de la tendencia al abandono anticipado del mercado laboral que se había observado en el pasado. No obstante, hay que señalar que esos datos no se computan a los parados de edades próximas a la jubilación y que, en muchos casos, no volverán a ocupar un puesto de trabajo, en especial aquéllos de larga duración. Por otra parte, cuando se pregunta a los trabajadores de más edad, son una escasa minoría los que piensan que van a mantenerse activos más allá de los 65 años (menos del 10%) y, en cambio, son muchos los que desearían jubilarse lo antes posible, en particular los españoles (Gráfico 9).

Gráfico 9. Porcentaje de personas a las que les gustaría jubilarse tan pronto como puedan, 2004



Nota: Trabajadores de 50 a 59 años (Sin ponderar).
Fuente: The Survey of Health, Ageing and Retirement in Europe. Proyecto SHARE.

3.2. La población jubilada

El análisis de la población inactiva pone de manifiesto el alto grado de institucionalización que la jubilación tiene entre la población de 65 años y más. Según datos de la Encuesta de Condiciones de Vida de las Personas Mayores, 2006 (ECVM) que elabora el IMSERSO, sólo el 0,7% de los mayores de 65 años españoles sigue trabajando. La gran mayoría de ellos (el 63,6%) ha trabajado anteriormente y hoy está jubilado, particularmente los varones (el 98,3%). Entre las mujeres, la condición de jubiladas sólo alcanza al 38,3% del total, en las que la mayor parte tiene la condición de pensionista que no ha trabajado con anterioridad (35,6%) o la de ama de casa (24,6%).

La realidad de Castilla y León respecto a la situación laboral de las personas mayores podemos describirla teniendo en cuenta los resultados del estudio estadístico denominado "Situación económica y social de las personas mayores de Castilla y León. Año 2007" (JCYL) realizado con carácter específico para la población anciana de la región. Dicho estudio indica que apenas el 0,4% de las personas de 65 o más años que residen en dicha Comunidad Autónoma se encuentran actualmente en activo, mientras que el 61,1% del total son jubilados después de haber trabajado a lo largo de toda su vida. Los varones de más de 65 años son los que en mayor medida ostentan la condición de jubilados que han trabajado siempre (el 95,2% frente al 34,1% de las mujeres). Por su parte, el 45,8% de la población femenina nunca ha trabajado fuera del hogar (Tabla 8).

Tabla 8. Situación laboral de la población de 65 años y más según sexo en España y en Castilla y León, 2006-2007

	España			Castilla y León		
Ambos sexos	**Hombre**	**Mujer**	**Total**	**Hombre**	**Mujer**	**Total**
Trabajo remunerado por cuenta ajena o propia	1,1%	0,4%	0,7%	0,5%	0,2%	0,4%
Está jubilado/a y ha trabajado siempre	98,3%	38,3%	63,6%	95,2%	34,1%	61,1%
Nunca ha trabajado fuera de casa	0,1%	24,6%	14,3%	0,1%	45,8%	25,6%
Otras situaciones	0,5%	36,5%	21,3%	3,9%	19,6%	12,6%
Ns/Nc	0,0%	0,2%	0,1%	0,3%	0,3%	0,3%
Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

Fuente: Elaboración propia a partir de datos de ECVM-2006 (IMSERSO) y del estudio "Situación económica y social de las personas mayores de Castilla y León. 2007" (JCYL).

Pese a que los resultados de las dos fuentes estadísticas analizadas no sean totalmente homogéneos, se pueden obtener conclusiones importantes que se desprenden de su conjunto. En este sentido, se puede identificar la mayor presencia relativa en Castilla y León de las personas de 65 años y más que nunca han trabajado fuera de casa (25,6%), siendo éstas fundamentalmente de sexo femenino (45,8%), así como la menor proporción de mujeres jubiladas que han trabajado siempre (34,1%) y el inferior grado de participación en la actividad laboral de las personas mayores de la región (0,4%), todo ello en relación a la situación para el conjunto de España.

Ambos estudios también ponen de manifiesto cambios en la consideración de la jubilación como experiencia vital, de modo que predomina cada vez más una concepción positiva de la jubilación, entendida como una etapa de liberación del trabajo y de merecido descanso. En el contexto nacional, dicha valoración es más frecuente entre los varones (40,5%) que entre las mujeres (33,6%), lo que puede tener relación con la distribución de las tareas en el hogar según sexo, en la medida en que muchas mujeres no se liberan nunca de las tareas domésticas. Asimismo, la situación económica también influye de forma directa en una positiva valoración de la jubilación, sobre todo en el caso de los varones (43,4%). En las mujeres, en cambio, esta influencia es menos evidente, quizá porque la mayoría de ellas no son las que sustentan económicamente el hogar (33,3%). En este sentido, la relación particular de cada individuo con el mercado de trabajo constituye, normalmente, un fuerte condicionante de su nivel de ingresos posterior y, por tanto, de su situación económica durante esta etapa de la vida de las personas mayores.

En cuanto a Castilla y León, el estudio sobre la situación socioeconómica de las personas mayores de la región revela que el 58,1% de las actualmente jubiladas se muestran muy o bastante de acuerdo con respecto a que estaban deseando hacerlo, con una proporción algo mayor entre los varones (59,6%) que entre las mujeres (55,8%); el 70,9% lo está con la idea de que la jubilación presenta hoy en día muchas ventajas; y el 94,4% con que la de que la jubilación constituye un merecido descanso, sin que dentro del ámbito regional se detecten diferencias apreciables según sexo en este aspecto (Tabla 9).

Tabla 9. Grado de acuerdo de la población jubilada de 65 años y más con respecto a la jubilación según sexo en Castilla y León, 2007

	España		
Ambos sexos	**Hombre**	**Mujer**	**Total**
Estaba deseando jubilarse	59,6%	55,8%	58,1%
hoy en día la jubilación tiene muchas ventajas	71,8%	69,7%	70,9%
La jubilación constituye un merecido descanso	95%	93,6%	94,4%

Fuente: Elaboración propia a partir de datos del estudio "Situación económica y social de las personas mayores de Castilla y León. 2007" (JCYL).

4. CONCLUSIONES.

La Organización Mundial de la Salud sostiene que si un mayor número de personas disfrutaran de la oportunidad de tener un trabajo digno (apropiadamente remunerado, en entornos adecuados, protegido contra los riesgos) desde la juventud, se llegaría a la vejez siendo capaz de seguir participando en la fuerza laboral y, en consecuencia, toda la sociedad saldría beneficiada. De acuerdo con ello, debe de propiciarse un reconocimiento cada vez mayor de la necesidad de apoyar la contribución activa y productiva que las personas mayores puedan realizar, y de hecho realizan, a la sociedad (OMS, 2002). La consecución del objetivo anterior requiere conocer, previamente, la situación en la que se encuentran y los rasgos que caracterizan al colectivo de personas mayores en relación con la actividad laboral y su experiencia en la jubilación. Posteriormente, se podrán platear actuaciones encaminadas a incrementar la participación económica de los mayores y se podrá afrontar satisfactoriamente la realidad de la etapa vital que constituye la vejez. A continuación, se sintetizan algunas conclusiones extraídas del análisis

efectuado en los epígrafes anteriores dedicados a describir la participación en el mercado laboral y en la actividad económica de las personas mayores en Castilla y León.

La tasa de actividad de Castilla y León creció en 2007 hasta el 53,7% de la población de 16 y más años, sin embargo, sigue siendo una de las regiones españolas con menor tasa de actividad, estando bastante alejada de la media para España (el 58,9%). Las tasas de actividad de Castilla y León para los varones y las mujeres se situaron en ese año en el 64,0% y el 43,6% respectivamente, siendo también ambos valores inferiores a los registrados por el conjunto nacional (el 69,3% y el 48,9%, respectivamente).

La realidad de la participación laboral y en la actividad económica de las personas mayores muestra amplias diferencias por sexo y grupos de edad. Así, considerando el colectivo que integra a los trabajadores de 55 años y más de Castilla y León, la tasa de actividad regional se sitúa por debajo de la media nacional (el 15,30% frente al 16,78%, respectivamente) y, aunque dicha tasa ha experimentado un ligero crecimiento en los últimos años en paralelo a la evolución positiva del mercado laboral español, el aumento en la región ha sido de menor cuantía (el 0,6%) que el acontecido a escala nacional (el 0,8%). De hecho, la Comunidad Autónoma de Castilla y León registra actualmente una de las más bajas tasas de actividad entre los trabajadores de este tramo de edad de toda España (el 16,8% en 2007).

La experiencia de la vida cotidiana de las personas mayores (65 años y más) de Castilla y León se caracteriza fundamentalmente por su retirada de la actividad laboral, siendo una escasa minoría los que todavía se mantienen activos más allá de los 65 años (el 1,4%, frente al 2% de media en España).

Se observa que la retirada del mercado de trabajo se ha adelantado temporalmente a la edad de los 65 años, pues ya se constata un brusco descenso en las tasas de actividad entre los trabajadores de 55 a 64 años de la región, de modo que solo el 47% siguen siendo laboralmente activos en 2007. Este resultado se corrobora también en la distinción según sexo, siendo siempre las tasas de actividad femeninas muy inferiores a las masculinas y también siempre menores que los valores promedio a escala nacional.

La tasa de empleo regional fue del 49,8%, en 2007, frente a una tasa media del 54% en España. Por sexos, la tasa de empleo masculina fue del 61% en Castilla y León y la femenina del 38,9%, frente a unas tasas medias del 64,9% y del 43,6% en España.

Entre el colectivo laboral de 55 y más años, al igual que ocurría con las tasas de actividad, la tasa de empleo de Castilla y León la sitúa entre las regiones españolas con menor proporción de población ocupada de mayor edad (el 16,1% en 2007, frente a una media del 18,8% a escala nacional). Dentro de este tramo de edad, la tasa de empleo masculina de Castilla y León fue del 23,8% y la femenina del 9,5%, ambas se encuentran muy por debajo de los valores medios nacionales (el 27,4% y el 11,8%, respectivamente). No obstante, se aprecia un mayor dinamismo en el empleo que a nivel nacional entre las mujeres trabajadoras de más edad durante los tres últimos años.

Los servicios (con el 58,9%) y la industria (con el 15,9%) son las actividades que concentran a la mayoría de los ocupados dentro del grupo de trabajadores de más edad, si bien destaca el mayor peso proporcional de los empleos agrarios en relación con el contexto a escala nacional (13,6% y 7,8%, respectivamente) lo que resalta la relevancia que aún tienen las actividades agrarias entre los trabajadores ocupados de más edad de Castilla y León. Sobresale también la fuerte presencia de mujeres de más edad de la región ocupadas en el sector terciario (el 85%), que casi duplica la de los varones, y el mayor porcentaje de los trabajadores varones ocupados en el sector agrario de Castilla y León en comparación con el conjunto nacional (el 15,9% y 8,9% del total, respectivamente).

El colectivo de trabajadores ocupados de 55 años y más se caracteriza por mostrar una proporción muy baja de empleo asalariado (el 9,8%, frente al 9,2% en España), lo que sugiere que las personas que siguen trabajando son, fundamentalmente, trabajadores por cuenta propia con mayor capacidad para decidir cuándo retirarse del mercado laboral.

En 2007, la tasa de paro en Castilla y León fue inferior a la media nacional (el 7,2% y 8,3%, respectivamente), hecho que también se constata para los trabajadores de 55 y más años (el 4,1% en Castilla y León, frente al 5,6% de promedio en España) y la sitúa entre las regiones españolas

con menor nivel de paro en este grupo de edad. Las tasas de paro femeninas de la región superan siempre a las de los varones para todos los grupos de edad, siendo similares a las que se registran en media para España en el caso de las mujeres y algo más bajas entre los varones de la región.

Una comparación de las tasas de actividad, de ocupación y de paro de la población trabajadora de más edad (55 años y más) de Castilla y León permite apreciar que existe una amplia disparidad entre sexos pues las tasas de actividad y de ocupación de los varones superan el doble de las de las mujeres, mientras que cuando se trata de las tasas de paro ocurre todo lo contrario. Por su parte, si se comparan los valores regionales y nacionales de dichas tasas también se pone de manifiesto diferencias entre la situación de los trabajadores de 55 años y más en Castilla y León y en España, que se plasman en que las tasas de actividad, de ocupación y de paro regionales son siempre inferiores a las nacionales, distancia que se estrecha en el caso de la referida al desempleo.

El 98,6% de la población de 65 años y más de Castilla y León en 2007 se encuentra en situación de inactividad laboral, que es una proporción que supera ligeramente el promedio para el caso de España (98%). El total de varones inactivos de la región ha aumentado durante el período 2005-2007, afectando particularmente al grupo de edad de 55 a 64 años. En cuanto a las mujeres inactivas, éstas representan la práctica totalidad de las que tienen 65 años y más (el 99,3%) en 2007. Aunque existen datos que indican que en España, se está produciendo una inversión de la tendencia al abandono anticipado del mercado laboral que se había observado en el pasado, cuando se pregunta a los trabajadores de más edad, menos del 10% piensa mantenerse activo tras cumplir los 65 años y una amplia mayoría desearía jubilarse lo antes posible (el 68% en España).

Los datos más actualizados referidos al caso de Castilla y León, correspondientes al año 2008, indican que sólo el 0,4% de las personas de 65 o más años siguen activas, estando el 61,1% jubilado tras haber trabajado toda su vida. La inmensa mayoría de los varones ha trabajado siempre y de forma regular, sin embargo, entre las mujeres sólo la tercera parte ha tenido una vida laboral continuada (95,2% varones, frente al 34,1% de mujeres) y casi la mitad (el 45,8%) nunca ha trabajado fuera del hogar. En comparación con el conjunto de España, en Castilla y León existe una mayor presencia relativa de personas de 65 años y más que nunca han trabajado fuera de casa (el 25,6%), fundamentalmente mujeres (45,8%) y, en cambio, una menor proporción de mujeres jubiladas que han trabajado siempre (34,1%) y de actividad laboral de las personas mayores (el 0,4%).

Las personas mayores de Castilla y León conciben la jubilación como una fase más de la vida, socialmente legitimada y con una visión positiva. Así, el 58,1% de las personas actualmente jubiladas se muestran muy o bastante de acuerdo con respecto a que estaban deseando poder abandonar la actividad laboral (el 59,6% en varones y el 55,8% en mujeres); el 70,9% lo está con la idea de que la jubilación presenta hoy en día muchas ventajas; y el 94,4% con la opinión de que la jubilación constituye un descanso merecido.

Para finalizar, es preciso señalar que los anteriores resultados y conclusiones se centran únicamente en la participación de las personas mayores dentro del mercado laboral formal, por lo que no toman en consideración la valiosa contribución que estas personas llevan a cabo de manera informal (actividades económicas ocasionales, trabajos domésticos no remunerados, ayudas familiares en la atención y cuidado de nietos o familiares enfermos, etc.) y las muchas otras actividades que desarrollan de forma voluntaria en centros educativos, comunidades locales e instituciones no lucrativas, así como en empresas y en organizaciones políticas, entre otras.

Si de lo que se trata es de que, en la medida de lo posible, las personas mayores puedan seguir siendo independientes y activas a medida que envejecen, habrán de diseñarse programas específicos que fomenten los hábitos saludables y que mejoren la calidad de vida de este colectivo de personas de edad avanzada, pero también otros que faciliten su participación activa en la sociedad, no sólo con medidas tendentes a favorecer su continuidad en el mercado laboral más allá de la edad legal de jubilación, sino también a través de un creciente reconocimiento social y apoyo institucional de la labor que nuestras personas mayores desempeñan en el sector informal, en el seno de nuestras familias y de muchas organizaciones sociales. En este sentido, las posibilidades y ventajas que puede ofrecer el trabajo voluntario para aquéllas personas

mayores que lo deseen son muchas, no solo porque supone una contribución positiva para la sociedad en su conjunto, sino porque beneficia también a las propias personas mayores en el marco de sus relaciones sociales y en el ámbito de su bienestar personal.

5. REFERENCIAS BIBLIOGRÁFICAS.

CONSEJO ECONÓMICO Y SOCIAL DE CASTILLA Y LEÓN (2008): "Situación Económica y Social de Castilla y León en 2007". CES de Castilla y León. Valladolid.

CONSEJO ECONÓMICO Y SOCIAL DE CASTILLA Y LEÓN (2009): "Perspectivas del Envejecimiento Activo en Castilla y León". Informe a iniciativa propia 1/2009. CES de Castilla y León. Valladolid.

EUROSTAT (2007): "Sustainable Development Indicators". Oficina Estadística de la Unión Europea (Eurostat). Luxemburgo.

INE (Varios años): "Encuesta de Población Activa". Instituto Nacional de Estadística. Madrid.

IMSERSO (2007): "Las personas mayores en España. Informe 2006. Datos estadísticos Estatales y Comunidades Autónomas. Vol. I y II". IMSERSO. Madrid.

IMSERSO-CIS (2006): "Encuesta de Condiciones de Vida de las Personas Mayores" (ECVM). Estudio 2.647. IMSERSO. Madrid.

JUNTA DE CASTILLA Y LEÓN (2008): "Situación económica y social de las personas mayores de Castilla y León. Año 2007". Dirección General de Estadística. Consejería de Hacienda. Junta de Castilla y León. Valladolid.

MINISTERIO DE TRABAJO E INMIGRACIÓN (Varios años): Anuarios y Boletines de Estadísticas Laborales. MTIN. Madrid.

SISTEMAS DE APOIO À CRIAÇÃO DE NOVAS EMPRESAS NO CONCELHO DA COVILHÃ

MARIA JOSÉ AGUILAR MADEIRA SILVA
E-mail: msilva@ubi.pt
NECE – Núcleo de Estudos em Ciências Empresariais
Universidade da Beira Interior
RUTE MARIA FERNANDES ANTUNES
E-mail: rute.mf.antunes@sapo.pt
Universidade da Beira Interior

Resumo

A presente investigação visa identificar e analisar factores institucionais formais do meio envolvente que condicionam a criação de empresas industriais no Concelho da Covilhã. Na elaboração deste estudo considera-se os organismos, os serviços, os programas e as medidas de apoio à criação de empresas, tanto do ponto de vista da oferta, como do lado da procura por parte dos empresários. Os resultados obtidos, do ponto de vista da oferta, mostram que, no Concelho da Covilhã, existem muitos serviços e programas, pouco diversificados, que oferecem principalmente medidas não económicas e não financeiras. Do ponto de vista da procura, constatou-se que os empresários podem ter amplo conhecimento dos organismos, mas tal facto não implica que estes utilizem os serviços/programas oferecidos pelos organismos. Finalmente, são feitas recomendações tendo em vista uma melhor coordenação entre as diferentes instituições envolvidas, para que os programas e serviços oferecidos sejam mais eficaz no futuro.

Palavras-chave: Criação de Empresas, Empreendedorismo, Instituições, Teoria Económica Institucional.
Área Temática: Economia Urbana, Regional e Local

Abstract

The current investigation aims to identify and analyse the formal institutional factors of the environment that confines the creation of industrial enterprises in Covilhã County. The environment formal institutional is analyse in terms of the offer point of view (institutions, the services, the programs and support measures to enterprise creation), and the demand point of view, by entrepreneurs. The results, from the point of view of the services/programs offer, shows that, in Covilhã County, there are a lot of services and programs, but poorly diversified, that mainly offer non-economic and non-financial measures. From the point of view of demand, we noticed that entrepreneurs may have wide knowledge of the institutions, but that doesn't imply that they use their programs and services. Finally recommendations are made having in view a better coordination among the different institutions involved so that the programs they offer can be more effective in the future.

Key words: Entrepreneurship, Enterprise Creation, Institutions, Institutional Economic Theory
Thematic Area: Urban, Regional and Local Economics

1. INTRODUÇÃO

A criação de empresas é referida na literatura sobre empreendedorismo como um dos factores responsáveis pelo crescimento económico, criação de emprego, fomento de inovação, renovação da estrutura económica, e melhoria da competitividade global das regiões e dos países. Estes contributos fazem com que a criação de empresas assuma um papel de destaque e uma elevada importância não só ao nível socioeconómico, como também ao nível político.

Em Portugal, o interesse por esta temática, por parte das Administrações Públicas, tanto europeias como nacionais, têm estipulado medidas e programas de apoio para a criação de novas

empresas e para a promoção do espírito empresarial. Perante a presente conjuntura económica em que se encontra Portugal e face às actuais taxas de desemprego[1], que ultimamente têm vindo a aumentar, o fomento do empreendedorismo capaz de levar à criação de empresas, parece ser uma das medidas que pode dar o seu contributo para a minimização dos problemas económicos e sociais que têm assolado o país nos últimos anos. Assim, no contexto português revela-se fundamental analisar os factores que possam contribuir para a promoção da criação de empresas.

Tomando como referência as abordagens actuais sobre a temática, destacam-se a teoria económica institucional de Douglas North (1990), que proporciona um enquadramento teórico adequado à análise da criação de empresas. Este facto é corraborado Veciana (1999:25) ao considerar que a teoria económica institucional "actualmente proporciona o marco conceptual mais consistente e apropriado para o estudo da influência dos factores do meio envolvente na função empresarial e na criação de empresas". O estudo recentemente publicado por Veciana e Urbano (2008) vem, mais uma vez, reforçar a selecção deste enquadramento teórico pela adequação que este apresenta para o estudo dos factores que influenciam a criação empresas.

Constatada a importância da criação de empresas e dado que a análise no enquadramento institucional formal é um fenómeno ainda pouco estudado, sobretudo em Portugal, considerou-se oportuno, quer em termos académicos, quer em termos práticos, investigar esse fenómeno, através de uma região em concreto, nomeadamente, o Concelho da Covilhã.

Assim, o objectivo principal da investigação visa analisar a valorização dos organismos, serviços e medidas de apoio à criação de empresas por parte dois empresários, procurando analisar se os factores institucionais formais do meio envolvente condicionam a criação de empresas industriais no Concelho da Covilhã. A questão central de investigação que se coloca é a seguinte: quais são os factores institucionais formais, do meio envolvente, que são valorizados e utilizados pelos empresários e condicionam a criação de empresas industriais, no Concelho da Covilhã? Para dar resposta a esta questão formula-se a hipótese de investigação que será objecto de testes empíricos. Esta hipótese relaciona a valorização e a utilização dos serviços/programas dos próprios organismos, por parte dos empresários, com a actividade empreendedora no âmbito da criação de empresas.

Os resultados da investigação revelam-se importantes por duas razões. Por um lado permite conhecer o ponto de vista da oferta dos serviços/programas de apoio à criação de empresas. Por outro lado, visando o ponto de vista da procura, pretende-se de avaliar qual a valorização e utilização dos serviços/programas dos próprios organismos, por parte dos empresários.

O artigo está estruturado da seguinte forma: na secção dois efectua-se uma revisão da literatura relevante que estuda a relação entre os factores institucionais formais e a actividade empreendedora. Nessa mesma secção propõem-se um modelo conceptual e formula-se a hipótese de investigação. Na secção seguinte se descreve o desenho de investigação utilizado para testar a hipótese formulada. Na quarta secção apresentam-se e discutem-se os resultados. Finalmente, na quinta secção apresentam-se as considerações finais referentes à oferta e à procura institucional e sugerem-se futuras investigações a realizar sobre esta temática.

2. REVISÃO DA LITERATURA

Actualmente existe, cada vez mais, uma constante preocupação ao nível da sociedade e, principalmente, das administrações públicas, no que diz respeito a aspectos, tais como: o crescimento económico, a criação de novos postos de trabalho e inovação empresarial. O facto da criação de novas empresas se apresentar como uma via para a solução dos problemas de desemprego (Birch, 1979, 1987; Birley, 1987; Kirchhoff e Phillips, 1988, 1994; Storey, 1982, 1994; White e Reynolds, 1996), de desenvolvimento económico e regional (Kent, 1982; Sexton, 1986; Dubini, 1989; Storey, 1994; Wennekers e Thurik, 1999; Reynolds *et al.*, 2003; Acs *et al.*, 2004; Parker, 2004, entre muitos outros) e do fomento da inovação (Acs e Audretsch, 1988; Drucker, 1964, 1986; Hisrich e Peters, 1989) tem despertado, com maior intensidade, a

[1] *Segundo dados do INE, no 4º trimestre de 2008 situava-se nos 7,8% (http://www.ine.pt acedido em 17/02/2009)*

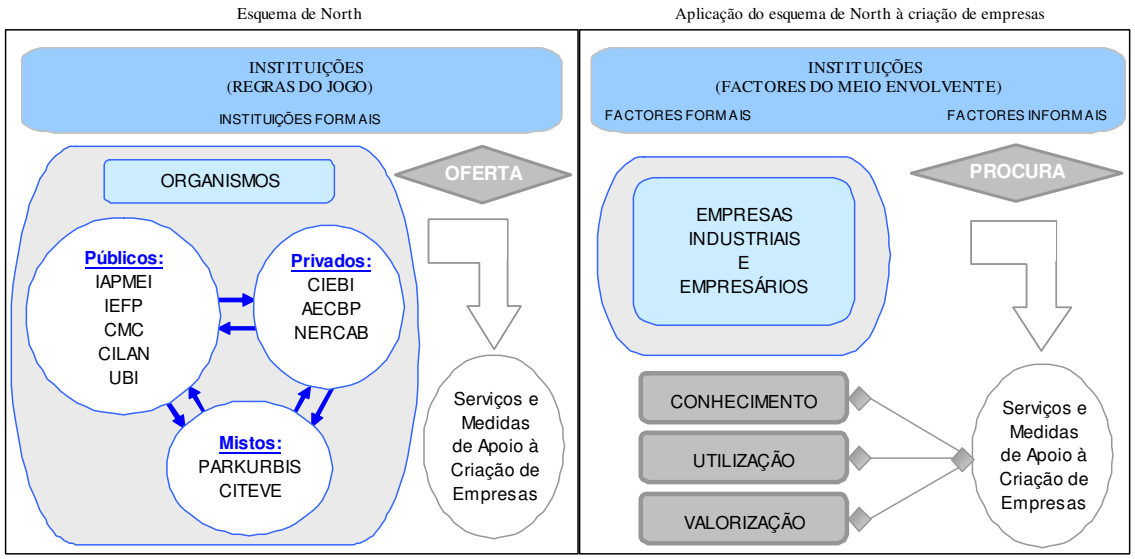
curiosidade, o interesse e a sensibilidade de várias entidades e membros da sociedade, em geral, para lidarem com este tema.

Pode referir-se concretamente, questões relacionadas com o fomento e estímulo de criação de novas empresas têm despertado o interesse de vários investigadores, que assumem que a utilização de mecanismos de apoio pode ter um efeito positivo tanto no número de empresas criadas como na sua taxa de sobrevivência (White e Reynolds, 1996; Storey, 2000; Boter e Lundström, 2001; Reynolds et al., 2003, Urbano, Diáz e Hernández 2007).

Dada a contribuição que as novas empresas têm no desenvolvimento económico e social de uma região ou de um país, o estudo dos factores que condicionam a criação de empresas revela-se de maior importância. Neste contexto, é de destacar algumas das investigações que contemplam tanto a análise do enquadramento formal como informal da criação de empresas (Aponte, 2002; Urbano, 2003; Díaz, 2003; Díaz, Urbano e Hérnandez, 2005). Existindo, também, outros estudos que se basearam na análise do enquadramento institucional formal que tiveram origem na Catalunha (Urbano, 2003); Porto Rico (Veciana, Aponte e Urbano, 2001; Aponte, 2002) e no Caribe Colombiano (Gómez, Veciana e Urbano, 2004) e que, focaram o seu estudo especialmente na identificação e descrição de instituições que tinham um papel mais significativo para a criação de empresas, assim como na análise de medidas de apoio oferecidas pelas instituições, o grau de conhecimento, utilização e valorização que os potenciais empresários tinham sobre os programas de apoio, e grau de adequação entre as medidas oferecidas e as reais necessidades dos futuros empresários.

Contudo, convém referir que apesar de existirem múltiplos enfoques teóricos que abordam a criação de empresas (Shane e Venkataraman, 2000; Gartner, 2001; Ucbasaran *et al.,* 2001; Aldrich e Martínez, 2001; Davidson e Wiklund, 2001, Trigo, 2003), a análise da intervenção das instituições no desenvolvimento económico e social, apenas aparece evidenciado na teoria institucional (North, 1990, 2005). Salientando que o objectivo desta investigação, seguindo o enfoque também adoptado por outros autores (Gibb, 1993; Smith, 1994; Malecki, 1997; Kolfsten e Scharberg, 1998; Young *et al.*, 1999; Lerner Harber, 2000; Urbano 2006), consiste na análise através da teoria económica institucional, do enquadramento formal da criação de empresas, identificando os organismos e programas de apoio existentes no Concelho da Covilhã. Deste modo, o foco da investigação centrou-se na análise do enquadramento institucional formal, de acordo com o modelo teórico proposto na figura seguinte:

Figura 1 - Aplicação do Modelo de North sobre as Instituições para a Análise dos Factores do Meio Envolvente que Condicionam a Criação de Empresas



Fonte: *Elaboração Própria*

Através da observação da Figura 1, de um lado, temos a Oferta (Instituições Formais), representada pelos Organismos (Públicos, Privados e Mistos) que oferecem serviços e medidas de apoio à criação de empresas e, do outro, as empresas industriais, em que os seus empresários procuram esses mesmos serviços e medidas de apoio à criação de empresas.

Relativamente aos tipos de medidas de apoio à criação de empresas, pode fazer-se a distinção entre medidas gerais e medidas específicas. As medidas de carácter geral são aquelas que se referem fundamentalmente a normas e regulamentos governamentais que afectam a criação de empresas. E, medidas específicas são aquelas que se consideram medidas de carácter não financeiro (informação/orientação, assessoria/acompanhamento, formação, seguimento, incubação) e de carácter financeiro (empréstimos, capital de risco e investimento a fundo perdido), como aliás, se pode verificar através da Tabela 1.

Tabela 1 – Tipologia das Medidas de Apoio à Criação de Empresas

MEDIDAS NÃO FINANCEIRAS	
Informação/ Orientação	Informação e encaminhamento do futuro empresário sobre todos os aspectos relacionados com a criação de empresas, sobre a forma jurídica da nova empresa e sobre as medidas específicas para a criação de uma empresa
Assessoria/ Acompanhamento	Assessoria e acompanhamento ao futuro empresário na elaboração do plano da empresa, na procura de financiamento e outras questões relacionadas com o processo de criação de empresa.
Seguimento	Seguimento da evolução das empresas que tenham recebido algum tipo de apoio por parte de algum Organismo, aquando do seu processo de criação.
Formação	Cursos de formação, sessões de esclarecimento, seminários e conferências sobre as temáticas empreendedorismo e criação de empresas.
Incubação	Oportunidade dos novos empresários se poderem instalar nos chamados Centros de Empresas, Centros Tecnológicos, Parques Tecnológicos e Zonas Industriais, sob condições mais favoráveis que no mercado.
Outras medidas	Prémios e concursos para fomentar a criação de empresas
MEDIDAS FINANCEIRAS	
Empréstimos	Linhas de crédito para os novos empresários, a partir de acordos celebrados entre administrações públicas e entidades financeiras, os quais permitem obter empréstimos com condições mais favoráveis.
Capital de Risco	Actividade financeira que consiste na tomada de uma participação, a médio/longo prazo, geralmente minoritária, no capital social de uma empresa, assegurando suporte financeiro ao seu desenvolvimento.
Investimento a Fundo Perdido	Investimento realizado sem expectativa de retorno do montante investido. Este tipo de investimento é realizado geralmente pela Administração Pública no sentido de melhorar as condições de existência de determinados sectores e regiões.

Fonte: *Elaboração Própria*

Assim, ao nível do Concelho da Covilhã, as estruturas de serviços/medidas que poderão constituir importantes bases de apoio à actividade empresarial estão evidenciadas na Tabela 2.

Tabela 2 – Organismos com maior Influência na Criação de Empresas no Concelho da Covilhã

		TIPO DE ORGANISMOS									
		ORGANISMOS PÚBLICOS					ORGANISMOS PRIVADOS			ORGANISMOS MISTOS	
	MEDIDAS	IAPMEI	IEFP	CMC	CILAN	UBI	CIEBI-BIC	AECBP	NERCAB	PARKURBIS	CITEVE
FINANCEIRAS	Informação/ Orientação	X	X	X		X	X	X	X	X	X
	Assessoria/ Acompanhamento	X	X	X		X	X	X	X	X	X
	Seguimento	X	X	X			X			X	
	Formação		X		X	X	X	X	X		X
	Incubação de Micro e PME's			X		X	X			X	
	Outra	Centro Empresas				Concursos de Ideias; Planos de Negócios				Bolsa de Ideias; Concursos	Centro Tecnológico
FINANCEIRA	Empréstimos	X									
	Capital de Risco	X									
	Outra		Investimento a Fundo Perdido								

Fonte: *Elaboração Própria*

Da análise do ponto de vista da oferta dos organismos, serviços/programas de apoio à criação de empresas, pode dizer-se que, no Concelho da Covilhã existem várias instituições e programas, pouco diversificados, que oferecem principalmente medidas não financeiras, que se baseiam fundamentalmente em serviços de informação/orientação, assessoria/acompanhamento, o que leva, muita das vezes, a uma duplicação de recursos e de esforços, por parte dos organismos da região, implicando, por sua vez, uma descoordenação ao nível da procura (empresário).

Consoante o grau de valorização que o próprio empresário tenha relativamente a tais medidas, oferecidas pelas Instituições Formais, assim irá agir e reagir perante o meio envolvente, o que influenciará, de certa forma, a criação de empresas e, consequentemente, o desenvolvimento económico.

Através da análise do modelo teórico proposto e das relações conceptuais que dele derivam, bem como da revisão da literatura, formula-se uma hipótese de investigação a testar. Nesta hipótese visa dar resposta à seguinte questão: será que quanto maior for o grau de conhecimento do organismo, maior será o grau de utilização do serviço/programas à criação de empresas?

Assim, pretende-se testar o nível de dependência existente entre as variáveis valorização/utilização. Esta hipótese visa apurar se todos os empresários que utilizaram os serviços/programas para a criação de empresas oferecidos pelos organismos de apoio, valorizaram positivamente a prestação de tais serviços, atribuindo elevada importância à sua utilização.

H_1: A valorização que os empresários atribuem aos serviços/programas oferecidos pelos organismos de apoio à criação de empresas está positivamente relacionada com a utilização das mesmas medidas.

A hipótese teórica de investigação exposta será testada empiricamente, tendo em conta os dados obtidos no inquérito efectuado às empresas, conforme se apresenta nas secções seguintes.

3. DESENHO DA INVESTIGAÇÃO

Esta investigação baseia-se no estudo de duas populações. Do lado da oferta, os organismos de apoio à criação de empresas existentes no Concelho da Covilhã, em que o método de recolha de informação utilizado foi a entrevista pessoal real a peritos dos referidos organismos. O tipo de estudo foi qualitativo e o procedimento amostral não aleatório, uma vez que se baseou nas instituições mais relevantes de apoio à criação de empresas, através de fontes secundárias de informações relacionadas com cada organismo (documentos internos, folhetos informativos, acesso à página web do próprio organismo). Quanto à análise estatística e tratamento informático dos dados, relativamente à oferta, procedeu-se a uma análise descritiva dos mesmos.

Do lado da procura, as empresas sedeadas no Concelho da Covilhã, nomeadamente as Micro, Pequenas e Médias Empresas do sector industrial (CAE 10 a 33), em que no método de recolha da informação se recorreu ao Questionário Pessoal entregue em mão aos empresários da referida região. O tipo de estudo foi quantitativo e o procedimento amostral aleatório, em que a base de dados foi construída pela investigadora com a colaboração de alguns Presidentes de Junta de Freguesia do Concelho da Covilhã e pelos carteiros do CDP - Centro de Distribuição Postal da Covilhã. Em síntese, todo o trabalho de investigação desenvolvido permitiu no final, ter acesso aos dados expostos na Tabela 3.

Tabela 3 – Resumo da Ficha Técnica da Investigação Desenvolvida

	ENTREVISTA PESSOAL A PERITOS DE ORGANISMOS DE APOIO À CRIAÇÃO DE EMPRESAS	QUESTIONÁRIO PESSOAL REALIZADO A EMPRESÁRIOS
Universo (Sector de Actividade)	Instituições mais relevantes de apoio à criação de empresas	Micro, Pequenas e Médias Empresas do sector industrial (CAE 10 a 33)
Âmbito (Área Geográfrica)	Concelho da Covilhã	Concelho da Covilhã
Método de Recolha da Informação	Entrevista Pessoal	Questionário Pessoal entregue em mão
Recolha de Dados	Através de fontes secundárias de informações relacionadas com cada instituição (documentos internos, folhetos informativos, acesso à página web da própria instituição)	Base de Dados construída pela própria investigadora com a colaboração de alguns presidentes de Junta de Freguesia do Concelho da Covilhã e pelos Carteiros do Centro de Distribuição Postal (CDP) da Covilhã
Procedimento Amostral	Não Aleatório	Aleatório
Data de Realização do Trabalho de Campo	Abril a Maio de 2008	Maio a Junho de 2008
Tamanho da População	10 Instituições	165 Empresas
Tamanho da Amostra	-------- (*)	59 Empresas
Medidas de Controlo	-------- (*)	Realização de 2 pré-teste
Taxa de Resposta	-------- (*)	35,80%
Análise Estatística e Tratamento Informático de Dados	Análise descritiva dos dados (Microsoft Excel)	Análise descritiva dos dados e Inferência Estatística (SPSS 16.0)

-------- (*) Neste tipo de procedimento amostral não é possível estimar

Fonte: *Elaboração Própria*

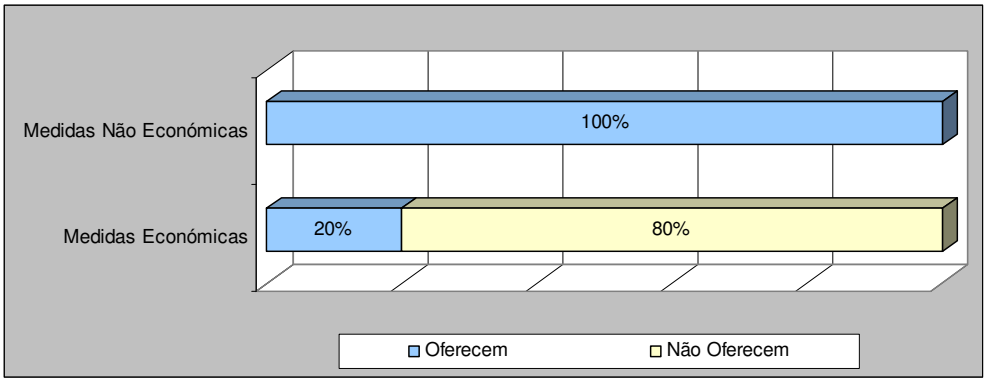
A análise estatística e tratamento informático dos dados, fez-se recorrendo a uma análise descritiva dos dados e inferência estatística, utilizando o programa estatístico SPSS, Versão 16.0.

4. ANÁLISE E DISCUSSÃO DE RESULTADOS

Na presente secção pretende-se efectuar o estudo da criação de empresas industriais e a influência que os organismos de apoio à actividade empreendedora, existentes no Concelho da Covilhã, podem exercer sobre as mesmas, a partir da informação obtida.

Da análise dos dados dos representantes dos organismos da amostra, foram extraídas todas as informações cedidas e procedeu-se neste ponto a uma explanação dos seus resultados, através da construção de seguintes figuras.

Figura 2 – Organismos que Oferecem Medidas de Apoio Financeiras e Não Financeiras



Como mostra a figura anterior, apenas 2 dos 10 organismos (20%) que foram entrevistados oferecem algum tipo de medidas financeira de apoio à criação de empresas, enquanto que os mesmos, e todos os outros oferecem algum tipo de medidas não financeiras (100%). O tipo de medidas oferecido pelos vários organismos está repartido conforme a divisão reflectida nas figuras seguintes.

Figura 3 – Tipos de Medidas Não Financeiras Oferecidas pelos Organismos de Apoio

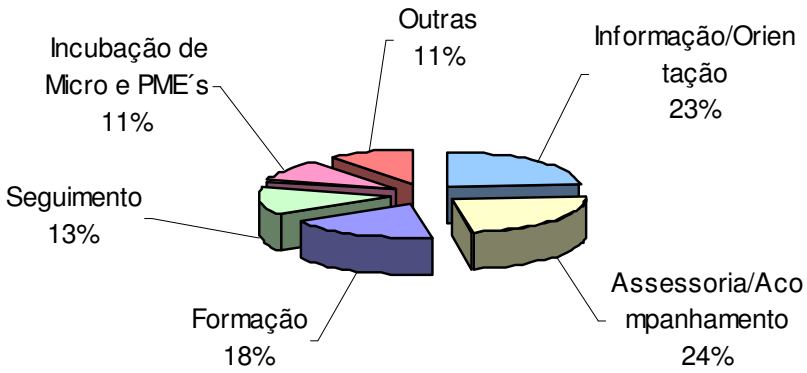
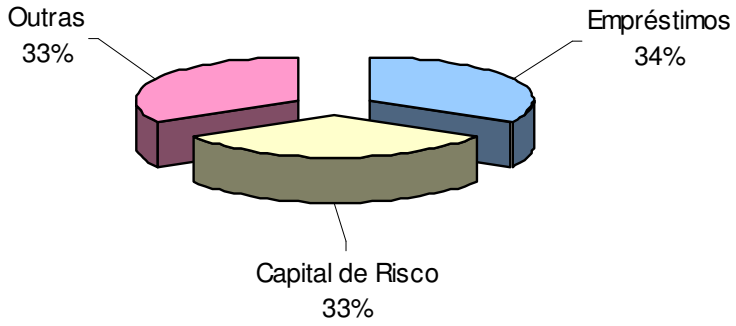


Figura 4 – Tipos de Medidas Financeiras Oferecidas pelos Organismos de Apoio



Relativamente a valorização que os empresários atribuem, aos organismos e programas de apoio à criação de empresas em estudo, demonstra que, em média, os organismos mais valorizados são o IAPMEI (37,5%) e o IEFP (37,5%). O menos valorizado pelos empresários inquiridos foi o CIEBI (10,8%). Relativamente aos organismos mais valorizados, destacam-se

com índices de valorização mais elevados os serviços/programas de informação/orientação no NERCAB (31,3%) e na AECBP (20,8%), e no âmbito do programa de formação é valorizado o CITEVE (18,8%).

Pela análise da Tabela 4 é possível constatar que os serviços/programas financeiros têm uma menor valorização relativamente a serviços/programas não financeiros, como seja a informação/orientação. No IAPMEI todos os serviços/programas não financeiros têm maior valorização atribuída pelos empresários do que o programa FINICIA.

Tabela 4. – Valorização dos Organismos e programas de apoio à criação de Empresas

Serviços/Programas	*Valorização dos Serviços/Programas para a Criação de Empresas(%)*									
	IAPMEI	*IEFP*	*CMC*	*CILAN*	*UBI*	*CIEBI*	*AECBP*	*NERCAB*	*PARKURBIS*	*CITEVE*
Programas não Económicos										
Informação/Orientação	39,6	45,8	16,7		18,8	12,5	8,3	14,6	14,6	10,4
Assessoria/Acompanhamento	37,5	33,3	16,7		14,6	10,4	8,3	14,6	14,6	12,5
Seguimento	37,5	33,3	10,4			10,4			14,6	
Formação				18,8	12,5	10,4	20,8	31,3		18,8
Incubação de micro e PME's			10,4		12,5	10,4			14,6	
GAPI-CITEVE										12,5
Programas Económicos										
Programa FINICIA	35,4									
ILE's		39,6								
ACPE		35,4								

Fonte: *Elaboração própria*

A quarta hipótese procura testar a existência de uma relação entre a variável "valorização" e a variável "utilização dos serviços/programas de apoio à criação de empresas" – *H1: A valorização que os empresários atribuem aos serviços/programas oferecidos pelos organismos de apoio à criação de empresas está positivamente relacionada com a utilização das mesmas medidas.*

Os resultados obtidos e apresentados na Tabela 5, foram feitos através da análise da correlação entre as variáveis "importância da utilização de serviços/programas para a criação de empresas" e "importância da valorização dos programas utilizados", no qual se verifica que existe uma correlação positiva forte (medida pelo coeficiente de Spearman) e estatísticamente significativa ao nível de 0,01 entre a utilização que os empresários fazem dos serviços/programas de apoio à criação de empresas e a valorização que atribuem após usufruir desses benefícios. Pelo que se poderá concluir que os empresários ao utilizarem a oferta disponibilizada pelos organismos, valorizam bastante os serviços/programas utilizados no apoio à criação de empresas.

Tabela 5 – Correlação entre a importância de utilização e valorização dos serviços/programas para a criação de empresas segundo o Organismo

Utiliza/Valoriza	*Coeficiente de Correlação Spearman*
IAPMEI	0,904 (**)
IEFP	0,884 (**)
CMC	0,871 (**)
CILAN	0,996 (**)
UBI	0,998 (**)
CIEBI	1,000 (**)
AECBP	0,999 (**)
NERCAB	0,955 (**)
PARKUBIS	0,998 (**)
CITEVE	1,000 (**)

(**) Correlação é significativa ao nível de 0.01

Na Tabela 6 estão registados os valores descritivos do grau de concordância dos empresários inquiridos face à adequação das medidas de apoio existentes.

Os valores obtidos sugerem que mais de 50% dos empresários inquiridos apresentam índices de concordância elevados com o facto das medidas de apoio: colocarem demasiados entraves burocráticos para obtê-las; não deixarem os empresários actuar com liberdade; não serem suficientemente conhecidas; e não se promoverem de forma eficiente. Ainda, 50% dos empresários manifestam um grau de discordância relativamente ao facto de as medidas de apoio para a criação de empresas actuais serem adequadas às necessidades dos empresários e serem acertadas e úteis.

Quadro 5.21. – Adequação das Medidas de Apoio Existentes

Medidas de Apoio	*Mediana*
Acertadas e úteis	2,50
Adequadas às necessidades dos empresários	2,00
Incidem sobre o êxito do processo de criação de uma empresa	3,00
São insuficientes	3,00
Existem demasiados entraves burocráticos para conseguir obtê-las	4,50
Seria melhor menos ajudas oficiais e deixar ao empresário mais liberdade de actuação	4,00
Não são suficientemente conhecidas	4,00
Não se promovem de forma eficiente	4,00
Não se utilizam demasiado	3,50

Fonte: *SPSS*

Um índice de concordância face às medidas de apoio existentes foi calculado. Começou-se por colocar todas as afirmações relativas às medidas de apoio na negativa e inverter a pontuação (isto é, foram invertidos os primeiros três itens). Somando o nível de importância atribuído a cada medida por cada individuo obteve-se um *score* final de concordância que teoricamente pode variar entre 9 e 45 valores. Valores elevados deste *score* sugerem a não adequação das medidas existentes.

Tabela 7 – *Score* final de adequação das medidas de apoio existentes

N	Válido	58
	Omisso	1
Média		32,7593
Mediana		33,0000
Desvio padrão		5,32681
Mínimo		15,00
Máximo		44,00
Percentis	25	30,0000
	50	33,0000
	75	36,0000

Fonte: *SPSS*

De acordo com a Tabela 7 é possível verificar que o valor mínimo de pontuação obtido foi 15 e o valor máximo foi 44. Em média, o índice de discordância face às medidas de apoio existentes (32,7) foi elevado (acima da média empírica 27), o que denota uma desadequação das medidas às reais necessidades dos empresários na criação de empresas.

5. CONSIDERAÇÕES FINAIS

A presente investigação analisou os factores institucionais formais do meio envolvente que condicionam a criação de empresas industriais no Concelho da Covilhã, considerando, na elaboração deste estudo, os organismos, os serviços e as medidas de apoio à criação de

empresas. O presente estudo seleccionou teoria económica institucional como o enquadramento teórico que se apresentou mais adequado ao estudo dos factores que influenciam a criação empresas.

A análise dos factores institucionais formais do meio envolvente que condicionam a criação de empresas é elaborada tanto do ponto de vista da oferta (organismos, medidas e serviços de apoio), como do lado da procura de programas e serviços por parte dos empresários.

Do ponto de vista da oferta a análise que foi efectuada, a partir da amostra dos representantes dos organismos de apoio à criação de empresas, podem extrair-se que no Concelho da Covilhã existem muitas instituições que oferecem principalmente medidas não financeiros pouco diversificadas, que se baseiam fundamentalmente em serviços de informação/orientação, assessoria/acompanhamento, o que leva, muita das vezes, a uma duplicação de recursos e de esforços, por parte dos organismos da região, implicando, por sua vez, uma descoordenação ao nível da procura (empresário). De acordo com a análise dos dados efectuada neste investigação, considera-se que os serviços/programas que os organismos existentes no Concelho da Covilhã têm para oferecer aos empresários não são elaborados com base em pressupostos específicos para a criação de empresas e adequados à região na qual a empresa se insere.

Do ponto de vista da procura, da análise que foi efectuada a partir dos dados da procura institucional por parte dos empresários, podem extrair-se as seguintes conclusões: (1) no que se refere à utilização atribuída pela amostra de empresários, no IAPMEI, os serviços/programas financeiros são mais valorizados do que os serviços/programas não financeiros. Ao contrário do IEFP, em que os empresários utilizam em maior número programas não financeiros, em detrimento de programas financeiros. (2) em termos de valorização, em média, os serviços/programas não financeiros (37,8%) são mais valorizados pelos empresários covilhanenses, do que os serviços/programas financeiros (36,8%). Destacam-se os serviços/programas de informação/orientação com índices de valorização mais elevados do IEFP (45,8%) e IAPMEI (39,6%), como o de formação do NERCAB (31,3%).

Os empresários que utilizam os serviços/programas de apoio à criação de empresas atribuem um elevado nível de valorização a esta mesma utilização. Pode deduzir-se que, o empresário ao utilizar algum tipo de serviço/programa oferecido por algum dos organismos, ficou bastante satisfeito, considerando significativa e importante a sua utilização, valorizando muito positivamente os benefícios que daí advieram.

Os resultados da investigação revelam a sua importância, dado que por um lado permite conhecer o ponto de vista da oferta dos serviços/programas de apoio à criação de empresas e, por outro lado, permite analisar o ponto de vista da procura por parte dos empresários.

A investigação efectuada sugere que haja uma coordenação entre os vários organismos visando melhorar a oferta/utilização destas instituições no futuro. Neste âmbito os organismos poderão tomar as seguintes medidas propostas, nomeadamente: (i) trocar informação de modo a evitar duplicação de programas iguais; (ii) redefinir os seus campos de acção por forma a que diferentes organismos oferecem programas diferentes e complementares; (iii) criar redes visando a potenciar a ligação em sistema para evitar desperdício de recursos e fomentar a melhor utilização dos serviços/programas oferecidos por esses mesmos organismos.

Contudo, esta investigação abordar apenas o estudo dos factores formais relativamente à criação de empresas, o que se considera uma limitação da própria investigação, deste modo, considera-se que a incorporação de factores informais iria ampliar o trabalho realizado. Propõem-se para futuras investigações realizar este mesmo estudo, noutros concelhos do País, com o objectivo, de num futuro próximo, se poderem fazer análises comparativas relativamente a Portugal, e entre Portugal e outros países.

BIBLIOGRAFIA

ACS, Z.J e AUDRETSCH, D.B. (1988), "Innovation in Large and Small Firms: An Empirical Analysis", *The American Economic Review*, Vol. 78, 4, pp.678-690.

ACS, Z.J; ARENIUS, P.; HAY, M. e MINNITI, M. (2004): *Global Entrepreneurship Monitor. 2004 Executive Report,* Babson College and London Business School.

ALDRICH, H.E. e MARTÍNEZ, M.A. (2001): "Many are Called, but Few are Chosen: An Evolutionary Perspective for the Study of Entrepreneurship", *Entrepreneurship Theory and Practice,* Vol. 25, 4, pp.41-56.

APONTE, M (2002): "*Factores Condicionantes de la Creación de Empresas en Puerto Rico: Un Enfoque Institucional",* Tesis Doctoral, ES: Universidad Autónoma de Barcelona.

BIRCH, D. (1979): *The Job Generation Process*, Cambridge, MA.

BIRCH, D. (1987): *The Job Creation in America*, The Free Press, New York.

BIRLEY, S. (1987): "New Ventures and Employment Growth", *Journal of Business Venturing*, Vol. 2, nº 2, pp.155-165.

BOTER, H. e LUNDSTRÖM, A. (2001): "Support Services to SMEs. Analysis of support Actors and Company Characteristics", Conference proceedings, Vol.2, Paper presented at RENT XV 2001.

CAE (2008): *Classificação das Actividades Económicas – Rev. 3,* Decreto-Lei nº 381/2007 de 14 de Novembro de 2007.

DAVIDSON, P. e WIKLUND, J. (2001): "Levels of Analysis in Entrepreneurship Research: Current Research Pratice and Suggestions for the Future", *Entrepreneurship Theory and Practice*, Vol. 25, 4, pp.81-99.

DÍAZ, C. (2003): "*La Creación de Empresas en Extremadura – Un Análisis Institucional",* Tesis Doctoral*,* ES: Universidad de Extremadura.

DÍAZ, C.; URBANO, D. e HERNÁNDEZ, R. (2005): "Teoria Económica Institucional y Creación de Empresas"*, Investigaciones Europeas de Dirección y Economia de la Empresa*, Vol. 11, 3, pp.209-230.

DRUCKER, P. (1964): *Managing for Results,* Harper and Row, New York.

DRUCKER, P. (1986): La *Innovation y el Empresário Innovador. La Pratica y los Principios,* Edhasa, Barcelona.

DUBINI, P. (1989): "The Influence of Motivation and Environment on Business Starts-Ups. Some Hints for Public Policies", *Journal of Business Venturing*, Vol. 4, 1, pp. 11-26.

GARTNER, W.B. (2001): "Is There na Elephant in Entrepreneurship? Blind Assumptions in Theory Development"*, Entrepreneurship Theory and Pratice*, Vol. 25, 4.

GIBB, A.A. (1993): "Key Factors in the Design of Policy support for the Small and Medium Enterprise (SME) Development Process: An Overview"*, Entrepreneurship and Regional Development*, Vol. 5, 1, pp.1-24.

GÓMEZ, L.; VECIANA, J.M. e URBANO, D. (2004): "Medidas de Apoyo a la Creación de Empresas en el Caribe Colombiano. Análisis de la Oferta y la Demanda de Servicios"*, Investigación y Desarrollo,* Vol. 12, 2: pp.372-395.

HISRICH, R. e PETERS, M. (1989): "Entrepreneurship: Starting, Developing and Managing a New Enterprise", Cap.1, Richard D. Irwin, Inc., Boston, Pp.3-23.

KENT, C.A. (1982): "Entrepreneurship in Economic Development", In Kent, C.; Sexton, D. e Vesper, K. (eds): *Encyclopaedia of Entrepreneurship*, Cap. 12, Prentice Hall, Inc. Englewood Cliffs, N.J., pp.237-256.

KIRCHHOFF, B.A. e PHILLIPS, D.B. (1988): "The Effect of Firm Formation and Growth on Job Creation in United States", *Journal of Business Venturing*, Vol. 3, 4, pp.261-272.

KLOFSTEN, M. e SCHARBERG, C. (1998): "Teaching and Training Small Businesses: Attitudes Among the Organisers", Paper for the 10th Nordic Small Business Research Conference, Váxjó University, Váxjó, Sweden, June, 14-16.

LERNER, M. e HABER, S. (2000): "Performance Factors of Small Tourism Ventures: The Interface of Tourism, Entrepreneurship and the Environment"*, Journal of Business Venture*, Vol. 16, pp.77-100.

MALECKI, E.J. (1997): "Entrepreneurs, Networks and Economic Development: A Review of Recent Research"*, Advances in Entrepreneurship, Firm Emergence and Growth*, Vol. 3, pp. 57-118.

NORTH, D.C. (1990): *Institutions, Institutional Change and Economic Performance,* Cambridge University Press: Cambridge (UK).

NORTH, D.C. (2005): "Understanding the Process of Economic Change", Princeton, University Press.

PARKER, S.C. (2004): *The Economics of Self-Employment and Entrepreneurship,* United Kingdom.

REYNOLDS, P.; HAY, M. e CAMP, R.M. (2003): *Global Entrepreneurship Monitor: 2003 Executive Report*, Babson College, Kauffman Center for Entrepreneurial Leadership; London School Business (eds.). London.

SEXTON, D.L. (1986): "Role of Entrepreneurship in Economic Development", em Hisrich, R.D. (1986): *Entrepreneurship, Intrapreneurship and Venture Capital*, Cap.2, pp.27-39.

SHANE, S. e VENKATARAMAN, S. (2000): "The Promise of Entrepreneurship as a field of Research", *Academy of Management Review*, Vol. 25, 1, pp.217-226.

SMITH, F.W. (1994): "Venture Capital", In Hisrich, R.D. e Peters, M.P. *Entrepreneurship. Starting, Developing, and Managing a New Enterprise*, Cap. 10, pp.266-296, Routledge: London.

STOREY, D.J. (1982): *Entrepreneurship and the New Firm*, Croom Helm: London.
STOREY, D.J. (1994): *Understanding the Small Business Sector*, London/New York Routledge.
STOREY, D.J. (2000): "Six Steps to Heaven: Evaluating the Impact f Public Policies to support Small Business in Developed Countries", In D. Sexton e H. Landstrom (eds.). *Handbook of Entrepreneurship*, pp.176-193, Oxford: Blackwell.
TRIGO, V. (2003): *Entre o Estado e o Mercado, Empreendedorismo e a Condição do Empresário na China*, Ad Litteram, Lisboa.
UCBASARAN, D.; WESTHEAD, P. e WRIGHT, M. (2001): "The Focus of Entrepreneurial Research: Contextual and Process Issues", *Entrepreneurship Theory and Practice*, Vol. 25, 4, pp.57-80.
URBANO, D. (2003): "*Factores Condicionantes de la Creación de Empresas en Catalunya: Un Enfoque Institucional*", Tesis Doctoral, ES: Universitad Autónoma de Barcelona: Barcelona.
URBANO, D. (2006): *La Creación de Empresas en Catalunya: Organismos de Apoyo y Actitudes hacia la Actividad Emprendedora*, Colección de Estúdios *CIDEM*, Generalitat de Catalunya: Barcelona.
URBANO, D; Díaz, J.C; e Hernández, R. (2007): "Evolución y Princípios de la Teoria Económica Institucional. Una Propuesta de Aplicación para el Análisis de los Factores Condicionantes de La Creación de Empresas", *Investigaciones Europeas de Dirección y Economia de la Empresa*, Vol. 13, 2, pp. 183-198.
VECIANA, J.M. (1999): "Creación de empresas como Programa de Investigación Científica", *Revista Europea de Dirección y Economia de la Empresa*, Vol. 8, 3; pp.11-36.
VECIANA, J.M. e URBANO, D. (2008): "The Institutional Approach to Entrepreneurship Research. Introduction", *The International Entrepreneurship and Management Journal*, Vol. 4, 4, pp.365-379.
VECIANA, J.M.; APONTE, M. e URBANO, D. (2001): "Instituciones y Medidas de Apoyo a la Creación de Empresas: Un Estúdio Empírico comparativo entre Cataluña y Puerto Rico", 3º Seminário Hispano-Luso de Economia Empresarial, Trujillo: Cáceres.
WENNEKERS, S. and THURIK, R. (1999): "Linking Entrepreneurship and Economic Growth", *Small Business Economics*, Vol. 13, pp.27-55.
WHite, S.B. e REYNOLDS, P.D. (1996): "Government Programs and High Growth New Firms", *Frontiers of Entrepreneurship Research*, Babson College.
YOUNG, M.; WYMAN, S.M. e BRENNER, C.T. (1999): "Assessment of Small Business Perception of Needed Information and Assistance", *Journal of Business and Entrepreneurship*, Vol. 11, 1, pp.99-105.

MEDIDA DE LA ESPECIALIZACIÓN ECONÓMICA DE LAS PROVINCIAS ESPAÑOLAS A TRAVÉS DEL USO DEL SUELO URBANO

MIGUEL ÁNGEL FAJARDO CALDERA
E-mail: fajardo@unex.es
Departamento de Economía Aplicada
Universidad de Extremadura
MIGUEL ÁNGEL FAJARDO BULLÓN
E-mail: mafajardo@mcx
Diplomado comercial del estado
PABLO ALONSO GONZÁLEZ
E-mail: pablo.alonsog@uah.es
Departamento de Estadística, Estructura Económica y O.E.I.
Universidad de Alcalá
IRENE ALBARRÁN LOZANO
E-mail: ialbarra@est-econ.uc3m.es
Departamento de Estadística
Universidad Carlos III de Madrid

Resumen

El objetivo de este trabajo es analizar la estructura productiva provincial, medida a través de la tipología de bienes inmuebles urbanos que posee la provincia, así como de la calidad de estos bienes inmuebles medida por su valoración catastral y su eficiencia económica a través de indicadores económicos. Para realizar este análisis, los autores han considerado los porcentajes de la distribución del suelo total entre algunos tipos de usos, como el industrial, cultural, religioso, comercial, oficinas y edificación singular. Tomando estos valores como variables y utilizando como metodología de trabajo las técnicas del Análisis de Componentes Principales y los Modelos de Ecuaciones Estructurales, que nos permita reducir este conjunto de variables a unas pocas (componentes factoriales) y las relaciones entre ellas.

Palabras clave: Uso del suelo urbano. Bienes inmuebles. Análisis de Componentes Principales (ACP).
Área Temática: Economía Urbana, regional y local

Abstract

The objective of this work is to analyze the provincial economic activity and your quality by using the use of urbanizeable urban ground. In order to make this analysis the authors have considered some variables as the percentage of the distribution of the total ground between industrial, cultural, religious, commercial, offices and singular construction uses. From this variables and using the Principal Components Analysis (PCA) technique, the authors reduce this set of variables to two factors that explain the economic behaviour and your quality of the Spanish provinces.

Keywords: Use of the urban ground. Real estate. Principal Components Analysis (PCA).
Thematic Area: Urban, Regional and Local Economics.

1.- INTRODUCCIÓN.

Las dos últimas décadas han sido testigos de una evolución en la que el Catastro Inmobiliario se ha convertido en una gran infraestructura de información territorial disponible para todas las Administraciones públicas, fedatarios, empresas y ciudadanos en general, y así lo han ido entendiendo todo tipo de usuarios públicos o privados que han ido utilizándolo según sus legítimas necesidades.

Por otro lado, el catastro viene siendo utilizado por investigadores de muchas áreas de conocimiento, como fuente de información primaria para sus trabajos, o como hipótesis en sus propias investigaciones, para ello baste un simple vistazo a las publicaciones que emplean catastro como palabra clave en sus trabajos, o el dato de 61 tesis doctorales relacionadas con el catastro, leídas en España en los últimos 25 años.

Actualmente, para conocer la riqueza patrimonial del territorio del país, tanto por lo que se refiere a su suelo rústico como al urbano, la más completa fuente de información es la proporcionada por la Dirección General del Catastro. La estadística del Catastro, incluida dentro del Plan Estadístico Nacional y difundida a través de Internet y de sus publicaciones anuales sobre los Catastros Urbano y Rústico, recoge todos los datos necesarios para obtener un certero conocimiento de cuál es este patrimonio, así como de su valor administrativo. Todos estos datos se proporcionan a nivel de municipio, incluyendo todos los integrantes del ámbito nacional, a excepción de los correspondientes a las Diputaciones Forales de Álava y Vizcaya y, respecto del Catastro de Parcelas, de todo el País Vasco y Navarra.

En cuanto al Catastro Urbano, la Dirección General del Catastro confecciona dos tipos de publicaciones: estadísticas básicas por municipios y estadística nacional de parcelas catastrales. Las primeras incluye el total de unidades inmobiliarias, su valor catastral agregado y el valor catastral medio de cada municipio; el segundo, por su parte, permite conocer el número de solares y parcelas edificadas existentes en un determinado municipio, así como la distribución según la superficie del suelo de ambos tipos de inmuebles y una clasificación por tipos de propiedad de las parcelas edificadas.

El objetivo de este trabajo es analizar la especialización provincial del uso del suelo urbano a través de sus variables características y tipologías de usos residenciales, comerciales, industriales, equipamientos, tanto para santuarios, iglesias, industrias, almacenes, centros cívicos y comerciales, espectáculos, e instalaciones de seguridad.

El conocimiento de la distribución del tipo de uso del territorio de una región, país, provincia o municipio de los bienes inmuebles es un instrumento eficaz, no sólo como un instrumento de gestión territorial, sino también como un indicador de la especialización de su actividad económica, como recogemos en este trabajo.

Así, se considera *suelo de naturaleza urbana* el clasificado por el planeamiento urbanístico como urbano. También los terrenos que tengan la consideración de urbanizables y aquel que reúna las características del artículo 8 de la Ley 6/1998, de 13 de abril, sobre Régimen del Suelo y Valoraciones y por último aquéllos en que puedan ejercitarse facultades urbanísticas equivalentes a los anteriores según legislación autonómica.

Dentro del suelo urbano, entendemos por *bien inmueble* a la parcela o porción de suelo de una misma naturaleza, enclavada en un término municipal y cerrada por una línea poligonal que delimita, a tales efectos, el ámbito espacial del derecho de propiedad de un propietario o de varios pro indiviso, y en su caso, las construcciones emplazadas en dicho ámbito, cualquiera que sea su dueño, y con independencia de otros derechos que recaigan sobre el inmueble.

Los tipos de uso del suelo urbano que hemos considerado en este artículo han sido ***industrial, oficinas, comercial, cultural, religioso y edificación especial*,** y a partir de la distribución (en porcentaje) del total del suelo urbano en los tipos de uso de los bienes inmuebles, hemos construido las variables que especificamos a continuación:

Uso industrial de bienes inmuebles (% sobre el total del suelo): Se entiende por uso industrial las naves de fabricación y almacenamiento (fabricación en una planta, en varias plantas y almacenamiento), garajes y aparcamientos y servicios de transporte (estaciones de servicio y estaciones).

Uso oficinas de bienes inmuebles (% sobre el total del suelo): corresponde a edificios exclusivos (oficinas unitarias y múltiples), edificio mixto (unido a vivienda o a industria) y Banca y Seguros (en edificio exclusivo o en mixto).

Uso comercial de bienes inmuebles (% sobre el total del suelo)**:** comprende comercios en edificio mixto (locales comerciales y talleres, galerías comerciales); comercios en edificio exclusivo (en una planta o varias plantas); mercados y supermercados (mercados, hipermercados y supermercados).

Uso cultural de bienes inmuebles (% sobre el total del suelo)**:** Comprende aquellos edificios culturales con residencias (internados, colegios mayores), culturales sin residencia (Facultades, colegios, escuelas), bibliotecas y museos).

Uso religioso de bienes inmuebles (% sobre el total del suelo)**:** Comprende el conjunto de actividades relacionadas con la celebración de los diferentes cultos (iglesias y capillas) y el alojamiento de los miembros de las comunidades religiosas (conventos y centros parroquiales).

Uso edificios singulares de bienes inmuebles (% sobre el total del suelo)**:** Son aquellos de carácter histórico-artísticos (monumentales, ambientales o típicos); de carácter oficial (administrativos, representativos) y de carácter especial (penitenciarios, militares y varios; obras de urbanización interior; campings, campos de golf; jardinería; silos y depósitos para sólidos; depósitos líquidos; depósitos de gases).

2.- DATOS.

Para alcanzar el objetivo propuesto de analizar la especialización económica, hemos considerado un estudio transversal de los datos al considerar como referencia temporal el año 2006. Los valores de las variables en estudio corresponden al porcentaje del suelo urbano que se ha dedicado a bienes cuyo uso esta categorizado por industrial, comercial, oficinas, cultural, religioso, edificación singular de las provincias españolas.

Para analizar la calidad de esa especialización económica hemos considerado los valores catastrales totales (construcción y suelo) medios provinciales de los bienes inmuebles de cada una de las tipologías de suelo consideradas como variables. Dado que no disponemos de valores catastrales medios provinciales del suelo y de la construcción de los bienes inmuebles asignados a cada tipología del suelo, hemos utilizado como una aproximación los valores catastrales medios provinciales para todos los bienes inmuebles del suelo y de la construcción de dichos bienes.

Los datos se han obtenido de las estadísticas disponibles del Catastro Inmobiliario español correspondientes al año 2006.

3. METODOLOGÍA.

Para realizar este estudio vamos a utilizar *el análisis de componentes principales,* técnica que fue propuesta por Karl Pearson como parte del análisis factorial. El objetivo de esta técnica consiste en representar medidas numéricas de varias variables en un espacio de pocas dimensiones (componentes principales) de tal forma que la pérdida de información sea mínima. Los nuevos componentes principales o factores serán una combinación lineal de las variables originales, y además serán independientes entre sí.

Supongamos pues que tenemos un vector $X = (X_1, X_{2,}, ..., X_p)$ de dimensión p, distribuidas $N(\mu,\Sigma)$. Para que un análisis de componentes principales tenga sentido es necesario que existan altas correlaciones entre las variables, ya que esto es indicativo de que existe información redundante y, por tanto, pocos factores explicarán gran parte de la variabilidad total.

Si analizamos las correlaciones entre las variables en estudio, podemos observar que estas son en general suficientemente altas y significativas, como se expresa en la siguiente matriz de correlación dada en el cuadro 1.

Cuadro 1
Matriz de correlaciones

Variables	comercial	cultural	industrial	oficinas	Ed. Singular	religioso
comercial	1	-,393(**)	-,707(**)	,719(**)	-,544(**)	-,377(**)
Cultural	-,393(**)	1	,314(*)	-,023	,295(*)	,449(**)
Industrial	-,707(**)	,314(*)	1	-,570(**)	,612(**)	,369(**)
Oficinas	,719(**)	-,023	-,570(**)	1	-,373(**)	-,141
ed. Singular	-,544(**)	,295(*)	,612(**)	-,373(**)	1	,347(*)
Religioso	-,377(**)	,449(**)	,369(**)	-,141	,347(*)	1

Fuente: Área de Estadística. Dirección General del Catastro. Ministerio de Economía y Hacienda.

Analíticamente, podemos comprobar el grado de correlación con la siguiente prueba o test: **Test de esfericidad de Bartlett.**

Para ello es necesario analizar la normalidad de las variables. A través de la prueba de Kolmogorov-Smirnov comprobamos que las variables siguen una distribución normal con un nivel de significación del 1%, como podemos comprobar en el siguiente cuadro 2:

Cuadro 2
Prueba de normalidad de Kolmogorov-Smirnov.

Variables		comercial	cultural	industrial	oficinas	ed. singular	religioso
N		48	48	48	48	48	48
Parámetros normales (a,b)	Media	,03121	,00144	,08839	,00572	,00074	,00206
	Desviación típica	,012639	,000363	,063128	,002359	,000424	,001717
Z de Kolmogorov-Smirnov		,430	,457	1,535	,622	1,140	1,311
Sig. asintót. (bilateral)		,993	,985	,018	,833	,148	,064

Fuente: Área de Estadística. Dirección General del Catastro. Ministerio de Economía y Hacienda.

Una vez analizada la normalidad de las variables, contrastamos el test de Bartlett. Este test contrasta la hipótesis $\mathbf{H_0}$ de que la matriz de correlaciones es una matriz identidad (incorrelación lineal entre las variables). Si, como resultado del contraste, no pudiésemos rechazar esta $\mathbf{H_0}$, y el tamaño de la muestra fuese razonablemente grande, deberíamos reconsiderar la realización de un **ACP**, ya que las variables no están correlacionadas.

El estadístico de contraste del test de Bartlett es: $\chi^2_{15} = -\,(\,48 - 1 - (2*6 + 5)/6\,)\,\ln|\,0.071\,| = 116.800$, con un *p-value* = 0.000, lo que nos permite rechazar la hipótesis Ho.

Para el cálculo de los autovalores, consideramos que Σ es una matriz simétrica y definida positiva, podemos encontrar una matriz L, tal que $\Sigma = L\Delta L^{-1}$, donde Λ es una matriz diagonal formada por los valores propios de Σ que denotamos por $\lambda_1, \lambda_2, ..., \lambda_p$.

Es posible ordenar, de acuerdo con sus magnitudes, los valores propios de Σ de tal manera que podemos reorganizar las columnas de la matriz L, que pueden estar formadas por vectores propios normalizados, es decir, perpendiculares entre sí y de longitud igual a la unidad. Al vector propio asociado con el mayor valor propio se le denomina primera dirección principal, al segundo valor propio en magnitud se le denomina segunda dirección principal, y así sucesivamente.

A cada combinación lineal $F_i = u_{i1}X_1 + u_{i2}X_2 + u_{i3}X_3 + ... + u_{ip}X_p$ de las variables originales centradas se les denomina componente principal.

Por otra parte, la traza de Σ, por ser la suma de las varianzas de las variables originales X, recibe el nombre de *varianza total*, y se tiene que:

$Traza(\Sigma) = Traza(L\Sigma L^{-1}) = \sum_{i=1}^{p} \lambda_i$, siendo la $Var(F_i) = \lambda_i \;\; i = 1, 2, ...p$. y que

$\text{cov}(F_i, F_j) = 0$

Por tanto, la varianza total es igual a la suma de los valores propios de Σ e igual a la suma de las varianzas de las componentes principales. Además, las componentes principales son variables aleatorias incorrelacionadas entre si y sí además todas las variables son normales, entonces las componente principales son normales.

Basándonos en lo anterior, la elección de los factores se realiza de tal forma que el primero recoja la mayor proporción posible de la variabilidad original; el segundo factor debe recoger la máxima variabilidad posible no recogida por el primero, y así sucesivamente. Del total de factores se elegirán aquéllos que recojan el porcentaje de variabilidad que se considere suficiente. A éstos se les denominará componentes principales. En la siguiente tabla se exponen los valores de los autovalores y la proporción de varianza explicada.

Tabla 3

Porcentaje de la varianza total explicada por los autovalores.

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3,164	52,741	52,741	3,164	52,741	52,741	2,653	44,210	44,210
2	1,189	19,823	72,564	1,189	19,823	72,564	1,701	28,354	72,564
3	,593	9,876	82,440						
4	,555	9,258	91,698						
5	,318	5,293	96,990						
6	,181	3,010	100,000						

Fuente: Área de Estadística. Dirección General del Catastro. Ministerio de Economía y Hacienda.

Mediante el criterio de Káiser, que consiste en conservar aquellos factores cuyo autovalor asociado sea mayor que 1, tenemos la siguiente selección de 2 componentes principales, que explican un 72.56% de la varianza total, como podemos observar en la tabla 3.

Una vez seleccionados los componentes principales, se representan en forma de matriz. Cada elemento de ésta representa los coeficientes factoriales de las variables (indican relaciones entre las correlaciones entre las variables y los componentes principales). La matriz tendrá tantas columnas como componentes principales y tantas filas como variables. Solo se ofrecen aquellas que tengan un valor superior al 0.5, como exponemos en la siguiente tabla 4.

Tabla 4

Matriz de componentes

	Componente	
Variables	F1	F2
Comercial	-,892	
Industrial	,858	
ed. Singular	,748	
Oficinas	-,694	,582
Religioso	,573	,572
Cultural	,515	,687

Fuente: Área de Estadística. Dirección General del Catastro. Ministerio de Economía y Hacienda.

Un aspecto clave en ACP es la interpretación de los factores, ya que ésta no viene dada a priori, sino que será deducida tras observar la relación de los factores con las variables iniciales (habrá, pues, que estudiar tanto el signo como la magnitud de las correlaciones).

Así, si con la solución inicial no se consiguiese una fácil interpretación de los factores, éstos pueden ser rotados de manera que cada una de las variables tenga una correlación lo más próxima a 1 con un factor y a 0 con el resto de factores. Los tipos de rotaciones más habituales son la ortogonal y la oblicua. La rotación ortogonal permite rotar los factores estimados inicialmente, de manera que se mantenga la incorrelación entre los mismos. El método más utilizado de rotación es el Varimax (Varianza máxima), ideado por Káiser, cuyo objetivo es minimizar el número de variables que tienen saturaciones altas en cada factor, el cual nos ofrece los siguientes resultados (para una mejor comprensión no se dan las cargas factoriales < 0.5) que se exponen en la tabla 5:

Tabla 5

Matriz de componentes rotados

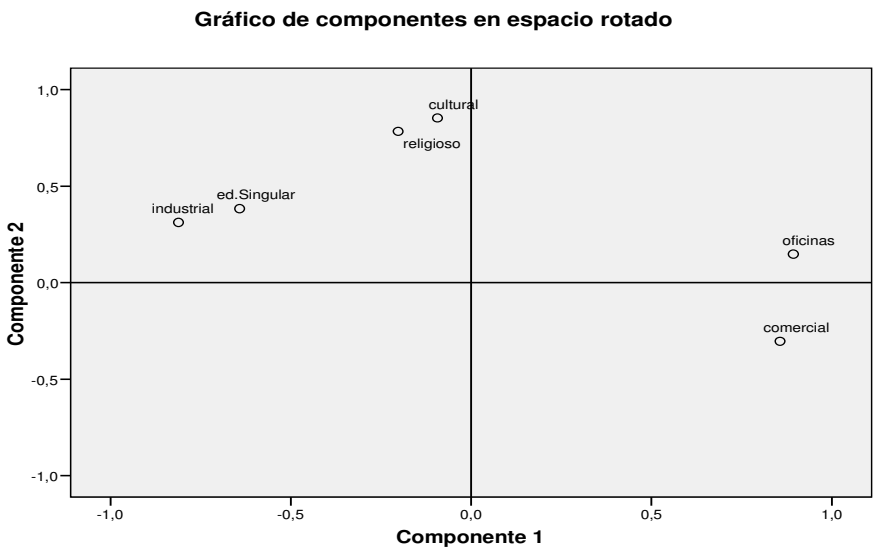
	Componente	
	1	2
oficinas	,893	
comercial	,856	
industrial	-,812	
ed. Singular	-,642	
cultural		,853
religioso		,784

Fuente: Área de Estadística. Dirección General del Catastro. Ministerio de Economía y Hacienda.

En la matriz de componentes rotados, observamos que las variables *oficinas* y *comercial* están correlacionadas positivamente con la componente 1 y las variables *industrial* y *edificación singula*r están correlacionadas negativamente con la componente 1, por lo que este factor nos indica que los individuos que presenten puntuaciones positivas altas, estarán muy especializados en aspectos *comerciales (oficinas y comercios)* y aquellos que presenten puntuaciones altas pero negativas están especializados en aspectos *industriales* (*industria e infraestructuras*). Por otra parte, puntuaciones altas y positivas en la componente 2, nos indica especialización en actividades *turísticas y culturales* (*cultura y religión*).

Una visión de la localización de las variables, según las cargas factoriales obtenidas en las dos componentes seleccionadas anteriormente, podemos observarla en el gráfico 1 en dos dimensiones de PC1 y PC2, donde se recoge la variabilidad de las observaciones, y si existe alguna que ofrezca un valor extrañamente alto o bajo en cada eje y que ofrecemos a continuación:

Gráfico1



Una vez alcanzada la solución factorial final, es decir, estimados los factores principales, es importante obtener una estimación de las puntuaciones de los sujetos (individuos u objetos) investigados para saber cuánto puntúan en cada factor, las cuales se obtienen de $X = FL^{-1}$, obteniendo la siguiente matriz de coeficientes, para el cálculo de las puntuaciones en las componentes, y tal que las estimaciones resultantes tienen una media cero y una varianza igual al cuadrado de la correlación múltiple entre las puntuaciones factoriales estimadas y los valores factoriales verdaderos; además, una representación gráfica de la localización de cada provincia en unos ejes cartesianos (componentes), nos permitirá comparar el comportamiento de las provincias según sus valores en las componentes 1 y 2 y que expondremos a continuación en la tabla 6 y en el gráfico 2:

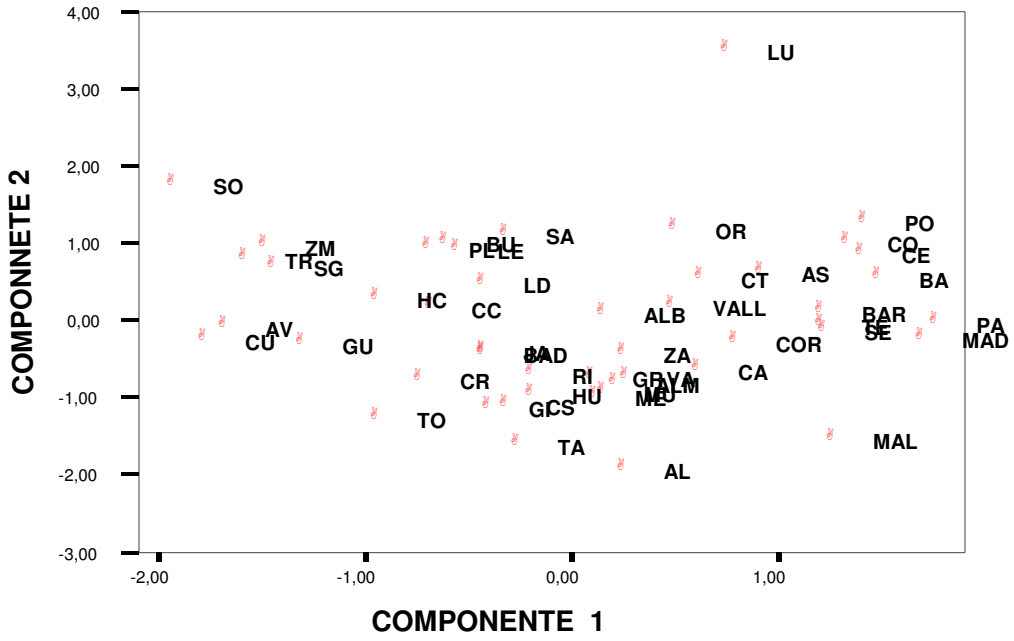
Tabla 6

Matriz de coeficientes para el cálculo de las puntuaciones en las componentes.

Variables	Componentes	
	1	2
Comercial	,317	-,017
Cultural	,154	,580
Industrial	-,295	,033
Oficinas	,438	,310
Religioso	,089	,506
ed. Singular	-,202	,123

Fuente: Área de Estadística. Dirección General del Catastro. Ministerio de Economía y Hacienda.

Gráfico 2



Por último exponemos en la siguiente tabla 7 los valores ordenados que toman las provincias en las componentes dadas:

Tabla 7

Puntuaciones factoriales de las provincias en las dos componentes.

Provincias	F1: Oficinas + comercial)-(industrial +ed. Singular)	Provincias	F2: Cultural + religioso.
Soria	-1,95590	Alicante	-1,93665
Cuenca	-1,79664	Tarragona	-1,63000
Ávila	-1,70206	Málaga	-1,54294
Teruel	-1,60129	Toledo	-1,29235
Zamora	-1,49978	Girona	-1,12643
Segovia	-1,46561	Castellón	-1,11894
Guadalajara	-1,32877	Melilla	-,98766
Toledo	-,96439	Huelva	-,96526
Huesca	-,96346	Murcia	-,94722
Ciudad Real	-,75475	Almería	-,83087
Palencia	-,70566	Ciudad Real	-,77515
Cáceres	-,70330	Valencia	-,75697
Burgos	-,62377	Granada	-,75644
León	-,56780	Rioja (La)	-,70652
Jaén	-,45309	Cádiz	-,65853
Lleida	-,45050	Zaragoza	-,44202
Badajoz	-,44929	Badajoz	-,42371
Girona	-,41661	Jaén	-,41755
Salamanca	-,34250	Guadalajara	-,32154
Castellón	-,33526	Córdoba	-,27878
Tarragona	-,28447	Cuenca	-,27116
Huelva	-,21257	Madrid	-,23640
Rioja (La)	-,20881	Sevilla	-,13219
Granada	,07761	Ávila	-,09039
Melilla	,09198	Tenerife	-,06921
Albacete	,12948	Palmas (Las)	-,05171
Murcia	,13446	Albacete	,06574
Almería	,19567	Barcelona	,10282
Alicante	,22799	Cáceres	,15083
Zaragoza	,23040	Valladolid	,16973
Valencia	,24252	Huesca	,26108
Valladolid	,46693	Lleida	,46371
Ourense	,48035	Balears (Illes)	,54192
Cádiz	,60007	Cantabria	,54805
Cantabria	,61487	Asturias	,61501
Lugo	,73652	Segovia	,68943
Córdoba	,78035	Teruel	,77536
Asturias	,90549	Ceuta	,86641
Barcelona	1,19104	León	,90586
Tenerife	1,19799	Palencia	,91736
Sevilla	1,21297	Zamora	,94916
Málaga	1,25305	Coruña (A)	1,00126
Coruña (A)	1,32034	Burgos	1,00296
Ceuta	1,39059	Salamanca	1,09320
Pontevedra	1,39772	Ourense	1,16259
Balears (Illes)	1,47225	Pontevedra	1,27201
Madrid	1,68344	Soria	1,74407
Palmas (Las)	1,75220	Lugo	3,46802

Fuente: Elaboración propia. Área de Estadística. Dirección General del Catastro. Ministerio de Economía y Hacienda.

4. CALIDAD DE LA ESPECIALIZACIÓN.

Una vez analizadas las provincias y obtenidos sus valores correspondientes en las dos componentes estimadas, vamos a considerar su calidad de la especialización (dada por los valores obtenidos en las componentes) por el valor medio total (suelo más construcción) catastral de los bienes inmuebles provinciales comparado con el valor medio total catastral de los bienes inmuebles español.

Para ello vamos a considerar primeramente a aquellas provincias que están más especializadas en bienes industriales e infraestructura que en bienes comerciales y oficinas, considerando los valores de las provincias en la componentes 1 (valores negativos especialización industrial, valores positivos especialización comercial) y vamos a cruzarla con los valores medios totales catastrales de los bienes inmuebles.

En segundo lugar, vamos a considerar de las provincias anteriores las valoraciones medias de construcción y de suelo de todos los bienes inmuebles yÿÿu comparación con el valor medioÿÿe construcción y suelo de todos los bienes inmuebles de España.

De las dos consideraciones anteriores obtenemos la siguiente tabla 8, donde se refleja la combinación de las magnitudes consideradas en los dos aparados anteriores y cuyos resultados exponemos a continuación en la tabla 8:

Tabla 8

Provincias Especializadas en suelo industrial + ed. Singular. (Valor Componente 1 < 0)

Valor catastral		***Edificación singular***	
		≤ media de España	***> media de España***
Industrial	***≤ media de España***	Ávila Badajoz Burgos Cáceres Ciudad Real Cuenca Guadalajara Huelva Huesca Jaén León Lleida Rioja (La) Palencia Salamanca Segovia Soria Teruel Toledo Zamora	
	> media de España	Castellón **Girona** Tarragona	

Las provincias señaladas en negrita tienen un valor catastral medio de construcción de todos los bienes inmuebles superior al valor medio correspondiente español.

Obsérvese en el cuadro anterior que excepto Castellón, Girona y Tarragona, las demás provincias especialistas en bienes inmuebles de naturaleza industrial y de edificación singular, tienen valores catastrales medios totales por debajo de la media española, en ambos tipos de edificación industrial y singular; lo que nos puede dar idea de la no alta calidad de sus bienes inmuebles. Además, excepto Girona, que tiene el valor medio de todos los bienes inmuebles en la construcción superior a la media española, mientras que todas las provincias especialistas en este tipo de construcción industrial y edificación singular están por debajo del valor catastral medio español del suelo y de la construcción de bienes inmuebles; por lo que nos confirma la poca calidad (en valor monetario) de los bienes inmuebles de las provincias especialistas en este tipo de inmuebles.

Si Consideramos ahora a aquellas provincias especializadas en bienes inmuebles comerciales y de oficinas, es decir, que han obtenido puntuaciones positivas en la componente 1, obtenemos los siguientes resultados al considerar los valores medios totales catastrales en comercios y oficinas comparados con los correspondientes al estado español, los cuales exponemos a continuación en la tabla 9:

Tabla 9
Provincias especializadas en suelo comercial y de oficinas

		Oficinas	
		≤ media de España	***> media de España***
Comercial	***≤ media de España***	*Alicante* Almería *Balears (Illes)* Cádiz Córdoba **Coruña (A)** Granada Lugo *Málaga Murcia* Ourense **Asturias** Pontevedra Cantabria *Valencia* Valladolid Melilla	
	> media de España	Albacete *Palmas (Las)* Tenerife *Sevilla* **Zaragoza** Ceuta	*Barcelona Madrid*

Las provincias no escritas ni en cursiva ni en negritas son aquellas que tienen un valor catastral por bienes inmuebles por debajo de la media española, tanto en valor catastral del suelo como en el valor catastral de la construcción. Las que están escritas en letras negritas indican que el valor promedio de todos los bienes en la construcción es mayor que el español. Aquellas que están escritas en letra cursiva indican que el valor promedio de todos los bienes es superior al español en construcción y suelo.

En el cuadro anterior, en la parte izquierda superior, podemos observar que las Islas Baleares, La Coruña, Málaga, Murcia, Asturias y Valencia tienen valores catastrales medios mas bajo que la media española en bienes de equipo comerciales y de oficinas, pero tienen un valor catastral del suelo o de la construcción superior a la media española, en contra de las otras provincias que tienen un valor catastral medio en suelo y construcción menor que la media española; por lo que nos indica la poca valoración de estos bienes inmuebles.

En la parte izquierda inferior de la tabla anterior, podemos observar que Las Palmas, Sevilla y Zaragoza tienen un valor catastral medio por bienes inmuebles comerciales superior a la media española, pero más bajo en oficinas y además superan el valor medio español en construcción o en suelo, por lo que nos ofrecen una valoración superior a las otras provincias que las acompañan.

Por último no podemos dejar de citar el caso de las provincias de Madrid y Barcelona, en las cuales el valor catastral medio de los bienes inmuebles comerciales y oficinas superan a la media española y además su valor catastral de bienes inmuebles es superior al español en valor medio del suelo y en valor medio de la construcción.

Si atendemos a las provincias especialistas en bienes inmuebles religiosos y culturales, tenemos que:

Provincias Especializadas en suelo cultural + religioso. (Valor Componente 2 > 0)

		Religioso	
		≤ media	***> media***
Cultural	***≤ media***	*Balears (Illes)* Burgos Cáceres **Coruña (LA)** Huesca León Lugo Ourense **Asturias** Palencia Pontevedra Cantabria Segovia Soria Teruel Zamora	Salamanca Valladolid
	> media		Albacete *Barcelona* Ceuta

Las provincias escritas sin negritas ni cursivas son aquellas que tienen un valor catastral por bienes inmuebles por debajo de la media española, tanto en valor catastral del suelo como en el valor catastral de la construcción. Aquellas que están escritas en negritas indican que el valor promedio de todos los bienes en la construcción es mayor que el español. Aquellas que están escritas en letras cursivas indican que el valor promedio de todos los bienes es superior al español en construcción y suelo.

Podemos observar en el cuadro anterior que las provincias de Las Islas Baleares, La Coruña y Asturias tienen un valor catastral medio por debajo de la media española en bienes inmuebles culturales y religiosos, pero toman un valor catastral medio superior en suelo o en construcción a las otras provincias situadas en el cuadrante superior izquierdo. Por el contrario, Salamanca y Valladolid tienen un valor medio inferior a la media española en bienes culturales pero no así en bienes religiosos, aunque el valor medio de todos los bienes es inferior a la media española en suelo y en construcción.

Por último, Albacete y Ceuta tienen valores medios catastrales en bienes culturales y religiosos que la media española al igual que Barcelona, pero sus valores catastrales medios de todos los bienes en suelo y construcción están por debajo de la media española, no ocurriendo de esta forma a Barcelona.

5.- CONCLUSIONES.

Del análisis realizado de las provincias españolas según la especialización otorgada por el peso que tiene en la tipología de uso del suelo urbano en industria, comercial, oficinas, cultura, religioso y edificación singular, destacamos las siguientes conclusiones:

1º. Que las provincias españolas que están especializadas en industria e infraestructura: Ávila Badajoz Burgos Cáceres Ciudad Real Cuenca Guadalajara Huelva Huesca Jaén León Lleida Rioja (La) Palencia Salamanca Segovia Soria Teruel Toledo Zamora tienen un valor catastral medio en bienes inmuebles industriales y de edificación singular por debajo de la media española respectivamente, lo que nos indica la propia calidad (monetaria) de estos bienes inmuebles. Las provincias de Castellón **Girona** y Tarragona, por el contrario tienen un valor catastral medio superior al valor español en los bienes inmuebles industriales, aunque no en la edificación singular; además Girona tiene en todos sus bienes inmuebles un valor de construcción superior al nacional.

2º. Las provincias especializadas en actividades comerciales: Alicante, Almería, Islas Baleares, Cádiz, Córdoba, Coruña (LA), Granada, Lugo, Málaga, Murcia, Ourense, Asturias, Pontevedra, Cantabria, Valencia, Valladolid y Melilla tienen valores medios catastrales que España en bienes inmuebles comerciales y oficinas. Además, La Coruña y Asturias tienen un valor catastral medio en todos los bienes inmuebles superior al español en la construcción y las provincias Alicante, Islas Baleares, Málaga, Murcia y Valencia tienen valores medios en construcción y suelo superiores al español al considerar todos los bienes inmuebles.

Por otra parte, las provincias también especializadas en actividades comerciales: Albacete, Las Palmas, Tenerife, Sevilla, Zaragoza y Ceuta tienen un valor catastral medio superior al español en bienes inmuebles comerciales y además Zaragoza presenta un valor medio de todos los bienes inmuebles superior al español en la construcción de dichos bienes. Las provincias de Las Palmas y Sevilla tienen un valor medio superior al español en construcción y suelo al considerar todos los bienes inmuebles.

Además, las provincias especializadas en actividades comerciales: Barcelona Madrid tienen valores catastrales medios superiores a España en bienes comerciales y oficinas y además superan a España en valor medio en construcción y suelo al considerar todos los bienes inmuebles.

3º. La provincias especializadas en bienes culturales y religiosos: Islas Baleares, Burgos, Cáceres, La Coruña, Huesca, León, Lugo, Ourense, Asturias, Palencia, Pontevedra, Cantabria, Segovia, Soria, Teruel y Zamora presenta valores medios catastrales por debajo del español en bienes inmuebles culturales y religiosos. Además, La Coruña y Asturias tienen valores catastrales medios de todos los bienes inmuebles en construcción superior al español y la

provincia Islas Baleares tienen valor catastral superior al español en construcción y suelo al considerar todos los bienes inmuebles.

Las provincias especializadas en bienes culturales y religiosos: Salamanca y Valladolid tienen un valor catastral medio más alto que el español en bienes culturales.

Por último, las provincias de Albacete, Barcelona y Ceuta presentan valores catastrales superiores al español en bienes culturales y religiosos respectivamente, siendo Barcelona la provincia que supera a España en valor medio en construcción y suelo al considerar todos los bienes inmuebles.

6. BIBLIOGRAFÍA.

Área de Estadística. Dirección General del Catastro. Ministerio de Economía y Hacienda.

Dillon W., M. Goldstein (1984) "Multivariate Análisis", John Wiley & Sons.

Escofier B., J. Pages (1992) "Análisis Factoriales Simples y Múltiples", Universidad del País Vasco, Bilbao.

Lebart, Morineau et Piron (1998) "Statistique Exploratoire Multidimensionnelle", Dunod. París.

http://mitago.net/Proyectos/provincias/mapa

PROGRAMAS FINANCEIROS EM ZONAS POBRES: MICROFINANCIAMENTO NO ALENTEJO

CONCEIÇÃO REGO[1]
E-mail: mcpr@uevora.pt
ISABEL VIEIRA
E-mail: impvv@uevora.pt
CARLOS VIEIRA
E-mail: cvieira@uevora.pt
Departamento de Economia & CEFAGE
Universidade de Évora

Resumo

A implementação e a sustentabilidade de actividades produtivas em regiões de baixa densidade demográfica e económica são desafios fundamentais neste tipo de territórios. Dada a fraca capacidade de atracção de empreendedores externos, o combate ao chamado 'ciclo vicioso da pobreza', através da promoção e dinamização de iniciativas criadoras de emprego e de riqueza, requer a existência de mecanismos de microfinanciamento compatíveis com as especificidades dos pequenos empresários locais. Neste estudo analisam-se os efeitos ao nível da promoção de iniciativas empresariais, no Alentejo, decorrentes do funcionamento de dois instrumentos financeiros direccionados para pequenos empresários: o programa FAME (dinamizado pela Associação de Desenvolvimento Regional do Alentejo) e o microcrédito (da responsabilidade da Associação Nacional de Direito ao Crédito). A análise comparativa destes programas permite a obtenção de conclusões com implicações políticas e práticas que podem ser úteis para potenciar os efeitos positivos do microfinanciamento na região Alentejo.

Classificação JEL: G21,R11, R12
Palavras-Chave: Desenvolvimento territorial, Microfinanciamento, Regiões de Baixa Densidade, Restrições de liquidez
Área temática: Economia Urbana, Regional e Local

Abstract

Regions of low demographic and economic density face major development challenges that greatly increase the struggle of local productive agents for subsistence. Given the relative unattractiveness of such territories to external entrepreneurs, strategies to combat poverty via activities capable of generating employment and to add value to local communities require the support of microfinance mechanisms, designed to encompass the specificities of local poor agents. In this study, an analysis of the effects promoted by two microfinance programs (Microcredit and program FAME), implemented in the Portuguese region of Alentejo, is developed. The comparative assessment of the two programs, in view of the region's poverty profile, has policy and practical implications that may be of use to enhance the positive effects of microfinance in this region.

JEL Classification: G21; R11; R12.
Key words: Territorial development, Microfinance, Low density regions, Liquidity constraints
Thematic Area: Local, regional and urban economy.

[1] Autor para correspondência: Conceição Rego, Departamento de Economia, Universidade de Évora, Largo dos Colegiais, 2, 7000-803 Évora, Portugal.

1. INTRODUÇÃO

O nível de sofisticação financeira dos países condiciona o seu desenvolvimento económico e social, pois as áreas mais desenvolvidas dispõem normalmente de maior diversidade de meios para afectação das poupanças ao financiamento das actividades produtivas. Assim, as empresas aí localizadas estão em situação de vantagem relativa, tendo à disposição muitas alternativas para a captação dos recursos financeiros essenciais à prossecução dos seus objectivos. Contudo, independentemente do nível de desenvolvimento já atingido, existem em todos os países pequenos empresários, ou potenciais empreendedores, que não têm acesso às fontes de intermediação financeira institucionais, por não preencherem os requisitos que estas estabelecem como condições mínimas para financiamento. A incapacidade de apresentação de fiadores e garantias formais, a ausência de um historial de crédito, baixos níveis de instrução, ou simplesmente porque os montantes pedidos são considerados demasiado baixos para que o sistema financeiro tradicional os considere interessantes e compatíveis com os custos operacionais de um empréstimo, são as justificações habituais. Além destes factores, em muitos países é também frequente a discriminação de género, que se traduz no afastamento efectivo das mulheres dos mecanismos formais de financiamento.

A exclusão do direito ao crédito condena os indivíduos a situações de pobreza permanente, por terem que desistir de planos para a implementação ou a consolidação de negócios próprios, ou ser obrigados a associar-se a redes informais de agentes que disponibilizam financiamento mas que, simultaneamente, os exploram e impedem o regular desenvolvimento dos seus projectos. Tradicionalmente, e em especial nos países que podem assegurar sistemas de protecção social eficientes, as reacções a situações de pobreza e exclusão materializam-se na atribuição de subsídios a quem não tem capacidade para satisfazer as necessidades básicas. Contudo do ponto de vista da eficiência e da sustentabilidade, pelo menos uma parte do dinheiro canalizado para subsidiar o desemprego em países desenvolvidos devia ser desviado para iniciativas de criação do próprio emprego e de promoção do empreendedorismo. Esta estratégia permitiria a criação de condições para reduzir gastos e aumentar receitas públicas, pois enquanto os desempregados recebem subsídios, os microempresários pagam impostos e geram o seu próprio rendimento. Trata-se de um verdadeiro 'Ovo de Colombo' com o qual até os políticos mais obstinados são obrigados a concordar. Porque razão não é então esta ideia mais explorada? Talvez por falta de 'espírito empreendedor' em sociedades que conseguiram implementar redes de segurança social bem sucedidas.

Tais redes não estão, no entanto, ao alcance da maioria dos países em desenvolvimento e, provavelmente por isso, o microcrédito surgiu precisamente num dos países mais pobres do mundo, o Bangladeche. Foi lá que, na década de 70 do século passado, um professor de economia chamado Muhammad Yunnus chegou à conclusão que, muitas vezes, para erradicar situações de pobreza basta garantir o acesso a montantes de dinheiro praticamente irrelevantes, mas fundamentais para financiar um projecto de negócio. O enorme sucesso da implementação de programas de microfinanciamento no Bangladeche popularizou o conceito de microcrédito. A ideia de conceder crédito a pessoas pobres, desde que apresentem planos de negócio sensatos, ou desenvolvam actividades potencialmente lucrativas, independentemente da sua história pessoal ou do montante necessário para tais fins, ultrapassou fronteiras e conquistou uma escala global. A sua consolidação universal aconteceu com a proclamação pelas Nações Unidas do ano de 2005 como Ano Internacional do Microcrédito e com a atribuição, em 2006, do Prémio Nobel da Paz a Muhammad Yunus e ao seu *Grameen Bank* (Banco da Aldeia). Nessa data já o número de instituições de microfinanciamento era de 3133 e o crédito por elas disponibilizando chegava a mais de 113 milhões de pessoas em todo o mundo (Daley-Harris, 2006).

A experiência do Bangladeche demonstra que o microcrédito pode ter resultados positivos nas circunstâncias mais difíceis. Mostra também que a sustentabilidade desta forma de concessão de crédito pode exigir a evolução das estruturas iniciais, para permitir capacidade de resposta a necessidades financeiras dos clientes, para além da concessão de um primeiro empréstimo. Vários projectos de microfinanciamento têm sido implementados em todo o mundo, em países desenvolvidos e em desenvolvimento, com resultados diferenciados. Os distintos contextos geográficos, económicos e sociais, bem como os conhecimentos adquiridos

com a experiência, têm conduzido à adopção de diversas estratégias, adaptadas às características específicas dos indivíduos alvo e do respectivo território.

Actualmente, as experiências de microfinanciamento são consideradas como elementos importantes, que devem ser tidos em conta pelos interessados na definição e implementação de políticas de desenvolvimento sustentado e de estratégias de erradicação da pobreza. Justifica-se a consideração do microfinanciamento como variável nas análises de desenvolvimento, de acordo com as teorias de desenvolvimento endógeno que realçam a necessidade de valorizar os recursos das regiões como instrumentos de promoção do seu desenvolvimento económico e social. De facto, não só os resultados dos projectos de microfinanciamento podem indiciar o nível de desenvolvimento já alcançado, como o perfil de cada região tem que ser tido em consideração nos processos de implementação destes programas financeiros.

Contudo, independentemente da área geográfica e do contexto socio-político em que são implementados, os projectos de microfinanciamento têm sempre como público-alvo a população pobre. Não se aborda obviamente da mesma forma o combate à pobreza em regiões desfavorecidas da OCDE ou em zonas remotas da Ásia, mas o facto comum a estas situações tão díspares é que em nenhuma delas os pobres podem recorrer às fontes institucionais para financiar iniciativas empreendedoras. Assim, se não existirem mecanismos que permitam o microfinanciamento, ou as sociedades recorrem ao subsídio e à caridade para ajudar os mais desfavorecidos, ou estes são deixadas à sua sorte e condenadas a existências de mendicidade

Portugal é particularmente interessante do ponto de vista da implementação de programas de microfinanciamento, pois apresenta características de desenvolvimento mistas. De facto, este é o país mais pobre da muito desenvolvida zona euro e, apesar de registar uma tendência de convergência desde a integração europeia em 1986, a aproximação da economia portuguesa aos níveis médios da União Europeia (UE) tem sido muito mais lenta e irregular do que seria desejável e expectável. Estas dificuldades de convergência estão provavelmente relacionadas com as assimetrias regionais que continuam a persistir, apesar do significativo esforço financeiro destinado à recuperação estrutural e à dinamização das áreas mais carenciadas.

De entre as regiões portuguesas, o Alentejo é a mais pobre, apresentando algumas características de subdesenvolvimento semelhantes às identificadas nas zonas mais desfavorecidas do mundo. Neste contexto, não é imediatamente óbvio se os programas de microfinanciamento implementados nesta região devem seguir os modelos adequados aos países desenvolvidos, ou se o subdesenvolvimento relativo da região faz dela um espaço onde podem ser melhor sucedidos modelos alternativos. O presente estudo procura responder a esta questão, recorrendo a análises comparativas de padrões de pobreza e programas de microfinanciamento implementados no Alentejo, à luz dos consensos actuais relativos aos principais determinantes do sucesso deste tipo de programas financeiros.

A análise compreende, além desta secção introdutória, 4 partes organizadas da seguinte forma: na secção 2 apresenta-se o conceito de microcrédito, caracterizam-se os principais modelos de microfinanciamento e desenvolve-se uma breve revisão da bibliografia relativa à aplicação dos diferentes modelos em países com distintos níveis de desenvolvimento; na secção 3 analisa-se o perfil de pobreza da população portuguesa, contrastando-o com o típico de áreas mais e menos desenvolvidas, e enquadra-se a situação particular dos pobres residentes no Alentejo; a secção 4 compreende a descrição de dois programas de microfinanciamento implementados no Alentejo, o microcrédito e o programa FAME, e a análise comparativa dos mesmos; na secção 5 são enunciadas as principais conclusões do estudo e as suas implicações políticas e práticas, sendo sugeridas estratégias para potenciar os efeitos positivos do microfinanciamento na região Alentejo.

2. MODELOS DE MICROFINANCIAMENTO

O microfinanciamento tem origem nos programas de microcrédito que, de forma simples e genérica, pode ser definido como uma provisão de crédito do sector informal direccionado para pessoas pobres e pequenas empresas sem acesso a outras fontes de financiamento. Na prática, actualmente, o microcrédito assume uma enorme variedade de modelos, adaptados às

características específicas dos locais onde é implementado ou dos respectivos promotores. Até mesmo a expressão 'sector informal' pode ser questionada, se utilizada para designar empréstimos oriundos de fontes não registadas, pois é frequente a intermediação de instituições formais e 'semi- formais', como por exemplo o *Grameen Bank*. Os objectivos dos programas de microfinanciamento são também muito diversos, podendo compreender desde a simples oferta de pequenos empréstimos a empresários excluídos dos processos de intermediação financeira institucionais, até programas mais ambiciosos de desenvolvimento integrado, que compreendem também formação e assistência técnica.

Os vários modelos distinguem-se pela principal fonte de financiamento, que pode ser externa e, intermediada por exemplo por ONGs ou directamente canalizada por instituições financeiras convencionais, com programas especificamente direccionados para o sector informal, maioritariamente em bancos públicos ou impostos ou subsidiados pelo Estado. Em alternativa, o financiamento pode ser angariado internamente, como acontece no caso das uniões de crédito e nas designadas ROSCAs (Rotating Savings and Credit Associations[2]) ou ASCRAs (Accumulating Savings and Credit Associations[3]). Algumas estruturas de microfinanciamento podem também funcionar simultaneamente com fundos externos, originários de diversas fontes, e com recursos próprios, resultantes das poupanças da comunidade.

Podem assim existir várias classificações alternativas para os diferentes modelos de microfinanciamento.[4] Uma delas distingue entre iniciativas que emergem no seio das comunidades e as que são sugeridas do exterior. No primeiro caso, podem constituir-se associações, cooperativas, uniões de crédito, ROSCAs, bancos comunitários ou bancos de aldeia, dependendo por exemplo das regulamentações locais ou das vantagens fiscais. Estes modelos podem ser apenas uma das múltiplas componentes de programas mais abrangentes de desenvolvimento da área em que são implementados. As iniciativas externas à comunidade podem ser direccionadas para os indivíduos, através da concessão de crédito ou oferta de garantias bancárias, ou destinar-se a grupos de pessoas, colectivamente responsáveis pelos empréstimos concedidos a cada uma delas. Este modelo de grupo foi inicialmente adoptado no Bangladeche, pelo pioneiro Grameen Bank, onde as tradicionais garantias bancárias eram substituídas pela responsabilidade colectiva e pela pressão dos pares, e compreende actualmente a maioria dos beneficiários de microcrédito à escala mundial (Hermes e Lensink, 2007).

Nos esquemas de empréstimo em grupo, a instituição responsável pelos fundos financia conjuntos de indivíduos, formados pelos próprios ou constituídos pelo organismo financiador, que são colectivamente responsáveis pelo pagamento e serviço de todas as dívidas. Cada elemento do grupo está naturalmente motivado para supervisionar e apoiar os colegas e procura assegurar o cumprimento das obrigações inerentes a todos os empréstimos. De facto, se houver incumprimento da parte de um dos elementos, todos os outros terão que suprir a falta em conjunto, pois não o fazer significa comprometer o acesso de todo o grupo a fundos adicionais. Esta estrutura funcional reduz os problemas tradicionais de assimetria de informação inerentes aos mercados financeiros e, consequentemente, também as questões de selecção adversa e risco moral, tal como se demonstra em Arnott e Stiglitz (1991) ou, mais recentemente, em Chowdhury (2005). O apoio do grupo pode inclusivamente aumentar os índices de coragem, motivação e inspiração dos seus membros, bem como constituir uma excelente rede social, fundamental para o marketing e comercialização dos bens e serviços por eles produzidos. Tem igualmente a vantagem de substituir as garantias formais ou os historiais de crédito favoráveis, que de outra forma seriam exigidos como forma de acesso a financiamento.[5]

Uma vez que é grande a variedade de modelos de microfinanciamento, é importante tentar identificar, para cada caso, aquele que, teoricamente, maximiza a probabilidade de sucesso das iniciativas específicas de desenvolvimento estratégico, independentemente da orientação de género ou de serem destinados a regiões rurais ou urbanas, localizadas em países desenvolvidos

[2] Onde um grupo informal de participantes contribui periodicamente para um fundo comum, que é emprestado à vez, em cada período, a um dos participantes, seleccionado aleatoriamente, por meio de negociação, ou de outra forma previamente estabelecida.
[3] Onde a participação dos membros se vai acumulando até um deles decidir tomar parte dos fundos acumulados sob a forma de empréstimo.
[4] Muhammad Yunus sugere uma classificação lata em dez categorias (em *www.grameen-info.org/bank/whatismicrocredit.htm*), que compreende desde os tradicionais credores locais aos sistemas de microcrédito do tipo do Grameen Bank.
[5] Stiglitz (1990) designa este fenómeno como 'garantia-social' (social-collateral).

ou em desenvolvimento. Tal exercício é crucial para a definição de políticas, instituições e programas de microfinanciamento eficientes.

Alguns estudos mostram que os sistemas de empréstimos em grupo são mais difíceis de implementar em países desenvolvidos (por exemplo Frankiewicz, 2001), registando custos operacionais e taxas de não cumprimento mais elevados e resultados globais mais fracos. Nestes países, o empréstimo individual parece ter maior probabilidade de sucesso, não só por ser mais flexível e ajustar-se melhor ás necessidades de financiamento dos candidatos, mas também por garantir maior confidencialidade e ser menos exigente em termos do tempo gasto em reuniões de grupo e na monitorização dos pares.

Contudo, Gomez e Santor (2003) mostram que os esquemas de grupo apresentam taxas de não cumprimento mais baixas do que as registadas para os empréstimos individuais, em resultado dos processos de selecção do grupo e dos incentivos no seio do mesmo. Os autores analisam duas grandes instituições canadianas de microfinanciamento, recorrendo a técnicas de *credit ranking* e à estimação de um modelo probit, obtendo taxas de não cumprimento elevadas, mas bastante distintas, com valores de 21.2% para os esquemas de grupo e de 41.4% para os credores individuais. No entanto, de acordo com a bibliografia existente, as taxas de não cumprimento são habitualmente muito mais baixas nos países em desenvolvimento, onde valores de um só dígito não são raros.

Do ponto de vista da eficiência funcional, o microcrédito devia ser gerido por bancos comerciais, dada a sua longa tradição na oferta de crédito e os conhecimentos, recursos e infra-estruturas disponíveis, nomeadamente as vastas redes de balcões especialmente vocacionados para a prestação de serviço bancário aos potenciais utilizadores. Contudo, para que pudessem de facto prestar um serviço não discriminatório, os bancos comerciais teriam que alterar substancialmente alguns dos seus hábitos mais profundamente enraizados, como por exemplo o favorecimento consistente dos homens no acesso ao crédito. Como mostra por exemplo Berger (1989), ainda que as mulheres sejam normalmente maioritárias nos esquemas de microfinanciamento, os homens predominam sempre em programas geridos directamente por bancos comerciais.

De acordo com Remenyi (2000), os programas de microfinanciamento são normalmente bem sucedidos em situações onde os bancos comerciais habitualmente falham porque os promotores dos primeiros aprenderam com a experiência 6 lições fundamentais. Estas dizem respeito (1) à substituição das garantias tradicionais por formas alternativas e mais eficientes de selecção de clientes; (2) à ligação dos interesses próprios de cada cliente aos objectivos de redução de riscos e das taxas de incumprimento; (3) à minimização dos custos operacionais na concessão de pequenos empréstimos e constituição de pequenos depósitos; (4) à utilização da pressão dos pares para incentivar o cumprimento das obrigações contratuais; (5) à oferta de produtos e serviços financeiros compatíveis com as necessidades dos pobres; (6) à segmentação dos mercados de forma a não perder de vista o público alvo – os pobres - mas procurando atrair apenas os realmente motivados.

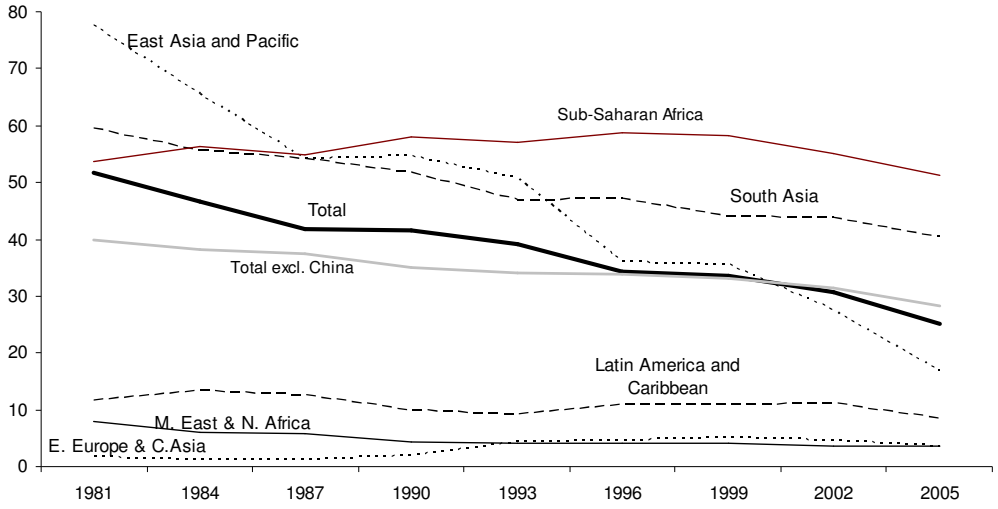
A fidelização aos princípios fundamentais do microfinanciamento potencia o sucesso deste tipo de programas, em termos do enfraquecimento e eventual abolição de ciclos viciosos de pobreza, mas também a viabilidade deste processo de financiamento. Como refere Yunus em entrevista à *The Economist* (19-03-2009), o microcrédito não foi afectado pela crise financeira actual porque está profundamente enraizado na economia real: *'When we give a loan of $100, behind the $100 there are chickens, there are cows. It is not something imaginary'*.

3. A POBREZA EM PORTUGAL E NO ALENTEJO

Almeida Garrett, o mais proeminente autor português do período romântico, coloca numa das suas obras mais relevantes a seguinte questão: *'E eu pergunto aos economistas políticos, aos moralistas, se já calcularam o número de indivíduos que é forçoso condenar à miséria, ao trabalho desproporcionado, à desmoralização, à infâmia, à ignorância crapulosa, à desgraça invencível, à penúria absoluta, para produzir um rico?'* Esta pergunta tem permanecido sem resposta, provavelmente porque, tal como muitas vezes acontece, a resposta dos economistas políticos é *'depende'*. Depende, por exemplo da forma como são definidos os conceitos de pobreza e de riqueza.

A definição mais comum de pobreza é a privação de bens essenciais, como alimentação, habitação ou vestuário, exacerbada pela falta de emprego e de oportunidades de educação que poderiam permitir a sua erradicação. O Banco Mundial define como linha de pobreza o valor de 1,25 dólares americanos por dia, medidos em paridade de poder de compra.[6] De acordo com este valor, necessariamente arbitrário, a percentagem de pobres nos países em desenvolvimento diminuiu nas últimas décadas, ainda que a taxas significativamente mais baixas se excluirmos a China, país onde a tendência de queda tem sido significativamente mais pronunciada. A diminuição da percentagem de pobres tem sido assim limitada ao grupo dos países asiáticos, tendo todas as outras regiões registado melhorias pouco significativas.

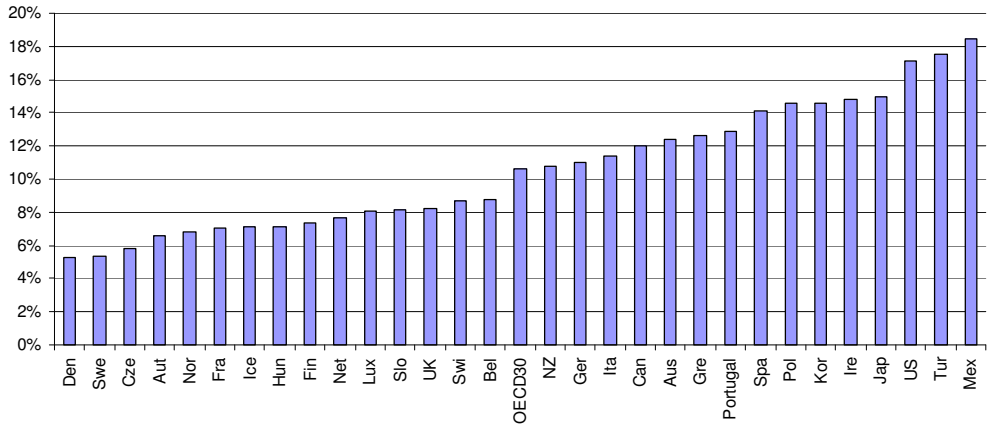
Figura 1: População abaixo da linha de pobreza (%)



Fonte: Chen e Ravaillon (2008)

Se não associarmos a pobreza a um nível mínimo de rendimento compatível com a subsistência, o conceito passa a ser essencialmente relativo. A introdução da palavra pobreza num motor de busca da Internet, como por exemplo o Google, dá origem a uma miríade de definições diferentes, obviamente adaptadas às diferentes circunstâncias. Um 'pobre' nos países economicamente mais desenvolvidos não enfrentará as mesmas dificuldades de um 'pobre' que vive em nações do designado Terceiro Mundo. No entanto, a pobreza relativa também tem consequências muito sérias ao nível do bem-estar individual e da coesão social.

Figura 2: Percentagem da população com menos de 50% do rendimento disponível mediano (2005)

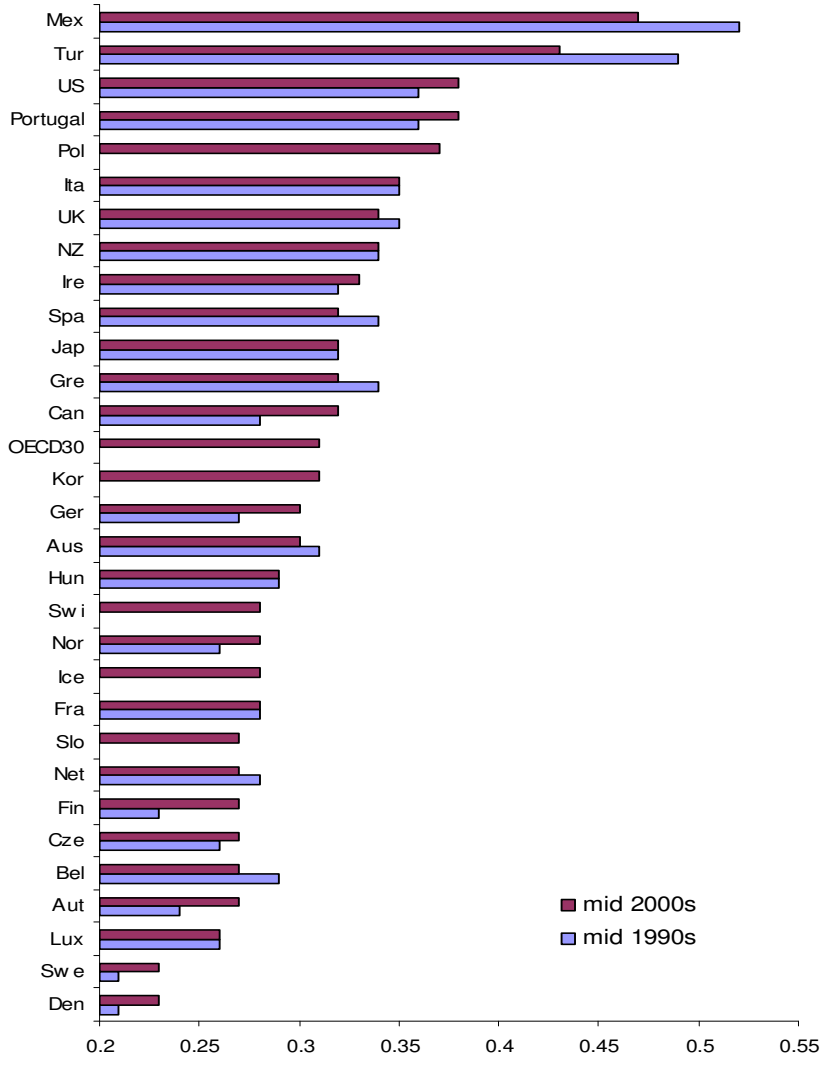


Fonte: OCDE (2008)

[6] Valor actualizado em 2008 da anterior estimativa de 1 dólar por dia como padrão internacional de pobreza.

Adoptando um conceito de pobreza mais relativo, por oposição ao de pobreza absoluta estimado pelo Banco Mundial, encontramos em alguns dos países relativamente desenvolvidos desigualdades substanciais na distribuição do rendimento. A figura 2 apresenta os valores do coeficiente de Gini para o rendimento disponível. Na lista de países da OCDE, Portugal aparece no terceiro pior lugar, com pior desempenho que todos os outros membros da UE, tendo inclusivamente piorado neste domínio ao longo da última década.

Figura 3: Coeficiente de Gini, baseado no rendimento disponível líquido de impostos e transferências (população total)

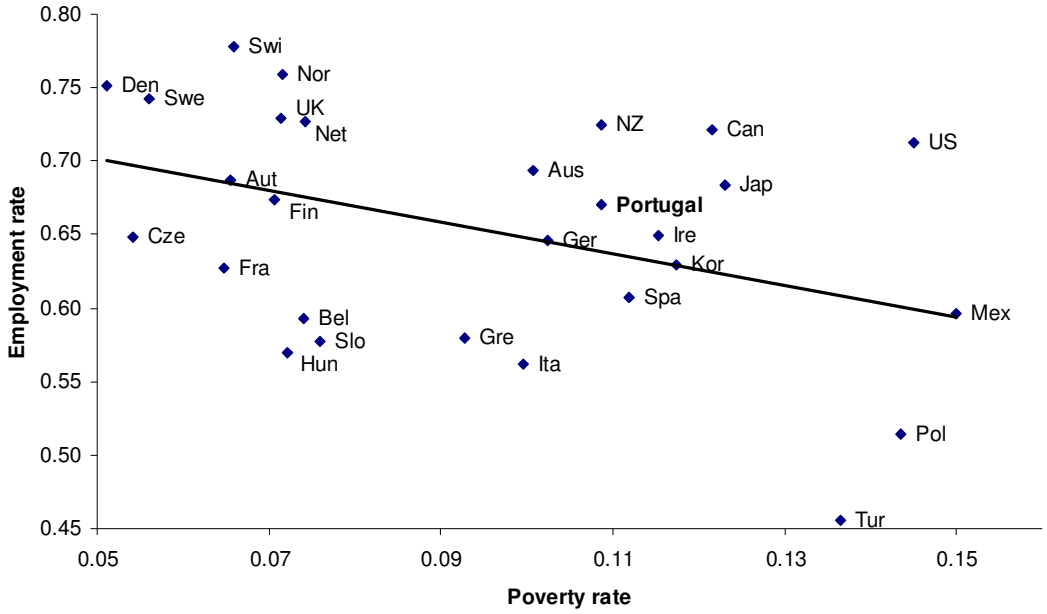


Fonte: OCDE (2008).

Nota: os valores do coeficiente de Gini variam ente 0, que corresponde a ausência de desigualdade, e 1, para desigualdade máxima.

Tendo como referência os valores do PIB per capita em paridade de poder de compra, Portugal é um país desenvolvido que pertence ao grupo dos 25% melhores do mundo. Apesar disso, em 2007, o valor deste parâmetro correspondia apenas a 74,8% da média da UE27. Existem igualmente em Portugal graves problemas de desigualdade na repartição do rendimento, que fazem com que deva ser prestada muita atenção à formulação e implementação de estratégias de erradicação da pobreza, e em particular aos programas de microfinanciamento. De facto, dada a expressiva relação negativa entre emprego e pobreza, explicitada na figura 4, se o financiamento contribuir para resolver problemas de desemprego, estará também, necessariamente, a desagravar a pobreza.

Figura 4: Taxas de emprego e de pobreza na população activa (2005)



Fonte: OCDE (2008)

De acordo com Alves (2009), a pobreza resulta, entre outras, de uma série de causas pessoais, económicas, sociais, com impacte geográfico variável. Entre estas causas, o autor realça a importância das decisões familiares face a choques específicos ou sistemáticos, a estrutura socio-demográfica das regiões, a sua dotação de capital humano, o grau de eficiência dos mercados de trabalho, de crédito e de bens, bem como o papel das políticas públicas e a sua influência nos incentivos e escolhas individuais. Cada um destes factores, ainda que por si só não explique o perfil de pobreza de uma região, contribui para a sua diferenciação e requer uma atenção especial por parte dos responsáveis pelas políticas sociais.

Os estudos realizados com o objectivo de identificar os principais determinantes da pobreza em Portugal, como por exemplo Soares e Uva (2002), sugerem que o fenómeno está associado à idade (sendo as crianças e os idosos os mais afectados), ao género (o risco de pobreza aumenta quando a pessoa de referência num agregado familiar é uma mulher), à composição dos agregados familiares (aumentando o risco de pobreza de forma muito significativa para agregados com 3 ou mais dependentes), aos baixos níveis de escolaridade, à ruralidade e ao isolamento. No caso do Alentejo, a ruralidade, a estrutura demográfica e a escolaridade são provavelmente os factores que têm vindo a contribuir mais para o relativo atraso da região e para a sua aparente incapacidade de convergência para valores mais consentâneos com a sua circunstância de território da zona euro.

Ainda que as condições de vida em Portugal e no Alentejo tenham sofrido melhorias muito significativas na sequência da instauração do regime democrático, em 1974, e da integração nas comunidades europeias, em 1986, ambos têm classificações relativas muito desfavoráveis em termos de pobreza. O risco de pobreza em Portugal diminuiu 6 pontos percentuais numa década (de 24% em 1997 para 18% em 2007, de acordo com dados do INE), mas o país continua a ser o mais desigual da UE27 em termos da distribuição do rendimento. Por outro lado, o Alentejo, ainda que seja actualmente a região mais pobre em Portugal (com uma taxa de pobreza de 17%), é a mais igualitária a nível nacional.

O nível de pobreza de um território pode ser avaliado tendo em conta os valores médios do rendimento ou os da despesa dos agregados familiares que o constituem. De acordo com os estudos existentes, a óptica da despesa é mais adequada para a identificação de situações de pobreza persistente. Independentemente da perspectiva adoptada na identificação da pobreza, o Alentejo tem tido sempre um desempenho mau no contexto português. Houve uma melhoria relativa entre 1995 e 2000, período em que a região deixou de ocupar a última posição, chegando a ser terceira, mas a partir de então o seu desempenho relativo deteriorou-se levando-a, de novo, ao último lugar que agora ocupa (ver quadro 1).

Quadro 1: Níveis médios de despesa e de rendimento dos agregados familiares

	Portugal	Lisbon	Alentejo	Ranking
		Despesa Média dos Agregados		
1989	6506	7409	4942	antepenúltimo
1995	11569	11993	8449	último
	(78%)	(62%)	(71%)	
2000	13828	15164	11022	penúltimo
	(20%)	(26%)	(30%)	
2005	17607	20715	14067	último
	(27%)	(37%)	(28%)	
		Rendimento Médio dos Agregados (preços correntes)		
1989	6352	6947	4952	último
1995	12615	13885	9904	último
	(99%)	(100%)	(100%)	
2000	16189	18203	13562	antepenúltimo
	(28%)	(31%)	(37%)	
2005	22136	27463	18276	último
	(37%)	(51%)	(35%)	

Fonte: INE (taxas de crescimento entre parêntesis)

A baixa densidade populacional do Alentejo é frequentemente apontada como causa do seu fraco desempenho económico. Ocupando uma correspondente a 34% do território português e absorvendo apenas 7% da sua população, a densidade populacional da região é de apenas 24 hab/Km^2, a mais baixa do país, e relativamente distante do valor médio nacional de 115 hab/Km^2. Por um lado, a baixa densidade populacional não permite a escala mínima compatível com a viabilidade de muitas iniciativas empreendedoras vocacionadas para mercados locais. Por outro lado, o facto do Alentejo ser escassamente povoado torna-o menos interessante do ponto de vista eleitoral e, por isso, tem beneficiado menos de investimentos públicos que são implementados quando os governos procuram maximizar a probabilidade de reeleição.

O Alentejo partilha com o conjunto do país uma série de características sócio demográficas e enfrenta problemas da mesma natureza, embora alguns sejam sentidos de forma particularmente aguda nesta região. A população residente no Alentejo tem diminuído, ao contrário do que acontece a nível nacional. De 1991 a 2001 registou-se um decréscimo de 0,7% no Alentejo e um aumento de 5% a nível nacional. Para além da diminuição, a população da região apresenta sinais expressivos de envelhecimento, bem patentes na evolução do Índice de Envelhecimento que passou de 107 em 1991 para 168 em 2004. Actualmente, 23% da população residente no Alentejo tem mais de 65 anos (o valor é de 17% para Portugal), correspondentes a 9% do total do país para este nível etário. A percentagem de jovens (até 14 anos) é 6%, valor substancialmente inferior aos 15% para o total nacional.

Os agregados familiares onde uma mulher é a pessoa de referência enfrentam risco de pobreza acrescido, que é agravado se o local de residência for o Alentejo. De facto, enquanto o rendimento anual médio nos agregados femininos em Portugal atinge apenas 88% do recebido nos agregados masculinos, no Alentejo esse valor é de ainda menor: 85%. O agregado médio no Alentejo recebe 83% do valor médio nacional, com os agregados alentejanos masculinos e femininos a auferir respectivamente 84% e 80% dos recebidos pelos agregados portugueses médios.

Os níveis de qualificação dos activos residentes no Alentejo são baixos, particularmente quando comparados com os padrões médios da UE. A maioria dos indivíduos frequentou apenas o primeiro ciclo do ensino básico (50,4% dos residentes no Alentejo e 46,2% a nível nacional, em 2001). Apenas 6,1% da população do Alentejo conclui o ensino superior enquanto para o conjunto do país a média é 8,4%.

A ruralidade é igualmente um factor associado à pobreza e o Alentejo é a região do país onde esta tem mais expressão. O peso relativo do sector primário é significativo na região,

sendo a única do país onde ele ocupa ainda a segunda posição. 20,8% das empresas locais estão classificadas no sector da agricultura, silvicultura e pescas, mas o seu contributo é apenas de 6,4% em termos nacionais. O sector terciário é dominante e orientado para actividades de baixo valor acrescentado, principalmente em serviços públicos como a educação, administração pública ou segurança social. As actividades industriais são as menos significativas à escala regional e as instituições públicas locais são, em muitos locais, os principais empregadores. Nesta região uma pequena minoria de indústrias de elevada produtividade, como as localizadas no complexo industrial de Sines e algumas agro-indústrias, coabitam com uma grande variedade de actividades pouco produtivas, como a agricultura de subsistência, o comércio a retalho e segmentos não qualificados da administração pública.

Outra característica típica de territórios menos desenvolvidos, que se verifica igualmente na generalidade do país, é a elevada proporção de microempresas e o autoemprego. 87,7% das empresas locais empregam menos de 10 trabalhadores (a nível nacional são 84,5%) e 70% menos de 5 (65,2% a nível nacional). Nos países desenvolvidos as microempresas são responsáveis por 8 a 20% do emprego, enquanto em Portugal esse valor é de 30%.[7]

As dinâmicas do mercado de trabalho contribuem também para a diferenciação da região Alentejo. Nesta zona a taxa de actividade é 3 pontos percentuais mais baixa do que a nacional (49% e 52,2%, respectivamente), sendo este indicador ainda mais desfavorável para as mulheres (41,9% no Alentejo e 46,7% para o total do país). A taxa de desemprego na região tem sido consistentemente mais elevada do que a média nacional. A título de exemplo, referem-se os valores do 4º trimestre de 2008, de 10% para o Alentejo e 7,8% para Portugal. Tal como acontece a nível nacional, o desemprego afecta mais as mulheres e mais os jovens, independentemente do sexo.

4. MICROFINANCIAMENTO NO ALENTEJO: O MICROCRÉDITO E O PROGRAMA FAME

As desvantagens relativas do Alentejo, acima retratadas, fazem dele um território adequado à implementação de projectos de microfinanciamento. O nível educacional médio da respectiva população, os seus níveis de pobreza, a escassez de empresas de grande dimensão com capacidade para absorver o desemprego local e os fluxos migratórios provenientes maioritariamente da Europa de Leste, têm contribuído para a criação de bolsas de desempregados e dependentes que habitualmente ficam sob a alçada da segurança social. Tais situações de precariedade dificilmente evoluem no sentido da sustentabilidade pessoal e, por arrastamento, das economias locais, se os mais necessitados não tiverem acesso ao ingrediente indispensável à resolução da maioria dos problemas económicos: capital financeiro.

4.1. O microcrédito

Em Portugal o microcrédito foi introduzido pela Associação Nacional de Direito ao Crédito (ANDC), organização não lucrativa, criada em 1998, que funciona em estreita ligação com o IEFP (Instituto do Emprego e Formação Profissional) e com o Banco Comercial Português, instituição financeira privada que garante a disponibilidade dos financiamentos e assume o risco da maioria das potenciais perdas. Os empréstimos são concedidos apenas a indivíduos seleccionados pela ANDC e os primeiros créditos datam de 1999. Inicialmente, o financiamento disponível por indivíduo, e projecto, não ultrapassava os 5000 euros, reembolsáveis em 36 mensalidades. Actualmente, os empréstimos podem atingir o valor máximo de 7500 euros e estão já envolvidas outras entidades bancárias privadas: Caixa Geral de Depósitos e Banco Espírito Santo. A ANDC continua a ser a entidade responsável pela certificação da qualidade das propostas de negócio, bem como pelo acompanhamento dos microempresários durante a existência do empréstimo. De acordo com um estudo da Universidade Católica Portuguesa (2007), em Janeiro de 2007, a 'taxa de delinquência' – ou incumprimento – era de 18,1%. Até

[7] Cf. Frankiiewicz, 2001 e DEEP (2004).

Agosto de 2006 a ANDC registou atrasos de pagamento em 141 projectos, no valor de 167978,54 euros.

Com o objectivo de divulgar o conceito e a prática do microcrédito em Portugal, a ANDC tem implementado diversas estratégias. Primeiro procurou desenvolver colaborações com instituições locais, que na maioria dos casos não foram bem sucedidas, principalmente porque os interlocutores tinham dificuldade em apreender a natureza 'não caritativa' do microcrédito. Posteriormente, foram feitas tentativas para encontrar agentes locais capazes de promover o microcrédito mas, mais uma vez, sem resultados positivos por falta de competências adequadas por parte dos interlocutores contactados.

Como qualquer outro mecanismo de financiamento, o microcrédito, enquanto instrumento de combate à pobreza, só é bem sucedido se canalizado para actividades empresariais potencialmente lucrativas. Assim, é fundamental que os agentes de microcrédito tenham capacidade de análise que lhes permita avaliar e seleccionar os clientes. Têm que saber distinguir entre indivíduos capacitados para implementar uma ideia de negócio coerente e outros que não têm aptidão para vir a gerar o seu próprio posto de trabalho e, consequentemente, o seu rendimento. Actualmente existem em Portugal um pequeno de agentes de microcrédito, qualificados, que assumem a responsabilidade por áreas geográficas extensas, nas quais podem de forma efectiva e consequente dar apoio aos potenciais empreendedores, trabalhando em conjunto com associações de âmbito local.

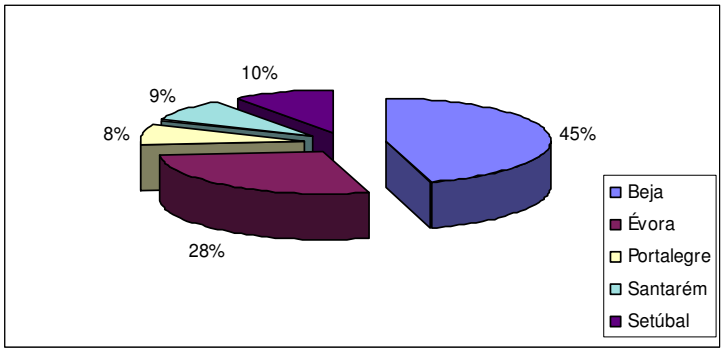
Em sete anos de actividade integral, a ANDC possibilitou o acesso ao crédito a cerca de 700 microempresários[8], com um efeito multiplicador no emprego de 1,2. Seria de esperar, e de acordo com o baixo nível de desenvolvimento do Alentejo, que a proporção local de projectos financiados fosse relativamente elevada. Contudo, a região representa apenas cerca de 10% do total de microempresários, valor que, apesar de baixo, é mais expressivo do que o peso relativo da população residente no Alentejo, ou que a contribuição do PIB regional para o nacional (cerca de 7% em ambos os casos).

Em Portugal, tal como nos países mais desenvolvidos, o microcrédito é basicamente um fenómeno urbano. Metade dos microempresários financiados reside nas áreas de Lisboa e do Porto, a primeira representando 38,3% do total. Ao contrário, no Alentejo o microcrédito assume uma natureza mais rural, o que provavelmente justificaria a adopção de um modelo de microcrédito mais específico, em linha com o que tem sido aplicado noutras regiões com o mesmo tipo de perfil, noutras áreas menos desenvolvidas do mundo. Esta sugestão é reforçada pela análise da distribuição geográfica dos projectos financiados no Alentejo. Como se constata pela figura 5, o distrito de Beja absorve cerca de 45% dos projectos e, dentro deste, o concelho de Moura, local predominantemente rural, retém cerca de 31,7% do total do distrito. Neste caso, o sucesso relativo do microcrédito está directamente relacionado com a actividade desenvolvida por uma associação de desenvolvimento local (ADC Moura), a qual tem desempenhado um papel fundamental, não apenas na promoção do conceito de microcrédito, mas também no constante apoio aos promotores dos projectos, ao longo do processo de candidatura e durante a fase de reembolso do financiamento. Noutras localidades, sejam cidades ou pequenas vilas, existe menor número de projectos financiados, embora algumas registem um peso maior de indivíduos com baixos níveis de vida bem e outras tenham níveis mais elevados na actividade económica. No primeiro caso haveria maior necessidade de ajuda com vista à recuperação de situações de pobreza relativa, no segundo poderiam configurar-se cenários mais interessantes do ponto de vista do desenvolvimento dos micronegócios.

A análise dos dados, e em particular a comparação do caso de Moura com outros concelhos alentejanos, sugere que os financiamentos de microcrédito apresentados por entidades ou indivíduos estranhos ao local de residência dos potenciais clientes têm menor sucesso. A proximidade física e relacional entre os promotores de microcrédito e os potenciais empreendedores parece ser essencial para o sucesso dos projectos e para a multiplicação dos mesmos.

[8] Neste trabalho são analisados dados relativos à actividade da ANDC até Março de 2007.

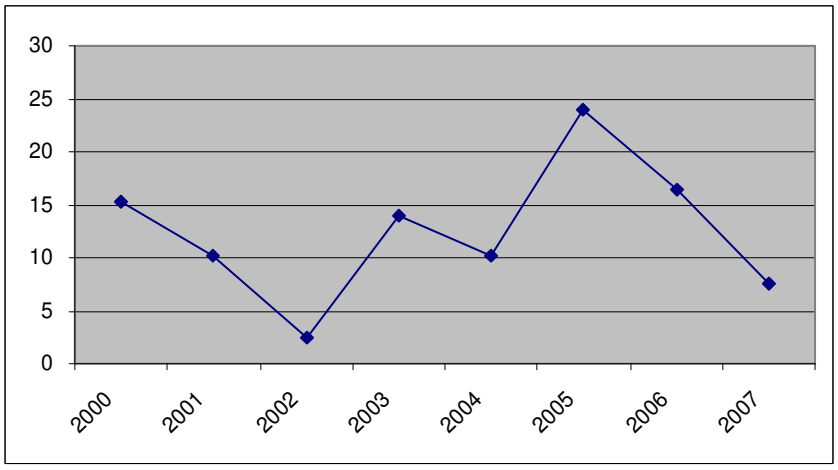
Figura 5: Projectos de microcrédito no Alentejo



Fonte: Elaboração própria a partir de dados recolhidos na ANDC.

A figura 6 regista a evolução do número de projectos financiados no Alentejo, que tem sido irregular, embora com tendência crescente (o valor para 2007 corresponde ao total de projectos financiados durante o primeiro trimestre, apenas). Os anos de 2005 e 2006 foram os mais dinâmicos, facto a que não terá sido estranho o abrandamento em 2005 da taxa de crescimento anual do PIB, que passou de 4% para 3%, e a aceleração da taxa de desemprego, que atingiu então o valor máximo dos 20 anos anteriores.

Figura 6: Evolução dos projectos financiados no Alentejo



Fonte: Elaboração própria a partir de dados recolhidos na ANDC.

No quadro 2 sintetizam-se as principais características do microcrédito no Alentejo e em Portugal. Em ambos os casos, a maioria dos clientes financiados são portugueses, casados, vivendo em famílias com mais de 3 membros (figura 7). Apesar destes aspectos comuns, existem diferenças substanciais nos dois territórios que se traduzem em perfis de microcrédito distintos. No Alentejo, os beneficiários deste tipo de crédito são maioritariamente homens, ao contrário do que acontece no resto do país onde, na classe etária dos 35-45 anos, cerca de 80% dos projectos são promovidos por mulheres. Portugal está assim em sintonia com o expectável, pois a pobreza, o desemprego e a exclusão afectam principalmente mulheres. No Alentejo deve ser prestada mais atenção às questões de género, tal como acontece nas áreas menos desenvolvidas onde as mulheres são especialmente visadas pelos agentes de microcrédito e são escolhidas como responsáveis pelo financiamento, em vez dos homens, mesmo nos agregados onde ambos estão desempregados. Tal escolha decorre, por um lado, do facto das mulheres estarem expostas a riscos mais elevados de pobreza e, por outro, da experiência prática que sugere serem as mulheres consistentemente mais capazes de gerir de forma equitativa e produtiva os orçamentos familiares, em especial em situações de pobreza extrema.

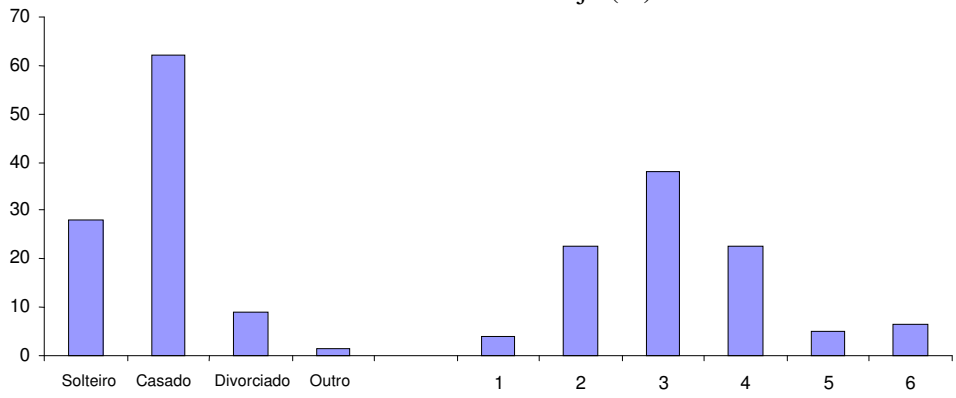
Quadro 2: Principais características comparativas dos micro-empresários beneficiários de microcrédito: Portugal e Alentejo (%)

Empresários		Portugal[(1)]	Alentejo[(2)]
Género	Homens	47,1	50,6
	Mulheres	52,8	49,4
Idade	18-25 anos	11,6	13,9
	26-35 anos	30,4	41,8
	36-45 anos	34,9	26,6
	46-55 anos	19,8	13,9
	> 56 anos	4,2	3,8
Nacionalidade	Portuguesa	n.a.	97,47
	Outra	n.a.	2.53
Nivel de Escolaridade	Nenhum nível de educação	1,5	
	1° ciclo	22	29,11
	2º ciclo	22	20,25
	3º ciclo	22	21,52
	Secundário	22	26,58
	Ensino Superior	10	2,53
Outras qualificações	p.e., formação profissional	n.a.	37,97
	Não	n.a.	62,03

Fontes: [(1)]Mendes et al. (2007). [(2)]Dados recolhidos pelos autores junto da ANDC .

No Alentejo, a maioria dos beneficiários tem, no início do empréstimo, 26 a 35 anos e muitos já tinham experiência anterior como empresários em nome individual, embora tais negócios pudessem não estar a funcionar no momento em que se inicia o processo de microfinanciamento. É este o grupo etário mais atingido pelo desemprego na região, e, provavelmente por esta razão, é nesta idade que se verifica maior predisposição para o empreendedorismo e para a assumpção do risco de criação do próprio emprego. Por comparação com o Alentejo, em Portugal os beneficiários de microcrédito são em média mais velhos e melhor qualificados. Em ambos os casos, apenas 1/3 está envolvida na frequência de cursos profissionais, promovidos principalmente pelos Centros de Emprego e Formação Profissional.

Figura 7: Estado civil e dimensão dos agregados familiares dos beneficiários de microcrédito no Alentejo (%)



Estado civil (esquerda) e Número de pessoas por agregado familiar (direita)

Fonte: Elaboração própria a partir de dados recolhidos na ANDC.

Mais de 2/3 dos beneficiários de microcrédito no Alentejo utilizam esta fonte de financiamento com os objectivos de iniciar um novo negócio e criar o seu próprio posto de trabalho. Muitos viveram períodos de desemprego, em alguns casos prolongado (quadro 3). Mais de 80% obteve também outras fontes de financiamento, maioritariamente proveniente de fundos destinados à promoção da criação do próprio emprego.

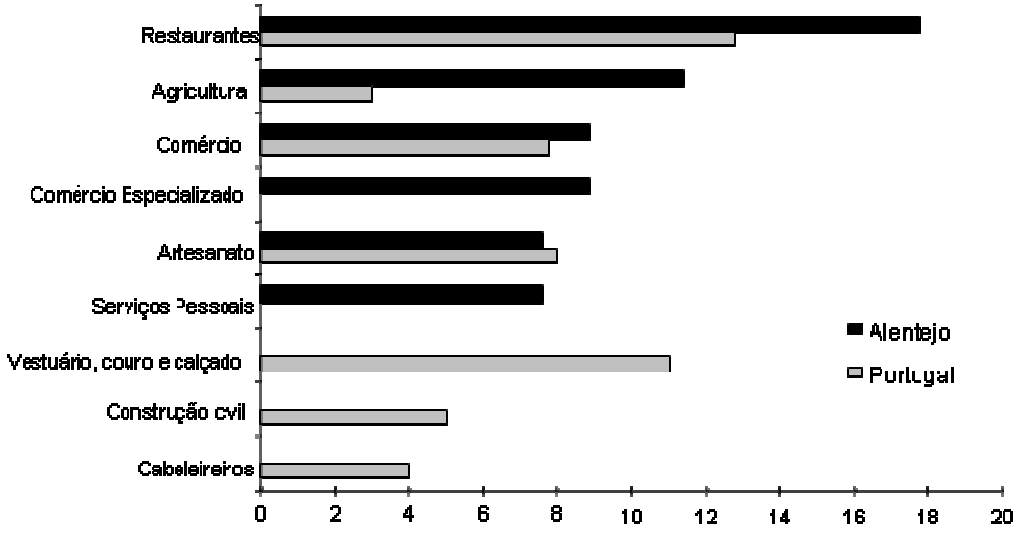
Quadro 3: Motivação, financiamento e emprego criado pelos projectos de microcrédito no Alentejo

Questão	Resposta	%
Razões para utilizar microcrédito	Iniciar negógio	67,09
	Reforçar negócio existente	31,65
	Outro	1,27
Situação laboral anterior	Empregado por conta de outrém	8,86
	Empregado por conta própria	50,63
	Desempregado	27,85
	Desempregado de longa duração	12,66
Número de postos de trabalho criados	1	77,22
	2	20,25
	Outro	2,53
Outras fontes de financiamento usadas no projecto?	Sim	81,01
	Não	18,99

Fonte: Elaboração própria a partir de dados recolhidos na ANDC.

A figura 8 mostra que as actividades objecto de financiamento são maioritariamente serviços de proximidade, por exemplo relacionadas com alimentação e bebidas, comércio a retalho e artesanato. No Alentejo, e reflectindo a ruralidade relativa da região, a agricultura e actividades relacionadas têm um peso importante. Uma vez que um dos problemas de desenvolvimento da região pode estar relacionado com a sua reduzida escala enquanto mercado, os microempresários naturalmente vocacionados para este tipo de actividades, que apresentam muitas vezes características específicas apenas desta zona do país e que são muito valorizadas quando adequadamente promovidas, devem ser orientados no sentido de tentar conquistar procura externa para os seus produtos. Em muitos casos, existem nichos de mercado inexplorados mas que são potencialmente lucrativos, desde que ao esforço produtivo seja aliado o esforço de distribuição.

Figura 8: Sector de actividade dos projectos de microcrédito, Portugal[1] e Alentejo[2] (%)



Fontes: [(1)]Mendes et al. (2007). [(2)]Dados recolhidos pelos autores junto da ANDC.

O valor médio dos empréstimos concedidos no Alentejo é ligeiramente mais elevado do verificado a nível nacional (4804 euros e 4375 euros, respectivamente), mas em ambos os casos tais valores são significativamente mais baixos que os respectivos PIB *per capita.* Esta é uma característica que diferencia o modelo de microcrédito português, e também o alentejano, dos modelos típicos dos países em desenvolvimento, pois em muitos países asiáticos, africanos e latino-americanos, os empréstimos médios são substancialmente mais elevados que o PIB *per capita*.

4.2. O Programa FAME

O programa FAME (Fundo de Apoio às Micro-Empresas) é um instrumento de financiamento criado com o objectivo de apoiar empresas com um número máximo de 10 empregados (microempresas) e promover os seus ganhos de competitividade, dotando-as de maior capacidade para a prestação de serviços de qualidade. Os promotores deste programa seleccionam para apoio financeiro a projectos com características inovadoras ou que promovam a diferenciação no contexto da região. Privilegiam o apoio ao investimento nas áreas da qualidade, ambiente, segurança e higiene.

A entidade promotora deste instrumento financeiro é a Agência de Desenvolvimento Regional do Alentejo (ADRAL) que, por via de parcerias público-privado, estabelece acordos com as diversas autarquias e com entidades bancárias (neste caso, com o Banco Espírito Santo). Nesta parceria participam também o Instituto de Apoio às Pequenas e Médias Empresas (IAPMEI) e Sociedades de Garantia Mútua que partilham o risco com as restantes entidades financiadoras. Actualmente, os fundos são disponibilizados pelas autarquias (20%) e pela entidade bancária (80%). Sobre os empréstimos concedidos pelas autarquias não são cobrados juros. O juro cobrado pelo banco está indexado, com *spread* reduzido, à EURIBOR a 180 dias. Os empréstimos podem financiar até 100% do investimento (no máximo de 45 mil euros), sendo o prazo de reembolso variável, entre 3 a 6 anos, com um ano de carência.

A ADRAL e as outras entidades envolvidas nas parcerias acompanham os projectos após a sua implementação, através de auditorias que visam a aplicação do apoio financeiro e o comportamento do negócio, procurando avaliar a necessidade de apoios complementares ou outros. As áreas de negócio abrangidas por este programa incluem a indústria, comércio, serviços, turismo e construção. Uma das características diferenciadoras do FAME é o facto de ser estruturado a nível concelhio e ser também este o seu âmbito de actuação. Em cada concelho onde o programa está disponível só podem candidatar-se a apoio empresas aí sedeadas. A ADRAL e as autarquias são responsáveis pela promoção deste instrumento de financiamento, inicialmente através de sessões públicas institucionais, que visam dar a conhecer a existência e as respectivas regras de funcionamento e, posteriormente, em acções dirigidas especificamente aos potenciais interessados na sua aplicação, que têm que ser empresas já constituídas ou em formação.

Em cada concelho este instrumento financeiro assume características específicas, particularmente no que respeita aos montantes disponíveis nos fundos FAME municipais, que provavelmente reflectem as diferenças de densidade económica e populacional nos vários territórios (ver quadro 4).

Quadro 4: Concelhos onde existe o programa FAME e financiamentos disponíveis (€)

	Concelhos do Alentejo com FAME Activo	Montante do Fundo	Máximo Elegível por Projecto
NOVO FORMATO (Eixo III - FINICIA)	Alcácer	75.000,00	45.000,00
	Aljustrel	150.000,00	45.000,00
	Almodôvar	200.000,00	45.000,00
	Alter do Chão	250.000,00	45.000,00
	Alvito	100.000,00	45.000,00
	Arraiolos	125.000,00	45.000,00
	Avis	250.000,00	45.000,00
	Beja	500.000,00	45.000,00
	Borba	250.000,00	45.000,00
	Campo Maior	250.000,00	45.000,00
	Castelo de Vide	50.000,00	45.000,00
	Castro Verde	500.000,00	45.000,00
	Crato	250.000,00	45.000,00
	Cuba	75.000,00	45.000,00
	Évora	250.000,00	45.000,00
	Ferreira do Alentejo	250.000,00	45.000,00
	Gavião	250.000,00	45.000,00
	Marvão	100.000,00	45.000,00
	Mértola	250.000,00	45.000,00
	Moura	125.000,00	45.000,00
	Nisa	200.000,00	45.000,00
	Redondo	225.000,00	45.000,00
	Reguengos de Monsaraz	100.000,00	45.000,00
	Santiago do Cacém	125.000,00	30.000,00
	Serpa	100.000,00	45.000,00
	Sines	125.000,00	45.000,00
	Sousel	50.000,00	45.000,00
	Vendas Novas	250.000,00	45.000,00
	Viana	250.000,00	45.000,00
	Vidigueira	150.000,00	45.000,00
Formato Inicial	Montemor-o-Novo	150.000,00	15.000,00
	Portel	100.000,00	15.000,00
	TOTAL do FUNDO no ALENTEJO	**6.075.000,00**	

Fonte: ADRAL.

Este programa de apoio ao investimento empresarial tem um cariz marcadamente regional, actualmente com implantação em 32 concelhos de todas as NUTS III do Alentejo.[9] Nos dois últimos anos registou-se um alargamento geográfico muito significativo do mesmo, no Alentejo e no conjunto do país, pois o IAPMEI passou a também a aplicá-lo, integrando-o no âmbito mais vasto do Programa FINICIA (no seu Eixo III). No caso particular do Alentejo, o facto da

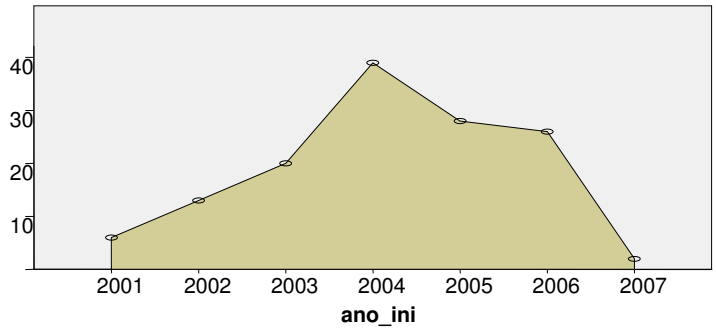
[9] Os dados reportados neste estudo foram recolhidos em 2007, e respeitam ao 1º trimestre desse ano. Na altura, o programa FAME estava em funcionamento em 14 concelhos onde haviam sido financiados 134 projectos. Os concelhos aqui analisados são Évora, Redondo, Ferreira do Alentejo, Beja, Vidigueira, Nisa, Portalegre, Moura, Montemor-o-Novo, Mértola, Aljustrel, Borba, Vendas Novas e Portel.

Direcção Geral das Autarquias Locais ter produzido um parecer que reconhecia a legalidade da criação deste tipo de instrumentos de apoio financeiro, bem como o efeito de imitação por parte dos concelhos vizinhos daqueles que já o disponibilizavam, foram determinantes para a sua disseminação. Os promotores da iniciativa também consideram que a maior cobertura territorial está relacionada com a alteração nas taxas de comparticipação, diminuindo o peso suportado pelas autarquias (inicialmente, estas e a entidade bancária participavam os fundos a 50%). Paradoxalmente, o expressivo aumento da cobertura territorial não tem sido acompanhada por um correspondente acréscimo no número de projectos financiados, provavelmente porque o tecido empresarial local ser muito deficiente e estes apoios destinarem-se a actividades já existentes.

Desde a sua instituição, em 2001, e até ao final do 1º trimestre de 2007, foi prestado apoio a cerca de 130 empresários no âmbito do programa FAME, sendo 2004 o ano mais produtivo (figura 9). A maior escala deste programa de microfinanciamento, por comparação com o de microcrédito analisado anteriormente, decorre provavelmente da forma como o FAME tem vindo a ser implementado, sempre em parceria com autarquias e com apoio sistemático após a concessão do empréstimo, mas também do próprio público-alvo. Visar apenas empresários já em actividade significa que se trabalha com pessoas que efectivamente já demonstraram a sua capacidade para implementar projectos próprios e que, por isso, têm algo a perder, estando necessariamente mais motivados para procurar ajuda para os seus planos de expansão ou de consolidação empresarial.

Os empréstimos concedidos no âmbito do programa FAME são em média de 16.000 euros, mas a amplitude de variação é elevada: enquanto o apoio máximo rondou os 49.000 euros, o mínimo foi de 1970 euros. O dinheiro pode ser aplicado em obras de adaptação, remodelação e conservação, na aquisição de equipamento básico (máquinas e equipamento específico), informático e de higiene e segurança.

Figura 9: Número de projectos FAME

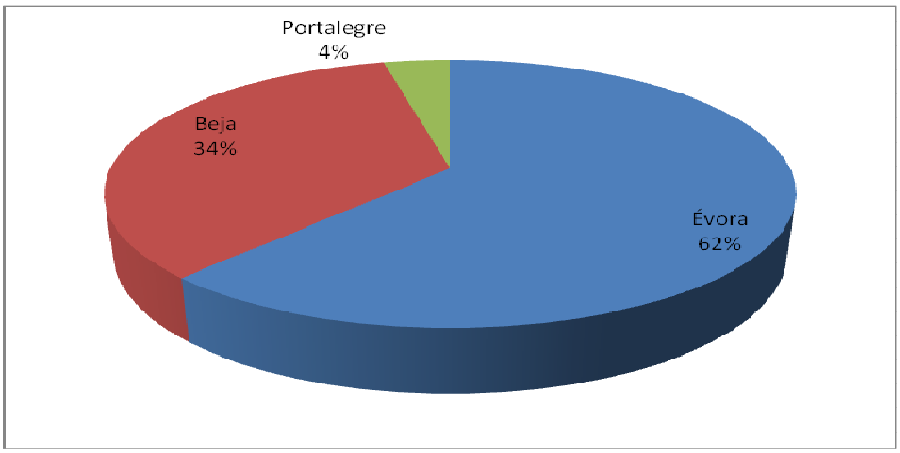


Fonte: Elaboração própria (dados recolhidos na ADRAL até ao 1º trimestre de 2007).

A análise dos dados relativos aos projectos financiados mostra que o distrito de Évora é dominante na distribuição dos financiamentos aprovados por distrito, concentrando mais de 60% dos projectos (figura 10). O facto do concelho de Évora apresentar valores significativamente mais elevados que os restantes concelhos (cerca de 1/3 do total de projectos aprovados em toda a região), condiciona a estrutura anterior e evidencia a maior dinâmica, em termos económicos e empresariais, registada neste território. É também mais um reflexo das significativas assimetrias intra-regionais, que se traduzem na desigual capacidade de resposta aos estímulos comuns para o desenvolvimento empresarial.[10]

[10] Évora, com 56 mil habitantes, é a maior cidade do Alentejo. A sua actividade económica está profundamente ligada aos serviços, particularmente aos públicos, posicionando-a como o grande pólo de actividade terciária no Alentejo Central.

Figura 10: Distribuição geográfica dos projectos FAME



Fonte: Elaboração própria (dados recolhidos na ADRAL).

Os micro-empresários que beneficiaram de apoio do programa FAME são sobretudo homens, maioritariamente entre os 26 e os 45 anos, casados, de nacionalidade portuguesa (quadro 5). Em termos de formação, estes replicam a generalidade da população activa da região. Cerca de 1/3 dos micro-empresários apenas possuem como habilitações o 1º ciclo do ensino básico e quase 50% não foram além do 9º ano de escolaridade.

Quadro 5: Principais características dos microempresários do programa FAME (%)

Empresários		Alentejo
Género	Homens	63,4
	Mulheres	36,6
Idade	18-25 anos	4,5
	26-35 anos	26,9
	36-45 anos	33,6
	46-55 anos	18,7
	> 56 anos	16,4
Estado Civil	Solteiro	15,7
	Casado	78,4
	Divorciado	5,3
Nacionalidade	Portuguesa	98,5
	Outra	1,5
Nível de Escolaridade	Nenhum nível de educação	
	1° ciclo	27,6
	2º ciclo	4,5
	3º ciclo	14,9
	Secundário	17,2
	Superior (Licenciatura)	15,7
	Bacharelato	1,5
	Outro	7,5

Fonte: Dados recolhidos junto da ADRAL.

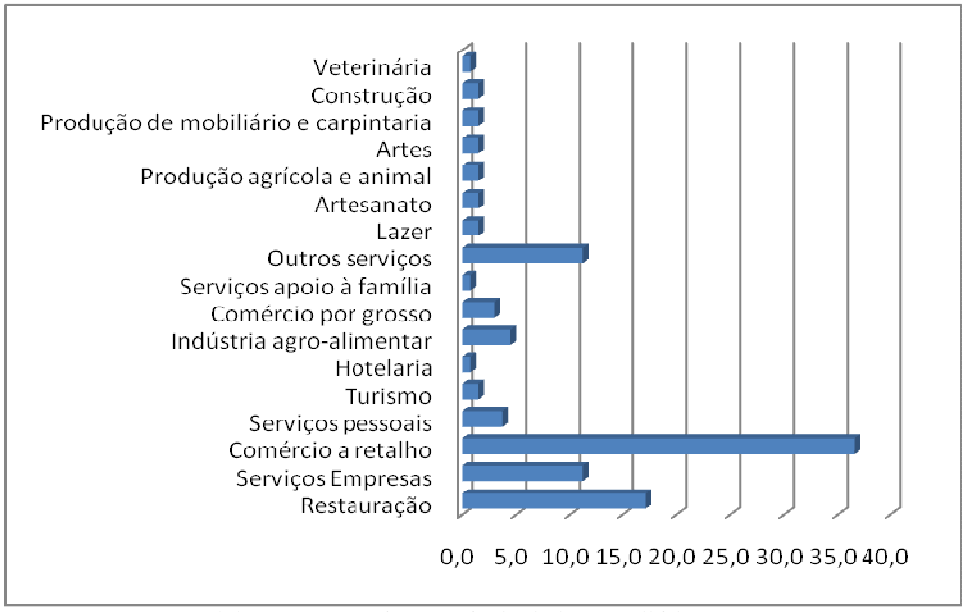
Como este programa é destinado apenas a iniciativas empresariais, já constituídas ou em fase de arranque, predominando as empresas em nome individual, verifica-se que todos os indivíduos apoiados afirmavam estar empregados no momento do contrato de crédito e, de um modo geral, desenvolviam já a sua actividade na área onde concretizaram o novo investimento. Como objectivo para a aplicação do empréstimo foi referida maioritariamente a criação de mais postos de trabalho. Apesar disto, o número médio de postos de trabalho criados por esta via é em média 2,5, um número baixo mas superior ao que se consegue com os programas de microcrédito.

Quadro 6: Forma jurídica e Número de postos de trabalho criados pelos projecto FAME

Questão	Resposta	%
Número de postos de trabalho criados	1	41
	2	25,4
	3	16,4
	4	3
	5	1,5
	6	6
	7	0,7
	8	2,2
	10	1,5
Forma Jurídica	Sociedade por Quotas	45,5
	Empresas em nome Individual	52,2
	Cooperativa	1,5
	Outra	0,7

Fonte: Elaboração própria a partir de dados recolhidos na ADRAL.

Figura 11: Distribuição dos projectos FAME por sector de actividade (%)



Fonte: Elaboração própria a partir de dados recolhidos na ADRAL.

Os projectos aprovados ao abrigo do programa FAME seguem o conjunto de actividades predominantes na região: o sector terciário é aquele onde se concentram mais investimentos, predominando as actividades de comércio a retalho e a restauração (figura 11). Os serviços de apoio, às empresas e outros de natureza diversa, com carácter mais ou menos especializado, também têm expressão significativa. É de salientar a diversidade de áreas de intervenção que abrangem também a indústria agro-alimentar e actividade de produção agrícola e animal.

4.2. Microcrédito vs Fame

Ainda que os dois programas de financiamento analisados tenham como público-alvo as micro organizações empresariais (que envolvem até 10 empregados), e sejam objecto de incentivo institucional, por parte da ANDC em associação com o IEFP, no caso do microcrédito, e da ADRAL em parceria com as autarquias, no caso do FAME, nos dois casos em conjunto com entidades bancárias, eles distinguem-se na forma como são implementados, no valor que acrescentam a nível pessoal e económico e no tipo indivíduos a que, efectivamente, chegam.

A génese dos dois instrumentos de financiamento é diferente. O programa de microcrédito está vocacionado para uma população relativamente mais desfavorecida, contribuindo com pequenos montantes de dinheiro (que não ultrapassam actualmente 7500 euros) para instituir ou consolidar pequenos projectos empresariais. Não exige garantias formais e depende da avaliação pessoal dos promotores. O microcrédito é muitas vezes acompanhado de outros apoios financeiros oriundos do IEFP, com o objectivo de auxiliar a criação do próprio emprego, pois muitos clientes são desempregados. O FAME, por outro lado, destina-se apenas a empresários pobres, e por isso já efectivamente envolvidos em actividades empreendedoras, em fase inicial ou não, não havendo assim um envolvimento directo com desempregados. O alcance financeiro deste programa é substancialmente maior, podendo os empréstimos atingir 45000 euros.

Esta diferença no valor médio dos empréstimos concedidos (figura 12) tem tradução directa no número de postos de trabalho criados no âmbito de cada mecanismo de microfinanciamento (figura 13). Ainda que não tenha por objectivo a substituição de subsídios de desemprego por rendimentos gerados pelas próprias pessoas, o FAME acaba por ter a este nível um efeito directo mais significativo, não só por conseguir financiar maior número de projectos, mas também porque cada um destes cria mais emprego do que nos casos em que o financiamento é concedido via ANDC.

Figura 12: Valor médio dos empréstimos

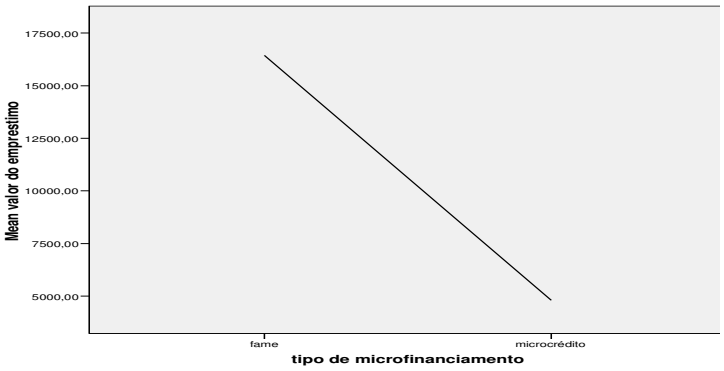
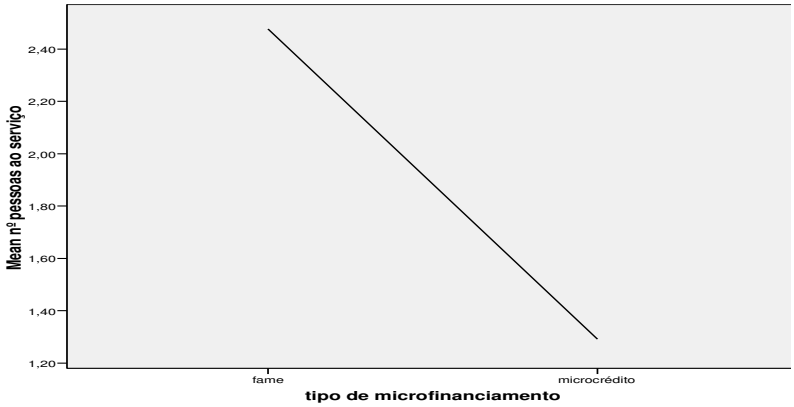


Figura 13: Postos de trabalho criados



Fonte: Elaboração própria a partir de dados recolhidos na ADRAL e na ANDC

A idade e o nível de instrução podem condicionar o perfil individual face ao risco e, consequentemente, o espírito empreendedor e o tipo de projectos criados por cada pessoa. As figuras 14 e 15 apresentam os contrastes dos dois programas de microfinanciamento nestes domínios.

Figura 14: Idade do empresário

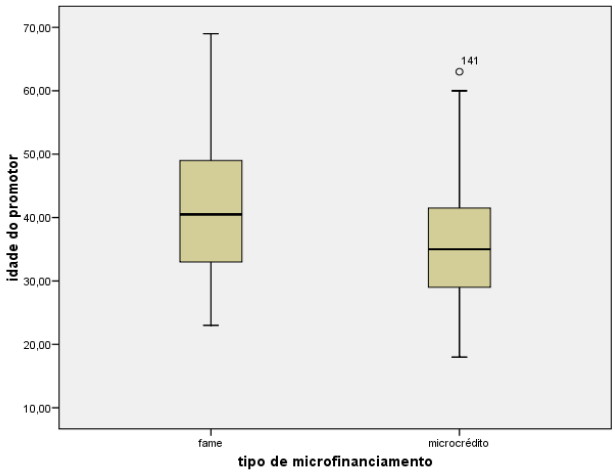
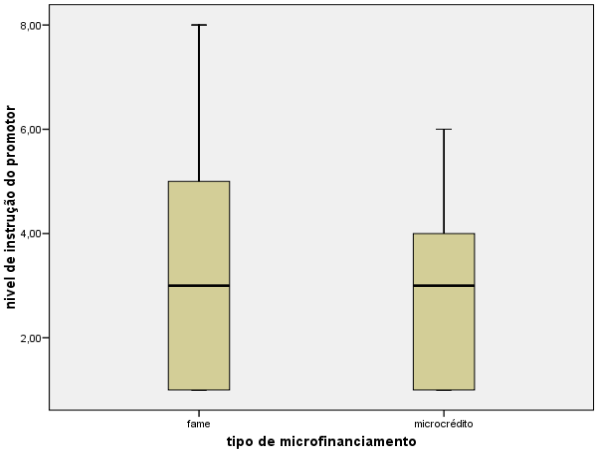


Figura 11: Nível de instrução



Fonte: Elaboração própria a partir de dados recolhidos na ADRAL e na ANDC

Os empresários envolvidos no programa FAME são em média mais velhos que os indivíduos que recorrem ao microcrédito, sendo também mais elevada a dispersão desta variável no primeiro caso. Os níveis de instrução são em ambos os casos, em média, relativamente baixos, mas o FAME consegue chegar a pessoas com mais formação. Em média os clientes do microcrédito frequentaram a escola até ao 2º ciclo do ensino básico (6º ano de escolaridade), enquanto os do FAME atingiram o 9º ano de escolaridade. Também no caso das habilitações existe maior dispersão entre os beneficiários do programa FAME, provavelmente devido ao peso mais significativo dos indivíduos licenciados.

A análise qualitativa de contraste dos dois programas desenvolvida até aqui é agora complementada com uma avaliação formal da significância estatística da diferença entre as médias (σ). A hipótese nula em teste é a ausência de diferenciação significativa entre os dois programas de microfinanciamento:

$$H0: \sigma fame = \sigma microcrédito$$
$$H1: \sigma fame \neq \sigma microcrédito \quad ,$$

Como decorre da análise da informação contida no quadro A.2 (em anexo), a hipótese nula é rejeitada, confirmando assim a conclusão informal anterior de que, no que diz respeito à idade dos clientes, à sua formação, aos montantes concedidos em empréstimo e ao número de postos de trabalho criados por cada projecto, estes dois programas de microfinanciamento são distintos.

Para consubstanciar esta abordagem de contraste quantitativo, recorre-se também à análise factorial, com o objectivo de identificar factores determinantes em ambos os casos. Os resultados são apresentados nos quadros A.3 e A4 (em anexo) e sugerem que, no caso do programa FAME, os testes de verificação da possibilidade de análise factorial são estatisticamente significativos, ainda que com um nível de correlação parcial relativamente baixo (Bartlett com signficância 0,000 e 0,5 < KMO < 0,6). Os factores comuns, isto é, a proporção da variância explicada pelas componentes principais, registam valores aceitáveis, em torno dos 60% para a maioria das variáveis (quadro A.3). A análise realizada permite-nos concluir que os três primeiros factores gerados explicam cerca de 54,2% do total da variância registada. Efectivamente, trata-se de um caso em que o poder explicativo do modelo é limitado. Ainda assim, quando verificamos as características das variáveis que estruturam os factores, constata-se que o primeiro factor está associado ao 'concelho' e ao 'ano' em que os contratos tiveram lugar. O peso deste factor reflecte a importância do concelho de Évora nos financiamentos concedidos, uma vez que este representa 1/3 do total de contratos analisados e foi o local onde se iniciou o projecto FAME. O segundo factor está associado às características dos promotores, idade e grau de instrução, que são determinantes na predisposição para o investimento. O terceiro factor é explicado pelo sexo dos proponentes, evidenciando o facto de serem homens os responsáveis por cerca de 2/3 dos projectos apoiados.

A análise factorial relativa ao microcrédito é também estatisticamente significativa, mas apenas razoavelmente adequada (Bartlett com significância 0,000 e 0,5 < KMO < 0,6). Neste caso, os factores comuns apresentam valores mais significativos (entre os 0,7 e 0,8, o que significa que as componentes principais têm aqui maior poder explicativo). Nesta análise são identificados 5 factores, que explicam cerca de 72% da variância total (quadro A.4). O primeiro factor está associado aos motivos para recurso ao microcrédito, bem como ao número de postos de trabalho criados. Preferencialmente, os contratos de microcrédito deram origem a novos negócios e, em média, criaram 1,2 postos de trabalho. O segundo factor é o estado civil dos clientes (sendo a esmagadora maioria casada), e o terceiro factor realça a importância do montante dos empréstimos e do concelho de origem de quem os solicita. O volume médio dos empréstimos é relativamente homogéneo, inferior a 5000 euros e apresenta concentração territorial no concelho de Moura. Os quarto e quinto factores estruturam-se em torno do contexto familiar, que avalia a dimensão média das famílias dos proponentes, e do negócio desenvolvido. Este, evidencia maior concentração nos domínios da restauração e da agricultura, em linha com o sugerido pela análise informal desenvolvida inicialmente.

5. CONCLUSÕES E IMPLICAÇÕES

Neste trabalho analisa-se o papel do microfinanciamento numa região portuguesa, o Alentejo, que apresenta características mistas de desenvolvimento. A exclusão das fontes tradicionais de financiamento, por incapacidade de apresentação de garantias formais, ou por necessidade de quantias ou de condições incompatíveis com a escala habitual nas instituições bancárias, impedem em todo o mundo uma percentagem substancial de indivíduos pobres de desenvolver actividades produtivas que poderiam assegurar a sua sustentabilidade económica e pessoal. Os programas de microfinanciamento procuram colmatar esta lacuna dos mercados financeiros, que descartam sistematicamente clientes e projectos potencialmente viáveis, mas revestem-se de características não caritativas que não são imediatamente óbvias, principalmente no mundo mais desenvolvido. Este facto exige um enorme esforço de adaptação de cada programa aos contextos específicos em que são aplicados, sob pena de não se produzir a totalidade dos efeitos esperados.

Os problemas de pobreza não são iguais em todo o mundo e as abordagens que devem ser adoptadas para o seu combate também diferem. Assim, o Alentejo é um espaço que desafia os promotores de microfinanciamento porque, apesar de se situar numa das zonas económicas mais ricas do mundo, apresenta níveis de pobreza elevados e características de exclusão semelhantes às que podem ser observadas em localizações geográficas menos favorecidas. Estão em funcionamento no Alentejo dois programas, o microcrédito e o FAME, que visam públicos alvo diferentes. O primeiro está direccionado para os pobres que têm capacidade de trabalho e de empreendedorismo não concretizado. O FAME visa os empresários pobres que procuram ajuda para a consolidação dos seus negócios.

Qualquer destes programas tem uma expressão relativamente reduzida no território em análise, embora em ambos os casos sejam muitas as histórias individuais que atestam o valor deste tipo de iniciativas e sugerem o potencial ainda por explorar, desde que seja possível fazer chegar a ajuda de forma efectiva àqueles que dela mais necessitam. A eficácia do microcrédito e do FAME é melhorada pela proximidade dos agentes promotores. Este aspecto, que é comum aos esquemas de microfinanciamento em todo o mundo, decorre do facto dos potenciais clientes deste tipo de serviço terem formação escolar reduzida, ou viverem situações de isolamento ou exclusão social. Qualquer destas circunstâncias diminui substancialmente a sua capacidade de obter e utilizar informações relevantes para a melhoria da sua qualidade de vida pessoal e familiar. A proximidade é também importante após a divulgação dos instrumentos disponíveis pois, principalmente no caso dos potenciais clientes do microcrédito, existem quase sempre experiências de vida povoadas de insucessos de natureza variada que tendem a aumentar os receios de implementar estratégias que envolvem risco.

A formação dos promotores de microfinanciamento é fundamental pois deles depende a triagem de projectos e de pessoas. No caso do Alentejo, a prática mostra que boa parte dos empresários financiados dedica-se a actividades com forte cariz regional. Isto é, as pessoas fazem aquilo que sabem fazer e aquilo que sempre viram fazer. É assim vital, dada a fraca densidade populacional da região e a crise que tem deprimido substancialmente a procura, que exista orientação por parte dos promotores, no sentido de realçar a importância de buscar mercados alternativos para os produtos, apostando na certificação e na publicidade.

Talvez os esquemas de grupo na concessão de microfinanciamento não façam muito sentido no Alentejo, ao contrário do que acontece noutras zonas relativamente rurais em países menos desenvolvidos. Contudo, especialmente nas áreas onde outro tipo de projectos comunitários tenham sido bem sucedidos, poderiam ser tentadas experiências piloto, que contribuiriam para aumentar o conhecimento dos promotores e das entidades responsáveis e para a sua divulgação junto de potenciais interessados.

Não é possível suprir a curto prazo as carências estruturais evidenciadas pela região Alentejo e pelas pessoas que nela residem. A realidade mostra que 20 anos de fundos comunitários, direccionados para a redução das assimetrias regionais em Portugal, não conseguiram evitar que este continue a ser o território mais pobre do país. Contudo, se as estratégias de implementação de microfinanciamento forem mais acompanhadas e complementadas com serviços de consultoria gratuitos, disponibilizados, por exemplo, através

de parcerias com as instituições de ensino superior espalhadas pelo Alentejo, é possível alargar o âmbito e melhorar os efeitos dos programas já existentes. Experiências em todo o mundo mostram como o acesso ao crédito pode transformar a vida das pessoas pobres, conferindo-lhes a autoconfiança e a sustentabilidade que faz delas seres humanos mais felizes e mais capazes de acrescentar valor. Neste domínio, o sucesso gera sucesso e o Alentejo não seria uma excepção. Sem necessidade de mais meios, mas com mais empenho e acompanhamento, também aqui será possível substituir o subsídio e a caridade pelo empreendedorismo e pela auto sustentabilidade.

5. REFERÊNCIAS

ALVES, MANUEL BRANDÃO (2006), O microcrédito, mais uma panaceia?, comunicação adaptada (Novembro de 2006) da versão apresentada no 12º Congresso da Associação Portuguesa de Desenvolvimento Regional, APDR,Viseu.

ARNOTT, R. AND STIGLITZ, J. (1991) Moral hazard and nonmarket institutions: dysfunctional crowding out or peer monitoring?, The American Economic Review, 81(1), March, pp. 179-190.

BERGER, M. (1989) Giving Women credit: the strengths and limitations of credit as a tool for alleviating poverty, World Development, Vol. 17, No. 7, pp. 1017-1032.

CHOWDHURY, P. (2005) Group-lending: sequential financing, lender monitoring and joint liability, Journal of Development Economics, 77, pp. 415-439.

COMISSÃO DE COORDENAÇÃO E DESENVOLVIMENTO REGIONAL DO ALENTEJO (2007), Plano Operacional Regional do Alentejo 2007-2013, disponível a partir de www.qren.pt (retirado em Abril de 2007).

COUTINHO, A. AND SILVA, P. (2005) Microcrédito em Portugal: uma nova oportunidade, Millennium BCP.

DALEY-HARRIS, S. (2006) State of the microcredit summit campaign report 2006, Washington D.C.: Microcredit summit campaign.

DEEP (2004) Emprego no Sector Estruturado, Ministério da Segurança Social e do Trabalho.

FRANKIEWICZ, C. (2001) Calmeadow Metrofund: A Canadian experiment in sustainable microfinance, Calmeadow Case Study, Toronto.

GOMEZ, R. AND SANTOR, E. (2003) Do peer group members outperform individual borrowers? A test of peer group lending using Canadian micro-credit data, Bank of Canada WP 2003-33.

HERMES, N. AND LENSINK, R. (2007) The empirics of microfinance: what do we know? The Economic Journal, 117, February, pp. F1-F10.

INSTITUTO NACIONAL DE ESTATÍSTICA (2007), Contas Regionais Definitivas de 2000-2003 e Preliminares de 2004, Destaque – Informação à Comunicação Social, INE.

PESTANA, MARIA H. E GAGEIRO, JOÃO N. (2000), Análise de dados para ciências sociais, a complementaridade do SPSS, 2ª edição, Lisboa, Edições Sílabo.

REGO, CONCEIÇÃO; VIEIRA, ISABEL E VIEIRA, CARLOS (2007), "Credit constraints in low density regions: microcredit in Alentejo", comunicação apresentada no XIII Congresso da APDR (Julho, Angra do Heroísmo, Portugal)

STIGLITZ (1990) Peer monitoring and credit markets, World Bank Economic Review, 4(3), pp. 351-366.

UNIVERSIDADE CATÓLICA PORTUGUESA (2007), Estudo de avaliação do sistema do microcrédito em Portugal, Mendes, Américo (coord), Universidade Católica Portuguesa e Quartenaire Portugal.

CHEN AND RAVAILLON (2008) *The developing world is poorer than we thought, but no less successful in the fight against poverty*, The World Bank Development Research Group, Policy Research Working Paper 4703.

OECD (2008) Growing Unequal? Income Distribution and Poverty in OECD Countries.

'Sítios' consultados:

http://www.oecd.org/dataoecd/48/56/41494435.pdf

http://siteresources.worldbank.org/DATASTATISTICS/Resources/WDI08supplement1216.pdf

http://www.globalissues.org/article/4/poverty-around-the-world

http://www.oecd.org/document/53/0,3343,en_2649_33933_41460917_1_1_1_1,00.html

http://web.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTPROGRAMS/EXTPOVRES/EXTPOVCALNET/0,,contentMDK:21867101~pagePK:64168427~piPK:64168435~theSitePK:5280443,00.html

Anexos

Quadro A.1: Estatísticas para as variáveis Valor do empréstimo, Idade do Promotor, Nível de Instrução do Promotor e Nº de Pessoas ao Serviço

Tipo de microfinanciamento		Valor do empréstimo	Idade do promotor	Nível de instrução do promotor	Nº postos de trabalho criados
fame	N	134	134	119	132
	Mean	16435,9663	42,3209	3,3445	2,4773
	Std. Error of Mean	923,72348	,97463	,19149	,18486
	Std. Deviation	10692,87236	11,28216	2,08886	2,12389
microcrédito	N	79	79	79	79
	Mean	4804,6110	35,9494	2,5823	1,2911
	Std. Error of Mean	101,58416	1,14778	,15031	,07461
	Std. Deviation	902,89977	10,20168	1,33597	,66317

Fonte: Elaboração própria a partir de dados recolhidos junto da ADRAL e da ANDC

Quadro A.2: Estatísticas para o teste t de comparação das médias entre as variáveis em análise

Equal variances

Quadro A.3: Análise Factorial - FAME

Communalities			Total Variance Explained				Rotated Component Matrix(a)				
	Initial	Extraction	Component	Initial Eigenvalues				Component			
				Total	% of Variance	Cumulative %		1	2	3	
valor	1	0,6356726	1	1,740689	21,75861251	21,75861251	valor	-0,7835657	-0,0706282	0,1292635	
ano_inic	1	0,5030787	2	1,4419166	18,02395715	39,78256966	ano_inic	0,6471849	0,0518472	0,285556	
concelho	1	0,6037583	3	1,1506711	14,38338836	54,16595802	concelho	0,7344822	-0,2507478	-0,0376795	
negocio	1	0,6749755	4	0,9698112	12,12263978	66,2885978	negocio	-0,0245354	0,098669	-0,8152533	
datnasc	1	0,6081231	5	0,8899802	11,12475214	77,41334994	datnasc	-0,0038643	0,7798028	0,0039673	
sexo	1	0,4475879	6	0,6778355	8,472944138	85,88629407	sexo	0,016255	0,1426821	0,653426	
nivinstr	1	0,5885846	7	0,6595605	8,244505637	94,13079971	nivinstr	-0,062782	0,7632059	0,0464737	
empreg	1	0,2714959	8	0,469536	5,869200289	100	empreg	-0,3806459	-0,3528994	-0,0454597	

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Rotation converged in 4 iterations.

Quadro A.4: Análise Factorial - Microcrédito

Communalities			Total Variance Explained				Rotated Component Matrix(a)					
	Initial	Extraction	Component	Initial Eigenvalues				Component				
				Total	% of Variance	Cumulative %		1	2	3	4	5
volempres	1	0,7951027	1	2,4119203	24,11920325	24,11920325	volempres	0,2469688	0,0661705	0,8359265	-0,0371895	0,1719721
concelho	1	0,6714713	2	1,4584283	14,58428295	38,70348621	concelho	-0,4030773	-0,0325619	0,6712175	-0,0226559	-0,2385233
negócio	1	0,79078	3	1,2368939	12,36893905	51,07242525	negócio	-0,0197288	0,090438	-0,004175	0,1467329	0,8721604
datnasc	1	0,6687676	4	1,0714577	10,71457715	61,78700241	datnasc	-0,2228072	-0,7353175	0,0591141	0,0496961	-0,2691998
nivinstrc	1	0,7255375	5	1,0086555	10,08655492	71,87355732	nivinstrc	-0,2234534	-0,4450469	0,3976519	-0,5631012	0,048263
outform	1	0,6250593	6	0,7177864	7,177863596	79,05142092	outform	0,5148407	0,2539119	-0,4312504	-0,2419289	0,2258773
estadcivil	1	0,7786971	7	0,6504233	6,504232895	85,55565382	estadcivil	-0,1641587	0,8469159	0,0690996	0,1467133	-0,0904601
sitfamil	1	0,8332378	8	0,5490109	5,490108655	91,04576247	sitfamil	0,0591856	0,021078	0,0555987	0,8935606	0,1665796
motivos	1	0,6259053	9	0,5291843	5,291843379	96,33760585	motivos	0,7486185	-0,0489639	0,0388192	0,1832211	0,1673356
npesserv	1	0,6727973	10	0,3662394	3,66239415	100	npesserv	0,7425894	0,0574421	-0,0135615	0,0329416	-0,341745

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Rotation converged in 6 iterations.

ESTIMACIÓN EN PEQUEÑAS ÁREAS EN EL ÁMBITO DE LA COMUNIDAD VALENCIANA

SANTIAGO MURGUI IZQUIERDO
E-mail: Santiago.Murgui@uv.es
CRISTINA AYBAR ARIAS
E-mail: Cristina.Aybar@uv.es
Mª CONSUELO COLOM ANDRÉS
E-mail: Consuelo.Colom@uv.es
Mª CRUZ MOLÉS MACHÍ
E-mail: Cruz.Moles@uv.es
Departamento de Economía Aplicada
Universitat de València

Resumen

En este trabajo, promovido y desarrollado con la colaboración del Instituto Valenciano de Estadística de la Generalitat Valenciana, se plantea una propuesta de estimación en pequeñas áreas de la Comunidad Valenciana, en el entorno contextual de la Encuesta Industrial. Considerando la información disponible en la propia encuesta y otra información auxiliar obtenida a partir del DIRCE, ambas referidas al año 2005, la propuesta se concibe bajo el dominio de varios modelos alternativos. De forma particular se han contemplado modelos de efectos fijos y de efectos aleatorios. Los resultados que se ofrecen se refieren a estimaciones provinciales en el contexto de las variables de más relevancia, destacando las preferencias de un modelo sobre otro y los correspondientes errores asociados. Las estimaciones obtenidas se armonizan con los resultados globales para la Comunidad Valenciana obtenidos por el INE a partir del diseño aleatorio de la Encuesta Industrial.

Palabras Clave: estimación en pequeñas áreas, modelo de efectos fijos y efectos aleatorios.
Área temática: Economía Urbana, Regional y Local.

1. INTRODUCCIÓN

Entre los distintos productos que habitualmente viene desarrollando el Instituto Nacional de Estadística –INE- con referencia a estadísticas económicas, cabe destacar la Encuesta Industrial. Se trata de una encuesta de periodicidad anual, basada en un diseño muestral aleatorio estratificado, que considera unidades estadísticas a las empresas cuya actividad principal se incluye dentro de las Secciones C y E de la CNAE-93. Su objetivo es la obtención de estimaciones anuales para los agregados de las principales variables estructurales y de actividad de los sectores industriales.

La necesidad de obtener estimaciones para las distintas ramas de actividad desagregadas entre las distintas regiones o comunidades autónomas, ha conducida a contemplar junto a las empresas otras unidades en la encuesta, principalmente los establecimientos. Se define así a las unidades situadas en un lugar geográficamente delimitado en las que se realizan actividades económicas por cuenta de una única empresa.

El diseño y el tamaño de la muestra soporte de la Encuesta Industrial han sido fijados con el objetivo de obtener estimaciones fiables desagregadas, tanto en el ámbito espacial de las comunidades autónomas como por sectores o ramas de actividad. Sin embargo, el creciente interés que manifiestan los organismos de naturaleza pública y privada por acceder a información más específica y localizada, plantea la necesidad de disponer de estimaciones con una referencia espacial menor a la de las comunidades autónomas. Este es el reto que se propone la metodología estadística que analiza la estimación en pequeñas áreas, la obtención de resultados fiables referidos a las provincias, comarcas o agrupaciones comarcales.

Como es de prever, una mayor desagregación de las estimaciones implicará en general una pérdida no deseada de precisión estadística en los resultados. Este inconveniente puede ser eliminado, o cuanto menos controlado, recurriendo a dos posibles estrategias: incrementar los tamaños muestrales inicialmente fijados y/o incorporar información complementaria relevante al proceso inferencial.

La primera de las alternativas señaladas, la ampliación de la muestra, necesariamente conlleva un incremento de los costes de la operación, los que se generarán en el trabajo de campo y los que se deriven del cálculo de nuevos factores de elevación.

En este trabajo, promovido y desarrollado con el apoyo del Instituto Valenciano de Estadística de la Generalitat Valenciana, se persigue la estimación en pequeñas áreas de la Comunitat Valenciana respetando el actual diseño de la Encuesta Industrial realizado por el INE, por lo que ineludiblemente será necesario apelar a técnicas estadísticas capaces de aprovechar las ventajas que reporte la incorporación de información complementaria.

La variable auxiliar que se ha considerado para incorporarla al proceso inferencial es el número de empleados remunerados de los establecimientos, cuyos valores censales son accesibles a través de la información que contiene el DIRCE.

Aunque el proyecto tiene unos objetivos más amplios, esta comunicación analiza la estimación referida al ejercicio 2005 para tres variables básicas, agregadas para las tres provincias que forman la comunidad: Alicante, Castellón y Valencia. Las citadas variables son: cifra de negocio, total consumo y gastos de personal. Asimismo, se van a considerar estimaciones específicas para cada uno de los 13 sectores industriales siguientes:

r1 energía, aguas, extractivas
r2 alimentación, bebidas, tabaco
r3 textil, confección, cuero, calzado
r4 madera
r5 papel y cartón, artes gráficas y edición
r6 química
r7 caucho y plástico
r8 productos minerales no metálicos
r9 metalurgia y productos metálicos
r10 maquinaria y equipo mecánico
r11 material y equipo eléctrico, electrónico y óptico
r12 material de transporte
r13 manufacturas diversas

2. MODELO DE EFECTOS FIJOS Y ESTIMACIONES

Las particulares connotaciones de tipo teórico y práctico anteriormente señaladas han conducido a centrar el estudio en las técnicas de estimación basadas en modelos, y de forma particular en técnicas de estimación con modelos de efectos fijos y con modelos de efectos mixtos.

Conviene señalar, antes de pasar a la especificación de los modelos, que la información disponible acerca de la variable auxiliar viene referida a los establecimientos que realizan actividad industrial, y que razones que apelan al secreto estadístico hacen imposible conocer la identificación de las empresas a las que pertenece cada uno. En estas circunstancias, independientemente de que el diseño muestral de la Encuesta Industrial utiliza como unidad de muestreo la empresa y que los consiguientes factores de elevación están también referidos a estas entidades, en lo que sigue y de cara a la estimación en pequeñas áreas adoptaremos como unidad de muestreo los establecimientos. Lógicamente esta forma de proceder hace inviable apelar a la estructura estocástica derivada del diseño aleatorio de la muestra para determinar la fiabilidad de los estimadores que se propongan.

En realidad, la imposibilidad de apelar al diseño aleatorio no supone ningún inconveniente, puesto que para una gran parte de los sectores que se consideran los tamaños muestrales restringidos a cada provincia son excesivamente pequeños y en consecuencia no permiten garantizar un nivel suficiente de precisión para las estimaciones que podrían calcularse a partir de la estructura estocásticas asociada con el diseño muestral.

Formularemos en primer lugar el modelo y los principales resultados inferenciales asociados cuando únicamente se consideran efectos fijos.

Si denotamos por Y a cualquiera de las variables de interés y por X la variable auxiliar que define el número de empleados remunerados, el modelo que se propone es el que determina la ecuación:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \beta + \boldsymbol{\varepsilon} \quad \text{con} \quad \boldsymbol{\varepsilon} \rightarrow N\left[\mathbf{0}, \sigma^2 \begin{pmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_N \end{pmatrix} \right]$$

O bien de forma equivalente $\mathbf{y} = \mathbf{x}\beta + \quad$, donde β y σ^2 son parámetros a estimar.

La formulación del modelo puede presentarse de forma compuesta distinguiendo sobre el universo global U el conjunto de establecimientos *s* que constituyen la muestra, para los cuales se dispone de información acerca de la variable de interés, y el resto de los establecimientos *r*:

$$\begin{pmatrix} \mathbf{y}_s \\ \mathbf{y}_r \end{pmatrix} = \begin{pmatrix} \mathbf{x}_s \\ \mathbf{x}_r \end{pmatrix} \beta + \begin{pmatrix} \boldsymbol{\varepsilon}_s \\ \boldsymbol{\varepsilon}_r \end{pmatrix}$$

En este modelo, las componentes del vector $\mathbf{y}_s$ son proporcionadas por la Encuesta Industrial, mientras que todas las componentes de $\mathbf{x}$, tanto las de $\mathbf{x}_s$ como las de $\mathbf{x}_r$, se obtienen a partir del DIRCE.

Conviene ahora hacer una matización acerca de las diferencias en la configuración de la muestra que utiliza la Encuesta Industrial y la que se considera para la estimación en base al modelo. En el caso de la Encuesta Industrial, la investigación sobre los estratos formados por las empresas con 20 o más trabajadores se hace de forma exhaustiva, de manera que la estimación de agregados poblacionales se construye sumando a los agregados determinados sobre estos estratos la estimación obtenida sobre el estrato de empresas con menos de 20 trabajadores.

Cuando las estimaciones se determinan a partir del modelo, la unidad de muestreo es el establecimiento y no pueden acumularse los datos para reconvertir dichas unidades en empresas porque la información proporcionada desde el DIRCE no permite su identificación. De esta forma, está garantizado que todos los establecimientos con 20 o más trabajadores forman parte de una empresa con 20 o más trabajadores, esto es, pertenecen a estratos investigados exhaustivamente. Sin embargo, es posible que parte de los establecimientos con menos de 20 trabadores pertenezcan a empresas que en su conjunto sí posean 20 o más trabajadores por lo que en realidad forman parte de estratos exhaustivamente analizados.

En la inferencia construida a partir del modelo la estructura estocástica que se considera es la que define el propio modelo, no teniendo ninguna relevancia el diseño muestral. En estas circunstancias, poco importa el que un establecimiento aparezca en la muestra con probabilidad 1 o inferior a 1, lo único relevante es conocer sobre qué establecimientos se dispone de información.

Denotando por a=1,2,3 a las tres provincias de la Comunitat Valenciana, en la estimación de los totales asociados con la variable Y a nivel provincial T_{ya} se demuestra que el estimador lineal insesgado con mínimo error cuadrático es uno de tipo predictivo, siendo su expresión algebraica:

$$\hat{T}_{ya} = \sum_{i=1}^{n_a} y_{ai} + \sum_{i=1}^{N_a - n_a} x_{ai}\,\hat{\beta} = \sum_{i=1}^{n_a} y_{ai} + \sum_{i=1}^{N_a - n_a} x_{ai} \frac{\sum_{i \in s} y_i}{\sum_{i \in s} x_i} = t_{yas} + t_{xar} \frac{t_{ys}}{t_{xs}}$$

Donde N_a denota el número total de establecimientos de la provincia a, n_a el número de establecimientos incluidos en la muestra de a, t_{yas} el total agregado muestral para la variable Y en la provincia a, t_{xar} el total agregado de X en la provincia a fuera de la muestra, t_{ys} el total agregado de Y en todo el conjunto muestral de las tres provincias y t_{xs} el total agregado de X en todo el conjunto muestral de las tres provincias.

El error cuadrático asociado con este estimador viene determinado por $E\left[\left(\hat{T}_{ya} - T_{ya}\right)^2\right] = \sigma^2 t_{xar}\left(\frac{t_{xar}}{t_{xs}} + 1\right)$, donde el parámetro σ^2 admite un estimador insesgado definido por la expresión $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i\in s}\frac{\left(y_i - \hat{\beta}\, x_i\right)^2}{x_i}$, donde $\hat{\beta}$ es el estimador lineal insesgado óptimo del parámetro β determinado por la expresión $\hat{\beta} = \frac{\sum_{i\in s} y_i}{\sum_{i\in s} x_i} = \frac{t_{ys}}{t_{xs}}$.

Antes de ofrecer los resultados provinciales obtenidos con la mitología descrita, en la Tabla 1 se ofrecen los tamaños del universo y la muestra de referencia para cada sector y provincia.

TABLA 1

Universo y muestra de referencia para la estimación con establecimientos

r13	emp>=20 censales (1) EI	emp<20 muestra (na) (2) EI	L1+L2+L3 referencia (Na) (3) DIRCE	universo de referencia global (1)+ (3)	fracción muestral
ALACANT					
1	24	**34**	186	**210**	27,62
2	97	**59**	634	**731**	21,34
3	401	**262**	3146	**3547**	18,69
4	35	**35**	406	**441**	15,87
5	58	**37**	525	**583**	16,30
6	25	**18**	99	**124**	34,68
7	72	**22**	442	**514**	18,29
8	100	**58**	534	**634**	24,92
9	59	**57**	1184	**1243**	9,33
10	43	**19**	369	**412**	15,05
11	14	**46**	177	**191**	31,41
12	12	**11**	94	**106**	21,70
13	67	**44**	612	**679**	16,35
CASTELLON					
1	11	**18**	203	**214**	13,55
2	15	**35**	226	**241**	20,75
3	33	**12**	175	**208**	21,63
4	14	**24**	239	**253**	15,02
5	17	**3**	146	**163**	12,27
6	40	**12**	54	**94**	55,32
7	6	**0**	51	**57**	10,53
8	249	**54**	338	**587**	51,62
9	35	**22**	379	**414**	13,77
10	48	**13**	240	**288**	21,18
11	7	**11**	58	**65**	27,69
12	9	**4**	39	**48**	27,08
13	56	**16**	183	**239**	30,13
VALENCIA					
1	27	**45**	351	**378**	19,05
2	169	**95**	878	**1047**	25,21
3	166	**108**	1320	**1486**	18,44
4	117	**84**	968	**1085**	18,53
5	113	**49**	1026	**1139**	14,22
6	85	**40**	321	**406**	30,79
7	92	**14**	388	**480**	22,08
8	138	**100**	760	**898**	26,50
9	233	**111**	1786	**2019**	17,04
10	107	**53**	675	**782**	20,46
11	78	**115**	503	**581**	33,22
12	43	**32**	130	**173**	43,35
13	242	**113**	1874	**2116**	16,78

En una primera columna se recoge el número de establecimientos con 20 o más trabajadores en las tres provincias, para cada uno de los sectores considerados. En segundo lugar se recoge el número de establecimientos con menos de 20 empleados, obtenido desde la información muestral que ofrece la Encuesta Industrial. En la tercera columna se indica el número total de establecimientos con menos de 20 trabajadores que contiene el DIRCE. De esta forma, puesto que los datos de las columnas 1 y 3 son censales, su suma recogida en la columna 4 ofrecerá el tamaño del universo global de establecimientos. Finalmente, la suma de los datos contenidos en las dos primeras columnas ofrecerá la muestra completa y su cociente con respecto al tamaño del universo expresará la fracción de muestreo global recogida en la última columna.

En la Tabla 2 se ofrecen las estimaciones para cada sector económico y provincia de las tres variables analizadas. Los cálculos se han efectuado a partir de la expresión algebraica anterior, utilizando el programa informático SAS, y vienen expresados en millones de euros.

TABLA 2

Estimaciones provisionales provinciales basadas en el modelo de efectos fijos

ALACANT

r13	Cifra negocios	Total consumo	Gastos personal
1	1320	580	129
2	1780	1128	211
3	3447	2226	576
4	406	254	73
5	820	389	141
6	223	110	40
7	785	433	150
8	1895	984	276
9	1251	725	208
10	383	201	98
11	208	115	43
12	104	61	19
13	697	379	128

CASTELLON

r13	Cifra negocios	Total consumo	Gastos personal
1	4436	4034	78
2	477	359	40
3	315	178	71
4	175	113	31
5	273	151	51
6	1710	966	200
7	67	37	16
8	4603	2013	916
9	493	278	97
10	392	214	99
11	83	44	17
12	168	83	38
13	555	252	121

VALENCIA

r13	Cifra negocios	Total consumo	Gastos personal
1	2755	1986	203
2	5595	3713	579
3	1854	1115	369
4	1290	797	237
5	1313	650	284
6	1352	771	217
7	1363	757	277
8	2484	1226	388
9	3442	2013	579
10	1171	644	286
11	1160	713	236
12	7006	5630	532
13	2021	1119	466

Las estimaciones provinciales obtenidas deben considerarse provisionales, a falta de efectuar una corrección posterior. El objeto de tal corrección será asegurar la armonización entre las estimaciones provinciales que finalmente se propongan y las estimaciones obtenidas por el INE para toda la Comunitat Valenciana en su conjunto.

Agregando para cada variable y sector los resultados provisionales de las tres provincias se obtienen estimaciones sectoriales para toda la Comunitat Valenciana. En la Tabla 3 se ofrecen dichos valores agregados.

TABLA 3
Estimaciones provisionales agregadas para toda la Comunitat Valenciana

r13	Cifra negocios	Total consumo	Gastos personal
1	8511	6600	410
2	7851	5200	831
3	5616	3519	1017
4	1871	1164	342
5	2406	1190	475
6	3285	1848	457
7	2215	1227	443
8	8983	4223	1580
9	5186	3016	885
10	1946	1058	484
11	1451	872	297
12	7277	5774	588
13	3274	1750	715

Planteemos ahora la corrección a las estimaciones provinciales a las que anteriormente se hacía referencia. La Encuesta Industrial a empresas que desarrolla el INE está diseñada para conseguir en cada sector estimaciones fiables a nivel nacional y autonómico. En estas circunstancias, resulta obligado asumir la validez de las estimaciones que proporciona para todas las variables en la Comunitat Valenciana. Por otra parte, es perfectamente normal que las estimaciones recogidas en la Tabla 3, también referidas a toda la comunidad autónoma, difieran de las que se derivan de la Encuesta Industrial. La desviación entre ambas estimaciones esta justificada en base a dos razones, la más importante es la diferente estructura estocástica que da soporte a los cálculos, basada en el diseño aleatorio en un caso y en el modelo de superpoblación en el otro. La otra razón, la forma de acceder a los agregados autonómicos,

directa en el caso de la Encuesta Industrial y a través de la agregación de estimaciones provinciales cuando nos apoyamos en el modelo.

En la Tabla 4 se recogen las estimaciones para la Comunitat Valenciana obtenidas a partir de la Encuesta Industrial utilizando los factores de elevación que esta proporciona.

TABLA 4

Estimaciones para toda la Comunitat Valenciana obtenidas a partir de la EI

r13	Cifra negocios	Total consumo	Gastos personal
1	8109	6400	380
2	7149	4588	819
3	5213	3132	1049
4	1681	1020	330
5	2293	1090	464
6	3198	1822	449
7	2153	1189	422
8	8037	3664	1566
9	4377	2610	848
10	1951	1033	487
11	1422	838	297
12	7281	5784	588
13	3107	1639	690

En la Tabla 5 se han calculado la desviaciones relativas entre ambas estimaciones, las derivadas a partir de la Encuesta Industrial y las calculadas mediante la agregación de las estimaciones provinciales con base en el modelo. El resultado, en términos porcentuales, es un cociente con numerador la diferencia entre ambas estimaciones y como denominador la estimación con base en el modelo.

TABLA 5

Desviaciones relativas porcentuales entre las estimaciones agregadas para toda la Comunitat Valenciana obtenidas desde los agregados provinciales provisionales en base al modelo y las estimaciones de la EI

r13	Cifra negocios	Total consumo	Gastos personal
1	4,73	3,03	7,25
2	8,95	11,75	1,45
3	7,17	11,01	-3,11
4	10,16	12,34	3,52
5	4,71	8,33	2,41
6	2,65	1,39	1,84
7	2,79	3,07	4,78
8	10,52	13,23	0,88
9	15,60	13,47	4,18
10	-0,30	2,40	-0,55
11	1,39	3,39	-0,96
12	0,06	-0,10	0,39
13	5,10	6,33	3,53

En general las estimaciones del INE adoptan un valor inferior al que se obtiene a partir del modelo. La magnitud de las desviaciones relativas encontradas, aunque en algún caso no es despreciable, en general puede ser justificada a partir del error de muestreo. De hecho, la mayor parte de tales desviaciones se encuentran comprendidas dentro del intervalo del 95% de probabilidad.

Al objeto de preservar la armonización de los resultados que se va a ofrecer para áreas pequeñas con los calculados por el INE a nivel de toda la comunidad, se ha procedido a corregir convenientemente las estimaciones provinciales recogidas en la Tabla 2. De esta forma se garantiza la igualdad entre las cifras globales obtenidas para cada sector y variable a nivel de toda la Comunitat Valenciana por una y otra vía de estimación.

Si se denota por $\hat{T}_y$ la estimación obtenida para toda la comunidad en base al modelo y por T_y^e la obtenida a partir de la Encuesta Industrial, la estimación final corregida del agregado de cada provincia *a* se determina mediante la expresión $\tilde{T}_{ya} = \hat{T}_{ya} \frac{T_y^e}{\hat{T}_y}$.

En la Tabla 6 se ofrecen las estimaciones finales corregidas para las tres provincias, los 13 sectores y las tres variables analizadas. Asimismo, para facilitar su interpretación se han calculado las respectivas fracciones porcentuales con respecto a la cifra global de toda la Comunitat Valenciana.

TABLA 6
Estimaciones provinciales corregidas en base al modelo de efectos fijos

ALACANT

r13	Cifra negocios	% provincial	Total consumo	% provincial	Gastos personal	% provincial	% empleo
1	1257	15.51	562	8.79	119	31.42	33.08
2	1620	22.67	995	21.69	208	25.44	26.30
3	3199	61.38	1981	63.25	594	56.66	59.88
4	365	21.70	223	21.86	71	21.46	22.28
5	782	34.09	356	32.69	137	29.61	30.20
6	217	6.78	109	5.98	39	8.72	10.93
7	763	35.45	419	35.28	143	33.91	36.64
8	1695	21.09	853	23.29	273	17.46	17.41
9	1056	24.12	628	24.05	199	23.51	24.81
10	384	19.67	196	18.95	99	20.34	21.19
11	205	14.36	111	13.21	44	14.56	15.10
12	104	1.42	610	1.05	18	3.15	5.36
13	662	21.30	355	21.67	123	17.86	17.32
total	12309	21.99	6850	19.68	2069	24.67	27.94

CASTELLON

r13	Cifra negocios	% provincial	Total consumo	% provincial	Gastos personal	% provincial	% empleo
1	4226	52.12	3911	61.12	72	18.95	18.94
2	434	6.07	317	6.90	40	4.85	6.02
3	292	5.61	159	5.06	74	7.02	6.95
4	157	9.36	99	9.69	30	9.17	9.32
5	260	11.35	138	12.70	49	10.65	9.97
6	1665	52.07	953	52.30	197	43.79	38.54
7	65	3.00	36	3.03	15	3.50	3.64
8	4119	51.25	1747	47.68	908	57.96	56.03
9	417	9.52	241	9.22	93	10.99	10.17
10	393	20.14	209	20.23	100	20.53	20.16
11	81	5.69	42	5.03	17	5.82	6.00
12	168	2.30	83	1.44	38	6.46	8.54
13	527	16.96	236	14.40	117	16.93	14.90
total	12804	22.88	8171	23.47	1749	20.86	18.33

VALENCIA

r13	Cifra negocios	% provincial	Total consumo	% provincial	Gastos personal	% provincial	% empleo
1	2625	32.37	1926	30.09	189	49.62	47.97
2	5094	71.26	3277	71.41	571	69.70	67.68
3	1721	33.02	992	31.68	381	36.32	33.18
4	1159	68.94	698	68.45	229	69.37	68.41
5	1251	54.56	596	54.61	277	59.74	59.83
6	1316	41.15	760	41.72	213	47.49	50.54
7	1325	61.55	734	61.69	264	62.59	59.72
8	2223	27.66	1064	29.03	385	24.58	26.55
9	2905	66.36	1742	66.73	555	65.50	65.02
10	1174	60.19	628	60.82	288	59.13	58.65
11	1137	79.95	685	81.76	237	79.62	78.90
12	7010	96.27	5640	97.52	532	90.39	86.10
13	1918	61.74	1048	63.93	450	65.21	67.78
total	30858	55.13	19790	56.85	4570	54.48	53.72

La interpretación de los resultados que recoge la Tabla 6 a partir de las fracciones porcentuales es inmediata. Para el conjunto de los 13 sectores, la provincia de Valencia acapara el 53% del empleo y esto se traduce en proporciones del mismo orden para los agregados de las tres variables consideradas, Alicante aglutina el 28% del empleo y Castellón el 18%. Sin embargo, esta distribución no resulta uniforme para los 13 sectores, de hecho Valencia acumula prácticamente la totalidad regional en el sector r12 (material de transporte), Alicante domina claramente en el sector r3 (textil, confección, cuero, calzado) y Castellón domina en el sector r8 (productos minerales no metálicos). Lógicamente, la importancia de los resultados no se centra en las cifras de empleo que ya eran conocidas, sino en las de las variables analizadas.

3. FIABILIDAD DE LAS ESTIMACIONES BASADAS EN MODELOS DE EFECTOS FIJOS

Para determinar la fiabilidad de las estimaciones anteriores se ha procedido al cálculo del error cuadrático y el correspondiente coeficiente de variación, definido mediante el cociente entre la raíz del error cuadrático medio y el propio valor de la estimación. Los cálculos se han efectuado con el mismo programa informático ya señalado y los resultados se recogen en la Tabla 7.

TABLA 7
Error relativo (corrigiendo predicción)

ALACANT

r13	Cifra negocios	Total consumo	Gastos personal
1	4,19	4,16	0,94
2	5,04	7,20	1,01
3	4,75	6,99	0,83
4	4,89	6,62	2,44
5	2,40	3,58	1,64
6	7,53	7,15	2,39
7	5,34	8,54	2,76
8	3,24	4,12	1,18
9	11,50	9,13	1,58
10	6,86	12,36	1,91
11	2,59	3,79	1,27
12	8,13	9,01	3,85
13	3,20	4,54	1,05

CASTELLON

r13	Cifra negocios	Total consumo	Gastos personal
1	0,89	0,43	1,11
2	8,68	10,44	2,44
3	6,34	10,64	0,82
4	7,28	9,58	3,66
5	2,24	2,86	1,41
6	0,46	0,38	0,22
7	11,50	18,10	4,87
8	0,99	1,49	0,26
9	10,97	8,97	1,27
10	5,43	9,39	1,54
11	3,80	5,78	1,84
12	3,20	4,21	1,20
13	1,81	3,07	0,50

VALENCIA

r13	Cifra negocios	Total consumo	Gastos personal
1	2,62	1,58	0,77
2	2,20	2,99	0,50
3	4,11	6,49	0,60
4	3,41	4,69	1,67
5	2,82	4,03	1,53
6	3,29	2,71	1,16
7	2,71	4,30	1,32
8	3,11	4,16	1,06
9	6,77	5,33	0,92
10	4,46	7,66	1,31
11	1,27	1,67	0,63
12	0,15	0,12	0,17
13	3,53	4,91	0,92

Las cifras obtenidas para el error de estimación en la provincia de Valencia se sitúan en valores bastante contenidos para lo que es habitual en las investigaciones muestreo, por lo que en términos generales, la fiabilidad de las estimaciones puede considerarse aceptable. En el caso de Alicante y de manera particular en Castellón, aparecen algunos valores aislados ya más elevados. En ocasiones, el motivo de este incremento en el error está en el escaso tamaño de la muestra, sin embargo otras veces la causa parece estar relacionada con el grado de ajuste del modelo propuesto.

Cuando el modelo propuesto no permite asignar niveles reducidos para el error de la estimación, caben dos posibles alternativas: que el modelo sea el adecuado pero presente una elevada varianza, o bien que existan valores anómalos. En el primer caso, salvo que se decidiera incrementar el tamaño de la muestra no parece posible conseguir un incremento de la fiabilidad de los resultados. En el segundo, la detección de los posibles valores anómalos y el consiguiente tratamiento diferencial de los mismos sí puede repercutir en el error de las estimaciones agregadas.

En la actualidad estamos ensayando la aplicación de un test de detección de valores anómalos para poblaciones finitas y evaluando las posibles alternativas de estimación que se ofrecen, dependiendo del tratamiento que se conceda a estos valores.

4. MODELO DE EFECTOS MIXTOS Y EXTENSIONES

En la literatura habitual sobre estimación en pequeñas áreas es muy frecuente la utilización de un modelo que junto a los efectos fijos incorpora un efecto aleatorio asociado con cada área.

El modelo mixto admite la existencia de unos efectos fijos determinados por dos parámetros β_0 y β_1 comunes a todos los establecimientos y un efecto aleatorio, específico para todos los establecimientos de una misma área. Siguiendo la notación introducida en el modelo de efectos fijos, denotemos por y_{ai} al valor que toma la variable de interés en el establecimiento *i* del área *a*, por η_a el efecto aleatorio del área *a* y por ε la perturbación aleatoria.

En forma analítica el modelo puede especificarse a través de las ecuaciones:

$$y_{ai} = \beta_0 + \beta_1 x_{ai} + \eta_a + \varepsilon_{ai} \quad \text{para} \quad a = 1,\ldots,A \quad i = 1,\ldots,N_a$$

asumiendo las siguientes hipótesis sobre las componentes aleatorias:

$$\eta_a \to N\left[0, \sigma_1^2\right] \quad \text{y} \quad \varepsilon_{ai} \to N\left[0, \sigma_2^2\, x_{ai}\right]$$

En lenguaje matricial el modelo anterior se puede expresar como:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Z\eta} + \boldsymbol{\varepsilon}$$

siendo:

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_A \end{pmatrix}_N \quad \mathbf{X} = \begin{pmatrix} \mathbf{1}_1 & \mathbf{x}_1 \\ \vdots & \vdots \\ \mathbf{1}_A & \mathbf{x}_A \end{pmatrix}_{N\cdot 2} \quad \mathbf{Z} = \begin{pmatrix} \mathbf{1}_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{1}_A \end{pmatrix}_{N\cdot A}$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \quad \boldsymbol{\eta} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_A \end{pmatrix}_A \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_A \end{pmatrix}_N$$

donde:

$$\mathbf{y}_a = \begin{pmatrix} y_{a1} \\ \vdots \\ y_{aN_a} \end{pmatrix}_{N_a} \quad \mathbf{x}_a = \begin{pmatrix} x_{a1} \\ \vdots \\ x_{aN_a} \end{pmatrix}_{N_a} \quad \mathbf{1}_a = \begin{pmatrix} 1_{a1} \\ \vdots \\ 1_{aN_a} \end{pmatrix}_{N_a} \quad \boldsymbol{\varepsilon}_a = \begin{pmatrix} \varepsilon_{a1} \\ \vdots \\ \varepsilon_{aN_a} \end{pmatrix}_{N_a}$$

Siguiendo la metodología general de minimización del error cuadrático, es posible determinar el estimador lineal óptimo asociado con este modelo, así como su error cuadrático medio, cuyo valor únicamente es accesible de manera aproximada. Sus expresiones son bastante complejas y pueden verse recogidas en la literatura citada en las referencias bibliográficas en diversas formas. Hemos optado por no incluirlas en esta comunicación por una razón fundamental, las estimaciones provinciales en el apartado anterior obtenidas a partir de un modelo de efectos fijos no mejoran, al contrario, pierden precisión, si se considera como soporte el modelo mixto.

La forma analítica en que se ha constatado esto último es a través de la estimación de la varianza σ_1^2. Se comprueba que, prácticamente en todos los casos, la citada varianza es no significativa. En otras palabras, sobre el modelo estimado para cada variable y sector, no parece existir un efecto aleatorio real asociado a las provincias.

Una posible interpretación de esta singularidad estaría relacionada con las características propias de las tres provincias de la comunidad, las tres tienen una parte litoral y otra de interior, y las tres presentan un nivel de desarrollo económico bastante similar. Tal como hemos comprobado posteriormente, al calcular estimaciones para diferentes comarcas de la Comunitat

Valenciana, esta singularidad desaparece en muchos casos y el modelo mixto adquiere un papel relevante que deberá ser evaluado.

Siguiendo con el proyecto global de estimación para pequeñas áreas de la Comunitat Valenciana, en estos momentos estamos analizando resultados por comarcas y agrupaciones comarcales. Es inmediato comprobar que en muchos sectores y comarcas la muestra disponible es muy reducida o incluso carece de representación. Lógicamente, la estimación de agregados en estos casos es improcedente.

La necesidad de hacer agregaciones sobre la estructura espacial y sectorial abre una doble vía: o bien se agregan comarcas para asegurar una representación muestral mínima, o bien, se agregan sectores. En ambos casos, existen distintos criterios para efectuar las agregaciones y es este uno de los problemas que conjuntamente con los técnicos del Instituto Valenciano de Estadística venimos analizando

5. CONCLUSIONES

a) Utilizando las técnicas de estimación en pequeñas áreas, se han obtenido resultados provinciales agregados para cada uno de los trece sectores industriales descritos y tres variables relevantes.

b) El procedimiento empleado se apoya en la información censal que proporciona el DIRCE acerca de la variable que indica el empleo remunerado.

c) La utilización conjunta de información procedente de la Encuesta Industrial y del DIRCE presenta algunos problemas de homologación que se han resuelto buscando los puntos convergentes.

d) Los primeros resultados obtenidos presentan ligeros problemas de compatibilidad con los que ofrece el INE para toda la Comunitat Valenciana. Algo lógico si se considera que el primero basa sus argumentos en el diseño aleatorio de la muestra, mientras que los segundos apelan a un modelo de superpoblación que se apoya en información auxiliar.

e) Las correcciones introducidas en las estimaciones están dirigidas a resolver este problema, asegurando la armonización entre las estimaciones provinciales y las globales ofrecidas por el INE.

f) Aunque es Valencia la provincia que aglutina más del 50% de la actividad en el conjunto de la comunidad y en el contexto global de los 13 sectores analizados, la distribución entre los sectores no es uniforme, de hecho existe un sector en el que respectivamente domina cada una de las otras provincias.

REFERENCIAS BIBLIOGRÁFICAS

EUSTAT (2006). Estimación de áreas pequeñas en la Encuesta Industrial de la C. A. de Euskadi.

KACKAR, R.N. Y D.A. HARVILLE (1984). Approximations for standard errors of estimators of fixed and random effects in mixed linear models. */Journal of the American Statistical Association/*, vol. 79, núm. 388, pp. 853-862.

PRASAD, N.G.N. Y J.N.K. RAO (1990). The estimation of the mean squared error of small area estimators. /Journal of the American Statistical Association/, vol. 85, núm. 409, pp. 163-171.

RAO, J.N.K. (2003). Small Area Estimation. Wiley Series in Survey Methodology.

TURISMO E DESENVOLVIMENTO ECONÓMICO LOCAL: SEGMENTAÇÃO DO MERCADO TURÍSTICO DA COVILHÃ DE ACORDO COM AS DESPESAS DOS VISITANTES NO DESTINO

JOANA INÊS SILVA DE LIMA
E-mail: jisl@ua.pt
MARIA CELESTE DE AGUIAR EUSÉBIO
E-mail: maguiar@egi.ua.pt
Universidade de Aveiro

Resumo

O objectivo deste estudo consiste na segmentação do mercado turístico da Covilhã, segundo as despesas que os visitantes realizam no destino.
O estudo inclui uma reflexão teórica acerca do papel das despesas dos visitantes para estimular o desenvolvimento económico local e a apresentação dos resultados da aplicação de um inquérito a 205 visitantes do concelho da Covilhã, em Março e Maio de 2008.
Foram identificados, através de uma análise de *clusters* hierárquica, quatro segmentos de visitantes da Covilhã distintos em termos de importância para o desenvolvimento económico do destino: "baixo", "médio", "alto" e "muito alto". Estes segmentos apresentam diferenças estatisticamente significativas em termos de importância económica para o destino, perfil socio-demográfico e comportamento de viagem.
O presente estudo pretende contribuir para a obtenção de informação que permita elaborar estratégias que integrem uma visão de maximização dos benefícios que o turismo poderá proporcionar para o desenvolvimento económico local.

Palavras-chave: Despesas turísticas, segmentação, desenvolvimento económico local, concelho da Covilhã.
Área temática: Economia Urbana, Regional e Local.

Abstract

This study aims to segment the tourist market of a Portuguese destination – the municipality of Covilhã - based on the level of visitors' expenditures during the visit. In order to accomplish this objective a theoretical reflection about the relevance of tourist expenditures in the local economic development was undertaken, followed by a visitor survey, during the months of March and May of 2008, yielding a total of 205 visitors interviewed.
Through a hieratical cluster analysis, four distinct segments emerged in terms of expenditures - "light", "medium", "heavy" and "very heavy". The segments identified show significant statistical differences in respect to their economic relevance to the destination, socio-demographic profile and trip behaviour.
This study contributes to obtaining information in order to define more appropriate strategies to increase the influence of tourism in the local economic development.

Keywords: visitors' expenditure, segmentation, local economic development, municipality of Covilhã.
Thematic area: Urban, Regional and Local Economics.

1 - INTRODUÇÃO

O Turismo é uma das actividades que invoca elevado potencial para estimular o desenvolvimento das economias locais, através das componentes da Procura Turística (consumo turístico individual, consumo turístico colectivo e formação bruta de capital fixo turístico). Acréscimos nestas componentes irão estimular a Produção local, o Emprego, o VAB, o

Rendimento das Famílias, as Receitas Públicas e, também, a diversificação da estrutura produtiva local.

A revisão de literatura permite concluir a existência de uma grande quantidade de publicações sobre a temática dos impactes económicos do turismo e do desenvolvimento económico. No entanto, verifica-se uma lacuna em termos de investigação sobre a avaliação e planeamento de estratégias que permitam maximizar a relevância económica do turismo para os destinos turísticos locais. Essa lacuna torna-se ainda mais visível no que concerne ao reconhecimento do papel que os modelos de segmentação dos visitantes, segundo as despesas que realizam nos destinos, podem assumir como *inputs* para aumentar a contribuição do Turismo para o desenvolvimento económico dos destinos.

As despesas dos visitantes são a principal força económica que impulsiona os impactes do turismo na economia dos destinos, ao representar um fluxo de moeda adicional para essas economias (Fredman, 2008). Assim, assume-se que, quanto maiores forem as despesas dos visitantes, maiores serão os impactes económicos positivos directos para o destino.

Segmentar o mercado de acordo com os seus consumos e caracterizar esses segmentos em termos de perfil socio-demográfico, comportamento de viagem e avaliação do destino é um passo fundamental, que permitirá definir estratégias de desenvolvimento que estimulem o contributo do turismo para o desenvolvimento económico local.

Com a realização deste trabalho pretende-se segmentar o mercado turístico da Covilhã, de acordo com as despesas que os visitantes realizam neste concelho e, desta forma, contribuir para a definição de estratégias de *marketing* turístico orientadas de forma a contribuírem positivamente para o desenvolvimento económico do destino. A prossecução deste objectivo pressupõe contribuir para a fundamentação da importância de estratégias de segmentação do mercado turístico na definição de propostas de desenvolvimento turístico que, por sua vez, contribuam para maximizar os efeitos económicos positivos que o turismo provoca numa determinada economia local. A importância deste tipo de investigação torna-se ainda mais relevante ao considerar o caso dos concelhos do Interior de Portugal, onde se insere o concelho da Covilhã, por se tratarem de concelhos economicamente menos desenvolvidos.

Para concretizar os objectivos descritos anteriormente, apresenta-se uma reflexão teórica acerca do Turismo como factor impulsionador de desenvolvimento económico local, através do estudo dos impactes económicos do turismo para os destinos. Essa reflexão teórica explorará, ainda, o conceito de segmentação, bem como as bases de segmentação mais utilizadas na literatura para o mercado turístico, para se passar a fundamentar a utilidade da adopção de uma estratégia de *marketing* que passe por segmentar o mercado turístico com base nas despesas feitas pelos visitantes no destino.

Posteriormente, apresentam-se os resultados obtidos no estudo de segmentação do mercado turístico da Covilhã. Inicialmente descreve-se a metodologia escolhida para a aplicação do modelo de segmentação ao concelho da Covilhã, detalhando a metodologia de recolha e de análise estatística dos dados. De seguida, identificam-se os segmentos de visitantes da Covilhã de acordo com as despesas realizadas durante a visita a este concelho e apresentam-se as diferenças existentes no que respeita às características desses segmentos, em termos de perfil socio-demográfico e económico, de comportamento em viagem e da avaliação do destino. Por fim, serão discutidas as implicações destes resultados para aumentar o contributo do Turismo no desenvolvimento económico da Covilhã.

2 - O TURISMO COMO FACTOR IMPULSIONADOR DE DESENVOLVIMENTO ECONÓMICO DOS DESTINOS.

Apesar da importância e dinâmica económicas que o turismo alcançou, ainda não existem certezas acerca do seu real papel no processo de desenvolvimento económico.

As teorias de crescimento e desenvolvimento económicos endógenos consideram a diversificação e diferenciação como motores destes processos. Geralmente, a actividade turística, além de resultar do aproveitamento dos recursos dos destinos, aumenta a diversificação e a diferenciação dos produtos locais (através do aumento da variedade e qualidade dos produtos

disponíveis). Desta forma, a actividade turística pode desempenhar um papel fundamental na diversificação e diferenciação da economia local. Exemplo disso é o facto de muitas cidades, outrora cidades industriais, terem encontrado no turismo a solução para se regenerarem e modernizarem, depois do declínio da actividade industrial.

Observa-se, através da literatura, que as principais razões para que o turismo possa influenciar positivamente o desenvolvimento económico dos locais são: valorização económica de bens livres; geração de rendimento; criação de infra-estruturas colectivas; aproveitamento de equipamentos não utilizados; modernização da estrutura económica local; aumento da produção local; dinamização das actividades económicas tradicionais; e aumento do emprego (Cabugueira, 2005; Sharpley e Telfer, 2002; Kadt, 1979).

No que diz respeito à valorização económica de bens livres, a actividade turística, na sua maioria, desenvolve-se devido à presença no local de uma atracção que representa um recurso natural, patrimonial ou cultural. Esse recurso endógeno passará a ser produtivo para a economia local (Cabugueira, 2005). Por outro lado, o facto de ser uma actividade que pode gerar ganhos a partir de "infra-estruturas" naturais, já existentes no local (aproveitamento de recursos endógenos), faz com que os custos de investimento para o desenvolvimento do turismo possam ser, comparativamente, mais baixos (Sharpley e Telfer, 2002).

Quando um visitante se desloca para um destino, gera uma procura adicional nesse destino, transferindo para aí parte do seu rendimento, sob a forma de despesas turísticas. Essas despesas irão ter um efeito "em cascata" na economia do destino (efeitos directos, indirectos e induzidos), devido à interligação entre os diversos sectores económicos locais que o turismo proporciona (Cabugueira, 2005). Deste aspecto resulta a geração de rendimento para a economia local em consequência do seu desenvolvimento turístico.

Relativamente à criação de infra-estruturas colectivas, o desenvolvimento do turismo depende muito da existência de alguns equipamentos sociais e infra-estruturas colectivas (que servem visitantes e população local), como infra-estruturas que satisfaçam necessidades básicas (saneamento, electricidade, entre outros) ou infra-estruturas de transportes e comunicações. Essas infra-estruturas serão construídas para possibilitar o desenvolvimento turístico, mas irão contribuir para melhorar as condições de vida da população local (Kadt, 1979).

O turismo pode, também, aproveitar instalações obsoletas ou abandonadas nos destinos (como fábricas desmobilizadas ou aldeias desertificadas), dando-lhes nova função, e funcionar como justificação para a preservação e aproveitamento do património local. Este aproveitamento de equipamentos não utilizados é um factor de regeneração dos locais que desenvolvem o sector turístico como sector estratégico para o seu desenvolvimento económico.

A modernização da estrutura económica local tem a ver com o facto do desenvolvimento turístico de um local aumentar a dimensão do mercado e alterar as exigências desse mercado. Estas modificações fazem com que a economia seja dinamizada, podendo mesmo aumentar a modernização de toda a estrutura económica local.

O turismo cria a necessidade de aumentar a produção local, tal como defende Kadt (1979), através do aumento da procura de produtos alimentares, produtos típicos variados e outros produtos consumidos pelos visitantes. Constitui exemplo dessa realidade o facto de, na maioria dos destinos de pequena dimensão, a procura turística se interessar pela arte e pelas actividades artesanais locais. Este interesse não só permitirá um aumento dos benefícios que os artistas e artesãos locais obtêm do turismo, como fará com que estas actividades tradicionais se revitalizem e modernizem.

O aumento da procura que o turismo gera, não origina apenas efeitos sobre a produção local, mas também traz a possibilidade da economia local aumentar o seu nível de emprego. Isto é, o turismo trará benefícios para sectores económicos locais (por exemplo, para os artesãos e artistas locais) mas também para os residentes que poderão vir a beneficiar do aumento de emprego neste sector (Kadt, 1979).

Estes benefícios que o turismo pode desencadear nas economias resultam dos impactes económicos positivos (directos, indirectos e induzidos) do turismo.

As despesas dos visitantes assumem papel de relevo no estudo dos impactes positivos do turismo para os locais de destino. Estas despesas permitem identificar as actividades onde são efectuados mais gastos e os impactes económicos que são gerados e que integram alterações no nível de emprego, rendimento, produção local, receitas públicas e na estrutura produtiva local.

Os gastos dos visitantes, ao representarem despesas extra introduzidas na economia local, geram um primeiro *round* de efeitos para as actividades económicas que fornecem, directamente, bens e serviços aos visitantes, como hotéis e restantes actividades características do turismo (Sharpley e Telfer, 2002). Estes benefícios económicos directos das despesas dos visitantes, bem como a sua contabilização dependem dessas mesmas despesas (Eusébio, 2006). Parte deste valor sairá, nesse momento, da economia local, de forma a cobrir as importações de consumos intermédios efectuados pelas actividades que fornecem os visitantes. Outra parte das despesas iniciais servirá para essas mesmas actividades pagarem as compras de *inputs* intermédios feitas a outros sectores da economia local e para pagarem dos *inputs* primários, dando origem a impactes positivos indirectos. Por fim, quando os acréscimos verificados no rendimento das famílias, em consequência dos impactes directos e indirectos do turismo, entram novamente na economia através do consumo ou investimento ocorrerão benefícios categorizados como benefícios induzidos.

Considerando o mecanismo dos efeitos económicos positivos referidos, facilmente se deduz que quanto maior forem as importações necessárias para satisfazer os sucessivos aumentos de procura na economia local, maiores serão as fugas de rendimento e menores serão os efeitos indirectos e induzidos positivos para o local. Por outro lado, quanto mais interligação e dependência existir entre os sectores económicos locais, maiores serão os efeitos indirectos e induzidos positivos para o desenvolvimento económico local.

No entanto, os impactes totais do turismo para os destinos são de difícil contabilização, já que as despesas dos visitantes vão despoletar uma série de efeitos em cadeia que são interdependentes (efeitos directos, indirectos e induzidos). Para tentar avaliar a intensidade destes efeitos utiliza-se o conceito de multiplicador turístico. O multiplicador turístico para um destino traduz o efeito global em termos de rendimento, de emprego, de *output* ou de outra variável económica, causado por um estímulo directo (as despesas dos visitantes) na economia do destino. A aplicação prática dos multiplicadores turísticos permite, então, estimar o efeito directo, indirecto e induzido duma variável económica (por exemplo, o rendimento, emprego ou *output*), como consequência de uma alteração na procura turística final.

Apesar da aceitação generalizada relativamente aos benefícios ultrapassarem em larga escala os custos económicos do turismo, considera-se que quantificar esses custos (desde as variações no nível geral de preços da economia local, aos custos de oportunidade e à forte dependência do turismo) é um aspecto que deve merecer importância e estudo para um planeamento e gestão conscientes e adequados dos destinos.

Desta forma, se o local tiver feito um planeamento adequado do desenvolvimento do turismo, este efeito de criação de procura do turismo estará previsto e diminuirá a necessidade de recorrer a importações para suprir as necessidades da procura adicional gerada, ou seja, a economia local irá aumentar a sua produção, a capacidade de criar emprego e sentir necessidade de se modernizar quer a nível tecnológico, quer ao nível da qualidade e variedade de produtos que oferece.

Ao afirmar que o turismo pode ser um elemento decisivo no processo de desenvolvimento económico local, não se pretende olhar esta actividade como a panaceia que resolverá todos os problemas das comunidades, nem que as economias se devem concentrar apenas neste sector. Diversificação económica deve ser prioridade para o processo de desenvolvimento económico do destino. O desenvolvimento turístico deve ser planeado de forma a contribuir para essa diversificação dentro da economia local e nunca para a substituir (Inskeep, 1991). Em virtude da importância que as repercussões e efeitos "arrastadores" dos potenciais benefícios e custos do turismo assumem para o local, é essencial uma avaliação integral desses efeitos e a definição de planos de desenvolvimento do turismo a nível local.

3 - SEGMENTAÇÃO DA PROCURA TURÍSTICA COM BASE NAS DESPESAS REALIZADAS PELOS VISITANTES NO DESTINO.

Os destinos turísticos são, antes de mais, territórios. São espaços com características e dimensões variadas, cujo objectivo é o de aproveitar essas características como atractivos para determinados visitantes, estimulando, desta forma, o desenvolvimento turístico e económico do local. Hoje em dia, os locais têm de ser vendidos como um produto composto e complexo, necessitando, para tal, de conhecer o comportamento dos consumidores que pretende alcançar (Plog, 1974), identificando grupos homogéneos de visitantes.

Cooper *et al* (2005) atribuem a especificidade do *marketing* turístico ao facto de este processo de *marketing* ter de considerar que o produto que pretende desenvolver é um território com características únicas, composto por inúmeros serviços e que os consumidores que pretende atingir, para lá de serem envolvidos num processo de decisão específico e complexo relacionado com percepções, risco e níveis de qualidade elevados, também irão ter comportamentos, enquanto consumidores, diferentes daqueles que é habitual adoptarem face ao consumo de outros produtos. Daqui decorre a importância de planear o desenvolvimento dos locais enquanto destinos turísticos. Uma fase crucial para o sucesso do planeamento de destinos turísticos é saber onde o destino quer chegar. Esta fase envolve conhecer o conceito de segmentação de mercado.

A segmentação de mercado pode ser entendida como a agregação de consumidores individuais em grupos ou a desagregação do mercado global em partes mais pequenas (McKercher *et al*, 2002). Isto é, segmentar o mercado consiste em, dentro de um mercado caracterizado pela heterogeneidade, identificar um determinado número de mercados de menor dimensão, que registam um certo grau de homogeneidade interior. Esta homogeneidade verifica-se relativamente a alguns critérios (bases de segmentação) que podem ir desde as características demográficas, geográficas, até a características psicológicas e de comportamento dos consumidores.

Kotler *et al* (2006) dividem o processo de segmentação em duas etapas. A primeira etapa é a de identificar as bases para agrupar os potenciais visitantes (bases de segmentação) para, na segunda etapa, se descreverem os perfis dos subgrupos (segmentos) encontrados. Este processo permitirá não só medir a atractividade de cada segmento e, a partir daí, seleccionar os segmentos-alvo para o destino, como também permitirá desenvolver o posicionamento do destino e escolher um programa de *marketing* específico para esses alvos.

A utilidade e importância da segmentação do mercado têm sido cada vez mais reconhecidas na literatura sobre turismo (por exemplo, Bryant e Morrison, 1980; Mudambi e Baum, 1997; Plog, 1974; Andereck e Caldwell, 1994; Mok e Iverson, 2000; Koc e Altinay, 2007).

Koc e Altinay (2007) realçam esta importância, afirmando que desenvolver e manter vantagens competitivas nos, cada vez mais, competitivos mercados turísticos, depende fortemente de compreender os consumidores em termos de quem consome, o que se consome, quando, onde, porquê e como se consome. Partindo deste conhecimento dos consumidores ou potenciais consumidores dos produtos turísticos, podem ser desenvolvidas estratégias de *marketing* mais apropriadas para cada destino.

Para lá da competitividade existente no mercado turístico, a diminuição dos custos inerentes ao processo de *marketing* e a satisfação a procura turística de modo mais adequado aos seus gostos, desejos e necessidades, são vantagens importantes que os destinos poderão retirar de estratégias assentes na segmentação de mercado.

Assim, segmentar o mercado permite aos locais não só seleccionar os segmentos que entende serem mais vantajosos de atrair (segmentos-alvo), para alcançar os seus objectivos enquanto destino turístico e ganhar vantagens competitivas no mercado turístico, mas também conhecer melhor as necessidades, percepções e vontades relevantes dos seus visitantes (Dolnicar, 2002 e Laesser e Crouch, 2006).

Kotler *et al* (1993) resumem a importância que a segmentação assume para os destinos afirmando que os destinos turísticos necessitam de decidir não só quantos visitantes pretendem atrair, mas também que tipo de visitantes lhes é favorável atrair e reter.

Apostar num processo de segmentação do mercado turístico significa, então, aceitar e compreender as seguintes quatro assumpções: a) o mercado de um determinado produto turístico é constituído por diferentes segmentos, formados por indivíduos com necessidades e gostos diferentes relativamente ao produto; b) estes potenciais visitantes podem ser agrupados em segmentos de mercado cujos elementos têm características semelhantes e identificáveis; c) cada produto turístico será mais apelativo para alguns segmentos de mercado do que para outros; d) os destinos e as organizações podem tornar a sua estratégia de *marketing* mais efectiva se desenvolverem produtos específicos para determinados segmentos de mercado (Mill e Morrison, 1992).

Para definir estratégias que permitam aumentar o impacte económico do turismo para os destinos turísticos, utilizando para o efeito modelos de segmentação do mercado turístico, é necessário definir qual ou quais os elementos que permitirão distinguir os segmentos ou grupos de visitantes, ou seja, há que definir as bases de segmentação.

Dibb *et al* (2001) define bases ou variáveis de segmentação como sendo as dimensões ou características individuais ou de grupo que irão ser utilizadas para dividir o mercado global em segmentos. Existe um conjunto diversificado de variáveis que poderá ser utilizado como base de segmentação do mercado turístico. Porém, Dibb *et al* (2001:210) e Kotler *et al* (2006) alertam para o facto de raramente existir uma única ou melhor maneira de segmentar o mercado, devendo ser os técnicos de *marketing* a escolher a variável de segmentação a utilizar, em função dos objectivos pretendidos.

Moutinho (2000) refere que as variáveis de segmentação mais utilizadas em turismo são as características geográficas, demográficas, geo-demográficas, psicográficas, taxa de utilização dos recursos do destino e sensibilidade ao preço. Eusébio *et al* (2008) concluem, de forma mais geral, que as características demográficas dos visitantes têm sido a base de segmentação mais utilizada, pela facilidade de quantificação e pelas necessidades e desejos dos consumidores estarem fortemente correlacionadas com as características demográficas, tal como afirma Kotler *et al* (1999). Juaneda e Sastre (1999) sistematizam as variáveis de segmentação do mercado turístico mais utilizadas nos seguintes grupos: i) socio-demográficas (género, idade, dimensão da família, região de origem dos visitantes ou nível social); ii) socio-económicas (rendimento ou profissão); iii) motivo principal da viagem; iv) comportamento em viagem (actividades praticadas ou despesas dos visitantes no destino); e v) psicográficas (estilo de vida, benefícios procurados ou actividades e interesses).

Eusébio *et al* (2008) concordam com as variáveis de segmentação identificadas e defendem uma segmentação do mercado turístico com base em variáveis que traduzam o comportamento de viagem do visitante.

Na Tabela 1 identificam-se as bases de segmentação utilizadas em 40 estudos que foram objecto de análise neste estudo.

Tabela 1: Bases de segmentação do mercado turístico utilizadas em alguns dos estudos publicados sobre segmentação do mercado turístico.

Categorias de segmentação	Variáveis utilizadas	Artigo (referência)
Percepções do destino	- Emoções associadas à visita ao destino	Bigné e Andreu (2004).
	- Satisfação	Tsiotsou e Vasioti (2006).
	- Expectativas acerca da qualidade dos hotéis	Juwaheer (2007).
Perfil	- Características psicológicas (auto-confiança, curiosidade, variedade de interesses e espírito aventureiro)	Plog (1974).
	- Género	Kim, Lehto e Morrison (2007).
	- Origem geográfica	Andereck e Caldwell (1994).
	- Origem geográfica	Juaneda e Sastre (1999).
	- Origem geográfica	Moscardo, Pearce e Morrison (2001).
	- Valores pessoais: tradicionalistas *vs* modernos	Thrane (1997).

Tabela 1 (cont.)

Categorias de segmentação	Variáveis utilizadas	Artigo (referência)
Comportamento de viagem	-Variação sazonal das despesas dos visitantes	Koc e Altinay (2007).
	- Fontes de informação utilizadas	Bieger e Laesser (2001).
	- Motivações	Andereck e Caldwell (1994).
	- Motivações	Andreu, Kozak, Avci e Cifter (2005).
	- Motivações	Beh e Bruyere (2007).
	- Motivações	Kau e Lim (2005).
	- Motivações	Kim e Agrusa (2008).
	- Actividades praticadas durante a visita	Mehmetoglu (2007).
	- Actividades praticadas durante a visita	Moscardo, Pearce e Morrison (2001).
	- Actividades praticadas durante a visita	McKercher *et al* (2002).
	- Actividades Praticadas durante a visita - Época da visita	Spencer e Holecek (2007).
	- Familiaridade com o destino (primeira visita ou visita repetida)	Litvin (2007).
	- Benefícios procurados	Sarigöllü e Huang (2005).
	- Benefícios procurados	Frochot (2005)
	- Benefícios procurados	Kastenholz, Davis e Paul (1999).
	- Benefícios procurados	Lee, Morrison e O'Leary (2006).
	- Benefícios procurados	Molera e Albaladejo (2007).
	- Comportamento face ao meio-ambiente	Weaver e Lawton (2002).
	- Comportamentos anteriores face ao meio-ambiente no destino	Dolnicar e Leisch (2008).
	- Critérios de selecção do artesanato e grau de envolvimento na compra	Hu e Yu (2007).
	- Despesas turísticas	Pizam e Reichel (1979)
	- Despesas turísticas	Spotts e Mahoney (1991)
	- Despesas turísticas	Mudambi e Baum (1997).
	- Despesas turísticas	Agarwal e Yochum (1999b)
	- Despesas turísticas	Goldsmith e Litvin (1999)
	- Despesas turísticas	Mok e Iverson (2000)
	- Despesas turísticas	Pérez e Sampol (2000)
	- Despesas turísticas	Díaz-Pérez *et al* (2005).
	- Despesas turísticas	Legohérel e Wong (2006)
	- Despesas turísticas	Wang, Y. *et al* (2006)
	- Despesas turísticas	Laesser e Crouch (2006)

Uma das possíveis variáveis de segmentação com base no comportamento dos visitantes durante a viagem é a despesa turística. Esta variável tem sido muito utilizada desde finais da década de 70 até à actualidade (Pizam e Reichel, 1979; Spotts e Mahoney, 1991; Agarwal e Yochum, 1999b; Mok e Iverson, 2000; Pérez e Sampol, 2000; Díaz-Pérez *et al*, 2005; Legohérel e Wong, 2006; Laesser e Crouch, 2006; Wang *et al*, 2006). Estes estudos demonstram a utilidade e validade desta variável enquanto base de segmentação do mercado turístico.

Pizam e Reichel (1979), Spotts e Mahoney (1991) e Mok e Iverson (2000) declaram que a segmentação segundo as despesas turísticas satisfaz todos os requisitos relativos às características dos segmentos de mercado estabelecidos por Kotler *et al* (2006).

Considerando a relação entre o turismo (e, mais especificamente, as despesas dos visitantes no destino) e o desenvolvimento económico local, este tipo de segmentação surge com destaque enquanto instrumento de elaboração de estratégias que potenciem o desenvolvimento económico dos locais, atraindo os visitantes que efectuam mais despesas nos destinos.

Compreender o padrão de despesas dos visitantes, durante a sua visita a um determinado destino, é um elemento estratégico no mercado competitivo actual. A importância estratégica desta compreensão decorre da relevância que as despesas dos visitantes no destino podem assumir no processo de desenvolvimento económico local. Em consequência da natureza abrangente das despesas feitas durante uma viagem, para segmentar o mercado turístico com base nesta variável, é crucial perceber o comportamento de consumo dos visitantes durante a sua viagem e os principais factores que influenciam esse mesmo comportamento.

A compreensão acerca de quem gasta e onde gasta, permitirá aos agentes dos destinos desenvolver planos de *marketing* adequados para captar os segmentos que efectuam mais despesas no destino, potenciando desta forma o desenvolvimento económico local (Kotler *et al*, 1993, Mok e Iverson, 2000, Díaz-Pérez *et al*, 2005 e Fredman, 2008).

Atendendo ao facto de os recursos financeiros serem recursos limitados, especialmente no que respeita a financiamentos públicos, estratégias de *marketing* efectivas e eficazes em termos de maximização do retorno e rendimento, ganham ainda maior relevância e aparecem como instrumentos chave para o sucesso económico dos destinos na actualidade. Uma das possíveis abordagens para aumentar a eficácia dessas estratégias de *marketing*, no que respeita à maximização do retorno financeiro, reside na prossecução de uma segmentação cujo objectivo seja o de concentrar esforços para atrair e fidelizar os segmentos que, comparativamente, geram maior retorno e rendimento para o destino (Laesser e Crouch, 2006). No entanto, Legohérel e Wong (2006) alertam para o facto de que, apesar deste objectivo, pouca atenção tem sido prestada pelos agentes da oferta turística à segmentação dos visitantes baseada nas despesas que efectuam, a fim de maximizar o retorno dos seus esforços de planeamento e de *marketing*.

A segmentação com base nas despesas turísticas deve ser resultado de um estudo bem definido através de uma análise cuidadosa e detalhada do comportamento turístico, em termos de quantidade de produtos consumida e de níveis de despesa turística realizada (Legohérel e Wong, 2006). Depois, segundo os referidos autores, caberá aos agentes da oferta turística atrair os visitantes desejados e/ou convencer os visitantes a efectuarem mais despesas no destino. Deste ponto de vista, o nível de despesa será mais significativo que a quantidade de produtos consumida.

Concluída a segmentação, será necessário avaliar o potencial de cada segmento de mercado, para que o destino possa escolher atrair e reter aqueles segmentos que evidenciem maior potencial. Kotler *et al* (2006) afirmam que o potencial do segmento de mercado é um dos critérios para avaliar a atractividade de um segmento. Legohérel e Wong (2006) complementam esta afirmação, dizendo que este potencial/atractividade pode ser medido através da despesa total efectuada durante a viagem no destino ou da despesa diária por visitante no destino.

Legohérel e Wong (2006) acrescentam que foi demonstrado, através de estudos realizados anteriormente, que a despesa total muitas vezes se apresenta menos pertinente do que a despesa diária para identificar o segmento de visitantes mais atractivo para o destino turístico. Essa pertinência vai depender dos objectivos da estratégia de segmentação do mercado turístico, isto é, dependerá dos destinos quererem ter visitantes que efectuam mais despesas, independentemente da duração da sua estadia, ou se preferem atrair visitantes que efectuem mais despesas em períodos de estadia curtos. Os mesmos autores referem que atrair visitantes que efectuem mais despesas em períodos de estadia curtos será favorável para os destinos, visto que possibilita a atracção de um maior número de visitantes.

Conhecer as características dos visitantes que realizam mais despesas no destino permitirá a todos os agentes do destino turístico, delinear estratégias para atrair os segmentos de visitantes que se mostram mais atractivos, tendo em conta os objectivos que se pretendem alcançar em termos de desenvolvimento local. Como anteriormente foi referido, as despesas dos visitantes vão funcionar como dinamizadores da economia local e, assim, construir uma estratégia de *marketing* que atraia os segmentos de visitantes que gastam mais no destino, irá potenciar essa dinâmica económica e possibilitar um maior desenvolvimento económico local (Legohérel e Wong, 2006).

Deste modo, uma segmentação de mercado assente nas despesas dos visitantes, num contexto de maximização do retorno de estratégias de segmentação, apresenta-se como um instrumento de planeamento e de gestão imperativo para o desenvolvimento económico local (Laesser e Crouch, 2006).

4 - IDENTIFICAÇÃO DOS SEGMENTOS DE MAIOR VALOR ECONÓMICO PARA A COVILHÃ

4.1 – Metodologia

4.1.1 – Recolha de dados

Após a análise de todas as fontes de informação disponíveis que fornecem informação estatística sobre a procura turística em Portugal, verificou-se que não existia informação que permitisse caracterizar o mercado turístico da Covilhã e identificar os segmentos de maior valor económico para este destino. A ausência desta informação implicou o recurso a métodos directos de recolha de dados, tendo-se optado por utilizar o método de inquérito por questionário para proceder a essa recolha.

Na impossibilidade de inquirir todos os visitantes do Concelho da Covilhã, que no ano de 2008 se descolaram a este destino turístico, recorreu-se a uma amostragem por *clusters* multi-etapas. Após uma consulta das possíveis fontes que poderiam fornecer informação sobre a dimensão da população em estudo, observou-se que não existe informação disponível que permitisse a sua quantificação. Perante esta realidade optou-se por utilizar o número de visitantes que contactaram os postos de turismo da Covilhã, para se obter alguma informação e a sua distribuição temporal. Com base nestes dados dividiu-se o ano em análise (2008) em duas épocas – época alta e época baixa – e, dentro de cada uma destas épocas, definiu-se, aleatoriamente, um mês durante o qual se aplicariam os questionários, dentro destes meses foram definidos aleatoriamente os dias em que iriam ser administrados os questionários. Numa segunda etapa, identificaram-se os pontos de inquirição a utilizar tendo-os categorizado em três grupos: meios de alojamento, infra-estruturas de informação turística e atracções. Dentro de cada um dos grupos identificados foram seleccionados aleatoriamente os pontos de inquirição. Finalmente, num dia X e no ponto de inquirição Y eram inquiridos todos os visitantes.

O questionário foi aplicado durante 10 dias no mês de Março e 12 dias no mês de Maio de 2008, durante a semana e fim-de-semana, tendo-se inquirido 205 visitantes que responderam sobre o comportamento de viagem e padrão de despesas turísticas de 504 visitantes.

A construção do questionário foi elaborada com base nas recomendações da OMT (1995) sobre a recolha de informação estatística em Turismo e a quantificação das despesas turísticas. Utilizou-se, igualmente, como referência o questionário elaborado por Eusébio (2006). O questionário elaborado no âmbito deste estudo, integra questões relacionadas com o perfil socio-demográfico e económico do visitante; comportamento em viagem e avaliação do destino.

A aplicação do questionário aos visitantes da Covilhã foi realizada utilizando, essencialmente, uma abordagem pessoal directa. No entanto, foi também utilizada uma abordagem pessoal via intermediário e contou com a colaboração esporádica de duas pessoas qualificadas para a função, para além da constante presença da investigadora responsável pelo trabalho.

4.1.2 – Análise de dados

Numa fase posterior à aplicação dos inquéritos, procedeu-se à análise estatística dos dados recolhidos, a qual permitiu retirar as conclusões, discutir resultados e fazer uma reflexão crítica sobre todo o trabalho de investigação.

Recorreu-se ao programa informático SPSS (*Statistical Package for the Social Sciences*) – versão 14.0, para realizar essa mesma análise estatística.

Numa primeira fase, efectuou-se uma análise univariada para caracterizar a amostra dos visitantes e efectuar uma análise exploratória dos dados que permitisse compreender melhor a sua estrutura e quais as técnicas estatísticas mais adequadas ao seu tratamento.

Numa segunda fase, e depois de uma análise gráfica dos dados, realizou-se uma análise de *clusters* para identificar grupos homogéneos de visitantes, em termos de despesa diária *per capita*. Os *clusters* encontrados irão corresponder aos segmentos de visitantes a analisar.

Utilizou-se o método de *Ward*, um método hierárquico, para obter esses *clusters*. Este tipo de método identifica pares de visitantes semelhantes entre si de acordo com o atributo ou característica que se seleccionou. Assim, a utilização deste método permite obter grupos com dimensão semelhante (Arimond e Elfessi, 2001).

Por último, recorreu-se a testes do qui-quadrado (χ^2), testes ANOVA e ao teste não-paramétrico *Kruskal Wallis* para se verificar a existência de diferenças entre os segmentos encontrados, de acordo com o seu perfil socio-demográfico e económico, o seu comportamento em viagem e a avaliação que realizam do destino. O teste do qui-quadrado foi utilizado para verificar a existência de associação entre os segmentos identificados e características correspondentes a variáveis nominais. O teste da ANOVA (quando se verificavam os pressupostos de normalidade e homocedasticidade) e o teste não-paramétrico *Kruskal Wallis,* foram efectuados para verificar as diferenças que envolviam características correspondentes a variáveis quantitativas.

4.2 – Apresentação dos resultados

4.2.1 – Identificação dos segmentos com base nas despesas dos visitantes no destino.

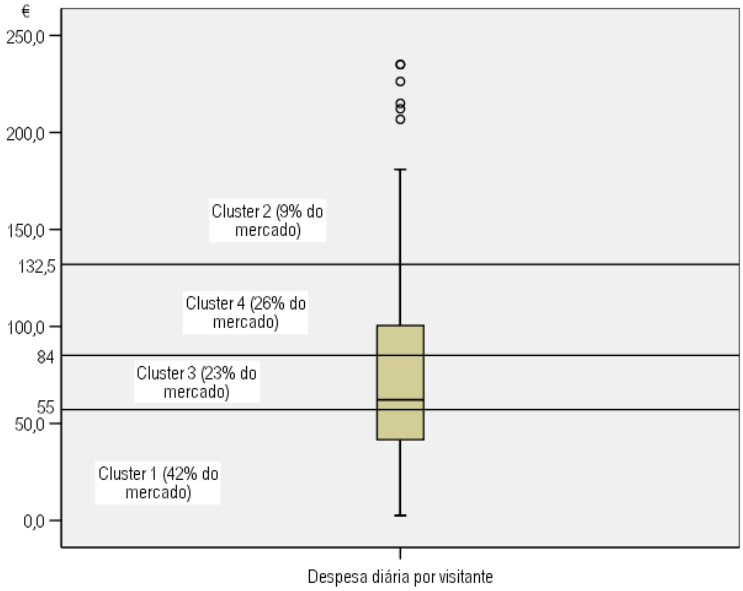
Como resultado da análise de *clusters* obteve-se uma combinação de quatro *cluster*s. Após a análise dos segmentos identificados, decidiu-se designar cada segmento de acordo com a sua importância económica para o destino. Assim, designou-se como segmento "baixo" o segmento que regista menor nível de despesa diária, segmento "médio" o segmento com nível moderado de despesa diária, segmento "alto" o segmento com níveis de despesa relativamente elevados e como segmento "muito alto" o segmento que regista níveis de despesa relativamente mais elevados (Tabela 2). Constatou-se, ainda, que o segmento com maior número de elementos é o segmento "baixo" (representa 42% dos visitantes), seguido dos segmentos "alto" e "médio" (com uma representatividade de 26% e 23%, respectivamente) e que o segmento com menos elementos é o "muito alto" (engloba apenas 9% dos visitantes) (Tabela 2 e Gráfico 1).

Tabela 2: Segmentos de visitantes, segundo as despesas diárias por pessoa que efectuam no destino

Nº do Segmento	Limites de despesa diária *pc*	Designação (nível de despesa *pc*)	Elementos		Despesa média *pc*	
			N	%	Média (€)	Desvio padrão
1	≤ 55€	"Baixo"	86	42	35,29	14,47
3	56€-84€	"Médio"	48	23	67,84	7,92
4	85€-132,5€	"Alto"	53	26	107,51	12,38
2	≥ 133€	"Muito alto"	18	9	175,47	35,92

Através da observação dos limites de cada segmento, verifica-se que os visitantes do segmento "muito alto" são visitantes que realizam, no mínimo, 133€ de despesas por dia, no destino. Deste modo, e considerando que o segmento "baixo" engloba visitantes cuja despesa diária por pessoa seja inferior ou igual a 55€, observa-se que, em termos de importância económica dos visitantes para o destino, o segmento muito alto será muito mais atractivo para o destino do que o segmento "baixo" (Gráfico 1).

Gráfico 1: *Boxplot* relativo à despesa diária por visitante, com as linhas limite dos *cluster*s/segmentos e a representatividade de cada um deles no mercado turístico da Covilhã



Desta distribuição dos visitantes pelos segmentos de despesa diária, conclui-se que a representatividade de cada um dos segmentos no mercado global do destino não é a desejada para maximizar o contributo do turismo para o desenvolvimento do destino. Para esse contributo ser maximizado deveriam estar a ser atraídos os segmentos "muito alto" e "alto".

Numa situação deste tipo, recorrer a um modelo de segmentação como o que se propõe nesta dissertação pode ser um instrumento-chave para permitir maximizar o papel do turismo no desenvolvimento económico do destino. Segmentar o mercado turístico do destino e delinear estratégias baseadas no conhecimento das características da procura turística e dos segmentos-alvo que essa segmentação permitiu, irá criar as condições ideais para atracção dos visitantes que realizam mais despesas diárias no destino. Estas despesas são o motor do desenvolvimento económico que o turismo pode trazer aos destinos.

4.2.2 – Caracterização dos segmentos identificados

Nesta secção apresentam-se os resultados obtidos no que respeita à caracterização dos segmentos de visitantes identificados para a Covilhã. Esta caracterização passa pela identificação de diferenças estatisticamente significativas entre os segmentos de visitantes encontrados em termos de:

a) perfil sócio-demográfico e económico;
b) comportamento de viagem;
c) avaliação e fidelização ao destino.

a) Perfil socio-demográfico e económico dos visitantes.

A análise das diferenças existentes em termos de perfil dos visitantes foi concretizada recorrendo às características demográficas, socio-culturais e económicas dos visitantes. Considerou-se essencial realizar esta análise, uma vez que este grupo de características influencia fortemente o nível de despesa que os visitantes realizam no destino (Pizam e Reichel, 1979; Spotts e Mahoney, 1991; e Wang *et al*, 2006).

O quadro que se segue, resume as diferenças em termos de perfil socio-demográfico e económico dos segmentos de visitantes.

Tabela 3: Diferenças entre os segmentos, em termos de características socio-demográficas dos visitantes que os integram

Variáveis	*Cluster* "Baixo" (41,95%)	*Cluster* "Médio" (23,41%)	*Cluster* "Alto" (25,85%)	*Cluster* "Muito alto" (8,78%)	**Teste χ^2** [Valor (Sig.)]
Estado Civil:					7,96 (0,047)
Casado	45,3%	26,6%	22,7%	5,5%	
Outro	36,4%	18,2%	31,2%	14,3%	
Idade:					25,29 (0,000)
[18-30[	42,37%	23,73%	20,34%	13,56%	
[30-50[	47,83%	29,35%	15,22%	7,61%	
+ de 50	31,48%	12,96%	50,00%	5,56%	
Nível de Habilitações:					19,09 (0,000)
Ensino Básico ou Secundário	56,3%	22,90%	16,7%	4,2%	
Ensino Superior	29,4%	23,9%	33,9%	12,8%	

Relativamente ao estado civil dos visitantes, conclui-se que os visitantes cujo estado civil é "casado", estão relativamente mais concentrados no segmento de despesas "baixo" e "médio", enquanto os visitantes com outro estado civil se concentram nos segmentos "alto" e "muito alto".

No que concerne à idade existem as seguintes diferenças entre os segmentos:

- Segmento com despesas diárias por pessoa "muito altas" concentra relativamente mais visitantes no escalão etário dos 18 aos 30 anos;
- Segmento de visitantes com despesas "altas" integra maioritariamente visitantes com idades superiores a 50 anos; e
- Segmentos "baixo" e "médio" são constituídos por mais visitantes com idades entre os 30 e 50 anos.

Os segmentos de despesas "altas" e "muito altas" são constituídos por relativamente mais visitantes com ensino superior, enquanto o segmento "baixo" integra mais visitantes com ensino básico ou ensino secundário.

b) Comportamento de viagem.

As características do comportamento de viagem dos visitantes que diferenciam os segmentos identificados categorizam-se em:

i) motivo principal da visita;
ii) duração da visita;
iii) características do grupo de viagem;
iv) época da estadia;
v) principal tipo de alojamento e meio de transporte utilizados;
vi) planeamento da viagem;
vii) tipo de actividades praticadas;
viii) familiaridade com o destino;
ix) organização da viagem;
x) padrão de despesa (distribuição percentual das despesas por tipo de produto adquirido).

A

Tabela Tabela 4 invoca as diferenças estatisticamente significativas no comportamento de viagem dos visitantes que fazem parte dos quatro *clusters* identificados. Não foi possível utilizar o teste da ANOVA uma vez que as variáveis quantitativas que foram testadas não cumpriam os pressupostos da normalidade e/ou da homocedasticidade. Assim, para essas variáveis e em alternativa à ANOVA (Pestana e Gageiro, 2005), utilizou-se o teste não-paramétrico de *Kruskal-Wallis* (KW), tal como foi referido no ponto anterior.

Tabela 4: Diferenças estatisticamente significativas entre os segmentos, de acordo com o comportamento de viagem

Variáveis	*Cluster* "Baixo" (41,95%)	*Cluster* "Médio" (23,41%)	*Cluster* "Alto" (25,85%)	*Cluster* "Muito alto" (8,78%)	Teste χ^2 [Valor do teste (Sig)]	Teste KW [Valor do teste (Sig)]
Motivo viagem:						
Lazer/Férias	38,24%	28,68%	24,26%	8,82%	27,53 (0,000)	-
VFA	73,68%	5,26%	15,79%	5,26%		
Outros	19,35%	22,58%	45,16%	12,90%		
Duração média da estada dos turistas:	3 noites	2 noites	2 noites	2 noites	-	10,96 (0,012)
Grupo de viagem:					-	12,93 (0,005)
Dimensão média do grupo	3	2	3	2		
Nº de crianças	1	0	0	0	-	6,73 (0,081*)
Principal meio de alojamento:					33,50 (0,000)	-
Emp.Tur./Aloj.Local	28,39%	28,39%	32,26%	10,97%		
Outro	85,71%	3,57%	7,14%	3,57%		
Principal meio de transporte:					9,26 (0,026)	-
Automóvel próprio	40,40%	27,15%	21,85%	10,60%		
Outro meio de transporte	42,00%	14,00%	40,00%	4,00%		
Planeamento:						
Fontes informação comerciais:					29,11 (0,000)	-
Sim	26,13%	28,83%	36,94%	8,11%		
Não	60,64%	17,02%	12,77%	9,57%		
Fontes informação não comerciais:					22,49 (0,000)	-
Sim	57,43%	18,81%	14,85%	8,91%		
Não	26,92%	27,88%	36,54%	8,65%		
Actividade praticada:						
Visitar infra-estruturas recreativas:					10,71 (0,013)	-
Sim	41,86%	27,91%	11,63%	18,60%		
Não	41,98%	22,22%	29,63%	6,17%		
Visitar áreas protegidas:						
Sim	39,08%	34,48%	22,99%	3,45%	13,58 (0,004)	-
Não	44,07%	15,25%	27,97%	12,71%		
Participar em congressos, seminários, conferências e cursos de formação:					20,24 (0,000)	-
Sim	9,09%	18,18%	63,64%	9,09%		
Não	45,90%	24,04%	21,31%	8,74%		
Familiaridade:						
1ª visita	27,54%	36,23%	30,43%	5,80%	14,35 (0,002)	-
Visita repetida	49,26%	16,91%	23,53%	10,29%		
Pacote de viagem:						
Sim	8,82%	23,53%	52,94%	14,71%	23,86 (0,000)	
Não	48,54%	23,39%	20,47%	7,60%		

**Estatisticamente significante a um nível de 10%.*

Os segmentos "alto" e "muito alto" são constituídos por relativamente mais visitantes cujas motivações se encontram na categoria "Outros" ("Outros" integra os motivos de "Negócios e Razões Profissionais" e "Outros Motivos"), enquanto o segmento "médio" integra, maioritariamente, visitantes que se deslocaram à Covilhã por motivos "Lazer/Férias" e os visitantes do segmento "baixo" têm motivações relacionadas com "VFA".

No que respeita à duração da estadia, dentro do grupo turistas, verificou-se que existem diferenças entre os segmentos de visitantes, observando-se uma relação inversa entre a duração média da estadia e as despesas médias diárias por visitante.

O segmento "baixo" integra visitantes inseridos em grupos de viagens com mais elementos do que os que pertencem ao segmento "muito alto" e é também o segmento "baixo" aquele que integra visitantes com um número médio de crianças superior no seu grupo de viagem (uma criança por grupo de viagem) comparativamente com os restantes segmentos cujo número médio de crianças no grupo de viagens é zero.

Os resultados demonstram que os visitantes que pernoitam principalmente em "Empreendimentos Turísticos" ou "Alojamento Local" integram os segmentos de visitantes "médios", "altos" ou "muito altos".

Os visitantes que se deslocam em automóvel próprio concentram-se nos segmentos "muito alto" e "médio", enquanto os que viajam em outro meio de transporte pertencem relativamente mais aos segmentos "baixo" e "alto". O facto do segmento "alto" ter esta característica pode dever-se ao facto deste segmento concentrar relativamente mais visitantes que utilizaram o automóvel alugado como meio de transporte.

No que respeita às fontes de informação[1] utilizadas, os visitantes do segmento "médio" e "alto" são os que recorrem relativamente mais a fontes de informação comerciais (dentro das quais a internet é a mais utilizada). Por outro lado, os segmentos extremos utilizam relativamente mais fontes de informação não comerciais (dentro das quais a experiência pessoal é a mais referida pelos visitantes). Estas conclusões podem estar relacionadas com os resultados verificados relativos às diferenças de familiaridade dos segmentos com o destino (Tabela 4), que indicam que os segmentos "baixo" e "alto" são os segmentos que integram relativamente mais visitantes em visita repetida, concentrando os outros dois segmentos mais visitantes em primeira visita.

As diferenças que se mostraram estatisticamente significantes entre segmentos no que respeita ao tipo de actividades praticadas no destino são ao nível da visita a infra-estruturas recreativas, visita a áreas protegidas e participar em congressos, seminários, conferências e em cursos de formação. O segmento "baixo" constitui-se maioritariamente por visitantes que não praticam as actividades mencionadas. O segmento "médio" engloba, maioritariamente, visitantes que visitam infra-estruturas recreativas e áreas protegidas. O segmento "alto" representa, maioritariamente, visitantes que participam em congressos, seminários e conferências no destino. Por fim, o segmento "muito alto" engloba relativamente mais visitantes que apenas se distinguem dos do segmento "alto" por, para além de participarem em congressos, seminários e conferências, também visitam infra-estruturas recreativas (Tabela 4). Estes resultados permitem concluir que os segmentos que efectuam maior volume de despesa no destino são segmentos que englobam visitantes com motivações profissionais ou de negócios mas que durante a sua visita também usufruem de infra-estruturas recreativas no destino (actividade de lazer).

O segmento "muito alto" concentra relativamente mais visitantes em visita repetida do que o segmento "baixo" (Tabela 4). Este facto pode prender-se com o facto de que este segmento integra visitantes cujo motivo principal da visita é negócios ou razões profissionais, para as quais se sabe que o factor familiaridade não assume o mesmo tipo de importância que para outros grupos de motivações.

Os segmentos com maiores despesas viajam relativamente mais com aquisição de pacote de viagem, ao contrário do segmento "baixo", que viaja mais de forma independente.

Realizou-se, ainda, uma análise da despesa total por visitante. Relativamente à despesa total pode verificar-se que seriam necessários mais de quatro visitantes do segmento "baixo" (cada visitante realiza cerca de 80,33€ de despesa total) ou cerca de três visitantes do segmento "médio" (cada visitante realiza em média, 136,05€ de despesa total) para ter o mesmo efeito, em termos de despesa no destino, que um visitante do segmento "muito alto" (cada visitante realiza, em média, 357,74€ de despesa total). Relativamente ao segmento "alto", seriam necessários cerca de 2 visitantes deste segmento (cada visitante realiza, em média, 184,17€) para o destino obter o mesmo efeito, em termos de despesa, que obtém com um visitante do segmento "muito alto" (Tabela 5).

[1] Dividiram-se as fontes de informação em fontes comerciais (agências de viagens, internet, jornais, revistas, TV, rádio e catálogos de empresas/instituições) e em fontes não comerciais (experiência pessoal, recomendações de familiares e amigos e outras fontes de informação).

Tabela 5: Diferenças entre os segmentos, em termos de despesa total por visitante

Tipo de despesa	*Cluster* "Baixo"	*Cluster* "Médio"	*Cluster* "Alto"	*Cluster* "Muito alto"	Teste KW [Valor do teste (Sig)]
Despesa média total por visitante (€)	80,33	136,05	184,17	357,74	72,656 (0,000)
Estrutura da despesa (%):					
Alojamento	29,23%	40,18%	37,94%	43,17%	68,067 (0,000)
Alimentação	40,21%	33,31%	32,26%	25,17%	39,668 (0,000)
Transporte	10,77%	6,32%	9,02%	13,35%	28,778 (0,000)
Actividades culturais, recreativas e desportivas	6,14%	5,91%	7,56%	11,35%	15,113 (0,002)
Compras	10,28%	13,77%	12,81%	6,19%	23,282 (0,000)
Outras	3,37%	0,51%	0,41%	0,78%	2,441 (0,486)

Relativamente à estrutura das despesas realizadas no destino, observa-se que a maioria da despesa total por visitante pertencente ao segmento "baixo" é realizada em alimentação, enquanto nos outros segmentos essa maioria é afecta à despesa em alojamento (Tabela 5). O segmento "médio" é o que afecta relativamente menos despesa total por pessoa à categoria de transportes e, a par do segmento "alto", é o que mais despesa total relativa efectua em compras. Por fim, o segmento "muito alto" é, de todos os quatro segmentos, o que afecta uma maior parcela das suas despesas a actividades culturais, recreativas e desportivas (Tabela 5).

Ainda dentro da estrutura das despesas realizadas no destino, testou-se também a existência de diferenças entre os segmentos de acordo com a sua estrutura de despesa em compras de produtos típicos (Tabela 6).

Os segmentos "médio" e "alto" são os segmentos para os quais o peso das despesas em produtos típicos na despesa total por visitante em compras é mais elevado. O segmento "alto" realiza a maioria das suas despesas totais em compras em produtos típicos de artesanato. Por fim, o segmento "muito alto" é o segmento cujas despesas nos produtos típicos têm menor importância na despesa total em compras realizadas no destino (Tabela 6).

Tabela 6: Diferenças entre os segmentos de acordo com o peso das despesas em produtos típicos nas despesas totais, por pessoa, realizadas em compras

Tipo de Produto	% na Despesa total, por visitante, em Compras				Teste KW [Valor do teste (Sig)]
	Cluster "Baixo"	*Cluster* "Médio"	*Cluster* "Alto"	*Cluster* "Muito alto"	
Total Produtos Típicos	84%	95%	91%	72%	18,715 (0,000)
Produtos Típicos – Alimentação	69%	63%	41%	52%	9,865 (0,020)
Produtos Típicos – Artesanato	19%	31%	49%	37%	13,173 (0,004)

A análise das diferenças entre segmentos em termos de estrutura das despesas efectuadas, conjugada com as outras características diferenciadoras dos segmentos (em especial as percepções dos visitantes acerca da *performance* do destino), é um elemento importante para delinear estratégias de captação e fidelização, uma vez que tem subjacente as escolhas dos visitantes em termos de produtos que preferem consumir nos destinos.

c) Avaliação e fidelização ao destino.

Pretendeu-se verificar a existência de diferenças estatisticamente significativas entre os segmentos de acordo com a avaliação que os visitantes fazem deste local enquanto destino turístico e a sua fidelização a este destino, analisando-se a relação existente entre os segmentos de visitantes e as variáveis que se seguem:

i) avaliação holística;
ii) fidelização ao destino.

Em relação à avaliação holística não se verificou a existência de diferenças entre os segmentos, sendo que todos os segmentos apresentam um elevado grau de satisfação com a visita à Covilhã (Tabela 7).

Também não foi encontrada nenhuma diferença estatisticamente significativa entre os segmentos, de acordo com a intenção de fidelização ao destino (probabilidade de voltar a visitar a Covilhã e de recomendar este destino ao seu grupo de familiares e amigos). Verificou-se, ainda, que a intenção de voltar a visitar a Covilhã e de recomendar este destino ao seu grupo de familiares e amigos é bastante elevada em todos os segmentos (Tabela 7).

Tabela 7: Diferenças entre os segmentos de acordo com a avaliação holística e fidelização ao destino.

Média atribuída a cada variável[2]	*Cluster* "Baixo" (41,95%)	*Cluster* "Médio" (23,41%)	*Cluster* "Alto" (25,85%)	*Cluster* "Muito alto" (8,78%)	**Teste KW** [Valor (Sig.)]
Nível global de satisfação com a visita	3,99	3,98	4,13	3,72	4,755 (0,191)
Probabilidade de voltar a visitar	4,44	4,25	4,32	4,50	4,494 (0,213)
Probabilidade de recomendar visita ao local a familiares e amigos	4,44	4,44	4,53	4,50	0,752 (0,861)

O facto de não se terem encontrado diferenças estatisticamente significativas entre os segmentos, pode indicar que a Covilhã está a conseguir fidelizar todos os segmentos de visitantes, uma vez que a probabilidade de voltar e de recomendar este destino ao seu grupo de familiares e amigos é elevada em todos os segmentos.

No quadro seguinte sistematizam-se as principais características que diferenciam os segmentos de visitantes identificados.

[2] Os visitantes classificaram o seu nível global de satisfação com a visita numa escala de 1 (muito mau) a 5 (muito bom) e a probabilidade de voltarem a visitar e de recomendarem a amigos e familiares, numa escala de 1 (muito improvável) a 5 (muito provável).

Tabela 7: Caracterização dos segmentos de visitantes, em termos de despesa diária por pessoa

Segmento de visitantes com nível de despesa "baixo":	Segmento de visitantes com nível de despesa "médio":
• idade entre os 18 e os 50 anos • ensino básico ou secundário concluído • estado civil "casado" • recurso a fontes de informação não comerciais • aquisição de pacotes de viagem • deslocação através de "outros meios de transporte" que não o automóvel próprio • preferência por outros meios de alojamento que não empreendimentos turísticos ou alojamento local para passarem a noite • motivação da visita é VFA • elevado grau de familiaridade com o destino • duração média da estadia é 3 noites • viagem em grupos de 3 pessoas, em média, incluindo 1 criança com menos de 15 anos • não participa em actividades no destino, como visita a áreas protegidas e infra-estruturas recreativas ou participação em congressos, seminários • maior parte da despesa diária no destino é realizada em produtos de alimentação	• idade entre os 18 e os 50 anos • ensino superior completo • estado civil "casado" • recurso a fontes de informação comerciais • aquisição de pacotes de viagem • deslocação através de automóvel próprio • preferência por empreendimentos turísticos ou alojamento local para passarem a noite • motivação da visita é Lazer/Férias • reduzido grau de familiaridade com o destino; • duração média da estadia é 2 noites • viagem em grupos de 2 pessoas, em média, sem crianças com menos de 15 anos • participa em actividades no destino como visita a áreas protegidas e infra-estruturas recreativas no destino • maior parte da despesa diária no destino é realizada em serviços de alojamento e produtos de alimentação
Segmento de visitantes com nível de despesa "alto":	Segmento de visitantes com nível de despesa "muito alto":
• idade superior a 50 anos • ensino superior completo • outro estado civil que não seja "casado" • recurso a fontes de informação comerciais • aquisição de pacotes de viagem • deslocação através de "outros meios de transporte" que não o automóvel próprio • preferência por empreendimentos turísticos ou alojamento local para passarem a noite • motivação da visita é outra razão que não Lazer/Férias ou VFA • reduzido grau de familiaridade com o destino • duração média da estadia está entre 1 e 2 noites • viagem em grupos de 2 a 3 pessoas, em média, sem crianças com menos de 15 anos • participação em congressos, seminários e conferências • maior parte da despesa é realizada em serviços de alojamento	• idade entre os 18 e 30 anos • ensino superior completo • outro estado civil que não seja "casado" • recurso a fontes de informação não comerciais • aquisição de pacotes de viagem • deslocação através de automóvel próprio • preferência por empreendimentos turísticos ou alojamento local para passarem a noite • as motivações da visita são, maioritariamente, razões profissionais • elevado grau de familiaridade com o destino • duração média da estadia está entre 1 e 2 noites • viagem em grupos de 1 a 2 pessoas, em média, sem crianças com menos de 15 anos • participação em congressos, seminários e conferências e visitam infra-estruturas recreativas no destino • maior parte da despesa diária é realizada em serviços de alojamento

5 – DISCUSSÃO DOS RESULTADOS E IMPLICAÇÕES PARA A DEFINIÇÃO ESTRATÉGIAS DE DESENVOLVIMENTO TURÍSTICO

No que respeita às estratégias de *marketing* turístico futuras para a Covilhã que integrem uma visão de maximização dos benefícios económicos para o destino, os resultados apresentados permitem conhecer os segmentos de visitantes que podem potenciar esses benefícios para o desenvolvimento local e distinguir esses segmentos do restante mercado turístico.

Uma estratégia de *marketing* que tenha como base a segmentação de mercado de acordo com as despesas dos visitantes será uma estratégia adequada para potenciar o desenvolvimento económico local, pois permitirá, acima de tudo, o uso mais eficiente dos recursos cada vez mais escassos. Este uso eficiente dos recursos, partirá da caracterização dos segmentos de maior valor económico para a Covilhã que, por sua vez, permitirá a este destino perceber em que direcção deve desenvolver o turismo, maximizando o retorno financeiro da estratégia de *marketing* formulada e dos recursos nela aplicados.

A caracterização dos segmentos economicamente mais atractivos para o destino permite retirar linhas orientadoras para as estratégias de *marketing* turístico supramencionadas, adaptando os produtos turísticos e escolhendo os melhores meios de comunicação a utilizar com o mercado-alvo.

O destino, numa situação como a Covilhã, que está a conseguir fidelizar todos os segmentos de visitantes, e visando atingir o objectivo de maximizar o contributo económico do turismo, deverá orientar as suas estratégias de *marketing* para a finalidade de conseguir aumentar as despesas que os segmentos "baixo" e "médio" realizam no destino ou direccionar essas estratégias para a concentração de esforços para atracção de um maior número de visitantes dos segmentos que trarão maior valor económico para o destino ("alto" e "muito alto"). Independentemente de uma ou outra das linhas de orientação referidas, o destino, para conseguir delinear estratégias coerentes e eficazes, deverá utilizar a informação relativa às características dos segmentos-alvo para perceber as preferências dos seus visitantes-alvo e, assim, conseguir melhorar ou criar os aspectos que influenciam as despesas desses segmentos.

Adicionalmente, as conclusões retiradas deste trabalho podem mostrar aos planeadores locais que estratégias de promoção dos produtos típicos ou de aproveitamento de produtos típicos ainda não explorados pode ser um instrumento para aumentar as despesas que os visitantes com "alto" nível de despesa no destino realizam durante a sua estadia. Deixam também transparecer que a aposta em produtos típicos de artesanato poderá levar ao aumento das despesas dos segmentos "médio", "alto" e "muito alto", já que representam uma parcela nunca inferior a 30% das despesas realizadas em compras. Será ainda de salientar que o facto de o segmento "muito alto" ser aquele em que os produtos típicos representam uma menor parcela das despesas realizadas em compras, permite considerar como possibilidade de estratégia do destino a promoção destes produtos direccionada para as características específicas deste segmento, de forma a que esses produtos se tornem mais atractivos para estes visitantes.

Deste modo, a diversificação da oferta turística local (por exemplo, o desenvolvimento de produtos gastronómicos locais, uma maior valorização dos produtos alimentares e de artesanato típicos, a organização de visitas aos diversos pontos de atracção do concelho ou a melhoria da animação nocturna), o aumento da informação turística ou o aumento da qualidade dos serviços turísticos prestados, são algumas das propostas que podem ser retiradas deste estudo e adoptadas por estratégias de *marketing* para este destino. Estas propostas baseiam-se nas características diferenciadoras dos segmentos e que, se aplicadas, permitirão uma diversificação dos produtos onde os visitantes (quer os visitantes dos segmentos economicamente mais atractivos quer os visitantes dos restantes segmentos) podem efectuar despesa, aumentando assim a sua despesa que efectuam no destino.

Tanto as estratégias referidas como, por exemplo, a aposta nos produtos turísticos relacionados com o turismo de negócios podem, ainda, conduzir a uma maior captação de visitantes que se insiram nos segmentos de despesas "altas" e "muito altas".

6 – REFERÊNCIAS BIBLIOGRÁFICAS

AGARWAL, V. B. E YOCHUM, G. R. (1999b). "Tourist spending and race of visitors". *Journal of Travel Research*, 38(4), pp.173-176.

ANDERECK, K. L. E CALDWELL, L. L. (1994). "Variable selection in tourism market segmentation models". *Journal of Travel Research*, 33(2), pp.40-46.

ANDREU, L., KOZAK, M., AVCI, N. E CIFTER, N. (2005). "Market Segmentation by Motivations to Travel: British Tourists Visiting Turkey". *Journal of Travel & Tourism Marketing,* 19(1), pp.1-14.

ARIMOND, G. E ELFESSI, A. (2001). "A clustering method for categorical data in tourism market segmentation research". *Journal of Travel Research*, 39 (4), pp.391-397.

BEH, A. E BRUYERE, B.L. (2007). "Segmentation by visitor motivation in three Kenyan national reserves". *Tourism Management*, 28, pp.1464–1471.

BIEGER, T. E LAESSER, C. (2002). "Market Segmentation by Motivation: The Case of Switzerland". *Journal of Travel Research*, 41, pp.68-76.

BIGNÉ, J.E. E ANDREU, L. (2004). *Emotions in segmentation – An empirical study.*Annals of Tourism Research, 31(3), pp.682-696.

BRYANT, B.E. E MORRISON, A.J. (1980). "Travel Market Segmentation and the Implementation of Market Strategies". *Journal of Travel Research*, 18(3), pp.2-8.

CABUGUEIRA, A. (2005). "A importância económica do turismo". *Revista de Turismo e Desenvolvimento*, 2 (2), pp.97-104.

COOPER, C., FLETCHER, J., GILBERT, D. E WANHILL, S. (2005). "Tourism: Principles and Practice", 3ª ed., Prentice Hall.

DÍAZ-PÉREZ, F.M., BETHENCOURT-CEJAS, M. E ÁLVAREZ-GONZÁLEZ, J.A. (2005*).* "The segmentation of canary island tourism markets by expenditure: implications for tourism policy". *Tourism Management*, 26 (6), pp. 961-964.

DIBB, S. SIMKIN, L., PRIDE, W. M. E FERRELL, O.C. (2001). "Marketing –Concepts and Strategies". Houghton Mifflin Company, 4th European Edition.

DOLNICAR, S. (2002). "A Review of Data-Driven Market Segmentation in Tourism". *Journal of Travel & Tourism Marketing*, 12 (1), pp.1-22.

DOLNICAR, S. E LEISCH, F. (2008). "Selective marketing for environmentally sustainable tourism". *Tourism Management*, 29(4), pp. 672-680.

EUSÉBIO, C.A., CASTRO, E. A. E COSTA, C. (2008). "Diversidade no Mercado Turístico da Região Centro de Portugal: Identificação dos segmentos de maior valor económico em termos de Actividades Turísticas praticadas". *Revista de Turismo e Desenvolvimento*, 10, pp. 09-24.

EUSÉBIO, M. C. (2006). "Avaliação do impacte económico do turismo a nível regional – O caso da Região Centro de Portugal". Tese de Doutoramento, Universidade de Aveiro, 2006.

FREDMAN, P. (2008). "Determinants of visitor expenditure in mountain tourism". *Tourism Economics,* 14(2), pp.297-311.

FROCHOT, I. (2005). "A benefit segmentation of tourists in rural areas: a Scottish perspective". *Tourism Management*, 26, pp.335–346.

HU, B. E YU, H. (2007). "Segmentation by craft selection criteria and shopping involvement". *Tourism Management*, 28, pp.1079–1092.

INSKEEP, E. (1991). "Tourism Planning. An integrated and sustainable development approach". New York, Van Nostrand Reinhold.

JUANEDA, C. E SASTRE, F. (1999). "Balearic islands tourism: a case study in demographic segmentation". *Tourism Management*, 20, pp.549-552.

JUWAHEER, T.D. (2007). "Using Service Quality Expectations as a Criterion to Segment International Tourists in the Hospitality Industry: An Outlook of Hotels of Mauritius". *Journal of Travel & Tourism Marketing*, 21(2/3), pp.1-18.

KADT, E. DE (1979). "Tourism: Passport to Development?*"*. Oxford, Oxford University Press.

KASTENHOLZ, E., DAVIS, D., E PAUL, G. (1999). "Segmenting tourists in rural areas: The case of North and Central Portugal". *Journal of Travel Research*, 37, pp. 353–363.

KIM, D. LEHTO, X.Y. E MORRISON, A.M. (2007). "Gender differences in online travel information search: Implications for marketing communications on the internet". *Tourism Management*, 28, pp.423–433.

KIM, S.S. E AGRUSA, J. (2008). "*Segmenting* Japanese Tourists to Hawaii According to Tour Purpose". *Journal of Travel & Tourism Marketing*, 24(1), pp.63-80.

KOC, E. E ALTINAY, G. (2007). "An analysis of seasonality in monthly per person tourist spending in Turkish inbound tourism from a market segmentation perspective". *Tourism Management*, 28, pp.227-237.

KOTLER, P., HAIDER, D. E RAIN, I. (1993). "Marketing Places". The Free Press Maxwell Maxmillan International.

KOTLER, P. E ARMSTRONG, G. (2002). "Principles of Marketing". Prentice Hall, 3ªedição. Madrid.

KOTLER,P., BOWEN, J. T. E MAKENS, J. (2006), *Marketing for Hospitality and Tourism*, 4ª ed., Prentice Hall, New Jersey.

LAESSER, C. E CROUCH, G.I. (2006). "Segmenting markets by travel expenditure patterns: the case of international visitors to Australia". *Journal of Travel Research*, 44(4), pp.397-406.

LEE, G., MORRISON, A.M. E O'LEARY, J.T. (2006). "The economic value portfolio matrix: A target market selection tool for destination marketing organizations". *Tourism Management* 27, pp.576–588.

LEGOHÉREL, P. E WONG, K.K.F. (2006). "Market Segmentation in the Tourism Industry and Consumers' Spending: What About Direct Expenditures?". *Journal of Travel & Tourism Marketing*, 20(2), pp.15-30.
LITVIN, S.W. (2007). "Marketing Visitor Attractions: A Segmentation Study". *International Journal of Tourism Research*, 9, pp.9–19.
MCKERCHER, B., HO, P.S.Y., CROS, H. DU E SO-MING, B. C. (2002). "Activities-Based Segmentation of the Cultural Tourism Market". *Journal of Travel & Tourism Marketing*, 12(1), pp.23-46.
MEHMETOGLU, M. (2007). "Typologising nature-based tourists by activity—Theoretical and practical implications". *Tourism Management*, 28, pp.651–660.
MILL, R.C. E MORRISON, A.M. (1992). "The tourism system: an introductory text". Prentice Hall, 2ª ed., New York.
MOK, C. E IVERSON, T.J. (2000). "Expenditure-based segmentation: Taiwanese tourists to Guam". *Tourism Management*, 21, pp.299-305.
MOLERA, L. E ALBALADEJO, I.P. (2007). "Profiling segments of tourists in rural areas of South-Eastern Spain". *Tourism Management*, 28, pp.757–767.
MOSCARDO, G. PEARCE, P. E MORRISON, A. (2001). "Evaluating Different Bases for Market Segmentation: A Comparison of Geographic Origin versus Activity Participation for Generating Tourist Market Segments". *Journal of Travel & Tourism Marketing*, 10(1), pp.29-49.
MOUTINHO, L. (ed) (2000). "Strategic Management in Tourism". CABI Publishing, Glasgow.
MUDAMBI, R. E BAUM, T. (1997). "Strategic segmentation: an empirical analysis of tourist expenditures in Turkey". *Journal of Travel Research,* 36(1), pp.29-34.
OMT (1995). "Collection of tourism expenditures statistics". Organização Mundial de Turismo (OMT), Madrid.
PEREZ, E. A. E SAMPOL, C. J. (2000). "Tourist expenditures for mass tourism markets". *Annals of Tourism Research*, 27(3), pp.624-637.
PESTANA, M. H. E GAGEIRO, J. N. (2005). "Análise de dados para ciências sociais: a complementariedade do SPSS". Edições Sílabo. Lisboa:
PIZAM, A. E REICHEL, A. (1979). "Big Spenders and Little Spenders in U.S. Tourism". *Journal of Travel Research*, 18(1), pp.42-43.
PLOG, S. (1974). "Why destination areas rise and fall in popularity". *Cornell Hotel and Restaurant Quarterly*, 14(4), pp.55-58.
SARIGÖLLÜ, E. E HUANG, R. (2005). "Benefits Segmentation of Visitors to Latin America". *Journal of Travel Research*, 43, pp.277-293.
SHARPLEY, RICHARD E TELFER, DAVID J. (eds) (2002). "Tourism and Development: Concepts and Issues". Channel View Publications, UK.
SPENCER, D.M. E HOLECEK, D.F. (2007). "Basic characteristics of the fall tourism market". *Tourism Management*, 28, pp.491–504.
SPOTTS, D.M. E MAHONEY, E. (1991). "Segmenting visitors to a destination region based on the volume of their expenditures". *Journal of Travel Research*, 2, pp.24-31.
THRANE, C. (1997). "Values as segmentation criteria in tourism research: the Norwegian Monitor approach". *Tourism Management*, 18(2), pp.11-13.
TSIOTSOU, R. E VASIOTI, E. (2006). "Satisfaction: A Segmentation Criterion for "Short Term" Visitors of Mountainous Destinations". *Journal of Travel & Tourism Marketing*, 20(1), pp.61-73.
WANG, Y., ROMPF, P., SEVERT, D. E PEERAPATDIT, N. (2006). "Examining and Identifying the Determinants of Travel Expenditure Patterns". *International Journal of Tourism Research*, 8, pp.333–346.
WEAVER, D.B. E LAWTON, L.J. (2002). "Overnight Ecotourist Market Segmentation in the Gold Coast Hinterland of Australia". *Journal of Travel Research*, 40, pp.270-280.

ACERCA DOS IMPACTES DA UNIVERSIDADE DE ÉVORA NO SEU MEIO ENVOLVENTE

CONCEIÇÃO REGO
E-mail: mcpr@uevora.pt
ANTÓNIO CALEIRO
E-mail: caleiro@uevora.pt
Departamento de Economia & CEFAGE
Universidade de Évora

Resumo
As instituições de ensino superior, em geral, e as universidades, em particular, exercem sobre o meio envolvente inúmeros efeitos (multiplicadores), quer no lado da procura (por exemplo, por via dos rendimentos associado à sua actividade) quer no lado da oferta (por exemplo, ao nível do conhecimento associado aos seus graduados), os quais interagem entre si. Este conjunto de efeitos é de crucial importância para regiões económica (e demograficamente) deprimidas, como é o caso do Alentejo (Central), onde a Universidade de Évora se encontra localizada. Tendo em conta esta perspectiva, o presente trabalho pretende contribuir para a detecção da real importância da Universidade de Évora no meio envolvente ao mostrar como, em termos qualitativos, o mercado de trabalho dos indivíduos com formação superior, na região de Évora, se comporta de forma diversa do restante mercado de trabalho.

Palavras-chave: Análise de impactes regionais, Instituições de ensino superior, Universidades.
Área Temática: Economia Urbana, Regional e Local.

Abstract
Higher education institutions, in general, and universities, in particular, exert many regional (multiplier) effects, whether on the demand side (for instance, through the income related to their activity) or on the supply side (for instance, through the level of knowledge related to their graduates), which interact with each other. This set of effects is of crucial importance for economic (and demographically) depressed regions, such as the (central) Alentejo, where the University of Évora is located. Given this perspective, the paper contributes to the detection of the real importance of the University of Évora for the local region by showing how, in qualitative terms, the graduate labor market in the region of Évora behaves differently from the rest the labor market.

Key words: Higher education institutions, Regional impacts analysis, Universities.
Thematic Area: Urban, Regional and Local Economics.

1. INTRODUÇÃO

Nos últimos 30 anos a expansão do ensino superior foi um dos fenómenos sociais mais marcantes em Portugal. Esta expansão concretizou-se, fundamentalmente, por três formas: criação de novas universidades públicas, criação de ensino politécnico – público e privado – e desenvolvimento das instituições privadas. Esta evolução alterou de forma substancial o panorama do ensino superior em Portugal. Assim, este subsistema de ensino, inicialmente elitista, massificou-se; o número de alunos, docentes e estabelecimentos de ensino aumentou exponencialmente; os estabelecimentos que inicialmente apenas se localizavam nas grandes cidades do litoral (Lisboa, Coimbra e Porto) passaram a fazer parte do quotidiano das cidades do interior, alterando-o profundamente e permitindo o acesso a este grau de ensino por parte de numerosos estudantes que, de outra forma, não teriam condições económico-financeiras para o frequentar. Da instalação das instituições de ensino superior nas cidades do interior, em

particular, resultam efeitos decorrentes do exercício das funções tradicionais de ensino, investigação e extensão, bem como outros, muito diversos, ao nível económico, social, demográfico ou urbano. Em Évora, a marca da Universidade é indisfarçável, quer por via da presença de população estudantil quer através das estruturas que permitem o funcionamento desta instituição de ensino superior (IES), assistindo-se à consolidação do conceito de "cidade universitária" na medida em que as actividades da academia decorrem, fundamentalmente, em diversos espaços dentro do perímetro do centro histórico de Évora, fomentando a natural convivência entre os residentes e os estudantes.

As instituições de ensino superior, em geral, e as universidades, em particular, exercem sobre o meio envolvente inúmeros efeitos (multiplicadores), quer no lado da procura (por exemplo, por via dos rendimentos associado à sua actividade) quer no lado da oferta (por exemplo, ao nível do conhecimento associado aos seus graduados), os quais interagem entre si. Este conjunto de efeitos é de crucial importância para regiões económica (e demograficamente) deprimidas, como é o caso do Alentejo (Central), onde a Universidade de Évora (U.E.) se encontra localizada. Tendo em conta esta perspectiva, o presente trabalho pretende contribuir para a detecção da real importância da Universidade de Évora no meio envolvente ao mostrar como, em termos qualitativos, o mercado de trabalho dos indivíduos com formação superior, na região de Évora, se comporta de forma diversa do restante mercado de trabalho.

Este trabalho, além da introdução, contém uma secção 2 com uma breve revisão da bibliografia relativa à avaliação dos efeitos das IES no meio envolvente, incidindo exclusivamente na análise dos efeitos gerados pelo 'lado da oferta', uma secção 3 em que procederemos à apresentação das principais características da Universidade de Évora bem como da região envolvente e na secção 4 procedemos a uma avaliação dos efeitos da U.E. no mercado de trabalho regional dos quadros superiores em geral e, em particular, no mercado de trabalho de indivíduos com formação superior. Concluiremos com uma breve secção de conclusões.

2. REVISÃO DA BIBLIOGRAFIA

"Os recursos humanos serão sempre a parcela mais rica e promissora do potencial endógeno, tanto mais rica e tanto mais promissora quanto mais qualificada. Daí que a educação, a formação em geral assuma função de carácter estratégico, estruturante, no processo de desenvolvimento" (Simões Lopes, 1996: 4). Como bem sabemos, numerosos trabalhos têm permitido revelar a existência de uma ligação significativa e positiva entre educação e desenvolvimento económico: os países mais desenvolvidos são, de um modo geral, os que têm níveis de instrução mais elevado ou os que gastam relativamente mais com a educação e com I&D e, correlativamente, toda a insuficiência nestes domínios constitui um obstáculo ao desenvolvimento. A educação é um legado de uma geração para a seguinte, assentando aqui a sua principal função social (Thomas, 1995). Trata-se de um investimento a médio prazo, feito pela sociedade em geral e pelas famílias, na medida em que se abdica do contributo produtivo actual dos jovens na expectativa de um contributo reforçado no futuro (Lopes, 2001). O valor económico da educação reflecte-se de duas formas: por um lado, por via da rendibilidade privada da educação, sob a forma de salários mais elevados, reflectindo o respectivo nível de produtividade relativa e, por outro lado, através da existência de externalidades positivas. A educação aumenta a produtividade através do aumento de competências e dos conhecimentos colocados no mercado de trabalho. Este efeito torna-se relativamente mais fácil de quantificar se partirmos do pressuposto de que os empregadores vão pagar mais aos trabalhadores com maiores níveis de instrução apenas na medida em que estes são mais produtivos (Brown & Heaney, 1997; Turner, 1997).

A acumulação de capital, não só humano, como também físico e relacional, é um factor fundamental para a existência de crescimento económico sustentável, a longo prazo, e para a diminuição das divergências de rendimento entre países. A inter-acção entre a aquisição quotidiana de saberes e o ensino institucional é decisivo. A formação de competências decorre, por um lado, da formação institucional adquirida em diversos estabelecimentos de ensino

(escolas e universidades) e, por outro, da aprendizagem realizada no decurso da actividade profissional.

As IES, de um modo geral, e as universidades, em particular, são agentes promotores do desenvolvimento das regiões onde estão inseridas. As universidades são entidades polarizadoras do processo de desenvolvimento, capazes de gerar externalidades positivas fundamentais para a melhoria do capital humano através do exercício das funções de ensino, investigação e extensão. As universidades são pólos de desenvolvimento na medida em que através da sua actividade, da formação de recursos humanos, da difusão do conhecimento, da informação e da inovação, transmitem impulsos de crescimento para o tecido económico e social envolvente. Por via do desempenho da função ensino, estas instituições têm um papel fundamental na qualificação da população activa, que se reflecte na melhoria dos níveis de produtividade e de competitividade da economia bem como no aumento dos rendimentos do trabalho, designadamente nos salários, com a consequente melhoria do nível de vida.

No quadro da União Europeia salienta-se a importância crescente do papel das IES no domínio da Economia do Conhecimento. A criação da 'Europa do Conhecimento' foi definida como um dos principais objectivos a alcançar no âmbito da Cimeira de Lisboa (Março de 2000). Do ponto de vista da Agenda de Lisboa, a existência de crescimento económico está associada à produção de novos conhecimentos, da sua transmissão através do exercício das funções de ensino e formação bem como através da sua disseminação por via das tecnologias de informação e comunicação. Neste contexto, as universidades desempenham um papel crucial, particularmente nos domínios da promoção da investigação e da divulgação dos respectivos resultados, na cooperação com as empresas bem como no ensino e formação, particularmente a formação de investigadores, com vista à promoção do desenvolvimento local e regional. De acordo com a Comissão Europeia [COM (2003): 58][1], as universidades formam um número crescente de estudantes com níveis mais elevados de qualificações, os quais contribuem para o fortalecimento da competitividade da economia europeia. As universidades também promovem a prossecução dos objectivos da Estratégia de Lisboa, particularmente nos domínios do emprego e da coesão social bem como na melhoria do nível médio de educação na Europa: no conjunto da população com idades compreendidas entre os 25-64 anos, a taxa de emprego dos indivíduos com qualificações ao nível do ensino superior atingia, em 2001, os 84%. Por outro lado, a taxa de desemprego entre os diplomados permanecia em 3,9%, cerca de 1/3 do total dos indivíduos que estavam desempregados mas que possuíam um nível menor de habilitações académicas.

A análise dos efeitos locais e regionais das IES relaciona-se com o respectivo contexto territorial. Neste sentido, as IES constituem-se como instituições muito relevantes na promoção do desenvolvimento regional bem como no fortalecimento da coesão europeia. A promoção de centros tecnológicos e de centros de ciência, a proliferação de estruturas regionais de cooperação entre as empresas e as universidades, o alargamento das estratégias de desenvolvimento regional das IES bem como as redes regionais de universidades são algumas das dimensão da actividade das IES com vista à promoção do desenvolvimento territorial.

Os contributos das IES, para as economias regionais, têm vindo a assumir características diferenciadas. Drucker & Goldstein (2007:22) identificaram 8 funções diferentes, ou 'produtos', passíveis de serem associados às actuais universidades, os quais podem fomentar processos de desenvolvimento económico:

1. Criação de conhecimento,
2. Criação de capital humano,
3. Transferência de 'know-how',
4. Inovação tecnológica,
5. Investimento em capital,
6. Liderança regional,
7. Infra-estruturas de conhecimento,
8. Influência no meio envolvente.

[1] Comunicação da Comissão, *The role of the universities in the Europe knowledge*, Comissão Europeia, Bruxelas, 05.02.2003.

Em particular, no caso da difusão de novos conhecimento para o território envolvente, inovadores em termos de tecnologia ou de organização – um dos impactes gerados pelas instituições de ensino superior –, pode acontecer de formas variadas, entre as quais destacamos os estudantes, através da realização de estágios ou de colaboração com projectos de investigação mas também através da inserção de diplomados no mercado de trabalho local ou a partir dos investigadores, através da realização de projectos de investigação pura ou aplicada e da publicação dos respectivos resultados. Deste ponto de vista é inegável a preponderância do capital humano no processo de desenvolvimento económico local e regional.

3. ALGUMAS CARACTERÍSTICAS DO ALENTEJO E DA UNIVERSIDADE DE ÉVORA

O Alentejo é a região portuguesa menos densamente povoada (no final dos anos 90, Portugal apresentava uma densidade de 112 habitantes por km^2 enquanto no Alentejo existiam apenas 19,8 habitantes por km^2). A população, duplamente envelhecida, vive fundamentalmente em aglomerados urbanos de pequena e média dimensão, enquanto as áreas rurais se desertificam.

Depois do fim do predomínio da actividade agrícola, o sector terciário passou a ser o que ocupa mais activos, sendo baixo o seu nível médio de qualificações bem como a sua escolaridade. Claramente, este facto é fortemente condicionador de qualquer estratégia de desenvolvimento que se procure implementar na região o facto de cerca de 1/3 dos residentes apenas terem quatro anos de escolaridade (dados para 2001, obtidos através do Recenseamento Geral da População, INE). Aos baixos nível de qualificações escolares e profissionais estão associados, naturalmente, índices remuneratórios também menores que a média nacional. A debilidade da demografia regional repete-se na frágil dinâmica da actividade económica, onde, com excepções pontuais, as principais iniciativas de natureza empresarial são exógenas à região. Em termos de rendimentos, da região e dos seus residentes, os dados disponíveis para o PIB *per capita* revelam que o Alentejo é uma das 25 mais pobres regiões da Europa comunitária, sendo a que mais se afasta da média do país. No entanto, o Alentejo dispõe de um ambiente natural e cultural satisfatoriamente preservado e de uma dotação em infra-estruturas também razoável, atractivas em termos de fixação populacional e de actividades produtivas.

Évora, com os seus 56 mil habitantes, é a maior cidade do Alentejo. A sua actividade económica está profundamente ligada aos serviços, particularmente aos públicos. Évora posiciona-se como o grande pólo de terciário de todo o Alentejo Central, oferecendo um razoável conjunto de serviços de apoio à população, o qual tem vindo a evoluir no sentido do preenchimento de funções menos banais bem como na qualificação dos serviços já existentes, a que se alia também a localização na cidade das principais instituições da Região no apoio ao tecido empresarial e à população, entre as quais instituições públicas e associativas de que são exemplos a Universidade, associações sectoriais, instituições ligadas ao emprego e à formação, ou ao planeamento territorial.

A Universidade de Évora é a principal instituição de ensino superior existente no Alentejo e a única de cariz universitário público presente na região. Se a sua importância, no contexto do ensino superior público português é diminuta – os alunos desta instituição representam cerca de 4% do total de estudantes matriculados em Portugal neste grau de ensino –, já o mesmo não se pode dizer do seu peso nos cenários regional e concelhio: na U.E. estudam cerca de 50% dos alunos matriculados em todos os estabelecimentos de ensino superior público do Alentejo e estes estudantes representam cerca de 14% dos residentes no concelho de Évora.

Esta instituição tem ao seu serviço cerca de mil funcionários, predominando os docentes. A melhoria verificada, fundamentalmente ao longo dos anos 90, no nível de habilitações permite que actualmente a maioria dos docentes seja doutorada. Assim, a Universidade de Évora configura-se como uma das principais entidades empregadoras da cidade e a que apresenta níveis médios de qualificações mais elevados. Contudo, estima-se que o seu efeito na economia local e regional seja bem mais vasto que o resultante da criação de emprego directo: o conjunto da actividade económica da U.E. bem como dos seus estudantes deverá ter um contributo de

cerca de 1,5% no PIBpm do Alentejo, gerando um efeito multiplicador, quer no rendimento quer no produto do concelho de Évora que oscila entre 1,2 e 1,3 unidades monetárias, e deverá potenciar a existência de mais de dois mil empregos induzidos, concentrados maioritariamente no sector dos serviços (Rego, 2003).

Este estudo permitiu concluir ainda que, para além deste efeito, o principal contributo da U.E. para o desenvolvimento da região tem lugar através do desempenho da função ensino, ou seja, por via dos diplomados que integram o mercado de trabalho local, na medida em que estes contribuem para a melhoria da qualificação do emprego da cidade e da região (Rego, 2003). Neste domínio em concreto a principal forma de interacção entre a U.E. e o tecido socioeconómico regional decorre da inserção de diplomados em estágios ou nos quadros das instituições e empresas da cidade e da região, fundamentalmente nas áreas científicas das ciências económicas e empresariais, ciências exactas e ciências agrárias. As actividades relacionadas com o ensino e com os serviços são as que mais contribuem para a inserção profissional dos diplomados por esta instituição de ensino.

4. UM TIPO ADMISSÍVEL DE IMPACTES DA UNIVERSIDADE DE ÉVORA NO MERCADO DE TRABALHO LOCAL

A Universidade de Évora, enquanto uma das principais entidades geradoras do *número* de licenciados que se têm empregado na região do Alentejo, tem desempenhado, obviamente, um papel crucial no desenvolvimento do mercado de trabalho em geral, naquela região, e, em particular, no mercado de trabalho do seu principal *output*, ou seja dos quadros superiores em geral e, em particular, no mercado de trabalho de indivíduos com formação superior.

A nossa tese, aqui, resulta da vontade de ir para além da evidência *numérica*, atrás realçada. Na verdade, estudos recentes, têm demonstrado que as diversas características de funcionamento do mercado de trabalho como, por exemplo, a sua maior ou menor flexibilidade (em termos de legislação laboral ou em termos da mobilidade espacial da oferta) ou, ainda, o próprio comportamento dos sindicatos são decisivas para explicar a maior ou menor *persistência* no nível de desemprego desse mercado. Assim, tendo em conta que a região Alentejo, onde a Universidade de Évora se situa, se encontra deprimida, em termos económicos mas também demográficos, é de esperar que, pelas características particulares do mercado de trabalho de indivíduos com uma formação superior, o impacte seja também de natureza *qualitativa*.

Em termos *qualitativos*, julgamos ser tanto maior o impacte da Universidade de Évora, quanto mais distinta for a evolução do mercado de trabalho dos indivíduos com formação superior relativamente à registada no mercado de trabalho em geral, ao nível das principais zonas de influência da Universidade, ou seja o concelho de Évora.[2] Em trabalho anterior (Rego & Caleiro, 2003), tendo por base o período 1985-1999, mostrou-se que em termos do distrito de Évora, o nível de quadros superiores empregados evoluiu de forma assumidamente distinta entre 1992 e 1998. Em termos do concelho de Évora, a subida sustentada do número de quadros superiores empregados que se registou desde 1994, ainda que não tão pronunciada como aconteceu em termos distritais, não foi acompanhada pelo total do emprego.

De acordo com Rego & Caleiro (2003), a distinta evolução cronológica, quer em termos concelhios quer em termos distritais, do número de quadros superiores empregados face ao volume total de emprego justificou que se aceitasse ser a Universidade de Évora geradora de impactes, de *natureza quantitativa*, no mercado de trabalho mas também, pelas características particulares do mercado de trabalho dos quadros superiores, geradora de *eventuais* impactes de *natureza qualitativa*. A *eventualidade* destes impactes foi, então, de acordo com a metodologia

[2] Naturalmente, estamos com isto a admitir que aquele impacte será medido de forma *indirecta*, tal como a natureza agregada dos dados assim o exige, ao considerarmos que o efeito atribuível aos graduados por outras instituições (embora residam no concelho de Évora) se equipara ao efeito atribuível aos graduados pela Universidade de Évora (embora residam noutros concelhos), sendo certo que os estudos prévios sobre a questão (Rego & Caleiro, 2003; Rego, 2003; Rego & Caleiro, 2004; Rego & Caleiro, 2008) indiquem ser o concelho de Évora a zona de maiores impactes relativos da Universidade de Évora.

de modelos autoregressivos de primeira ordem, a qual, por ser de importância para o presente trabalho, se apresenta de seguida.

Alguns dos recentes desenvolvimentos teóricos sobre o mercado de trabalho, nomeadamente a chamada "hipótese da histerese", resultam considerar que a taxa ou o nível de desemprego do momento presente, u_t, é gerado por um processo auto-regressivo como, por exemplo:

$$u_t = \rho\, u_{t-1} + \varepsilon_t\,, \qquad (\mathbf{1})$$

onde ρ se assume como uma medida de *persistência* enquanto ε_t mede todos os factores explicativos do nível ou taxa de desemprego do momento presente para além do nível ou taxa de desemprego do momento anterior.

De forma a verificar a validade da hipótese considerada em Rego & Caleiro (2003) procedeu-se à estimação econométrica (em termos recursivos) de uma generalização do modelo em (**1**). Assim, foram estimados modelos do tipo:

$$L_t = \alpha + \beta\, L_{t-1} + \varepsilon_t\,, \qquad (\mathbf{2})$$

onde a estimativa, b, do parâmetro auto-regressivo, β, mede a persistência no volume de emprego, *L*.

Conforme se pôde verificar, o volume de emprego no mercado de quadros superiores exibia, quer em termos do distrito quer em termos do concelho de Évora, uma maior persistência do que a que caracterizava o emprego em termos totais. Mais importante que este facto é a constatação que o emprego de quadros superiores se caracterizava por um fenómeno de "sobre-histerese", quer em termos distritais quer concelhios, enquanto o total do emprego se caracterizava por um fenómeno de "sob-histerese". Significa isto que o mercado de quadros superiores no concelho e distrito de Évora, de cuja oferta a Universidade de Évora é uma das principais fontes, reagiria, de forma *qualitativamente* (e obviamente quantitativamente) diferente aos vários *choques* que afectarem o mercado de trabalho no seu todo.

Em suma, de acordo com Rego & Caleiro (2003), o mercado de quadros superiores funciona(ria), em termos do concelho e do distrito de Évora, de forma qualitativamente diferente do mercado de trabalho na sua generalidade. Para tal, admitia-se ser a Universidade de Évora a principal responsável, sendo aquele um dos admissíveis impactes da Universidade de Évora, cuja importância não devia passar despercebida.

Os modelos autoregressivos do tipo (**2**) têm, efectivamente, sido utilizados para medir a persistência numa série. A bondade dos seus resultados no entanto depende da natureza dos dados quanto à sua estacionaridade. Assim, é preferível a utilização de uma abordagem não paramétrica para medição da persistência (Marques, 2004; Dias & Marques, 2005). De acordo com esta abordagem, uma medida de persistência de uma série é dada por:

$$\gamma = 1 - \frac{n}{T}\,, \qquad (\mathbf{3})$$

onde *n* representa o número de vezes que a série cruza a sua media/tendência durante um intervalo com $T + 1$ observações.[3]

Assim, dado que uma série persistente é aquela que converge mais lentamente para o seu valor de equilíbrio (média ou tendência) depois de um 'choque', então quanto mais persistente será, menor será *n* em (**3**) e logo maior será γ.

De forma a complementar os resultados de Rego & Caleiro (2003) em termos do *nível de emprego*, assim como, de uma forma indirecta, verificar a sua robustez, consideramos os *níveis de desemprego*, em termos mensais, desde Janeiro de 2004 até Fevereiro de 2009, de acordo com os diversos níveis de escolaridade, para o concelho de Évora e região Alentejo.[4]

De acordo com a metodologia proposta, necessita-se de encontrar, para cada série, a sua tendência (enquanto valor médio ou de equilíbrio). Para o cálculo desta é habitual fazer-se uso do chamado filtro de Hodrick-Prescott. A utilização deste filtro permite também determinar a componente cíclica, a qual, para o cálculo do valor de *n* em (3), pode ser analisada

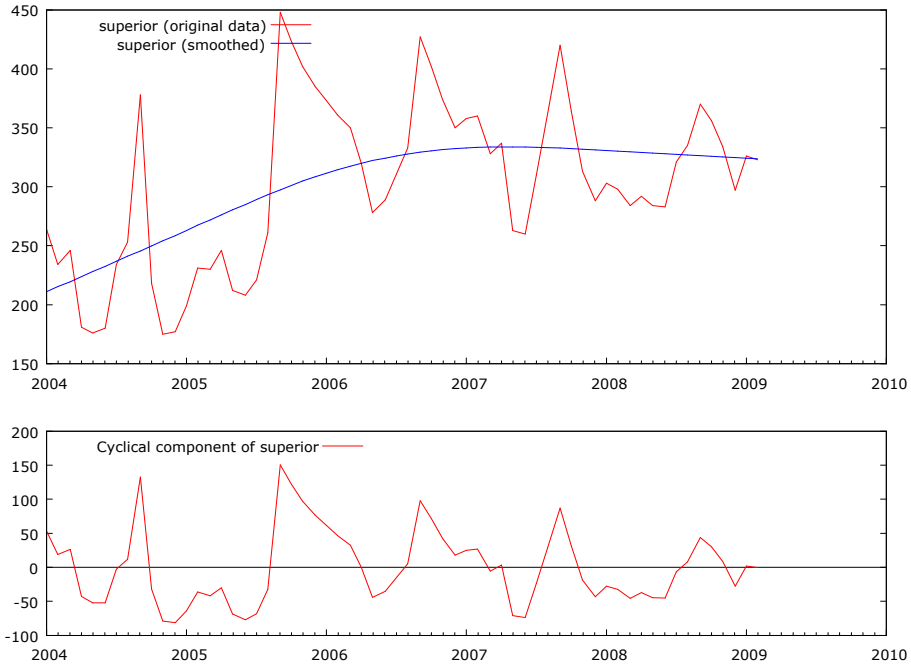
[3] Tal como se reconhece em Marques (2004) e em Dias & Marques (2005), γ mede a probabilidade incondicional de a série não cruzar a sua tendência/média no período.

[4] Os dados, cuja fonte é o Instituto do Emprego e Formação Profissional, encontram-se no Anexo a este trabalho.

determinando o número de vezes que passa de negativa (valores abaixo da tendência) a positiva (valores acima da tendência) e vice-versa. Assim, obtiveram-se os seguintes resultados:[5]

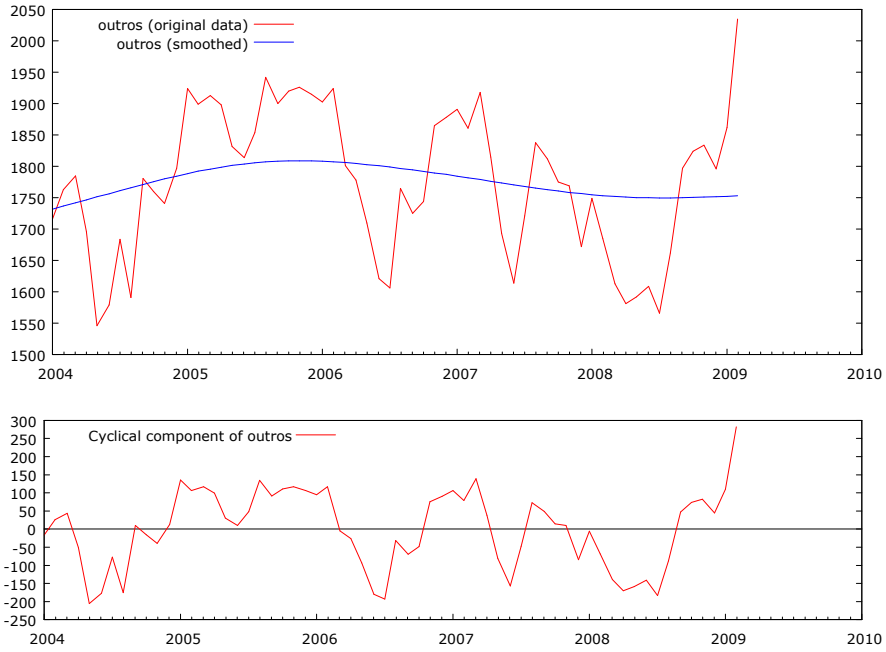
- Para o nível de habilitações correspondente ao ensino superior no concelho de Évora, $n = 15$, logo $\gamma = 0{,}7541$. A figura 1 mostra estes resultados.

Figura 1: Os resultados para o nível superior (concelho de Évora)



- Para o nível de habilitações correspondente aos outros níveis de ensino no concelho de Évora, $n = 11$, logo $\gamma = 0{,}8197$. A figura 2 mostra estes resultados.

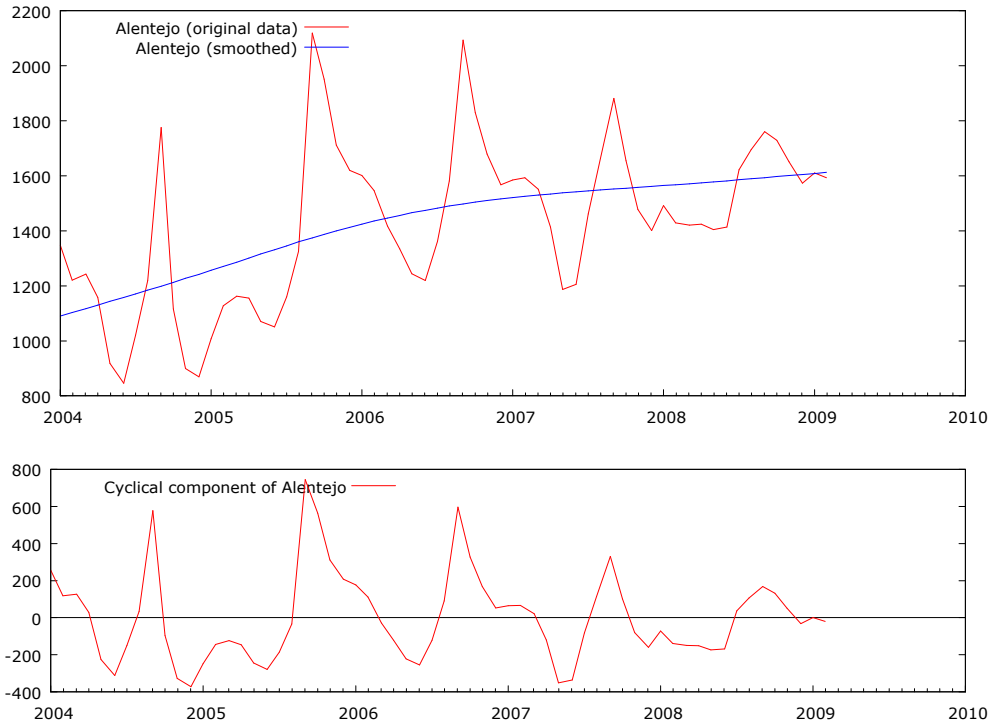
Figura 2: Os resultados para os outros níveis de escolaridade (concelho de Évora)



[5] Os resultados foram obtidos recorrendo ao programa livre *gretl*, o qual se pode encontrar em http://gretl.sourceforge.net.

- Para o nível de habilitações correspondente ao ensino superior na região Alentejo, n = 13, logo γ = 0,7869. A figura 3 mostra estes resultados.

Figura 3: Os resultados para o nível superior no Alentejo



Assim, podemos concluir que a série do desemprego para os indivíduos com nível de escolaridade correspondente ao ensino superior no concelho de Évora é a menos persistente, enquanto a série mais persistente é a correspondente aos indivíduos com outros níveis de escolaridade no concelho de Évora. Sendo certo que a Universidade de Évora não poderá ser considerada como a única responsável por esta diferença a favor do desemprego (por ser menos persistente) dos indivíduos com formação superior, é nosso parecer que certamente não pode ser excluída do conjunto de elementos com maior relevância na explicação do facto de o desemprego dos indivíduos com formação superior apresentar menor inércia que o correspondente aos outros níveis de ensino. A corroborar este nosso parecer apresenta-se o facto de, ao nível de toda a região Alentejo, onde a influência da Universidade de Évora ser, naturalmente, mais 'diluída' que no concelho de Évora, o nível de persistência no desemprego dos indivíduos com formação superior ser maior (mas menor que o nível de persistência do desemprego para os restantes níveis de escolaridade no concelho de Évora).

5.CONCLUSÃO

A bibliografia de referência no domínio da análise de impactes territoriais das instituições de ensino superior, particularmente os gerados 'pelo lado da oferta', mostra de forma inequívoca a importância destas instituições como entidades que potenciam o aumento dos níveis de desenvolvimento económico, quer no que respeita à criação de condições propiciadoras à melhoria do capital humano bem como através do fomento da criação e transferência de conhecimento inovador.

Um dos principais contributos da Universidade de Évora para o desenvolvimento da região faz-se através da função ensino, ou seja, por via dos diplomados que integram o mercado de trabalho local, na medida em que estes contribuem para a melhoria da qualificação do emprego da cidade e da região. Em Rego & Caleiro (2003) foram fornecidos claros indícios que, para além daquela melhoria, a Universidade de Évora influencia *qualitativamente* o funcionamento

do mercado de trabalho no concelho e distrito onde se implanta, ao tornar o nível de emprego de quadros superiores mais persistente (ou seja, no caso, mais robusto à subida) do que as restantes faixas do mercado de trabalho naquelas regiões. No presente trabalho, aqueles indícios confirmaram-se ainda mais ao concluir-se que a série do desemprego para os indivíduos com nível de escolaridade correspondente ao ensino superior no concelho de Évora é a menos persistente (de entre as séries relevantes consideradas), enquanto a série mais persistente é a correspondente aos indivíduos com outros níveis de escolaridade no concelho de Évora. Sendo certo que se trata de um método *indirecto* – como não pode deixar de ser – de 'medição' dos reais impactes da Universidade de Évora no seu meio envolvente, julga-se aceitável não se poder excluir do conjunto dos principais factores explicativos da realidade observada em Évora e no Alentejo aquela instituição de ensino superior.

Claramente, uma forma de contornar a necessidade de inferir, de forma indirecta, acerca da real importância da Universidade de Évora, consiste no recurso a dados de natureza micro. Esta, caso seja possível, será uma via (com evidentes potencialidades) a explorar em trabalhos futuros.

REFERÊNCIAS BIBLIOGRÁFICAS

BROWN, R.H., & M. T. HEANEY (1997), "A note on measuring the economic impact of institutions of higher education", *Papers in Regional Science*, Vol. 80, pp. 255-273

DIAS, DANIEL, & CARLOS ROBALO MARQUES (2005), "Using Mean Reversion as a Measure of Persistence", *Working Paper Series* N.º 450, European Central Bank, Março.

DRUCKER, J., & H. GODSTEIN, (2007), "Assessing the regional economic development impacts of universities: a review of current approaches", International Regional Science Review, 30, 1: 20-46 (January 2007).

MARQUES, CARLOS ROBALO (2004), "Inflation Persistence: Facts or Artefacts?", *Working Paper* 8, Banco de Portugal, Junho.

LOPES, RAUL (2001), *Competitividade, Inovação e Territórios*, Lisboa, Celta Editora

REGO, CONCEIÇÃO (2003), "Impactes da Universidade de Évora: Estudo de alguns efeitos no território envolvente", Tese de Doutoramento, Universidade de Évora.

REGO, CONCEIÇÃO, & ANTÓNIO CALEIRO (2003), "Impactes das Instituições de Ensino Superior no Território: Estudo do caso da Universidade de Évora", *Documento de Trabalho* n.º **2003/02**, Departamento de Economia, Universidade de Évora. (disponível em http://www.decon.uevora.pt/get_file.php3?id=429)

REGO, CONCEIÇÃO, & ANTÓNIO CALEIRO (2004), "A Atracção das Universidades em Regiões Economicamente Deprimidas: O caso da Universidade de Évora", *Revista Portuguesa de Estudos Regionais*, Vol. 7, 3.º Quadrimestre, pp. 19-40.

REGO, CONCEIÇÃO, & ANTÓNIO CALEIRO (2008), "On the spatial diffusion of knowledge by universities located in small and medium sized towns", Comunicação apresentada na conferência: Knowledge in Small and Medium Sized Towns: Towns as a place of knowledge and diffusion, Universidade do Algarve, Portugal, 04-05/12/2008.

SIMÕES LOPES, A. (1996), "Educação, formação e desenvolvimento regional", comunicação apresentada ao IV Encontro Nacional da APDR, Covilhã

THOMAS, D.J. (1995), "Education and the role of the university in economically developing regions", *Higher Education Policy*, Vol. 8, nº 2, pp. 51-62

TURNER, P. (1997), "The economic impact of a university on its environment", *European Ed.*, Vol. 29, n.º 3, pp. 88-95.

Anexo – Os dados

Tabela 1: Número de desempregados por níveis de escolaridade

	Concelho de Évora		Alentejo
	Superior	Outros níveis	Superior
Jan-04	264	1716	1347
Fev-04	234	1763	1221
Mar-04	246	1785	1243
Abr-04	181	1696	1158
Mai-04	176	1546	918
Jun-04	180	1579	846
Jul-04	234	1684	1021
Ago-04	253	1591	1219
Set-04	378	1781	1776
Out-04	218	1760	1116
Nov-04	175	1741	900
Dez-04	177	1797	869
Jan-05	199	1924	1008
Fev-05	231	1899	1127
Mar-05	230	1913	1163
Abr-05	246	1898	1155
Mai-05	212	1832	1071
Jun-05	208	1814	1051
Jul-05	221	1854	1160
Ago-05	261	1942	1326
Set-05	448	1900	2119
Out-05	423	1920	1951
Nov-05	402	1926	1712
Dez-05	385	1915	1620
Jan-06	373	1903	1601
Fev-06	361	1924	1546
Mar-06	350	1801	1417
Abr-06	320	1778	1335
Mai-06	278	1710	1244
Jun-06	289	1621	1220
Jul-06	311	1606	1361
Ago-06	333	1765	1581
Set-06	427	1725	2094
Out-06	402	1744	1831
Nov-06	374	1865	1679
Dez-06	350	1878	1568
Jan-07	358	1891	1585
Fev-07	360	1861	1593
Mar-07	328	1918	1551
Abr-07	337	1815	1414
Mai-07	263	1693	1187

Jun-07	260	1614	1206
Jul-07	310	1721	1462
Ago-07	362	1838	1663
Set-07	420	1812	1882
Out-07	364	1775	1655
Nov-07	313	1769	1478
Dez-07	288	1672	1401
Jan-08	303	1749	1492
Fev-08	298	1685	1429
Mar-08	284	1613	1421
Abr-08	292	1581	1424
Mai-08	284	1592	1405
Jun-08	283	1609	1414
Jul-08	321	1566	1622
Ago-08	335	1662	1696
Set-08	370	1797	1761
Out-08	356	1824	1729
Nov-08	334	1834	1651
Dez-08	297	1796	1573
Jan-09	326	1862	1610
Fev-09	323	2035	1593

Fonte: IEFP (Desemprego Registado por Concelhos -- Estatísticas Mensais)
(http://www.iefp.pt/estatisticas/MercadoEmprego/ConcelhosEstatisticasMensais/Paginas/Home.aspx)

VENTAJAS ECONÓMICAS EN LA DECLARACIÓN DE "MUNICIPIO TURÍSTICO" A DETERMINADAS ZONAS DE LA REGION DE MURCIA.

M. CARMEN PASTOR ÁLVAREZ
E-mail: carmen.pastor@upct.es
Departamento de Ciencias Jurídicas
Facultad de Ciencias de la Empresa
Universidad Politécnica de Cartagena

Resumen

La adquisición por parte de una determinada zona de la consideración de "Municipio Turístico" juega un papel muy relevante dentro de la ordenación territorial de los recursos, espacios y actividades turísticas, ya que tal denominación puede llevar aparejados unos efectos de vital importancia para esos municipios. Analizaremos dichos efectos, así como la legislación autonómica que justifica dicha denominación básicamente, sin olvidar el análisis de otras Comunidades Autónomas que también se hacen eco de esta figura.
Asimismo se hace necesario delimitar dicho concepto de otros términos que también tienen una amplia trascendencia en el sector turístico como son los de "Zonas de gran afluencia turística" y "Áreas turísticas saturadas". Dichos términos, que también examinaremos, responden a ideas diferentes, pero que igualmente desembocan en medidas, también diversas, que hayan de adoptarse en el sector turístico por parte de la administración competente.

Palavras-Chave: Municipio turístico / Región de Murcia / actividades turísticas.
Área temática: Economia Urbana, Regional e Local.

Abstract

For a certain area, being declared "Municipio Turístico" plays a very relevant role regarding the territorial planning of resources, spaces and tourist activities, since such a denomination may bring about some consequences of vital importance for those municipalities. We will analyse these effects, as well as the regional laws that justify this denomination, without neglecting the analysis of other Autonomous Communities that also cover this concept.
It is also becoming necessary to distinguish this concept from other terms of similar importance in the tourist sector, such as "Zonas de gran afluencia turística" and "Áreas turísticas saturadas". These terms -which will also be examined- correspond to different ideas, and, accordingly, they lead to different measures that must be adopted in the tourist sector by the authorised administration agency.

Thematic Area: Local, regional and urban economy.

1. LA DECLARACIÓN DE MUNICIPIO TURÍSTICO.

1.1. El concepto de "Municipio Turístico".

El análisis de las diferentes leyes autonómicas que recogen este término carecen de un concepto en sentido estricto de lo que se ha venido en llamar "Municipio Turístico", el cual viene dado con carácter descriptivo, es decir, en tanto reúna una serie de criterios, características o indicadores que permitan su catalogación como tal, además de destacarse cuál es la finalidad esencial que cada Comunidad Autónoma (que lo tenga regulado), persigue con su implantación y declaración como "Municipio Turístico" a zonas determinadas de su territorio.

Existen grandes diferencias en el tratamiento jurídico que las legislaciones de las CCAA otorgan al concepto de "Municipio Turístico"; unas con más detalle que otras, e incluso pasando por las que desconocen este término, como la Ley de Ordenación del Turismo de la Rioja de 2001. Sólo nueve CCAA han regulado de alguna manera esta declaración en sus leyes de Turismo: Murcia, Galicia, Navarra, Valencia, Cantabria, Andalucía, Cataluña, Aragón y Canarias.

Pero debemos comenzar por averiguar el concepto de "Municipio Turístico", y nada mejor que contemplarlo a la luz de las leyes autonómicas que lo enuncian:

- Ley 3/1998, de Ordenación del Turismo en Valencia, en su art. 5[1], que a nuestro modo de ver resulta de las más acertadas.

- Ley 5/1999, de Ordenación del Turismo de Cantabria, en su art. 39, aunque habla de "Comarcas y Municipios Turísticos"[2].

- Ley 12/1999, de Ordenación del Turismo de Andalucía, en sus arts. 6 (definición de Municipio Turístico y finalidad de su declaración) y 7 (Declaración)[3].

- Ley 13/2002, de Ordenación del Turismo de Cataluña, en su art. 18[4].

- Ley 6/2003, de Ordenación del Turismo de Aragón, en su art. 15[5].

---

[1] **Art. 25**: "*Podrán alcanzar la consideración de Municipio Turístico de la Comunidad Valenciana aquellos que puedan identificarse con alguno de los siguientes supuestos:*
Municipio de Destino Turístico: "Aquellos que a lo largo de todo el año mantienen una afluencia de visitantes, pernoctando en los mismos, superior al número de personas inscritas en su padrón municipal de residentes, suponiendo esta actividad la base de su economía y en los que la capacidad de sus alojamientos turísticos resulte superior a la de sus viviendas de primera residencia".
Municipio de Destino Vacacional: "Aquellos que en algún momento del año tienen una afluencia de visitantes, pernoctando en los mismos, superior al número de personas inscritas en su padrón municipal de residentes, pudiendo tener como complemento para su economía otras actividades y que la capacidad de sus alojamientos turísticos, añadida a las de vivienda de segunda residencia resulte superior a la de sus viviendas de primera residencia; en dicho cómputo, el número de plazas en alojamientos turísticos deberá representar al menos el 1% de su oferta".
Municipio de Destino de atracción turística: "Aquellos que por sus atractivos naturales, monumentales, socioculturales o por su relevancia en algún mercado turístico específico, reciben en un determinado momento del año un número significativo de visitantes en relación a su población de derecho, sin que los mismos pernocten necesariamente en ellos".

[2] **Art. 39**: "1. *La Consejería competente en materia de turismo establecerá Comarcas Turísticas, que agruparán territorios identificados por la unidad de su recurso turístico y que podrán llevar a cabo funciones de promoción, gestión y comercialización de sus productos turísticos.*

2. *Podrán conseguir la consideración de Municipios Turísticos de Cantabria aquellos que puedan acreditar: a) Parámetros con afluencia considerable de visitantes y pernoctaciones, la existencia de un número significativo de alojamientos turísticos de segunda residencia y en los que la actividad turística represente la base de su economía o, como mínimo, una parte importante de ésta. b) Parámetros cualitativos. El Municipio deberá acreditar un programa de promoción y fomento del turismo sostenible. La calidad y la innovación serán aspectos de la máxima importancia en la formulación del programa, detallando como mínimo las siguientes cuestiones: La necesidad de desarrollar un turismo que responda a las expectativas económicas y a las exigencias de protección del entorno./ Toda actuación turística ha de tener incidencia en la mejora de la calidad de vida de la población, contribuir a su enriquecimiento sociocultural y a la creación de un trabajo digno./ El mantenimiento de la cultura, historia y personalidad del municipio./ La adaptación del urbanismo y la edificación al paisaje local y la ordenación integral del patrimonio./ La existencia de reservas de agua estratégicas, así como su depuración y reutilización./ El nivel de generación de residuos estará ligado a la existencia en la zona de planes de recogida , reducción, reutilización y reciclaje./ Planes de formación y reciclaje profesional para los trabajadores y gestores del sector turístico./ Planes para la reconvención de los recursos potenciales en productos turísticos y su ordenación del territorio.*

[3] **Art. 6**: "1. *Se consideran Municipios Turísticos, y como tales podrán solicitar su declaración aquellos que cumplan los criterios que reglamentariamente se establezcan y entre los cuales deberán figurar la población turística asistida, el número de visitantes y la oferta turística*".

2. *Constituye la finalidad esencial para la declaración de Municipio Turístico el fomento de la calidad en la prestación de los servicios municipales al conjunto de la población turística asistida, mediante una acción concertada de fomento.*

3. *A los efectos de esta ley, se considera población turística asistida la constituida por quienes no ostenten la condición de vecinos del municipio pero tengan estancia temporal en el mismo por razones de visita turística, segunda residencia o alojamiento turístico. Su determinación se efectuará por los medios de prueba que reglamentariamente se establezcan*".
Art. 7: "1. *Para la declaración de Municipio Turístico se tendrá en cuanta, en especial, las actuaciones municipales en relación a: a) Los servicios mínimos que presta el Municipio respecto a los vecinos y la población turística asistida .b) Los servicios específicos en materia de salubridad pública e higiene en el medio urbano y natural y de protección civil y seguridad ciudadana y cuantos sean de especial relevancia turística*".

[4] **Art. 18**: "*Pueden tener la consideración de Municipio Turístico, a efectos de la presente Ley, los municipios que cumplan como mínimo una de las siguientes condiciones: a) Que la media anual ponderada de población turística sea superior al número de vecinos y el número de plazas de alojamiento turístico y de plazas de segunda residencia sea superior al de vecinos. b) Que el término municipal incluya un área territorial que tenga la calificación de recurso turístico esencial*".

[5] **Art. 15**: "*1 Podrán solicitar la declaración de municipios turísticos aquellos en los que concurran, al menos, dos de los siguientes requisitos: a) Que se trate de municipios cuya población de hecho al menos duplique la población de derecho en las temporadas turísticas. b) Que se trate de términos municipales en los que el censo de viviendas sea superior al doble de las viviendas habitadas*

- Ley foral 7/2003, de Ordenación del Turismo de Navarra, en su Disposición Adicional Única[6].

Del análisis conceptual de todas ellas podemos entresacar de forma principal que la declaración de "Municipio Turístico" se hace con carácter descriptivo, vinculándolo casi de forma unánime a tres factores o criterios:

a) La relación entre la población turística y el número de residentes y/o entre el número de alojamientos turísticos y el de viviendas de residencia primaria.
b) La existencia de un producto o servicio turístico que genera gran afluencia turística.
c) El cumplimiento de parámetros cualitativos: protección del entorno, programas de calidad en la oferta turística, políticas de promoción turística del municipio, etc.

1.2. El concepto de "Municipio Turístico en la CA de la Región de Murcia.

La Ley 11/1997, de 12 de Diciembre, de Turismo de la Región de Murcia regula el "Municipio Turístico" en el Tít. IV, titulado "De la acción pública en materia de turismo", en concreto en su art. 51, bajo la siguiente redacción:

1. En atención a las circunstancias señaladas en el apartado 3 de este artículo, los Municipios de la Región de Murcia podrán obtener la denominación de "Municipio Turístico", beneficiándose de modo preferencial de las acciones de fomento y promoción de la Administración Regional. Reglamentariamente podrán establecerse distintos tipos de municipios turísticos.

2. En la elaboración y modificación de los instrumentos de planteamiento urbanístico correspondientes a estos municipios, será oída la Consejería competente en materia de turismo de la Región de Murcia, en los aspectos de su competencia.

3. Para la concesión de la denominación de Municipio Turístico, que será decretada por el Consejo de Gobierno de la Comunidad Autónoma de la Región de Murcia, a propuesta del Consejero competente en la materia, y oído el parecer del Consejo Asesor Regional del Turismo, se valorarán las siguientes condiciones:

a) Oferta turística local, alojativa y complementaria que pueda justificar la citada denominación.
b) Porcentaje significativo del Presupuesto Municipal anual destinado a promoción e infraestructuras turísticas, excluidas las cantidades destinadas a fiestas locales.
c) Realización de programas que incidan en la calidad de la oferta turística.
d) Existencia de Ordenanzas de Medio Ambiente donde figuren debidamente recogidas medidas para la preservación de los valores medioambientales, así como para el respeto de los derechos a la intimidad, a tranquilidad y el bienestar general de los usuarios turísticos, con el alcance que se determine reglamentariamente; debiendo abarcar aspectos tales como salubridad, seguridad, control de ruidos y olores, y cualquier otro que resulte procedente a los citados fines.
e) El incremento significativo de población que se genere en ese municipio, con ocasión de los períodos vacacionales.
f) Disponer de oficinas de turismo convenientemente señalizadas y dotadas.
g) Aquellas otras especiales circunstancias que, debidamente acreditadas, aconsejen la concesión de la denominación.

---

por sus habitantes de derecho. c) Que se trate de poblaciones en las que el número de plazas turísticas hoteleras o extrahoteleras duplique, al menos, la población de derecho.

2. *Para la declaración de Municipio Turísticos se tendrán en cuanta los siguientes elementos: a) La existencia de planeamiento urbanístico, con especial valoración del sistema de espacios libres. b) La existencia de zonas verdes y espacios libres que sirvan de protección del núcleo histórico edificado. c) El esfuerzo presupuestario realizado por el municipio en relación con la prestación de los servicios municipales obligatorios y de todos aquellos servicios con especial repercusión en el turismo. d) La adopción de medidas de protección y recuperación del entorno natural y del paisaje. e) La adopción de medidas de defensa y restauración del patrimonio cultural y urbano. f) La relevancia de los recursos turísticos existentes en el término municipal*".

[6] **Disposición Adicional Única** (Municipios y zonas de interés turístico): "*Tendrán la consideración de municipios turísticos y de zonas de interés turístico aquellos que sean reconocidos como tal en le Plan Plurianual de Actuación*".

4. Reglamentariamente se determinarán los índices a aplicar para la determinación de los requisitos establecidos en los apartados anteriores, así como el procedimiento de concesión y revocación de la denominación.
5. La concesión de la denominación de "Municipio Turístico", así como la revocación de la misma, será publicada en el Boletín Oficial de la Región de Murcia.

La consideración que nos debe ofrecer la regulación legal de esta figura por parte de nuestro legislador autonómico resulta más que positiva (aunque también contenga ciertos defectos que también expondremos), no sólo porque fue una de las primeras leyes autonómicas en regularla, sino además por los siguientes aspectos que destacamos seguidamente:

- La consideración de "Municipio Turístico" juega un papel relevante dentro de la ordenación territorial de los recursos, espacios y actividades turísticas, ya que tal denominación suele llevar aparejada unos efectos de gran importancia para estos municipios, entre ellos el acceso preferente a medidas de fomento económico y promoción, y participación de la Consejería competente en materia de turismo en la elaboración y modificación de los instrumentos de planificación urbanística de este tipo de municipios previstas en los planes y programas autonómicos, el establecimiento de convenios con las administraciones supramunicipales, etc.

- La regulación legal contenida en el art. 51 de la Ley 11/1997, obedece a un criterio de contenido intermedio, con cierto detalle, si la comparamos con las otras normas autonómicas. Haciendo un análisis más detenido del mismo destacamos los siguientes aspectos:

1) El apartado 1 regula las ventajas o beneficios de modo preferencial que se persiguen con la declaración de "Municipio Turístico" en relación a las acciones de fomento y promoción de la Administración regional, siempre y cuando se reúnan las circunstancias señaladas en el apartado 3 de este mismo precepto. Curiosamente la norma acaba de forma abierta, previendo la posibilidad de que reglamentariamente puedan establecerse diferentes tipos de municipios turísticos, ¿acaso en sentido similar a la Ley valenciana?.

Pero, el aspecto fundamental de este precepto es precisamente la referencia a las ventajas o beneficios que supondría para el "Municipio Turístico" las acciones de fomento y promoción. En concreto, la ley 11/1997 regula ambas: el art. 48 se refiere a la promoción[7] y el art. 49 al fomento de la actividad turística[8]. En cuanto a la promoción, el precepto nada nos dice en que consiste, tan sólo que deben elaborarse Planes de Promoción y Comercialización de la oferta turística de la Región; y en este sentido debemos destacar la elaboración del Plan Director de Turismo de la Región de Murcia 2006/2012, una de cuyas líneas estratégicas de actuación (nº6) recoge tres Planes: de información, promoción y comercialización..

En cuanto al fomento de la actividad turística se lleva a cabo, como dice el precepto, a través de Programas (o Planes) de fomento del Turismo en la Región, lo cuales se configuran en torno a cuatro núcleos fundamentales: formación (impulso de la formación turística y creación de los Centros de Cualificación Turística), información (afianzar y homogeneizar la imagen turística de la Región en la publicidad genérica y creación de la Red de Información Turística), promoción (consolidar la articulación y coordinación de las actuaciones promocionales así

[7] **Art. 48**: "1. *Corresponde a la Consejería competente en materia de turismo la planificación de la política de promoción turística de la Región de Murcia en el territorio nacional y de modo concertado con la Administración del estado, fuera del territorio nacional.*

2. *Bajo la superior dirección del Consejo de Gobierno, la política de promoción turística se diseñará en colaboración con la Corporaciones Locales y los agentes implicados, a través del Consejo Asesor Regional de Turismo*.

3. *Anualmente la Consejería competente en materia de turismo elaborará el Plan de Promoción y Comercialización de la oferta turística de la Región de Murcia, previa consulta la Consejo Asesor Regional de Turismo*".

[8] **Art. 49**: "1. *La Consejería competente en materia de Turismo, y en los términos que reglamentariamente se determine, podrá elaborar programas de fomento de la actividad turística que tendrán entre otras finalidades, las de concentración, modernización y saneamiento de empresas, la difusión de la imagen de la Región de Murcia y la formación técnica y profesional de actividades turísticas. Dichos programas podrán incluir el apoyo técnico o jurídico, la concesión de ayudas económicas o cualquier otra técnica de fomento respecto de las empresas que se estimen merecedoras de apoyo público.*

2. *La Consejería competente en materia de turismo podrá calificar determinados proyectos empresariales como de "interés turístico regional", dotándolos de ayudas excepcionales para su puesta en marcha, para la realización de estudios de mercado, y para su promoción de introducción en mercados turísticos. Dichos proyectos contarán con el apoyo institucional de la Administración Autónoma, quien a su vez recabará el de la Administración Estatal su fuera preciso*".

como afianzar sus fundamentos metodológicos para mejorar la eficacia) y calidad (difusión de la importancia de la calidad en la prestación de los servicios de desarrollo de un plan para implantar sistemas de calidad en las empresas).

2) El apartado 2 se refiere a los instrumentos de planteamiento urbanístico que les correspondan a los "Municipios Turísticos". En tal caso, aunque la elaboración y modificación de los mismos corresponda al propio Municipio, se establece una cierta intervención de la Administración regional: será oída la Consejería competente en materia de Turismo de la Región de Murcia, en los aspectos que sean de su competencia.

3) El apartado 3 ha de considerarse el eje central de este precepto, pues es el que recoge de forma detallada las condiciones que se valorarán para que un Municipio consiga su catalogación como "Municipio Turístico"; condiciones que consideramos un poco rigurosas, y por tanto de difícil cumplimiento en su totalidad. A ello se añade la evidente indeterminación cualitativa y cuantitativa de algunos de los requisitos, por ejemplo, ¿qué se entiende por "porcentaje significativo de del Presupuesto Municipal, previsto en el apartado b)?, o ¿cuál es el incremento significativo de la población, recogido en el apartado e)?. Todo ello lleva a que la Administración competente en materia de turismo disponga de un amplio margen de discrecionalidad; tanto es así que, por el momento, ningún Municipio de la Región ha conseguido tal calificación.

Por otro lado, las condiciones o requisitos enumerados no son taxativos, pues en el último de ellos (apartado g) se deja abierta la norma para que puedan valorarse otras circunstancias que, debidamente acreditadas, aconsejen su concesión, ¿acaso significa, que un Municipio, aún reuniendo los seis requisitos precedentes, queda a voluntad de la Administración competente exigir algunos más en "especiales circunstancias", como expresa la norma?. Pensamos que la enumeración, ya de por sí exhaustiva debiera haber sido taxativa y no dejar abierta la posibilidad a otras opciones con cierto margen a la arbitrariedad, en su caso.

La valoración que ha de llevarse a cabo para la concesión de "Municipio Turístico" es realizada por tres órganos autonómicos: el Consejo de Gobierno de la Comunidad Autónoma de la Región de Murcia, que concederá o no dicha denominación; el Consejero competente en la materia, quien hará la propuesta; y el Consejo Asesor Regional de Turismo, que emitirá su parecer al respecto, pero sólo a efectos de ser oído.

4) El apartado 4 deja a la determinación por vía reglamentaria de los índices para determinar los requisitos anteriores. También quedará por determinar cómo será el procedimiento de concesión y revocación de los mismos.

5) El apartado 5 regula la necesaria publicidad de la concesión de Municipio Turístico una vez realizada, para lo cual se publicará en el Boletín Oficial de la Región de Murcia.

Por último, y pese a lo expuesto, echamos en falta, por un lado, una definición de "Municipio Turístico", pues lo único que contiene la norma es una descripción detallada de requisitos para su concesión contenidos en el párrafo 3 y, por otro lado, la remisión a una reglamentación futura tanto en cuanto a los tipos de municipios turísticos como respecto de los índices para determinar los requisitos de concesión. Ambas cuestiones podían haberse reflejado en la norma directamente

2. DELIMITACIÓN DE OTRAS FIGURAS AFINES: "ÁREAS TURÍSTICAS SATURADAS" Y "ZONAS DE GRAN AFLUENCIA TURÍSTICA".

Dada la similitud entre algunos términos que se vienen utilizando en el ámbito turístico, en muchas ocasiones sin demasiada claridad, se hace necesario deslindar la figura de "Municipio Turístico" de las llamadas, por un lado, "áreas turísticas saturadas" y "zonas de gran afluencia turística". Realmente la distinción más evidente es la que se produce entre "Municipio

Turístico" y "zona de gran afluencia turística, al tratarse de dos figuras bien diferenciadas, aunque puedan coincidir en algún aspecto, como seguidamente vamos a analizar:

2.1. "Áreas Turísticas saturadas" y "Municipios Turísticos".

La aparición del concepto de "áreas turísticas saturadas" responde a una idea muy puntual derivada precisamente de la saturación turística en zonas muy concretas, en las que previamente a la autorización de nuevas actividades, se debieran realizar actuaciones tendentes a solucionar los problemas derivados de la saturación de una oferta con insuficientes infraestructuras para satisfacerla. Por ello, las actuaciones en materia turística, evidentemente, no pueden limitarse sólo al fomento de la oferta sino también a la adecuación de la misma a unos estándares de calidad[9].

Las "Áreas turísticas saturadas" se regulan en el art. 52 de la Ley 11/1997 en el sentido siguiente: "*La Administración Regional podrá proceder a la declaración temporal de 'área turística saturada' en los términos y con los efectos que reglamentariamente se determinen*". No se trata de una norma que nos diga mucho, pues únicamente se ha limitado a acoger la figura, a establecer su carácter temporal (propuesta que deberá formular el servicio de Planificación de la Dirección General de Infraestructuras de Turismo, como órgano directivo al que está encomendada esta función) y a diferir los términos de su contenido y el alcance de sus efectos a un desarrollo reglamentario posterior; aunque en este último caso deberá existir una coordinación legal a la hora de realizar ese desarrollo reglamentario, habiendo de tener en cuenta toda la legislación del suelo, en tanto que las actividades turísticas suponen y constituyen un tipo de uso del suelo, que viene condicionado asimismo por la normativa aplicable en estas materias y por los instrumentos de planeamiento en ella regulados. Por tanto habrá de tenerse en cuenta tanto la Ley del Suelo regional (texto refundido de 10 de Junio de 2005), así como las Directrices, Planes de Ordenación territorial o los Planes urbanísticos municipales. Y en las zonas costeras, el decreto 57/2004, de 18 de Junio, aprobatorio, ya que contiene la regulación de las Directrices y Plan de Ordenación Territorial del Litoral de la Región de Murcia, que podrían verse afectadas por la futura normativa reglamentaria[10].

Con independencia de las críticas que pueda suponer la redacción de este precepto, tanto por la falta de una definición clara como por la remisión en blanco a una futura e hipotética reglamentación posterior, lo cierto es que su inclusión supone una muestra de la preocupación del Gobierno regional por cumplir con la premisa establecida como principio rector en el número 11 del art. 4 de la Ley 11/1997 (acomodar su actuación a la contribución al medio ambiente y conservación de la naturaleza y el paisaje). Pues, lo que resulta evidente es que el fenómeno de la saturación en determinadas zonas tiene una incidencia notable en la calidad de los destinos turísticos. El exceso de visitantes, la congestión, la masificación en definitiva, lleva a que los destinos puedan sufrir un cúmulo de repercusiones negativas: menoscabo en el disfrute del visitante, daños a la flora, la fauna, los valores paisajísticos, etc., incremento de basuras y contaminación, disminución de la eficacia de los servicios turísticos, aprovechamiento al límite de la capacidad de infraestructura local, etc.

Las diferencias más importantes en relación con el concepto de Municipio Turístico serían básicamente tres:

a) El carácter temporal que aparece expresamente previsto para las áreas turísticas saturadas en su art. 52, mientras que nada se dice al respecto con relación al municipio turístico.

b) No se establece requisito alguno para catalogar a una zona como área turística saturada, lo que sí se hace respecto al municipio turístico. Ello lleva a la concesión de un cheque en blanco a la administración turística competente para dicha catalogación.

c) La remisión por vía reglamentaria en cuanto a sus términos y efectos confunde aún más el tema, al carecer de criterio legal alguno.

[9] Dictamen nº 9/1997 del Consejo Económico y Social, sobre el Anteproyecto de la Ley del Turismo de la Región de Murcia, y gracias al cual se incluyó esta figura en el texto definitivo de la Ley 11/1997.

[10] Vid. LÓPER PELLICER, J.A. *Ley del Turismo de la Región de Murcia. Comentarios y legislación complementaria* (Comentario al art. 52). Thomson-Aranzadi, Pamplona 2006, págs. 154 y ss.

2.2. "Zonas de gran afluencia turística" y "Municipios Turísticos".

Las llamadas "zonas de gran afluencia turística", al contrario de las dos figuras anteriores, no se regulan en la Ley 11/1997, sino que quedan amparadas por la Ley 11/2006, de 22 de Diciembre, sobre Régimen del Comercio Minorista y Plan de Equipamientos Comerciales de la Región de Murcia, en concreto en su art. 41[11], que a la vista de su contenido podemos afirmar que se trata de un concepto íntimamente ligado al régimen de la libertad de horarios (arts. 36 a 39). Las diferencias más relevantes entre esta figura y la de "Municipio Turístico" son, a nuestro modo de ver, las siguientes:

a) Observamos que dentro de las características que definen al "Municipio Turístico" no se hace referencia alguna al tema de los horarios comerciales, cuestión que es la predominante en la figura de "zonas de gran afluencia turística". La diferencia más importante entre ambas deriva precisamente de esa cuestión, pues mientras que en el concepto de "Municipio Turístico" predomina la idea de un beneficio derivado de las acciones de fomento y promoción del turismo por parte de la Administración regional, así como la posibilidad de intervenir en el planteamiento urbanístico del Municipio reconocido como turístico; en cambio, en las "zonas de gran afluencia turística" lo que prima es que su determinación va íntimamente ligada al tema de los períodos a que se limite la aplicación del régimen de libertad de horarios.

b) La determinación de una "zona como de gran afluencia turística" no se produce a instancia únicamente de la Administración regional, como ocurre en le caso del Municipio Turístico, sino que dicha solicitud puede provenir también de otras entidades, como los municipios correspondientes, las asociaciones o entidades más representativas del sector. Pero, cabe también la posibilidad de que se produzca a instancia sólo de la Administración, previa consulta de dichas entidades. En estos términos se pronuncia el párrafo 2 del art. 41.

c) El art. 41.1 de la Ley 11/2006 refiere la "zona de gran afluencia turística" a un ámbito territorial más diverso, al reconocer igualmente tanto al término municipal en sí como a una parte del mismo, lo que no ocurre en el "Municipio Turístico", que se corresponde con el ámbito del propio Municipio.

d) El art. 41.1 también refleja un ámbito temporal más restringido para las "zonas de gran afluencia turística" circunscrito a determinados períodos del año en función de la mayor afluencia de turistas, lo que no ocurre con el "Municipio Turístico" que, aunque nada dice en este sentido, ha de presumirse la vigencia del mismo de forma continuada y no periódica, en tanto se cumplan los requisitos previstos en el párrafo 3 del art. 51, y no exista una revocación expresa por parte de la Administración regional en su caso, como prevé el párrafo 5 del art. 51.

e) Y, por último, para la determinación de las "zonas turísticas saturadas" se incluye dos parámetros que permitan apreciar el incremento de población: uno estadístico, al considerar como incremento de la población por turismo la media ponderada anual de población significativamente superior al número de residentes y, otro, totalmente variable e indeterminado cuando la norma se refiere a aquellos términos municipales "*…en los que tenga lugar una gran afluencia de visitantes por motivos turísticos*". Precisamente este último parámetro es la única semejanza entre las figuras que venimos analizando, pues para la definición de Municipio Turístico del art. 51, se habla de "*incremento significativo de la población… con ocasión de los períodos vacacionales*", aunque los términos no sean idénticos, responden a una misma idea, que es el incremento de turistas o visitantes.

Ahora bien, pero, mientras en las "zonas de gran afluencia turística", la afluencia de visitantes ha de ser en momentos determinados o por temporadas como criterio determinante, en le caso del "Municipio Turístico" sólo es uno más de los requisitos a tener en cuenta junto con el resto de requisitos a los que se refiere el párrafo 3 del art. 51, pero no es el criterio determinante.

---

[11] **Art. 41**: "1. *Se considerarán zonas de gran afluencia turística a los efectos de esta Ley los términos municipales o parte de los mismos en los que, en determinados períodos del año, la media ponderada anual de población sea significativamente superior al número de residentes o en los que tenga lugar una gran afluencia de visitantes por motivos turísticos.*

2. *La determinación de las zonas de gran afluencia turística, así como el período o períodos a que se limite la aplicación del régimen de libertad de horarios, será establecido mediante Orden de la Consejería competente en materia de comercio, a solicitud o previa consulta de los municipios correspondientes y de las asociaciones o entidades más representativas del sector*".

3. VENTAJAS ECONÓMICAS.

Tras todo lo expuesto anteriormente, existe un conjunto de indicadores necesarios para el cumplimiento de los fines que la consideración de esta figura conlleva (indicadores estrictamente turísticos, medioambientales, demográficos, económicos, etc.), siendo preciso establecer cuáles serán los que determinen el que un Municipio adquiera la consideración de "Municipio Turístico".

Por consiguiente y a modo de resumen, consideramos que este concepto juega un papel muy relevante dentro de la ordenación territorial de los recursos, espacios y actividades turísticas ya que tal denominación puede llevar aparejados unos efectos económicos de vital importancia para esos municipios, como son, entre los más importantes, los siguientes:

- La incorporación a criterios de calidad a la gestión de empresas y los servicios turísticos.
- El acceso preferente a las medidas de fomento económico previstas en los planes y programas provinciales y/o autonómicos.
- La capacidad para establecer tributos o recargos específicos.
- El establecimiento de convenios con las administraciones supramunicipales.
- La posibilidad de solicitar la declaración de un recurso como recurso turístico esencial y disfrutar de los beneficios que se establezcan a favor de su promoción y protección.
- La participación de las políticas de implantación o mejora de infraestructuras y servicios que sean impulsadas desde las Administraciones autonómicas, etc.

4. CONCLUSIONES.

a) La consideración de un determinado municipio o parte de él como "zona de gran afluencia turística" no está supeditada al cumplimiento de requisitos sustanciales estrictos, ya que la definición legal de dichas zonas aparece formulada en términos notablemente ambiguos, siendo suficiente con que en las mismas *tenga lugar una gran afluencia de visitantes por motivos turísticos*.

b) Por el momento, la declaración como "zona de gran afluencia turística" ha sido formulada respecto a la práctica totalidad de la zona costera, tanto de las provincias limítrofes (Alicante y Almería), como de la propia Región de Murcia (los municipios de Águilas, Los Alcázares, San Javier y San Pedro del Pinatar, así como amplias zonas de Mazarrón), con la única excepción de Cartagena, cuyo casco urbano no cuenta con tal declaración –pese a que otras partes del municipio sí la tienen-, lo que sitúa a dicha zona en una posición de clara desventaja desde el punto de vista de su atractivo turístico-comercial, poniendo así en riesgo el éxito de las cuantiosas inversiones que se vienen realizando en los últimos años con objeto de reforzar dicho atractivo y propiciar, así, la reconversión de la economía de la ciudad.

c) La declaración como "zona de gran afluencia turística" ha de hacerse por Orden de la Consejería competente en materia de comercio de la Comunidad Autónoma, a solicitud o previa consulta de los municipios correspondientes y de las asociaciones o entidades más representativas del sector. La decisión de la Consejería tiene carácter discrecional, y por tanto sólo ha de incluir una motivación sucinta en la que se valore la solicitud desde el punto de vista del interés público.

d) La declaración como "zona de gran afluencia turística" puede ser delimitada, en función de las finalidades que se persigan en cada caso, atendiendo a criterios muy variados: criterio espacial (todo el municipio o una parte de él), temporal (todo el año o períodos determinados, indefinida o para un determinado número de años, con o sin limitación de horarios de apertura y cierre), por tipos, superficies o volúmenes de facturación de los establecimientos, etc.

e) La declaración como "Municipio turístico" no está directamente relacionada con la "zona de gran afluencia turística", pues la primera está regulada por la legislación sobre turismo, que la supedita al cumplimiento de requisitos notablemente rigurosos, pese a que la verdadera trascendencia de la declaración –más allá de lo que pueda suponer en términos de imagen-, se

limita, en la práctica, al reconocimiento de un trato preferencial por parte de la Administración Autonómica en aspectos relacionados con el fomento y promoción del turismo, y a la posibilidad de intervención por parte de aquella en el planeamiento urbanístico del Municipio reconocido como turístico. Por otra parte, la declaración como municipio turístico no lleva consigo, por sí misma, la aplicación de un régimen de libertad de horarios comerciales. En todo caso, la declaración como "Municipio turístico" no ha sido reconocida, por el momento, a ningún Municipio de la Región. Ello debería dar qué pensar a las autoridades autonómicas, máxime a la vista del tiempo transcurrido desde la entrada en vigor de la Ley del Turismo.

BIBLIOGRAFÍA.

CONSEJO ECONÓMICO Y SOCIAL DE LA REGIÓN DE MURCIA (1997), "Dictamen nº 9/1997, sobre el Anteproyecto de la Ley del Turismo de la Región de Murcia", Murcia.

LÓPEZ PELLICER, J.A. (2006), "Ley del Turismo de la Región de Murcia. Comentarios y legislación complementaria", Thomson-Aranzadi, ed, Pamplona.

PASTOR ÁLVAREZ, M. CARMEN, y Otros (2007), "Informe sobre los presupuestos jurídicos que condicionan la posible declaración de Cartagena como zona de gran afluencia turística", Universidad Politécnica de Cartagena.

SISTEMA DE INDICADORES DE TURISMO SOSTENIBLE PARA DESTINOS CONSOLIDADOS

FRANCISCO JAVIER BLANCAS PERAL
Email: fjblaper@upo.es

MACARENA LOZANO OYOLA
Email: mlozoyo@upo.es

FLOR Mª. GUERRERO CASAS
Email: fguecas@upo.es

Departamento de Economía, Métodos Cuantitativos e Historia Económica
Universidad Pablo de Olavide

RAFAEL CABALLERO FERNÁNDEZ
Email: rafael.caballero@uma.es

MERCEDES GONZÁLEZ LOZANO
Email: m_gonzalez@uma.es

FÁTIMA PÉREZ GARCÍA
Email: f_perez@alu.uma.es

Departamento de Economía Aplicada (Matemáticas)
Universidad de Málaga

Resumen

En este trabajo proponemos un sistema de indicadores de turismo sostenible que pueda ser utilizado en los procesos de planificación a nivel local en destinos consolidados. El objetivo es ofrecer un instrumento que proporcione la información necesaria para evaluar la situación de un destino y definir las actuaciones necesarias para mejorar su competitividad. En la definición de los elementos del sistema hemos seguido las directrices de la Organización Mundial del Turismo y las investigaciones que definen a nivel local sistemas de indicadores de sostenibilidad.
Los indicadores del sistema propuesto se dividen en tres dimensiones (social, económica y ambiental), que proporcionan información de tipo clave (a considerar en cualquier tipo de destino) y específica (ligada al segmento turístico de mayor presencia). Para su cuantificación hemos creado una base de datos propia. Finalmente, realizamos análisis exploratorios de la sostenibilidad de los destinos de mayor relevancia de Andalucía.

Palabras clave: Indicadores – Turismo Sostenible – Andalucía.
Área temática: Economía Urbana, Regional y Local.

Abstract

In this paper we propose a system of sustainable tourism indicators to be used in planning processes at a local level in established destinations. This system is aimed to provide suitable information to assess the state of destinations and define new public measures that improving its competitiveness. The elements of this system have been defined following World Tourism Organization guidelines as well as other works presenting systems of local sustainability indicators.
The indicators of the proposed system are related with three dimensions (social, economic and environmental). They provide key information (to be considered in any destination) and specific information (related to the most important tourist segment for destination). We have created a new database in order to quantify the proposed indicators system. Finally, we carry out several tourist sustainability analysis of the most relevant destinations of Andalusia.

Thematic Area: Urban, Regional and Local Economics.

1. INTRODUCCIÓN.

En los últimos años el mercado turístico internacional se ha desarrollado en un contexto de creciente competencia, ya que frente a los destinos tradicionales surgen destinos emergentes, especialmente en países en desarrollo, que intentan obtener ventajas competitivas mediante modelos turísticos tradicionales. Estos modelos se caracterizan por un alto volumen de demanda, bajos precios, estancia media elevada y reducido gasto turístico, una actividad estacional e intensiva, con una alta concentración espacial y, como consecuencia, con importantes impactos negativos sobre el medio cultural y natural.

La creciente importancia de estos nuevos destinos se está traduciendo en una pérdida de competitividad de los destinos consolidados. Para hacer frente a esta situación, los gobiernos regionales han puesto en marcha actuaciones para definir un nuevo modelo turístico diversificado, de calidad y sostenible. Se intenta así configurar una actividad turística que se caracterice por: una demanda más cualificada, menos masificada y no concentrada espacialmente, una estacionalidad decreciente de la actividad, estancias en el destino más fraccionadas y un mayor gasto turístico por parte del visitante. Desde el punto de vista de la oferta, el objetivo es mejorar la calidad y la diversidad de su oferta complementaria, garantizando la conservación de los recursos naturales y culturales al minimizar los impactos que el turismo produce sobre el medio en el que se desarrolla y del que es ampliamente dependiente. En definitiva, los esfuerzos van dirigidos a incrementar la competitividad de los destinos, mejorando la sostenibilidad de la actividad turística.

Por otro lado, debemos tener en consideración que, independientemente del grado de consolidación de su actividad turística, todos los destinos necesitan contar con algún instrumento que permita evaluar su situación y, tomando como base el diagnóstico obtenido, definir las actuaciones necesarias para su transformación en destinos más sostenibles. Para ello, uno de los instrumentos más utilizados en la práctica desde la década de los noventa es la definición de un sistema de indicadores. No obstante, las investigaciones se han centrado en destinos donde el desarrollo de la actividad es más incipiente (Farsari y Prastacos, 2002; Vyas y Kumaranayake, 2006), dejando de lado zonas más consolidadas en las que la gestión de los impactos del turismo resulta más compleja. Asimismo, gran parte de estos trabajos formulan teóricamente el sistema sin llegar a cuantificarlo totalmente, lo que les resta operatividad y dificulta su utilización en la práctica.

La definición de un sistema de indicadores es aún más necesaria en el marco de planificación que empieza a instaurarse en los destinos, donde se da un mayor protagonismo a la dimensión territorial. Así, serán los entes públicos y privados los que promoverán iniciativas de turismo sostenible, de forma que las intervenciones que se lleven a cabo en el sector se plantearán desde el propio territorio al que van destinadas.

Hasta el momento las actuaciones desarrolladas a nivel local no han modificado de forma significativa las pautas de producción y consumo insostenibles propias de los modelos turísticos imperantes. En la práctica, parece no existir ni la información ni los incentivos suficientes para poner en marcha nuevas actuaciones que permitan definir prácticas más sostenibles en el destino. En este contexto, consideramos que una de las principales causas de esta situación es el lenguaje complejo empleado en los informes y comunicaciones relativas a la planificación turística regional, lo que dificulta la actuación a nivel local.

Para solventar esta situación, consideramos que es necesario mejorar el conocimiento de los agentes locales mediante instrumentos de medición que les permitan visualizar los aspectos relacionados con la sostenibilidad turística sobre los que es necesario actuar. Asimismo, estos instrumentos permiten establecer metas y plazos reales para cumplirlas, de forma que puedan transmitirse al colectivo de ciudadanos los resultados obtenidos.

Por todo ello, en este trabajo nos planteamos como objetivo la definición de un sistema de indicadores de turismo sostenible pensado para los destinos cuya actividad turística está más consolidada, considerando la dimensión territorial propia de los recientes procesos de planificación turística. Así, el sistema que proponemos pretende ser una herramienta de análisis para ser utilizada en los procesos de planificación locales, que permita comparar en el tiempo y en el espacio la sostenibilidad de la actividad turística de los destinos. Consideramos que la

utilización de este instrumento permite obtener un mejor conocimiento de la situación del sector para detectar y prevenir los problemas en su inicio, mejorar la toma de decisiones por parte de los agentes implicados en la planificación del sector, diseñar medidas correctivas adecuadas si fuera necesario, etc. Además, el sistema de indicadores es una herramienta que puede adaptarse a las características del segmento turístico que se esté analizando, permitiendo así la formulación de una actuación diferenciada para cada ámbito territorial del espacio turístico.

Este trabajo se estructura como sigue: en el siguiente apartado describimos el sistema de indicadores de turismo sostenible para destinos consolidados que hemos elaborado. En el tercer epígrafe presentamos un análisis estadístico de la base de datos que permite cuantificar el sistema. En la siguiente sección, utilizamos el sistema definido para realizar un estudio exploratorio de la sostenibilidad de los municipios turísticos de mayor relevancia de un destino consolidado como es la Comunidad Autónoma de Andalucía. Por último, se recogen las conclusiones extraídas del trabajo.

2. PROPUESTA DE UN SISTEMA DE INDICADORES DE TURISMO SOSTENIBLE PARA DESTINOS CONSOLIDADOS.

Como hemos mencionado, en este trabajo proponemos como instrumento de evaluación de la sostenibilidad turística un sistema de indicadores. Adoptamos así un enfoque analítico para la medición del grado de sostenibilidad (Organization for Economic Cooperation and Development, 2000) consistente en definir un conjunto de indicadores de turismo sostenible, cuyos niveles absolutos y la dirección en la que cambian, tengan como finalidad indicar si el sector turístico de una determinada región presenta una situación más o menos sostenible. En el ámbito que nos ocupa, consideramos como indicadores de turismo sostenible aquellas *"…medidas que proporcionan la información necesaria para comprender mejor los vínculos y los impactos del turismo con respecto al entorno cultural y natural en el que se desenvuelve la actividad y del que es ampliamente dependiente"* (World Tourism Organization, 1996).

Para la formulación de un sistema de indicadores apropiado para nuestro estudio realizamos las siguientes tareas. En primer lugar, era necesario delimitar dos cuestiones básicas que determinan la definición del sistema: el concepto a medir (turismo sostenible) y el conjunto de destinos que se desean evaluar (los destinos turísticos de mayor relevancia turística de Andalucía).

Por lo que respecta al concepto de turismo sostenible, dado el objetivo de este trabajo, adoptamos una definición institucional basada en *la satisfacción de las necesidades* de las generaciones presentes y futuras como eje fundamental, siguiendo así las directrices de instituciones internacionales como la Organización Mundial del Turismo (OMT). Asimismo, adoptamos una posición de convergencia para interpretar el concepto (Clarke, 1997; Hardy et al., 2002) procediendo a su formulación de forma que ésta permita establecer claramente la estrategia a seguir (Hunter, 2002). De esta forma, consideramos como turismo sostenible aquél que: "*...satisface las necesidades de los turistas y regiones anfitrionas presentes, al mismo tiempo que protege y mejora las oportunidades del futuro. Está enfocado hacia la gestión de todos los recursos de tal forma que se satisfagan todas las necesidades económicas, sociales y estéticas al tiempo que se respeta la integridad cultural, los procesos ecológicos esenciales, la diversidad biológica y los sistemas de apoyo a la vida*" (World Tourism Organization, 1993). Esta formulación general, inspirada en el Informe Brundtland y elaborada por el consenso internacional, es la que tomamos como referencia para nuestro estudio. Fijada esta definición, en los años siguientes se intentó dotarla de contenido estableciendo directrices generales a seguir en los procesos de desarrollo, fijándose de forma clara el el camino a seguir en la práctica (Clarke, 1997; Hardy et al., 2002). En este sentido, se deben desarrollar políticas turísticas que garanticen la protección de los recursos naturales, sociales y culturales que sustentan la actividad y la capacidad de los mismos para la satisfacción de las necesidades de los turistas y las poblaciones residentes presentes y futuras (Sharpley, 2000; Liu, 2003). Así, un modelo turístico sostenible requiere (World Tourism Organization, 2004):

- Dar un uso óptimo a los recursos ambientales, manteniendo los procesos ecológicos esenciales y ayudando a conservar los recursos naturales y la diversidad biológica.
- Respetar la autenticidad sociocultural de las comunidades de residentes, conservando su patrimonio cultural y sus valores tradicionales, así como contribuir al entendimiento y a la tolerancia intercultural.
- Asegurar unas actividades económicas viables a largo plazo, que generen beneficios socioculturales bien distribuidos y que contribuyan a la reducción de la pobreza.
- Una amplia participación informada de todos los agentes implicados en el proceso de planificación y gestión del destino, así como un liderazgo político firme para lograr una colaboración amplia y establecer un consenso.
- Alto grado de satisfacción de los turistas, que propicie una mayor conciencia por parte de éstos sobre los problemas de la sostenibilidad y fomente por su parte prácticas más sostenibles.

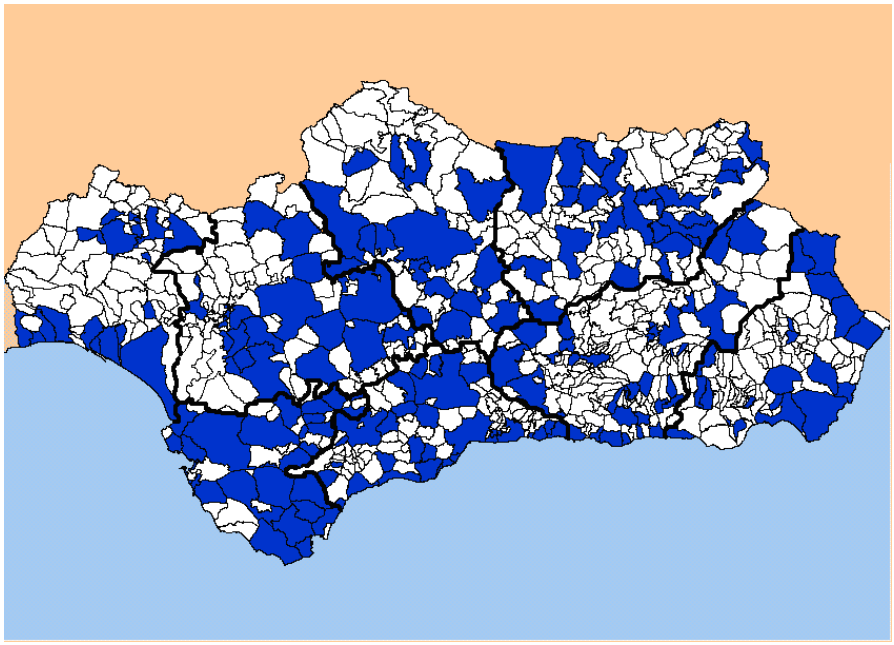
Delimitado el concepto que el sistema debe permitir evaluar, era necesario fijar el conjunto de destinos a analizar puesto que sus características determinan en parte la definición de los elementos del sistema propuesto. En este sentido, siguiendo lo establecido por el Plan General de Turismo Sostenible de Andalucía 20082011, el espacio turístico andaluz puede dividirse en tres grandes ámbitos: costero, urbano y rural, diferenciados por las características de su entorno y el tipo de actividad o segmento turístico de mayor presencia. Para asegurar el carácter ilustrativo de nuestro estudio nos hemos centrado en aquel ámbito en el que se concentra de forma significativa un mayor volumen relativo de demanda, de forma que nos asegure la relevancia turística de la zona estudiada. Así, analizamos el conjunto de destinos de turismo cultural que conforman el ámbito urbano el cual recibe el 55,82% de la demanda total de la región. El ámbito turístico territorial urbano se corresponde con aquellos territorios en los que se desarrolla la actividad turística cuya principal motivación es la cultura. Para delimitar el conjunto de municipios de este ámbito, ante la inexistencia de datos desagregados a nivel local indicativos de la motivación de la visita de los turistas recibidos, decidimos aplicar criterios de oferta. Así, consideramos que en un destino se realiza una actividad turística de tipo cultural si se cumplen las siguientes condiciones:

- La existencia dentro del término municipal de recursos patrimoniales catalogados como Bienes de Interés Cultural, considerando como tales el denominado patrimonio inmueble. Así, seleccionamos en función del número de monumentos, jardines históricos, conjuntos históricos, sitios históricos, zonas arqueológicas y lugares de interés etnológicos presentes en el municipio.
- La inclusión de la visita al término municipal dentro de rutas e itinerarios de tipo cultural que permitan el aprovechamiento turístico de los recursos patrimoniales presentes. En concreto, seleccionamos los destinos que estén incluidos en, al menos, dos rutas calificadas como de tipo cultural por la Consejería de Turismo, Comercio y Deportes de la Junta de Andalucía. La fijación de este umbral no es arbitrario sino que es el resultado de un análisis exploratorio del número de rutas en las que está incluido cada uno de los municipios de la zona de mayor relevancia turística de Andalucía, intentando buscar un umbral que permita descartar aquellos municipios que, aún estando incluidos dentro de alguna ruta, su principal atractivo no fueran sus recursos culturales.

Asimismo, se incluyeron como destinos de turismo cultural aquellos municipios que disponen de un núcleo de población superior a los 20.000 habitantes, cuya actividad turística se considera calificada como de tipo urbana.

De esta forma, el conjunto de destinos que conforman la zona de mayor relevancia turística desde un punto de vista cultural se encuentra formado por 181 municipios que aparecen en el Mapa 1.

Mapa 1. Municipios de turismo cultural de mayor relevancia en Andalucía.



Fuente: Elaboración propia.

Teniendo en consideración las cuestiones anteriores, procedemos en segundo lugar a identificar los aspectos básicos sobre los que es necesario actuar para alcanzar el objetivo de sostenibilidad. Esta tarea nos permite igualmente vencer la limitación impuesta por la ambigüedad de la definición institucional anterior, desagregando el concepto de turismo sostenible en varios componentes. En concreto, siguiendo otros estudios en la materia (Ávila et al., 2002; Dachary y Arnáiz, 2002; Fullana y Ayuso, 2002), presentamos el turismo sostenible dividido en tres dimensiones: social, económica y ambiental. La asignación de los aspectos a cada dimensión la realizamos en función de su naturaleza tomando como base estudios precedentes, cubriendo así la falta de directrices al respecto.

Para identificar estos aspectos básicos procedemos a seleccionar las cuestiones que son necesarias para poder planificar un destino turístico bajo objetivos de sostenibilidad según lo establecido por la OMT (World Tourism Organization, 2004). En concreto, los trabajos elaborados por esta institución ofrecen un listado indicativo de aspectos relativos a la sostenibilidad turística, de forma que la consideración o no de cada uno de ellos dependerá del análisis que se quiera realizar y del tipo de destino. Asimismo, para cada aspecto básico, en estas publicaciones se ofrece el conjunto de indicadores teóricos alternativos más usados y relevantes para su medición, señalando en cada caso la variabilidad indicativa de una situación más sostenible en el destino, y también el carácter clave o específico de los indicadores. En este sentido, se consideran claves los indicadores que evalúan cuestiones esenciales para gestionar el turismo en cualquier destino (tales como la intensidad turística, la estacionalidad de la demanda o los efectos del turismo sobre la comunidad residente); por lo que respecta a los específicos, se definen como aquellos que permiten gestionar factores importantes para el destino ligados a sus propias características (zonas costeras, islas, destinos culturales, destinos de montaña, etc.).

Con toda esta información de partida, nuestra tarea consistía en fijar los criterios más adecuados que nos permitieran realizar una selección de aspectos e indicadores que conformaran un sistema con las características buscadas, mencionadas anteriormente, siendo los criterios seguidos:

- El cumplimiento de los criterios de selección que garanticen el avance hacia una situación de mayor sostenibilidad. Estos criterios son los de validez científica, representatividad, relevancia, fiabilidad, sensibilidad, carácter predictivo, comprensibilidad, comparabilidad, cuantificación, costeeficacia, transparencia y cobertura geográfica (Nardo et al., 2005a; Nardo et al., 2005b).

- La importancia del indicador para la planificación y gestión del destino turístico.
- La disponibilidad de fuentes estadísticas de las que poder obtener datos para cuantificar el indicador o bien la posible elaboración propia de la información estadística necesaria.
- El nivel espacial del análisis, de forma que los indicadores hagan referencia a cuestiones que afecten en distinto grado a las zonas analizadas y no tengan el mismo efecto sobre cada municipio.

Asimismo, dado el carácter indicativo de la información proporcionada por la OMT, al definir nuestro sistema hemos valorado la posibilidad de incluir nuevos aspectos e indicadores apropiados para el estudio de sostenibilidad turística planteado en este estudio. Para ello hemos tomado como referencia investigaciones ya existentes, donde se aborda la definición de sistemas de indicadores de sostenibilidad para territorios de ámbito local (Romagosa y Cuétara, 2001; Castro, 2004; Gallego y Moniche, 2005; Frausto et al., 2006; Sancho y García, 2006). En cualquier caso, los nuevos indicadores incluidos debían cumplir los criterios de selección mencionados anteriormente.

Teniendo en consideración todo lo anterior, presentamos en las Tablas 1, 2 y 3 los 85 indicadores que componen el sistema, indicando en cada caso el aspecto básico y la cuestión específica evaluada, así como la dimensión conceptual a la que pertenecen. Debemos destacar que los indicadores que afectan a aspectos incluidos en diferentes dimensiones se incluyen en ambas. De esta forma, se evita la pérdida de información que se produciría al incluir el indicador sólo en una de las dimensiones sobre las que tiene influencia. Asimismo, en cada dimensión distinguimos entre indicadores claves e indicadores específicos asignando a cada uno de ellos un símbolo que utilizaremos en el resto del análisis. En concreto, los indicadores claves son denotados como I_{Ci} mientras que para los indicadores específicos utilizamos una notación que hace referencia al segmento turístico de mayor presencia en el ámbito: turismo cultural $I_{ECU\,i}$. En cuanto a la variabilidad deseada, en el sistema distinguimos entre dos tipos de indicadores: positivos y negativos. Consideramos que un indicador es positivo si un mayor valor representa una mejora de la sostenibilidad turística y negativo si ocurre lo contrario.

Así, los indicadores contemplados dentro de la dimensión social son los que aparecen en la Tabla 1.

Tabla 1. Indicadores de turismo sostenible de la dimensión social.

ASPECTO BÁSICO	CUESTIÓN EVALUADA	I_j	SIGNO	INDICADOR
Efectos socioculturales del turismo en la comunidad residente	Capacidad de servicios para actividades deportivas	I_{C1}	Positivo	Dotación de espacios deportivos
	Capacidad de servicios sanitarios	I_{C2}	Positivo	Dotación de centros sanitarios
	Capacidad de servicios de transportes	I_{C3}	Positivo	Dotación de vehículos de transporte de viajeros
	Capacidad de servicios financieros	I_{C4}	Positivo	Dotación de establecimientos de servicios financieros
	Capacidad de otros servicios	I_{C5}	Positivo	Dotación de establecimientos de actividades del sector servicios
	Capacidad de servicios farmacéuticos	I_{C6}	Positivo	Dotación de farmacias
Seguridad pública local	Valoración de la seguridad en el destino	I_{C7}	Positivo	Valoración de la seguridad en el destino por parte del turista
Conservación del patrimonio cultural	Patrimonio bajo protección cultural	I_{C8}	Positivo	Número de sitios designados bajo figura de protección
	Aportaciones voluntarias para protección del patrimonio	I_{C9}	Positivo	Número de voluntarios culturales
	Intensidad de uso del patrimonio	I_{C10}	Negativo	Presión sobre el patrimonio cultural
	Interpretación del patrimonio	I_{ECU3}	Positivo	Número de guías expertos en interpretación
	Protección de las tradiciones culturales	I_{ECU6}	Positivo	Número de fiestas y costumbres conservadas

Tabla 1 (cont.)

ASPECTO BÁSICO	CUESTIÓN EVALUADA	I_j	SIGNO	INDICADOR
Efectos sobre la estructura de la población local	Mantenimiento del nivel de población	I_{C11}	Negativo	Variación de la población
	Incremento de la población joven	I_{C12}	Positivo	Porcentaje de población joven
	Envejecimiento de la población	I_{C13}	Negativo	Porcentaje de población mayor no activa
	Densidad de población	I_{C14}	Negativo	Concentración de población por superficie
	Mantenimiento del nivel de población	I_{C15}	Negativo	Saldo migratorio
	Mantenimiento del nivel de población	I_{C16}	Negativo	Crecimiento vegetativo
Carga social del destino	Imposición de la cultural extrajera (presión sobre la cultura de los residentes)	I_{C17}	Negativo	Porcentaje de población extranjera
	Capacidad de carga social	I_{C18}	Negativo	Población turística por habitante local
Efectos sobre el bienestar de la población Efectos sobre el bienestar de la población	Efecto de los condicionantes sociales sobre la longevidad de la población	I_{C19}	Positivo	Esperanza de vida
	Efectos sobre la renta disponible	I_{C20}	Positivo	Variación de la renta disponible
	Efectos sobre los niveles educativos de la población	I_{C21}	Positivo	Porcentaje de población escolarizada en niveles no obligatorios
	Efectos sobre la situación de dependencia poblacional	I_{C22}	Negativo	Índice de dependencia demográfica general
	Efectos sobre el acceso a la vivienda	I_{C26}	Negativo	Valor catastral de los inmuebles por habitante
Mejora del paisaje urbano	Rehabilitación de edificios	I_{ECU1}	Positivo	Porcentaje de edificios rehabilitados
	Rehabilitación de edificios (coste de la protección cultural)	I_{ECU4}	Positivo	Fondos destinos a la rehabilitación de edificios
	Mejora del entorno urbano	I_{ECU5}	Positivo	Fondos destinados a la mejora del entorno urbano físico

Fuente: Elaboración propia.

Con los indicadores contemplados en esta dimensión se evalúan aspectos señalados como básicos por la OMT relacionados con los efectos socioculturales de la actividad turística, la conservación del patrimonio cultural y la seguridad pública en el destino. Tras la revisión de otras investigaciones realizadas en el campo de la sostenibilidad local mencionadas anteriormente, pudimos incluir en el sistema tres aspectos no contemplados por la OMT. En primer lugar, se seleccionaron indicadores que permitieran analizar los efectos que el desarrollo turístico del destino puede tener sobre la estructura de su población. En concreto, hacemos especial hincapié en la variación, tanto en número como en composición, que la población del municipio puede sufrir ante una excesiva e inadecuada explotación turística de la zona, la atracción de turistas que deciden residir de forma definitiva en el municipio, etc. Un segundo aspecto, está compuesto por indicadores que permitan tener en cuenta la capacidad de carga social del destino, de forma que pueda controlarse el nivel de actividad turística para evitar situaciones negativas en cuanto al rechazo de la población local. Finalmente, un tercer aspecto incluido en el sistema está formado por indicadores que permiten evaluar los posibles efectos indirectos que la actividad turística del destino pueda provocar sobre el nivel de bienestar de su población.

En el caso de la dimensión económica, el sistema de indicadores teóricos propuestos son los que aparecen en la Tabla 2.

Tabla 2. Indicadores de turismo sostenible de la dimensión económica.

ASPECTO BÁSICO	CUESTIÓN EVALUADA	I_j	SIGNO	INDICADOR
Beneficios económicos para la comunidad y el destino	Volumen de demanda turística	I_{C23}	Positivo	Número de turistas recibidos
	Duración de la estancia	I_{C24}	Positivo	Estancia media
	Ingresos generados por el turismo	I_{C25}	Positivo	Gasto turístico
	Inversión en inmuebles	I_{C26}	Positivo	Valor catastral de los inmuebles
	Empleo terciario generado	I_{C27}	Positivo	Proporción de empleados en el sector servicios
	Desempleo en el destino	I_{C28}	Negativo	Tasa de desempleo
	Inversiones en el sector servicios	I_{C29}	Positivo	Volumen de inversiones terciarias registradas
	Comunicaciones telefónicas	I_{C30}	Positivo	Dotación de líneas telefónicas en servicio
	Comunicaciones Internet	I_{C31}	Positivo	Dotación de líneas RDSI en servicio
	Comunicaciones Internet	I_{C32}	Positivo	Dotación de líneas ADSL en servicio
	Renta disponible por habitante	I_{C33}	Positivo	Renta neta declarada por habitante
Mantenimiento de la satisfacción de los turistas	Satisfacción global del visitante	I_{C34}	Positivo	Nivel de satisfacción de los turistas
	Satisfacción del turista por la relación calidadprecio	I_{C35}	Positivo	Percepción de la relación calidadprecio
	Fidelidad de la demanda	I_{C36}	Positivo	Porcentaje de turistas que vuelven a visitar el destino
	Satisfacción de la visita de los sitios culturales del destino	I_{ECU8}	Positivo	Nivel de satisfacción de los visitantes de los sitios culturales del destino
Control del desarrollo	Ordenación del espacio turístico	I_{C37}	Positivo	Existencia de algún plan de ordenación territorial que afecte al turismo
Oferta de variedad de experiencias	Oferta de alojamiento turístico reglado	I_{C38}	Positivo	Plazas de alojamiento reglado ofertadas
	Calidad de la oferta de alojamiento turístico reglado	I_{C39}	Positivo	Plazas de alojamiento reglado de alta categoría
	Oferta de alojamiento no reglada			
	Oferta de servicios de restauración	I_{C41}	Positivo	Plazas de restauración ofertadas
	Promoción de experiencias turísticas	I_{C42}	Positivo	Dotación de oficinas de información turística
	Promoción de experiencias turísticas	I_{C43}	Positivo	Existencia de página web propia
Estacionalidad de la actividad turística	Estacionalidad de la oferta de alojamiento	I_{C44}	Positivo	Porcentaje de establecimientos de alojamiento abiertos todo el año
	Estacionalidad de la demanda	I_{C45}	Positivo	Proporción entre el número de turistas en periodos de baja afluencia respecto a los de afluencia máxima
	Estacionalidad del empleo turístico	I_{C46}	Positivo	Porcentaje de puestos de trabajo del sector turístico que son fijos
Empleo turístico	Volumen de empleo turístico	I_{C47}	Positivo	Número de empleados en el sector turístico
	Peso del empleo turístico dentro del empleo total del destino	I_{C48}	Positivo	Porcentaje de empleados en el sector turístico respecto al empleo total

Tabla 2 (cont.)

ASPECTO BÁSICO	CUESTIÓN EVALUADA	I_j	SIGNO	INDICADOR
Transportes relacionados con el turismo	Capacidad de los servicios de transportes públicos	I_{C3}	Positivo	Dotación de vehículos de transporte de viajeros
	Acceso a través de aeropuerto	I_{C49}	Positivo	Tiempo de acceso al aeropuerto más cercano
	Acceso a través de autopista	I_{C50}	Positivo	Tiempo de acceso a la autopista más cercana
	Acceso a través de carretera	I_{C51}	Positivo	Tiempo de acceso a la carretera nacional más cercana
	Acceso a través de ferrocarril	I_{C52}	Positivo	Tiempo de acceso a la estación de ferrocarril de acceso más cercana
	Red de carreteras para transporte público	I_{C74}	Positivo	Densidad de la red de carreteras
Competitividad del destino	Tasas de ocupación en alojamientos oficiales	I_{C53}	Positivo	Grado de ocupación medio en establecimientos de alojamiento reglado
	Rutas turísticas creadas	I_{ECU2}	Positivo	Número de rutas turísticas que incluyen al destino en su itinerario
	Explotación de las rutas	I_{ECU3}	Positivo	Número de guías expertos en interpretación
Inversión cultural	Coste de la protección cultural	I_{ECU4}	Positivo	Fondos destinos a la rehabilitación de edificios
Aglomeración	Distribución espacial de los sitios que se pueden visitar en el destino	I_{ECU7}	Positivo	Número de recorridos e itinerarios dentro del municipio

Fuente: Elaboración propia.

En este caso, el sistema propuesto contiene indicadores que hacen referencia a todos los aspectos calificados como básicos por la OMT. Además, el análisis de la literatura en materia de sostenibilidad local considerada no nos ha permitido seleccionar nuevos aspectos dentro de la dimensión económica que permitiera complementar la información contenida en el sistema, aunque sí identificar nuevos indicadores para evaluar los aspectos propuestos por la OMT.

Finalmente, la definición del sistema teórico presentado en este apartado se completa con los indicadores correspondientes a la dimensión ambiental que se muestran en la siguiente Tabla 3.

Tabla 3. Indicadores de turismo sostenible de la dimensión ambiental.

ASPECTO BÁSICO	CUESTIÓN EVALUADA	I_j	SIGNO	INDICADOR
Protección de los ecosistemas	Protección de los recursos naturales	I_{C54}	Positivo	Superficie natural protegida
Gestión energética	Energía	I_{C55}	Negativo	Consumo de energía
	Energía no renovable	I_{C56}	Positivo	Porcentaje de consumo energético de recursos renovables
Disponibilidad y gestión del agua	Consumo de agua	I_{C57}	Negativo	Volumen total diario de agua consumida
	Reutilización de agua	I_{C58}	Positivo	Volumen de agua reutilizada atribuible al turismo
Aguas residuales	Tratamiento	I_{C59}	Positivo	Volumen de aguas residuales que reciben tratamiento
	Instalaciones de tratamiento	I_{C60}	Positivo	Existencia de depuradoras de aguas residuales

Tabla 3 (cont.)

Gestión de residuos sólidos urbanos	Producción de residuos	I_{C61}	Negativo	Volumen de residuos producidos
	Residuos reciclados (vidrio)	I_{C62}	Positivo	Volumen de residuos reciclados con respecto al volumen total de residuos
	Valoración de la limpieza del destino	I_{C63}	Positivo	Percepción de la limpieza del destino por parte del turista
	Instalaciones de tratamiento	I_{C64}	Positivo	Existencia de instalaciones de tratamiento de residuos sólidos urbanos
	Recogida selectiva de papelcartón	I_{C65}	Positivo	Dotación de contenedores de papelcartón
	Residuos reciclados (papelcartón)	I_{C66}	Positivo	Volumen de papel y cartón recogido
	Recogida selectiva de vidrio	I_{C67}	Positivo	Dotación de contenedores de vidrio
Contaminación atmosférica	Contaminación acústica	I_{C68}	Negativo	Nivel diurno de contaminación acústica
	Contaminación acústica	I_{C69}	Negativo	Nivel nocturno de contaminación acústica
	Emisiones de contaminantes	I_{C70}	Negativo	Niveles de emisión de contaminantes a la atmósfera
Gestión del impacto visual de la infraestructura y las instalaciones	Impacto de las construcciones	I_{C71}	Negativo	Densidad de construcción por unidad de superficie
	Estado de erosión	I_{C72}	Negativo	Superficie total en estado de erosión
	Conservación del paisaje del destino	I_{C73}	Positivo	Superficie total destinada a paisajes
	Impacto de la red viaria	I_{C74}	Negativo	Densidad de la red de carreteras
Intensidad de uso	Intensidad turística	I_{C75}	Negativo	Turistas recibidos por superficie
	Uso del suelo	I_{C76}	Negativo	Construcciones desocupadas
Gestión ambiental	Organismo gestor en materia ambiental	I_{C77}	Positivo	Existencia de unidad administrativa de medio ambiente

Fuente: Elaboración propia.

Los aspectos contemplados en la Tabla 3 coinciden con los calificados como básicos en el sistema indicativo de la OMT con excepción de las cuestiones relacionadas con las políticas y prácticas de gestión ambiental en los negocios turísticos, puesto que la falta de información estadística nos ha imposibilitado considerar esta cuestión en nuestro estudio. Por lo que respecta a la inclusión de nuevos aspectos no considerados por la OMT destaca el que denominamos gestión ambiental. La inclusión del mismo la hemos realizado siguiendo otras investigaciones de sostenibilidad turística local (Sancho et al., 2007). Con la inclusión de este nuevo aspecto ambiental, nuestro objetivo es evaluar la respuesta administrativa del municipio para garantizar la consideración de las cuestiones ambientales dentro de las actuaciones públicas que se lleven a cabo. Asimismo, la revisión realizada de nuevo nos ha permitido la inclusión de indicadores adicionales para la cuantificación de los aspectos ya considerados por la OMT, permitiendo la evaluación de nuevas cuestiones específicas.

Identificado el conjunto de indicadores que componen el sistema propuesto, para su aplicación práctica procedimos a seleccionar las variables o medidas de cuantificación más adecuadas en cada caso teniendo en cuenta la información estadística existente. Asimismo hemos tenido en consideración trabajos especializados que nos ofrecieran una guía en la definición de las medidas de cuantificación más adecuadas. Así, por un lado, consideramos los trabajos cuyo principal objetivo era la definición de indicadores de sostenibilidad para territorios de ámbito local (Romagosa y Cuétara, 2001; Castro, 2004; Gallego y Moniche, 2005; Frausto et al., 2006; Sancho y García, 2006), y por otro lado, tomamos como guía trabajos elaborados por instituciones públicas que recogen experiencias regionales, nacionales e internacionales y directrices generales relativas a la cuantificación de indicadores de sostenibilidad, conocidas las fuentes de información estadísticas existentes. Dentro de este grupo podemos destacar los siguientes trabajos consultados:

- Los informes metodológicos elaborados por EUROSTAT sobre la medición del desarrollo sostenible del turismo (EUROSTAT, 2006).
- El Conjunto Básico de Indicadores desarrollado por la Agencia Europea de Medio Ambiente (Agencia Europea de Medio Ambiente, 2008).
- Los Indicadores elaborados por el Observatorio de Sostenibilidad de España (OSE, 2008).
- El Banco Público de Indicadores Ambientales del Ministerio de Medio Ambiente (Ministerio de Medio Ambiente, 2008).
- Las experiencias internacionales sobre la medición de la sostenibilidad en ciudades recogida en las Bases para un Sistema de Indicadores de Medio Ambiente Urbano de Andalucía elaborado por la Consejería de Medio Ambiente (Consejería de Medio Ambiente, 2001).
- El Informe sobre los indicadores locales de sostenibilidad utilizados por los municipios españoles firmantes de la Carta de Aalborg (Hernández, 2003).
- Los informes de sostenibilidad de Andalucía elaborados en los últimos años por el Observatorio Ambiental de Andalucía y la Fundación EOI Medio Ambiente (Fundación EOI de Andalucía y Observatorio Ambiental de Andalucía, 2008).
- El Informe sobre los indicadores locales de sostenibilidad elaborado por la Dirección General de la Vivienda, la Arquitectura y Urbanismo del Ministerio de Fomento (Ministerio de Fomento, 2004).
- El Sistema municipal de Indicadores de Sostenibilidad elaborado para el caso de Barcelona por su Diputación Provincial (Diputación Provincial de Barcelona, 2000).
- El informe sobre indicadores para la sostenibilidad elaborado por la Diputación Provincial de Jaén en el marco de su Agenda 21 (Diputación Provincial de Jaén, 2002).

De esta forma, en función de la medida finalmente utilizada para la cuantificación, hemos distinguido entre:

- *Indicadores de evaluación directa*, entendiendo por tales los que pueden ser cuantificados a través de los datos extraídos de fuentes estadísticas, bien en términos absolutos o bien mediante ratios definidos a partir de una variable de referencia. En este último caso, se obtiene una medida de cuantificación relativa que facilita las comparaciones entre los destinos analizados.
- *Indicadores de evaluación indirecta*. Dentro de este grupo consideramos los indicadores para los que no existe un dato estadístico que permita de forma directa cuantificar el aspecto evaluado. Es decir, son indicadores cuyo valor es inferido a partir de los datos disponibles para territorios que engloban a los destinos considerados. Este tipo de cuantificación resulta especialmente relevante para la evaluación de los aspectos en los que es necesario estimar la cuantía atribuible a la actividad turística de la zona. Asimismo, se incluyen dentro de este grupo los aspectos que sólo pueden ser evaluados mediante las aproximaciones que ofrecen los datos provinciales disponibles.

Realizado todo lo anterior, era necesario definir una base de datos adecuada que nos permitiera cuantificar la totalidad del sistema propuesto. Al análisis relativo a la base de datos asociada al sistema de indicadores dedicamos el siguiente apartado.

3. BASE DE DATOS: ANÁLISIS ESTADÍSTICO.

Dado el carácter desagregado del análisis planteado en este trabajo, ha sido necesario elaborar una base de datos propia que nos permitiera cuantificar totalmente el sistema

propuesto. Para ello, dado que nuestro sistema de indicadores recoge aspectos de muy diversa naturaleza (social, económica y ambiental), ha sido necesario utilizar múltiples fuentes elaboradas por diferentes instituciones. No obstante, en todo momento hemos intentado utilizar una misma fuente para cuantificar un mismo indicador en todos los destinos considerados, de forma que se asegure la homogeneidad en las evaluaciones obtenidas.

Por otra parte, dado el ámbito espacial del análisis, era necesario recopilar información estadística desagregada a nivel local o municipal. Actualmente, a pesar de los avances que se han producido en los últimos años, la información correspondiente al nivel municipal es muy escasa y resulta de difícil acceso. De hecho, en algunos municipios determinadas instituciones elaboran series de datos desagregados que no se hacen públicos, dándoles un uso interno dentro de la institución. Por todo ello, nos marcamos como objetivo aprovechar al máximo la información estadística oficial disponible al cuantificar el sistema planteado.

En este sentido, el Instituto de Estadística de Andalucía (IEA) ha sido una fuente de información fundamental en nuestra investigación a través del denominado *Sistema de Información Multiterritorial de Andalucía* (SIMA). En concreto, esta herramienta permite acceder a datos sobre el entorno físico, demográfico, económico y social de cualquier ámbito territorial de la Comunidad Autónoma de Andalucía, ya sea regional, provincial o municipal. Toda esta información se presenta organizada a través de una estructura jerárquica para garantizar un fácil acceso a las variables que conforman el Banco de Datos del sistema. Otras publicaciones del IEA que han servido como soporte a la investigación han sido los Anuarios Estadísticos de Andalucía, que suministran información relevante para realizar estimaciones a nivel local a partir de los datos disponibles para niveles territoriales de mayor dimensión (provincia, regional, etc.).

Otros organismos a destacar de los que se ha obtenido información estadística son los Observatorios Turísticos Provinciales y las Consejerías de Medio Ambiente, Cultura y Turismo, Comercio y Deportes de la Junta de Andalucía. Los estudios y estadísticas elaboradas por los mismos y consultadas online o bien a través de sus respectivos Centros de Documentación han resultado fundamentales para la obtención de la información necesaria para cuantificar algunos indicadores.

No obstante, la búsqueda de información no se ha reducido a organismos de ámbito regional. Publicaciones y estadísticas elaboradas por organismos como el Instituto Nacional de Estadística, el Ministerio del Interior, el Ministerio de Medio Ambiente o el Ministerio de Sanidad han sido consultadas igualmente para complementar la información no elaborada por los organismos regionales.

Otra cuestión a destacar en este proceso es el uso de fuentes estadísticas secundarias internas. Para ello, tuvimos que ponernos en contacto con los Servicios de Difusión de las Consejerías de la Junta de Andalucía solicitando información estadística desagregada a la que se le daba un uso interno, sin proceder a su publicación. En todos los casos, los datos recibidos no proporcionaban el valor de la variable solicitada para cada destino, por lo que tuvimos que proceder a agregar la información municipio por municipio.

Entre la información solicitada podemos destacar las estimaciones suministradas por la Consejería de Turismo, Comercio y Deportes relativa a las variables de demanda turística tales como el número de turistas recibidos, la estancia media, el grado de ocupación, etc., todas ellas obtenidas de la explotación de las Encuesta de Ocupación en Establecimientos Hoteleros y la Encuesta de Ocupación en Acampamentos Turísticos. Las últimas estimaciones disponibles estaban referidas al año 2006 dado el retraso existente en cuanto a la elaboración estadística municipal. Por la importancia de la información referente a la demanda turística en un estudio como el planteado en este trabajo, toda la información estadística con la que se trabajó se homogeneizó en torno al ejercicio correspondiente a ese año.

Finalmente, la cuantificación completa del sistema de indicadores definido requería la elaboración de nueva información estadística no disponible hasta este momento. El trabajo realizado para la generación de nueva información estadística desagregada a nivel municipal constituye una de las principales aportaciones de este trabajo. La búsqueda y elaboración de

nueva información se ha llevado a cabo realizando tres tipos de tareas. En primer lugar, cuantificamos la información de tipo cualitativo procedente de fuentes como, por ejemplo, el Directorio de Empresas y Entidades Relacionadas con el Medio Ambiente o los mapas de carreteras, de forma que la información desagregada contenida en las mismas pudiera agregarse para cada destino. Una segunda tarea consistió en realizar encuestas telefónicas a las entidades con sede en los distintos municipios de las que necesitábamos obtener la información necesaria para cuantificar los indicadores, como por ejemplo, el número de afiliados en asociaciones de voluntariado cultural. Finalmente, en tercer lugar se procedió a recopilar y cuantificar información de Ayuntamientos, Webs de información turísticas, Observatorios Turísticos, Diputaciones Provinciales, Patronatos Provinciales de Turismo y Consorcios Turísticos, con los que se mantuvo un contacto constante a lo largo de todo el proceso de elaboración de la base de datos.

De esta forma, la obtención de la información estadística que compone la base de datos se ha realizado de tres formas diferentes, distinguiéndose entre:

- *Información directa*, proporcionada por datos estadísticos disponibles y publicados por el organismo correspondiente.
- *Información indirecta*, hace referencia a datos estadísticos que están disponibles de forma interna en el organismo correspondiente y que se han solicitado y agregado para obtener el valor de las variables para cada uno de los municipios.
- *Información propia*, que incluye los datos estadísticos que hemos elaborado durante nuestra investigación.

Para obtener una visión general del comportamiento de los destinos en cada indicador hemos realizado un análisis descriptivo desde un punto estadístico de la base de datos creada. Junto a los principales medidas estadísticas descriptivas hemos utilizado el *Coeficiente Alpha de Cronbach* (Cronbach, 1951) para estimar la consistencia interna de la base de datos en cada dimensión, analizando si el conjunto de indicadores considerados en cada una de ellas mide el mismo concepto subyacente.

Los resultados obtenidos en este análisis nos permiten concluir que los indicadores económicos considerados en el sistema propuesto presentan una variabilidad elevada entre los destinos analizados. Prácticamente la totalidad de los indicadores presenta un coeficiente de variación superior al 80% confirmándose un alto grado de heterogeneidad entre los destinos, aún presentando tamaños poblacionales similares. Atendiendo a la forma de las distribuciones de cada indicador, podemos afirmar que predominan las distribuciones leptocúrticas y asimétricas positivas. Esto implica que los valores de los indicadores para los destinos analizados presentan una alta concentración alrededor de los valores medios, encontrándose el mayor grado de concentración en los valores menores a la media. No obstante, encontramos indicadores cuyas distribuciones no cumplen estas características. Así, por un lado, se obtienen distribuciones asimétricas con una mayor concentración de los valores de los indicadores en los valores mayores a la media y, por otro, distribuciones platicúrticas con una baja concentración de los valores de los indicadores alrededor de los valores medios.

Para analizar la consistencia interna de los indicadores de cada dimensión mediante el coeficiente Alpha de Cronbach, dada la alta variabilidad de los datos que hemos comentado, procedemos previamente a estandarizar los indicadores restándole su media y dividiendo por su desviación típica como es práctica habitual en estos casos (Nardo et al., 2005b). Así, el valor obtenido para el coeficiente una vez estandarizados los datos supera en todas las dimensiones el valor 0,6 considerándose, por tanto, que los indicadores están midiendo el mismo concepto subyacente (Nardo et al., 2005a).

Llegados a este punto, para facilitar la utilización de la información contenida en el sistema de indicadores propuesto, nos planteamos realizar en el siguiente apartado análisis exploratorios de sostenibilidad identificando grupos de destinos con características similares.

4. ANÁLISIS EXPLORATORIOS DE LA SOSTENIBILIDAD DE LOS DESTINOS ANDALUCES.

Cuantificado el sistema propuesto, en este apartado llevamos a cabo un estudio exploratorio de sostenibilidad turística de los destinos considerados. Para ello realizamos un análisis cluster con el objetivo de identificar grupos homogéneos de destinos con unas características similares. En concreto, planteamos un análisis cluster de tipo jerárquico diferente para cada una de las dimensiones de turismo sostenible, utilizando la distancia euclídea al cuadrado y, como método de conglomeración, el método de Ward para que la pérdida de la información resultante al formar los conglomerados fuese mínima.

Para determinar los rasgos de cada grupo obtenido, analizamos la situación que presentan los municipios del mismo por término medio en los indicadores de cada uno de los aspectos básicos contemplados. En las tablas que aparecen a continuación se señalan los aspectos básicos en los que cada grupo muestra una mejor situación. Esta información puede permitir a los gestores regionales determinar las líneas prioritarias de actuación en el sector turístico. En este marco, serán los agentes locales quienes, a la vista de la situación particular de la actividad turística del municipio de su competencia, definan las actuaciones necesarias para mejorar su situación.

Como muestra la Tabla 4, en el caso de la dimensión social se definen ocho grupos de destinos. Las debilidades mostradas por cada grupo señalan que las líneas de actuación en términos de sostenibilidad social deben dirigirse a: reforzar los efectos positivos que el turismo tiene sobre la dotación de servicios sociales a disposición de la población residente, potenciar la conservación de los recursos culturales que sustentan la actividad turística y la mejora del entorno urbano, todo ello vigilando los niveles de carga social que ejerce la población visitante.

Tabla 4. Grupos de destinos: dimensión social.

DIMENSIÓN SOCIAL	GRUPO 1	GRUPO 2	GRUPO 3	GRUPO 4	GRUPO 5	GRUPO 6	GRUPO 7	GRUPO 8
Efectos socioculturales en la comunidad residente			X	X			X	
Seguridad	X		X	X	X		X	X
Conservación del patrimonio					X	X		X
Estructura de la población local	X	X	X	X				
Carga social	X				X	X		
Bienestar de la población	X	X			X	X	X	X
Mejora del paisaje urbano			X	X	X	X		

Fuente: Elaboración propia.

En el caso de la dimensión económica, el análisis cluster identifica seis grupos de destinos (Tabla 5). Las debilidades mostradas por cada uno de ellos nos permiten concluir que las principales líneas de actuación regional deben potenciar la definición de una actividad turística que genere mayores beneficios económicos, sobre todo en términos de empleo directo e indirecto. Asimismo, desde el punto de vista de la oferta, es recomendable mejorar la dotación de recursos destinos a la inversión y protección cultural; distribuir más uniformemente el volumen de demanda entre los distintos sitios visitables del destino y, finalmente, asegurar el incremento de la oferta turística complementaria mediante la creación y explotación de nuevas rutas e itinerarios turísticos de tipo cultural.

Tabla 5. Grupos de destinos: dimensión económica.

DIMENSIÓN ECONÓMICA	GRUPO 1	GRUPO 2	GRUPO 3	GRUPO 4	GRUPO 5	GRUPO 6
Beneficios económicos					X	X
Satisfacción de los turistas		X	X			X
Control del desarrollo	X	X		X	X	X
Oferta turística	X		X			X
Estacionalidad		X	X		X	
Empleo						X
Transportes		X			X	X
Competitividad	X				X	X
Rutas turísticas					X	
Inversión cultural				X		
Aglomeración	X				X	

Fuente: Elaboración propia.

Por lo que respecta a la dimensión ambiental, el número de grupos resultante del análisis cluster se reduce a cinco como se muestra en la Tabla 6.

Tabla 6. Grupos de destinos: dimensión ambiental.

DIMENSIÓN AMBIENTAL	GRUPO 1	GRUPO 2	GRUPO 3	GRUPO 4	GRUPO 5
Protección ambiental		X	X		
Gestión energética	X		X		
Gestión del agua			X		X
Aguas residuales				X	X
Residuos sólidos urbanos				X	X
Contaminación atmosférica		X	X		
Impacto visual de las infraestructuras		X			
Intensidad de uso	X		X		
Gestión ambiental	X			X	X

Fuente: Elaboración propia.

En términos generales, en comparación con las dimensiones anteriores, podemos afirmar que los destinos considerados presentan menos fortalezas en cuanto a las cuestiones ambientales. Se hace necesario así potenciar la mejora de todos los aspectos contemplados en el sistema, destacando especialmente la gestión del impacto visual de las infraestructuras y las instalaciones presentes en el destino. En este sentido, los destinos deben buscar alcanzar fortalezas en cuanto a los problemas de erosión, la conservación del paisaje, el impacto de la red viaria o la densidad de construcción.

5. CONCLUSIONES.

En el presente trabajo definimos y cuantificamos totalmente un sistema de indicadores de turismo sostenible pensado y adaptado a la realidad de destinos consolidados, cuya principal actividad turística sea de tipo cultural. Para poner en práctica el sistema hemos utilizado como ejemplo ilustrativo el caso de los municipios turísticos andaluces de mayor relevancia dentro del espacio turístico urbano. Así, hemos elaborado una base de datos propia que no sólo sirve de guía para la cuantificación del sistema propuesto en otros destinos, sino que muestra en cierta medida la capacidad de la información estadística actual para realizar un análisis multidimensional de sostenibilidad.

Configuramos así una herramienta que ofrece información desagregada a los gestores locales, que puede contribuir a solventar el problema actual de inacción por parte de los mismos en los procesos de planificación del sector, como ya hemos comentado. De esta forma, como ilustra este trabajo, el sistema de indicadores puede ser usado para establecer líneas generales de actuación en los procesos de planificación regional. Al mismo tiempo, los gestores locales pueden utilizar la información que el sistema ofrece sobre su situación para formular una estrategia adecuada para solventar sus debilidades y aprovechar al máximo sus fortalezas. Dicha estrategia será diferente en función del grupo al que pertenece cada municipio en cada dimensión y de los valores de los indicadores en cada uno de los aspectos.

6. BIBLIOGRAFÍA.

AGENCIA EUROPEA DE MEDIO AMBIENTE (2008): *Conjunto Básico de Indicadores de la AEMA*. Agencia Europea de Medio Ambiente, Bruselas.

ÁVILA, R., INIESTA, A., HERRERO, D., de JUAN, J., AGUIRRE, G., GUEREÑA, A., MORERA, C., BELUCHE, G., AGUILAR, P., ZAMBRANO, D., RUIZ, R., BUGLASS, L., KAMP, C. y GIRALDO, A. (2002): *Turismo Sostenible*. Ed. EIPALA, Colección problemas internacionales, Madrid.

CASTRO, J.M. (2004): *Indicadores de Desarrollo Sostenible Urbano: Una Aplicación para Andalucía.* Consejería de Economía y Hacienda. Junta de Andalucía, Sevilla.

CLARKE, J. (1997): "A Framework of Approach to Sustainable Tourism". *Journal of Sustainable Tourism*, 5, 3, pp. 224233.

CONSEJERÍA DE MEDIO AMBIENTE (2001): *Bases para un Sistema de Indicadores de Medio Ambiente Urbano en Andalucía: Experiencias Internacionales en la Medición de la Sostenibilidad en las Ciudades*. Junta de Andalucía y Fondo Europeo de Desarrollo Regional, Sevilla.

CRONBACH, L.J. (1951): "Coefficient Alpha and the Internal Structure of Tests". *Psychometrika*, 16, pp. 297334.

DACHARY, A.C. y ARNÁIZ, S.M. (2002): *Globalización, Turismo y Sustentabilidad*. Ed. Universidad de Guadalajara, México.

DIPUTACIÓN PROVINCIAL DE BARCELONA (2000): *Sistema Municipal de Indicadores de Sostenibilidad*. Diputación Provincial de Barcelona, Barcelona.

DIPUTACIÓN PROVINCIAL DE JAÉN (2002): *Indicadores para la Sostenibilidad*. Diputación Provincial de Jaén, Jaén.

EUROSTAT (2006): *Methodological Work on Measuring the Sustainable Development of Tourism. Part 2: Manual on Sustainable Development Indicators of Tourism*. Working paper. Office for Official Publications of the European Communities, Luxembourg.

FARSARI, Y. y PRASTACOS, P. (2002): "Sustainable tourism indicators. Casestudy for the municipality of Hersonissos". Proceedings of the International Scientific Conference on "Tourism on islands and specific destinations", University of the Aegean, Chios.

FRAUSTO, O., ROJAS, J. y SANTOS, X. (2006): "Indicadores de Desarrollo Sostenible a Nivel Regional y Local: Análisis de Galicia, España, y Cozumel, México". *Estudios Multidisciplinarios en Turismo*. *SECTUR*, 1, pp. 175201.

FULLANA, P. y AYUSO, S. (2002): *Turismo Sostenible*. Rubes Ed., Barcelona.

FUNDACIÓN EOI DE ANDALUCÍA Y OBSERVATORIO AMBIENTAL DE ANDALUCÍA (2008): *Informe de Sostenibilidad en Andalucía 2008*. Colección EOI Medio Ambiente, Fundación EOI de Andalucía, Sevilla.

GALLEGO, I. y MONICHE, A. (2005): *Sistema de Indicadores Territoriales para un Destino Turístico*. Conferencia de la OMT Iguazú, Argentina/Brasil/Paraguay, 36 de octubre de 2005: La Cuenta Satélite de Turismo (CST): Comprender el Turismo y Diseñar Estrategias.

HARDY, A., BEETON, R.J.S. y PEARSON, L. (2002): "Sustainable Tourism: An Overview of the Concept and its Position in Relation to Conceptualisations of Tourism". *Journal of Sustainable Tourism*, 10, 6, pp. 475496.

HERNÁNDEZ, A. (2003): *Informe sobre los Indicadores Locales de Sostenibilidad Utilizados por los Municipios Españoles Firmantes de la Carta Aalborg*. Edición electrónica a texto completo en http://habitat.aq.upm.es/indloc/aindloc.html.

HUNTER, C. (2002): "Aspects of the Sustainable Tourism Debate from a Natural Resources Perspective". En Harris, B., Griffin, T. y Williams, P. (Eds.): *Sustainable Tourism. A Global Perspective*. Elsevier ButterworthHeinemann, Burlington, pp. 323.

LIU, Z. (2003): "Sustainable Tourism Development: A Critique". *Journal of Sustainable Tourism*, 11, 6, pp. 459475.

MINISTERIO DE FOMENTO (2004): *Informe sobre los Indicadores Locales de Sostenibilidad*. Dirección General de la Vivienda, la Agricultura y el Urbanismo, Ministerio de Fomento, Madrid.

MINISTERIO DE MEDIO AMBIENTE (2008): *Banco Público de Indicadores Ambientales*. Ministerio de Medio Ambiente.

NARDO, M., SAISANA, M., SALTELLI, A., TARANTOLA, S., HOFFMAN, A. y GIOVANNINI, E. (2005a): *Handbook on Constructing Composite Indicators: Methodology and User Guide*. OECD Statistics Working Papers.

NARDO, M., SAISANA, M., SALTELLI, A. y TARANTOLA, S. (2005b): *Tools for Composite Indicators Building*. Institute for the Protection and Security of the Citizen, European Commission.

ORGANIZATION FOR ECONOMIC COOPERATION AND DEVELOPMENT (2000): *Frameworks to Measure Sustainable Development*. Ed. Organization for Economic Cooperation and Development Publications Service, Paris.

OSE (Observatorio de Sostenibilidad de España) (2008): *Indicadores del Observatorio de Sostenibilidad de España*. Ministerio de Medio Ambiente.

ROMAGOSA, F. y CUÉTARA, L. (2001): "El Desarrollo Sostenible en Destinos Turísticos: Propuesta de un Sistema de Indicadores de Sostenibilidad". *Papers de Turisme*, 30, pp. 124139.

SANCHO, A. y GARCÍA, G. (2006): "¿Qué Indica un Indicador? Análisis Comparativo en los Destinos Turísticos". *Revista de Análisis Turístico*, 2, pp. 6985.

SANCHO, A., GARCÍA, G. y ROZO, E. (2007): "Comparativa de Indicadores de Sostenibilidad para Destinos Desarrollados, en Desarrollo y con Poblaciones Vulnerables". *Annals of Tourism Research*, 9, pp. 150177.

SHARPLEY, R. (2000): "Tourism and Sustainable Development: Exploring the Theoretical Divide". *Journal of Sustainable Tourism*, 8, 1, pp. 119.

VYAS, S. y KUMARANAYAKE, L. (2006): "Constructing SocioEconomic Status Indices: How to Use Principal Components Analysis". *Health Policy and Planning*, 21, pp. 459468.

WORLD TOURISM ORGANIZATION (1993): *Tourism: The Year 2000 and Beyond Qualitative Aspects*. Discussion paper, World Tourism Organization.

WORLD TOURISM ORGANIZATION (1996): *What Tourism Managers Need to Know: A Practical Guide to the Development and Use of Indicators of Sustainable Tourism*. Ed. World Tourism Organization, Madrid.

WORLD TOURISM ORGANIZATION (2004): *Indicators of Sustainable Development for Tourism Destinations. A Guidebook*. Ed. World Tourism Organization, Madrid.

THE ROLE OF FIRMS AND ENTREPRENEURSHIP ON LOCAL DEVELOPMENT IN THE REGION OF *VALE DO SOUSA*

NELSON JORGE RIBEIRO DUARTE
E-mail: duarte.nelson@gmail.com
ESTGF/IPP/CETRAD
FRANCISCO JOSÉ LOPES DE SOUSA DINIZ
E-mail: fdiniz@utad.pt
DESG/UTAD/CETRAD

Abstract

On the present paper it is considered and analyzed the role of management and entrepreneurship on firm and region sustainability. In order to analyze this relation it was considered the three factors associated to sustainable development (economic, social and environmental) and managers' behaviour on each of these factors.
Through a questionnaire presented to 251 micro, small and medium firms on this region, the results showed that on the contrary of what most literature presents, managers are putting in first place the environmental factor, followed by the social and the economic as the last one. These results lead to a characterization of firms as being sustainable on what regards sustainable development factors, but with a risk-averse and non innovative behaviour on firms' management.

Key words: small firms, entrepreneurship, sustainability.
Thematic Area: Urban, Regional and Local Economics.

1. INTRODUCTION - THEORETICAL OVERVIEW

The role played by firms nowadays on the economy is undoubtedly of great importance. The present situation that the world economy is facing is also showing the importance of firms, since governments are putting some efforts in order to help them due to their economic and social role. Among firms there is a category that deserves special attention: the micro and small firms. Together they represent a large percentage of firms in every country.

Table 1. UE-25 Firm indicators by class, except financial sector (%)

	Firms	Employment	Turnover	Added Value
Micro (0-9)	91,5	29,8	19,4	20,5
Small (10-49)	7,3	20,8	19,3	19,1
Medium (50-249)	1,1	16,5	19,2	17,8
Large (250+)	0,2	32,9	41,9	42,7

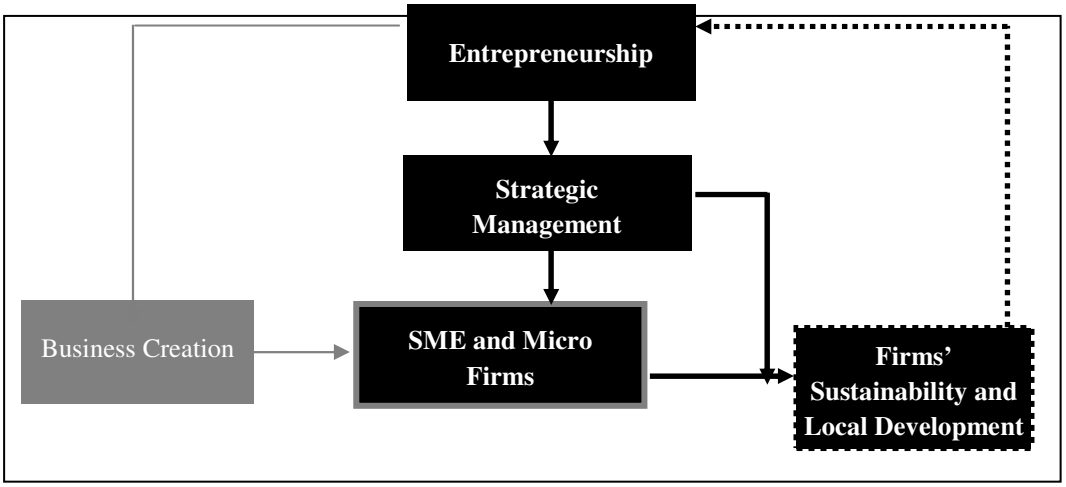
Source: Schmiemann (2006) pp.2 & EUROSTAT (2006)

According to a IAPMEI (*Instituto de Apoio às Pequenas e Médias Empresas e à Inovação – SMEs and Innovation Support Institute*) study, [IAPMEI (2007)], micro and SMEs represent 99,6% of total firms in Portugal, 78,2% of total employment and 55,6% of turnover on the Portuguese economy. These figures show the importance of small firms in Portugal.

The main idea on this paper is to study and analyze the role of firms on local sustainability and the existent relation among the concepts presented on Figure 1. Entrepreneurship is undoubtedly related with firm creation but it should be present on firms' management at all the time, especially on strategic management. Todays large firms where, sometime in the past, small ones. According to Magretta (2004) they became large by being the best small ones. The question is: What did they do in order to become the best ones? There is not only an answer to this question; however one of the factors that certainly influenced it was their strategy. In order to reach a goal is necessary to have a plan. Sailing according the wind direction won't take us anywhere, but with a goal, the wind might be a powerful tool to reach the proposed goal. By being the best ones, or at least by achieving their goals, these firms, in particular small firms, will be themselves sustainable, and they might also work as an engine for local development.

In order to close this cycle and following the idea presented by Schumpeter [Schumpeter (1934)] of creative destruction, it is expected a contribution from these firms to economic growth and development. Nowadays, most governments, especially those from developed countries are concerned not only with economic development, but also with social and environmental development, or in other words, with sustainable development. And since this is a cycle, this sustainable development will contribute for new and better firms fostering local entrepreneurship with new ideas and more efficient.

Figure 1. The relation among some important firm concepts



In order to analyze these relations a study was taken on a Portuguese region in the north of the country known as *Vale do Sousa*. This region is formed by a group of six municipalities with a large number of firms. Since the firms in this region, as all over the world, belong to different activities sectors it was analyzed the distribution among sectors, and we realized that firms on construction and manufacturing business represents nearly 50% of total firms in this region.

Since it would be difficult to study at the same time the strategies of manufacturing, services, or retailing businesses, the study was taken on manufacturing and construction businesses. It was then studied on these sectors the attitude of managers/owners towards some sustainable development concepts. After this analyze some ideas about the relation between local sustainability and firms' entrepreneurship will be present. The last section presents some final comments and considerations, and some suggestions for further research.

2. ENTREPRENEURSHIP AND STRATEGIC MANAGEMENT

The concept of entrepreneurship is nowadays frequently used, on academic, scientific, managerial and political fields. At the same time is a key concept on development discussions because is through entrepreneurial actions that is possible to create added value and to promote better economic and social conditions that will benefit both the individuals involved and the community as a whole. In order to foster local development many governments are trying to promote and support innovative actions from the existent firms and new firms creation.

The concept of entrepreneurship is related to firm creation, creativity, innovation, organization, cooperation, and among others, to firm management. In this last case the most frequent used concept is Intrapreneurship, a term that was first suggested by Pinchot (1985). Later, Carree, Stel, Thurik, & Wennekers (2000) defined entrepreneurship as firms' creation and intrapreneurship as new ideas and responsibilities implemented on the existent organizations. Both of them are assumed as essential for the economic activity.

The concept of intrapreneurship also plays an important role on the establishment of networks [Greeve (1995); Minguzzi & Passaro, (2000)] and it is close to the concept of strategic management innovation [Bruyat & Julien (2000); Antoncic & Hisrich (2001)]. According to Jong & Marsili (2006) these relations are valid not only for large firms, but also for small ones.

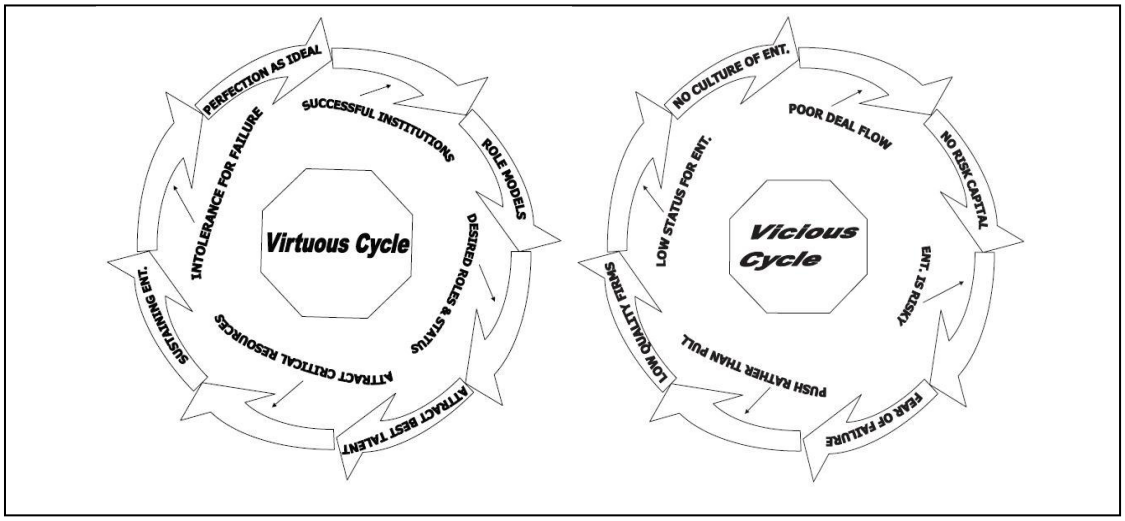
Based on the ideas presented by these and some other authors it is possible to say that there is a close relation between entrepreneurship and firm management, in other words, it can be said that innovative strategic management is very similar to the concept of intrapreneurship, and it should be present in all firms regardless its size. The concepts of entrepreneurship and strategic management are so close, that as Venkataraman & Sarasvathy (2001) refers to them, to study one of them disconnected from the other is like a representation of *Romeo and Juliette* with just one of them.

Bearing in mind the idea that entrepreneurship and strategic management are concepts that must be present, side by side, in firms' management, it is also important to notice the relations between firms (creation and management), entrepreneurship and local development (or the territory). As Melo (2002) presents, the capacity for a territory to become competitive relays on the dynamic behaviour of the firms based there. The lack of this ability or the lack of the entrepreneurial spirit is the background for most of problematic situations that some territories seem not to be able to overcome. Also Mezias & Kuperman (2000) and Lee & Peterson (2000) defend that entrepreneurship does not appear as an isolated action, but it comes from a community creating like that a social system that leads firms and entrepreneurs to success.

The existent relation between the region and the firms/entrepreneurs lead to another important concept that is the embeddedness effect. This effect can be related both with the entrepreneurs [Jack & Anderson (2002)] and with the organizations [Dacin, Ventresca, & Beal (1999)]. But as OECD (2000) this relation is composed by many factors, such as social, cultural or political, and if in one hand some conditions may contribute for an increase of entrepreneurship actions, from another hand they can also be a restriction for those actions. All the participants on this "network" must be working and aiming the same goals.

As Venkataraman (2004) suggests entrepreneurship in a region may occurs in two different ways, or two different cycles as presented on Figure 2.

Figure 2. Entrepreneurship Cycles



Source: Venkataraman (2004)

"A region exists in a state of ''virtuous equilibrium'' when it has been conducting economic and cultural activities for long periods and has settled into a predictable and comfortable position. Such a state exists when patterns of activity have formed and evolved through historical and local contingencies and through ceaseless competition. In these cases, competition is not limited to the realm of economics: social and political competition also contributes to the establishment of these patterns. People in power drive out other people in power; while people who are well-connected drive out those people who are not so well-connected. Indeed, in these regions, equilibrium is defined as much by product–market competition as it is by social and political competition" Venkataraman (2004).

On the other hand *"…when developments based in old industries no longer valued occur, no role models exist in the imitating regions that lack the cultural mechanisms for adaptation. The role models are elsewhere, in other regions where these new models originated and are taking shape. When this exogenously driven change occurs, the previously workable virtuous cycle turns out to be a vicious cycle under which change cannot easily occur. The culture of Schumpeterian entrepreneurship, which involves trying ''new'' things, making ''bold bets'' is nonexistent. Rather, there is only a very narrow, sustaining, and weak approach to entrepreneurship. The typical reaction of economic and political leaders in such environments is one of diffidence and inertia. They might know that they have to choose a new economic model, but are uncertain as to what the next model should be or how to create it. This is when a region becomes trapped in a vicious cycle: great ideas and bold bets cannot and does not emerge"* Venkataraman (2004).

The differences presented on those cycles are a portrait of what happens in some regions and may explain why while some are in a constant development some others cannot leave the undeveloped status that are bearing for a long time. The main idea to present from this section is that entrepreneurship, either on firm creation or management, cannot be disconnected from strategic management in order to plan at the short and long time, but the firms by themselves even following all the theoretical rules, and keeping innovation and management together, need to be embedded in the region. Only in that way is possible to exploit the synergies available from an interesting network that will bring better results for all participants.

3. SOME THEORETICAL CONCEPTS AND ITS RELATION WITH DEVELOPMENT

In order to fight poverty and the existent disequilibrium among regions small and micro firms are frequently presented as a useful tool. However according to Cheshire & Malecki (2004) the way regional, and consequently local, growth occurs is a process not sufficiently understood. So in order to best understand this process, is necessary to understand the concepts, at least some of them, related to local development.

Starting by the concept of economic growth, Romer (2001) presents the idea that it occurs when someone uses some resources turning them into something valuable. At the same time economic growth is presented by Islam, Munasinghe, & Clarke (2003) as a goal in many countries as a mean to improve population living standards and to eradicate poverty, but they also ask how can it be measured. Some may suggest the use of macroeconomic indicators such as GDP, some others like Apolinário (2005) describe it as a result of the human and technological capital, and the organization that manages both. The integration of these factors leads us to the concept of economic development.

The relations between growth and development are very close to each other. For instance, as Silva & Silva (2002) present it, the models of economic growth use as variables macroeconomic aggregates, trying to find quantitative relations between inputs and outputs possible to be explained by mathematical expressions or models. Some other models such as *The Agropolitan* or *The Local Initiative Model*, as presented by Weaver (1988) both related to territory or local development present SMEs and entrepreneurship as an important tool for development through job creation and natural resources management. Other tools presented at both models, are related to economic, social and environmental aspects, which allow us to establish a relation with the concept of sustainable development.

Sustainable development as presented above includes three dimensions, economic, social and environmental, and an interesting definition can be presented as: "the satisfaction of present need without compromising future needs". This means that economic growth is necessary, but social and environmental aspects must be also considered. Islam, Munasinghe, & Clarke (2003) present the idea that not only development must be sustainable but growth also needs sustainability, which means that not only economic indicators must be considered to analyze economic growth, but also social and environmental aspects must be taken into account.

Arvanitidis, Petrakos, & Pavleas (2007) identified some economic growth determinants, such as: (1) High quality of human capital; (2) High Technology, innovation and R&D; (3) Stable political environment; (4) High degree of openness (networks, links), among others. Most of them can be easily connected with the entrepreneurial fabric, which supports the idea that economic growth, whether sustainable or not, is promoted by firms either at a micro or macroeconomic level.

Economic growth is also related with the concept of innovation, as Mccann (2006) presents, innovation is the spark for regional economic growth that occurs in locations where firms (normally small) are based with a good labor force and specialized services. At the same time, as Vargas (2000) argues, micro firms have been seen for some time as an alternative to macroeconomic policies in order to fight poverty.

However the relation between economic growth or development and firms also presents some negative aspects. For instance the increase in the number of firms may create rivalry instead of cooperation [Visser (1999); Boari, Odorici, & Zamarian (2003); Narula, R. (2004)]. And even on the concept of sustainable development is not understood on the same way by everyone. Giddings, Hopwood, & O'Brien (2002) argue that even with governments and some business sectors being concerned about sustainability issues, the separation of the concept in three different dimensions may be used to justify the focus in one of these dimensions, usually the economic, as justified by some other authors such as Korton (1996) or Monbiot (2000).

Nevertheless it is consensual that macro-economic growth *per se*, is not a solution to increase population life standards, neither to promote sustainability. Nowadays, firms are presented as a more efficient solution, in particular, micro and small, since they not require large initial investments but they can support economic growth, creating jobs and supporting

community, and they can also play an important role on environmental levels, because of the costs or benefits that may result from their actions towards to environment.

4. FIRMS AND LOCAL DEVELOPMENT

In order to achieve economic development it is consensual that firms are an efficient, or at least, a widely acceptable solution. Vargas (2000) defends that some United Nations responsible have been supporting the creation of micro firms as a mean to achieve economic and sustainable development. According to Craig, Jackson, & Thomson (2005) there is a general idea about small firms, that they are a kind of growth incubator, being the place where innovation occurs and new ideas become viable businesses.

In order to justify the existent gap among countries and regions on what concerns economic growth the Organization for Economic Co-operation and Development (OECD), [OECD (2003)] identifies factors, institutions and policies that may increase growth on the long-term. On that study it can be identified as basic factors for growth: (1) Physical capital accumulation; (2) Human capital accumulation; (3) Research & Development. It is interesting to notice, that all of them are related to firms.

Even though firms are considered an important player on the development game, they are not the only players. Local development may be supported by firms, but is also supported by local agents. It is common to find entrepreneurs that started a business in their region and this relation lead us back to the concept of embeddedness, which as we have seen on the previous section, can be defined as intangible resources that result from close relations and from the knowledge that these entrepreneurs have with and from the region. These local entrepreneurs having economic goals, most likely are available to put some efforts, in order to also achieve social and environmental goals. Like that it can be said that these agents are closest to the goals of sustainability, both for the firm and the region.

The importance given to small firms is not recent, Pecqueur (1989) suggested a new framework to analyze development putting the firms (SMEs) as a central figure, a kind of pivot element on what concerns development agents or institutions. The relations established among all these participants will contribute for the establishment of networks at various levels that can contribute for firms development and sustainability as well as for local development, as argued for many authors [Ahern (1993); Narula (2004); Arend (2006); Acquaah (2007)]. More recently, some authors suggest alternative solutions for management models aiming sustainability [Kerr (2006); Rocha, Searcy, & Karapetrovic (2007); Espinosa, Harnden, & Walker (2008)], and some others are trying to introduce the concept of sustainable entrepreneurship [Young & Tilley (2006)].

5. FIRMS AND LOCAL DEVELOPMENT IN THE REGION OF VALE DO SOUSA

In order to analyze some of this concepts through a practical approach, it was studied the region of *Vale do Sousa*, considering the construction and industry sectors. The main results presented on this section result from the presentation of a questionnaire to 251 micro, small and medium-sized firms, on the region of *Vale do Sousa*. This is a sub region of NUTE III – Tâmega located on the North of Portugal.

The region where this study took place is a region composed by 6 municipalities (*Castelo de Paiva, Felgueiras, Lousada, Paços de Ferreira, Paredes, Penafiel*) that compose the *Vale do Sousa Urban Community*. This region is located in the North of Portugal, and for statistical purposes is a region within NUTE III – *Tâmega*. According to INE (2007) in this region there are 337.380 inhabitants with a relatively high percentage of young people.

Economically as most of the country, the primary sector was in the past the main activity. Other activities such as manufacturing or services have been assuming a more relevant role. Nowadays the main activities in this region are: shoes making, textile, wood furniture and construction. In four of these municipalities it is even possible to identify some industrial clusters as referred by Bessa (2004) and DHVMC (2004): *Felgueiras*: Shoes production;

Lousada: Textile; *Paços de Ferreira* and *Paredes*: Wood furniture.

To describe the entrepreneurial fabric of this region, it was necessary to collect information from different institutions, since the available information is not the same from every source. By using data from INE in 2005 were registered 34.049 firms, from all sectors. However information from CofaceMOPE reveals 11.973 firms, and according to the Work Ministry there are 10.231 firms. After analysing these differences, and some conversations with local authorities, it was realized that there is no valid information about the exact number of firms, and it was assumed, that a value of 12.000 firms should be very close to the reality.

After a decision about the number of firms to consider for the present study was done an analysis of firm distribution according to the activity sectors. This distribution, considering the data from the three institutions is more or less similar pointing as main activities the retailing, manufacturing, and construction sectors. Together these sectors represent around 75% of total firms in this region. However to analyse management strategies, entrepreneurial/innovative actions, and behaviour towards to sustainability from these type of firms is difficult to do using a single approach to all of them. In order to find more significant results it was decided to limit this study to industrial and construction businesses. In order to consider the industrial sector as a whole, the study analysed both the manufacturing and mining and quarrying firms.

This choice was done, since structurally there is not such a big difference among these three sectors. By their nature they are much closer to each other than anyone of them to the retailing sector. By that reason, and since these two sectors (industrial and construction) represent around 50% of total firms, the study was taken on these sectors. According to the data provided from the three institutions it was verified that the number of firms in industrial and construction sectors is around 5.000, and since this value is close to the exact number of firms, this will be used as the total population in study on this work.

On what regards firms' dimension, according to the data provided by CofaceMOPE, it is possible to say that this region does not present the usual distribution, where micro firms present an overwhelming percentage. In this region they are still the largest class of firms with 62% (in Portugal this figure is around 80%) and small firms represent 35%. Together they account for 97% of total firms which is within the class distribution found in Portugal. The remaining 3% are classified as medium-sized firms. This distribution is a typical one for a region that presents a relevant manufacturing sector, which, by its nature, presents a large amount of firms counting with 10 or more employees. This value leads to a change on firm class reducing like that the percentage of micro firms.

The questionnaire was elaborated in order to collect information not only for the subjects here discussed, but for a wider research on strategic entrepreneurship and sustainability. Since there was no viability to question the total population (5.000 firms) the study was taken by using a valid sample. In order to calculate the sample size Saunders, Lewis, & Thornhill (2003) present a formula that considers the variability of the factors to be studied, the confidence interval required, and the error margin. The formula is:

$n=p\%*q\%*[z/e\%]^2$

where:

n: minimum sample size required;

p%: proportion belonging to the specified category;

q%: proportion not belonging to the specified category;

z: z value corresponding to the level of confidence required:

e: margin of error required;

According to Saunders, since the population is less than 10.000 a smaller sample can be used without affecting the accuracy.

The adjusted formula is:

$n'=\{n/[1+(n/N)]\}$

Where:

n': adjusted minimum sample size;

n: the minimum sample size (as calculated above);

N: total population;

Considering as the main factor the strategic entrepreneurship, and considering a variability of 80%-20% (which was corroborated later with the results) it was obtained a n = 245,86 and a n' = 235,47. With this minimum required sample it was verified that to study the total population through a reduced number of questionnaires should be done 236 questionnaires.

After this result the questionnaires were distributed according the distribution of firms by sectors, classes, and municipalities. The questionnaire presented to firms, as referred already included a large number of questions in order to evaluate different aspects of firms' management. The questions considered for this paper were mainly related with sustainability and sustainable development aspects.

The questions presented on the questionnaire aiming to get some information about the firms behaviour to local development and sustainability, where written on a simple and clear language in order to get a spontaneous answer from the interviewee. Even though the question presented where not referring the scientific terms, they where elaborated in order to be easily understood and all of them based in some scientific concept derived from a previous literature review. The answers where asked based on a Likert scale (1 to 5).

In order to analyze the aspects related to local development it was elaborated a section composed by nine questions related with the three main areas of sustainable development. The first question about development was a general approach to the role of firms on local development. It was asked the interviewees to classify on a Likert Scale the sentence: *"The firms play an important role on local development"*. On a possible classification from 1 to 5 the average answers reached a value of 4,38, which means that the interviewees, most of them with managerial responsibilities in the firms, have an evident perception that firms are important agents on local development. It is interesting to notice that from all the answers the minimum value obtained was 2, which means that no one is in totally disagreement with the importance of firms on local development.

The following questions where analysed in groups of variables, following the methodology proposed by Hill & Hill (2002): latent variables. These variables are not directly observed but are rather inferred from other variables that are observed and directly measured. The values of the latent variables can be inferred from measurements of the observable variables. To analyze both the economic and social dimension it was used two groups of three questions, in order to analyze the environmental dimension a group of two questions.

The economic dimension was analyzed through the questions:

- Small firms are those who may have the largest contribution for local development;
- Firm profits must be reinvested on the firm;
- Firm profits must be reinvested on the region;

While the answers to the first question *"The firms play an important role on local development"* showed a significant agreement from the interviewees with a result of 4,38, the answers to the question focusing the small firms did not present such a significant degree of acceptance since the average value is 3,63. Analyzing both results it is possible to say that managers/owners agree that firms play an important role for local development, but they change their opinions when it comes to the small firms. It means that they attribute to their larger incumbents the responsibility to act on local development.

The next questions where related to profit reinvestments either on firm or in the region. The results obtained are similar, however there is an "egocentric" vision about reinvestment, since the average for firm reinvestment is 3,98 while the average for region reinvestment is 3,31. These results also show that firm factors present higher values than those related on the role of small firms in development. This might mean that managers are more concerned with the economic dimension.

By doing an analyzes of all the questions it is possible to have an idea about the behaviour that firms present towards to the economic dimension as a whole and the average result is of 3,64. Since this is the first result, some final and more valuable comments will be done after analyzing the other two dimensions. Like that it will be possible to compare the results from the three dimensions.

The social dimension was analyzed through the questions:

- Firm workers must be recruited on the firm municipality or close to the municipality borders;
- All the firm works, even those subcontracted, are done by adults;
- Firm must support society through sponsorship actions and developing social and cultural actions on a regularly basis;

The questions presented on the social dimension are related with employment, child labor and community support. From these question one that deserves special attention is the one related to child labor since the average result is 4,29. This value shows a social concern (that is mandatory by law) from the employees in order to have the work done by adults. Considering the restrictions existents, it was not possible to develop some other ways to verify if this is effectively happening, so the analyzes will be based on the obtained results. However, there still exists a significant level of young people leaving the school to get into the labor market. This may suggest some under covered child work that may be studied in further researches.

Considering the remaining questions, local employment presents an average of 3,74 which means that at social levels, firms are local employers contributing like this for an increase on social life conditions. The support to community through sponsorship and other support actions still presents positive values, but is less significant than the others factors, since the average result is 3,27. This allow us to conclude, that firms are concerned with child labor, and try to promote local employment, however less attention is given to the community support.

Analyzing the questions as one variable, we will get the results about the social dimension that gives a result of 3,76. This result is higher than the obtained at the economic dimension one. According to the literature on sustainable development this is not a normal result, since most of the authors defend that usually firms and even governments sometimes are putting more efforts on the economic dimension.

The environmental dimension was analyzed through the questions:

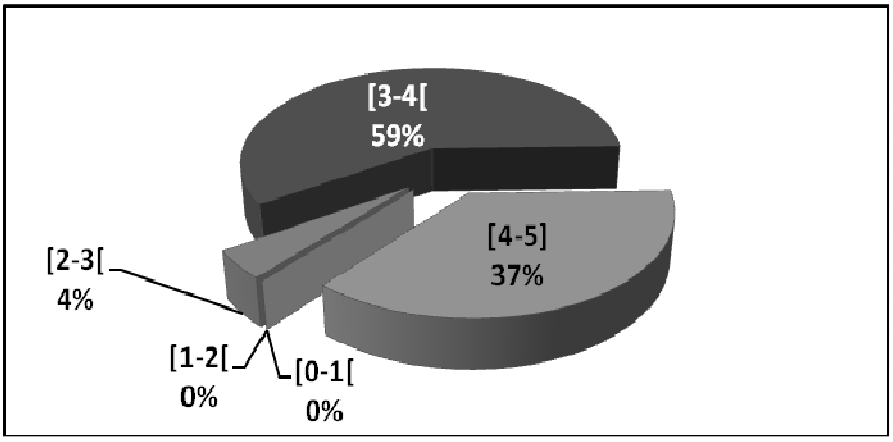
- Firm knows and tries to reduce the impacts cause on the environment (soil, air and water);
- If some environmental rules will not be respected by this firm, there is no impact on the environment;

In order to evaluate all these questions in the same scale, the last question was recoded to be analyzed together with the first from this group. For the question related to the environmental impacts the average result was 4,1, while the results for the last question presented an average of 3,9. This may lead us to the discussion of the Tragedy of the Commons presented by Garret Hardin in 1968. The values obtained on these questions may be inflated because of the social concern by answering the questionnaire. Even with a guarantee that the questionnaires are just for academic research some of the interviewees may have tried to give a social acceptable answer. This idea is presented because if a firm is concerned about their impacts on the environment as the results from the first question show, it was expected to get a higher average on what regards to environment rules (second question).

Considering the result from both questions the average result is 4,03 that is the highest value from the three dimensions. This demonstrates the importance that firm managers and owners are giving to the green dimension. The question that arises is: “Are the actions of these firms really in agreement with these results?” This question also opens new ways for further research.

Trying to get some results about the concept of sustainable development it will be assumed that the three dimensions have a similar importance on the contribution to sustainable development. Considering that all dimensions assume an equal importance the average result presents a value of 3,81 which means that on average the firms in the region of *Vale do Sousa*, present a positive, let’s say, proactive attitude towards sustainable development. Considering each firms results and forming classes where the lower means a weak approach to sustainable development and the highest a strong approach we can see that most of firms, 96%, present a proactive attitude to sustainable development (Figure 3).

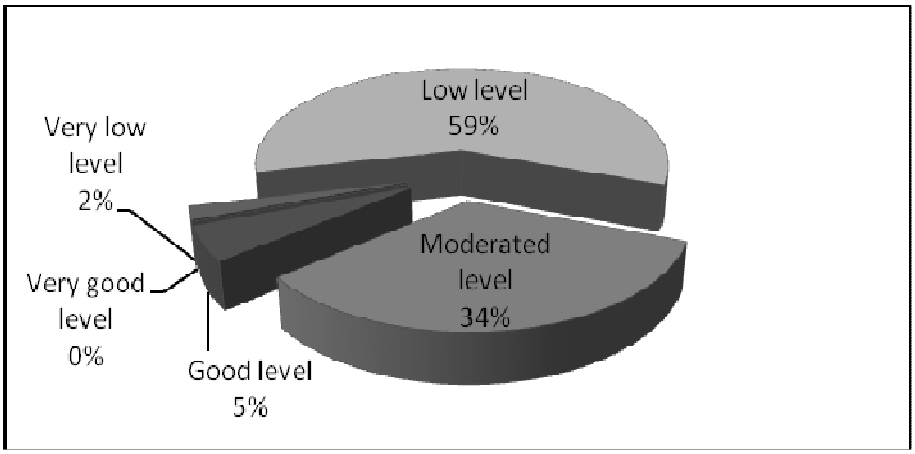
Figure 3. Percentage of Firms by Sustainability Classes



The two lowest classes do not include any firm, which means that all firms present at least a moderate approach to sustainable development. These results allow us to conclude that firms on the region of *Vale do Sousa*¸ promote a development that can be classified as sustainable. Disagreeing with the idea presented by the literature review, the economic dimension is the one that deserves less attention from the managers/owners.

On the same questionnaire where also presented some questions in order to evaluate the entrepreneurship degree on the firms and it was expected to find a relation between this degree and the sustainable development approach. However it was not possible to find any relation. As we can see on Figure 4 most firms present a low level of entrepreneurship, and even with this results the same firms presented a proactive approach to sustainable development.

Figure 4. Percentage of firms according to the level of entrepreneurship



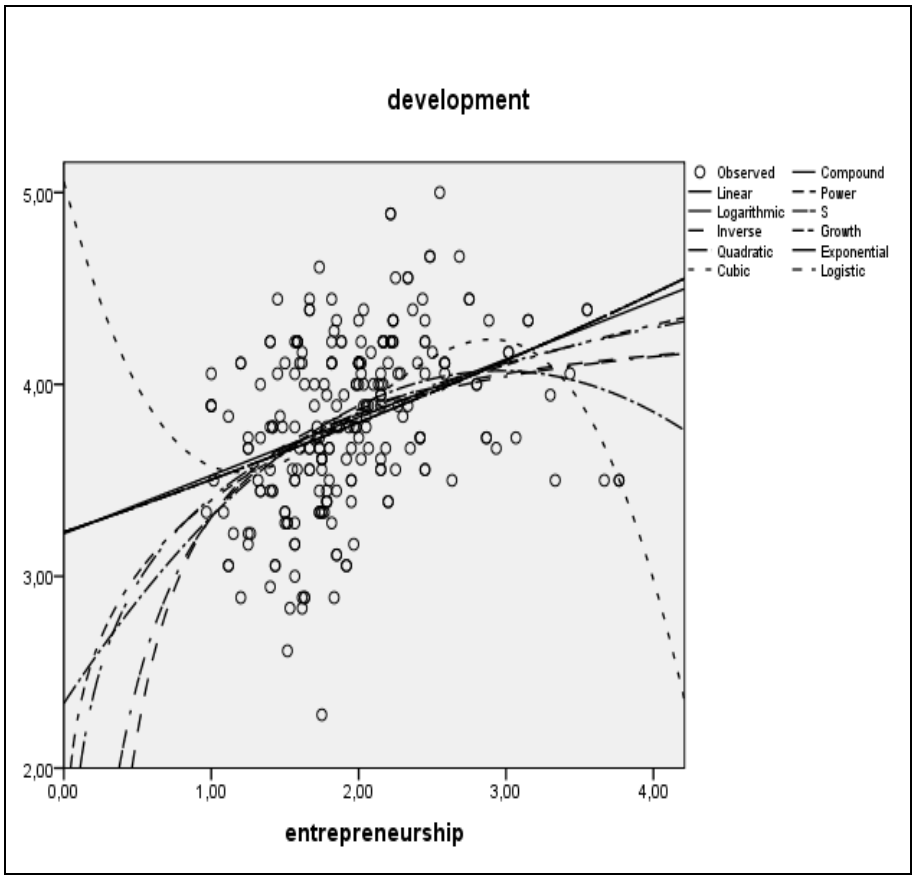
Considering an average degree of 1,94 to the entrepreneurship level and an average degree of 3,81 for local sustainability by using some econometric tools through a statistical program (SPSS) it was analyzed the relation between them. In first place it was done a bivariate scatter plot with both variables (entrepreneurship and development). The graph presented a dispersion that did not indicate any pattern or typical behaviour. After that, and since we were looking for a pattern it was tested some regressions on these variables, but the results pointed on the same way that the previous analysis. Considering the following regressions, Linear, Logarithmic, Inverse, Quadratic, Cubic, Compound, Power, S, Growth, Exponential and Logistic, it was verified that the highest value obtained for R^2 was 20,9% with the Cubic Regression, however this is an extremely low value. If this result were considered as acceptable we would get the following regression expression:

$y= -2{,}934x^3+1{,}732x^2-0{,}282x+5{,}057$

where y is development and x is entrepreneurship

Even being one of the most flexible regression models, with such a low R^2 it is worthless to analyze the regression parameters, since the conclusions would not be valid. As it is possible to see on Figure 5 none of the regression lines fit on the scatter plot that considers entrepreneurship and development.

Figure 5. Regression lines on the scatter plot (entrepreneurship – development)



The conclusion that can be drawn out from these results is that firms in the region of *Vale do Sousa*, do not present an entrepreneurial management but they still present a proactive approach to sustainable development. In this case is not possible at all to talk about the concept of sustainable entrepreneurship.

6. CONCLUSION

Considering the results from the three dimensions analyzed in order to measure sustainable development: Economic: 3,64; Social: 3,76; Environmental: 4,03, some general comments may be presented. According to the literature review some authors argue that the division of sustainable development concept into three dimensions is an excuse in order to focus the attention on the economic dimension leaving the two others for second plan. In the present study the results pointed in the opposite way, apparently firms in the region of *Vale do Sousa* are more concerned with the environmental and social dimensions than with the economic.

This may be what is really happening in this region, however, and since this shows an opposite tendency to the literature review, it requires some further research both on these sectors and in some other sectors that exist in this region. In order to analyze each dimension there are some other variables that can be taken into account that were not considered on the present study. A further research suggestion is to build a battery of indicators for each dimension according to the literature review in order to analyze those variables on the region of *Vale do*

Sousa, to accept or reject the results here presented. Among many other factors, the study of firms' stakeholders must be taken into consideration in order to get a different perspective from the firm.

Another important conclusion is that even though the manager/owners that were interviewed recognize the importance of firms' contribution on local development, they see this contribution as a responsibility of large firms. When the question is presented about firms in general they agree, but the agreement is not so strong when the question refers specifically the small firms.

On what regards entrepreneurship and sustainability it was possible to conclude, that in this case there is not a relation, which means that in this region is difficult to talk about sustainable entrepreneurship. On the contrary to what could be expected, the entrepreneurship degree (strategic entrepreneurship) is very low, while local sustainability present very good results. In this last situation, it will be important to develop further research in order to evaluate the sustainability of this region.

REFERENCES

ACQUAAH, M. (2007). Managerial social capital, strategic orientation, and organizational performance in an emerging economy. *Strategic Management Journal* , 1235-1255.

AHERN. (1993). Implication of strategic alliances for small R&D intensive firms. *Environment and Planning*, 1511-1526.

ANTONCIC, B., & HISRICH, R. (2001). Intrapreneurship: Construct refinement and cross cultural validation. *Journal of Business Venturing* , 495-527.

APOLINÁRIO, M. J. (2005). Desenvolvimento: O factor imaterial. *Revista Dirigir* , 44-48.

AREND, R. J. (2006). SME-Supplier alliance activity in manufacturing: Contingent, benefits and perceptions. *Strategic Management Journal* , 741-763.

ARVANITIDIS, P., PETRAKOS, G., & PAVLEAS, S. (2007). Determinants od economic growth: the experts' view. *Congress of European Regional Science Association.* Paris.

BESSA, D. (2004). *PRASD - Programa de Recuperação de Áreas e Sectores Deprimidos.* Lisboa: Ministério da Economia.

BOARI, C., ODORICI, V., & ZAMARIAN, M. (2003). Clusters and rivalry: Does localization really matter? *Scandinavian Journal of Management* , 467-489.

BRUYAT, C., & JULIEN, P. (2000). Defining the field of research in entrepreneurship. *Journal of Business Venturing* , 165-180.

CARREE, M., STEL, A., THURIK, R., & WENNEKERS, S. (2000). Economic development and business ownership: An analysis using data of 23 OECD countries in the period 1976-1996.

CHESHIRE, P. C., & MALECKI, E. J. (2004). Growth, development and innovation: A look backward and forward. *Papers in Regional Science* , 249-267.

CRAIG, B., JACKSON, W., & THOMSON, J. (2005). On SBA-guaranted lending and Economic Growth. In S. Shane, *Economic Development Through Entrepreneurship* (pp. 127-150). Cheltenham: Edward Elgar.

DACIN, M., VENTRESCA, M., & BEAL, B. (1999). The embeddedness of organizations: dialogue and directions. *Journal of Management* , 317-356.

DHVMC. (2004). *Plano de acção para a implementação e dinamização turística e cultural da Rota do Românico do Vale do Sousa.* Porto: DHV FBO Consultores.

ESPINOSA, A., HARNDEN, R., & WALKER, J. (2008). A complexity approach to sustainability - Stafford Beer revisited. *European Journal of Operation Research* , 636-651.

EUROSTAT. (2006). *Key Figures on European Business.* Luxembourg: European Commission.

GIDDINGS, B., HOPWOOD, P., & O'BRIEN, G. (2002). Environment, economy and society: Fitting them together into sustainable development. *Sustainable Development* , 187-196.

GREEVE, A. (1995). Networks and entrepreneurship - An analysis of social relations, occupational background, and use of contacts during the establishment process. *Scandinavian Journal of Management* , 1-24.

HILL, M., & HIL, A. (2002). *Investigação por Questionário.* Lisboa: Edições Sílabo.

IAPMEI. (2007). *Sobre as PME em Portugal.* Lisboa: Instituto de Apoio às Pequenas e Médias Empresas e à Inovação.

INE. (2008, July). Estatísticas Territoriais - N.º de Empresas. Lisboa, Lisboa, Portugal; [http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0001543&Contexto=pi&selTab=tab0].

INE. (2007). *Estimÿÿÿÿas provisórias de população residente Portugal, Nuts II, Nuts III, Municípios.* Lisboa: Instituto Nacional de Estatísticas.

ISLAM, S. M., MUNASINGHE, M., & CLARKE, M. (2003). Making long-term economic growth more sustainable: evaluating the costs and the benefits. *Ecological Economics* , 149-166.

JACK, S., & ANDERSON, A. (2002). The effects of embeddedness on entrepreneurial process. *Journal of Business Venturing* , 467-487.

JONG, J., & MARSILI, O. (2006). The fruit flies of innovations: A taxonomy of innovative small firms. *Research Policy* , 213-229.

KERR, C. D. (2006). Leadership strategies for sustainable SME operation. *Business Strategy and the Environment* , 30-39.

KORTON, D. (1996). *When Corporations Rule the World.* London: Earthscan.

LEE, S., & PETERSON, S. (2000). Culture, entrepreneurial orientation and global competitiveness. *Journal of World Business* , 401-416.

MAGRETTA, J. (2004). O grau zero da gestão. *Ideias & Negócios:* , 66-67.

MCCANN, P. (2006). On the supply-side determinants of regional growth. *Construction Management and Economics*, 681-693.

MELO, J. (2002). A problemática e as políticas de desenvolvimento local. In J. Costa, *Compêndio de Economia Regional* (pp. 515-533). Coimbra: Colecção APDR.

MEZIAS, S., & KUPERMAN, J. (2000). The community dynamics of entrepreneurship: the birth of american film industry: 1895-1929. *Journal of Business Venturing* , 209-233.

MINGUZZI, A., & PASSARO, R. (2000). The networks of relationships between the economic environment and the entrepreneurial culture in small firms. *Journal of Business Venturing* , 181-207.

MONBIOT, G. (2000). *Captive State.* London: Macmillan.

NARULA, R. (2004). R&D collaboration by SMEs: New opportunities and limitations in the face of globalisation. *Technovation* , 153-161.

NARULA, R. (2004). R&D collaboration by SMEs: New opportunities and limitations in the face of globalisation. *Technovation* , 153-161.

NARULA, R. (2004). R&D collaboration by SMEs: New opportunities and limitations in the face of globalization. *Technovation* , 1-16.

OECD. (2000). *Small and Medium Enterprise Outlook.* Paris: Organization for Economic Co-operation and Development.

OECD. (2003). *The Source of Economic Growth.* Paris: Organization for Economic Co-operation and Development.

PECQUEUR, B. (1989). *Le Dévelopment Local.* Paris: Syros - Alternatives Économiques.

PINCHOT, G. (1985). *Intrapreneuring.* New York: Harper and Row.

ROCHA, M., SEARCY, C., & KARAPETROVIC, S. (2007). Integrating sustainable development into existing management systems. *Total Quality Management & Business Excellence* , 83-92.

ROMER, P. (2001). Economic Growth.

SAUNDERS, M., LEWIS, P., & THORNHILL, A. (2003). *Research Methods fos Business Students.* Essex: Prentice Hall: Financial Times.

SCHMIEMANN, M. (2006). SMEs and Entrepreneurship in the EU. In *Statitics in Focus* (p. Luxembourg). Eurostat.

SCHUMPETER. (1934). *The Theory of Economic Development.* Cambridge.

SILVA, & SILVA. (2002). In Costa, *Compêndio de Economia Regional* (pp. 159-234). Lisboa: Colecção APDR.

VARGAS, C. M. (2000). 2000. *Community Development and Micro-enterprises: Fostering Sustainable Development* , 11-26.

VENKATARAMAN. (2004). Regional transformation through technological entrepreneurship. *Journal of Business Venturing* , 153-167.

VENKATARAMAN, S., & SARASVATHY, S. (2001). Strategy and entrepreneurship: outlines of an unfold story. In A. Hitt, E. Freeman, & J. Harrison, *Blackwell Handbook of strategic Management.* Oxford: Blackwell Publishing.

VISSER, E. J. (1999). A comparison of clustered and dispersed firms in the small-scale clothing industry of Lima. *World Development* , 1553-1570.

WEAVER, C. (1988). Concepts and theories of regional development planning: The state of the art. *Regional Development Planning in Developing Countries.* Dortmund.

YOUNG, W., & TILLEY, F. (2006). Can business move beyond efficiency? The shift toward effectiveness and equity in the corporate sustainability debate. *Business Strategy and the Environment* 402-415.

A FERRAMENTA *CAF* COMO INSTRUMENTO DE MELHORIA DA ADMINISTRAÇÃO LOCAL

RITA SANDRA BARROS RIBEIRO
E-mail: ritasribeiro@gmail.com
Câmara Municipal do Fundão
LUÍS ANTÓNIO NUNES LOURENÇO
E-mail: lourenço@ubi.pt
Universidade da Beira Interior

Resumo

A CAF (Estrutura Comum de Avaliação) permite a auto-avaliação organizacional do sector público. Esta ferramenta é simples e gratuita e possibilita a elaboração de estudos comparativos a nível europeu, de forma a promover o uso de técnicas de Gestão da Qualidade, rumo à melhoria contínua, procurando a Excelência Organizacional.
O objectivo central deste estudo foi o de verificar a consistência interna da ferramenta CAF, aferindo a interdependência entre os seus critérios, expressa através das hipóteses de investigação. Utilizaram-se dados empíricos, recolhidos através de questionário, junto de trabalhadores da na Câmara Municipal do Fundão.
Os resultados obtidos permitem considerar a CAF como uma importante ferramenta de auto-avaliação, propiciando a implementação, avaliação e desenvolvimento de um sistema de Qualidade em entidades públicas. Inerente à natureza do modelo está subjacente a ideia de que um bom desempenho nos meios conduzirá a bons resultados.

Palavras-chave: CAF, Qualidade, Administração Local.
Área temática: Economia Urbana, Regional e Local.

Abstract

CAF (Common Assessment Framework) provides an organizational self-assessment of the public sector. This tool is simple and costless and it also makes possible an elaboration of a comparative study at the European level, in a way of encouraging the use of quality management techniques and achieving a continuous improvement, in order to reach organizational excellence.
The main objective of this study is to access the internal consistence of CAF, by estimating the interdependency between its criteria, expressed by the research hypotheses. Data was through a questionnaire filled out by personnel of Câmara Municipal do Fundão.
Results make it possible to consider the CAF as an important tool of self-assessment, providing an implementation, assessment and development of a quality system in any public entity. Good performance on enablers will lead to good results.

Key words: CAF, Quality, Local Administration.
Thematic Area: Urban, Regional and Local Economics.

1. INTRODUÇÃO

Seguindo de perto Gadden (1972), a profissão de funcionário público pode ser encarada como a primeira, uma vez que a Administração Pública é considerada tão antiga como a própria Humanidade. Ao organizarem-se em comunidades, os seres humanos desde logo sentiram a necessidade de estruturarem alguns serviços em comum. Desde cedo, a evolução das organizações públicas teve de acompanhar os progressos sociais. Porém, os primeiros anos do século XXI não escondem a profunda antítese entre o ritmo das alterações à escala do globo, ao nível da sociedade global, e a lenta mudança da Administração Pública, a qual mantém o seu

recorte centenário, não obstante o diferente papel do Estado e as exigências crescentes dos cidadãos.

Mesmo após alguns processos de mudança, a Administração Pública não utilizou com sucesso as metodologias de gestão, pese embora a liderança política que promove a reforma defender a gestão e conceitos como a "orientação para o cliente", a "gestão por objectivos" ou a "avaliação do desempenho". Contudo, será que as práticas de sucesso nas empresas privadas são aplicáveis ao sector público?

Pode afirmar-se que, de uma administração clássica, baseada numa burocracia centralizada que implementa autonomamente programas de cima para baixo e que evita a participação dos cidadãos, transitou-se para uma administração gestionária, descentralizada, que importou técnicas de gestão do sector privado, orientando-se para os resultados, o mercado e o cliente. Porém, e acompanhando Bilhim (2001), deve conceber-se, de momento e ainda que de forma embrionária, uma administração enriquecida pelas teorias da cidadania democrática, que transforma o modelo gestionário numa administração receptiva à cidadania.

Reconhecida a necessária implementação de um processo de mudança no seio da Administração Pública, desde logo se coloca a questão da Reforma da Administração Pública Portuguesa, uma vez que com um olhar no passado, percebe-se o presente e projecta-se o futuro.

Porque qualquer Processo de Reforma da Administração Pública pressupõe um modelo de gestão desburocratizado e simplificado, com decisões eficazes e eficientes que correspondam às necessidades e expectativas dos clientes, a adopção de práticas de Qualidade é inquestionável.

Assumida a necessidade de consolidar na Administração Pública os princípios da Gestão pela Qualidade Total é apresentada a ferramenta da *Common Assessment Framework* (CAF), em Portugal designada por "Estrutura Comum de Avaliação".

Na CAF os resultados excelentes em desempenho, cidadãos-clientes, pessoas e sociedade são atingidos através de uma liderança que conduz a estratégia e o planeamento, as pessoas, as parcerias e recursos e os processos (Andrade, 2007). Atendendo a que esta ferramenta pressupõe uma interligação entre os seus critérios, o trabalho de investigação teve como principal objectivo analisar a consistência interna do modelo CAF, de acordo com as hipóteses de investigação definidas, as quais procuram avaliar o modo como o desenvolvimento das diferentes actividades ao nível de cada critério influencia o desenvolvimento da organização nos outros critérios.

Face ao interesse demonstrado pela Administração Local na implementação de programas de qualidade, e em virtude dos modelos de Gestão da Qualidade utilizados pelas autarquias serem suportados no modelo em estudo, entendeu-se ser oportuno recolher os dados empíricos numa Câmara Municipal, neste caso a do Fundão.

2. REFORMA ADMINISTRATIVA E QUALIDADE NA ADMINISTRAÇÃO LOCAL

Qualquer Processo de Reforma e Modernização da Administração Pública deve ter como meta uma necessária simplificação e desburocratização de procedimentos e processos salvaguardando, designadamente, o respeito pelos princípios da legalidade, da igualdade e da transparência, objectivando-se que a máquina administrativa saiba satisfazer as necessidades e expectativas dos cidadãos/clientes, administrativamente designados, ao nível da realidade autárquica, por munícipes.

A prestação de serviços públicos pressupõe uma série de intervenientes e, por conseguinte, de interacções com um contributo decisivo para a qualidade do serviço prestado, devendo assumir-se, claramente, que todo o Processo de Reforma Administrativa pressupõe, de forma inevitável, a adopção de "práticas de qualidade".

Embora as autarquias, se encontrem numa posição de monopólio na prestação de serviços indispensáveis ao cidadão/cliente, entenda-se munícipe, deve aquela, no entanto, procurar seguir um modelo de gestão interna/externa orientado para a adopção de práticas de Gestão pela Qualidade Total. Uma vez que a prestação de serviços envolve mais desempenho organizacional do que produção, a "mão-de-obra" é uma componente essencial, devendo a estratégia orientada

para o cliente pressupor uma capacidade dos colaboradores oferecerem empatia, simpatia e confiança.

A necessidade de equacionar a importância da Qualidade em todo este Processo de Reforma Administrativa está perfeitamente justificada, atendendo ao facto de que apenas uma diferente filosofia de gestão permitirá levar a bom porto qualquer processo de Modernização Administrativa.

Falar sobre Qualidade equaciona, desde logo, a necessidade de apresentar uma adequada definição para um conceito que tem suscitado inúmeras enunciações, subscritas quer por autores e teóricos do tema, quer por organizações internacionais. A evolução do conceito de Qualidade partiu da característica de um produto e chegou à característica de um sistema de gestão, responsável pelo produto final e extensível à qualidade dos serviços.

Embora com filosofias de gestão diferentes, todos os "gurus" da qualidade têm pontos em comum (Cortada e Quintella, 1993):

- A Qualidade deve ser abordada de forma específica e sistemática;
- A implementação da Qualidade requer tempo;
- Todas as pessoas devem estar envolvidas e devem poder tomar decisões referentes à Qualidade em todos os níveis;
- A Qualidade requer concentração;
- A Qualidade é um processo de gestão que gera produtos mais baratos e de qualidade superior;
- Os clientes são o ponto central em todos os processos de melhoria.

Considerando que as autarquias são entidades públicas prestadoras de serviços também estas devem apostar na Qualidade. A nível mundial, a Qualidade é uma prioridade estratégia que se traduz em adquirir vantagem competitiva. No caso da administração pública central, e das autarquias locais em particular, e considerando a posição monopolística que ocupam por inerência, a procura pela Qualidade traduz não tanto uma vantagem competitiva mas sim a procura da Excelência Organizacional.

Localmente, os órgãos executivos (Câmaras Municipais) têm cada vez mais a preocupação de satisfazer as necessidades dos munícipes objectivando, não apenas um proveito político, mas também o propiciar de um diferente patamar de qualidade de vida.

O conceito de Qualidade deve ser abordado na perspectiva da sua relevância e adequação no seio da Administração Pública Portuguesa, considerando o facto de certas especificidades do sector público comprometerem a aplicação directa de metodologias de Qualidade importadas do sector privado.

Com o que antecede, é inevitável concluir que não existe processo de Reforma Administrativa sem objectivos de Qualidade. Por seu turno, os objectivos da Qualidade identificam-se com os da Reforma das Administrações Públicas, pelo que falar em Reforma Administrativa e Qualidade na Administração Pública gera uma análise de conceitos e objectivos coincidentes.

De facto, o Processo de Reforma é indissociável do movimento internacional da Qualidade e da respectiva integração nos modelos de gestão global. Em Portugal, podem ser identificados como momentos marcantes a criação, em 1986, do SMA, a instituição do Prémio de Qualidade em Serviços Públicos e, no decurso do ano de 1999, a publicação de diplomas legais conexos com as medidas de Reforma e Modernização Administrativas (Decreto-Lei n.º 135/99, de 22 de Abril, alterado pelo Decreto-Lei n.º 29/00, de 13 de Março) e a instituição do Sistema de Qualidade em Serviços Públicos (Decreto-Lei n.º 166-A/99, de 13 de Maio). A institucionalização de uma nova cultura de gestão na Administração Pública é reforçada em 2000 com a Declaração de Lisboa sobre a Política da Qualidade das Administrações Públicas da União Europeia e a difusão do modelo da Estrutura Comum de Avaliação (*Common Assessment Framework* ou CAF), modelo de auto-avaliação objecto de estudo e apreciação no presente estudo.

A trave mestra do Sistema de Qualidade em Serviços Públicos é a certificação da qualidade. Os certificados são concedidos, após a avaliação do preenchimento de certos requisitos,

concretizada pela entidade gestora (o SMA). Porém, o Decreto-Lei n.º 166-A/99 nunca foi regulamentado e nunca foram igualmente criados os serviços de auditoria de qualidade, não tendo sido emitidos certificados. Assim, as entidades públicas interessadas em obter a certificação apenas podem recorrer às normas ISSO.

Segundo Carapeto e Fonseca (2005) podem ser identificadas três fases na evolução da Qualidade no Sector Público:

1. Qualidade no sentido do respeito pelas normas e pelos procedimentos (que corresponde, de certa forma, à conformidade com as especificações técnicas na indústria);
2. Qualidade equiparada ao conceito de eficácia, com ausência de erros e em estrita relação com os objectivos do serviço, divulgando-se a noção de gestão por objectivos (fase vigente a partir do anos 60 do século passado). Esta fase é equivalente à noção de aptidão para o uso defendida por Juran;
3. Qualidade enquanto satisfação do cliente, assimilando-se a ideia de Gestão pela Qualidade Total (a partir dos anos 80).

Porém, o conceito de Gestão pela Qualidade Total só tardiamente começou a ser adoptado nos países da Europa Continental, uma vez que, até há bem pouco tempo, a Qualidade cingia-se à mera conformidade com a lei. No entanto, os programas de modernização das administrações públicas dos governos europeus das últimas décadas privilegiam, indiscutivelmente, as teorias da Qualidade (Carapeto e Fonseca, 2005).

Assim, está em curso uma fase em que a Reforma Administrativa se confunde com o conceito da Qualidade Total, uma vez que em ambas as realidades se procura atingir a Excelência da gestão pública, a definição de uma estratégia envolvendo todas as partes interessadas e a racionalização e simplificação de processos com recurso à utilização de novas tecnologias.

Pese embora os programas de Reforma e Modernização da Administração Pública lançados pelos governos europeus terem como suporte a implementação das teorias da Qualidade, determinadas particularidades do sector público dificultam a aplicação das metodologias da Qualidade importadas directamente do sector privado. Algumas dessas especificidades estão relacionadas com o facto de enquanto que empresas procuram o lucro, vocacionando a sua actividade para alcançar uma vantagem competitiva no mercado, a Administração Pública serve os cidadãos, maximizando o interesse público, o que por vezes leva a que valores como a equidade, a igualdade e a democracia se sobreponham à concorrência e à economia.

Acresce ainda que os serviços públicos pautam primordialmente a sua actividade pelo cumprimento de programas político-partidários e estão sujeitos a preceitos legais específicos. Aspectos como o orçamento, o estatuto de carreiras, as remunerações, o sistema de avaliação do desempenho e as regras do recrutamento e selecção de pessoal são definidos pelo governo e comprometem eficazes processos de mudança da cultura organizacional (Carapeto e Fonseca, 2005).

Outros obstáculos estão associados à falta de planeamento estratégico direccionado para a mudança, à ausência de liderança, à falta de motivação e formação dos funcionários, à cultura organizacional ultrapassada (desde logo porque em causa está uma desestruturação das hierarquias e um *empowerment* dos funcionários e dos grupos de trabalho), à dificuldade de orientação para o cliente e ao facto de, frequentemente, os serviços prestados pela Administração Pública o serem em regime de monopólio. Saliente-se ainda que os clientes (os que pagam) nem sempre coincidem com os utentes, não obstante os serviços públicos de Qualidade terem de dar satisfação a ambos.

Para a maioria dos autores a introdução, com sucesso, da Gestão pela Qualidade Total nas organizações públicas assenta em três pontos críticos (Carapeto e Fonseca, 2005):

1. Na selecção dos indicadores de Qualidade, presentes nas ISO 9000 e em Modelos de Excelência (*European Foundation for Quality Management* - EFQM e CAF);
2. No uso de uma ferramenta de avaliação indicada para avaliar o progresso da organização em direcção à qualidade;

3. No envolvimento dos *stakeholders* pertinentes (políticos, dirigentes de topo e intermédios, clientes internos e externos).

Como factor crítico, Carapeto e Fonseca (2005) acrescenta ainda o encarar a Qualidade como uma estratégia de longo prazo e salienta que as organizações que pretendem orientar a sua conduta por padrões de Qualidade Total não devem contentar-se com a certificação, ou seja, com a emissão de um certificado que atesta a conformidade com os requisitos pré-definidos por normas ou parâmetros formais que quantificam a Qualidade. A observância das normas ISO, por exemplo, é um mínimo que permite alcançar a certificação mas nunca garante a Excelência nos resultados da actividade.

Todas estas particularidades propiciaram a criação de Modelos de Qualidade específicos para o sector público, tais como a vertente do modelo EFQM para o sector público e a ferramenta CAF. Fazendo uma referência específica à Administração Local, é importante referir que as preocupações com a Qualidade apenas se registam no começo da década de 90, objectivando-se aumentar a participação dos munícipes e a eficiência dos serviços (Rocha, 2006). De acordo com Berman e West (1995), as aplicações da Gestão pela Qualidade Total são mais frequentes nos serviços de polícia, jardinagem e asseio, parques e recolha de lixo, ou seja, em serviços prestados directamente aos cidadãos.

Segundo Rocha (2006), só nos finais dos anos 90 se pode falar da implementação de um modelo de Gestão pela Qualidade Total, seja ele a certificação através das ISO, seja um qualquer Modelo de Excelência. Importante é também perceber, desde logo, que a introdução da Gestão da Qualidade no governo local deve ter em conta que o sistema político-administrativo tem contornos especiais (Appleby e Clark, 1997):

- O facto das autoridades locais estarem sujeitas a eleições implica que a Qualidade não seja apenas um problema de gestão mas também uma questão política;
- Os municípios são organizações com diferentes funções, designadamente ao nível da educação, habitação social;
- Os municípios são moldados pela cultura local.

De acordo com um estudo sobre municípios portugueses, elaborado por Sá (2001), é importante apresentar as seguintes conclusões:

- Os municípios revelam estar conscientes da importância de adoptar uma gestão da Qualidade;
- Recentemente tem ganho relevo a auto-avaliação com base no EFQM ou na CAF;
- A implementação da Qualidade não é generalizada a todo o município, considerando que estes são organizações multifuncionais;
- A publicação do Decreto-Lei n.º 166-A/99, desprovida da respectiva regulamentação, motivou os municípios a recorrerem à certificação com base nas normas ISO.

Considerando que é fundamental consolidar na Administração Pública os princípios da Gestão pela Qualidade Total, direccionando-a, progressivamente, para um patamar de Excelência Organizacional, numa perspectiva de melhoria contínua, é analisada de seguida a ferramenta CAF. De facto o objectivo fundamental deste estudo é testar a sua coerência interna, apreciando a interdependência entre os seus critérios, em conformidade com as hipóteses de investigação formuladas, e avaliar a oportunidade da sua aplicação ao nível da Câmara Municipal do Fundão.

3. A CAF – ESTRUTURA COMUM DE AVALIAÇÃO

Enquanto Modelo de Gestão da Qualidade, a CAF possui como objectivo principal a melhoria contínua dos Serviços Públicos, como princípios o ênfase no cliente, a melhoria dos processos e o envolvimento total das partes interessadas, e como elementos de apoio a liderança, a aprendizagem e a formação, a equipa de apoio, a comunicação, a avaliação de resultados e o reconhecimento público.

Após a aprovação da Carta Europeia para a Qualidade, a Comissão Europeia comunicou, em 1996, aos ministros responsáveis pelas administrações públicas dos quinze, a necessidade premente de criar e potenciar uma Política Comum da Qualidade para as Administrações Públicas dos Estados Membros, assumindo que quanto melhor for a Administração Pública Europeia melhor será a Europa e maior a Cidadania Europeia. Espera-se que Administrações Públicas com qualidade contribuam para a coesão europeia, reforçando a cidadania, para a competitividade das empresas, para um melhor investimento e para fortalecer a Europa.

No decurso da presidência austríaca da União Europeia, ou seja, na segunda metade de 1998, os Directores-Gerais das Administrações Públicas Europeias equacionaram a possibilidade de desenvolver uma estrutura comum europeia que permitisse a auto-avaliação organizacional do sector público. Essa ferramenta de auto-avaliação deveria ser simples, gratuita, permitir um estudo comparativo ao nível europeu e promover o uso de técnicas de Gestão da Qualidade, rumo à melhoria contínua, procurando a Excelência.

Ficou definido que essa ferramenta seria desenvolvida por um grupo director, constituído por peritos nacionais, para organizar a cooperação entre os 16 membros (os 15 e a Comissão) ao nível da cooperação internacional. Este grupo, criado pelos Directores-Gerais, veio a ser posteriormente apelido de IPSG (*Innovative Public Services Group*), e supervisionou o desenvolvimento dos trabalhos de concepção e avaliação da CAF, concretizados pelo *CAF Working Group*. Saliente-se que o esboço da CAF foi ainda fruto dos trabalhos de cooperação entre o EIPA (*European Institute of Public Administration*)[1], a Academia Speyer e a EFQM, no intuito de desenvolver uma ferramenta de auto-avaliação para as administrações públicas europeias que pretendessem melhorar a Qualidade dos seus serviços.

Na 1ª Conferência da Qualidade das Administrações Públicas da União Europeia, sob o tema "Partilhar as Melhores Práticas", realizada em Lisboa, sob a presidência portuguesa, em Maio de 2000, a CAF foi apresentada pela primeira vez. Nova versão da CAF foi posteriormente apresentada, em 2002, na 2ª Conferência Europeia realizada em Copenhaga, Dinamarca (dias 2, 3 e 4 de Outubro), subordinada ao tema "Inovação, Mudança e Parcerias". Esta versão plasmou melhorias resultantes de um questionário efectuado, em 2001, aos Estados Membros que adoptaram a CAF entre 2000 e 2001, pelo que, em Maio de 2002, foi acordada uma nova estrutura e que se traduziu, basicamente, na alteração do critério 5 e na diminuição do número de sub-critérios.

Nos dias 27, 28 e 29 de Setembro de 2006 teve lugar a 4ª Conferência da Qualidade das Administrações Públicas da União Europeia, sob a presidência da Finlândia, e subordinada ao tema "*Building Sustainable Quality*". Nesta conferência, e à semelhança das anteriormente realizadas, foram apresentadas as melhores práticas dos diferentes Estados Membros[2] e divulgada a nova versão da CAF – A CAF 2006, sendo conveniente enumerar as principais alterações introduzidas:

- Atribuição de um maior ênfase à modernização e inovação, designadamente vincando o primeiro conceito nos critérios da Liderança e do Planeamento e Estratégia e o segundo no Processos;
- Possibilidade de pontuar de duas formas, ou seja, através do sistema de pontuação clássico e com recurso ao denominado sistema de pontuação avançado. O sistema de pontuação possibilita focalizar, de forma precisa, as acções de melhoria;
- Linhas de orientação para aplicação da CAF mais detalhadas, uma vez que a introdução e ou reformulação de muitos exemplos explicitam melhor a contribuição para uma melhor governação resultante da gestão da Qualidade no sector público;

[1] O Centro de Recursos CAF é coordenado pelo EIPA e foi criado com o objectivo de centralizar a informação relativa à CAF e à sua evolução, monitorizando a sua utilização nos Estados-Membros. No sítio da internet do EIPA pode-se encontrar uma base de dados dos casos de aplicação da CAF já registados.

[2] No caso de Portugal foram seleccionados, através de métodos e processos de selecção independentes, as boas práticas Bolsa de Emprego Público (BEP), Programa de Compras Electrónicas e Sistema "Empresa na Hora".

- Existência de orientações para desenvolver *benchlearning* com a CAF.

De forma sistemática, e segundo o referido na parte introdutória da versão actual da CAF, esta tem quatro objectivos principais (DGAEP, 2007):

1. Introduzir na Administração Pública os princípios da Gestão pela Qualidade Total (*Total Quality Management*) e orientá-la progressivamente, através da utilização e compreensão da auto-avaliação, para um ciclo completo e desenvolvido – PDCA: Planear (fase de projecto), Executar (fase da execução), Rever (fase da avaliação) e Ajustar (fase da acção, adaptação e correcção);
2. Facilitar a auto-avaliação das organizações públicas com o objectivo de obter um diagnóstico e um plano de acções de melhoria. Possibilita que os gestores públicos melhorem o desempenho da sua organização, considerando que as organizações públicas podem-se auto-avaliar a custos reduzidos, permitindo a mudança e a melhoria contínua dos processos e procedimentos;
3. Servir de ponte entre os vários modelos utilizados na Gestão da Qualidade, uma vez que é compatível com os principais modelos usados pelas organizações públicas e privadas na Europa, através da introdução de alguns indicadores que possibilitam comparar os resultados produzidos por diferentes sistemas;
4. Facilitar o *benchlearning* entre organizações do sector público.

Definidas, ainda que de forma sucinta, as suas características chave, bem como os seus objectivos, chegou o momento de perguntar para que serve a CAF. De forma igualmente concisa pode-se afirmar que a CAF serve para identificar os pontos fortes e os pontos críticos das organizações públicas, para revelar as percepções das pessoas em relação à sua própria organização, para aumentar a mobilização interna para a mudança e desenvolver o sentido e auto-responsabilização dos gestores, para conhecer o grau de satisfação dos diferentes públicos que se relacionam com as organizações públicas (de entre os quais podemos salientar os cidadãos, as empresas, os fornecedores e a sociedade civil) e permite construir projectos de mudança sustentados no conhecimento e no estado de saúde da organização.

Tal como o Modelo de Excelência, qual deriva, o CAF pode ser entendido como uma ferramenta de implementação da Gestão pela Qualidade Total (Lourenço, 2000); adequado às características dos organismos públicos, que auxilia as administrações públicas da União Europeia a compreender e implementar as técnicas de Gestão da Qualidade e a auto-avaliar as organizações públicas numa perspectiva de melhoria contínua. A metodologia do CAF assenta no círculo de Deming e na ferramenta de auto-avaliação. A ferramenta de auto-avaliação propicia a elaboração de um diagnóstico organizacional onde são identificados os pontos fortes e as áreas de melhoria. Segue-se a implementação de acções de melhoria que permitirão alcançar melhores resultados, levando a que se efectue novamente um novo diagnóstico organizacional.

Analisados os critérios da ferramenta CAF, oferece-se perguntar o porquê deste modelo fazer sentido ao nível organizacional e, atendendo ao objecto de estudo, no âmbito de uma autarquia local? Desde logo, existem causas internas e externas que condicionam a necessidade de se concretizar uma mudança organizacional. Ao nível externo, existe uma sociedade complexa na qual surgem novas exigências dos cidadãos e das empresas e uma premente necessidade de reduzir, ou pelos menos rentabilizar, a despesa pública. A aquisição de novos conhecimentos e competências e a necessidade de alterar a imagem da Administração Pública Local são causas internas que impulsionam a eminente mudança. A CAF é assim uma importante ferramenta de mudança uma vez que tem três objectivos: modernizar os serviços públicos, orientá-los para os clientes e implementar, de forma participativa, Programas de Qualidade no seio da administração.

Com a CAF é então possível, através de exercícios de auto-avaliação a baixo custo e passíveis de serem utilizados por pessoas que não possuem formação especializada em análise organizacional, identificar pontos fortes e oportunidades de melhoria, elaborar uma plano estratégico para implementar essas melhorias, contribuir para a responsabilização dos órgãos de

gestão e servir de ponto de partida para um processo de melhoria contínua da Qualidade, sensibilizando para a mudança.

A CAF permite ainda conhecer a própria organização, sendo fácil de aplicar e com custos reduzidos, e, se de aplicação sistemática e com recurso ao *benchmarking*[3], faculta indicadores de desempenho para comparar departamentos de uma organização, podendo ser aplicada quer ao nível da macroestrutura (toda a organização) quer da microestrutura (departamento), e para comparar várias organizações. A aplicabilidade da CAF ao nível da microestrutura está dependente do facto de um determinado departamento constituir o que na terminologia CAF é designado por "unidade de gestão". Este departamento tem de possuir uma intervenção significativa na elaboração da política e da estratégia e uma quase autonomia na política de pessoal.

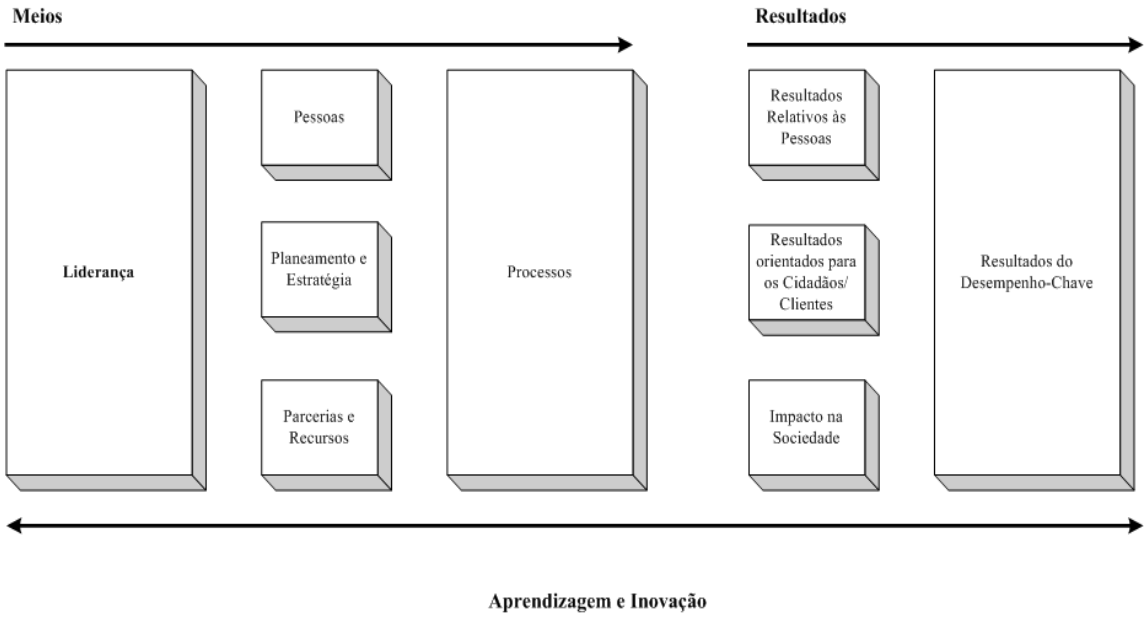
Porque qualquer processo de mudança pressupõe a presença de resistências, estas podem ser atenuadas se os colaboradores forem envolvidos e informados, perceberem os benefícios da mudança e, acima de tudo, se forem eliminados os seus aspectos ameaçadores. Este modelo envolve, desde logo, as pessoas da organização, pois será aplicado por um grupo de colaboradores que, ao avaliarem a organização em que se integram, expressam a sua percepção quanto aos pontos críticos que conduzirão a planos de acções de melhoria. Este envolvimento passará, necessariamente, por questionários aos clientes externos e internos, por entrevistas aos dirigentes, pela apresentação de um diagnóstico organizacional e de medidas de melhoria a implementar, pela negociação do programa de Qualidade a concretizar e seu acompanhamento pelas partes interessadas. A implementação da CAF está, em súmula, profundamente dependente de um forte compromisso da gestão de topo, de uma boa comunicação entre dirigentes e colaboradores, da vontade dos colaboradores em aumentar os níveis de desempenho pessoal e organizacional e da clarificação dos compromissos a assumir.

Podem-se assim identificar os seguintes factores de sucesso na aplicação da CAF e da auto-avaliação (Engel, 2003) apoio total da gestão de topo; aceitação da organização; grupo de auto-avaliação representativo da organização; apoio metodológico (devendo recorrer-se a consultores externos se se justificar); rapidez na execução de trabalhos; reunião de "provas" justificativas das pontuações; tradução de resultados em prioridades para acções de melhoria; reforço da comunicação nas acções de melhoria e repetição periódica do exercício para aferir o progresso.

Segundo a ferramenta CAF, a auto-avaliação de uma organização é orientada por dois tipos de critérios – os critérios de meios e os critérios de resultados, realidade expressa na Figura 1, os quais se subdividem 28 sub-critérios que identificam as principais questões a equacionar na avaliação de cada organização. Os critérios meios reportam-se à forma como as actividades da organização são desenvolvidas, designadamente quanto à utilização dos recursos disponíveis (como a organização actua para obter resultados), estando associados ao futuro da organização. Os critérios resultados avaliam o produto final das acções empreendidas, com utilização dos meios necessários, para atingir os objectivos fixados para a organização, reportando-se ao presente e/ou ao passado.

[3] A utilização da CAF permite, através do sítio da internet do EIPA (www.eipa.nl), que uma organização preencha o formulário de avaliação depois de ter efectuado o seu processo de auto-avaliação. Os resultados facultados permanecem anónimos e a entidade recebe o seu posicionamento em termos de perfil organizacional, comparativamente com a média das outras organizações que têm utilizado a CAF no mesmo país ou no mesmo sector de actividade.

Figura1 – A Ferramenta CAF



Fonte: CAF 2006 (DGAEP, 2007).

Os critérios e sub-critérios do modelo CAF são pontuados com vista a alcançar quatro objectivos basilares:

1. Fornecer informação e indicar o caminho a seguir para as acções de melhoria;
2. Medir o progresso da organização (caso a CAF seja aplicada com regularidade, aconselhando-se que o seja anualmente ou com intervalos de dois anos);
3. Identificar boas práticas que resultem de pontuações elevadas nos critérios de meios e de resultados;
4. Ajudar a identificar parceiros com os quais seja útil aprender (*benchmarking* – como comparar e aprender e *benchlearning* – o que se aprende com os outros).

Refira-se que este estudo empírico teve por base a versão 2002 da CAF a qual não é contudo muito diferente da de 2006.

Qualquer processo de auto-avaliação tem na sua origem o comprometimento activo dos gestores de topo, os quais devem escolher as unidades que vão ser avaliadas, preferindo aquelas que estejam voluntariamente empenhadas em participar, estabelecer as equipas de avaliação e o líder do projecto. Estas equipas devem ser muito heterogéneas, incluindo todos os níveis de gestão, e devem receber formação sobre o processo de auto-avaliação, existindo um líder de equipa. Após a recolha de toda a informação são estabelecidos os pontos fortes e identificados os pontos a melhorar. Concluída a auto-avaliação é definido o plano de acções de melhoria. Em síntese, as fases procedimentais da ferramenta CAF são as seguintes:

- Organização e Planeamento da Auto-Avaliação;
- Realização do Processo de Auto-Avaliação;
- Obtenção de Resultados e Elaboração do Plano de Melhorias;
- Monitorização dos Progressos e Repetição da Auto-Avaliação.

4. OBJECTO DE ESTUDO E PROCEDIMENTOS METODOLÓGICOS

As competências e atribuições das autarquias locais, bem como os modelos organizacionais dos serviços, estão definidas pelo normativo legal actualmente em vigor. Este normativo faculta um instrumento que permite configurar a estrutura hierárquica e funcional dos recursos organizacionais, que podem ser adaptáveis às características do município e dimensionáveis de

acordo com as necessidades, garantindo a concretização do processo de planeamento e gestão municipal.

Em termos orgânicos, as Câmaras Municipais têm vindo a alargar a sua estrutura organizacional, em virtude do número de competências transferidas pela Administração Central, no decurso da implementação de um processo de descentralização. Organizacionalmente, as autarquias estão estruturadas mediante hierarquias verticais sendo a cúpula composta pelo Presidente e pelos Vereadores, seguidos dos Directores de Departamento, Chefes de Divisão, Chefes de Secção, Encarregados, etc., tendo este conceito raiz na estratégia napoleónica de acordo com a qual as informações eram transmitidas desde os superiores até aos operacionais, assegurando-se que a informação de chegada fosse igual à de partida (Tenedório, 2003).

Em termos orgânicos, a Câmara Municipal do Fundão é composta por três departamentos, cinco divisões, três gabinetes e duas unidades operativas directamente dependentes do Presidente da Câmara.

A Câmara Municipal do Fundão contava em 30 de Abril de 2006 com 332 trabalhadores, cujas relações jurídicas de trabalho assentavam em nomeação, contrato administrativo de provimento, contrato de trabalho a termo resolutivo certo, requisição, destacamento e nomeação política. Os 332 trabalhadores encontravam-se, à data, distribuídos por diversos grupos profissionais (dirigente, técnico superior, técnico, informático, técnico profissional, chefia, administrativo, apoio educativo, auxiliar e operário), sendo os mais representativos os de pessoal auxiliar e operário, com 65,06% do total dos colaboradores.

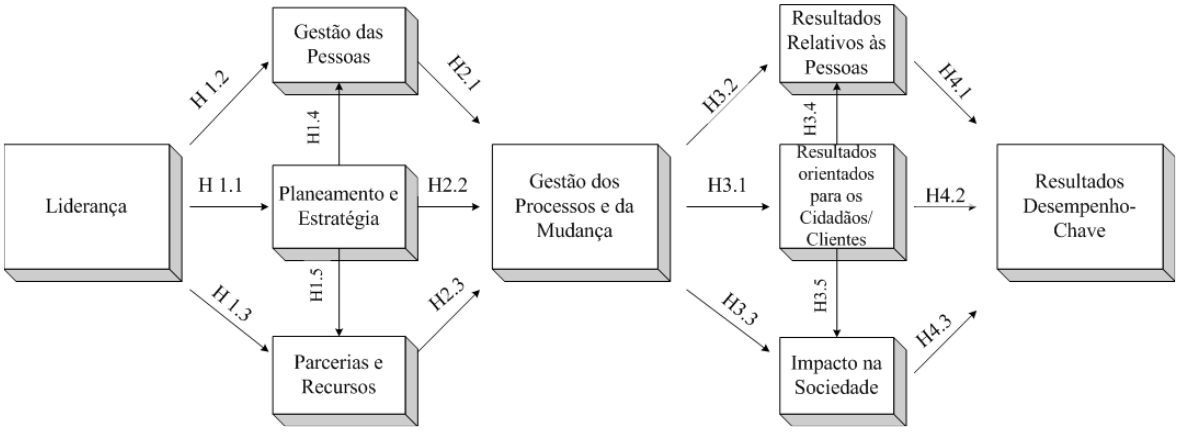
Uma vez que se pretende avaliar a consistência interna do modelo CAF e a relação de influência entre os critérios que o compõem, concluída a revisão da literatura relevante, definidos os objectivos e formuladas as hipóteses, tendo por base a revisão bibliográfica, elaborou-se um questionário modelo, entregue pessoalmente a todos os inquiridos. Aplicados os questionários, recolhidos os dados, e com vista a um adequado tratamento, foi utilizado o software estatístico *SPSS (Statistical Package for the Social Sciences) 15.0 for Windows* e o *Microsoft Office Excel*. A estruturação das informações recolhidas, e devidamente tratadas com recurso aos programas mencionados, permite retirar conclusões sobre as hipóteses previamente definidas.

Considerando que a Câmara Municipal do Fundão estabeleceu como prioridade estratégica a Modernização Administrativa, e uma vez que esta pressupõe, indiscutivelmente, uma adaptação permanente dos serviços municipais, de forma a satisfazer as crescentes exigências dos cidadãos/clientes, e dado que se equaciona a necessidade de implementar um modelo de qualidade vocacionado para o sector público – a ferramenta CAF, estabeleceu-se como objectivo basilar do estudo a avaliação da consistência interna deste modelo mediante a apreciação da interdependência entre os seus critérios, em função das hipóteses de investigação equacionadas, atendendo ao facto de que a promoção da Qualidade e a orientação para a Excelência são desafios inadiáveis que se colocam às organizações.

Pretende-se aplicar esta ferramenta, especialmente desenvolvida para as Administrações Públicas dos Estados Membros da União Europeia, a toda a Autarquia, de forma a identificar as boas práticas já existentes, os aspectos a melhorar, a avaliação regular que é concretizada e o grau de monitorização efectuado, de forma a permitir a introdução de processos de melhoria.

Em Ciências Sociais, as hipóteses de investigação surgem de várias formas com o intuito de serem confirmadas ou contestadas. As hipóteses a seguir representadas na Figura 2 têm a sua origem no modelo teórico CAF, já apresentado na Figura 1, ferramenta de avaliação de entidades públicas.

Figura 2 – Hipóteses de Investigação



Na CAF, os Resultados de Desempenho-Chave estão directamente dependentes dos Resultados Relativos às Pessoas, dos Resultados Orientados para os Cidadãos/Clientes e do Impacto na Sociedade, os quais são influenciados pela Gestão dos Processos e da Mudança. Por sua vez, também a Gestão dos Processos e da Mudança é directamente dependente da Gestão das Pessoas, do Planeamento e Estratégia e das Parcerias e Recursos, critérios estes igualmente dependentes da Liderança, critério que não depende de nenhum. Assim, a CAF baseia-se na premissa de que resultados excelentes em desempenho, cidadãos-clientes, pessoas e sociedade são atingidos através de uma liderança que conduz a estratégia e o planeamento, as pessoas, as parcerias e recursos e os processos (Andrade, 2007).

Para aferir a relação existente entre os critérios da ferramenta em estudo foram definidas as seguintes hipóteses de investigação:

H1.1: A *"Liderança"* está directamente relacionada com o *"Planeamento e Estratégia"*;

H1.2: A *"Liderança"* está directamente relacionada com a *"Gestão das Pessoas"*;

H1.3: A *"Liderança"* está directamente relacionada com as *"Parcerias e Recursos"*;

H1.4: O *"Planeamento e Estratégia"* estão directamente relacionados com a *"Gestão das Pessoas"*;

H1.5: O *"Planeamento e Estratégia"* estão directamente relacionados com as *"Parcerias e Recursos"*;

H2.1: A "*Gestão das Pessoas*" está directamente relacionada com a "*Gestão dos Processos e da Mudança*";

H2.2: O "*Planeamento e Estratégia*" estão directamente relacionados com a "*Gestão dos Processos e da Mudança*";

H2.3: As "*Parcerias e Recursos*" estão directamente relacionados com a "*Gestão dos Processos e da Mudança*";

H3.1: A "*Gestão dos Processos e da Mudança*" está directamente relacionada com os "*Resultados Orientados para os Cidadãos/Clientes*";

H3.2: A "*Gestão dos Processos e da Mudança*" está directamente relacionada com os "*Resultados Relativos às Pessoas*";

H3.3: A "*Gestão dos Processos e da Mudança*" está directamente relacionada com o "*Impacto na Sociedade*";

H3.4: Os "*Resultados Orientados para os Cidadãos/Clientes*" estão directamente relacionados com os "*Resultados Relativos às Pessoas*";

H3.5: Os "*Resultados Orientados para os Cidadãos/Clientes*" estão directamente relacionados o "*Impacto na Sociedade*";

H4.1: Os "*Resultados Relativos às Pessoas*" estão directamente relacionados com os "*Resultados de Desempenho-Chave*";

H4.2: Os "*Resultados Orientados para os Cidadãos/Clientes*" estão directamente relacionados com os "*Resultados de Desempenho-Chave*";

H4.3: O "*Impacto na Sociedade*" está directamente relacionado com os "*Resultados de Desempenho-Chave*".

4.1. Dados e amostragem

Neste estudo, o método escolhido para a recolha de informação é um método de observação indirecta e optou-se pela utilização de um questionário, que serviu para a recolha de dados respeitantes às diferentes variáveis em análise, face às vantagens que possui em relação à entrevista, uma vez que abrange uma porção mais ampla do universo em estudo, requer investimentos de tempo mais reduzidos e simplifica, ao permitir uma codificação, o tratamento das respostas (Fernandes, 1994). O questionário permite ainda alcançar um maior número de observações em simultâneo, abranger uma área geográfica mais ampla, obter um grande número de dados, dar maior liberdade nas respostas devido ao anonimato, correr menos riscos de distorção pela presença do investigador e obter respostas mais rápidas e mais precisas (Lakatos et al., 1996).

Seguindo a recomendação de Lakatos et al. (1996), a parte inicial do questionário inclui uma nota introdutória, explicando o objectivo da pesquisa e a necessidade de obter respostas, de modo a assegurar a colaboração do inquirido. O questionário utilizado é constituído por perguntas fechadas, uma vez que estas permitem o tratamento simplificado das respostas (as quais já se encontravam codificadas), e a dimensão do mesmo obriga a utilização deste tipo de perguntas sob pena do número de respostas não ser representativo se cada inquirido tivesse que responder a perguntas abertas.

O questionário foi estruturado em duas partes e baseou-se num modelo já elaborado por outros investigadores da mesma área (Mora, Leal e Roldán, 2005, e Fernandes, 2006). A primeira parte é composta por perguntas inerentes à identificação dos inquiridos: Unidade Orgânica, Categoria, Idade, Habilitações Literárias e Tempo de Trabalho na Câmara. A segunda é composta por nove conjuntos de questões agrupados pelos critérios da ferramenta CAF.

Embora em 30 de Abril de 2006 estivessem ao serviço da Câmara Municipal do Fundão 332 colaboradores, entendeu-se que os questionários apenas deveriam ser distribuídos a 250, uma vez que os demais trabalhadores ou possuíam um vínculo de cariz transitório ou prestavam serviço em Juntas de Freguesia e Jardins-de-Infância (vários com contratos de trabalho a termo resolutivo certo). Dos 250 questionários distribuídos tiveram retorno 159, mas só foram considerados como válidos 146, uma vez que o número de respostas dadas em 13 questionários ou era reduzido ou os mesmos estavam incorrectamente preenchidos.

Depois de terminado o prazo previamente estabelecido como horizonte temporal para a recepção dos questionários (1 mês), iniciou-se o tratamento estatístico.

Embora não exista consenso quanto à taxa de resposta considerada razoável, alguns investigadores (Lakatos e Marconi, 1991 e Kidder, Judd e Smith 1991) consideram como razoável uma taxa de 25%. Entregues 250 questionários retornaram 159 (63,6%) e foram considerandos válidos 146, os quais representam uma taxa de resposta de 58,4%.

5. RESULTADOS

Antes de se iniciarem os procedimentos estatísticos, foram analisados os *missing values*, com recurso ao software *SPSS*, a fim de ser detectado o problema das não respostas, tendo sido retiradas as variáveis que possuíam mais de 20% de valores em falta para evitar o enviesamento dos resultados (Guimarães e Cabral, 1997 e Pestana e Gageiro, 1998) ou seja:

- No critério Resultados Relativos às Pessoas a referente à "percepção quanto ao nível de satisfação actual dos colaboradores para com a Câmara" – com 45,2% de não respostas;
- No critério Resultados Orientados para os Cidadãos/Clientes a respeitante à "percepção quanto ao nível de satisfação actual dos cidadãos para com a Câmara" – com 43,2% de não respostas;

- No critério Impacto na Sociedade a relativa à "percepção quanto ao nível de satisfação actual da comunidade ou sociedade envolvente com a Câmara" – com 36,3% de não respostas.

A variável que tinha, de seguida, um maior número de não respostas era a afirmação "Redução do tempo efectivo de tratamento das reclamações apresentadas", pertencente ao critério Resultados Orientados para os Cidadãos/Clientes, com um valor de 8,9%.

De seguida, apresentam-se as médias globais das variáveis/critérios da ferramenta CAF, representadas no Quadro 1. As que possuem um valor mais elevado são o Impacto na Sociedade (3,71) e os Resultados de Desempenho-Chave da Câmara (3,60). As que possuem uma média mais reduzida são a Gestão das Pessoas (2,80) e o Planeamento e Estratégia (3,02). Considerando os valores apresentados, pode concluir-se que todas variáveis possuem valores acima do ponto médio, com excepção da Gestão das Pessoas com um valor de 2,80.

Quadro 1 – Média das Variáveis/Critérios da CAF

VARIÁVEIS /CRITÉRIOS DA CAF	MÉDIA
LIDERANÇA	3,34
GESTÃO DAS PESSOAS	2,80
PLANEAMENTO E ESTRATÉGIA	3,02
PARCERIAS E RECURSOS	3,34
GESTÃO DOS PROCESSOS E DA MUDANÇA	3,16
RESULTADOS RELATIVOS ÀS PESSOAS	3,34
RESULTADOS ORIENTADOS PARA OS CIDADÃOS/CLIENTES	3,48
IMPACTO NA SOCIEDADE	3,71
RESULTADOS DE DESEMPENHO-CHAVE DA CÂMARA	3,60

Previamente à análise de dados, no sentido de aceitar ou rejeitar as hipóteses anteriormente definidas, e uma vez que o método de recolha de dados, o questionário, procura medir percepções, é fundamental avaliar a consistência interna das perguntas efectuadas com recurso ao indicador *Alpha de Cronbach*, resultados que de seguida se apresentam no Quadro 2.

Quadro 2 – Consistência Interna entre Grupos de Questões – *Alpha de Cronbach*

GRUPO DE QUESTÕES	*ALPHA DE CRONBACH*
LIDERANÇA	0,957
GESTÃO DAS PESSOAS	0,963
PLANEAMENTO E ESTRATÉGIA	0,971
PARCERIAS E RECURSOS	0,912
GESTÃO DOS PROCESSOS E DA MUDANÇA	0,959
RESULTADOS RELATIVOS ÀS PESSOAS	0,974
RESULTADOS ORIENTADOS PARA OS CIDADÃOS/CLIENTES	0,956
IMPACTO NA SOCIEDADE	0,960
RESULTADOS DE DESEMPENHO-CHAVE DA CÂMARA	0,962

A consistência interna dos factores define-se como a proporção da variabilidade nas respostas que resultam de diferenças entre os inquiridos, ou seja, as respostas diferem não porque o questionário esteja confuso e leve a diferentes interpretações, mas porque os inquiridos têm diversas opiniões. Existem diversas opiniões sobre qual o valor que pode ser considerado

bom indicador dessa consistência. Van de Ven e Ferry, citados em Terziovski e Samson (1999) afirmam que o coeficiente *Alpha* deve assumir valores entre 0,7 e 0,9 e que o mínimo aceitável é 0,6. Por seu lado, Rowland, Arkkelin e Crisler (1991) consideram que o valor mínimo recomendado é de 0,7 e Cronbach, Pereira e Nunnally, citados em Sun (2000), referem que o *Alpha* é um bom indicador da consistência interna quando regista valores superiores a 0,8.

Todos os grupos de questões apresentam valores superiores a 0,9 o que, e atendendo aos valores de referência considerados pelos autores acima mencionados, indicia uma boa consistência interna.

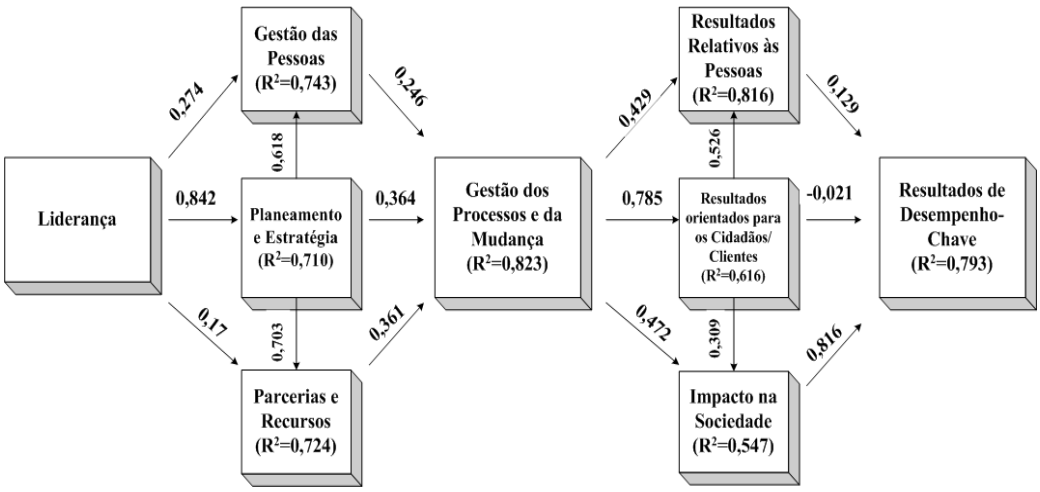
Para confirmar ou refutar as hipóteses de investigação definidas, recorreu-se ao Modelo de Regressão Linear. Foram verificados os pressupostos básicos necessários à utilização desta ferramenta estatística: homocedasticidade, independência, normalidade e inexistência de multicolinearidade (Pestana e Gageiro, 1998 e Neter et al., 1996).

Para confirmar ou refutar as hipóteses de investigação definidas recorreu-se, como já foi referido, ao Modelo de Regressão. Para a análise do modelo foi adoptada a hierarquia proposta por Jayaram e Ahire (1998), a qual estabelece quatro classes de níveis de significância para a análise dos coeficientes de regressão, a fim de determinar a relação existente entre as variáveis independentes e dependentes:

- Não significante: se $0,15 < p < 0,20$;
- Moderadamente significante: se $0,10 < p < 0,15$;
- Significante: se $0,05 < p < 0,10$;
- Altamente significante: se $p < 0,05$.
-

No modelo em estudo, representado na Figura 3, as setas indicam as relações existentes entre as variáveis do modelo, referenciando a intensidade das mesmas. Estão também representados os valores da variância explicada – R^2 – para cada variável dependente.

Figura 3. Avaliação da Ferramenta CAF



Para que a relação entre as variáveis seja considerada consistente é necessário que o valor do seu coeficiente possua valores superiores a 0,20, valor adoptado por outros investigadores na área da qualidade (Lourenço e Fernandes, 2006 e Pêgo, 2002).

Verifica-se que, à excepção de três casos, todas as relações apresentam valores superiores ao valor de referência. As relações que possuem um valor do coeficiente abaixo de 0,20 são as relações entre a Liderança e as Parcerias e Recursos, entre os Resultados Relativos às Pessoas e os Resultados de Desempenho-Chave e entre os Resultados Orientados para os Cidadãos/Clientes e os Resultados de Desempenho-Chave.

As relações consideradas com maior intensidade são as relações estabelecidas entre as variáveis Impacto na Sociedade e os Resultados de Desempenho-Chave, com um coeficiente igual a 0,816, e entre a Liderança e o Planeamento e Estratégia, com um coeficiente de 0,842.

Da observação dos valores dos coeficientes de determinação para cada variável dependente verifica-se a existência de valores bastante elevados, sendo o valor mais baixo de 0,547 que corresponde a variável Impacto na Sociedade. Pode assim concluir-se que o modelo é adequado para os dados existentes, já que todos os coeficientes possuem valores superiores ao valor de referência.

O modelo revisto permite concluir que apenas duas das hipóteses não foram consideradas altamente significantes, ou seja:

- A relação entre os Resultados Relativos às Pessoas e os Resultados de Desempenho-Chave é apenas significante;
- A relação entre os Resultados Orientados para os Cidadãos/Clientes e os Resultados de Desempenho-Chave é não significante.

Considerando os resultados apresentados, o modelo CAF faz todo o sentido não obstante o resultado da hipótese 4.2. Os resultados demonstram que existem relações entre os diferentes critérios que compõe a ferramenta CAF e, por conseguinte, a investigação suporta a consistência interna do modelo apesar de, estranhamente, não se ter confirmado a relação entre os Resultados Orientados para os Cidadãos/Clientes e o Resultados de Desempenho-Chave. O valor negativo obtido pode justificar-se pelo facto da avaliação ser efectuada através da percepção dos colaboradores e não pela avaliação dos clientes externos.

6. CONSIDERAÇÕES FINAIS

Neste estudo formularam-se hipóteses de investigação que reflectem as relações entre os critérios da ferramenta CAF. Conclui-se que os dados empíricos obtidos suportam a grande maioria das hipóteses equacionadas. Num universo de 16 hipóteses, apenas 1 não é suportada, pelo que é legítimo afirmar que os dados recolhidos na Câmara Municipal do Fundão dão suporte às inter-relações entre critérios que estão subjacentes ao modelo CAF. É de referir que qualquer extrapolação das conclusões tem de ser feita com precaução, dadas as características da amostra utilizada.

Os resultados obtidos na avaliação do modelo permitem considerar a CAF como uma importante ferramenta de auto-avaliação da realidade organizacional, propiciando a implementação, avaliação e desenvolvimento de um sistema de qualidade em qualquer entidade pública. Inerente à natureza do modelo está subjacente a ideia de que um bom desempenho nos meios conduzirá a bons resultados.

Observando-se os coeficientes das regressões, e face ao tratamento e análise da informação recolhida, deduzem-se algumas considerações pertinentes:

- As relações de influência entre a Liderança e os critérios Gestão das Pessoas, Planeamento e Estratégia e Parcerias e Recursos foram validadas e consideradas altamente significantes. Assim, pode-se dizer que a organização deverá considerar os líderes como os principais responsáveis pela implementação de práticas de Gestão pela Qualidade Total, com recurso à ferramenta CAF, devendo estes ter em atenção as características e competências individuais e de grupo, os objectivos organizacionais e os recursos e parcerias de que dispõem e de que necessitam para garantir o sucesso da execução de um Plano de Melhorias;
- Verificou-se que os critérios Gestão das Pessoas, Planeamento e Estratégia e Parcerias e Recursos influenciam a Gestão dos Processos e da Mudança. É importante salientar que estas variáveis possuem níveis de influência semelhantes sobre a Gestão dos Processos e da Mudança, sendo a Gestão das Pessoas aquela que possui um nível mais baixo, mas considerado significativo;
- Relativamente à Gestão dos Processos e da Mudança, verificou-se que existe uma forte relação entre esta variável e os Resultados Orientados para os

Cidadãos/Clientes. As relações directas entre a Gestão dos Processos e da Mudança e os Resultados Relativos às Pessoas e o Impacto na Sociedade são altamente significantes, no entanto não são tão fortes como a primeira;

- Finalmente, é de salientar a forte relação directa entre o Impacto na Sociedade e os Resultados de Desempenho-Chave. A relação existente entre os Resultados Relativos às Pessoas e os Resultados de Desempenho-Chave, apesar de significativa, é fraca. A relação directa entre os Resultados Orientados para os Cidadãos/Clientes e os Resultados de Desempenho-Chave não é significativa.

No contexto actual, e no seio da Administração Local em particular, a palavra de ordem é a mudança. Assim, o modelo CAF, enquanto instrumento de cooperação europeu e ferramenta de Gestão pela Qualidade Total, incita as organizações públicas a fornecer melhores serviços, sob a forma de prestações cada vez mais homogéneas. A Administração Local deve redefinir os seus modelos de gestão e de interacção com o seu cliente – o Cidadão. As autarquias devem criar e formular os seus circuitos procedimentais de forma simplificada e transversal aos seus serviços internos, estendendo-os, se possível, a todas as entidades externas, públicas e privadas, com as quais se relaciona.

BIBLIOGRAFIA

ANDRADE, ANA (2007), "CAF 2006 – Melhorar as organizações através da auto-avaliação", Qualidade, n.º 1, 2007, pp. 28-33.

APPLEBY, ALEX E CLARK, ANTHONY (1997), "Quality Management in Local Government, the same in the Private Sector, but Different", Leadership and Organization Development Journal, Vol. 18, n.º 1.

BERMAN, EVAN M. E WEST, JONATHAN P. (1995). "Municipal Commitment to Total Quality Management: A Survey of Recent Progress", Public Administration Review, Vol. 55, n.º 1, pp. 57-66.

BILHIM, J. (2001), "Gerir a Administração Pública como uma empresa", in Reforma do Estado e Administração Pública Gestionária, Instituto Superior de Ciências Sociais e Políticas da Universidade Técnica de Lisboa, Lisboa.

CARAPETO, CARLOS E FONSECA, FÁTIMA (2003), "A Ferramenta CAF", Gestão Pura, n.º 3, pp. 120-127.

CARAPETO, CARLOS E FONSECA, FÁTIMA (2005), "Administração Pública – Modernização, Qualidade e Inovação", Edições Sílabo, Lisboa.

CORTADA, JAMES E QUINTELLA, HEITOR (1993), "TQM Gerência da Qualidade Total", Editora McGraw-Hill.

DIRECÇÃO-GERAL DA ADMINISTRAÇÃO E DO EMPREGO PÚBLICO (2007). "Estrutura Comum de Avaliação: Melhorar as organizações públicas através da auto-avaliação", http://www.dgap.gov.pt/10_CAF_2007/01_CAF2006/ModeloCAF2006_edicao_portuguesa_27abril.pdf.Consulta efectuada em Março de 2007.

FERNANDES, A. J. (1994), "Métodos e Regras para a Elaboração de Trabalhos Académicos e Científicos", Colecção Educação, Porto Editora, Porto.

FERNANDES, ANTÓNIO A. C. M. (2006), "O Modelo da EFQM na Melhoria da Qualidade: O Estudo das Relações entre os Critérios do Modelo no Hospital Amato Lusitano". Dissertação de Mestrado, Universidade da Beira Interior, Covilhã

GADDEN, E. N. (1972), "A History of Public Administration", Frank Cass and Comp, Londres.

GUIMARÃES, RUI E CABRAL JOSÉ (1997), "Estatística", McGraw-Hill, Lisboa.

JAYARAM, J. E AHIRE, S. (1998), "Impact of Operations Management Practices on Quality and Time-Based Performance", International Journal of Quality & Reliability Management, Vol. 15.

LAKATOS, E. M. ET AL. (1996), "La metodología de los programas de investigación científica", Alianza Editorial, Madrid.

LOURENÇO, L. A. N. (2000), "Modelos de Excelência na Implementação da Gestão pela Qualidade Total", II Seminário Luso-Espanhol de Economia Empresarial, Covilhã.

LOURENÇO, L. A. N E FERNANDES A.(2007), "O Modelo de Auto-Avaliação da EFQM no Processo de Melhoria da Qualidade: o Hospital Amato Lusitano", II Seminário de Gestão em Saúde, Covilhã.

NETER, JOHN ET AL. (1996), "Applied Linear Statistical Models", 4ª ed., McGraw-Hill, Chicago.

PÊGO, VÍTOR N. A. (2002), "O Impacto das práticas da qualidade no desempenho e competitividade das empresas industriais dos distritos da Guarda e Castelo Branco". Dissertação de Mestrado, Universidade da Beira Interior, Covilhã.

PESTANA, MARIA H. E GAGEIRO, JOÃO N. (1998), "Análise de dados para ciências sociais: a complementaridade do SPSS", 4ª ed., Edições Sílabo, Lisboa.

ROCHA, J. A. OLIVEIRA (2006), "Gestão da Qualidade. Aplicação aos Serviços Públicos", Escolar Editora, Lisboa.

ROWLAND, D., ARKKELIN, D. E CRISLER, L. (1991) "Computer-based data analysis: using SPSS in the social and Behavioral Sciences", Nelson-Hall, Chicago.

SÁ, PATRÍCIA MOURA (2001), "Gestão da Qualidade nos Municípios Portugueses". Tese de Doutoramento, Sheffield..

SUN, H. (2000), "A Comparison of Quality Management Practices in Shanghai and Norwegian Manufacturing Companies", International Journal of Quality & Reliability Management, Vol. 17.

TENEDÓRIO, J. A. (2003), "Municípios, Ordenamento do Território e Sistemas de Informação Geográfica", Revista GeolNova, 7 pp. 201-220.

TERZIOVSKI, M. E SAMSON, D. (1999), "The Link Between Total Quality Management Practice and Organizational Performance", International Journal of Quality & Reliability Management, Vol. 16.

EL SECTOR PÚBLICO EMPRESARIAL VALENCIANO Y SU EXPANSIÓN ACTUAL

MARTÍN SEVILLA JIMÉNEZ
E-mail: Martin.sevilla@ua.es
TERESA TORREGROSA MARTÍ
E-mail: Teresa.torregrosa@ua.es
Dpto. Análisis Económico Aplicado
Universidad de Alicante

Resumen

La creación de las Comunidades Autónomas como forma de organización del Estado en España a partir de los años 80 del siglo pasado ha representado un cambio fundamental en la distribución del poder sobre el territorio español. Asimismo ha abierto la posibilidad de que, en virtud de la propia autonomía para determinar la forma de prestar los servicios a los ciudadanos, las Comunidades Autónomas hayan aprobado la creación una gran variedad de Empresas y organismos autónomos que, en la mayor parte de los casos no son coincidentes entre ellas ni presentan un claro criterio en su articulación y objetivos. Aunque este fenómeno es el más importante en los momentos actuales respecto a la importancia e incidencia económica territorial, también hemos de considerar como parte del Sector Público Empresarial Valenciano al dependiente de las Administraciones Locales Valencianas, Diputaciones y Ayuntamientos que, aunque de menor importancia, también forman parte del mismo.
Cuando a nivel general se ha abierto un amplio debate acerca del alcance de la intervención del Sector Público en la economía, en este apartado vamos a ver cómo, con independencia del debate teórico general, las CCAA y especialmente la valenciana han ido ampliando progresivamente no solo el peso del sector público sino el número y la importancia de sus organizaciones empresariales justificando este hecho en múltiples argumentos.

Palabras clave: comunidades autónomas, sector público empresarial, empresas públicas, organismos autónomos.
Área Temática: Economia Urbana, Regional e Local.

Abstract

The creation of the Autonomous Communities as form of the State organization in Spain from the 80's of the past Century has represented a fundamental change in the competencies and power distribution of the on the Spanish territory. It has also open the possibility of, based on the own Autonomous Community's autonomy to determine the provision of the services to the citizens, those territories have created a great variety of Companies and independent organisms that, in most of the cases are not coincident among them nor present a clear criterion on its articulation and objectives. Although this phenomenon is the most important at the present with respect to the importance and territorial economic incidence, we also have to consider as part of the Valencian Enterprise Public Sector, the Valencian Local Administrations, Delegations and City Councils that, although of smaller relevance, they are also part of the same.
When at general level an ample debate has been open about the reach of the Public Sector intervention in the economy, in this paper we are going to see how, regardless of theoretical debate, the Autonomous Communities and, especially the Valencian ones, have been progressively extending not only the weight of the public sector but the number and the importance of their enterprise organizations justifying this fact with multiple arguments.

Thematic Area: Urban, Regional and Local Economics.

1. INTRODUCCIÓN

La relación entre los ámbitos públicos y privados de la economía ha sido, es y será un tema ampliamente controvertido. Si bien durante los últimos 30 años parecía que los defensores de limitar la intervención del Estado en la economía habían triunfado, los dos últimos años han vuelto a poner sobre la mesa la importancia de la intervención del Estado en la economía en una situación de crisis. Nacionalizaciones de bancos, nuevas regulaciones de sectores económicos, subvenciones a las empresas en crisis, en definitiva, no sólo una vuelta vigorosa a la intervención estatal, sino la aparición simultánea de un nuevo paradigma económico.

Con independencia de cómo pueda evolucionar esta situación a futuro, no cabe ninguna duda que el debate entre lo privado y lo público seguirá abierto debido a que los excesos de estas intervenciones tendrán que ser contrarrestados por nuevas privatizaciones y desregulaciones que, si bien serán diferentes a las anteriores, volverán a poner en evidencia los frágiles límites entre los sectores públicos y privados.

A parte de las distintas posiciones teóricas, lo cierto es la gran importancia que representan las organizaciones del sector público en todas las economías avanzadas incluso antes de la crisis actual. Como ponen de manifiesto numerosos autores (Stiglitz, Krugman, Samuelson), no puede entenderse el funcionamiento de nuestras sociedades sin la permanente presencia del sector público en todos los ámbitos, no sólo en los económicos.

Si bien esta presencia es común a todos los países, también es cierto que cada uno de ellos tiene una forma concreta de desplegar el poder del Estado sobre el territorio. Dos grandes modelos de organización polarizan estos comportamientos: los modelos centralizados, cuyo representante más destacado sería Francia, y los modelos descentralizados a los que pertenecería España a partir de la Constitución de 1978.

La principal diferencia entre un modelo y el otro está referenciada al papel que en cada uno de ellos corresponde a los poderes de la Administración Central dependiente del Gobierno de la nación. Mientras que en el modelo centralizado la mayor parte de las responsabilidades corresponden a esos poderes, en el modelo descentralizado existe una distribución de competencias entre los órganos centrales y los autonómicos (denominados “estatales” en EEUU o “länder” en Alemania, por ejemplo) y, por lo tanto, la posibilidad de llevar a cabo políticas diferentes entre ellos.

El modelo descentralizado aplicado en España a partir de la aprobación de la Constitución de 1979 y de los sucesivos Estatutos de Autonomía de todas las CCAA españolas, ha generado la más amplia descentralización de poderes que se conoce a escala mundial, representando las CCAA un 34% del gasto público total en 2009 y motivando sus diversas regulaciones sectoriales motivos de preocupación respecto a la posible ruptura del mercado interior en España (Círculo de Empresarios, 2000).

La Comunidad Valenciana no ha sido ajena a este proceso. Desde la aprobación del Estatuto de Autonomía en 1983, la Generalitat Valenciana ha ido asumiendo competencias que anteriormente se ejercían de forma centralizada, aprobando normas propias sobre su gestión y creando organizaciones administrativas y empresariales específicas para la gestión de los servicios públicos de su competencia. Desde el punto de vista presupuestario, la Generalitat Valenciana gestiona un gasto consolidado en 2008 de 16.645 millones de €, de los que 14.286 millones de € corresponden a la Administración de la Generalitat, 740 millones de € a las 8 Entidades Autónomas y 2.882 millones de € a las Empresas públicas (17 Sociedades Mercantiles y 19 entidades de derecho público). Este gasto representa un 20% del PIB de la Comunidad Valenciana.

Aunque de menor importancia relativa, Diputaciones Provinciales y Ayuntamientos también poseen un importante papel en el sector público valenciano. Las primeras tienen un gasto consolidado de 922 millones de €, mientras que los segundos alcanzan los 4.808 millones de €, los dos en el ejercicio presupuestario de 2008

2. EL SECTOR PÚBLICO DE LA COMUNIDAD VALENCIANA.

El Sector Público en la Comunidad Valenciana está configurado principalmente por las Administraciones Públicas valencianas y los organismos y empresas dependientes de las mismas. Aunque también podría incluirse en este apartado la actividad de la Administración Central en el territorio, su dependencia de una Administración distinta y su configuración con otros criterios de organización y actuación, hace que no la consideremos en este apartado, aunque debamos resaltar su importancia económica en el territorio. A este respecto, cuestiones como el desarrollo de las infraestructuras del transporte (AENA respecto a los aeropuertos, ADIF respecto al ferrocarril, la Dirección General de Carreteras, etc.), los relacionados con el Medio Ambiente (Confederaciones Hidrográficas, Empresas Públicas de Agua, etc.) o con la Agricultura (SEIASA), entre otras, suponen una importancia clave en muchos aspectos de la actuación pública en materia económica en la Comunidad Valenciana.

En el Cuadro siguiente hemos recogido los datos presupuestarios de los tres principales ámbitos de las Administraciones Públicas valencianas. En el mismo se puede ver la importancia relativa que tiene la actuación presupuestaria de la Generalitat Valenciana respecto a las 3 Diputaciones y los 545 Ayuntamientos que componen la Comunidad.

Presupuestos Iniciales de Gasto. Administraciones Públicas Valencianas 2008

Presupuestos Iniciales de Gasto. Administraciones Públicas Valencianas 2008	GV Consolidado	Diputaciones CV	Ayuntamientos CV	Total CV
Gastos de personal	5.375.760.200	183.196.576	1.606.179.885	7.165.136.661
Gastos en bienes ctes. y	3.076.004.310	142.100.391	1.532.885.550	4.750.990.251
Gastos financieros	556.978.870	34.949.302	138.418.966	730.347.138
Transferencias corrientes	3.966.291.160	232.972.544	348.378.168	4.547.641.872
Inversiones reales	1.959.176.210	208.575.380	905.307.139	3.073.058.730
Transferencias de Capital	797.609.940	68.016.041	28.758.510	894.384.491
Activos financieros	95.723.030	3.783.995	13.554.444	113.061.469
Pasivos financieros	204.210.830	48.906.535	234.869.758	487.987.123
Total	16.031.754.550	922.500.764	4.808.352.420	21.762.607.734

Fuente: Ministerio de Economía y Hacienda y Generalitat Valenciana. Debe tenerse en cuenta que los Presupuestos de las Diputaciones Provinciales y los Ayuntamientos no están consolidados, por lo que sus cifras respectivas son algo superiores a las recogidas en el Cuadro.

Para el año 2008, los Presupuestos consolidados de la Generalitat Valenciana representan cerca de las tres cuartas partes del total del sector público valenciano.

Estos datos están en consonancia con lo comentado anteriormente acerca de la importancia de la gestión de las CCAA en España, al haber asumido las mismas una amplia gama de competencias que representan un elevado nivel de gasto público, especialmente los referidos a la gestión de la Educación y a la Sanidad. En contraste, las Diputaciones Provinciales no alcanzan el 5% del total del gasto total, al estar limitadas a la gestión de aspectos residuales de competencia pública provincial así como al desarrollo de los Planes Provinciales de cooperación de la Administración Central con los Ayuntamientos, de una cuantía muy inferior a la que pueda desarrollar la Generalitat.

Los Ayuntamientos gestionan una cifra cercana al 20% del total del gasto que, si bien es importante en volumen, representa una alta dispersión debido al elevado número de Ayuntamientos de la Comunidad Valenciana.

La estructura del gasto también permite distinguir el comportamiento entre unas y otras Administraciones. Así, mientras que la correspondiente a los Ayuntamientos se centra especialmente en el mantenimiento de los servicios públicos municipales finalistas, representando los gastos de personal y de funcionamiento las tres cuartas partes del total, en el caso de la Generalitat Valenciana, aparte de tener un elevado porcentaje de gastos de personal y

funcionamiento, tienen también una elevada cifra de transferencias corrientes y de capital debido a la modalidad de gestión que ha adoptado a través de empresas públicas.

Las únicas cifras en las que superan los Ayuntamientos a la Generalitat son las relativas a los Pasivos Financieros a pesar del conocido grado de endeudamiento de la Generalitat Valenciana. Ello es debido fundamentalmente a que los Ayuntamientos arrastran toda la deuda histórica de su actividad, mientras que la deuda de la Generalitat es relativamente reciente, por lo que aún no ha comenzado a amortizar el principal de sus créditos. Las cifras de Gastos Financieros corroboran este hecho, donde la Generalitat sí tiene que afrontar un elevado gasto en intereses de ese endeudamiento, siendo el de los Ayuntamientos mucho menor. A este respecto también hemos de tener en cuenta que, debido a la forma de contabilizar la Deuda Pública, una parte de la misma, si bien es contraída por algunos Entes Públicos, el pago de sus intereses los asume la Generalitat Valenciana, saldando los mismos como gastos financieros de la propia Generalitat o incluso a través de transferencias a dichos entes para que los mismos los realicen.[1]

En definitiva, los datos sobre el sector público valenciano ponen en evidencia la importancia que ha alcanzado la Generalitat Valenciana en la actividad económica de nuestra Comunidad siendo el agente más importante con diferencia.

3. EL SECTOR PÚBLICO EMPRESARIAL EN LA COMUNIDAD VALENCIANA: LA GENERALITAT

3.1. Introducción: De las formas de organizar la intervención pública

Mientras que la organización de las empresas privadas goza de de una amplia cobertura científica respecto a las razones que llevan a unas u otras formas de organización para buscar los máximos niveles de eficiencia privados (Coase, 1994), no podemos decir lo mismo en el caso de la organización de las Administraciones Públicas.

Preguntas como ¿Cuál es la estructura más eficiente para la prestación de un servicio público?, suelen llevar aparejadas una gran cantidad de respuestas y, en la mayor parte de los casos, los motivos que se pueden dar, están lejos de representar opiniones consistentes sobre las mismas. Las razones que se han argumentado para esta situación han insistido especialmente en dos aspectos básicos de las organizaciones públicas: su dependencia de los aspectos formales y jurídicos y el enorme peso que tienen las decisiones políticas en este ámbito.

Respecto al primer aspecto, se ha relacionado la organización de la intervención pública con lo que las leyes posibilitan a las Administraciones Públicas respecto a su propia Administración o autoorganización. En realidad, esto no es más que un primer paso del problema que estamos tratando ya que los preceptos sobre las posibilidades y alternativas que se le presentan a los gestores públicos, nada dicen acerca de la bondad de unas u otras formas de organización.

Así podemos ver que el Artículo 65 de la Ley 5/1983, de 30 de diciembre, de Gobierno Valenciano determina que "La Administración Pública de la Generalitat Valenciana se organiza y actúa con personalidad jurídica única, conforme a criterios de eficacia, publicidad, jerarquía, descentralización, desconcentración y coordinación, con sometimiento pleno a la Ley y a1 Derecho, dentro de la mayor economía de medios que permita la obtención de los fines que tiene encomendados". Como se ve, bajo este precepto se pueden amparar las más diversas formas de organización administrativa. (La Ley de la Generalidad Valenciana 4/1984, de 13 de junio, de Hacienda Pública de la Generalidad, también va en el mismo sentido)

Con respecto al segundo aspecto, la variante política tampoco facilita mucho las posibilidades de análisis de las organizaciones públicas para la prestación de los servicios. Si normalmente se ha atribuido a los partidos políticos de izquierda una mayor tendencia hacia la expansión del sector público así como a la creación de empresas y organismos para llevar a cabo la gestión pública, la evidencia reciente sobre este expansionismo no parece que se le pueda

[1] Un ejemplo de esta situación es el referido a las universidades públicas valencianas que se endeudaron para realizar sus Planes de Inversiones con el compromiso de que la Generalitat Valenciana fuera transfiriendo anualmente los recursos para hacer frente a su amortización y pago de intereses.

atribuir a ninguna tendencia política como lo pone en evidencia el comportamiento de las CCAA españolas (Gómez, M. 2000).

Algunos autores han tratado de enumerar los factores que hacen tan difícil el lograr algún patrón de comportamiento en las organizaciones públicas (aplicable también para las organizaciones territoriales) en contraste con las privadas, tratando de ver los factores que las condicionan y que son ciertamente amplios (Ramió, 1999, pg.37).

Como se ve, es en este panorama en el que se desenvuelven en la práctica las Administraciones Públicas y por ello se hace tan difícil en formalizar algún modelo que nos permita atribuir a una u otra forma de organización pública su superioridad sobre el resto.

Las CCAA y en concreto la Generalitat Valenciana se han enfrentado a esta problemática en una situación de competencias crecientes, por lo que no es extraño que estas indefiniciones hayan provocado una gran profusión de entes públicos descentralizados en su gestión[2].

3.2. La situación de las CCAA respecto al Sector Público Empresarial

La España de las Autonomías ha supuesto un profundo cambio en el panorama institucional. De un Estado centralista e intervencionista se ha pasado a otro que, si bien ha dejado de ser centralista, ha vuelto a recuperar las formas del intervencionismo de una manera descentralizada. No es extraño lo que ha sucedido teniendo en cuenta como se ha producido la "devolución" de las competencias a las CCAA.

Las dos primeras Comunidades Autónomas en ponerse en marcha fueron el País Vasco y Cataluña. Su interpretación de la autonomía no ha sido nunca el de entenderla como un mero acto de descentralización, sino que lo han interpretado como una recuperación de derechos de autogobierno "nacional". Se hacía evidente que eso conllevaba la capacidad de las mismas para poner en marcha una gran variedad de actividades, no solamente las transferidas por la Administración Central. Eso posibilitó que desde sus inicios, el Sector Público Empresarial del País Vasco y Cataluña fuera muy vigoroso, auspiciado por sus Gobiernos nacionalistas.

El resto de las CCAA, de una forma tímida al principio y creciente posteriormente, también se sumaron a esta dinámica, obteniendo como resultado de forma agregada el que hemos visto hasta los momentos actuales.

No podemos decir que este crecimiento haya tenido formas coincidentes ni haya afectado a todos los sectores por igual, pero si que ha sido un crecimiento común con independencia de los responsables políticos que hayan desempeñado el Gobierno de cada una de ellas.

Este proceso ha venido a coincidir con la desinversión y privatización de empresas públicas de responsabilidad estatal. Si bien las cifras presupuestarias de ambos procesos no son equiparables (las privatizaciones de empresas públicas estatales han sido muy superiores a las cifras totales de los presupuestos de las empresas autonómicas creadas), la evolución del número de empresas en una y otras Administraciones, dan un resultado creciente y superior a las empresas autonómicas

La importancia de este proceso, que lleva aparejada la salida de los Presupuestos Generales de las CCAA de estas actividades, es considerable y creciente. El propio Ministerio de Economía y Hacienda ha tratado de hacer frente al mismo a través del control de sus efectos financieros en las Leyes de Estabilidad Presupuestaria y del impulso a su conocimiento real con los denominados "Informes del Sector Público Empresarial Autonómico" y sus "Inventarios de Entes de las Comunidades Autónomas" desde 2007(MEH)

Este esfuerzo informativo ha permitido también ampliar el campo de observación del Sector Público Empresarial, incluyendo últimamente datos acerca de Fundaciones y Consorcios que anteriormente no existían, de los que podemos observar su importancia creciente en estos últimos años.

[2] A pesar del esfuerzo que se está haciendo por parte del MEH dentro del Consejo de Política Fiscal y Financiera, aún no hay una información completa sobre el sector ni coincidencia en lo que se debe considerar incluido en el mismo. Ver Consejo de Política Fiscal y Financiera de 10 de abril de 2003

Entes Públicos descentralizados y su evolución 2002-2006

Tipo de Ente	2002	2003	2004	2005	2006
Empresas	***1.530***	***1.769***	***1.957***	***2.135***	***2.255***
De ámbito Estatal	284	263	264	271	283
De ámbito Autonómico	542	588	616	719	787
De ámbito Local	704	918	1.077	1.145	1.185
Fundaciones	***106***	***163***	***262***	***348***	***491***
De ámbito Estatal	51	53	53	55	55
De ámbito Autonómico	47	101	134	183	316
De ámbito Local	8	9	75	110	120
Consorcios		***435***	***449***	***752***	***919***
De ámbito Estatal		-	1	8	10
De ámbito Autonómico		396	409	500	586
De ámbito Local		39	39	244	323
Total	***1.636***	***2.367***	***2.668***	***3.235***	***3.665***

Fuente: IGAE Informe 2006 pg.25

La actividad del Sector Público Empresarial Autonómico es muy diversa y difícilmente homologable en cada uno de los distintos territorios. Como se ha comentado en apartados anteriores, la creación de las empresas responde a diversos motivos que van desde la huida del derecho administrativo, la creación de INIs regionales, el desempeño de servicios públicos considerados preferentes por la CCAA o bien, para la gestión especializada de determinados servicios (Cuadrado Roura,J.R. y Carrillo, M. 2008).

La dinámica de estos procesos no responde a unas pautas muy concretas respecto a la mejora de los servicios públicos en general (salvo en contadas ocasiones como los relacionados con la gestión del agua o la depuración) sino que más bien se vincula con determinados impulsos de sus responsables públicos que destacan sobretodo su propia capacidad de autoorganización.

3.3. Las razones de la creación de organizaciones públicas descentralizadas por la Generalitat valenciana.

Ya hemos comentado anteriormente las dificultades que existen desde el punto de vista teórico para encontrar cuáles son las razones que están detrás de la creación de empresas u otros organismos por parte de las Administraciones Públicas. Esta evidencia afecta tanto a las Administraciones Centrales como a las Territoriales. Algunos autores (Cuadrado Roura,J.R. y Carrillo, M. 2008) han tratado de ver cuáles son las pautas que durante los últimos años han estado, en la práctica, detrás de esta creación de organismos y empresas públicas autonómicas (también incluye a los Ayuntamientos, aunque en este caso no está tan claro el argumento, al ser instituciones más antiguas). Entre las mismas destacan las siguientes: Para huir del control en la aplicación del derecho administrativo, más rígido que otros procedimientos legales; por precisar de una organización específica con un marco jurídico apropiado para la prestación de un servicio; por motivos asociados a las posibilidades de ampliación del endeudamiento; para absorber empresas con problemas, dando lugar a lo que se ha denominado "INIs" regionales o, por último, para la creación de empresas de especial interés para la región o el municipio.

En el caso de la Generalitat Valenciana, todas esas razones han estado tras las decisiones de crear entes y empresas públicas desde la aprobación y entrada en vigor del Estatuto de Autonomía en1983.

El objetivo de la primera etapa, entendiendo ésta entre 1983 y 1993, fue en primer lugar el de crear estas instituciones y empresas como soporte al desarrollo de los objetivos de la GV, especialmente en los aspectos relacionados con la Política Económica y la gestión cultural. De este periodo son el IMPIVA (1984) y la red de Institutos Tecnológicos (que adoptaron forma de organizaciones asociativas privadas) destinados a la dinamización del tejido industrial; Valencia Parque Tecnológico S.A. cuya finalidad era la creación de un parque tecnológico en Paterna.; el Instituto Valenciano de la Energía, para potenciar el uso de energías alternativas y que tuvo una importancia clave en el proceso de gasificación y cogeneración de la industria cerámica de Castellón; el Instituto Valenciano de Turismo, cuya finalidad era la promoción e innovación en el turismo valenciano; Procova (Promociones comerciales Valencianas S.A.) destinado a la promoción exterior, Sepiva, para la gestión de las inspecciones industriales y el desarrollo del suelo industrial. Respecto a la gestión cultural, se crearon el Instituto Valenciano de Artes Escénicas (IVAEM) y el Instituto Valenciano de Arte Moderno (IVAM) para la gestión del Museo de Arte Moderno de Valencia.

En segundo lugar se trató de ir creando organismos públicos con distinta forma jurídica que fueran capaces de gestionar los servicios que en aquellos momentos se estaban transfiriendo desde la Administración Central, como el Instituto Valenciano de la Vivienda (IVVSA) o Ferrocarrils de la Generalitat.

También de este periodo son los organismos que se crean para gestionar aspectos institucionales de la propia Generalitat, como el IVE o el CES.

La crisis económica que se inició en el año 1991 tuvo su impacto en esta dinámica, procediéndose en el periodo 1993-1995 a fusionar algunas de las empresas y organismos existentes, con el fin de reducir sus gastos de funcionamiento, una vez cumplidos sus cometidos iniciales. Una de las integraciones más importante fue la de la integración del IVEN y Valencia Parque Tecnológico en el IMPIVA.

Durante este periodo tuvo lugar también la creación en 1992 del Ente Público de Saneamiento de Aguas Residuales (EPSAR) que, a diferencia de los anteriores se creó asociado a un ingreso público: el canon de saneamiento, que funciona como un impuesto destinado al saneamiento de las aguas urbanas de la Comunidad Valenciana.

El cambio de gobierno en 1995 supuso también una modificación de criterios en la creación de empresas y organismos públicos. En un primer momento esta modificación se dirigió a una alteración de algunos nombres (Procova pasó a ser el IVEX; el ITVA fue cambiado a Agencia Valenciana de Turismo, con los mismos cometidos anteriores; la Sociedad Valenciana de Aprovechamiento Energéticos de Residuos S.A., pasó a denominarse VAERSA Sociedad Valenciana de Aprovechamiento Energéticos de Residuos S.A; Valencia, Ciencia y Comunicaciones, S.A. (VACICO), pasó a denominarse Ciudad de las Artes y de las Ciencias) y a iniciar un proceso de privatizaciones que, en la práctica, solo se aplicó a las instalaciones de ITV que estaban asignadas a SEPIVA.

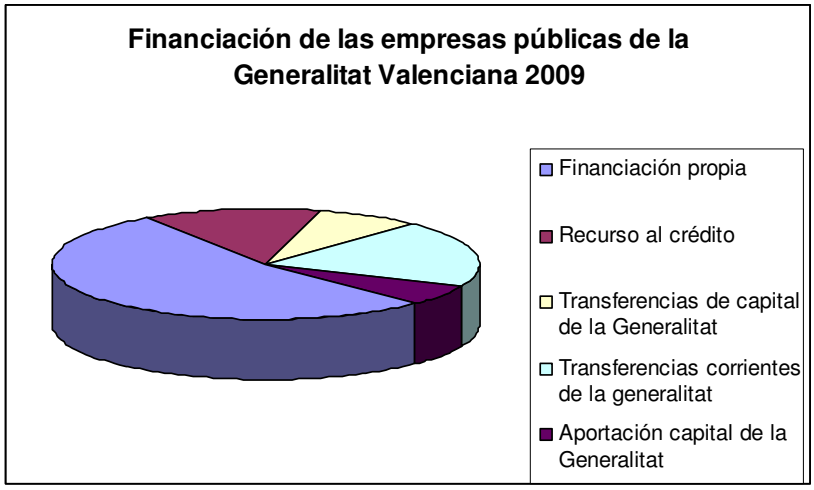
Desde 1995 a 2002 (periodo de la Presidencia de Eduardo Zaplana), el panorama de creación de empresas por parte de la Generalitat cambia notablemente, combinándose un discurso proclive al papel de la iniciativa privada con un afán intervencionista notable. Este comportamiento se ha continuado y reforzado en los Gobiernos siguientes hasta 2009, bajo las Presidencias de José Luis Olivas (2002-2003) y Francisco Camps (2003-…..), ampliando no sólo la cantidad de empresas públicas sino también con una profusa utilización de las Fundaciones dependientes de la Generalitat Valenciana.

3.4. Los sectores de intervención

En los momentos actuales, según la Ley de Presupuestos de la Generalitat Valenciana para 2009, existen en esta fecha 8 "entidades autónomas" con un gasto previsto de 740 millones de €; 18 "sociedades mercantiles" y 19 "entidades de derecho público" que juntas y bajo la denominación de "empresas públicas" suman 2.882 millones de €, un 22% del total del Presupuesto.

Según los Presupuestos de la Generalitat para 2009 las 45 Entidades Autónomas y empresas de la Generalitat se financiaban en un 52% con recursos propios, un 13 % con recursos

procedentes de créditos, un 8% con transferencias de capital, un 19 % con transferencias corrientes y un 6 % con aportaciones de capital procedente de la Generalitat



Fuente: Ley de Presupuestos de la Generalitat Valenciana para 2009

De la relación de estas empresas destacan algunas tanto por el volumen presupuestario que gestionan como por la apelación que las mismas hacen al crédito.

Millones de €	2009	
	Presupuesto de Gastos	Recurso al Crédito
Entidades Autónomas		
Servicio Valenciano de empleo y formación	473	
Agencia Valenciana de Fomento y Garantía Agraria	205	
Sociedades Mercantiles		
Televisión Autonómica Valenciana S.A.	207	
Instituto Valenciano de la Vivienda S.A.	239	8
VAERSA	132	8
Ciudad de las Artes y las Ciencias S.A.	119	
SEPIVA	57	17
Sociedad Proyectos Temáticos de la Comunidad Valenciana S.A.	146	37
CIEGSA	219	60
Ciudad de la Luz S.A.	42	
Aeropuerto de Castellón S.A.	103	
Entidades de Derecho Público		
Ferrocarrils de la G.V.	280	27
IMPIVA	122	
IVF	293	235
Agencia Valenciana de Turismo	72	
EPSAR	484	101
Agencia Valenciana de la Energía	54	
Ente Gestor de la Red de Transportes y Puertos de la G.V.	100	
TOTAL	3347	493

Fuente: Ley de Presupuestos de la Generalitat Valenciana para 2009

En el Cuadro anterior hemos relacionado las empresas públicas con un mayor presupuesto de gastos, así como con la apelación al crédito prevista. Hemos de tener en cuenta que esta apelación al crédito son las cifras correspondientes a este ejercicio ya que, según los datos del Banco de España (BE, Boletín Estadístico), la deuda de las empresas públicas de la Generalitat Valenciana alcanzaba a finales de 2008 los 2.003 millones de €, siendo la segunda en orden de importancia de todas las CCAA, solamente por detrás de la Generalitat de Cataluña.

Como se ve por la relación, la mayor parte de las mismas están dedicadas a gestionar servicios propios de la misma de una forma descentralizada, posibilitando de esta forma adaptarse a los requerimientos de una gestión especializada (como Ferrocarrils, el IVVSA o el IVF) o a posibilitar el conocer el destino de los fondos finalistas transferidos por otras administraciones (caso del SERVEF en materia de empleo o la Agencia Valenciana de Fomento y Garantía Agraria en materia de gestión de los Fondos procedentes de la Unión Europea para ayudas a la agricultura).

La gestión a través de la EPSAR permite a su vez vincular los recursos procedentes del canon de saneamiento con las inversiones y gastos destinados a gestionar el saneamiento de las aguas residuales de la Comunidad Valenciana, permitiendo el endeudamiento ampliar y adelantar estas gestiones para cumplir con los requerimientos de la Unión Europea en esta materia.

Otras empresas tienen que ver con las preferencias de la Política de la Generalitat Valenciana en sectores no ligados esencialmente con obligaciones básicas de los servicios esenciales de la Comunidad, como son las relativas a la Televisión Autonómica Valenciana S.A.; la Ciudad de las Artes y las Ciencias S.A.; la Ciudad de la Luz S.A.; o la Sociedad Proyectos Temáticos de la Comunidad Valenciana S.A.

Por último, esta relación incluye a Construcciones e Infraestructuras Educativas de la Comunidad Valenciana S.A.(CIEGSA), creada en su momento para sacar de los Presupuestos generales de la Generalitat las necesidades financieras asociadas a la construcción de centros escolares y, por lo tanto, que no se contabilizara su deuda dentro de las estadísticas del déficit excesivo a efectos del Pacto de Estabilidad y Crecimiento de la U.E.

4. CONCLUSIONES

El Sector público Empresarial valenciano ha experimentado un fuerte crecimiento desde el surgimiento de la Generalitat Valenciana en 1983. Si bien durante los primeros años se limitó a la creación de entidades autónomas y empresas vinculadas con sectores especializados de la gestión de la Generalitat, durante los años posteriores se ha ido ampliando tanto el campo de actuación como las formas de esta intervención.

Este comportamiento se ha dado de una forma coincidente con lo actuado con otras CCAA que han tratado de esta forma de utilizar estos instrumentos como un mecanismo para tanto demostrar su autonomía como para buscar procedimientos para ampliar su capacidad de actuación.

Esta forma de actuación ha venido a coincidir también con un periodo en el que frente a las teorías dominantes de disminución del sector público y primacía del sector privado, los Gobiernos autonómicos han considerado que dichas recomendaciones no iban con ellos, generando una gran actividad e intervencionismo en la economía. A esto ha contribuido también el intento de huida del estricto control que imponen las normas administrativas en materia laboral, de contratación, de inclusión de los límites del endeudamiento, etc.

La Generalitat Valenciana ha destacado en el conjunto de las CCAA españolas por el elevado nivel de endeudamiento de su gestión, tanto ordinaria como asociada a las empresas públicas que ha creado, sin que en términos de gestión, en los momentos actuales haya tenido actuaciones destacables, salvo en contadas excepciones (el caso más notable sería la EPSAR que ha posibilitado acercarnos a los estándares europeos en materia de saneamiento de las aguas a través de una financiación adecuada a través de un canon).

El resultado de esta actuación ha llevado en los momentos actuales a una gran cantidad de empresas públicas que se han visto reforzadas en los últimos años con un enorme elenco de

Fundaciones Públicas con un heterogéneo cometido y dependientes de la Generalitat que hacen enormemente difícil realizar un seguimiento adecuado de su gestión y que dificultan la búsqueda de la eficiencia que debe siempre perseguir toda Administración Pública.

La realidad política viene a decir que mientras que es muy fácil el ir creando nuevos organismos, el proceso inverso se hace casi imposible debido incluso a los compromisos endógenos que los mismos adquieren. La Generalitat Valenciana ha confirmado y reafirmado este proceso ya que, incluso con declaraciones políticas en contra del intervencionismo público, no ha soportado la tentación de ir ampliando tanto su ámbito de actuación como el número de organismos descentralizados y empresas para llevarlo a cabo.

Lo anterior, unido a una actividad administrativa y regulatoria no siempre justificada ha hecho levantar la voz de los agentes económicos empresariales sobre la necesidad de prestar atención a las posibles rupturas del mercado interior español que, a diferencia del de otros países, aparece muy descentralizado y con normas y actuaciones muy distintas en cada territorio.

Las CCAA han alcanzado en los momentos actuales un cierto grado de madurez por lo que, en estos momentos de crisis económica y de restricciones presupuestarias, se hace necesario un proceso de reflexión para dimensionar adecuadamente a las mismas, consolidando muchos de los organismos creados y volviendo a una senda más adecuada a los motivos que originaron su puesta en marcha a partir de la Constitución Española y sus Estatutos.

5. BIBLIOGRAFÍA

BANCO DE ESPAÑA. Boletín Estadístico, varios años

CARLES RAMIÓ (1999): *Teoría de la Organización y Administración Pública*. Tecnos

COASE, R.H. (1994): *La empresa, el mercado y la Ley*. Alianza Editorial

CUADRADO ROURA, J.R y MARTA CARRILLO (2008): *Expansión del Sector Público Empresarial en las Autonomías y Ayuntamientos. Evolución y análisis de las posibles causas*. DT Universidad de Alcalá.

GENERALITAT VALENCIANA. Leyes de Presupuestos de la Generalitat Valenciana de varios años. El último Presupuesto referenciado corresponde a 2009.

GENERALITAT VALENCIANA. Sindicatura de Cuentas. Informes varios años. El último Informe referenciado corresponde al ejercicio presupuestario de 2007

GÓMEZ, M. (2000): *El Renacimiento del INI*. Círculo de Empresarios. Madrid

http://www.circulodeempresarios.org/var/forum/storage/original/application/ba51d11b66bb83c30069e7c3b61c7182.pdf

MINISTERIO DE ECONOMÍA Y HACIENDA. (2008): *Inventario de Entes dependientes de las Comunidades Autónomas*. DIRECCIÓN GENERAL DE COORDINACIÓN FINANCIERA CON LAS CC.AA. Y EE.LL.

MINISTERIO DE ECONOMÍA Y HACIENDA. Consejo de Política Fiscal y Financiera. Informes varios años.

MINISTERIO DE ECONOMÍA Y HACIENDA. IGAE. (2009): *Sector Público Empresarial y Fundacional* 2006 http://www.igae.pap.meh.es/NR/rdonlyres/01083B44-1DC7-4299-86D9-E0CF7504285C/21612/OK1SPEyF_2006sinbioconda.pdf

MONASTERIO, C. y FERNÁNDEZ, R. (2008): Coordinación en el Estado Autonómico: Una agenda pendiente. *Papeles de Economía Española* nº 118

STIGLITZ, J.A. (2002): La Economía del Sector Público. 3ª edición. Antoni Bosch Editor.

ÁREA XI

ECONOMÍA SOCIAL Y DISCAPACIDAD

ECONOMIA SOCIAL E DISCAPACIDADE

WHY ARE DISABLED INDIVIDUALS NOT WORKING IN SPAIN? A JOB SEARCH APPROACH

JUDIT VALL CASTELLO
E-mail: judit.vallcastello@governance.unimaas.nl
Maastricht Graduate School of Governance
Maastricht University

Resumen

The aim of the research is to understand the incentives and disincentives to work provided by the disability scheme in Spain with special attention to the role played by the different degrees of disability pensions in the system which, in turn, entail different working restrictions and pension amounts. The interplay between these types of pensions as well as individual characteristics, labor market and economic conditions prove to be key determinants on individual's labor market decisions. The final goal is to draw some conclusions on the type of policy initiatives that could be more effective in promoting employment among disabled individuals.
We proceed in two steps. In the first step we design a job search model to understand the incentive structure that affects the labor market behavior of disabled individuals whereas in the second step we estimate a competing risk model in order to test the intuitions obtained from the job search model.

Palabras Clave: Disability Benefits, Work Incentives, Duration Analysis.
Thematic Area: Social Economics.

1 INTRODUCTION

1.1 Recent Developments

During the last years, disability policies have attracted particular attention in OECD countries both because they represent an important source of government expenditure and because societies are becoming more and more concerned about the need to strengthen the integration of disabled individuals in the society.

Countries in the OECD have tightened the conditions to access the unemployment and social assistance schemes introducing several work requirements while, at the same time, early retirement schemes have been limited (or even abolished) because of their well-documented work disincentives. These events have resulted in increasing numbers of pre-retirement individuals entering the disability schemes which, together with the ageing process of developed societies, raises concerns about the mid-term effects of disability expenses on the government's budget (OECD, 2007a, 2007b).

On the other hand, several studies and organizations of people with disabilities have stressed the need to promote the labor market integration of disabled individuals as a way of facilitating their broader integration in the society. The pre-conception that disability completely disables the individual for any kind of job is totally unfounded and some effort should be made in order to analyze how their abilities can be matched with labor market opportunities (OECD, 2003).

For these reasons, the possibility of increasing the number of disabled people that work is regarded as a good way to decrease the pressures on the financial stability of the social security system as well as to reach the social integration of disabled individuals.

As the OECD puts it; "Helping (disabled) people to work is potentially a "win-win" policy: it helps people avoid exclusion and have higher incomes while raising the prospect of more effective labor supply and higher economic output in the long term" (OECD, 2007a).

This seems to be a reasonable policy objective specially if we take into account two current developments that will provide new opportunities; on the one hand, medical advances will open up the possibilities for disabled people to stay healthier at older ages while, on the other hand, technological innovations will facilitate the adaptation of the working environment to the special requirements of the disabled.

However, there are several entry barriers to the labor market that individuals with disabilities face, which partially explain the low employment rates of this group of people:

1. Physical barriers: adaptation of the working space.
2. Stigma: prevents disabled individuals from receiving job offers.
3. Mismatch: the job search process is more complicated as it is more difficult to receive the "adequate" job offer; employers don't know the specificities of the disabling condition nor the way in which they will affect the type of job that they offer.

Therefore, there is a need for new specific policies that take into account all the specificities explained above in order to ensure their effectiveness and the real integration of the disabled into the labor market and into society.

The promotion of employment of disabled individuals is particularly relevant for the Spanish case and, in 2007, the OECD identified it as the most "formidable" challenge facing the Spanish government with respect to disability policies mainly due to two reasons: the significantly lower employment rates of disabled individuals in Spain with respect to other OECD countries and the higher proportion of older workers in the disability scheme (see table 1). In that sense, 70% of all disabled individuals receiving disability benefits in Spain are on the age range of 50-64 years old (see figure 6).

In terms of employment rates of disabled individuals, they have remained quite low at a 30% level[1] even if GDP growth has been quite high at around 2-6% during the last ten years and the general unemployment rate has dropped from 20% to 9,2%. Unemployment rates of disabled individuals have also remained constant at 15%. This employment rate of 30% is one of the lowest in the OECD, where employment rates for disabled individuals are 45% in the UK, 40% in Australia, 50% in Luxembourg, 45% in Norway or 52% in Switzerland (see figure 5).

Spending on disability benefits in Spain stood at 1,2% of GDP in 2005 which, together with expenditures on sickness and occupational injury benefits and services, increased this amount to 2,5% of GDP. This is the same level than unemployment-related expenditures although unemployment policies have traditionally raised much more attention among researchers and policy makers than disability policies.

Therefore, the aim of this paper is to understand the incentives and disincentives to work provided by the disability scheme in Spain both from the point of view of the individual and from the point of view of the employer and to identify the characteristics of disabled individuals that maximize their probability of being employed. When studying these issues, a crucial role will be played by the different degrees of disability pensions available in the Spanish system that entail different working restrictions as well as pension amounts. The interplay between these types of pensions as well as individual characteristics, labor market and economic conditions will prove to be key determinants of individual decisions of labor market status. The final goal is to draw some conclusions on the type of policy initiatives that could be more effective in promoting employment among disabled individuals and to provide some insides on the groups of disabled individuals that should be targeted by these policies in order to maximize the cost/effectiveness efforts.

In order to do that, we first design a theoretical job search model to capture the incentives for disabled individuals to accept or reject a job offer as well as the incentives to move between types of pensions. We derive the probabilities of accepting a job offer or moving to a higher

[1] The general employment rate has increased by 7% in the last 5 years. Employment rates stand at 65% (at the OECD average). 3,5% of the Spanish population is receiving disability benefits.

type of pension, conditional on the degree of disability, and we estimate a competing risk model with Spanish administrative data in order to test the predictions of the job search model.

The paper proceeds as follows: after some definitions and a brief discussion on the implications of using different definitions of disability, section 2 reviews the main literature focusing on the combination of a disability pension and a job and section 3 describes the specificities of the Spanish system of disability benefits. Section 4 introduces the job search model that has been designed to capture the incentives that shape individual's decisions on labor market status and provides some intuitions about the solutions of the model. Section 5 explains the econometric approach used in the analysis and presents the competing risk model. Finally, the main results are summarized in section 6 and some conclusions are derived in the last part of the paper.

1.2 Are disabled individuals able to work?

In Spain, permanent contributive[2] disability pensions are defined as the economic benefits that aim at compensating the individual for losing a certain amount of wage or professional earnings when the person is affected by a **reduction or a complete loss** of his/her working ability in a way that is assumed to be permanent due to the effects of a pathologic or traumatic process derived from an illness or an accident[3].

In order to capture the different situations in which a person can be after suffering from a disabling condition, the Spanish Social Security administration differentiates between four degrees of disability that depend on the amount of working capacity that has been lost:

Permanent Partial disability for the usual job: is the kind of disability that, without reaching the level of total disability, causes a decrease to the individual of at least 33% of the standard performance for his/her usual job but the individual is still able to develop the fundamental tasks of his/her usual job or professional activity.

Permanent Total disability for the usual job (TD): is the kind of impairment that disables the individual to develop all or the fundamental tasks of his/her usual job or professional activity but the individual is still capable of developing a different job or professional activity.

Permanent absolute disability for all jobs (AD): is the kind of impairment that disables the individual for the development of any kind of job or professional activity.

Severe disability: is the situation for which, as a result of anatomic or functional loses, the individual needs the assistance of a third person to develop the most essential acts of live such as eating, moving, etc…

Therefore, the degree of disability in which the individual is classified is set in relation to the working capacity lost as the goal of the benefits is to compensate for the reduction on wages or professional earnings caused by the disabling condition. The amount of pension received varies according to the disability degree in which the individual is classified. For example, individuals in the total disability scheme receive, in general, 55% of the regulatory base (which is an average of the last salaries) because they are assumed to get some income from work that would allow them to reach a similar amount of money than before becoming disabled. On the contrary, absolute disability pensions provide a 100% of the regulatory base as the individual is considered to have lost all his/her ability to work and, thus, is unable to get any extra income from work[4].

This kind of disability schemes that acknowledge several degrees of disabling conditions and that include in their definition the possibility of working while receiving a proportion of the

[2] Please note that, for the purpose of this paper, we only focus on contributive pensions, that is, economic benefits that require having contributed to the Social Security before becoming disabled (employed or receiving unemployment benefits). The reasons are; data availability problems for non-contributory pensioners and the idea that disabled individuals that have worked before are probably in a better position to find a new job.

[3] Own translation of the definition of permanent disability given by the Spanish Social Security administration at www.seg-social.es.

[4] More details on the way to compute these quantities are given in coming sections of the paper.

total amount of benefits are labeled in the literature as partial disability schemes and, apart from Spain, can also be found in a number of other European countries such as The Netherlands, Germany or Sweden.

The situation is very different in countries like the United States of America where the disability insurance program is defined as an all-or-nothing system for which the disabled individual can only be eligible and receive full disability benefits if he is found to have suffered a complete loss of work capacity. The alternative to that entails receiving zero benefits and occurs when the individual is found as still maintaining some capacity to work. It is widely recognized[5] that this dichotomous definition of disability entails huge work disincentives because partly disabled individuals that still keep some ability to work but that have suffered from an income loss due to the disabling condition, might have incentives to apply to the disability scheme if their reduced wage does not allow them to make ends meet.

Therefore, in systems with a binary definition of disability, the prohibition to develop a job once the individual is on the disability rolls is imposed by the same definition of disability[6]. Participation in the labor market would only be possible in the case of recovery from the disabling condition which would, in turn, entail losing the receipt of disability benefits (Benitez-Silva, 2008; Yin, 2008).

The puzzle that fosters this research lies in the observation that, even if total disability pensioners in Spain are "legally" given the option to combine the disability pension with a job and even if they have received a medical examination that classifies them as "able to engage in a job", the data show that most of the disabled individuals in Spain are not working.

For this paper, our interest lies on the incentives of disabled individuals to accept/reject a job and, for this reason, we are restricting our analysis to disabled individuals in the total and absolute disability schemes. We have included absolute disability holders because, even if these individuals are not allowed/able to work, we are interested in the transitions between these two types of pensions. We don't include partial disability holders because they don't have to change from their previous job and the benefit given is paid in only one installment so that there are really no incentives/disincentives to work to analyze.

2 DISABILITY AND WORK: A SURVEY

Several arguments can be put forward to explain these low employment rates. In the context of a country with a binary definition of disability, one of the main reasons of the low employment rates that are cited in the American literature is the negative effect of disability benefits on labor force participation rates of disabled individuals. Some authors go even further and try to establish a link between the size of the disability scheme and movements of the participation rate of disabled individuals in the USA (Duggan, 2006, 2007).

Following the same argument, a number of studies focus the attention on estimating the elasticity of the level of disability benefits on the labor force participation of disabled individuals. Although some variation of results exists among papers dealing with this topic, this elasticity has usually been found to be on the range of 0,21 to 0,93 in the USA depending on the empirical strategy used (see Parsons, 1980; Haveman et al., 1991; Gruber, 1996; kreider and Riphahn, 2000).

Another important part of the literature has focused on the existence of classification errors in the disability system. These types of studies were initiated by Nagi in 1969 after some of the beneficiaries in the USA system were found to retain some of their capacity to work. This line of research continued with the work of Smith and Lilienfeld (1971) and Benitez-Silva et al.

[5] The recognition of the disincentives to work provided by the USA disability system are even acknowledged by the Social Security Advisory Board, see Yin (2008).

[6] For a complete analysis of the disincentives embedded in the USA disability insurance system see all the work of Benitez-Silva. For a simulation of the effects of changing from the current binary disability definition in the USA to a partial definition see Yin (2008).

(2004, 2006). Quite surprisingly, even if these authors use different data sets and methodologies they all reach a similar conclusion and estimate the classification error to be around 20% for the USA system. For the Spanish case, Jimenez et al. (2006) estimate the award error defined as the difference of the probability of receiving a disability benefit and the probability of deserving it. Their results show that individuals between 55-59 years old, men, self-employed or working in the agricultural sector have a higher probability of receiving benefits without deserving them.

The lack of sufficient employer accommodation has also been identified as a potential determinant of the low employment rates of disabled individuals (Burkhauser et al., 1999, 2001, 2004).

Even if the topic of work incentives for disabled individuals has captured a lot of attention among USA researchers, in Spain the majority of studies that include an analysis of the disability scheme are centered on the identification of the alternative exit routes from the labor force that exist in the Social Security system for older workers approaching retirement. For example, Blanco (2000) estimates a competing risk model to determine the hazard rates of leaving employment and going to early retirement or disability for individuals aged 49 and over. Similarly, Jimenez et al. (2007) estimate a competing risk model for employed individuals between 45-59 who become unemployed, disabled or inactive with special attention to the effect of business cycle conditions on these transitions.

We are aware of only one paper by Malo et al. (2008) that uses Spanish data and focuses on the possibilities for disabled individuals to combine the receipt of the benefits with the development of a job. Using a cross-sectional sample for 2006, they estimate logistic regressions in order to analyze the variables that affect the probability of combining a disability pension with a job. Their findings suggest that individuals with a lower degree of disability, younger, more educated and with more previous work experience are more likely to be working while receiving the benefits. Surprisingly, they also find that the level of disability benefits does not have an effect on this probability.

The present work tries to extend the analysis of Malo et al. (2008) overcoming the limitations of cross-sectional data to include disability and disability and employment spells of all individuals in our sample and taking into account, not only the incentives of the individuals when accepting or rejecting a job offer, but also the interplay between the different levels of disability pensions in the Spanish system.

3 THE SPANISH SYSTEM

3.1 Awarding and revision process

In order to access permanent disability benefits in Spain, the individual has to come from a temporary disability situation (except in the cases where he/she does not have access to temporary disability benefits[7]).

This means that the individual has gone through the prescribed treatment while in the temporary disability situation but that, after being released from medical care, he/she is still found to have serious anatomical or functional reductions, which are susceptible to objective determination and presumably permanent, that diminish or impede his/her ability to work, which make him/her eligible to receive permanent disability benefits (Jimenez et al. 2006).

The individual can also access permanent disability benefits if, without being released from medical treatment, the disabling condition remains after extinguishing the maximum number of months of temporary disability benefits (which is usually set at 18 months but can get extended to 24 months in very special circumstances).

[7] The individual may have access to permanent disability benefits but not to temporary disability benefits if he/she is self-employed or if he comes from a situation of contributing to the system but not working (receiving unemployment benefits, maternity leaves, etc…).

The possibility of recovering the ability to work at some point in the future does not impede the classification of the individual as permanently disabled if such possibility is determined to be uncertain or very long term by the medical team (Jimenez et al., 2006).

The procedure of applying for permanent disability benefits can be initiated by the provincial office of the National Institute of Social Security (NISS), by the institutions that collaborate in the process (such as hospitals), or by the individual himself (in which case, more documentation is required).

The evaluation is done by the Disabilities Evaluation Team which analyses the medical report prepared by doctors in the provincial office of the NISS and a report on the professional background of the individual. The Disabilities Evaluation Team writes a report that will be evaluated by the directors of the provincial office of the NISS in order to take a decision on the type of disability pension granted (if any), the benefit level and the date of the next planned revision to check whether there has been an improvement or an aggravation of the condition.

Apart from these established revision dates, the administration also foresees the possibility of initiating a process of revision of the impairments either by the administration or following a request from the individual for 4 reasons:

1. Aggravation
2. Improvement
3. Diagnostic error
4. Development of a job

And the outcome of a revision process can be:

1. Confirmation of the degree of disability and the pension
2. Modification of the disability degree and of the pension
3. Extinction of the disability status and of the pension.

Therefore, the administration states very clearly that the development of a job can be a strong enough reason to start a revision process. However, when defining the compatibilities or incompatibilities of each pension type, the administration points out that:

Permanent Total disability pensions are compatible with the realization of any kind of job in the same firm or a different one with the only pre-requisite that the individual has to inform the administration if he/she decides to take up a job. This type of pension is incompatible with the development of the same job in the same firm than the one before becoming disabled.

For the absolute disability case, it is stressed that the benefit will not impede the development of those activities that are compatible with the disability status of the individual and that do not represent a change in his working ability. These rights and obligations are perfectly compatible with the right of the administration to revise the disability status if deemed necessary.

In reality, this contradiction seems to play a very important role for disabled individuals considering whether to accept a job or not because, to certain extend, the possibility of receiving a revision process if the job is accepted seems to work as a threat that disincentives employment.

3.2 Eligibility, pension amount and net income

As we are dealing with contributory pensions, there are a number of rules in terms of eligibility requirements that the individual has to fulfill in order to qualify for the pension. These requirements as well as the pension amount vary depending on the source of the disability, which can be an ordinary illness, a work-unrelated accident or a work-related accident or professional illness.

In the case of an ordinary illness, eligibility to the pension requires having contributed 1/3 of the time between turning 20 years old and the appearance of the disabling condition, with a

minimum of 5 years of contributions[8]. There is no contributive requirement when the disability is caused by an accident, whether or not work-related, or a professional illness.

The total amount of the pension is obtained from multiplying a percentage, which varies depending on the type of pension, to the regulatory base, which is different depending on the source of the disability.

The regulatory base of the pension in the case of an ordinary illness is computed as the average wage of the last 8 years of work. For a work-unrelated accident, is the average annual wage over a period of 24 consecutive months chosen by the person within the last 7 years of work. For work-related accident or professional illness, it is the average wage in the last year of work.

The percentage applied to the regulatory base is 55% under total disability pensions. This is increased to 75% when the individual is older than 55 years and, due to the lack of preparation of the individual and the disadvantaged social and labor market characteristics of the region, it is estimated to be very difficult for this disabled person to find a job in an economic activity different from the one previously developed.

For absolute disability pensions, it is equal to 100% of the regulatory base whereas for severe disability is 100% plus another 50% to cover for the person that is taking care of the disabled.

These percentages can be increased by 30-50% if the source of the disability is a work-related or professional accident and the employer is found guilty of breaking the safety regulations required for that job. This extra amount has to be paid by the employer.

All permanent disability pensions are automatically converted to old-age pensions once the individual turns 65 years old.

In terms of income taxes, total disability pensions are taxable under the general income tax rules except if the pensioner resides in the Basc country, where total disability pensions are exempted from income taxes if the individual does not have any other income from wages.

Absolute disability pensions are always exempted from income taxes.

Furthermore, if the individual works while receiving the pension, there is a reduction in the employment income used to calculate the income tax that equals 2.800 euros/year if the disability level is between 33% and 65% and 6.200 if the disability level is more than 65% or if the disabled has reduced mobility.

All this legislation is summarized in table 2 and is translated into the following income amounts for disabled individuals:

Absolute Disability:

$$y_i^A(p) = p_i$$

Total Disability:

$$y_i^T(p,a,b) = \begin{cases} 0{,}55\,p_i - r\times(0{,}55\,p_i) & if\ a < 55\ \&\ b \neq 1 \\ 0{,}75\,p_i - r\times(0{,}75\,p_i) & if\ a \geq 55\ \&\ b \neq 1 \\ 0{,}55\,p_i & if\ a < 55\ \&\ b = 1 \\ 0{,}75\,p_i & if\ a \geq 55\ \&\ b = 1 \end{cases}$$

Total Disability and work:

$$y_i^{TW}(t,w,p) = 0{,}55\,p_i - c\times t_i + w_i - r\times[0{,}55\,p_i + (w_i - k(t_i)]$$

p_i = regulatory base of the pension amount that depends on the contributions in the last job (salary) and on the source of the disability[9].

[8] If the individual is younger than 31 when becoming disabled, the requirement is to have contributed for 1/3 of the time between turning 16 years old and the appearance of the disabling condition with no minimum of years required.

$$k(t_i) = \begin{cases} 2.800 \, if \; 33 \le t_i \le 65 \\ 6.200 \, if \; t_i > 65 \end{cases}$$

a = age; b = dummy variable for Basc country; r = income tax rate

t_i = degree of disability

$c \times t_i$ = cost of working for a disabled individual which we model as a linear function of his/her degree of disability.

3.3 Database and sample selection

The study will use the Continuous Sample of Working Lives ("Muestra Continua de Vidas Laborales", MCVL) which is a microeconomic data set based on administrative records provided by the Spanish Social Security Administration. It contains a random sample of 4% of all the individuals who, at some point during 2007, had contributed towards the social security system (either by working or being in an unemployment scheme) or had received a contributory pension. The random sample selected contains over one million people.

There is information available on the entire employment and pension history of the workers, including the exact duration of employment, unemployment and disability pension spells, and for each spell, several variables that describe the characteristics of the job or the unemployment/disability benefits. There is also some information on personal characteristics such as age, gender, nationality and level of education. The macroeconomic variables used to capture the economic business cycle are derived from the Spanish "Instituto Nacional de Estadistica".

We select an inflow sample of all individuals that start receiving absolute or total disability benefits between 1996 and 2007 and we follow their disability pensions and employment trajectory until 2007 or until they are censored (if they stop receiving disability pensions and they don't work). The sample selected contains 31.839 individuals from which 20.985 are men and 10.853 are women.

4 A JOB SEARCH MODEL FOR DISABLED INDIVIDUALS

4.1 Overview

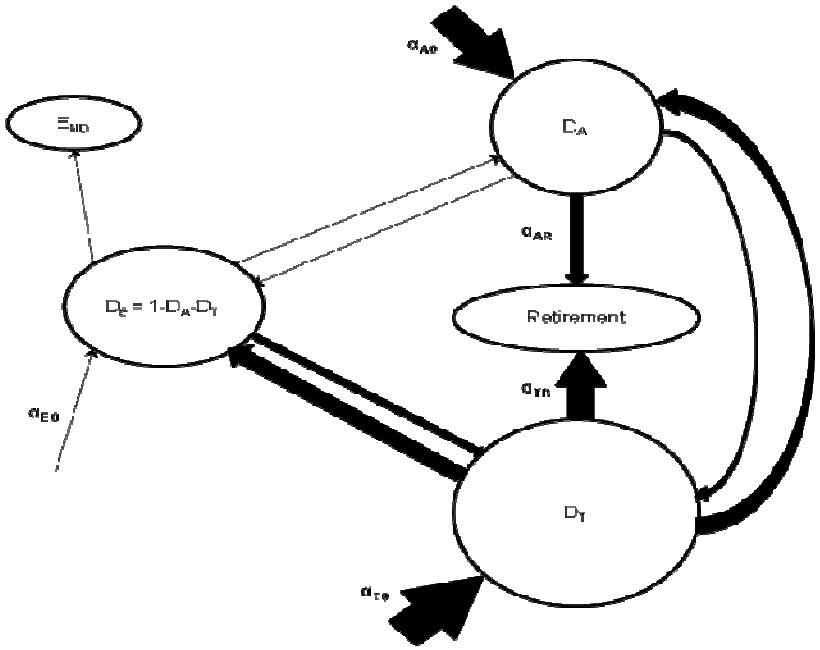
The total number of disabled individuals receiving a pension corresponds to the sum of the number of disabled individuals in absolute disability pensions, D_A, plus the number of disabled individuals receiving total disability pensions, D_T, plus the number of individuals receiving a pension and working, D_W. At the same time, W_{ND} represents exits from a situation of disability pension and working into an only employment situation.

Figure 1 below gives an overview of the possible transitions between these states. The size of the arrows is proportional to the number of individuals doing that transitions in our sample and the size of the circles is also proportional to the number of individuals in that state[10]. The majority of individuals enter the sample in D_T or D_A as: $\alpha_{T0} > \alpha_{A0} > \alpha_{W0}$

[9] We take it as a value; we don't calculate it as everybody in the sample receives always a pension and it would be very difficult to calculate as we don't know the source of the disability.

[10] The model does not include transitions from D_T or D_A to E_{ND}, as there are almost no observations in the sample doing that transition. Change the names and include arrow from Dw to retirement.

Figure 1. Disability pensioners by state and transitions between states.



4.2 Transitions

Each period, the individual starts in one of the following states: absolute disability, total disability, retired or combining total disability and work.

Retirement is a passive state that entails the receipt of a pension each period until death. All disabled individuals are automatically transferred to retirement pensions at the age of 65. This transfer is only a formality, as the amount and conditions of the pension are the same than in disability. However, once the individual is in the retirement scheme, transitions between disability degrees are not allowed anymore and the individual receives the same pension until death independently of the evolution of the degree of disability.

As the legislation specifically stresses that absolute disability pensions are given to individuals who have lost all of his/her ability to develop a job, we assume that individuals in the absolute disability scheme cannot combine the pension with a job[11].

Each individual in the sample has a degree of disability each period which is unobserved to the researcher but known by the individual and by the medical team (perfect screening)[12]. This disability degree is assumed to increase with age and to be exposed to both positive and negative random health shocks.

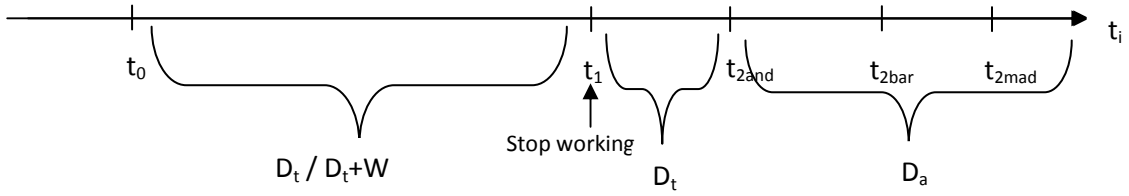
Therefore, individuals enter the sample with an initial degree of disability $t_{initial}$, which is between t_0 and t_2 if he is classified as total disability pensioner and above t_2 if he is classified as an absolute disability pensioner, and this level changes over time depending on age and a random shock.

If the individual starts the period in absolute disability pensions, there are three possible actions to take; if his degree of disability does not change and he is still below age 65, he will stay in the same state. On the contrary, if his degree of disability falls below the t_2 threshold as a result of a health shock, he will be moved to total disability pensions. Finally, when he turns 65, he will be transferred into retirement pensions but receiving the same amount of pension than before.

[11] In fact, in the sample that we have selected there are very few individuals in absolute disability that are observed as working (0,31%) and almost all of them keep the job for a very short period of time so that the assumption that individuals in absolute disability cannot work seems reasonable.

12 We assume perfect screening by the medical team of the true degree of disability of the individual so that there is no cheating in the model. This assumption is taken in order to keep the model simple. The possibility of having an imperfect screening process is left for future work.

Figure 1. Disability degree thresholds and possible states; stop working rule:



In employment, the individual receives a salary and has to pay a cost of working that depends on his/her degree of disability in a linear way (as stressed above). At the beginning of the period the individual can either continue working or stop the job to only receive the total disability pension. We do not distinguish between voluntary and forced employment terminations as it is quite difficult to defend the reliability of this variable in practice and our interest lies only on the effects of changes in the level of disability on employment status. Therefore, an employed individual will stop working (as a result of a lay off or a quit) if his degree of disability increases enough to make the cost of working too high. In the figure above this threshold is t_1.

At the age of 65, all employed disabled individuals are transferred to retirement pensions and get the same pension amount than in total disability pensions.

When the individual is receiving total disability pensions and is not working, there are four possible routes that can be taken: searching for a job, not searching for a job, moving into absolute disability benefits or retiring.

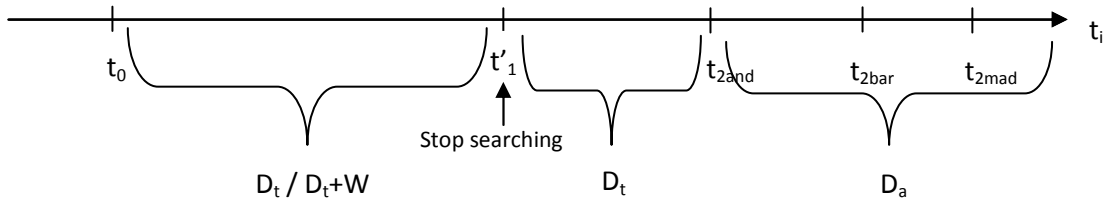
He is automatically moved to retirement when he turns 65 with the same amount of pension that he was receiving in total disability and he can only access absolute disability benefits if his degree of disability is high enough (t_2 threshold).

If he engages in active job search, he has to pay a cost of searching and has a probability λ of receiving a job offer. The arrival rate of offers and the costs of searching are fixed and independent of age, the degree of disability and other personal characteristics of the individual. If he decides not to search, he stays in total disability pensions.

There is a threshold t'_1 that determines the optimal searching rule for the individual. Below this level, the costs of working are sufficiently low to make the number of "acceptable" offers high enough to pay the search costs and engage in active job search. However, for disability levels above t'_1, the individual is better off not searching as his costs of working are too high to make searching an attractive alternative (given the distribution of offers and λ).

Therefore, there is an incentive for individuals in total disability pensions to stop looking for a job, avoiding the search costs, when their degree of disability increases above t'_1.

Figure 2. Disability degree thresholds and possible states; stop searching rule:



From figures 1 and 2 it can be observed that $t_1 > t'_1$ because when the individual is already working, he does not have to pay the search cost (as we don't include on-the-job search).

Note that, even if the move from total into absolute disability pensions and vice versa depends on the degree of disability of the individual t_i, the threshold level $t_2^*(pv_i)$ depends on the province in which he/she lives, as decisions on eligibility are taken at the provincial level. Therefore, movements from total into absolute disability pensions and vice versa will be determined by the degree of disability of the individual and the threshold level of his/her province of residence.

Figures 2, 3 and 4 below summarize the alternatives that the individual faces each period, given his/her degree of disability in that period. The absolute number of transitions is written next to the final exit destination and in brackets are percentages of initial state transitions that end in that particular final state.

For example, 94% of individuals that begin the period in absolute disability pensions are in the same type of pension in the following period whereas 1,6% move to total disability pensions and 4,7% of them move into retirement pensions as they turn 65.

From the individuals with a total disability pension, 4,5% of them are able to find a job in the following period and 87,6% are still in the same pension-only situation. Almost 3% of them move into an absolute disability pension and 4,85% go to retirement.

Figure 2. Decision tree of a disabled individual receiving absolute pension.

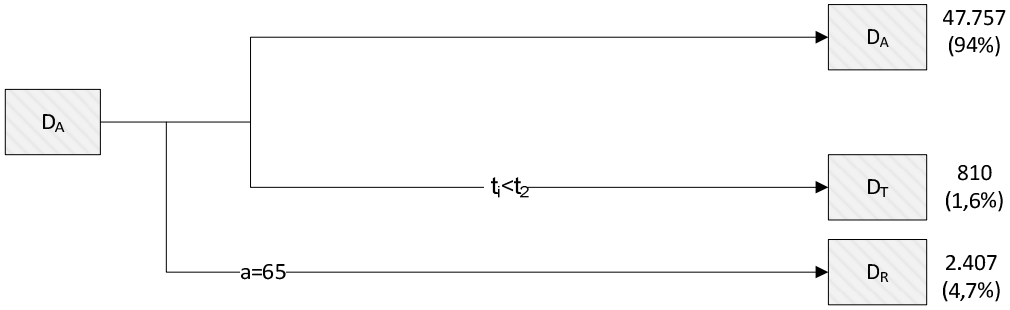
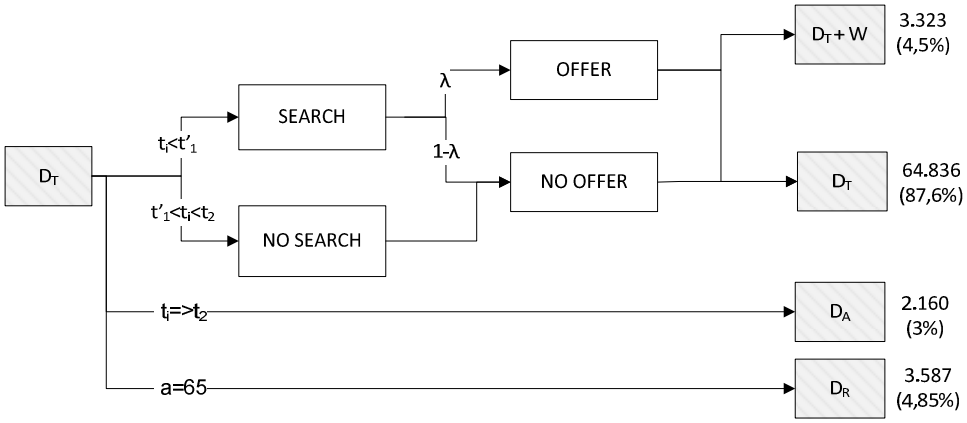
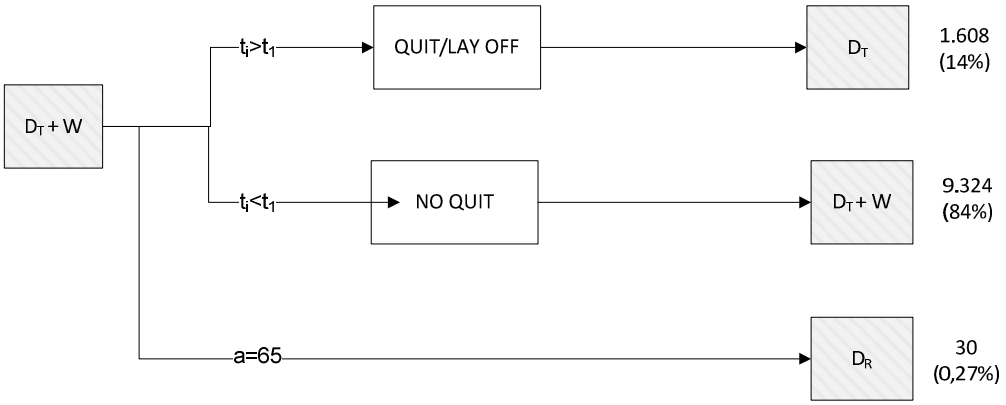


Figure 3. Decision tree of a disabled individual receiving total pension.



In the model, we rule out the possibility of completely recovering from the disabling condition so that we don't model transitions between total or absolute disability into inactivity or only employment basically because this is not the topic of this paper and also because we have observed that this type of movements only account for 0,15% of the individuals in our sample.

Figure 4. Decision tree of a disabled individual receiving total pension and working.



4.3 Value functions

The individuals in this model face three kinds of uncertainties with respect to their survival probability, the labor market environment (arrival of job offers) and health shocks. All individuals share the same probability of surviving to the next period conditioning on age.

This is a non-stationary model as age and disability are changing during the spell. The value functions are written in discrete time where each period is represented by an age, a.

Given the fact that everybody in the sample is automatically moved to the retirement scheme once they reach age 65, the value function at that age can take two values depending on whether the individual enters retirement from a total or an absolute disability situation:

$$VR_{65}^{T}(p,t_0) = \sum_{i=65}^{T} S_{65}(i)\beta^{i-65} y^{T}$$

$$VR_{65}^{A}(p,t_0) = \sum_{i=65}^{T} S_{65}(i)\beta^{i-65} y^{A}$$

From now on, we call the effective discount rate at age a, $\beta S_a(a+1)$, as β_a.

The value function of a disabled individual receiving an absolute disability pension corresponds to:

$$VA_a(p,t_0) = y^{A} + \beta_a\{\varphi_a VA_{a+1} + (1-\varphi_a)VT_{a+1}\} + VR_{65}^{A}$$

$\varphi_a =$ probability of having a health shock that is increasing with age.

Individuals receiving a total disability pension chose the maximum value between the options of searching and not searching if their degree of disability is below the threshold to move to absolute disability pensions. If they receive a health shock and their degree of disability increases above the threshold, they are moved to absolute disability pensions:

$$VT_a(p,t_0) = y_a^{T} + \beta_a\{\varphi_a VA_{a+1} + (1-\varphi_a)[Max\{VS_a(p,t_0), VT_{a+1}(p,t_0)\}]\} + VR_{65}^{T}$$

Value function of searching for a job where k is the fixed cost of searching:

$$VS_a(p,t_0) = -k + \beta_a\left\{\lambda \mathrm{E}_w\left[\max_{accept,reject}\{VTW_{a+1}, VT_{a+1}\} + (1-\lambda)VT_{a+1}\right]\right\}$$

If the individual decides to search for a job and he/she receives an offer, the reservation wage $\overline{w_a}$ defines the optimal rule for the decision of accepting/rejecting the offer and corresponds to the wage that equalizes:

$$VTW_a(\overline{w_a}(p,t_0), p, t_0) = VT_a(p,t_0)$$

Individuals receiving a total disability pensions and working have a probability φ_a, that increases with age, of having a health shock and stopping working and their value function is given by:

$$VTW_a(w, p, t_0) = y^{TW} + \beta_a [\varphi_a VT_{a+1} + (1-\varphi_a) VTW_{a+1}] + VR_{65}^T$$

The value of quitting the job corresponds to the value of only receiving total disability pensions.

We model the wage offer distribution function as:

$F(w) = \Pr(w_t \le w) = 1 - \left(\frac{w}{w_m}\right)^{-k}$ is the Pareto wage offer distribution[13] function for all $w_t > w_m$ where w_m is the (positive) minimum possible value of w and k is a positive parameter and with density function:

$$f(w) = \begin{cases} \frac{k w_m^k}{w^{k-1}} & \text{if } w > w_m \\ 0 & \text{otherwise} \end{cases}$$

Even if this distribution seems to state that salary offers below w_m are never observed, such a restriction is clearly false and a more appropriate interpretation is that salaries below this level are never acceptable given the empirically observed offer arrival rates. In this case, the Pareto distribution is capturing the tail of the offer distribution in the acceptable region[14].

5 ECONOMETRIC APPROACH

5.1 Description

For the estimation we consider two different models; one for individuals receiving only total disability pensions when entering the disability scheme and the other for individuals with total disability pensions and working when becoming disabled. We include multiple spells for both models.

In the first model of transitions from a total disability situation, there are 21.449 spells of total disability pensions (20.186 individuals in the sample). 9.158 of these spells end up in an exit either to absolute disability pensions, to total disability and working, to retirement or to only employment (no pension) and the rest of spells are censored and during our sample period they don't experience any exit from total disability pensions. From these 9.158, 2.160 exit to an absolute disability pension, 3.411 exit to total disability and working and 3.587 exit to retirement.

As exits to retirement are not very interesting as transfers from disability to retirement pensions happen automatically when the individual turns 65 years old, we have excluded the observations for individuals aged 65 or older. Therefore, our sample includes individuals from age 17 to 65 that start receiving a permanent total disability pension between 1996 and 2007 (inflow sample). We estimate a competing risk model to study transitions from a total disability situation to one of the two possible exit routes: a total pension plus work or an absolute disability pension.

The probability of applying to absolute benefits or of accepting a job for someone of a given disability degree who is currently receiving total benefits depends on a number of personal characteristics, the region where the individual lives, the pension amount received and business cycle conditions at the regional level.

Therefore, as independent variables we have included the level of education (nived), 6 duration dummies grouped in a two year interval, age of the individual, the amount of pension

[13] The Pareto distribution has been used for example in Jones (1988) and Lancaster and Chester (1983).

[14] If we want to calibrate the model, the variance parameter k and the scale parameter w_m should be fitted to correspond to the average monthly salaries of the sample period.

received[15] (the base category, penlow, corresponds to average pensions below the mean, 510 euros/month; penhigh to average pensions above the mean and below 1500 euros/month and pentop to average pensions above 1500 euros/month) .We also add dummy variables for female (Di), for each autonomous community (we leave out Madrid) and for each year included in the sample (from 1996-2007). The variable describing the education level is 1 if the individual is illiterate (does not have a school certificate), 2 if the individual finished and obtained a school degree (ESO, EGB completa, FP primer grado) and 3 if he got a high school diploma giving access to university or higher education (BUP, diplomados, licenciados, doctorados).

We have also downloaded macroeconomic data on yearly unemployment rates for each autonomous community and each gender from the database of the Spanish National Institute of Statistics and we assign, to each individual in our sample, the value of the unemployment rate corresponding to the autonomous community where he/she lives and to his/her gender in order to control for regional business cycle trends (there are 18 autonomous communities in Spain and two genders). In a similar way, we also assign to each individual a value that corresponds to the population in the autonomous community where they leave to control for the size of the population in the region.

We allow for interactions of the female dummy variable with age and with the unemployment rate and of the unemployment rate with individuals older than 55.

Table 3 contains a summary of the major reforms implemented by the Spanish central government in order to provide incentives to firms to hire disabled individuals. A number of these incentives were already in place since 1983 when the basic measures were introduced following what was ruled in 1982 by the important "Ley de Integracion Social de los Minusvalidos" law. The LISMI, as it was later called, set up the basic pillars of public policies towards disabled individuals in terms of anti-discrimination and employment promotion measures.

After the LISMI, only some small changes were introduced until 2004 when the deductions to the Social Security contributions to hire disabled women were increased from 70 to 90% for women below 45 years old and from 90 to 100% for women aged 45 or over. Apart from this reform in 2004, there was a substantial reorganization of the packet of firms' incentives in 2006 when deductions were changed from a percentage to a fixed amount of the monthly contribution that the employer has to pay to the administration for each disabled individual employed. As we don't have enough time periods to analyze the effect of the restructuration occurred in 2006, we focus our attention to the identification of the employment effects on disabled women of the reform introduced in 2004. We use two dummy variables, ditless45 and ditamore45, to identify the differential effect of the reforms on the probability of finding a job for women below or above 45 years old.

Our competing risk model does not have a constant term, as we include all duration dummies, and corrects the standard errors for clustering at the individual level because of the inclusion of multiple spells in the estimation.

For the second model, we consider individuals receiving disability pensions and working when entering the sample (becoming disabled) or later.

Again, we include multiple spells so that we have 4.776 spells of total disability plus working during our sample period that correspond to 4.331 individuals. 1.833 of these spells end up in an exit either to total disability pensions (they stop working), to absolute disability, to retirement or to only employment (no pension) and the rest of spells are censored as they don't experience any exit from total disability pensions plus working during our sample period. From these 1.833, 1.608 exit to a total disability pension, 109 exit to absolute disability, 86 to only working and 30 go to retirement.

As only the number of exits to total disability is relevant, we have decided to consider exits to all routes to be the same and we will estimate a single exit model (the competing risk model would not be identified because of few exits to absolute disability, working and retirement).

We include the same independent variables than in the previous model.

[15] This variable is lagged for one period in order to avoid including the amount of absolute disability pensions in the analysis of transitions out of the total disability state.

5.2 The competing risk model

With the assumption of independence of the destination-specific hazard rates[16], the discrete hazard rate for exit at time t to any of the two destinations is simply the sum of the destination-specific discrete hazard rates, $h_a(t)$ and $h_{tw}(t)$:

$$h(t) = h_a(t) + h_{tw}(t)$$

where

$h_a(t)$ is the hazard rate of experiencing a transition from total to absolute disability pensions

$h_{tw}(t)$ is the hazard rate of a transition to total disability and work

The discrete time hazard into one of the m states is equal to the probability of making a transition in interval t, conditional upon surviving up to the beginning of the interval.

We assume a specific form for the conditional destination-specific hazard rate h_{it}^m for individual i to destination m in interval t such that[17]:

$$h_{it}^m = \frac{\exp\left(\beta_0^m + \beta_1^m X_{it} + \theta_t^m\right)}{1 + \sum_{m=1}^{2} \exp\left(\beta_0^m + \beta_1^m X_{it} + \theta_t^m\right)}$$

Where X_{it} is a vector of covariates and θ_t^m is the baseline hazard function which is modeled as a piece-wise constant over a two year interval in order introduce the maximum flexibility to the baseline hazard specification.

For the given hazard rate described above, the individual worker's likelihood contribution has the same form of the likelihood of a standard multinomial logit model (Allison, 1982).

$$L_i = \left[\prod_{m=1}^{2} h_{it}^{m}\right]^{c_{it}^m} \left[h_{it}^{0}\right]^{1-\sum_{m=1}^{2} c_{it}^m} \left[\prod_{\tau=1}^{t-1} h_{i\tau}^{0}\right]$$

where c_{it}^m is a destination-specific censoring indicator which equals 1 if worker i exits to state m in interval t, and h_{it}^0 is the conditional probability of making no transition (reference category of staying in total disability).

$$h_{it}^0 = \frac{1}{1 + \sum_{m=1}^{2} \exp\left(\beta_0^m + \beta_1^m X_{it} + \theta_t^m\right)}$$

$$c_{it}^m = \begin{cases} 1 \text{ if } c_{it} = m \\ 0 \text{ if } c_{it} = 0 \end{cases} \Rightarrow c^m$$

[16] The assumption of independence may be questionable in the context of our analysis but it is still reasonable.

[17] We assume that the conditional destination-specific hazard rate is independent from the other competing risks and can be derived from a latent variable model in which the error term has a standard logistic distribution

6 RESULTS

6.1 Competing risk model for exits from total disability

Table 3 of the appendix reports the results of the competing risk model for exits from total disability into work or into absolute disability pensions.

All the variables describing individual characteristics have the expected sign: both age and the level of education are strong determinants of the probabilities of finding a job and of moving into absolute disability pensions. Individuals with a higher degree of education tend to move more to a job as well as to absolute disability pensions whereas lower educated individuals seems to be trapped into total disability pensions. On the other hand, as individuals get older, their level of disability increases and they have a higher probability to be transferred to absolute disability pensions and lower probability to find a job. Therefore, younger workers are the ones that are more able to find a job. This is illustrated in figures 11 and 12, which are plots of the empirical hazard rates by age.

Being a woman increases the probability of moving into absolute disability pensions and reduces the probability of finding a job. From the interaction term between woman and age, we can see that women have higher chances to move to absolute disability while young and higher chances to move to a job at relatively more advanced ages with respect to men.

Another important variable appears to be the amount of pension received. Individuals receiving the highest pensions of the pension distribution in our sample have a significant lower probability of finding a job or moving into absolute disability than the other two groups of individuals which receive a lower pension amount. On the other hand, individuals receiving a pension slightly over the pension mean of the sample are more likely to move to an absolute disability pensions than individuals with the lowest pensions but, at the same time, are considerably less likely to find a job. Therefore, this would reflect either the fact that only individuals with the lowest pension amounts have an incentive to look for a job because they cannot make ends meet or the fact that jobs are only available for lower educated individuals (which probably earned a lower salary before becoming disabled and are thus receiving the lowest pensions).

The unemployment rate of the Autonomous Community, which is the variable that captures the business cycle at the regional level, appears not to be significant although it has the expected negative sign for exits to a job. Furthermore, its interaction with the variable capturing age 55 or above, suggests that individuals of this advanced ages are more negatively affected by the economic situation of the region when looking for a job than younger workers. Differentiating between the two genders, we observe that women are more likely to move to absolute disability when the unemployment rate increases and less likely to find a job than men.

We also note that Madrid and Valencia seem to be the Autonomous Communities where it is less probable for disabled individuals in total disability to move to absolute disability pensions whereas Asturias and Cantabria appear to have the highest probability for this transition. On the other hand, Madrid holds the highest probability to find a job for disabled individuals.

Duration dependence decreases over time and, during the first two years of becoming disabled, individuals have the highest probability of being transferred into absolute disability whereas individuals have the highest chance to find a job after three and four years of entering the disability scheme. This is corroborated in figures 8 and 9 which depict the empirical hazard rates from total disability into absolute disability and work.

Finally, the variables that capture the effects of the policy reform in 2004 suggest a negative effect of the introduction of the incentive measures on the probability of women to find a job with respect to men. After the policy was implemented in 2004, women show a lower hazard rate of finding a job at all ages compared to men. Therefore, stronger incentives are needed in order to fight the double discrimination suffered by disabled women in the Spanish labor market.

6.2 Logistic hazard regression model: exits from total Disability + Work

Table 4 of the appendix reports the results of the logistic hazard regression model for exits from total disability and work into only disability pensions, that is, disabled individuals who stop working to receive only the pension.

All the variables describing individual characteristics have the expected sign and complement the results obtained in the previous section. Individuals with higher levels of education are less likely to stop working whereas women and older workers are much more likely to stop working and only receive the pension.

As before, the amount of pension received also affects the probability of stop working for disabled individuals. In particular, individuals receiving the highest pensions are less likely to stop working whereas disabled workers receiving above-average pension amounts are more likely to stop working than individuals with below-average pension amounts.

Not surprisingly, the unemployment rate of the Autonomous Community appears to affect positively exits from a job with this effect being stronger and more significant for older individuals and for women.

With respect to the dummy variables for the Autonomous Communities, Madrid is the region in which fewer jobs are lost for disabled individuals as opposed to the Canary Islands, where the probability of losing the job is the highest of all regions.

Duration dependence decreases over time so that the probability of losing the job reaches the highest level during the first year of being employed and decreases after that. This is illustrated in figure 10 which depicts the empirical hazard rate from total disability and work into total disability only.

Finally, the variables that capture the effects of the policy reform in 2004 suggest that women are less likely to lose their jobs after the introduction of the incentive measures as compared to men. After the implementation of the policy, women show a lower hazard rate of stop working at all ages vis-à-vis men. Therefore, maybe the policy did not foster employment of disabled women but it contributed to the maintenance of the already existing contracts of employment for disabled women.

7 CONCLUSIONS

In this paper we have analyzed the incentives provided by the Social Security system and the economy as a whole that affect both disabled individuals' decisions to work and firms' decisions to hire these disabled individuals. We have showed the importance of this question given the low employment rate of disabled individuals in Spain with respect to other OECD countries and the potential negative implications of this fact.

In order to understand the incentives of disabled individuals to work, we have first designed a job search model in which we identified all the possible choices available to the individual given his/her degree of disability and given the regulations embedded in the Spanish system.

In the empirical part of the paper, we have estimated a competing-risk model for exits from total disability pensions into a job or into absolute disability pensions in order to determine the variables that prove to have a stronger effect on this choice. At the same time, we have also introduced as explanatory variables business cycle conditions and we carried out a test of the effects of a policy reform launched in 2004 to promote employment of disabled women. This policy increased the deductions of the social security contributions for employers who hired disabled women. However, our results show that, not only the policy did not increase employment among disabled women but it also proved unable to stop the double negative discrimination suffered by disabled women as the hazard rate of finding a job was higher for men than for women on the years following the introduction of the policy initiative.

In general, our results show that younger and more educated men have higher probabilities of working and that the choice of accepting a job is highly correlated with lower amounts of pensions, which would reinforce the numerous studies available in the literature that highlight the disincentive effects to work provided the level of benefits received.

Not only disabled women have lower probabilities of finding a job but they also have a significantly higher probability of losing their jobs once they find one.

Business cycle conditions seem to only have a significant negative effect on workers aged 55 and over both for finding and losing a job.

The majority of the jobs found by disabled individuals don't last for a long period, with the hazard rate of losing the job reaching its maximum only one year after being hired. Therefore, disabled workers not only have difficulties in working but, when they do work, these jobs are of a temporary nature.

The findings of this paper reinforce the observation made by the report of the OECD released in 2007 which identified the ineffectiveness of employment regulations and supports to hire people with disabilities as one[18] of the main causes for the low employment rates of disabled individuals in Spain.

Therefore, a stronger political effort is needed in order to implement the principles of non-discrimination, participation and universal accessibility, together with increased and more attractive incentives to firms to hire disabled workers, as the existing ones have shown not to have any significant effect on employment of disabled individuals in Spain.

The reforms introduced in 2006 seem to be a first step in this direction and the evaluation of the results of this new policy will shed some more light on the necessary tools to be implemented in order to foster employment among disabled workers, with a special emphasis on disabled women.

REFERENCES

AARTS, L.J.M., AND P.R. DE JONG (1996): "European Experiences with Disability Policy", in J.L. Mashaw, V. Reno, R.V. Burkhauser, M. Berkowitz (eds.), Disability, Work and Cash Benefits. W.E. Upjohn Institute for Employment Research: Kalamazoo, Michigan.

ALBA, A., AND MORENO, F. (2004): "Discapacidad y Mercado de Trabajo", Obra Social Caja Madrid. Gente Interactiva. S.L. Madrid.

BLANCO, J.M. (ed.) (2002): Las Incapacidades Laborales. Un punto de vista práctico. 2ª edition, Editorial Lex Nova, Valladolid.

BENÍTEZ-SILVA, H., BUCHINSKY, M., CHAN, H., CHEIDVASSER, S., AND RUST, J. (1999): "An Empirical Analysis of the Social Security Disability Application, Appeal and Award Process", *Labour Economics*, Vol.6, 147-178.

BENÍTEZ-SILVA, H., BUCHINSKY, M., CHAN, H., CHEIDVASSER, S., AND RUST, J. (2004): "How Large is the Bias in Self-Reported Disability?" *Journal of Applied Econometrics*, Vol.19-6, 649-670.

BENÍTEZ-SILVA, H., BUCHINSKY, M., AND RUST, J. (2006): "How Large are the Classification Errors in the Social Security Disability Award Process?" manuscript, SUNY-Stony Brook, UCLA, and University of Maryland.

BENÍTEZ-SILVA, H., BUCHINSKY, M., AND RUST, J. (2008): "Induced Entry Effects of a $1 for $2 Offset in SSDI Benefits", manuscript, SUNY-Stony Brook, UCLA, and University of Maryland.

BENÍTEZ-SILVA, H., AND HEILAND, F. (2007): "The Social Security Earnings Test and Work Incentives", *Journal of Policy Analysis and Management*, Vol.26-3, 527-555.

BENÍTEZ-SILVA, H., AND HEILAND, F. (2008): "Early claiming of Social Security Benefits and Labor Supply Behavior of Older Americans", Forthcoming in *Applied Economics*

BOUND, J., SCHOENMBAUM. M., STINEBRICKNER, T.R., AND WAIDMAN, T. (1999): "The Dynamic Effects of Health on the Labor Force Transitions of Older Workers", *Labour Economics,* Vol.6, 179-202.

BOUND, J., AND R.BURKHAUSER (1999): "Economic Analysis of Transfer Programs Targeted on People with Disabilities", in O.Ashenfelter and D.Card (eds.), *Handbook of Labor Economics*, Vol.3C. Elsevier Science, North Holland: Amsterdam, The Netherlands.

BURKHAUSER, R.V., A.J. HOUTENVILLE, AND D.C. WITTENBURG (2003): "A User's Guide to Current Statistics on the Employment of People with Disabilities", in D.C. Stapleton and R.V.

[18] Two additional causes are highlighted: low work motivation due to the widespread use of temporary contracts and the stigmatization with regards to the ability to work of people with disabilities.

Burkhauser (eds.), *The Decline of People with Disabilities: A Policy Puzzle*. W.E. Upjohn Institute for Employment Research: Kalamazoo, Michigan.

HAVEMAN, R.H., AND B.WOLFE (2000): "The Economics of Disability and Disability Policy", in A.J. Culyer and J.P. Newhouse (eds.), *Handbook of Health Economics*, Volume 1. Elsevier Science, North Holland: Amsterdam, The Netherlands.

HOYNES, H.W., AND R.MOFFITT (1996): "The Effectiveness of Financial Work Incentives in Social Security Disability Insurance and Supplemental Security Income: Lessons from Other Transfer Programs", in J.L. Mashaw, V. Reno, R.V. Burkhauser, M. Berkowitz (eds.), *Disability, Work and Cash Benefits*. W.E. Upjohn Institute for Employment Research: Kalamazoo, Michigan.

JIMÉNEZ, S. (2006): "Evaluating the Labor Supply Effects of Alternative Reforms of the Spanish Pension System", *Moneda y Crédito*, 222, 271-312.

JIMÉNEZ-MARTÍN, S., J.M. LABEAGA, AND C. VILAPLANA PRIETO (2006): "Award Errors and Permanent Disability Benefits in Spain", Universitat Pompeu Fabra WP-966.

JIMÉNEZ-MARTÍN, S. (ed.) (2008): *Aspectos Económicos de la Dependencia y el Cuidado Informal en España*. Universitat Pompeu Fabra, Fedea y Fundación AstrZeneca.

LAFORTUNE, G., AND BALESTAT, G. (2007): "Trends in Severe Disability Among Elderly People: Assessing the Evidence in 12 OECD Countries and the Future Implications", *OECD Health Working Papers*, No.26, OECD Publishing.

LJUNGQVIST, L., AND SARGENT, T.J. (1995): "The Swedish Unemployment Experience", *European Economic Review*, Vol.39, 1043-1070.

MAGNAC, T. AND D. THESMAR (2002): "Identifying Dynamic Discrete Decision Processes", *Econometrica*, Vol.70-2, 801-816.

MALO, M.A. (2004): "Cómo Afectan las Discapacidades a la Probabilidad de ser Activo en España? Un Análisis Empírico con Datos de la Encuesta sobre Discapacidades, Deficiencias y Estado de Salud de 1999", *Cuadernos de Economía*, Vol.27, 75-108.

MALO, M.A., AND MUÑOZ-BULLÓN, F. (2006): "Employment Promotion Measures and the Quality of the Job Match for Persons with Disabilities", *Hacienda Pública Española*, 179, 79-111.

MALO, M.A., CUETO, B., AND RODRÍGUEZ, V. (2007): "Compatibilidad entre Pensiones Contributivas por Incapacidad y Empleo: El Caso Español", presented at the *Economic Well-Being and Social Protection in Southern Europe: Changing Policies and Trends*, Instituto de Estudios Fiscales, Madrid.

MCCALL, J.J. (1970): "Economics of Information and Job Search", *Quarterly Journal of Economics*, Vol.84, 113-126.

MORTENSEN, D.T. (1970): "Job Search, the Duration of Unemployment and the Phillips Curve", *American Economic Review*, Vol.60, 5, 847-62.

MORTENSEN, D.T. (1986): "Job Search and Labor Market Analysis", in O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics*, North Holland: Amsterdam, The Netherlands.

OECD (2003): *"Transforming Disability into Ability: Policies to Promote Work and Income Security for Disabled People"*, OECD, Paris.

OECD (2007): *"Sickness, Disability and Work: Breaking the Barriers"*, Vol.2: Australia, Luxembourg, Spain and the United Kingdom, OECD, Paris.

OECD (2007): "New Ways of Addressing Partial Work Capacity", OECD, Paris.

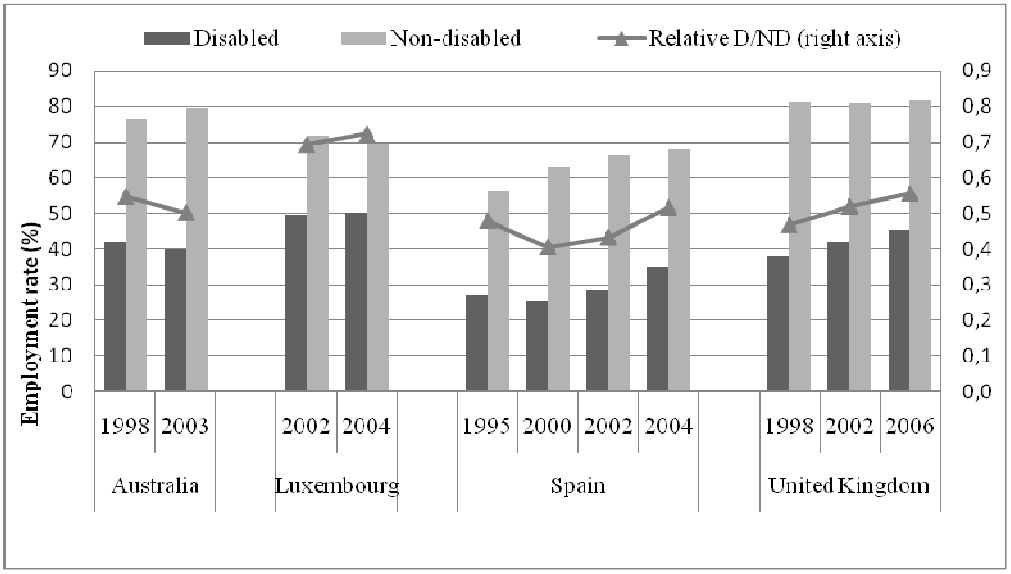
RISSMAN, E.R. (2006): "The Self-Employment Duration of Younger Men over the Business Cycle", *Economic Perspectives*, Vol.30, 3, 14-27.

RISSMAN, E.R. (2007): "Labor Market Transitions and Self-Employment", Federal Reserve Bank of Chicago, WP 2007-14.

YIN, N. (2008): "Partial Benefits in the Social Security Disability Insurance: A Policy Alternative to Foster Work among the Disabled", manuscript, SUNY-Stony Brook and Baruch College.

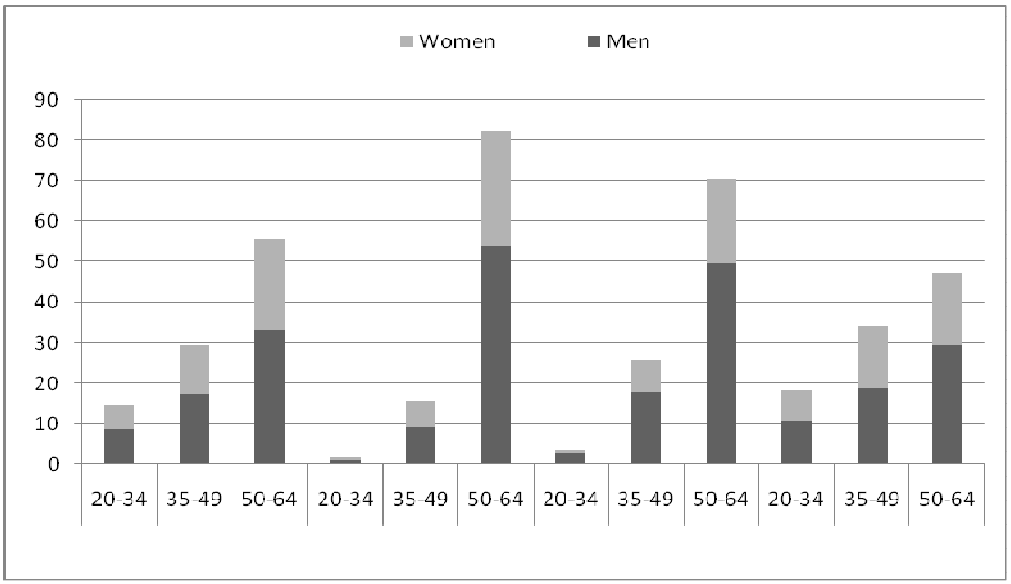
APPENDIX: TABLES & FIGURES

Figure 5. Employment rates for persons with disability are only half the level of those without disability, except in Luxembourg.
Employment rates of working-age people with and without disability, latest available years (percentage).



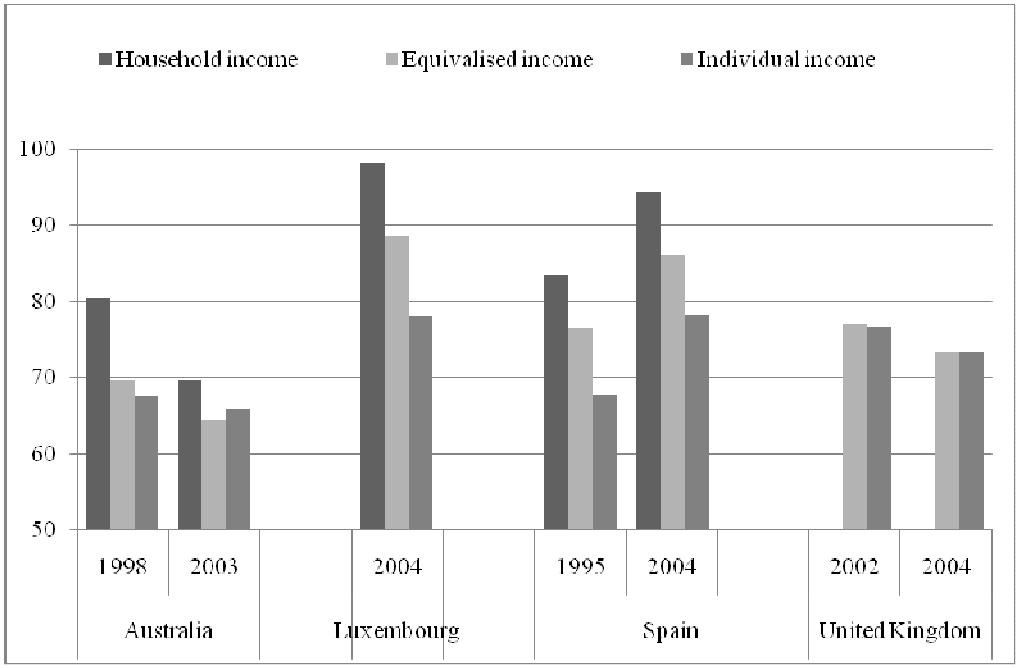
Definition of disability on self-assessment basis: existence of chronic health problem and long-term limitations in daily live activities.
Taken from OECD, 2007. Source: SDAC (Australia); national labour force surveys (Luxembourg and Spain, 2002; United Kingdom); ECHP (Spain 1995, 2000); and EU-SILC (Luxembourg and Spain 2004).

Figure 6. Disability population significantly biased toward older age groups, in particular in Luxembourg and Spain.
Disability recipients by age and gender, as a share of total disability recipients, 2005[a]



a. Beneficiaries: DSP customers (Australia); permanent and temporary disability pension recipients (Luxembourg); contributory and non-contributory incapacity benefits (Spain); Incapacity Benefit and Severe Disablement Allowance claimants, excluding IB short-term lower rate (United Kingdom)
Taken from OECD, 2007 Source: Data from: DEWR (Australia); IGSS (Luxembourg); MTAS (Spain); DWP (United Kingdom).

Figure 7. Relative income levels of persons with disability are higher in continental European countries.
Average incomes of persons of working age with disability over those without disability (percentage).



Total income of households with at least one member with disability; equivalised adjusted income of persons with disability, individual income of persons with disability. Income concept: disposable income, except Australia: gross income.
Taken from OECD, 2007. Source: SDAC (Australia); ECHP (Spain 1995, 2000); FRS (United Kingdom); and EU-SILK (Luxembourg and Spain 2004).

Table 1. Taken from OECD, 2007.

The magnitude of the policy challenge in Australia, Luxembourg, Spain and the United Kingdom, 2007				
Seven key policy challenges[a,b]	Australia	Luxembourg	Spain	United Kingdom
1) Controlling incapacity-related public spending	++	+++	++	++
2) Raising employment rates for people with health problems	+++	+	++++	++
3) Tackling lower incomes of households with disabled people	+++	+	++	+++
4) Reducing the inflow into sickness and disability benefits	++	++	++	++++
5) Addressing the increasing medicalisation of labour market problems	++	++	+	+++
6) Raising the outflow from usually permanent disability benefits	+++	+++	++	++
7) Strengthening co-ordination across different benefit schemes	++	+++	+++	++

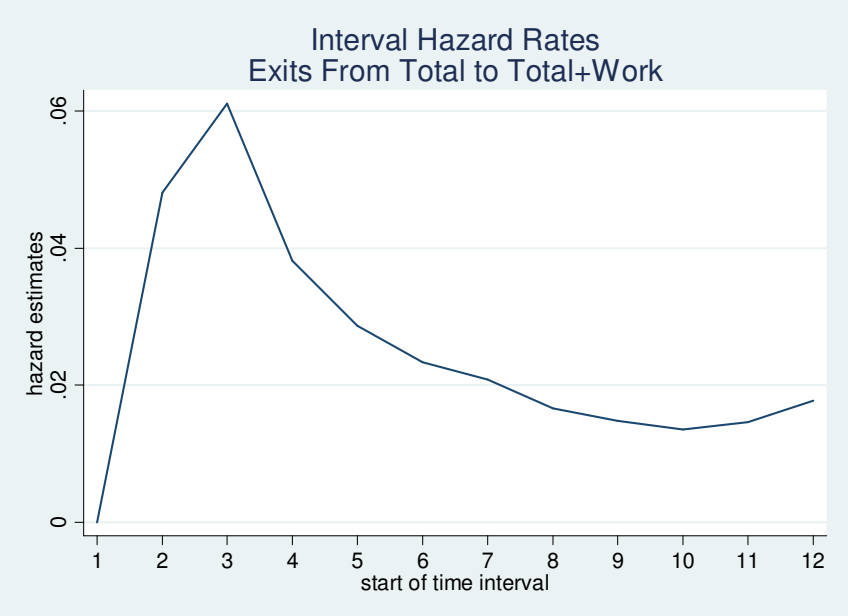
a) The scales should be interpreted as follows: + … minor challenge; ++ … moderate challenge; +++ ... substantial challenge; and ++++ … formidable challenge.
b) Public spending on sickness and disability benefits excludes employer-paid wage payments in the first weeks of absence. Relative poverty is defined as the share of persons with incomes below 50% of the median income of the entire population.
Source: OECD, 2007.

Table 2. Rules for the determination of the pension amount.

	TDP	ADP
Regulatory Base	Depends on the contributions in the last job (salary in the last job)	
	Depends on the source of the disability: common illness, non-work accident or working accident (or professional illness) [19].	
Percentage applied to Regulatory Base	Depends on age of the individual: 55% if <55 years old 75% if >55 years old & no job	100%
	Depends on the source of the disability: if working accident (or professional illness), 30%-50% more if employer broke safety regulations for the job[20].	
Income Taxes	Pay normal income taxes Exempted if Basc[21] country & no job	Exempted from income taxes
Reduction in employment income used to calculate the income tax	2.800 euros/year (if disability level between 33% and 65%) if working 6.200 euros (if disability more than 65% or below that but disabled with reduced mobility) if working	

EMPIRICAL HAZARD RATES BY DURATION

Figure 8. Exits from total disability to total disability + working:



[19] In our database we don't observe the source of the disability, we only observe the regulatory base.

[20] Please note that in the case of an ADP, the final percentage applied could be more than 100% of the regulatory base if the employer is found guilty of not preserving the required safety conditions for the job and has to pay between a 30%-50% increase in the percentage of the regulatory base (over the 100% already established for ADP pensions).

[21] Disabled individuals in the provinces of Vizcaya, Alava and Guipuzcoa, which constitute the Basc country, are exempted of paying income taxes on Total Disability Pensions if they don't work.

Figure 9. Exits from total disability to absolute:

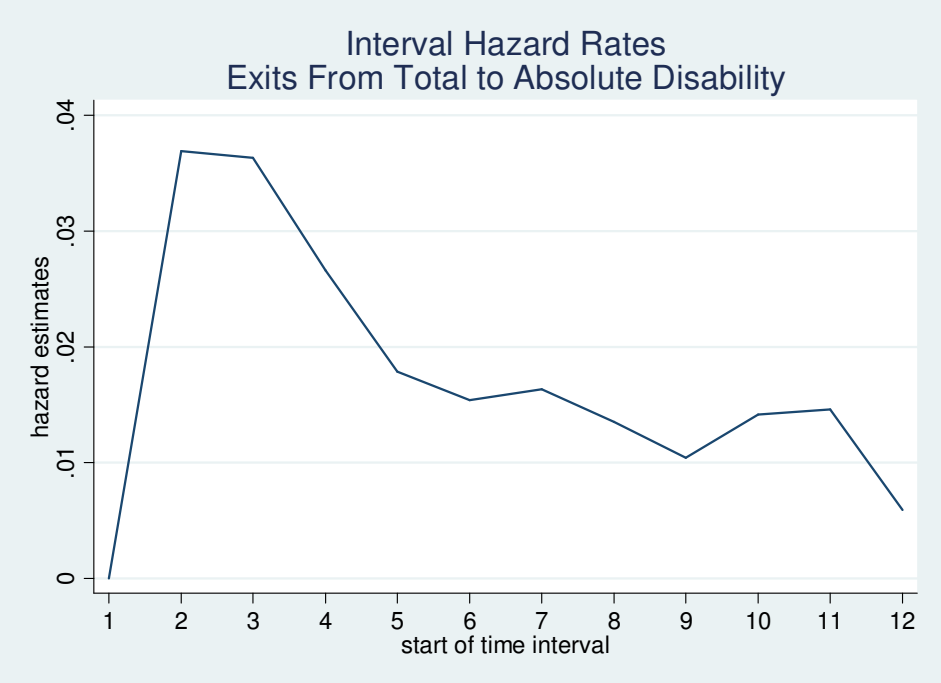
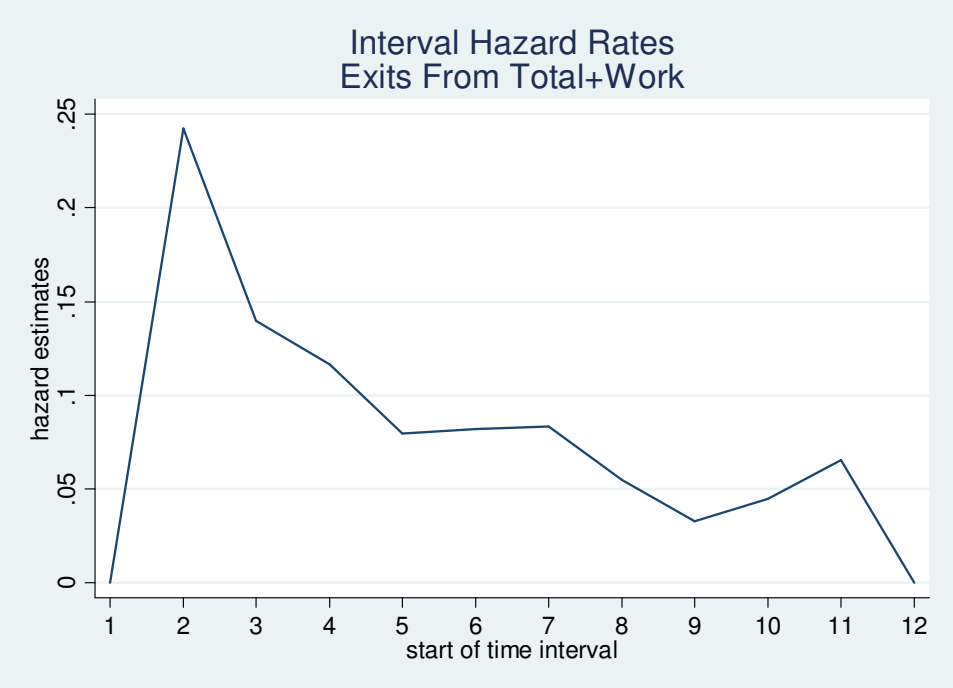


Figure 10. Single exits from total disability + working to all routes:



HAZARD RATES BY AGE

Figure 11. Exits from total to absolute by age:

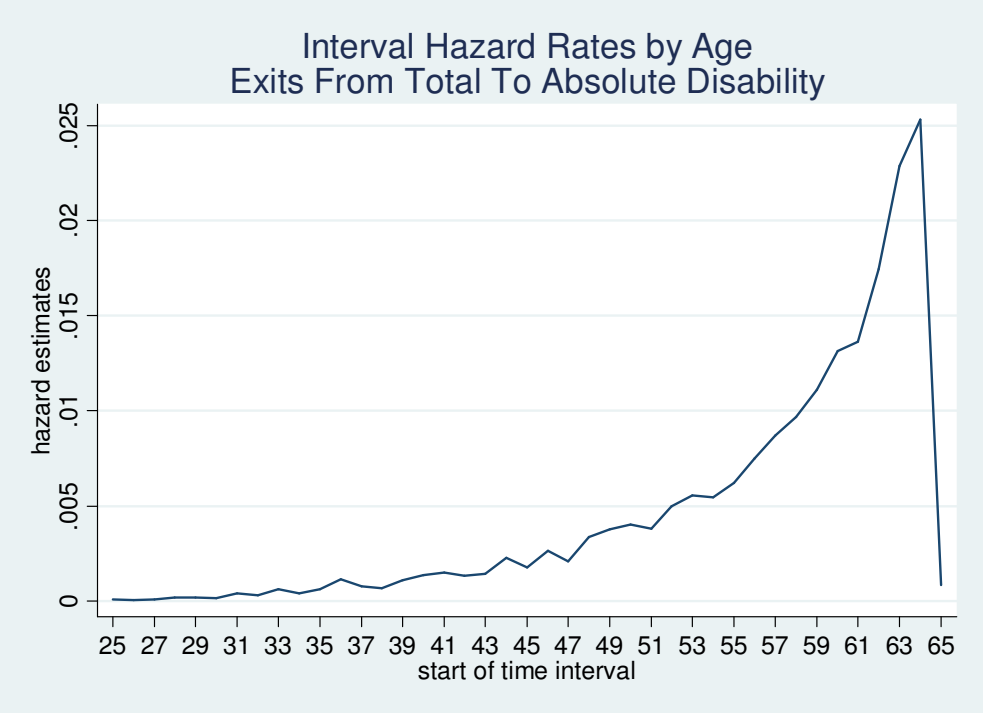


Figure 12. Exits from total disability to total + work by age:

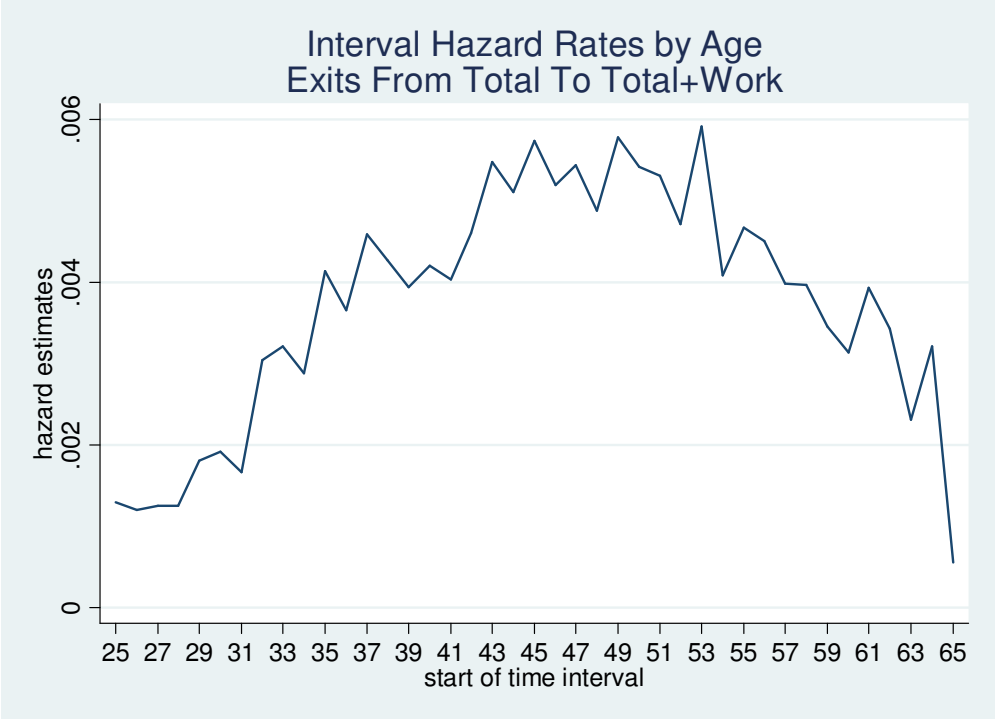


Table 3. Competing-risk hazard regression model, exits from total disability[22]

	Absolute	Work
nived	.1849473***	.2306866***
dur12	1.295549***	1.001956***
dur34	1.04121***	1.263339***
dur56	.3128476	.568397*
dur78	.1936696	.2782491
dur910	.0623797	-.0320684
Di	1.056996*	-1.041098**
age	.0277173***	-.0706337***
Diage	-.0286885***	.0098921
Ditless45	.1489229	-.1057449
Ditamore45	-.104439	-.1060071
penhigh	.0009395	-.1345997**
pentop	-2.823036***	-2.876131***
andalucia	.0792895	-.2379114
aragon	.4747029	-.2047013
asturias	1.692552*	-.2086874
baleares	.3363498	-.1652321
canarias	.8468795	-.546238
cantabria	1.374852	-.6103653
castleon	.3714658	-.1919812
castmancha	.8411825	-.2665731
extremadura	1.114773	-.2913881
galicia	.972108*	-.490373
murcia	.9382744	-.153999
navarra	1.065913	-.2441928
paisbasco	.968999	-.2444511
rioja	.85332	-.1495263
catalunya	.1786083	-.0684454
valencia	-.0527617	-.3617802*
ur	-.0222271	-.029285
ura55more	-.0083282	-.075713***
urDi	.0122553	-.0056094
popu	4.01e-07	-3.71e-08
log-likelihood		-1.75e+04
N		77894
* p<0.05, ** p<0.01, *** p<0.001		

[22] Year dummies are also included in the model.

Table 4. Exits from Total Disability+Work; Logístic hazard regression model

Coefficients	
nived	-.1031191*
dur12	-3.360404**
dur34	-3.814078***
dur56	-4.443666***
dur78	-4.43856***
dur910	-5.12191***
dur1112	-4.545604***
y3	.0255219
y4	-.3413059
y5	-.2789592
y6	-.4302366
y7	-.3101623
y8	-.6482444*
y9	-.446644
y10	-.4056313
y11	-.4456869
y12	-.767023*
Di	1.277997*
age	.0408603***
age55	.0272911
Diage	-.0174509
Ditless45	-.1183923
Ditamore45	-.0215109
penhigh	.1283097*
pentop	-2.988032***
andalucia	.3112771
aragon	.3756522
asturias	.636039
baleares	.6397798
canarias	.9577325
cantabria	.3763796
castleon	.3283024
castmancha	.2216167
extremadura	.3304128
galicia	.4922269
murcia	.1667779
navarra	.8573216
paisbasco	.3886442
rioja	.1335585
catalunya	.2678726
valencia	.3164146
ur	.0024863
ura55more	.0394328***
urDi	.0162748
popu	8.48e-08
log-likeliho	-3948.148
N	13255
* p<0.05, ** p<0.01, *** p<0.001	

Table 5. Reforms in firms incentives

Years	Type of Incentive	Permanent & Full-time contract	Temporary contract (part or full-time)	Permanent & Part-time contract	Self-employed
Status Quo: 1983	Subsidy1 Subsidy2	500.000 ptas/contract Avoid accidents			
	Deductions SS contributions	70% if < 45 90% if >=45			
1999 (Jan)	Subsidy1	650.000 ptas/contract		Proportional to h	
	Deductions SS contributions			70% if < 45 90% if >=45	
2002 (Dec)	Deductions in firm's taxes	6.000 euro for each person/year of increase over the mean of disabled workers wrt mean of non-disabled		6.000 for each person/year of increase over the mean of disabled workers wrt mean of non-disabled	
(April)	Deductions SS contributions				50% for 3 years
2004 (Jan)	Deductions SS contributions	90% if <45 & women 100% if >=45 & women		90% if <45 & women 100% if >=45 & women	
	Subsidy1	3.907 eur/contract (no change)		Proportional to h[23]	
	Subsidy2	Avoid accidents & adjust working space (max.901,52eur)	Avoid accidents & adjust working space		
2006 (Dec)	Deductions SS contributions	375 euro/month 425 if severe disability[24] +100 if >=45 years +70,83 if women	291,66 euro/month 341,66 if severe disability +50 if >=45 +50 if women	[25]100% of full-time if h=3/4 75% if 1/2>h<3/4 50% if 1/4>h<1/2 25% if h<1/4	50% for 5 years
2007 (June)	Subsidy3				Establish; max 10.000 euro External Services; 75% of cost, max 2.000 euro Education; 75% of cost, max 3.000 euro
	Interest rate reduction				4 points : loans for investment

[23] Number of working hours

[24] I will not be able to identify individuals with a severe disability, as I don't have information on the specific disability of each individual.

[25] Also for temporary and part-time contracts.

ANÁLISIS DE LA POBREZA Y LA EXCLUSIÓN SOCIAL EN LOS HOGARES CON DISCAPACITADOS, CON ESPECIAL ATENCIÓN AL CASO ESPAÑOL

JUANA DOMÍNGUEZ DOMÍNGUEZ
E-mail: juana.dominguez@uah.es
JOSÉ JAVIER NÚÑEZ VELÁZQUEZ
E-mail: josej.nunez@uah.es
Departamento de Estadística, Estructura Económica y O.E.I.
Universidad de Alcalá

Resumen
En este trabajo, se efectúa un análisis del grado de exclusión social de los hogares españoles pobres en los que alguno de sus miembros presenta algún tipo de discapacidad. Para ello, se considera un conjunto inicial de indicadores cualitativos que pueden implicar una situación de exclusión social y se construye un indicador sintético a partir del grado de exclusión social inherente en cada hogar. El indicador sintético propuesto presenta un buen comportamiento de acuerdo con las características más aceptadas en la literatura. Por otra parte, se incluye una comparación con la pobreza existente en la población en su conjunto, tanto desde el punto de vista de la incidencia como de su intensidad. Los datos utilizados proceden del Panel de Hogares de la Unión Europea (PHOGUE), correspondientes a España (1993-2000), y a las primeras oleadas de la Encuesta de Condiciones de Vida (2004-2006).

Palabras Clave: Discapacidad, Exclusión social, Pobreza, Condiciones de Vida.
Área temática: Economía Social y Discapacidad.

Abstract
This paper deals with an analysis of the social exclusion extent among the Spanish poor households with disabled people in. In doing so, a starting set of qualitative indicators is considered in such a way that each of them may imply a social exclusion state. Then, a synthetic indicator is built using the compounded social exclusion degree found in each of the studied households. This proposed synthetic indicator shows a good performance according with the most demanded properties. On the other hand, a comparison with poverty levels in the whole population of households is accomplished, taking into account both the incidence and the intensity aspects of poverty. Data used come from the Spanish waves of the European Community Household Panel (ECHP), corresponding to the 1993-2000 period, as well as from the available Spanish waves (period 2004-2006) of the European Union Statistics on Income and Living Conditions (EU-SILC).

Key Words: Disability, Social Exclusion, Poverty, Living Conditions.
Thematic Area: Social Economics.

1. INTRODUCCIÓN

El objetivo de este trabajo es el estudio comparativo de la evolución de la pobreza en la población de hogares españoles con la mostrada por la exclusión social en aquellos hogares pobres donde hay una o varias personas con algún tipo de discapacidad, asumiendo de esta manera implícitamente la pobreza como un indicador más de exclusión social. Por tanto, la primera cuestión que se plantea es como adoptar una definición válida de discapacidad[1].

[1] Para mayor detalle sobre las diversas definiciones relacionadas con la discapacidad, véase Malo (2007).

En general, se definen las discapacidades como las consecuencias provocadas por las deficiencias de un individuo desde el punto de vista del rendimiento funcional y de la actividad del mismo. La Clasificación Internacional de Deficiencias, Discapacidades y Minusvalías que hace la Organización Mundial de la Salud establece lo siguiente:

- *Deficiencia es toda pérdida o anomalía de la estructura o función psicológica, fisiológica o anatómica.*
- *Discapacidad es toda restricción o ausencia, debida a una deficiencia, de la capacidad de realizar una actividad en la forma o dentro del margen considerado como normal.*
- *Minusvalía es toda situación de desventaja sufrida por un individuo como consecuencia de una deficiencia o de una discapacidad, que limita o impide el desempeño de un rol dado en función de las características personales en un determinado entorno.*

Así, en relación con las definiciones anteriores, Malo (2003) resume: "El término deficiencia alude a cuestiones orgánicas, la discapacidad se refiere a las consecuencias (si las hay) de esos problemas orgánicos y la minusvalía se encontraría en el plano social, ya que se trata de los impedimentos *para ser como los demás* debido a una deficiencia o discapacidad".

En consecuencia, actualmente para el análisis desde el punto de vista estadístico, se utilizan sólo los términos de deficiencia y discapacidad, ya que el concepto de minusvalía plantea problemas de delimitación bastantes severos.

Las actividades a las que se refiere la definición de discapacidad consisten en ver, oír, comunicarse, aprender y aplicar conocimientos, desplazarse dentro del hogar, utilizar brazos y manos, desplazarse fuera del hogar, cuidarse de sí mismo, realizar las tareas del hogar y relacionarse con otras personas.

Las deficiencias se agrupan de manera genérica en tres categorías según que sean físicas, mentales o sensoriales. Se considera que una persona tiene una deficiencia física cuando padece alguna anomalía en la cabeza, columna vertebral, extremidades, parálisis de las extremidades, paraplejías y tetraplejías, trastornos de coordinación de los movimientos, alteraciones en los aparatos respiratorio, cardiovascular, digestivo o genitourinario, así como en el sistema endocrino-metabólico o el sistema inmunitario. En cuanto a las deficiencias mentales, se incluyen los retrasos mentales en sus distintos grados, el retraso madurativo, las demencias y los trastornos mentales tales como el autismo, las esquizofrenias y los trastornos psicóticos, somáticos y de la personalidad. Por último, las deficiencias sensoriales afectan a aquellas personas que presentan trastornos relacionados con la vista, el oído y el lenguaje.

Por una parte, este trabajo se encuadra en el marco general del estudio de la pobreza, en el son muchas las decisiones que deben tomarse y todas ellas con una incidencia importante sobre los resultados obtenidos. Una de las principales líneas de debate tiene su origen en relación con el carácter absoluto o relativo del concepto, tal y como se presenta en Townsend (1979), avivada por el trabajo de Sen (1983), al que siguieron réplicas de Townsend (1985) y el propio Sen (1985a). A modo de resumen, Subramanian (2004) concluye que podría considerarse la pobreza como absoluta en el espacio de las "funcionalidades" (*functionings*) o cosas que una persona valora ser o hacer, pero relativo en el espacio de los recursos, bienes e ingresos.

El estudio de la pobreza a través de las funcionalidades resalta el carácter multidimensional del fenómeno a través de indicadores no monetarios. Sin embargo, los problemas que plantea esta alternativa son importantes, comenzando con el de la selección y disponibilidad de datos adecuados (Laderchi, 1997), lo que ha llevado a algunos autores a utilizar modelos de clases latentes (Pérez Mayo y Fajardo, 2003 ó Ayala y Navarro, 2004, por ejemplo)[2]. La alternativa enlaza de nuevo con el concepto de pobreza económica (Sen, 1976), en el sentido de admitir el ingreso ó el gasto para fijar la posición económica de los hogares, pese a sus limitaciones (Ruiz-Castillo, 1987), que será la aproximación adoptada en el presente trabajo. En este contexto, el propio Sen (1976) señala la incidencia, la intensidad y la desigualdad como las tres dimensiones

[2] No obstante, recientemente, se están produciendo numerosos trabajos en esta línea, habitualmente a través del concepto de *privación multidimensional,* siguiendo las directrices de Brandolini y D'Alessio (2000), de Martínez y Ruiz-Huerta (2000) o de Bourguignon y Chakravarty (2003), aunque esta vía tampoco está exenta de dificultades como los propios autores señalan.

integradas en el fenómeno de la pobreza[3], todas ellas presentes en la elaboración del indicador propuesto por dicho autor, así como en la construcción de las curvas TIP/IID, que permiten efectuar comparaciones en pobreza, propuestas por Jenkins y Lambert (1997), de las que una aplicación al caso español puede verse en Casas, Domínguez y Núñez (2003), y al caso de diversos países de la Unión Europea en Domínguez y Núñez (2006, 2007), por ejemplo. Sin embargo, estas curvas no logran alcanzar una estructura de ordenación total, como tampoco lo consiguen otros tipos de curvas globales, como las propuestas por Atkinson (1987) ó por Foster y Shorrocks (1988a y b).

Por todo ello, en este trabajo, se optará por la aproximación tradicional a través de las dimensiones determinadas por la incidencia y la intensidad de la pobreza, que tampoco quedan exentas de problemas. En efecto, así ocurre incluso desde la identificación de los hogares pobres que permitirá estudiar la incidencia de la pobreza, a partir del valor cuya no superación determina la situación de pobreza, denominado umbral ó línea de pobreza, y del que su determinación ha generado un buen número de propuestas en la literatura[4].

Tampoco resulta fácil la selección de una medida que permita cuantificar la intensidad de la pobreza, habida cuenta de la estructura de orden parcial que las aproximaciones globales proporcionan. La gran variedad de propuestas ha movido a los investigadores a asumir una serie de condiciones mínimas ó axiomas de obligado cumplimiento (Foster, 1984; Núñez, 1990) que, no obstante, no permiten la selección de una alternativa que resulte claramente superior al resto (Foster y Sen, 1997), reproduciéndose la problemática de la cuasi-ordenación de intersección, propuesta por Sen (1973), en el marco de la desigualdad.

En lo que respecta al estudio de la exclusión social, Bossert, D'Ambrosio y Peragine (2007) apuntan que el concepto incluye preocupaciones por fenómenos sociales tan diversos como la pobreza, la privación, el bajo nivel educativo, el desempleo y otras desventajas relacionadas con el mercado laboral, las malas condiciones de la vivienda y la falta de acceso a instituciones políticas y sociales. Así pues, en sentido amplio, tal y como señalan Chakravarty y D'Ambrosio (2006), se dice que una persona está socialmente excluida si no puede participar en las actividades sociales y económicas básicas de la sociedad en que vive. Por lo tanto, como señala Poggi (2007), la exclusión social puede relacionarse con el concepto de *capacidad* de Sen (1985b) y puede definirse como un proceso que conduce a un estado de privación de funcionalidades relevantes.

Por lo tanto, de acuerdo con lo expresado, la exclusión social es un concepto más amplio que el de pobreza, ya que debe incluir otros aspectos que sean indicativos de marginación, además de la pobreza, con lo que su problemática deberá ser forzosamente multidimensional. Por otra parte, tal y como indica Atkinson (1998), debe admitirse que la exclusión social es un concepto relativo, ya que no puede hablarse de personas u hogares socialmente excluidos sin tener en cuenta la situación en que se encuentra el resto.

Por todo ello, en este trabajo, se toma en consideración el módulo de exclusión social incluido en la Encuesta de Condiciones de Vida (ECV), elaborada por el INE, en lo que tiene en común con el Panel de Hogares de la Unión Europea (PHOGUE), que será la otra fuente de datos utilizada, para construir la batería de síntomas sociales y económicos que resultan indicativos de vulnerabilidad a la marginación social. Estos síntomas se constituyen, generalmente, en variables cualitativas dicotómicas, lo que dificultará la construcción de un indicador que cuantifique la exclusión social padecida, como el que se propone en este trabajo.

Así pues, el trabajo se estructura de la siguiente manera. En la sección 2, se presentan los datos e indicadores de intensidad de pobreza utilizados como base, así como las principales decisiones relacionadas con la escala de equivalencia y el umbral de pobreza utilizados. La sección 3 resume la metodología inherente a la construcción del indicador sintético de exclusión social. La sección 4 presenta la selección de indicadores simples de exclusión social, además de la situación de pobreza, y analiza los resultados obtenidos, en relación con la evolución de la

[3] Con frecuencia, se señala también a la *duración de la situación de pobreza* como una cuarta dimensión, de indudable trascendencia. Esta línea de investigación encuentra un claro precedente en Bane y Ellwood (1986) y se desarrolla básicamente a través de modelos de duración ó de matrices de transición. Algunos de los trabajos en este sentido son Ayala y Rodríguez (2004), Bárcena, Fernández, Lacomba y Martín (2004) ó Cantó, del Río y Gradín (2003).

[4] Una panorámica puede verse en Hagenaars y Van Praag (1985) ó en Hagenaars (1986), con más detalle.

incidencia y la intensidad de la pobreza en los años 1993-2005, así como la evolución de la exclusión social de los hogares pobres con discapacitados en dicho periodo. Finalmente, se establecen las condiciones más relevantes del estudio.

2. DATOS UTILIZADOS Y MEDIDAS DE POBREZA EMPLEADAS.

La información más completa sobre las personas con discapacidad en España se encuentra en la Encuesta sobre Discapacidades, Deficiencias y Estado de Salud (EDDES) de 1999, realizada por el INE en colaboración con el IMSERSO y la Fundación ONCE. No sólo incluye información sobre las discapacidades y las deficiencias que las originan, sino también sobre empleo y sobre el hecho de ser beneficiario de distintas medidas de fomento del empleo dirigidas específicamente a las personas con discapacidad[5] y la intermediación especializada en el mercado de trabajo.

Esta encuesta tiene un tamaño muestral teórico de 79.000 hogares, pero se obtuvo una muestra efectiva de 70.402 hogares. Según García Ferruelo (2000), la falta de respuesta, de casi un 10%, se debió a la resistencia general de la población a abrir la puerta de sus hogares a los agentes entrevistadores y al carácter voluntario de la encuesta. El problema que plantea la EDDES, en relación con los estudios de pobreza, es que no incluye datos de renta, por lo que no se puede utilizar para analizar la pobreza desde el punto de vista económico[6].

Así pues, para evaluar la posición económica de los hogares, se ha utilizado el ingreso neto familiar (X), procedente del Panel de Hogares de la Unión Europea (PHOGUE), cuyos datos son representativos para los Países que componen la Unión Europea en el periodo en que se ejecutó, es decir, desde 1994 hasta 2001. Por otra parte, es conocido que la medición de la pobreza depende críticamente de la escala de equivalencia seleccionada, como indicador de las economías de escala en el seno del hogar (Duclos y Mercader-Prats, 1999), provocando patrones de comportamiento reconocibles[7]. Por ello, teniendo en cuenta la extensión que supondría la consideración de varias escalas de equivalencia y la persistencia de los patrones comentados, se ha optado por trabajar con la escala potencial con s=0.5, que es una de las opciones más habitual en los estudios de pobreza[8].

El PHOGUE[9] es una encuesta de carácter anual, cuya información, en el caso español, se obtiene de un panel de 7.200 hogares, elegidos en el año 1993. Para homogeneizar el proceso de comparación, dichos ingresos se han expresado en euros y se han deflacionado utilizando el IPC con base 1993.

En este trabajo, se ha tenido en cuenta la información de las 8 oleadas, que corresponden a los años 1994 a 2001. Como es bien conocido, los datos de ingresos de cada oleada están siempre referidos al año anterior y, por esta razón, la información se refiere al período comprendido entre los años 1993 y 2000. Sin embargo, en el PHOGUE, la variable indicadora de la presencia de discapacidad no puede obtenerse de una manera muy rigurosa. En concreto, se obtiene del fichero de personas adultas[10] a partir de las dos preguntas siguientes:

1. ¿Tiene alguna enfermedad crónica, física o mental, o alguna discapacidad o deficiencia?

En caso de que la respuesta sea afirmativa, se pasa a la siguiente pregunta,

2- ¿La enfermedad, incapacidad o problema de salud le impide desarrollar su actividad diaria?

Esta última pregunta presenta tres posibles respuestas:

1. Sí, intensamente.

[5] Contrato para la formación y en prácticas de minusválidos, incentivos a la contratación, cuota de reserva, empleo selectivo, subvenciones y empleo protegido en centros especiales de empleo.

[6] En relación con otros aspectos, hay disponibles gran cantidad de trabajos que toman como referencia la EDDES. Por ejemplo, Albarrán, Alonso y Fajardo (2007), entre los más recientes.

[7] Véase, por ejemplo, Burkhauser, Smeeding y Merz (1996), Coulter, Cowell y Jenkins (1992) o Buhmann, Rainwater, Schmaus y Smeeding. (1988) y, con aplicación al caso español, Casas, Domínguez y Núñez (2001) u Oliver, Ramos y Raymond (2001).

[8] Esta escala de equivalencia es la que recomienda EUROSTAT para efectuar estudios de desigualdad y pobreza.

[9] Un estudio detallado del PHOGUE puede encontrarse en Peracchi (2002).

[10] Este fichero excluye a los niños y a las personas que viven en centros especializados.

2. Sí, hasta cierto punto.
3. No.

Por lo tanto, consideraremos personas discapacitadas a todos los individuos que responden una de las dos primeras respuestas a la pregunta 2.

Como puede observarse, hay que destacar que el PHOGUE no sigue la definición de la Organización Mundial de la Salud. Sin embargo, puede considerarse que sigue criterios semejantes si tenemos en cuenta la segunda pregunta y la primera respuesta de ésta.

La Encuesta de Condiciones de Vida (ECV) es también de carácter anual y se obtiene de un panel rotante de 14.640 hogares, elegidos en el año 2004. Para homogeneizar el proceso de comparación, dichos ingresos han sido deflacionados utilizando el IPC con base 1993, ya se disponen de las tres primeras oleadas (años 2004-2006).

Las preguntas de la ECV que hemos utilizado para identificar la discapacidad son similares a las del Panel de Hogares de la Unión Europea:

1. ¿Tiene alguna enfermedad crónica o alguna incapacidad o deficiencia crónicas? Si la respuesta es afirmativa, entonces se pasa a la siguiente pregunta:

2. ¿Se ha encontrado limitado en el desarrollo de su actividad diaria durante los últimos 6 meses debido a un problema de salud?

1. Sí, intensamente.
2. Sí, hasta cierto punto.
3. No.

Para la selección del umbral de pobreza (z), se han seguido las directrices apuntadas en la metodología del PHOGUE y de la ECV, y por tanto, se opta por el 60% del ingreso neto mediano equivalente. Se ha elegido esta línea de pobreza relativa por ser mucho más robusta que las basadas en un porcentaje de la renta media, ya que está menos expuesta a los valores extremos de la distribución. Por otra parte, si se denota por *q* al número de hogares por debajo del umbral de pobreza (es decir, pobres) y por *n* al número total de hogares, las medidas de pobreza seleccionadas son[11]:

Medida de Foster, Greer y Thorbecke de orden 2 (FGT2):

$$FGT2(x,z) = \frac{1}{nz^2}\sum_{i=1}^{q}(z - x_i)^2$$

Medida de Foster, Greer y Thorbecke de orden 3 (FGT3 o pobreza severa):

$$FGT3(x,z) = \frac{1}{nz^3}\sum_{i=1}^{q}(z - x_i)^3$$

Lógicamente, la proporción de pobres (H=q/n) sólo se ha utilizado para estudiar la evolución de la incidencia de la pobreza a lo largo del periodo.

3. INDICADOR SINTÉTICO DE EXCLUSIÓN SOCIAL.

Para la selección del indicador de exclusión social, hemos optado por un indicador sintético que permite medir el grado de exclusión social de los hogares donde hay una o varias personas con discapacidad. En concreto, en este trabajo, se propone una adaptación de la medida

[11] Propuestas en Foster, Greer y Thorbecke (1984). Para mayor información sobre las propiedades de estas medidas de intensidad de la pobreza frente al resto, puede consultarse Domínguez y Martín (2006), entre otros.

propuesta por Dagum, Gambassi y Lemmi (1992) para medir pobreza a partir de conjuntos difusos.

Admitiendo un conjunto de k indicadores cualitativos sintomáticos de la exclusión social (Y_1, Y_2,..., Y_k), se define Y_{ij} como la modalidad del indicador Y_j correspondiente al hogar i de tal forma que:

$$Y_{ij} = \begin{cases} 1, & \text{situación de exclusión social del hogar i respecto a } Y_j \\ 0, & \text{situación de no exclusión social del hogar i respecto a } Y_j \end{cases}$$

Representamos por A al conjunto formado por los hogares excluidos socialmente de esa población. Así, para cada hogar i, se define la función de pertenencia $\mu_A(i)$ como su grado de exclusión social, donde $\mu_A(i) \in [0,1]$, de tal forma que, si $\mu_A(i) = 1$, entonces el hogar i vive en una situación de exclusión social extrema. En el caso en que $\mu_A(i) = 0$, dicho hogar i no estaría excluido socialmente. Esta función de pertenencia es similar a los *perfiles de exclusión* construidos en Chakravarty y D'Ambrosio (2006).

Así tenemos:

$$A = \left\{ \left(i, \mu_A(i) \right), i \in \mathbf{N} \quad \mu_A : \mathbf{N} \longrightarrow [0,1] \right\}$$

En tal caso, se plantean dos formas de obtener la función de pertenencia a la situación de exclusión social, dependiendo de que se considere que todos los indicadores tienen la misma importancia o que no la tienen. A continuación, se expone la forma que presenta tal función de pertenencia en ambos casos.

1) Todos los indicadores ponderan igual:

$$\mu_A(i) = \frac{1}{k} \sum_{i=1}^{k} Y_{ij}$$

De esta manera, $\mu_A(i)$ está entre 0 y 1. Efectivamente, si $\mu_A(i)=1$, entonces el hogar i vive en una situación de exclusión social extrema y, si $\mu_A(i)=0$, dicho hogar i no está en esa situación.

2) Todos los indicadores no ponderan igual:

$$\mu_A(i) = \frac{\sum_{j=1}^{k} Y_{ij}\, w_j}{\sum_{j=1}^{k} w_j}, \quad i=1,\ldots, n$$

Es decir, definimos $\mu_A(i)$ como la media ponderada de todos los indicadores Y_1, Y_2, ... , Y_k, en el hogar i, donde w_j es la ponderación correspondiente al indicador Y_j. Dada la manifiesta diferencia en el impacto y la repercusión de las diferentes causas posibles de exclusión social, esta será la opción elegida en este trabajo.

Para estimar las ponderaciones (w_j), se considera una relación funcional inversa entre el sistema de ponderaciones y las frecuencias correspondientes a los síntomas de exclusión social, enfatizando el carácter relativo del concepto de exclusión social al dar mayor importancia a aquellos síntomas menos frecuentes en la población considerada. Así pues, sea f_j la frecuencia relativa del indicador o síntoma de exclusión social Y_j, considerando los n hogares. Es decir:

$$f_j = \frac{\text{nº de veces que el indicador } Y_j \text{ toma el valor 1}}{n} = \frac{\sum_{i=1}^{n} Y_{ij}}{n}, \quad j=1,\ldots, k.$$

Entonces, consideramos la siguiente función, que presenta más estabilidad que la inversa de las frecuencias (Dagum, Gambassi y Lemmi, 1992):

$$w_j = \log\left(\frac{1}{f_j}\right), f_j>0 \quad j=1,\ldots,k$$

Así queda:

$$\mu_A(i) = \frac{\sum_{j=1}^{k} Y_{ij} \log\left(1/f_j\right)}{\sum_{j=1}^{k} \log\left(1/f_j\right)}, i=1,\ldots, n$$

Una vez que tenemos definida la función de pertenencia, se define el índice de exclusión social de la siguiente forma:

$$I_E = \frac{1}{n} \sum_{i=1}^{n} \mu_A(i),$$

donde $\mu_A(i)$ es la función de pertenencia del hogar i al conjunto difuso determinado por la exclusión social. Por lo tanto, $I_E \in [\,0,1\,]$ y representa la proporción de hogares que pertenecen al conjunto difuso de los excluidos socialmente. Así construida, esta medida cumple los axiomas de normalización, monotonía, no decrecimiento de la exclusión social marginal, de descomponibilidad en subgrupos y de anonimato y, por lo tanto, se configura como una medida de exclusión social con un buen comportamiento en relación con las propiedades exigibles a este tipo de medidas[12].

4. SELECCIÓN DE INDICADORES DE EXCLUSIÓN SOCIAL Y ANÁLISIS DE LOS RESULTADOS OBTENIDOS.

Tal y como indica Poggi (2007), el problema de la determinación de las funcionalidades que deben usarse para identificar si un hogar se encuentra excluido, o cómo seleccionarlas, está sujeto a amplia discusión, ya que una lista completa no puede recogerse de manera inequívoca.

En efecto, este es un problema común al intento de medir muchos aspectos sociales en los que sería deseable una aproximación multidimensional, como la pobreza o el bienestar. Sin embargo, como señala Ayala (2006), la política de convergencia de la Unión Europea ha llevado a los indicadores de inclusión social al primer plano de la agenda política comunitaria, de manera que el Comité de Protección Social de la Unión Europea encargó a un grupo internacional de expertos, dirigidos por el Profesor Atkinson, la realización de una propuesta de indicadores de exclusión social, aceptada finalmente en el Consejo de Laeken, en diciembre de 2001. El establecimiento de los criterios básicos para la selección de estos indicadores puede consultarse en Atkinson, Cantillon, Marlier y Nolan (2002). No obstante, pese a los indudables puntos fuertes que presenta la propuesta, por lo que los indicadores de Laeken deberán servir de punto de partida para otras propuestas ulteriores, Ayala (2006) concluye: "… son muchas las limitaciones de partida, siendo las más relevantes las que se refieren a la ausencia de dimensiones fundamentales de la exclusión social, como la vivienda, la inmigración o el acceso a otros bienes básicos relacionados con el bienestar social".

Teniendo en cuenta las consideraciones anteriores, en este trabajo, se opta por una definición de indicadores simples de exclusión social, basada en las fuentes de datos disponibles. Así, en primer lugar, se eligieron las variables cualitativas que, según la información disponible, puede considerarse que mejor definen el estado de exclusión social.

[12] La definición rigurosa de los axiomas referenciados y la comprobación del cumplimiento de los mismos por la medida presentada, ya que puede comprobarse fácilmente que coincide con E_1^n, pueden verse en Chakravarty y D'Ambrosio (2006).

Para ello, se han tomado aquellas preguntas que, formando parte del módulo de exclusión social de la ECV estaban disponibles también en el PHOGUE. En concreto, son las siguientes:

E1.- ¿Puede el hogar permitirse mantener la vivienda con una temperatura adecuada durante los meses de invierno?
E2.- ¿Puede el hogar permitirse pagar unas vacaciones fuera de casa, al menos una semana al año?
E3.- ¿Puede el hogar permitirse una comida de carne, pollo o pescado (o equivalentes para los vegetarianos) al menos cada dos días?
E4.- ¿Se han producido retrasos en el pago de la hipoteca o del alquiler del hogar en los últimos 12 meses?
E5.- ¿Se han producido retrasos en el pago de las facturas de la electricidad, agua, gas, etc., en los últimos 12 meses?
E6.- ¿Se han producido retrasos en el pago de compras aplazadas o de otros préstamos (deudas no relacionadas con la vivienda) en los últimos 12 meses?
E7.- ¿Dispone la vivienda de bañera o ducha?
E8.- ¿Dispone de inodoro con agua corriente en el interior de la vivienda para uso exclusivo del hogar?
E9.- ¿Tiene la vivienda problemas de ruidos producidos por vecinos o procedentes del exterior (tráfico, negocios fábricas, etc.)?
E10.- ¿Tiene la vivienda luz natural insuficiente en alguna habitación?
E11.- ¿Tiene la vivienda problema de goteras, humedades en paredes, suelos, techos o cimientos, o podredumbre en suelos, marcos de ventanas o puertas?
E12.- ¿Tiene la vivienda problemas de contaminación, suciedad, u otros problemas medioambientales producidos por la industria o el tráfico?
E13.- ¿Tiene la vivienda problemas de delincuencia, violencia o vandalismo en la zona?
E14.- Los gastos totales de la vivienda (incluyendo seguros, electricidad, comunidad, etc.) suponen para el hogar:

Una carga pesada *1*
Una carga razonable *2*
Ninguna carga *3*

E15.- ¿Tiene el hogar coche?
E16.- ¿Tiene el hogar televisión en color?
E17.- ¿Tiene el hogar teléfono (incluido móvil)?
E18.- ¿Tiene el hogar ordenador?

A estas características, incorporamos también la condición de pobreza, condicionando su presencia al hecho de que los hogares con discapacitados estén por debajo del umbral de pobreza. De esta manera, asumimos implícitamente la importancia que la pobreza tiene sobre la situación de exclusión social, como admiten la práctica totalidad de los estudios sociológicos que han analizado el fenómeno.

Además, a diferencia de los estudios realizados en relación con la exclusión social en la Unión Europea en Chakravarty y D'Ambrosio (2006) y en Poggi (2007), debe tenerse en cuenta que, en este trabajo, la unidad de análisis es el hogar en lugar del individuo y que se centra sólo en el caso español. Además, estos estudios utilizan perfiles de exclusión no ponderados a diferencia de las funciones de pertenencia a la exclusión social presentadas aquí, donde los síntomas de exclusión se ponderan en relación inversa a su incidencia en los hogares, como ya se ha defendido anteriormente.

Así pues, en la Tabla 1, se presenta el porcentaje de hogares que, siendo pobres, tienen entre sus miembros alguien con algún tipo de discapacidad.

Tabla 1: Porcentaje de hogares con discapacitados que tienen alguna de las carencias. Periodo 1993-2000.

	1993	1994	1995	1996	1997	1998	1999	2000
No Calefacción	0,8073	0,8074	0,7914	0,7553	0,7604	0,7587	0,7282	0,6888
No Vacaciones	0,7969	0,8029	0,8134	0,7958	0,7805	0,7685	0,7258	0,6664
No Comida de carne/pescado	0,1031	0,0704	0,0689	0,0527	0,0426	0,0548	0,0454	0,0404
Retraso Hipoteca/Alquiler	0,0760	0,0833	0,0608	0,0742	0,0489	0,0619	0,0317	0,0442
Retraso recibos Agua/luz	0,1051	0,1096	0,1169	0,1027	0,0810	0,0685	0,0737	0,0494
Retraso Compras Aplazadas	0,0680	0,0396	0,0279	0,0356	0,0216	0,0269	0,0234	0,0159
No Baño/Ducha interior	0,0492	0,0303	0,0298	0,0265	0,0145	0,0124	0,0167	0,0112
No Inodoro Interior	0,0319	0,0176	0,0219	0,0091	0,0113	0,0082	0,0135	0,0038
Ruidos Vecinos	0,2932	0,2630	0,3199	0,3374	0,3079	0,3007	0,2336	0,2466
Luz Insuficiente	0,1932	0,1590	0,2175	0,1769	0,1594	0,1733	0,1191	0,1265
Goteras y Humedades	0,4076	0,3338	0,3493	0,3430	0,3306	0,2943	0,2235	0,2568
Contaminación	0,1722	0,1470	0,1009	0,1180	0,1148	0,0981	0,0811	0,0667
Delincuencia	0,2476	0,2262	0,2063	0,2123	0,1691	0,1540	0,1340	0,1329
Pesada Carga Gastos	0,4778	0,4919	0,4558	0,4531	0,4196	0,4009	0,3421	0,3468
No Coche	0,2966	0,2978	0,2474	0,2533	0,2278	0,2140	0,1338	0,1471
No TV Color	0,0355	0,0283	0,0251	0,0148	0,0135	0,0086	0,0147	0,0092
No Teléfono/Móvil	0,2202	0,2353	0,1920	0,1653	0,1458	0,1075	0,0569	0,0453
No Ordenador	0,0000	0,0000	0,4429	0,4269	0,4139	0,3811	0,2950	0,3522

Fuente: Elaboración propia.

Puede observarse cómo la variable *tiene ordenador en casa* aparece con porcentaje nulo en 1993-1994, porque no se hacía esta pregunta en las dos primeras oleadas del PHOGUE. En la actualidad, sólo el 35% de los hogares no tienen ordenador en casa.

Por otra parte, en esta Tabla 1, se puede observar cómo ha mejorado el equipamiento de los hogares, ya que sólo un 1,1% no tiene baño fijo y el 0,3% no tiene inodoro interior. Los retrasos en los pagos han pasado de ser un 6,8% en el año 1993 a un 1,5% en el año 2000. Los problemas de contaminación han oscilado durante todo el periodo de estudio aunque, a partir de 1998, la tendencia es claramente decreciente, siendo en el año 2000 de 6,7%.

Los problemas de delincuencia han disminuido, aunque se observa en 1996 un repunte con respecto a las tres primeras oleadas del PHOGUE. Sólo el 14% de los hogares donde residen una o varias personas con algún tipo de discapacidad no tiene coche. También, se ha pasado de tener más de un 22% de hogares sin teléfono en 1993 a sólo un poco más del 4% en 2000.

Con respecto a la variable relacionada con las vacaciones, se observa que menos de un 21% de estos hogares podía permitirse pagar unas vacaciones fuera de casa al menos una semana al año en, mientras que en el año 2000 aproximadamente el 36% puede hacerlo.

En la actualidad, casi la totalidad de estos hogares poseen televisión en color y sólo el 4% no se puede permitir comer carne, pollo o pescado al menos cada dos días. Sin embargo, el 68% de los hogares no puede mantener una temperatura adecuada durante los meses de invierno y más de un 34% de ellos consideran una pesada carga los gastos relacionados con la vivienda.

En la Tabla 2, se muestra la evolución del porcentaje de pobres (incidencia de la pobreza), de la intensidad de la pobreza, así como de la exclusión social. Las medidas de incidencia e intensidad de pobreza se calcularon para toda la población[13], mientras que el indicador sintético de exclusión social se ha calculado sólo para aquellos hogares pobres en los que uno o varios de sus miembros presentaban alguna discapacidad, teniendo en cuenta que no se contabilizan los niños ni las personas que residen en centros especializados.

[13] Es decir, para todos los hogares, tuviesen o no personas con algún tipo de discapacidad.

Tabla 2: Evolución de la pobreza y la exclusión social en el periodo 1993-2000.

AÑO	1993	1994	1995	1996	1997	1998	1999	2000
H	0,1878	0,1762	0,1685	0,1857	0,1698	0,1755	0,1729	0,1918
FGT2	0,0309	0,0282	0,0288	0,0324	0,0283	0,0274	0,0253	0,0282
FGT3	0,0209	0,0190	0,0196	0,0220	0,0190	0,0184	0,0174	0,0190
Exclusión Social.	0,1420	0,1227	0,1245	0,1171	0,1033	0,0989	0,0872	0,0791

Fuente: Elaboración propia.

Cabe destacar, que se registra una gran disminución en los niveles de exclusión social para las familias con discapacitados en el periodo 1993-2000. La proporción se ha reducido desde el 14% hasta el 7,9%. Sin embargo, el porcentaje de pobres en toda la población de hogares ha tenido un movimiento oscilante produciéndose un repunte en el año 2000.

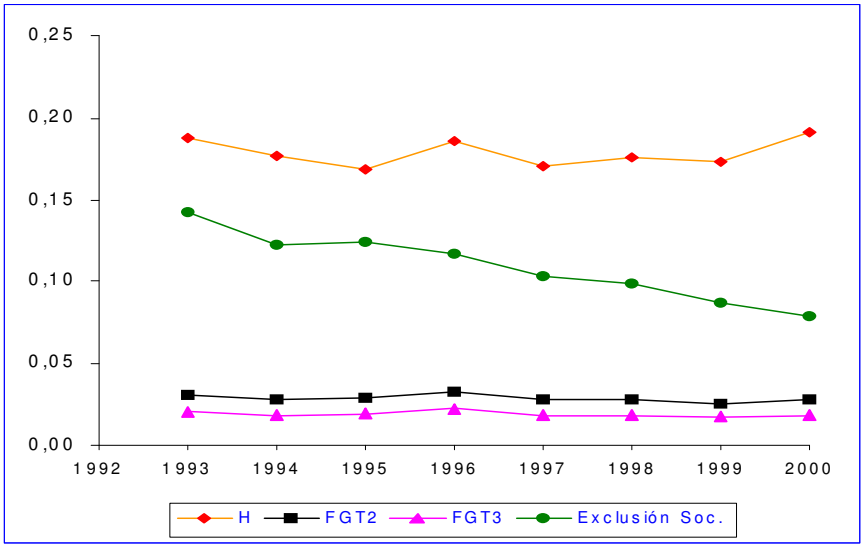
En lo que respecta a la proporción de pobres (H), se observa un apreciable descenso en los años 1994 y 1995, para producirse un repunte en el año 1996, volviendo a disminuir en el año siguiente. En la actualidad, se aprecia que más del 19% de hogares están por debajo del umbral de pobreza en España, que es la cifra más alta desde 1993.

Las medidas de intensidad de pobreza de Foster, Greer y Thorbecke, tanto FGT2 como FGT3, mantienen un movimiento oscilatorio a lo largo de todo el periodo, produciéndose un descenso de la primera en más de un 8,5%, en términos reales con respecto al año 1993, y superior a un 9%, para la segunda medida.

Si se observa ahora la evolución del indicador sintético de exclusión social en los hogares donde residen personas con alguna discapacidad, se aprecia un repunte en el año 1995 mientras que, a partir del año 1996, se muestra una clara tendencia decreciente. Efectivamente, en términos reales, se comprueba que se ha producido una disminución de más del 44% entre 1993 y 2000.

Todas las trayectorias comentadas de los indicadores de pobreza y exclusión social, durante 1993-2000, se visualizan con mayor comodidad en el gráfico que se adjunta a continuación, que corresponde a la Figura 1. Así, a modo de resumen, se observa cómo la proporción de pobres y las medidas de pobreza FGT2 y FGT3 presentan una tendencia oscilante, aunque creciente al final del periodo, mientras que la medida de exclusión social es claramente decreciente.

Figura 1: Evolución de la pobreza y la exclusión social, en el periodo 1993-2000.



Fuente: Elaboración propia.

Teniendo en cuenta las diferentes fuentes de datos y la dificultad para establecer una comparación global de todo el periodo considerado, por cuestiones de homogeneidad, se pasa ahora al análisis de los resultados correspondientes al periodo 2004-2006. Así, la Tabla 3 muestra la incidencia de los síntomas de exclusión social sobre los hogares pobres con discapacitados.

Tabla 3: Porcentaje de hogares con discapacitados que tienen alguna de las carencias. Periodo 2004-2006.

	2004	2005	2006
No Calefacción	0,2074	0,1748	0,2452
No Vacaciones	0,7923	0,7210	0,7230
No Comida de carne/pescado	0,0916	0,0460	0,0604
Retraso Hipoteca/Alquiler	0,0458	0,0288	0,0253
Retraso recibos Agua/luz	0,1029	0,0506	0,0446
Retraso Compras Aplazadas	0,0511	0,0198	0,0207
No Baño/Ducha interior	0,0225	0,0134	0,0095
No Inodoro Interior	0,0168	0,0097	0,0116
Ruidos Vecinos	0,2774	0,2782	0,2736
Luz Insuficiente	0,2046	0,1470	0,1543
Goteras y Humedades	0,3071	0,2658	0,2712
Contaminación	0,1566	0,1662	0,1344
Delincuencia	0,2151	0,1949	0,1873
Pesada Carga Gastos	0,6839	0,5829	0,6542
No Coche	0,1355	0,1040	0,1095
No TV Color	0,0119	0,0080	0,0050
No Teléfono/Móvil	0,0350	0,0348	0,0267
No Ordenador	0,1930	0,1367	0,1427

Fuente: Elaboración propia.

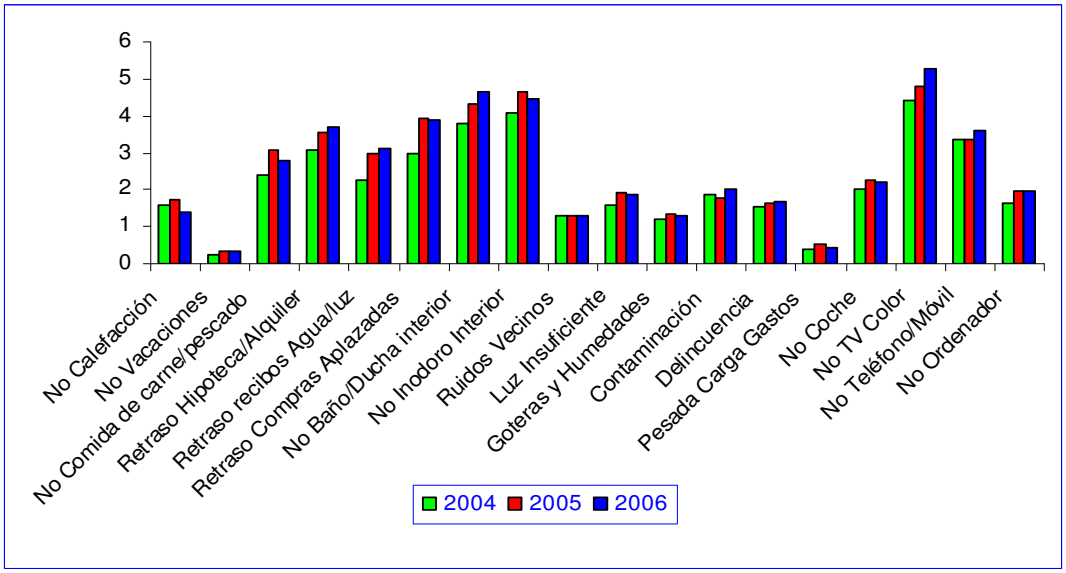
En esta Tabla 3, se puede observar cómo ha mejorado el equipamiento de los hogares, ya que sólo un 24,5% no pueden mantener una temperatura adecuada durante los meses de invierno, el 0,9% no tiene baño fijo y el 1,1% no tiene inodoro interior. Los retrasos en los pagos han pasado de ser algo más de un 5% en el año 2004 a un 2% en el año 2006. Los problemas de contaminación han oscilado durante todo el periodo de estudio, alcanzando en el año 2006 una incidencia de un 13,4%.

Los problemas de delincuencia han disminuido, pasando de más de un 21,5% en 2004 a menos de un 19% en 2006. Sólo el 11% de los hogares donde residen una o varias personas con algún tipo de discapacidad no tiene coche. También, se ha pasado de tener un 3,5% de hogares sin teléfono en 2004 a sólo un poco más del 2,5% en 2006.

Con respecto a la variable relacionada con las vacaciones, se observa que menos de un 21% de estos hogares podía permitirse pagar unas vacaciones fuera de casa al menos una semana al año en 2004, mientras que en el año 2006 aproximadamente el 27% puede hacerlo. Prácticamente la totalidad de los hogares tienen TV en color y sólo el 15% no tiene ordenador en casa.

La Figura 2 muestra las ponderaciones absolutas obtenidas para los síntomas de exclusión.

Figura 2: Evolución de las ponderaciones en los hogares donde hay una o varias personas con algún tipo de discapacidad. Periodo 2004-2006



Fuente: Elaboración propia.

En la Tabla 4, se muestra la evolución de la pobreza y la exclusión social en el período 2004-2006, teniendo en cuenta las observaciones efectuadas al analizar los datos del periodo anterior.

Tabla 4: Evolución de la pobreza y la exclusión social en el periodo 2004-2006.

AÑO	2004	2005	2006
H	0,2286	0,2414	0,2473
FGT2	0,0349	0,0368	0,0398
FGT3	0,0228	0,0239	0,0262
Exclusión Social.	0,1031	0,0814	0,0796

Fuente: Elaboración propia.

Cabe destacar que se registra una gran disminución en los niveles de exclusión social para las familias con discapacitados en el periodo 2004-2006. La proporción se ha reducido desde el 10% hasta el 7,9%. Sin embargo, el porcentaje de pobres en toda la población de hogares ha tenido un movimiento ascendente en todo el período.

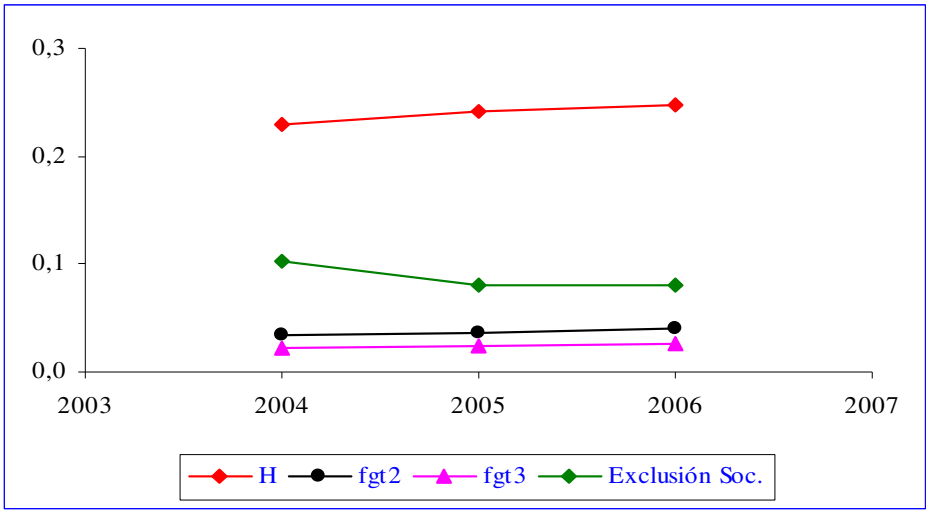
La medida de intensidad de pobreza FGT2 mantiene un movimiento ascendente a lo largo de todo el periodo, produciéndose un aumento de más de un 14%, en términos reales con respecto al año 2004. La misma tendencia se observa para la pobreza severa, aunque aquí el incremento es aún superior, alcanzando el 14,7% con respecto a 2004. Esto significa que las capas más bajas se están empobreciendo cada vez más.

Si se observa ahora la evolución del indicador sintético de exclusión social en los hogares donde residen personas con alguna discapacidad, se aprecia una clara tendencia decreciente. Efectivamente, en términos reales, se comprueba que se ha producido una disminución de más del 22% entre 2004 y 2006.

Todas las trayectorias comentadas de los indicadores de pobreza y exclusión social, durante 2004-2006, se visualizan con mayor comodidad en el gráfico que se adjunta a continuación, que corresponde a la Figura 3.

Claramente se observa como los indicadores de pobreza presentan trayectorias ascendentes, mientras que el de exclusión social muestra la tendencia contraria.

Figura 3: Evolución de la pobreza y la exclusión social, en el periodo 2004-2006.



Fuente: Elaboración propia.

Finalmente, a continuación, se presentan las conclusiones más relevantes del análisis efectuado.

5. CONCLUSIONES

Tras pasar revista a los diferentes conceptos ligados a la problemática de la discapacidad, en este trabajo se ha propuesto un indicador sintético que permite medir la exclusión social a partir de la pobreza de los hogares y de una batería de indicadores cualitativos sintomáticos de exclusión social. Por su especial significación, se ha otorgado un papel preponderante a las condiciones de pobreza de la población, mediante el condicionamiento de las variables cualitativas a la presencia de pobreza en la confección del indicador compuesto.

Además, se ha aplicado el indicador propuesto de exclusión social a los hogares españoles en los que alguno de sus miembros presenta discapacidades. Para ello, se han utilizado los datos provenientes del Panel de Hogares de la Unión Europea (PHOGUE) y de la Encuesta de Condiciones de Vida (ECV), habilitando un periodo de estudio que va desde 1993 hasta 2006, aunque admitiendo la falta de información correspondiente a los años 2001-2003 y las dificultades inherentes a la comparación de los resultados obtenidos entre los dos periodos involucrados, al provenir de distintas fuentes de datos.

Este estudio se ha complementado con el análisis de la incidencia y la intensidad de la pobreza de todos los hogares españoles, durante el mismo periodo 1993-2006. De esta manera, puede establecerse con mayor claridad la problemática de la exclusión social en los hogares con personas con discapacidad, si se compara con la evolución de la pobreza general. La intensidad de la pobreza se ha evaluado a través de las medidas de Foster, Greer y Thorbecke, de orden 2 y 3, siendo esta última especialmente adecuada para evaluar la pobreza severa.

De todo el análisis descrito, se pueden obtener gran número de conclusiones, como se ha ido destacando a lo largo del artículo. Entre las más sobresalientes, pueden destacarse las siguientes:

1.- El porcentaje de pobres en España ha aumentado notablemente desde el año 2000.

2.- Ha aumentado tanto la intensidad de la pobreza (FGT2) como la pobreza severa (FGT3) en los hogares españoles. Sin embargo, el incremento de la pobreza severa es bastante mayor. Esto significa que, aunque han aumentado los hogares en *situación límite,* no son tantos los que están en estado de *precariedad*.

3.- En términos absolutos, se aprecia una disminución de más de 7 puntos porcentuales en el indicador de exclusión social para hogares que incluyen personas con discapacidad. Esta disminución no se ha producido de manera uniforme en el periodo analizado.

Las conclusiones anteriores parecen indicar que, pese al incremento de la población pobre en España, los hogares que incluyen personas con discapacidad se encuentran menos excluidos socialmente en 2006 que al comienzo del periodo. Teniendo en cuenta que se ha analizado la pobreza desde un punto de vista relativo, este rasgo puede ser indicativo de la mejora en las condiciones de vida de la población pobre, en general, y de los hogares con personas con discapacidad, en particular.

Los indicadores cualitativos que han integrado el indicador sintético propuesto para medir la exclusión social han permitido, además, extraer gran cantidad de información útil sobre la evolución de la situación en que viven los hogares pobres con personas discapacitadas. Entre ellas, pueden extraerse las siguientes por su repercusión:

1.- Las viviendas de los hogares pobres que integran personas con discapacidad han ido mejorando paulatinamente sus condiciones de habitabilidad. Prácticamente todos los hogares tienen inodoro en su interior, baño fijo, televisión en color, teléfono, y sólo un 24% no puede mantener una temperatura adecuada durante los meses de invierno en 2006.

2.- Se ha producido una disminución en la frecuencia de retraso del pago de la hipoteca o del alquiler del hogar.

3.- Para más de la mitad de estos hogares, suponen una pesada carga los gastos totales de la vivienda (incluyendo seguros, electricidad, comunidad, etc.).

4.- En la actualidad, sólo el 28% de los hogares donde residen personas con algún tipo de discapacidad puede permitirse pagar una semana de vacaciones fuera de su hogar.

5.- Se ha producido una disminución de los retrasos en el pago de compras aplazadas o de otros préstamos (deudas no relacionadas con la vivienda), al igual que en el pago de las facturas de la electricidad, agua, gas, etc.

Finalmente, cabe reseñar que, en cuanto a los datos del Panel de Hogares de la Unión Europea, los resultados obtenidos deben tomarse con precaución, debido a la gran pérdida de datos muestrales que se ha producido en el periodo que dura esta encuesta (*attrition*), Sin embargo, el tamaño muestral de la Encuesta de Condiciones de Vida es notablemente superior, lo que dota de mayor fiabilidad a las estimaciones derivadas de su utilización.

BIBLIOGRAFÍA

ALBARRÁN, I; ALONSO, P. y FAJARDO, M.A. (2007). "Valoración global de la discapacidad. Propuesta de un índice y su aplicación a la población española recogida en la EDDES". *Estudios de Economía Aplicada*, 25(2), 523-549.

ATKINSON, A.B. (1987). "On the measurement of poverty". *Econometrica*, 55(4), 749-764.

ATKINSON, A.B. (1998). "Social exclusion, poverty and unemployment". En *Exclusion, employment and opportunity (*Atkinson, A.B. y Hills, J., eds.), CASE/4, Centre for Analysis of Social Exclusion, London School of Economics, 1-20.

ATKINSON, A.B.; CANTILLON, B.; MARLIER, E. Y NOLAN, B. (2002). *Social Indicators: The EU and Social Inclusion.* Oxford University Press.

AYALA, L. (2006). "La monitorización de la desigualdad y la exclusión social: hacia un sistema integrado de indicadores". En *V Informe FUHEM de políticas sociales: La exclusión social y el estado del bienestar en España* (Vidal, F.; dir.), FUHEM, Madrid, 43-59.

AYALA, L. y NAVARRO, C. (2004). "Multidimensional indices of housing deprivation with application to Spain". *Papeles de Trabajo* nº 12/04. Instituto de Estudios Fiscales, Madrid.

AYALA, L. y RODRÍGUEZ, M. (2004). "La probabilidad de reentrada en los programas asistenciales: ¿Cómo contribuye el primer episodio?". *VII Encuentro de Economía Aplicada.* Vigo. [Disponible en http://www.revecap.com/viieea/autores].

BANE, M. J. y ELLWOOD, D. T. (1986). "Slipping in and out of poverty: The dynamics of spells". *Journal of Human Resources*, 21(1), 1-23.

BÁRCENA, E. y FERNÁNDEZ, A.; LACOMBA, B. y MARTÍN, G. (2004). "Dinámica de la pobreza a corto plazo en España y Reino Unido a través de los datos del Panel de Hogares Europeo". *Estadística Española*, 46(157), 461-488.

BOSSERT, W.; D'AMBROSIO, C. y PERAGINE, V. (2007). "Deprivation and social exclusion". *Economica,* 74(296), 777-803.

BOURGUIGNON, F. y CHAKRAVARTY, S.R. (2003). "The measurement of multidimensional poverty". *Journal of Economic Inequality,* 1, 25-49.

BRANDOLINI, A. y D'ALESSIO, G. (2000). "Measuring well-being in the functionings space". *26th General Conference of the International Association for Research in Income and Wealth (IARIW).* Cracovia. [Disponible en http://www.iariw.org].

BUHMANN, B.; RAINWATER, L.; SCHMAUS, G. y SMEEDING, T.M. (1988). "Equivalence Scales, Well-Being, Inequality and Poverty: Sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) Database". *Review of Income and Wealth*, 42, 381-399.

BURKHAUSER, R.V.; SMEEDING, T.M. y MERZ, J. (1996). "Relative inequality and poverty in Germany and the United States using alternative equivalence scales". *Review of Income and Wealth*, 42, 381-399.

CANTÓ, O.; DEL RÍO, C. y GRADÍN, C. (2003). "La evolución de la pobreza estática y dinámica en España en el período 1985-1995". *Hacienda Pública Española / Revista de Economía Pública,* 167(4), 87-119.

CASAS, J.M.; DOMÍNGUEZ, J. y NÚÑEZ, J.J. (2001). "Sobre la utilización de las escalas de equivalencia en el estudio de la desigualdad y la pobreza. El caso de España". Ponencia. *Anales de Economía Aplicada. XV Reunión Anual de ASEPELT-España*. La Coruña.

CASAS, J.M.; DOMÍNGUEZ, J. y NÚÑEZ, J.J. (2003). "La pobreza en España: Estudio a partir de curvas I.I.D. y su sensibilidad frente a escalas de equivalencia". En *Información Económica y Técnicas de Análisis en el Siglo XXI* (Casas, J.M. y Pulido, A., eds.). INE, Madrid, 161-173.

CHAKRAVARTY, S.R. y D'AMBROSIO, C. (2006). "The measurement of social exclusion". *Review of Income and Wealth,* 52(3), 377-398.

COULTER, F.; COWELL, F. y JENKINS, S.P. (1992). "Equivalence scale relativities and the extent of inequality and poverty", *Economic Journal*, 102, 1067-1082.

DAGUM, C.; GAMBASSI,R. y LEMMI,A. (1992). "New Approaches to the Measurement of Poverty". En *Poverty Measurement for Economies in Transition in Eastern European Countries*, Polish Statistical Association and Central Statistical Office, Warsaw, 201-225.

DOMÍNGUEZ J, y MARTÍN, A.M. (2006). "Medición de la pobreza: una revisión de los principales indicadores." *Revista de Métodos Cuantitativos para la Economía y la Empresa*. 2**,** 27-66.

DOMÍNGUEZ, J. y NÚÑEZ, J.J. (2006). "La pobreza económica en los países del sur de la Unión Europea entre los años 1993 y 2000, a partir de curvas I.I.D.". *Anales de Economía Aplicada, 2006* (G. Guirao y V.J. Cano, eds.), Delta Publicaciones, Madrid, 439-469.

DOMÍNGUEZ, J. y NÚÑEZ, J.J. (2007). "The evolution of economic inequality in the EU countries during the nineties: A new methodological approach". *Inequality and Poverty. Research on Economic Inequality*, 14, 137-157.

DUCLOS, J.-Y. y MERCADER-PRATS, M. (1999). "Household needs and poverty: with application to Spain and the U.K.". *Review of Income and Wealth*, 45(1), 77-98.

EUROSTAT (1994-2001). *Panel de Hogares de la Unión Europea para España (Primer a Octavo Ciclos).* Madrid.

FOSTER, J.E. (1984). "On economic poverty: A survey of aggregate measures". *Advances in Econometrics*, 3, 215-251.

FOSTER, J.E.; GREER, J. y THORBECKE, E. (1984). "A class of decomposable poverty measures". *Econometrica,* 52(3), 761-766.

FOSTER, J.E. y SEN, A. (1997). *On Economic Inequality. Expanded Edition: Annexe.* Clarendon Paperbacks. Oxford.

FOSTER, J.E. y SHORROCKS, A. (1988a). "Poverty orderings and welfare dominance". *Social Choice and Welfare*, 5, 179-198.

FOSTER, J.E. y SHORROCKS, A (1988b). "Poverty orderings". *Econometrica*, 56(1), 173-178.

GARCIA FERRUELO, M. (2000). "Estudio de la salud, situación laboral y salarios en España". *Cuadernos Aragoneses de Economía*, 10(1), 233-245.

HAGENAARS, A.J.M. (1986). *The perception of poverty*. North Holland. Amsterdam.

HAGENAARS, A.J.M. y VAN PRAAG, B. (1985). "A synthesis of poverty line definitions". *Review of Income and Wealth*, 31, 139-154.

INE (2001). *Encuesta sobre Discapacidades, Deficiencias y Estado de Salud, 1999. Metodología.* Instituto Nacional de Estadística. Madrid.

INE (2004-2006). *Encuesta de Condiciones de Vida. Metodología.* Instituto Nacional de Estadística. Madrid.

JENKINS, S.P. y LAMBERT, P.J. (1997). "Three 'I's of poverty curves, with an analysis of UK poverty trends". *Oxford Economic Papers*, 49, 317-327.

LADERCHI, C.R. (1997). "Poverty and its many dimensions: The role of income as an indicator". *Oxford Development Studies*, 25(3), 345-360.

MALO, M. A. (2003). "Las personas con discapacidad en el mercado de trabajo español". *Revista del Ministerio de Trabajo y Seguridad Social*, 46, 99-126.

MALO, M. A. (2007). "La definición de la discapacidad en la investigación económica: Una reflexión necesaria sobre qué características debería cumplir". *Estudios de Economía Aplicada*, 25(2)**,** 407-428.

MARTÍNEZ, R. y RUÍZ-HUERTA, J. (2000). "Income, multiple deprivation and poverty: An empirical analysis using Spanish data". *26th General Conference of the International Association for Research in Income and Wealth (IARIW)*. Cracovia. [Disponible en http://www.iariw.org].

NÚÑEZ, J.J. (1990). "Una clase de índices estadísticos de pobreza". *Actas de las XV Jornadas Luso-Espanholas de Matemáticas*. Évora (Portugal), IV, 313-319.

OLIVER, J.; RAMOS, X. y RAYMOND, J.L. (2001). "La Mejora en la Distribución de la Renta en España, 1985-1996. Un Análisis de Robustez". En *Desigualdad, Redistribución y Bienestar: Una Aproximación a Partir de la Microsimulación de Reformas Fiscales*. (Labeaga, J.M. y Mercader-Prats, M., Coords.). Estudios de Hacienda Pública. Instituto de Estudios Fiscales. Madrid, 355-372.

PERACCHI, F. (2002). "The European Community Household Panel: A review". *Empirical Economics*, 27, 63-90.

PÉREZ MAYO, J. y FAJARDO, M. A. (2003). "Un método alternativo para identificar los hogares en situación de pobreza". *Estadística Española*, 45(152), 115-134.

POGGI, A. (2007). "Does persistence of social exclusion exist in Spain?". *Journal of Economic Inequalit,* 5, 53-72.

RUIZ-CASTILLO, J. (1987). *La medición de la pobreza y la desigualdad en España, 1980-81*. Estudios Económicos**,** 42. Banco de España.

SEN, A.K. (1973). *On economic inequality*. Clarendon Press. Paperbacks. Oxford.

SEN, A.K. (1976). "Poverty: An ordinal approach to measurement ". *Econometrica*, 44, 219-231.

SEN, A.K. (1983). "Poor, relatively speaking". *Oxford Economic Papers*, 35 (2), 153-169.

SEN, A.K. (1985a). "A sociological approach to the measurement of poverty: A reply to Professor Townsend"**.** *Oxford Economic Papers*, 37(4), 669-676.

SEN, A.K. (1985b). *Commodities and Capabilities.* North Holland, Amsterdam.

SUBRAMANIAN, S. (2004). "Indicators of inequality and poverty"**.** *WIDER Research Paper* 2004/25. United Nations.

TOWNSEND, P. (1979). *Poverty in the United Kingdom: A survey of household resources and standards of living*. Harmondworths. Eng.: Penguin Books. Middlesex. 33.

TOWNSEND, P. (1985). "A sociological approach to the measurement of poverty: A rejoinder to Prof. Amartya Sen". *Oxford Economic Papers*, 37(4), 659-668.

UN MODELO TEÓRICO DEL MERCADO DE TRABAJO DE PERSONAS CON DISCAPACIDAD: APLICACIONES POLÍTICAS

CARLOS CAZALLAS ALCAIDE
E-mail: ccalc@once.es
Dpto. Apoyo al Empleo
Organización Nacional de Ciegos Españoles (España)

ANTONIO GARCÍA LIZANA
E-mail: aglizana@uma.es
Dpto. Economía Aplicada (Política Económica)
Universidad de Málaga (España)

Resumen

El mercado de trabajo de personas con discapacidad se caracteriza principalmente por una baja tasa de actividad y una elevada tasa de desempleo, siendo necesario poder explicar esta situación desde la perspectiva del análisis económico.
Por lo que en este artículo se va a describir un modelo teórico general, desde una perspectiva económica, siguiendo las principales teorías del mercado de trabajo, adaptándolas a la realidad socioeconómica del colectivo de discapacitados, con la finalidad de identificar y describir los factores que pueden estar incidiendo en la actividad y el empleo de las personas con discapacidad y que puedan ser utilizados como marco teórico de análisis y diseño de políticas económicas.

Palabras Clave: personas con discapacidad, mercado de trabajo, teoría neoclásica, teoría institucionalista.
Área temática: Economía Social y Discapacidad.

Abstract

The main feature of the disabled people Labour Market is the low rate of activity and the high rate of unemployment. This situation needs to be justified and studied under a perspective of an economic analysis.
This paper aims to explain a general theoretical model under an economic perspective, following the main theories of the Labour market applied to the social and economic conditions of disabled labour force. The aim is to identify and describe all the elements that might be influencing or interfering in the activity and employment of disabled people in order to set some theoretical guidelines for the analysis and design of economic policies.

Key words: disabled people, labour market, Neoclassical theory, Institucionalist theory.
Thematic Area: Social Economics.

1.- INTRODUCCIÓN[1]

Si se observa el mercado de trabajo de personas con discapacidad, en España, en una primera aproximación, con datos del Módulo de 2002 de la Encuesta de Población Activa, se advierten unas tasas de actividad y empleo más reducidas que para la población sin discapacidad, tal y como se recoge en el siguiente cuadro comparativo:

[1] El contenido de esta comunicación está basada en la investigación elaborada para Fundación ONCE, Estudio de investigación sobre ECONOMÍA, DISCAPACIDAD Y EMPLEO desde la perspectiva de la Economía Aplicada, y realizado en el marco del Programa Operativo de Lucha contra la Discriminación.

Tabla 1. Tasas de actividad, empleo y paro de personas con y sin discapacidad (%)

Con discapacidad	Ambos sexos	Varones	Mujeres
Tasa de actividad	33,7	39,3	27,1
Tasa de empleo	28,5	34,3	21,7
Tasa de paro	15,3	12,8	19,7
Sin discapacidad	Ambos sexos	Varones	Mujeres
Tasa de actividad	70,0	84,4	55,8
Tasa de empleo	62,4	78,1	48,7
Tasa de paro	11,0	7,5	16,2

Fuente: Encuesta de Población Activa, Módulo de 2002.

En la sociedad actual, la obtención y desempeño de un puesto de trabajo va a suponer, para cualquier persona, un hecho determinante, ya que implica disfrutar de independencia económica y potenciar su desarrollo humano, lo que va a favorecer su integración en la comunidad, generando un sentimiento de utilidad social y mejorando la autoestima del individuo.

Para las personas con discapacidad, todo ello es particularmente importante, pues la ocupación de un puesto de trabajo va a suponer la potenciación de unas capacidades y habilidades sociolaborales, al mismo tiempo que un medio de incorporación más efectiva y plena a la sociedad, fortaleciendo el desarrollo personal y profesional, gracias a la satisfacción que se deriva de estar ocupado laboralmente, las relaciones sociales que se establecen con el entorno y, principalmente, por la imagen positiva que el individuo genera de si mismo (Barton, 1998). En consecuencia, su adecuada inserción laboral es fundamental para que se alcance el objetivo de su integración social, siendo el escenario contrario el inicio de la exclusión.

Ahora bien, si tal incorporación laboral no se está produciendo, resulta importante conocer cuáles son las circunstancias concretas que arrojan como resultado una deserción tan evidente del mercado laboral. Podría ocurrir que tales circunstancias se correspondan con determinadas características de los interesados, que los estén alejando del mercado laboral; o que se encuentren en el comportamiento de los demandantes de mano de obra, generando obstáculos para su integración; o que estén en el propio mercado, en sus reglas de funcionamiento y en sus mecanismos de ajuste. Así, algunos autores advierten sobre eventuales limitaciones personales que reducen la empleabilidad de este segmento de la población (ver Fernández 2003, Cabra de Luna 2004 y Villa 2007); o bien, descargan en las actitudes familiares y las ayudas estatales elementos que desmotivan el interés de los afectados para participar en el mercado. Pero con frecuencia, y al margen de ello, hay quienes aprecian en la demanda laboral serias restricciones derivadas de ciertas actitudes, motivaciones y prejuicios que inciden en la menor contratación de personas con discapacidad. O bien, se culpa a las condiciones concretas del mercado de trabajo de la existencia de una serie de desigualdades, ya sean en el acceso al mismo, la contratación, los ascensos, los salarios o las condiciones de empleo, que pueden derivar en una situación de discriminación económica.

Por todo ello, en los siguientes apartados se va a exponer un modelo teórico del mercado de trabajo de personas con discapacidad, considerando la oferta laboral, demanda de trabajo y condiciones del mercado, integrando, fundamentalmente aunque no exclusivamente, los dos principales enfoques teóricos, neoclásico e institucionalista, ya que los principales elementos de ambos pueden ser recogidos en un modelo consolidado explicativo mediante una función de oferta, una función de demanda y su interacción en el mercado bajo diferentes restricciones a aplicar; para posteriormente poder ofrecer unas medidas de respuesta desde la Política Económica.

2.- LA OFERTA DE TRABAJO DE PERSONAS CON DISCAPACIDAD.

En primer lugar se debe considerar cuáles son los factores que van a influir para que una persona con discapacidad se muestre en situación de actividad o inactividad ante el mercado de trabajo, para lo que se debe tener en cuenta los siguientes modelos:

- El modelo básico de oferta microeconómica de trabajo (McConnell y otros, 2003) del que se deduce que el número de horas a trabajar depende directamente del salario, existiendo diversos factores que pueden modificar dicha curva de oferta de trabajo, como pueden ser los gustos o preferencias, la utilidad o desutilidad de trabajar, las circunstancias personales del individuo, las rentas salariales y no salariales, etc.
- El modelo del tiempo de Becker (1965) del que se deben considerar como elementos de análisis la perspectiva del hogar[2] y los usos múltiples del tiempo (tiempo de trabajo, tiempo de producción doméstica y tiempo de consumo)[3], siendo el principio de la ventaja comparativa el mecanismo por el que se asigna el tiempo (familiar) a sus diferentes usos.
- La inversión en capital humano de Becker (1975) también es un factor que influirá en las preferencias de los individuos para asignar un mayor número de horas al tiempo de trabajo, reduciendo su tiempo de ocio; es decir: que sean oferentes de trabajo en el mercado laboral.

Con todo ello, la persona con discapacidad va a maximizar la siguiente función de utilidad, sujeta a diferentes restricciones:

$$\text{Maximizar } U = U(C_{ih}, O_{ih}), \qquad [1]$$
$$pC_{ih} = wL_{ih} + M_{ih} + M_e,$$
$$M_e = wL_h + M_h - pC_h$$
$$L_{ih}+O_{ih}=24$$

donde C = consumo, O = ocio (computando como tal las horas no dedicadas a trabajo), w = salario, L = horas de trabajo, M = rentas no salariales, para ih = individuo (persona con discapacidad) perteneciente al hogar h, y h = hogar (unidad familiar). Siendo M_e = rentas excedentarias de la unidad familiar derivadas a la persona con discapacidad, definidas como la renta monetaria remanente del hogar, una vez cubiertas las necesidades de la unidad familiar en su conjunto (pC_h). La suma del tiempo dedicado al trabajo (L_{ih}) y el tiempo dedicado al ocio (O_{ih}) debe ser igual a 24 (horas del día).

Así pues, operando con el sistema anterior, quedaría:

$$\text{Maximizar } U = U(C_{ih}, O_{ih}), \qquad [2]$$
$$pC_i = w(L_{ih} + L_h) + M_{ih} + M_h - C_h$$
$$L_{ih}+O_{ih}=24$$

Del anterior modelo se pueden extraer las siguientes conclusiones relativas a posibles variables que estarían afectando a la actividad de las personas con discapacidad:

- Aspectos personales de la persona con discapacidad: A través de su mapa de preferencias se recoge la utilidad que genera el trabajo, siendo las circunstancias personales (edad, sexo, tipo de discapacidad, necesidades de atención médica y periodos de rehabilitación y descanso, estudios, experiencia profesional, etc.) las que van a determinar que el individuo valore en mayor o menor medida su tiempo de ocio y esté dispuesto a trabajar un mayor o menor número de horas.

[2] Es decir: la toma de decisiones depende de la unidad familiar, estando interrelacionadas las decisiones individuales entre sí.

[3] En el modelo de Linder (1970) se consideran cinco usos del tiempo: tiempo de trabajo, tiempo de trabajo personal, tiempo dedicado al consumo, tiempo para la cultura y tiempo inactivo. En una comparativa con el modelo de Becker, estos tres últimos pueden reflejarse en el tiempo de consumo de Becker, ya que suponen una utilización del tiempo que no reporta un mayor nivel de ingresos (tiempo de trabajo) o de producción doméstica (tiempo de trabajo personal).

- Aspectos de la unidad familiar: La presencia de otros individuos en la unidad familiar con rentas salariales o no salariales que sean suficientes para satisfacer las necesidades del hogar y detraer una renta excedentaria para la persona con discapacidad (M_e); o bien que, por sus condiciones de salud, estudios o experiencia profesional, posean una ventaja comparativa en la producción de bienes, según el modelo de Becker, puede también afectar a la decisión de participar en el mercado de trabajo por parte de la persona con discapacidad. Obviamente, en caso de $M_e<0$, o muy reducida, la inclinación a participar en el mercado laboral por parte de la persona con discapacidad será mayor.

- Por otra parte, al depender dicha renta excedentaria, derivada del hogar hacia la persona con discapacidad, del consumo y del trabajo (y, por tanto, del ocio) del hogar, también va a influir en la decisión individual de la persona con discapacidad la función de utilidad entre consumo (C_h) y ocio (O_h) agregada de todas las personas que componen el hogar o unidad familiar.

- La inversión en capital humano de la persona con discapacidad merece una atención diferenciada, ya que si bien va a depender directamente de circunstancias individuales y familiares de diversa naturaleza, es interesante recordar que según el modelo de inversión en capital humano, una menor perspectiva de rentabilidad de dicha inversión[4] pueden conducir a las personas con discapacidad a no obtener formación o estudios cualificados en análoga proporción que el resto de los miembros de la familia, lo que repercutirá tanto en la función de preferencias, penalizando el tiempo de trabajo como ámbito de satisfacción y realización personal, como en la mayor ventaja comparativa para la producción de bienes de los otros familiares.

- Rentas no salariales: Las prestaciones económicas concedidas como consecuencia de sufrir una discapacidad, las cuales suponen un incremento especial de $M_{i,}$ permiten una elevación de las posibilidades de consumo y ocio, lo que puede conducir a un menor número de horas trabajadas para un nivel de salario similar; o, con otras palabras, a establecer un salario de reserva más alto a la hora de ofrecer trabajo en el mercado laboral.

De la maximización de la utilidad de la persona con discapacidad, se va a deducir la siguiente ecuación de oferta de trabajo:

$$L_i = f(w) \text{ para } |M_i \text{ constante.} \qquad [3]$$

Al igual que en el modelo microeconómico básico de oferta de trabajo, la curva de oferta va a depender del salario de equilibrio, que entrará en directa competencia con los demás oferentes de mano de obra, ya sean personas con discapacidad o sin discapacidad. La diferencia fundamental es que, suponiendo que las preferencias y las rentas no salariales fueran idénticas entre las personas con discapacidad y sin discapacidad, en la oferta de las primeras la incidencia de rentas excedentarias derivadas por el hogar o unidad familiar hacia las mismas (como consecuencia de los hábitos sociales existentes) ocasionará un desplazamiento de la curva de oferta de trabajo hacia arriba, provocando (como en el caso de las prestaciones por discapacidad) que para un mismo número de horas de trabajo el salario de reserva, por el cual se está dispuesto a trabajar, sea mayor.

Esto puede suponer, que las personas con discapacidad podrían quedarse fuera del mercado de trabajo, ya que el salario de equilibrio del mismo tenderá a ser inferior a su salario de reserva. Sin embargo, de acuerdo con la teoría del mercado dual de trabajo (Piore 1975 y Doeringer y Piore 1985), las personas con discapacidad estarían participando del mercado secundario, donde estarían aceptando salarios más bajos que en el mercado ordinario; lo que parece sugerir un salario de reserva inferior al salario de equilibrio.

Dicha controversia puede resolverse, de momento, mediante el establecimiento de dos subgrupos dentro del colectivo de personas con discapacidad. El primer subgrupo estaría

[4] Ya sea porque otros individuos del hogar al rentabilizar más la inversión del capital humano, los recursos escasos se desvíen a éstos, ya sea porque las dificultades existentes en el mercado de trabajo, les desanimen a acumular capital humano.

compuesto por aquellos cuyo salario de reserva es alto, como consecuencia de rentas no salariales o de rentas excedentarias derivadas por el hogar, que estarían desanimando al discapacitado, ya que el salario que percibiría en el mercado sería inferior a su salario de reserva.

Un segundo subgrupo sería aquel que no percibe rentas no salariales o que su unidad familiar no puede satisfacer sus necesidades de consumo y ocio (derivadas de su función de utilidad), teniendo que acudir al mercado de trabajo para conseguir un empleo remunerado. Dada la existencia de una segmentación del mercado, las características del mercado secundario, condicionadas por una demanda de algún modo sesgada, obligarían a la aceptación de retribuciones inferiores al salario de equilibrio general. De todos modos, dentro de este subgrupo podrían establecerse igualmente dos conjuntos, uno de personas con discapacidad cualificadas y otro de personas discapacitadas no cualificadas. Al tener que acceder al mercado de trabajo, las personas cualificadas tienen mayores oportunidades para estar participando en el mercado primario, mientras que las personas con discapacidad no cualificadas estarían concurriendo al segmento de trabajo secundario de la economía, tal y como se establecía en la teoría institucionalista del mercado dual de trabajo.

Aunque pudiese también suceder que las personas con discapacidad cualificadas tuviesen o hubiesen tenido rentas no salariales, derivadas de prestaciones económicas o de rentas excedentarias de la unidad familiar, y por tanto con un salario de reserva alto, si al mostrarse activos en el mercado de trabajo el salario a percibir fuese superior a éste, como consecuencia de su formación, la decisión racional sería participar igualmente en el mercado laboral.

Cabe destacar, en este sentido, que tal eventualidad plantea consecuencias interesantes desde el punto de vista de la actuación política, con respecto a la teoría del capital humano y su adquisición o inversión por parte de una persona con discapacidad, ya que si a priori puede parecer que su tasa de retorno interno (TIR) sea inferior al de una persona sin discapacidad, por las dificultades que pudiera derivarse de la discapacidad durante la vida laboral, a nivel individual siempre será preferible realizar la inversión en capital humano frente a no realizarla, en relación con la posibilidad de acceder al mercado primario de trabajo y percibir una remuneración por encima del salario de reserva. Obviamente, las expectativas existentes con respecto a esta eventualidad van a influir en las decisiones concretas que en la práctica se tomen, lo que debe ser tenido en cuenta a la hora de afrontar el diseño de medidas orientadas a modificar la situación.

Desde esta perspectiva, la actuación política podría ir dirigida a modificar las pautas de conducta familiar, de acuerdo con las cuales se toman las decisiones sobre inversión en capital humano en la familia desde el punto de vista de la ventaja comparativa que representa, comparando entre los individuos del hogar. En cambio, debería plantearse el análisis económico de lo que supone para la persona con discapacidad la inversión en capital humano en términos de utilidad personal y de lo que supone para la unidad familiar en términos de reducción de la renta excedentaria a derivar hacia el miembro con discapacidad. En este caso, aun cuando el valor actual de dicha inversión en capital humano fuese cero o negativo, el hecho de que el trabajador con discapacidad pueda acceder más fácilmente a un empleo, aunque sea en el mercado secundario de baja remuneración, va a reducir el montante de renta excedentaria de la familia que debe ser comprometido para la persona con discapacidad, lo que puede compensar parcialmente la rentabilidad negativa de la inversión. Todo ello con independencia de la utilidad positiva que el trabajo genera para la persona con discapacidad, en términos de autonomía personal e integración social. Lo cual, obviamente, nos lleva de nuevo al peso de los hábitos sociales como condicionantes de las actitudes y del comportamiento familiar, a la hora de determinar las estrategias de actuación, y, por tanto, de las preferencias agregadas entre consumo y ocio.

Con todo lo anterior, se podría establecer que la oferta de trabajo de las personas con discapacidad va a estar dependiendo, por tanto, del salario, de sus preferencias entre consumo y ocio, su capital humano adquirido, las rentas no salariales y la renta excedentaria de la unidad familiar derivada a la persona con discapacidad (que a su vez dependerá del consumo y ocio del hogar y de las rentas no salariales del hogar). Es decir:

$$L_{ih} = F[w, C_{ih}, O_{ih}, h_{ih}, M_{ih}, M_e(C_h, O_h, M_h)]. \quad [4]$$

Siendo h_i el capital humano.

3.- LA DEMANDA DE TRABAJO DE PERSONAS CON DISCAPACIDAD.

La demanda de trabajo es considerada como una demanda derivada de la demanda del producto (McConnell y otros, 2003); dependiendo, por consiguiente, de la productividad marginal del trabajo y del valor de mercado del bien o servicio. Todo ello deriva en que la empresa va a demandar factor trabajo dependiendo directamente del salario de mercado; demanda que se cruzará en el mercado con la oferta laboral. Sin embargo, es fundamental advertir que no puede aplicarse en nuestro caso inicialmente el supuesto de homogeneidad del factor trabajo, ya que surgen diferentes elementos para considerar dos grupos diferenciados, personas con discapacidad y sin discapacidad. Es decir: la demanda va a interaccionar en un entorno de heterogeneidad del factor trabajo.

Asimismo, del estudio del modelo microeconómico básico de demanda de trabajo, se deben recoger una serie de factores que podrían causar modificaciones o desplazamientos de la curva de demanda laboral. Siendo los siguientes de vital importancia, por su directa relación con el trabajo en sí, y que pueden determinar la elección de un trabajador con discapacidad o sin discapacidad:

- El ratio entre los costes laborales y los costes totales. En este sentido, los periodos de ausencias o bajas médicas del trabajador con discapacidad, los eventuales costes de adaptación del sistema de producción a las particulares exigencias de cada tipo de discapacidad, etc., provocarían un efecto negativo en dicho ratio. Obviamente, desde las perspectivas de la acción política podrían compensarse tales costes (mediante bonificaciones, subvenciones y desgravaciones); pero si éstas no se establecen adecuadamente, el efecto final sobre la demanda de trabajo resultará indeterminado.

- La posibilidad de sustitución por otros factores de producción. Entiéndase como la posibilidad de sustituir factor trabajo con discapacidad por factor trabajo sin discapacidad. Dicho intercambio puede resultar más fácil que el inverso. Además, la reducida proporción que el colectivo de personas con discapacidad representa en comparación con el de personas sin discapacidad no va a facilitar, por razones puramente matemáticas, la sustitución inversa señalada.

- La productividad. Obviamente, la decisión empresarial se inclinará preferentemente por la mano de obra que resulte más productiva. Dicha productividad está vinculada con el desempeño laboral, la formación y la experiencia. Pero la formación y experiencia no son activos que beneficien en demasía a las personas con discapacidad. Además, y en igualdad de circunstancias, a priori se puede tender a pensar que el desempeño de un puesto por una persona sin discapacidad es superior, olvidando que con la formación y experiencia necesaria para el puesto de trabajo y una adecuada adaptación del mismo, la productividad de la mano de obra con discapacidad puede ser equiparable a la del resto de la población, o incluso, en ciertos casos, aventajarla. Ello está condicionado por el desconocimiento de las habilidades sociales, laborales e instrumentales del colectivo de trabajadores con discapacidad por parte de los empresarios, fruto de una posible información asimétrica del mercado (Spence, 1973); así como por eventuales prejuicios que pueden distorsionar la percepción objetiva de la realidad.

De todos modos, las importantes desigualdades existentes en la práctica, incluso en igualdad de desempeño, formación y experiencia laboral, es lo que lleva a plantear la posible existencia de discriminación económica[5] en el mercado de trabajo de personas con discapacidad. Así pues, siguiendo el modelo del gusto por la discriminación de Becker (1957), los empresarios estarían

[5] McConnel, Brue y MacPherson (2003) definen discriminación económica cuando existiendo la misma capacidad (para el desempeño del puesto de trabajo), nivel de estudios, formación y experiencia, se recibe un trato inferior en la contratación, en el acceso a una ocupación, los ascensos, el salario o las condiciones de trabajo.

asumiendo un coste subjetivo a través del coeficiente de discriminación D, por el que no contratarían a personas con discapacidad, a menos que los salarios fuesen suficientemente bajos para compensar dicho coste.

Es decir, si consideramos: w_x, salario unitario o coste con el que el empresario discriminador remunera a un trabajador X perteneciente al grupo no discriminado; y $w_y + D$, salario o coste que le supone la contratación de un trabajador Y perteneciente al grupo sobre el que tiene prejuicios; entonces, tal demandante de empleo sólo contratará a trabajadores susceptibles de sufrir discriminación o prejuicios por el empresario si:

$$w_y + D < w_{x,} \qquad [5]$$

o lo que es lo mismo, el precio máximo que está dispuesto a pagar por un trabajador del grupo discriminado será:

$$w_y = w_x - D. \qquad [6]$$

No obstante, en un sistema de percepción de bonificaciones o subvenciones a la contratación de personas con discapacidad, dicho coeficiente de discriminación debería quedar parcial o totalmente anulado, con lo que Reformulando la ecuación anterior tendríamos:

$$w_y = w_x - D + B. \qquad [7]$$

Donde w_x = salario de la persona sin discapacidad, w_y = salario de la persona con discapacidad, D = coeficiente de discriminación y B = bonificación a la contratación. Consecuentemente, se deduce:

- Las bonificaciones o subvenciones a la contratación deben cubrir el coste subjetivo o coeficiente de discriminación D.
- Si aun cubriendo dicho coste, el desempleo persiste, debería plantearse la existencia de otras explicaciones alternativas, como barreras psicosociales (Fundación Encuentro, 2001) o información asimétrica que conducirían a la no contratación de discapacitados (Spence, 1973 y Weiss, 1996).

Debemos considerar también un segundo modelo, de la teoría de la discriminación estadística propuesto por Phelps (1972) y que consiste en que los empresarios, con el fin de reducir los costes de contratación, estarían juzgando a una persona con discapacidad en función de las características medias del colectivo de discapacitados y no en función de sus propias características personales. Los juicios serían correctos, reales u objetivos, en el sentido de que el grupo pudiera tener, de hecho, las características que se le atribuyen; pero son incorrectos con respecto a esa persona con discapacidad.

Como conclusión, se observa que los demandantes de trabajo, en general, no contratan a personas con discapacidad, que pueden desempeñar un puesto de trabajo acorde con su formación y experiencia con una productividad adecuada, por factores, tales como la falta de información, barreras psicosociales, desconocimiento de las habilidades sociolaborales del colectivo, posibles perfiles a demandar, etc, todos ellos agrupables bajo una misma variable: la información asimétrica (Spence, 1973). Con lo que se podría reformular la ecuación [12], para introducir dicha variable como un coste de oportunidad:

$$w_y = w_x - D - I_a + B. \qquad [8]$$

Donde I_a va a ser el coste de la información que no se puede obtener del mercado y que resulta de gran relevancia a la hora de contratar a una persona con discapacidad.

Por tanto, la demanda de trabajo de personas con discapacidad va a estar en función del salario, del coeficiente de discriminación (o prejuicios) de los empresarios, el coste de la información, los costes laborales específicos (CL_{in}) atribuidos a las personas con discapacidad i, en relación con el conjunto de la plantilla n y las bonificaciones y subvenciones.

$$L_d = F(w, D, I_A, B, CL_{i,}) \qquad [9]$$

Frente a tal situación, la respuesta política puede consistir en la modificación, incremento o supresión de las bonificaciones existentes para compensar los diferentes tipos de costes señalados e incentivar las decisiones de los demandantes de empleo; o bien en la aplicación de otras medidas más contundentes, que persigan la reducción o eliminación de los costes, etc.

4.- LAS CONDICIONES DEL MERCADO DE TRABAJO DE PERSONAS CON DISCAPACIDAD.

Desde la perspectiva del análisis microeconómico, es en el mercado de trabajo donde se va a cruzar la oferta de trabajo de personas con discapacidad [4] y la demanda de trabajo de personas con discapacidad [9]. Desde un punto de vista convencional, dicho mercado de trabajo de personas con discapacidad formaría parte del mercado laboral general, que a priori se caracteriza por los siguientes elementos:

- Un gran número de empresas que compiten entre sí para cubrir puestos de trabajo idénticos,
- Numerosas personas con una misma cualificación que ofrecen independientemente sus servicios de trabajo,
- Conducta salario-aceptante por parte de la oferta y de la demanda de trabajo,
- Una movilidad perfecta y sin coste,
- Transparencia del mercado.

El punto de equilibrio será aquel donde se corten las curvas de oferta y demanda, correspondiéndose con el salario en el que se contrata una determinada cantidad de trabajo, de tal modo que cualquier excedente de mano de obra siempre podría ser absorbido por el mercado como consecuencia de la aceptación del nivel de salarios pertinente, por bajo que sea. No obstante, observamos que existen múltiples factores que van a distorsionar el mercado, afectando a las características enumeradas, eliminando la transparencia del mismo (supuesto básico en las teorías neoclásicas), la movilidad perfecta y sin coste, el comportamiento independiente de las diversas unidades individuales que ofrecen y demandan trabajo, el comportamiento salario-aceptante...; todo lo cual incide claramente sobre el funcionamiento regular y la asignación de las cantidades de trabajo, por lo que es previsible la existencia permanente de desajustes internos entre oferta y demanda, provocando los siguientes efectos:

- Excesos de oferta laboral (paro).
- Expulsión o efecto desánimo de una parte de la oferta (inactividad).
- Retraimiento de la demanda de trabajo.

Realmente, la distorsión del mercado, que lleva a los efectos comentados, pueden ser achacada a factores institucionales, tal y como plantean las teorías institucionalistas del mercado de trabajo[6], como las interferencias políticas (con el establecimiento de niveles mínimos de salario, el establecimiento de controles y regulaciones inadecuados...), el papel de sindicatos y organizaciones empresariales, la existencia de mercados internos, etc.

Ahora bien, en la medida en que observamos en la práctica unos resultados diferentes entre los segmentos de personas con y sin discapacidad, es preciso aceptar determinadas peculiaridades dentro del mercado, susceptibles de explicar tal hecho.

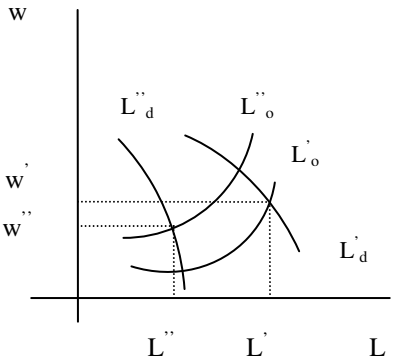
Tal como ha sido analizado, existen unas especificaciones propias de la oferta y demanda de trabajo de personas pertenecientes al colectivo de estudio, que las diferencian del caso general; lo que supone admitir la presencia de dos mercados con un comportamiento diferentes, de tal modo que las funciones correspondientes estarían desplazadas hacia la izquierda, de tal modo

[6] De hecho, algunos de los factores recogidos por las mencionadas teorías, tales como las bonificaciones o subvenciones a la contratación laboral, las rentas no salariales procedentes de la protección social existente, la información asimétrica, la discriminación, etc., han sido incorporados a las funciones de oferta y demanda, ya que son cuantificables y afectan de manera directa a las mismas, bien desplazándolas, modificando su elasticidad, etc.

que la posición de equilibrio se resolvería con un volumen menor de contratación y, previsiblemente, aunque no necesariamente, con menores salarios (ver figura 1). Si a ello unimos la heterogeneidad en la composición de la oferta de trabajo de personas con discapacidad, o la existencia de otros factores de carácter institucional, además de los que han sido señalados, que influyen sobre este segmento del mercado de manera muy especial, no puede extrañar el dispar comportamiento observado. Así, podemos señalar algunos de esos factores, como los siguientes:

- La regulación normativa del empleo,
- El entramado social de la discapacidad,
- Las costumbres y los hábitos sociales.

Figura 1



Pero, además, si incorporamos los condicionamientos macroeconómicos sobre el mercado de trabajo, podremos percibir nuevos elementos que favorecen en poco la situación del segmento laboral que nos ocupa. Evidentemente, dicha situación resulta menos problemática en un ambiente de expansión, si bien las condiciones laborales que han sido establecidas anteriormente, harán bascular las preferencias del mercado hacia las personas sin discapacidad en primer lugar, alcanzando a las personas con discapacidad más tarde. Pero sería en el supuesto de que aparezcan dificultades en la economía cuando la situación se convierta en más crítica. Desde el punto de vista de la Economía de demanda, la caída de las expectativas como consecuencia de la contracción de las mismas se traduce en una reducción del empleo; pero ello, dadas las características de las funciones de oferta y demanda de mano de obra analizadas, incidirá con mayor fuerza en las personas con discapacidad; las reducciones salariales o las disminuciones de contratación repercutirán con mayor fuerza. Desde el punto de vista de la Economía del lado de la oferta, los menores rendimientos esperados por los empresarios o los mayores costes laborales imputados a las personas con discapacidad, las sitúa en una posición particularmente vulnerable para prescindir de las mismas, como un modo de corregir, en alguna medida, las dificultades que en tales capítulos afectan a las empresas.

5. LA RESPUESTA POLÍTICA

Habiendo analizado el comportamiento del mercado laboral y confirmado desde el punto de vista racional las dificultades existentes para la plena integración en el mismo de las personas con discapacidad, sin que parezca factible, por otro lado, la resolución espontánea de dicha situación, corresponde ahora diseñar las respuestas políticas oportunas que cabe aplicar, desde el sector público y el privado, para afrontar tal estado de cosas, actuando sobre los problemas existentes en sí; pero también sobre los factores explicativos de los mismos, a tenor del análisis efectuado, de manera que se propicie el desplazamiento de las funciones de oferta y demanda de mano de obra con discapacidad, para corregir la situación.

5.1. Diagnóstico teórico de la situación.

Aun cuando no se trate de un diagnóstico completo, ya que no se basa en la observación sistemática de la realidad y en la interpretación de los valores y signos encontrados, resulta de gran utilidad establecer un diagnóstico previo de carácter racional, a partir del análisis teórico de la situación efectuado en los apartados precedentes, identificando los principales problemas que deben ser abordados, para encontrar respuestas suficientes, ante la debilidad de resultados que se han alcanzado hasta la fecha.

a) Limitaciones objetivas para la actividad económica de las personas con discapacidad.
b) Hábitos sociales existentes con respecto al empleo de las personas con discapacidad.
c) Efecto desmotivador de las prestaciones públicas asistenciales.
d) Proteccionismo familiar.
e) Costes elevados de acceso al mercado de las personas con discapacidad.
f) Costes imputables al puesto de trabajo.
g) Capital humano deficiente.
h) Expectativas laborales negativas.
i) Aversión al riesgo.
j) Prejuicios existentes.
k) Desconocimiento de la situación real, como consecuencia de la falta de información.
l) Tecnología inadaptada a las condiciones de las personas con discapacidad.
m) Desigualdades internas en la población potencialmente activa con discapacidad.
n) Asimetría de género, que se traduce en unas condiciones más desfavorables por tal motivo de la población femenina.

5.2. Algunas medidas de política económica.

El diagnóstico efectuado sitúa en el foco de atención algunas cuestiones que deberían ser abordadas por la política, teniendo en cuenta el fin general que debe ser perseguido: el incremento de la equidad dentro de la sociedad. En concreto, nos preocupa incrementar dicha equidad en el ámbito laboral, mejorando en dicho sentido el acceso al empleo de un segmento de la población particularmente desfavorecido. El objetivo final, por tanto, sería el aumento del empleo; llamando la atención sobre una variable intermedia particularmente significativa en este sentido: la tasa de empleo de las personas con discapacidad. En la medida en que dicha tasa de empleo aumente, mejorará el empleo del conjunto de la población, incrementándose en consecuencia la equidad. En tales circunstancias, podemos proponer las siguientes actuaciones políticas pertinentes, identificando las medidas a adoptar, a la vista de los problemas inmediatos que han sido señalados. Dichas actuaciones se recogen, de manera sucinta, en relación con cada uno de los problemas básicos diagnosticados, que se encontrarían en el origen de la situación existente.

a) Limitaciones objetivas a la actividad económica. Algunas de estas limitaciones son de carácter legal, y están vinculadas con la percepción de determinadas prestaciones públicas, por lo que su eventual reducción iría vinculada al tratamiento de esta variable. En el caso de las discapacidades más severas o en situaciones de discapacidades añadidas, podría apostarse, mediante políticas fiscales, monetarias y directas (desde los centros públicos especializados), por impulsar los avances de la tecnología y de la medicina para reducirlas, mejorando las condiciones de la población afectada para facilitar su incorporación al mundo laboral. Pero conviene advertir que nos encontramos fundamentalmente con un problema de delimitación estadística, que puede distorsionar las comparaciones con el segmento de las personas sin discapacidad, si pretenden obtenerse conclusiones de carácter estrictamente económico. Es evidente que los casos en que no tiene sentido buscar explicaciones en el funcionamiento del mercado de trabajo. Sencillamente no son activos, desde el punto de vista económico, porque no pueden serlo. Una correcta evaluación de la situación exigiría disponer de información sobre este particular[7]. Una cuestión diferente es la de

[7] Obsérvese que en el propio R.D. 2274/85, por el que se regulan los centros ocupacionales, se habla de que los mismos deben tender a favorecer la futura incorporación de las personas con discapacidad al trabajo, pero siempre que sea *posible*.

considerar que tales personas deben contar con los pertinentes sistemas de apoyo, lo que evidentemente supondría un estímulo para la creación de empleo.

b) Hábitos sociales: La modificación de las actitudes, motivaciones y usos sociales resulta clave en cualquier estrategia que pretenda modificar la oferta laboral. La cuestión estaría en determinar hasta qué punto se está dispuesto a avanzar en este camino. En la medida en que una persona con discapacidad se sienta legitimada por la sociedad para no concurrir al mercado laboral, y sea acreedora frente a la colectividad de determinados derechos reconocidos por su situación personal concreta, difícilmente se verá estimulada para participar en la población activa. A menos que la necesidad de sobrevivir la obligue a ello.

c) Prestaciones públicas asistenciales. Pero no sólo es un problema de sensibilización. Los hábitos sociales pueden verse modificados por las prácticas políticas, económicas y sociales reiteradas durante un tiempo, pero, además, tales prestaciones afectan directamente a la dependencia condicionada, en la medida en que vinculan la satisfacción de las necesidades a la generosidad externa y no al propio esfuerzo (con independencia de que ello sea legítimo, y, además, un derecho reconocido), limitando los incentivos para trabajar, sobre todo si el trabajo lleva aparejada la pérdida de las prestaciones. Aun cuando las expectativas de ingresos pudieran ser mayores en el mercado laboral, la "aversión al riesgo" podría jugar un papel fundamental para permanecer percibiendo los subsidios públicos, antes que aventurarse en un terreno no suficientemente conocido, y que, evidentemente, puede arrastrar la pérdida de toda clase de ingresos si la aventura finalmente fracasa. Es obvio que, aún así, no tiene sentido suprimir las prestaciones sociales; pero sí revisarlas profundamente, haciéndolas compatibles (e incluso incrementándolas) con la presencia en el mercado laboral del beneficiario (siempre que para éste sea factible, de acuerdo con sus concretas características personales). Igualmente, el establecimiento y reforzamiento (según los casos) de incentivos específicos vinculados con la inserción laboral del beneficiario, o con la preparación inmediata para el trabajo, de carácter fiscal, monetario o institucional, etc., podrían jugar un buen papel en este terreno.

d) Proteccionismo familiar. Posiblemente una de las piedras de toque fundamentales en relación con el tema que en este momento nos ocupa. Por lo que la familia puede representar como soporte financiero para atender las necesidades de la persona con discapacidad (lo que podría ser una versión modificada del punto anterior), por supuesto; pero, sobre todo, por las actitudes de sobreprotección, vinculadas con aversión al riesgo, en definitiva, que puede representar para el propio afectado y para el conjunto de la familia que aquél salga de casa, etc. Obviamente, ello se encuentra muy influido por los hábitos sociales existentes y encuentra cierto respaldo y legitimación con determinados tipos de prestaciones públicas; por lo que, en realidad, habría que remitir a uno y otro ámbito el tipo de respuestas a dar, considerando el papel de las familias como una cuestión específica a tener en cuenta en el diseño de las referidas políticas.

e) Costes de acceso al mercado. De todos modos, el trabajo en el ámbito de la dependencia condicionada fracasará con muy poco remedio, si no se abordan con rotundidad las asimetrías en la oferta laboral. Abaratar los costes de acceso (información, movilidad, empleo de tiempo, etc.), contribuyendo a la transparencia del mercado, y favoreciendo tanto la recepción de la información significativa como la transmisión de la misma, puede ser crucial para un mejoramiento de la situación. Podrían, ciertamente, compensarse tales costes mediante subsidios y subvenciones en la línea indicada en c), de manera que se incentivara el acceso al mercado; pero, igualmente, sería importante potenciar los canales existentes que faciliten dicho acceso, atendiendo a las circunstancias específicas de los subsegmentos más relevantes de personas con discapacidad; pues, parece obvio, que las necesidades y dificultades (y, por tanto los costes) son diferentes según los casos. Medidas de carácter institucional (creación o modificación de organismos e instituciones especializadas; formación de personal específicamente capacitado, más allá de las rutinas administrativas; etc.) podrían ser muy valiosas.

f) Costes imputables al trabajo. También en este punto la problemática es muy diferente, según el subgrupo de población; e, incluso, según cada persona; por lo que habría que modular detenidamente las actuaciones a acometer, teniendo en cuenta tales circunstancias. Las medidas podrían ir dirigidas tanto a reducir dichos costes, buscando alternativas, por ej., como a compensarlos. En el primer caso, puede ser importante facilitar la incorporación de una manera ágil y económicamente factible de recursos técnicos, métodos de trabajo, etc., que faciliten las tareas

laborales, mediante incentivos tales como desgravaciones fiscales, subvenciones, facilidades crediticias..., completado con políticas de formación de capital humano orientado hacia el manejo de dichos recursos. La compensación de dichos costes, para evitar algunos de los problemas señalados más arriba, sólo debería plantearse en la medida en que los recursos técnicos no estén disponibles, o sean insuficientes, mediante complementos monetarios por parte del sector público que refuercen los salarios percibidos, lo que permitiría modificar la curva de oferta, reduciendo en la práctica el salario de reserva.

g) Capital humano. Aumentarlo depende tanto de la educación formal como de la experiencia laboral; aunque si se incluyen dentro del concepto de capital humano cuestiones como salud, actitudes, valores, relaciones sociales, nos encontraríamos con un espectro más ambicioso. De todos modos, limitándonos al ámbito del conocimiento, y teniendo en cuenta que la experiencia laboral sólo puede venir del éxito en el conjunto de las políticas aplicadas, la cuestión estaría en cómo mejorar la educación formalizada de las personas con discapacidad. Posiblemente, la presencia de análogos costes que el caso del ámbito laboral están aquí presentes, por lo que cabría aplicar respuestas similares. Pero no debe perderse de vista la influencia de las expectativas laborales como incentivo para el estudio (las cuales van unidas a la experiencia laboral); así como de algunas otras variables ya tomadas en cuenta, como la protección familiar o los hábitos sociales. Una cuestión adicional, con todo, que debe ser subrayada es la atención específica a cada tipo de discapacidad, según sus necesidades concretas. Así como tener en cuenta los perfiles de trabajo demandados, y su adecuación con las características personales de las personas en cuestión. Todo esto puede abonar la oportunidad de un debate profundo acerca de los modelos educativos más idóneos para reforzar el capital humano bajo esta perspectiva, tanto en la etapa infantil como a lo largo de toda la vida. Tratándose de un problema que debe implicar necesariamente a los especialistas en tales cuestiones, nos limitamos a apuntarlo, subrayando la utilidad de una respuesta generalizada en este sentido, que modificaría la situación, claramente, incidiendo, incluso, en otras variables ya consideradas. De nuevo, políticas de carácter institucional, reforzadas por medidas fiscales y monetarias podrían contribuir significativamente a los objetivos intermedios planteados con respecto a esta variable.

h) Expectativas laborales. Las mismas están condicionadas por la imagen que cada persona tiene de sí misma y de las propias competencias para entrar en el mercado, así como del conjunto de las personas con discapacidad (o del grupo concreto de discapacidad con el que se identifica), de la experiencia propia y/o de otras personas con discapacidad, así como de la situación económica general. La propia experiencia personal es un dato de partida, que se podrá ir modificando en la medida en que la política dé sus frutos; algo parecido habría que decir de la experiencia compartida con otras personas con discapacidad. De todos modos, es posible que la percepción de tales experiencias, y, por supuesto, de la propia imagen y la del colectivo esté distorsionada como consecuencia de los hábitos sociales. Una adecuada política informativa puede ser fundamental en estas circunstancias, mientras que se modifican las propias expectativas como consecuencia de los resultados que se vayan logrando.

i) Aversión al riesgo. La aversión al riesgo, en cuando condicionante de las decisiones empresariales, suele reforzar las oportunidades de aquellos trabajadores ya incorporados al sistema o que ofrecen perfiles más seguros; pero lo primero no favorece la contratación de mano de obra con discapacidad, dada su menor presencia en los empleos ya consolidados; ni lo segundo, en la medida en que los prejuicios o las lagunas de información, la colocan en desventaja. Ante este panorama, podría optarse por una política de compensación de riesgos, recurriendo a subvenciones, desgravaciones fiscales, bonificaciones en la seguridad social, y otras análogas, que suelen ser recursos ampliamente utilizados. Pero ello, como se ha señalado más arriba, puede reforzar la imagen negativa de las personas con discapacidad; o pensar en su contratación sólo por el beneficio inmediato que supone, sin plantearse una política de asignación de recursos consistente, que permita sacar el provecho óptimo de las cualidades de las personas contratadas, con las consecuencias negativas que ello puede representar sobre su autoestima y sobre la propia experiencia laboral personal. La reserva de puestos de trabajo posiblemente se vea condenada al fracaso, como solución de amplio alcance, ya que los temores empresariales pueden hacerles preferir el pago de una sanción al riesgo de contratar en tales condiciones. En todo caso, dicha contratación puede verse más como una carga que afrontar, del mismo modo que los impuestos o

las medidas medioambientales; sin que ello, en tales circunstancias, pueda ser plenamente beneficioso para los propios interesados. Por el contrario, tal política podría generar hostilidad entre los compañeros de trabajo, y, sobre todo, entre las personas que se han visto desplazadas por el hecho de las cuotas. En resumen, sería de gran interés poder evaluar hasta qué punto dichas políticas contribuyen o no a mejorar la productividad de las empresas, o, por el contrario, a deteriorarla; pues ello sí que haría un flaco favor a la modificación de la situación. Más bien convendría afrontar la existencia de prejuicios y falta de conocimiento que pueden confluir con dicha aversión, para intentar modificar la situación.

j) Prejuicios. La existencia de prejuicios, que pueden ir desde la prevención ante un perfil concreto del oferente de trabajo que no resulta atractivo (sin generalizarlo a otros individuos) hasta la segregación sistemática por el mero hecho de la diferencia que incorporan tal tipo de personas, es, evidentemente, un elemento de discriminación que influye claramente en las restricciones de la demanda. Planteado de este modo, su incidencia puede ir más lejos de las meras expectativas de rendimiento, afectando directamente a la demanda laboral de personas con discapacidad, aun asumiendo costes por tal determinación, de acuerdo con el modelo del gusto por la discriminación de Becker. Ante un panorama extremo, sólo cabe una política sistemática de sensibilización que pudiera modificar tales prevenciones; con independencia de las penalizaciones en que pudiera incurrir el discriminador si las leyes persiguen tal tipo de conducta. No obstante, no parece razonable trasladar el gusto por la discriminación planteado hasta esos extremos al comportamiento de la demanda con respecto a las personas con discapacidad, de un modo significativo, desde el punto de vista del funcionamiento económico general. Sí, por el contrario, puede estar muy extendida la existencia de prejuicios puntuales que hacen percibir a los potenciales trabajadores con discapacidad menos rentables para la empresa, bien por su esperada menor productividad, por las reacciones previsibles de los clientes, por los mayores costes de aprendizaje y adaptación que se supone acarrean, etc. Las alternativas políticas serían similares a las indicadas en el punto anterior, con las mismas condicionantes que ahí se planteaban. En la medida en que los prejuicios puedan estar relacionados con la falta de información y de conocimiento suficiente, el tipo de respuesta nos remite al punto siguiente.

k) Desconocimiento de la situación. Como en otro lugar hemos visto, se suele atribuir parte al menos de responsabilidad en la existencia de restricciones de la demanda a la falta de información sobre las ofertas de las personas con discapacidad y a los limitados conocimientos de los demandantes sobre sus verdaderas posibilidades. De una parte, tal planteamiento apunta a las asimetrías de oferta, ya analizadas, y que tienen ahí su respuesta. Pero hay elementos que trascienden dichas asimetrías y que se alimentan de actitudes preconcebidas o, realmente, de una información sesgada y de un conocimiento limitado o parcial de la situación, generalizando a todo el conjunto de la población con discapacidad experiencias puntuales (como ocurre en el modelo de la discriminación estadística; aunque no exclusivamente). De nuevo se puede pensar en fórmulas compensatorias de los menores rendimientos esperados anejos a ese desconocimiento; se puede forzar la situación obligando a una cuota mínima de contratación (que rara vez se verá superada de un modo generalizado); o se puede plantear algún tipo de campaña de información específicamente diseñada, facilitar el acceso a fuentes de datos fidedignas, incentivar los contratos en prácticas de personas con discapacidad con el propósito de impulsar su introducción y conocimiento en el mercado, etc.; así como potenciar servicios de asesoramiento e información sobre las ventajas comparativas de determinados perfiles personales para desempeñar concretos puestos de trabajo.

l) Tecnología inadaptada a las condiciones de las personas con discapacidad. Uno de los problemas más frecuentes que condicionan las expectativas de rendimiento son los derivados de una tecnología inadaptada a las condiciones concretas de las personas con discapacidad, lo cual sí repercute, efectivamente, en el rendimiento de las mismas. La eventual apuesta por la adaptación del puesto de trabajo genera costes para la empresa, lo que igualmente repercute en el rendimiento. En este caso se trata de una limitación objetiva, que tal vez pudiera corregirse, si fuera el caso, mejorando el conocimiento sobre las eventuales ganancias de productividad que podría aportar la persona con discapacidad, absorbiendo los costes de adaptación. Pero también, en general, podría corregirse favoreciendo la introducción de tecnología adaptada en cualquier circunstancia (siempre que no implique un aumento de costes significativo, pues suscitaría un rechazo generalizado), así

como fomentando los programas de I+D+i orientados en tal dirección, mediante las correspondientes políticas de apoyo, a través de las fórmulas ya conocidas.

m) Desigualdades internas en la población potencialmente activa con discapacidad. Tal como se plantea la situación en muchas ocasiones, se alude a la población potencialmente activa con discapacidad de una manera agregada, ignorando las desigualdades internas; cuando es el caso que dichas desigualdades pueden ser especialmente significativas. Cuestiones como la edad o el género (al que dedicaremos un punto específico a continuación) pueden introducir (y de hecho introducen) problemas adicionales que deben ser evaluados y afrontados. Pero, con independencia de ello, cada tipo de discapacidad introduce matices y condicionantes específicos que están distorsionando la información estadística agregada y, por tanto, dificultando la resolución de los problemas que nos ocupan. Del mismo modo que cualquier persona perteneciente al segmento sin discapacidad no es apta para cualquier puesto de trabajo, ni sus posibilidades de rendimiento son las mismas en todos los puestos (en función de su perfil educativo, actitudes y aptitudes personales, posibilidades de comunicación, imagen, capacidad física...), del mismo modo habría que actuar con una precaución similar al analizar la situación de las personas con discapacidad. La situación no es la misma en cada puesto de trabajo; ni tampoco el mercado debe responder del mismo modo ante unos tipos de discapacidad y otros. El tratamiento político amorfo y generalizado podría introducir, en este sentido, elementos de discriminación política (y no precisamente positiva) aun con la mejor de las intenciones. De ahí la necesidad de efectuar análisis desagregados y, al mismo tiempo, diseñar políticas concretas adaptadas a los resultados de tales análisis.

n) Asimetría de Género. Es evidente que la confluencia del género con la discapacidad alimenta una situación diferencial matizada doblemente. Pero no parece que exista una incidencia particular en el ámbito de la discapacidad en perjuicio femenino, más allá de lo que ocurre, en general, en la sociedad. Mas aún; según se ha señalado más arriba, las distancias entre varones y mujeres podrían ser más reducidas en el caso del segmento de las personas con discapacidad. Si ello es así, cabría preguntarse por la alternativa más conveniente: una política diferenciada con relación a varones y mujeres, el tratamiento transversal del tema del género en toda las actuaciones políticas que venimos planteando, o no introducir ninguna atención específica, por entender que es suficiente con las políticas generales de género establecidas en nuestra sociedad para atender este problema. Esta última cuestión podría incluso defenderse alegando que podría difuminar los contenidos propios relativos a la discapacidad, restando fuerza y capacidad de impacto al diversificar los objetivos políticos en dos frentes, perjudicando la atención requerida.

6.- CONCLUSIONES.

De acuerdo con la evidencia empírica y la percepción del funcionamiento social suelen asignarse tres características fundamentales a la situación laboral de las personas con discapacidad, en comparación con la situación similar de las personas sin discapacidad: a) reducida tasa de actividad; b) alta tasa de desempleo; c) peores condiciones de contratación, en términos de retribución, acceso y promoción, etc. El resultado de todo ello es la existencia de una reducida tasa de empleo, con todas las consecuencias que se derivan, tanto en términos personales (integración social y desarrollo humano) como sociales (contribución al progreso general).

El análisis de los posibles factores que influyen en dicha situación no resulta una tarea fácil, toda vez que se acumulan explicaciones diversas y se hacen intervenir múltiples elementos, los cuales no pueden ser excluidos a priori. Dichos factores se encuentran tanto en el lado de la oferta como en el de la demanda específica de mano de obra con discapacidad, dependiendo del juego de ambas la situación final observada. Asimismo, debe valorarse la influencia de la situación general del mercado laboral, además de la coyuntura económica, la productividad general de la economía y otras consideraciones macroeconómicas.

Centrando la atención en los aspectos específicos que inciden en el lado de la oferta de mano de obra con discapacidad, pueden señalarse tanto elementos de carácter objetivo (expectativas laborales poco estimulantes, altos costes para acceder al mercado, menor capital humano, importancia de las rentas no salariales...), como subjetivos (relacionados con la función

de preferencias, fuertemente influida, entre otras razones, por los hábitos sociales, las actitudes de la propia población afectada, la protección familiar...); circunstancias todas ellas que presionan para que el salario de reserva tienda a elevarse, y a desplazar, en consecuencia, la función de oferta, presionando a la baja la tasa de actividad. Al mismo tiempo, las condiciones en el lado de la demanda, como la existencia de prejuicios, las expectativas de los demandantes sobre la productividad de la mano de obra con discapacidad, la aversión al riesgo, el funcionamiento de los mercados internos, etc., presionan para que el salario de referencia al que las empresas están dispuestas a contratar tienda a ser muy bajo, deprimiendo igualmente la demanda de mano de obra con discapacidad. De ahí se sigue un ajuste difícil de la situación con las consecuencias conocidas.

En definitiva, nos encontramos con un mercado laboral segmentado, caracterizado no sólo por las preferencias de la demanda (motivadas por la existencia de prejuicios, por ej.), sino también por una clara asimetría (externa) en la oferta, dadas las desiguales posiciones de partida de los dos segmentos de población, personas con y sin discapacidad. Tal segmentación pone en evidencia la paradoja (al menos aparente) de que coexistan unos salarios de reserva particularmente elevados y unos salarios efectivamente percibidos por personas con discapacidad empleadas más reducidos que los promedios del mercado, lo que sugiere la necesidad de afrontar la existencia, también, de una asimetría interna en el lado de la oferta, en el que podemos encontrar situaciones muy diferenciadas según tipo de discapacidad, género, formación, historia clínica, edad, etc.

En consecuencia, se necesitan respuestas políticas específicamente diseñadas para atender tal situación, realistas y no sujetas ellas mismas a prejuicios. Dichas políticas pueden ser meramente compensadoras (de los costes específicos vinculados con la discapacidad, de las menores expectativas de rendimiento por parte de los demandantes de empleo, de las necesidades insatisfechas de las personas con discapacidad y sus familias motivadas por la no inserción en el mercado laboral, de los menores salarios percibidos, etc.), consolidando, por tanto, la situación de fondo; o, por el contrario, afrontar la generación de un panorama nuevo, atacando las raíces del problema. En concreto, habría que señalar tres frentes: la reducción efectiva de costes anejos a la discapacidad, tanto en el ámbito de la oferta como de la demanda (estimulando la necesaria cualificación de empleados, empleadores y personal auxiliar, innovaciones tecnológicas adaptadas, etc.); mejorar los servicios de información y gestión laboral; modificar las actitudes y hábitos sociales. Todo ello teniendo como referente la unidad familiar, y no sólo individualmente a la persona con discapacidad; así como adoptando una perspectiva más desagregada, que tenga en cuenta la asimetría interna aludida.

Todo ello sugiere la utilidad de incorporar y tener en cuenta en el diseño y ejecución de las políticas las múltiples iniciativas sociales existentes, particularmente apta para captar el variado panorama de la discapacidad, afrontar los problemas de sensibilización social, identificar adecuadamente los costes y limitaciones existentes, etc. No en balde, si algo ha puesto de relieve la denominada crisis del Estado de bienestar es la necesidad de incorporar la iniciativa ciudadana, la sociedad civil, en definitiva, en el ámbito de la acción política, como un modo de completar el binomio Estado-mercado, ofreciendo una fórmula posiblemente más precisa y con menores costes. La existencia de una pluralidad tal de iniciativas, vinculadas de manera inmediata con las familias, sugiere la posibilidad de apostar por un modelo de planificación estratégica en red, que aúne las iniciativas y estimule la coordinación intersectorial en beneficio del aumento de la tasa de empleo de las personas con discapacidad.

7. BIBLIOGRAFÍA.

BARTON, L. [Comp.] (1998): *Discapacidad y Sociedad*. Madrid: Morata.

BECKER, G.S. (1957): *The economic of discrimination.* Chicago University Press, Chicago.

BECKER, G.S. (1965): "A Theory of Allocation of Time", *Economic Journal*, September, pp. 493-517.

BECKER, G.S. (1975): *Human Capital*. National Bureau of Economic Research, Nueva York.

CABRA DE LUNA, M.Á. (2004): "Discapacidad y aspectos sociales: la igualdad de oportunidades, la no discriminación y la accesibilidad universal como eje de una nueva política a favor de las personas con discapacidad y sus familias", *Revista del Ministerio de Trabajo y Asuntos Sociales*, nº 50, pp. 21-46.

DOERINGER, P. y PIORE, M. (1985): *Mercados internos de trabajo y análisis laboral*. Ministerio de Trabajo y Seguridad Social, Madrid.

FERNÁNDEZ GUTIÉRREZ, C.R. (2003): "La igualdad de oportunidades, la no discriminación y la accesibilidad universal como eje de una nueva política a favor de las personas con discapacidad y sus familias", *Documentación Social*, nº 130, pp. 25-40.

FUNDACIÓN ENCUENTRO (2001): "Integración Social de las personas con discapacidad", en *Informe España 2001: Una interpretación de su realidad social*. Fundación Encuentro, Madrid, pp.247-307.

GARCÍA LIZANA, A. y FAJARDO CALDERA, M. A. (coord.) (2008): *Estudio de investigación sobre economía, discapacidad y empleo desde las perspectivas de la Economía Aplicada*. Investigación para la Fundación ONCE. Policopiado.

INE (2002): *Las personas con discapacidad y su relación con el empleo: Encuesta de Población Activa del segundo trimestre de 2002*. Instituto Nacional de Estadística, Madrid.

LINDER, S. (1970): *The hurried leisure class*. Columbia University Press. New York.

McCONNELL, C.; Brue, S. y MacPherson, D. (2003): *Economía Laboral*. McGraw-Hill, Madrid.

PHELPS, E.S. (1972): "The statiscal theory of racis and sexs", *American Economic Review*, Sept., pp. 659-661.

PIORE, M.J. (1975) "Notes for a theory of labour market stratification", in Edward, R.C., Reich, M. and Gordon, D. [eds.]: *Labour market segmentation*. Lexington Books, Lexington, Massachusetts, pp. 129-149.

SPENCE, M. (1973): "Job Market Signaling", *Quarterly Journal of Economics*, Vol. 87, nº 3, pp. 355-374.

VILLA FERNÁNDEZ, N. (2007): *La inclusión progresiva de las personas con discapacidad intelectual en el mundo laboral (1902-2006)*. Ministerio de Trabajo y Asuntos Sociales, Madrid.

WEISS, A. (1996): "Human capital versus signally explanations of wages", Journal of Economic Perspectives, Vol. 9 (4), pp. 133-154.

EXPLAINING THE ENTREPRENEURIAL EFFORT OF LOW INCOME FAMILIES BASED ON CHOICE THEORY

MARTIN RAMIREZ-URQUIDY
E-mail: martinramirez@uabc.mx
MICHELLE TEXIS-FLORES
E-mail: mtexis@uabc.mx
Universidad Autónoma de Baja California

Abstract

Theories of entrepreneurship explain the choices confronted by individuals in becoming entrepreneur or remain as worker. The possibility that individuals mix both activities is recognized in literature but has not been sufficiently addressed. This paper proposes a model of choice to study the determinants of mixing activities and the supply of entrepreneurial effort. The model is tested by applying selection techniques on micro data from a sample of 603 individuals and households either running microenterprises as unique activity or as complementary one together with paid jobs. It is reported that individuals and household self-select to participate in entrepreneurship either as part or full time entrepreneurs according to a set of socioeconomic characteristics wherein the number of dependent children, the sex of the entrepreneur, and migratory status are found to be very important. In addition, it is reported that the supply of entrepreneurial effort depends negatively on education and positively on enterprise performance

Key-Words: Entrepreneurship, low income families, microenterpsies, choice theory.
Thematic Area: Social Economics.

1. INTRODUCTION

A prevailing feature of economies, particularly those of developing countries, is the high proportion of individuals running microenterprises with few workers or working as self-employees. This trend is more apparent when the economy exhibit weak dynamics and reduces significantly when the economy is vigorous, hence for developed nations these agents exhibit relatively low dynamics and contribute more modestly as compared to developing economies where they constitute a major source of formal and informal jobs and enterprises.

Microenterprises are generally defined as those employing up to 10 workers including the owner, although most of the sector is compounded by those employing up to 5, and averaging 2 workers. According to Lieadholm (2002), this type of enterprises employs between 17 and 27% of labor force in developing countries, figure that increases significantly to about 75% when considering only informal jobs in the economy (Perry et. al., 2007). A figure of the microenterprise overall contribution in jobs is provided in Orlando and Pollack (2000), who suggests 54% on average for Latin American countries, some reaching more than 60%. These authors also argue that the poverty rate of people engaged in microenterprise activity averages 25% of population, much higher than those in non-microenterprise activity which is about 12% in the same set of countries, and that more than one third of households have at least 50% of their income generated in the microenterprise sector, for some countries the rate may reach more than half of families.

These data suggest the link between microenterprise incidence and low income, and also the propensity of families to choose to mix activities between entrepreneurship and paid employment. Entrepreneurship in this paper is confined as firm or enterprise formation as

assumed in Acs and Kallas (2007), without any regard for the innovativeness entail in the concept and the complex relation between firm formation, innovation and economic growth. Given that most enterprises start as very small, this leaves room for considering entrepreneurs those individuals who self-employ or run microenterprises as defined earlier.

Modern economic theories of entrepreneurship exhibit the dominance of choice theory and focus on studying the reasons for enterprise creation or transition from employment to entrepreneurship, and on determining the characteristics of the entrepreneurs (Parker, 2006). This literature highlights the importance of the availability of financial capital, the earning differential of the transition from employ to entrepreneur, the labor market hardships, such as low wages or high and persistent unemployment; and the influence of some personal characteristics of individuals such as human capital, labor market experience, family background in business, the need for schedule flexibility and independence, migratory and marital status, among others (Georgellis, 2005). The possibility that individuals or households mix both activities, being employed in a paid job and entrepreneur, and if so, how much of both efforts to supply, is recognized but has not yet generated a significant body of literature in spite of the incidence of this phenomenon among low income families as suggested by the figures.

This paper addresses this issue by proposing a model of choice to study the determinants of mixing activities and beyond by explaining the decision of how much entrepreneurial effort to provide conditional on the individual having entered entrepreneurship. The model is tested by applying the Heckman two-steps selection model on micro data from a sample of 603 individuals either running microenterprises as unique activity or as complementary one together with paid jobs. It is reported that individuals and household self-select to participate in entrepreneurship either as part or full time entrepreneurs according to a set of socioeconomic characteristics wherein the number of dependent children, the sex of the entrepreneur, and migratory status are found to be important. In addition, it is reported that the supply of entrepreneurial effort depends negatively on education and positively on enterprise performance variables such as sales. The paper provides some policy implications related to entrepreneurship, poverty and gender issues.

2. THE THEORETICAL MODEL

Assume a household who chooses either to establish an enterprise selling a good at price (p), work for another company as paid workers with wage $w(h)$ depending on human capital (h), or mix these activities. Assume in addition a twice differentiable family utility function (1) with u'>0 and u''<0 and a twice differentiable family production function (2) with standard concavity properties f '> and f '' <0. Utility depends on income (I) and leisure (H), and output (Y) is a function of labor dedicated to entrepreneur activities running an enterprise (L_e), physical capital (k), and also human capital (h). Human capital includes work and entrepreneurial experience and school education.

(1) $U = u(I, H)$

(2) $Y = f(L_e, k, h)$

Total family income is $I = I_e + I_j$, where I_e and I_j are income from entrepreneurial activities and income from working as an employed respectively, so total family income is represented as in 3:

(3) $I = pf(L_e, k, h) - w(h)L_e + w(h)L_j$

Families assign some time to work (L) and some time to rest, spend time together and care for themselves (H), such that $L + H = 1$ so they choose a proportion of total available time to

both activities. Work time can be either assigned to work as paid worker (L_j), to work as an entrepreneur (L_e), or mix, such that $L = L_e + L_j$. With this, *H* equation is given by 4.

(4) $H = 1 - L_e - L_j + \phi L_e$

Paid jobs and entrepreneurship are different in that the former is liked to a boss, usually to a strict schedule and goals are exclusively the organization´s, whereas the later is independent and more flexible. Flexibility in this setting is defined as the ability of the household to adjust the enterprise organization as to achieve family goals and may take pursuing family and enterprise goals simultaneously. There is at least a two-folded role of families: caring for their members and income-making for livelihood, both taking time. Given this role, families manage to maximize wellbeing by allocating time to activities promoting their income and care for their members. Enterprises in this setting may be source of income and flexibility to achieve family goals. Flexibility is incorporated in the model by adding the proportion $\phi \in$ [0,1] of L_e ÿÿuch that this share of work time in the enterprise is added to *H* as written in equation 4.

If the value of this parameter is extreme toward 1 means that the enterprise provides total flexibility and work in the enterprise liberate proportionally leisure time to achieve family goals. This is the case of home work shops managed by parents who simultaneously are able to care for children or respond whenever required. The other extreme is the parameter tending to cero which means that the enterprise does not provide flexibility either because it is larger, located away from home or its organization is more complex.

Given these equations and assumptions, the household's optimization problem is choosing L_e and L_j as to maximize $E(U)$ with $L_e, L_j > 0$ for an interior solution or either $L_e = 0$ or $L_j = 0$ for a corner outcome. This problem is expressed on equation 5, where the arguments *I* and *H* of the utility function (1) have been replaced by their respective equations 3 and 4.

(5) $\underset{L_e, L_j}{\text{Max}} \quad E[pf(L_e, k, h) - w(h)L_e + w(h)L_j, 1 - L_e - L_j + \phi L_e] + \mu_1 L_e + \mu_2 L_j$

First order conditions for L_e and L_j can be written as the following equations 6 and 7 respectively:

(6) $$\frac{\partial E(U)}{\partial L_e} = \frac{\partial E(U)}{\partial I}[p\frac{\partial f(L_e, k, h)}{\partial L_e} - w(h)] + \frac{\partial E(U)}{\partial H}[\phi - 1] + \mu_1 = 0$$

(7) $$\frac{\partial E(U)}{\partial L_j} = \frac{\partial E(U)}{\partial I} w(h) + \frac{\partial E(U)}{\partial H}[-1] + \mu_2 = 0$$

Equation 6 and 7 suggests that the level of work as entrepreneur or employee is taken up to the point where the marginal expected utility of those activities exhaust, that is $\frac{\partial E(U)}{\partial L_e}, \frac{\partial E(U)}{\partial L_j} = 0$. The equilibrium level of both activities depends positively on the expected marginal utility of income $\frac{\partial E(U)}{\partial I}$ and negatively on the marginal utility of leisure $\frac{\partial E(U)}{\partial H}$ so the higher value put on leisure the lower overall work of the family, with the

parameter ϕ mitigating this effect; while the higher value put on income the higher working effort.

Equality between 6 and 7 $\frac{\partial E(U)}{\partial L_e} = \frac{\partial E(U)}{\partial L_j}$ determines the optimal choice L_e, L_j of families (equation 8), with a left-hand side of this equation greater than the right-hand encouraging the family to venture in entrepreneurial activities whereas the opposite generate incentives to look for jobs.

(8) $$\frac{\partial E(U)}{\partial I}[p\frac{\partial f(L_e,k,h)}{\partial L_e} - w(h)] + \frac{\partial E(U)}{\partial H}[\phi - 1] + \mu_1 = \frac{\partial E(U)}{\partial I}w(h) + \frac{\partial E(U)}{\partial H}[-1] + \mu_2$$

For corner solution we have two cases. The first where families allocate all the effort to entrepreneurship L_e>0 and L_j=0 [$\mu_1 = 0, \mu_2 > 0$], with the right-hand side greater than the left-hand side; and a second one where they allocate all the effort in paid jobs L_e=0 and L_j>0 [$\mu_1 > 0, \mu_2 = 0$], with the right-hand side less than the left-hand side. The conditions for mixing activities is equivalent to the interior solution L_e, L_j >0 where $\mu_1, \mu_2 = 0$ and the associated condition is given by the equality in 8. These conditions can be restated as follow:

(9) $$\frac{\partial E(U)}{\partial I}[p\frac{\partial f(L_e,k,h)}{\partial L_e} - w(h)] + \phi\frac{\partial E(U)}{\partial H} > \frac{\partial E(U)}{\partial I}w(h)$$

(10) $$\frac{\partial E(U)}{\partial I}[p\frac{\partial f(L_e,k,h)}{\partial L_e} - w(h)] + \phi\frac{\partial E(U)}{\partial H} = \frac{\partial E(U)}{\partial I}w(h)$$

(11) $$\frac{\partial E(U)}{\partial I}[p\frac{\partial f(L_e,k,h)}{\partial L_e} - w(h)] + \phi\frac{\partial E(U)}{\partial H} < \frac{\partial E(U)}{\partial I}w(h)$$

Equations 9-11 suggest that (i) households will be completely entrepreneurial if at the margin, expected utility of entrepreneurial effort is greater than that reported in employing as paid worker, (ii) household will mix activities by equalizing the expected utility reported by effort in each activity and (iii) households will be employees if expected utility derived from this activity exceeds that of entrepreneurship at the margin.

The marginal profits of L_e given by $[p\frac{\partial f(L_e,k,h)}{\partial L_e} - w(h)]$, increases the marginal utility of L_e and thus increases this effort within the family, whereas *w(h)* increases marginal utility of L_j and would tend to increase propensity to work as employees and reduce the entrepreneurial effort. As consequence, all the factors affecting marginal productivity of the household in business (marginal profits) would tend to affect *cetiris paribus* the entrepreneurial effort such as assets (*k*), entrepreneurial related human capital (*h*), and the price of the good entailing profitability of the activity in markets.

Other factors affecting entrepreneurial effort of individuals by pushing them into this activity are less obvious but are implied in the model such as those impacting marginal utility of leisure and the parameter ϕ representing the degree of flexibility of the entrepreneurial activity. The desire for leisure may be affected by the number of children or dependants in the family, but they also affect marginal utility of income which *cetiris paribus* increases propensity to

work in either activity. The extent the number of children affect entrepreneurial effort depends on the extent households, in search for turning caring and income-making compatible, select entrepreneurial activities providing flexibility in terms of the parameter ϕ being greater than zero. In addition, there are other factors making individuals self-select as entrepreneurs such as migratory status, sex, and marital status, which are in part implied by this model.

3. EMPIRICAL APPROACH AND RESULTS

THE DATA

The data base used in this paper is derived from a multipurpose survey applied once to households and individuals running microenterprises during 2001-2008 for two different states of Mexico: Baja California and Nayarit. Most of the enterprises perform traditional activities in commerce, service and manufacturing where no high-tech is involved. The survey was not designed to test the theory proposed in this paper. Yet, the survey includes data on the enterprise characteristics such as sales, capital invested, duration of the microenterprise, working hours in the enterprise; and characteristics of the individuals such as education, experience in business, sex of the owner, marital status, number of children, and percentage of the households total income provided by their entrepreneurial activities.

The sample descriptive statistics are shown on table 1 crossed by schooling level where several basic features of data are reported. First, average amount of sales and capital are positively related to schooling level, although great dispersion in each group exist. In general, the value of the activities in terms of these two indicators fluctuates between 3 and 10 2003 minimum wages. Another feature of the sample is size in terms of employees which averages 2, following an inverted U-shape with respect to schooling level with the minimum size of 1.33 in the non-educated group, which is similar to that of the college group. The employment peak is 2.47 found in the junior high group. One explanation for this comes out of the relation between the number of workers, the number of dependants and age of the owner. It seems that the number of workers behaves similarly to the number of dependants and both may be related to the age of the entrepreneur. For older people, the average number of dependants may be lowest because children are older and may not depend on parents, and for middle ages in the table these averages go up since children may be younger and still depend on parents, and finally go down for younger people who may be in the process of starting families. One explanation for the number of dependants being similar to the number of workers is that entrepreneurial households allocate relatives as workers in the enterprise, which is widely observed in enterprises of this type, but in this work is inferred not observed in this sample.

Finally, the most stable figure in the table is the percentage of household income provided by the entrepreneurial activity which averages 60%, and goes down from 65% for the lowest two schooling levels to 47.81% for the college group, which may suggest that the higher the education level, and thus the expected wage in labor markets, the lower the household dependency on an entrepreneurial activity. In different words, when expected wage is higher the opportunity cost of starting businesses is also higher. For lower educated groups and thus expected wage, the opportunity cost of starting business is lower, and consequently it may be expected these households tend to depend more on entrepreneurship *cetiris paribus* than higher educated households do.

Previous figure is useful to make a classification between fully entrepreneurship or mixing occupation by partly participating in entrepreneurship and in labor markets. Table two exhibit frequencies by schooling level and separate those whose income depends fully on the enterprise (Full entrepreneur) or partly (Mix occupation). According to this table, 184 sample units out of 558 are fully entrepreneurs (33%) while 374 (67%) mix occupation

Table 1: Descriptive statistics of the sample crossed by schooling level

		Sales	Capital	Workers	Duration	Age of owner	Dependants	% Sales/ Income
		Mexican pesos 2003=100						
No schooling	**Mean**	6097.00	2867.00	1.34	4.60	46.73	1.87	65.76
	Std Dev.	5719.00	4040.00	0.54	5.69	13.64	1.18	29.08
	n	30	32	35	33	30	31	33
Elementary	**Mean**	17204.84	14767.25	1.99	5.35	43.65	2.66	65.80
	Std Dev.	40690.93	42413.58	2.02	6.03	12.69	1.83	32.39
	n	126	170	185	179	169	149	169
Junior High	**Mean**	24605.62	25247.15	2.47	4.43	40.02	2.33	64.06
	Std Dev.	60569.84	145380.90	2.28	5.16	11.07	1.86	31.86
	n	138	163	180	176	166	147	165
Techical	**Mean**	19203.99	14687.45	1.64	2.94	39.82	2.26	53.93
	Std Dev.	36701.49	31573.84	1.03	3.92	9.91	1.36	31.59
	n	73	73	78	79	74	65	72
High School	**Mean**	21604.89	23362.74	1.64	2.82	33.79	1.74	49.63
	Std Dev.	52890.35	73788.27	1.00	3.23	10.71	1.43	33.67
	n	96	102	107	105	102	100	103
College	**Mean**	49893.38	26536.70	1.33	2.89	30.69	1.47	47.81
	Std Dev.	111108.80	58160.25	0.91	3.60	9.54	1.42	35.64
	n	16	18	18	18	16	17	16
Total	**Mean**	**20919.78**	**19086.60**	**1.97**	**4.19**	**40.05**	**2.24**	**60.25**
	Std Dev.	**51794.97**	**89178.78**	**1.81**	**5.10**	**12.12**	**1.70**	**32.85**
	n	**479**	**558**	**603**	**590**	**557**	**509**	**558**

Table 2: Degree of entrepreneurship crossed by schooling level

Schooling level	Frequencies	%	Full entrepreneur	%	Mix occupation	%
No schooling	33	100%	12	36%	21	64%
Elementary	169	100%	67	40%	102	60%
Junior High	165	100%	61	37%	104	63%
Technical	72	100%	18	25%	54	75%
High School	103	100%	22	21%	81	79%
College	16	100%	4	25%	12	75%
Total	**558**	**100%**	**184**	**33%**	**374**	**67%**

Reveling information is obtained by crossing by schooling level where is noticed that the higher education the fewer individuals or households being fully entrepreneurs and the more mixing occupation, whereas the lower schooling level the higher the amount of sample units being fully entrepreneurs and the less mixing occupation. As shown on the table, the percentage of those mixing occupation and fully entrepreneurs vary significantly across educational groups. On the one hand, the percentage of individuals mixing occupation is lower than average for the lower educated group (higher than average for the higher educated group); and on the other, the percentage of individuals being fully entrepreneurs is higher than average for the lower educated group (lower for the higher educated group). Hence, in general terms, these data suggests a negative relation between education and the degree at which individuals and households depend on entrepreneurial activities for a living. If schooling level is proxy for wealth, income, or reward in labor markets, then the data also suggests a link between entrepreneurship as defined earlier and low income.

EMPIRICAL MODEL

The theoretical foundations for the supply of entrepreneurial effort were introduced in the previous section. The set of conditions 9-11 suggest that individuals and households choose to participate as entrepreneurs as unique activity (Condition in 9) or as complementary one together with participation in labor markets (condition in 10), and participating only in labor markets as paid workers (Condition 11). The data used in this paper regard individuals and households that are observed only either under 9 or 10, data on individuals under condition 11 is not available. In addition, the sample under study may have not been drawn randomly from population, rather individuals and households self-select to participate as entrepreneurs as unique or complementary activity (occupation mixing) depending on some characteristics that make them more disposed to do so. This may create a bias, which requires special econometric treatment for correction by a selection method. The typical approach to correct selection is proposed in Heckman (1979) as the two-steps model. Thus, next we propose the selection equation and then the equation for the supply of entrepreneurial effort.

The selection equation is derived from conditions 9-10 in terms of a latent variable z^* such that if $U(L_e > 0, L_j > 0) = z^*$ then $L_e, L_j > 0$ for an interior solution; and $U(L_e > 0, L_j = 0) > z^*$ for a corner outcome $L_e > 0, L_j = 0$. Interior solution in this context means that individuals and families choose to have at least a minimum or more of supply of entrepreneurial effort (Participate in entrepreneurship) whereas the corner solution indicates that all the effort is focused on entrepreneurship (Full entrepreneurship). Denote the former as $D_e = 1$ and the later as $D_e = 0$, this binary variable is the selection dependent variable for the selection equation (Equation 12) whose argument is the vector *W* of independent variables determining individuals and households choices between mix activities (*De=1*) and fully venture in entrepreneur activities (*De=0*), γ is the vector of associated maximum likelihood coefficients and u_i is the error term of the Probit regression. Equation 13 is the supply of entrepreneurial effort conditional on individuals and families having decided to participate in entrepreneurship $U(L_e, L_j) \geq z^*$, whose argument is the vector *X* of independent variables, β the associated parameters and ε_i the error term of the OLS regression.

(12) $D_e = \gamma' W + u_i$

(13) $L_e = \beta' X + \varepsilon_i$

The two steps method assumes that u_i and ε_i are bivariate normal distributed with mean 0, u_i variance 1, ε_i variance σ_ε and correlation coefficient ρ between u_i and ε_i. If errors are uncorrelated so ρ is close to 0, consistent estimation of the parameters in 13 is possible by OLS with no need for selection bias correction. If errors are correlated so ρ is high suggesting dependency between the two variables, then correction is required and Heckit estimators needed to achieve consistency. The Heckit estimator defined by equation 14 augments the OLS regression by including an estimate of the omitted regressor for each observation from the Probit model 12 to take in account a possible selection bias. This estimate is the inverse Mills ratio $\lambda_i = \frac{\phi(\hat{\gamma} W_i)}{\Phi(\hat{\gamma} W_i)}$, where ϕ and Φ are the standard normal density function and the standard normal cumulative distribution function respectively.

$$(14)\ L_e = \beta' X + \beta_\lambda \lambda_i + v_i$$

Equation 12 is estimated in Heckman first step and 14 in the second. For the first step, the vector independent variables W is compounded by binary variables Dummy sex=1 if male (*w1*), Dummy marital status=1 if married (*w2*), Dummy migrant= 1 if migrant (*w3*), and the number of family dependants (*w4*). These variables are hypothesized to influence individuals to self-select in entrepreneurship. The second step regards the estimation of equation 14, the supply of entrepreneurial effort (L_e) measured in working hours per week and obtained from survey by multiplying hours of work per day by working days per week. The vector *X* of independent variables is compounded by Experience in business (*x1*) whose measure is given by the ratio years of operation of the enterprise and the age of the entrepreneur, the schooling level (*x2*) which was specified as an ordered variable such that no education was given 1 and college 6, Monthly Sales of the enterprise in monetary terms (*x3*) and a monetary measure of physical capital (*x4*).

ECONOMETRIC RESULTS

The results of the application of Heckman Two-Steps to equations 12 and 14 are exhibited on table 3. The expectation regarding a possible selection bias was materialized given high correlation of errors u_i and ε_i as measured by ρ suggesting dependency between errors, and the inverse mills ratio coefficient β_λ which turned out to be highly significant. This suggests that correction was made by the inclusion of the estimated omitted variable λ_i in equation 14.

Table 3: Econometric Results

Heckman selection model (two-steps estimates)

Dependent variable: ***Supply of entrepreneurial effort (Le)***

Variables		**Coefficients**		**Std. Error**	**Z**	**Prob**
Constant	*β0*	78.35	**	45.92	1.71	0.088
Experience in business (*x1*)	*β1*	84.82		74.64	1.00	0.250
Schooling level (*x2*)	*β2*	-23.96	*	5.77	-4.15	0.000
Sales (Monthly) (*x3*)	*β3*	0.0003378	*	0.0001582	2.14	0.033
Physical capital (*x4*)	*β4*	-0.0000162		0.0001184	-0.14	0.891
Selection variable	*De*					
De=1 if mix, *De*=0 not						
Constant		0.4070	*	0.1734	2.35	0.019
Dummy sex=1 if males (*w1*)	*γ1*	-0.2273	**	0.1405	-1.62	0.106
Dummy marital status=1 if married (*w2*)	*γ2*	0.0395		0.1742	0.23	0.82
Dummy migrant= 1 if migrant (*w3*)	*γ3*	-0.4133	*	0.1446	-2.86	0.004
Number of family dependants (*w4*)	*γ4*	-0.0693	**	0.044	1.61	0.105
	β_λ	106.02		47.58	2.23	0.26
	ρ	0.88				

Observations= 331, Censured 167, Uncensored 164

* 5% Significance, ** 10% Significance

The theoretical model predicted these four variables involved in 14 to explain the supply of entrepreneurial effort, although only two of them were significant. The supply of entrepreneur effort is explained positively by Sales (*x3*) and negatively by Schooling level (*x4*), whose coefficients were statistically significant at 5%. Experience in business and physical capital resulted not significant.

According to the theoretical model, sales as *proxy* of enterprise performance, increases the utility that is the reward of entrepreneurial effort thus deepening this effort within the

household. Contrarily to sales, education which proxies expected wage or the reward for participation in labor markets reduces the entrepreneurial effort of households by increasing the opportunity cost of this activity. In different words, low educated individuals who expect a low wage are more disposed to supply entrepreneurial effort than higher educated people. This finding is consistent with observations regarding the incidence of self-employment and very small microenterprises in developing countries where labor markets exhibit high and persistent unemployment and low wages.

As mentioned before, physical capital and experience in business were not significant in explaining the supply of entrepreneurial effort. In the first case, dispersion in terms of amount of assets across enterprises and measurement problems either under or over estimation may be involved. Problems in measurement may have been caused by errors in accounting for the amount of households assets associated to the enterprise and the lack of accuracy in determining what the business´s and what the household´s assets are. This is particular important for home shops or some streets vendors who tend to mix assets for home and business use. The survey has limits in digging for assets information since depends on entrepreneurs response. In addition, it is more difficult to remember and estimate this variable than are sales or other data. In the second case, although entrepreneurial related human capital has been found significant in determining enterprise performance, transition to entrepreneurship and longevity of businesses in most studies, here was not relevant for the supply of entrepreneurial effort. One explanation is the proxy used in this paper, the ratio business years of operation-entrepreneur age rather than using something else. The variable years of operation alone was correlated with sales in the supply equation so using it would have caused trouble and other possible entrepreneur related human capital variable were not available.

The other aim of the paper relates selection in the Heckit first step, which given the statistics in table 3, turned out to be present in determining participation in the entrepreneurial activity and effort. Individuals and households may have some characteristics that make them more disposed to entrepreneurship. The regression included sex, marital and migratory status, and the number of dependent children, all of them assumed to make people self-select in entrepreneurship according to the theory proposed earlier. According to the econometric report for equation 12, the dummy sex =1 if the entrepreneur is male (*w1*) and dummy migrant=1 for migrants (*w3*) turned out to be negative and statistically significant at 10% and 5% respectively, whereas being married (*w2*) resulted not significant. In addition, the number of dependent children (*w4*) exhibits statistical significance at 10% and negative sign.

The result for sex suggests that given a male entrepreneur, individuals and households are less likely to decide to mix activities and more likely to be full time entrepreneurs. In different words, being female entrepreneur increases the probability that the individual or household be part time entrepreneur or occupation mixer. The explanation for this relates the separation of roles in families, particularly parents, so when males are the entrepreneurs, they tend to perform entrepreneurial activities full time as family heads whereas females play the role of home wives. On the other hand, when females are the entrepreneurs, they may combine home wife's activities with entrepreneurial ones and males participate in labor markets.

A related results regard de number of dependent children (*w4*) which contributes negatively to the likelihood of people deciding to mix activities. Put in different words, this result suggests that *w4* increases the likelihood of individuals and households to decide to enter entrepreneurship as unique activity. Although the theory in this paper predicts that the number of children increases the household desire for leisure and at the same time increases the marginal utility of income with the possibility of leaving its impact on entrepreneurial effort ambiguous, it does provide an explanation by contending that children may impact decision to self-select entrepreneurial activities providing flexibility in terms of the parameter ϕ being greater than zero in search for turning caring and income-making compatible.

Finally, the explanation for the migratory status being significant in affecting positively the likelihood of selection of entrepreneurship as unique activity (or negatively the selection of mixing activities) may not be provided in the theory presented in this paper but by other sources in literature. It has been argued the propensity of migrants to select start businesses for a living once they move to another country or region on the grounds that they may face discrimination in

labor markets and lack of social capital to look for jobs. The inclusion of this variable in the empirical model was made for specification reasons, and no explanation is provided in the theoretical model.

4. CONCLUSIONS

This paper aimed to propose a model of choice to study individuals and family choices to enter entrepreneurship as complementary (mix occupation together with participation in labor markets) or unique activity (full entrepreneurship) and the determinants of the supply of entrepreneurial effort. It is reported that individuals and household self-select to participate in entrepreneurship either as part or full time entrepreneurs according to a set of socioeconomic characteristics wherein the number of dependent children, the sex of the entrepreneur, and the migratory status are found to be important. In addition, it is reported that the supply of entrepreneurial effort depends negatively on education and positively on enterprise performance variables.

The findings of the paper are useful to document the proliferation of self-employs and very small microenterprises in developing countries characterized by high unemployment and low wages, and emphasize the linkage between low income and microenterprise activity. The paper also provides some insights on entrepreneurship beyond of what is regularly considered by introducing the role played by dependent children in entrepreneurial household choices, an explicit concept of flexibility required in making compatible caring and income-making which are two important roles of families; and the role played by microenterprises. The importance of children incorporates the issue of gender, another important element in entrepreneurship particularly in low income environments, where families with female entrepreneurs where found more likely to mix occupation supporting the linkage between children-flexibility-microenterprise

Finally, the paper has implications for a social entrepreneurial policy involving microenterprise development, poverty alleviation, gender and households issues; and a set of innovative incentives to encourage this type of entrepreneurs targeting an appropriate balance between social pecuniary and no pecuniary returns from policy.

REFERENCES

ACS, ZOLTAND & KADRI KALLAS (2007) "State of Literature on Small to Medium-Size Enterprises and Entrepreneurship in Low Income-Communities" Discussion Papers on Entrepreneurship, Growth, and Public Policy, Paper # 0307. Germany: Max Planck Institute of Economics, pp 45.

GEORGELLIS, Y, J. G. SESSIONS, Y N. TSITSIANIS (2005) Self-Employment Longitudinal Dynamics: A Review of the Literature, Economic Issues, Vol. 10, Part 2, 2005.

HECKMAN J. (1979) Sample Selection Bias as a Specification Error, Econometrica, Vol. 47, (1), pp. 153-161

LIEDHOLM, CARL. (2002). Small Firm Dynamics: Evidence from Africa and Latin America. Small Business Economics (18): pp 227–242.

ORLANDO MARIA BEATRIZ & MOLLY POLLACK (2000) Microenterprise and Poverty: Evidence from Latin America, Washington DC: Inter-American Development Bank, pp 45.

PARKER SIMON (2006) The Economics of Self-Employment and Entrepreneurship Cambridge: Cambridge University Press, pp 323.

PERRY GUILLERMO E, WILLIAM F. MALONEY, OMAR S. ARIAS, PABLO FAJNZYLBER, ANDREW D. MASON, JAIME SAAVEDRA-CHANDUVI. (2007). Informality: exit and exclusion, Washington D.C: World Bank, pp 22.

LA LEY DE AUTONOMÍA PERSONAL Y ATENCIÓN A LAS PERSONAS EN SITUACIÓN DE DEPENDENCIA EN 2008. ¿ALGO MÁS QUE UNA EXPECTATIVA?

MARÍA CRESPO GARRIDO
E-mail: maria.crespo@uah.es
Universidad de Alcalá

Resumen

La Ley de Dependencia se configuró como el cuarto pilar del Estado del Bienestar, pues reconocía como derecho subjetivo las prestaciones que hasta el momento se percibían de manera netamente asistencial. No obstante, la aplicación de la misma, habiendo suscitado enormes expectativas, simultáneamente ha frustrado muchas esperanzas, por no llegar a más del 10% de los solicitantes tras dos años de aplicación. Y es que; los errores en la determinación de los dependientes existentes en España, la falta de acuerdo entre las Comunidades Autónomas, en lo que a financiación se refiere, la disparidad de los niveles de prestaciones entre territorios y en definitiva, la precipitación, están haciendo que lo que se configuró como un derecho subjetivo se esté plasmando como una gran frustración.

Palabras Clave: Dependencia, comunidades autónomas, prestaciones públicas.
Área Temática: Economía Social y Discapacidad.

Abstract

The Law of Dependence was formed as the fourth pillar of the State of the Well-being, since it was recognizing as a civil right the services that up to the moment were perceived in a net welfare way. Nevertheless, its application had created enormous expectations and simultaneously it has frustrated many hopes, for not coming to any more than 10 % of the solicitors after two years of application. And is that; the mistakes in the determination of the existing dependants in Spain, the lack of agreement among the Autonomous Communities, in what to funding the rainfall refers, the disparity of the levels of services among territories and definitively, the harry up are doing that what was formed as a civil right is taking form as a great frustration.

Key Words: Dependence, autonomous communities, public services.
Thematic Area: Social Economics.

1. INTRODUCCIÓN

Cumplidos dos años de la Ley 39/2006 de Promoción de la Autonomía Personal y Atención a las Personas en Situación de Dependencia es el momento adecuado para poder hacer una valoración de la protección de los derechos sociales que esta norma ha puesto en marcha.

El anuncio de la aprobación de la, vulgarmente denominada Ley de Dependencia ha creado enormes expectativas entre los potenciales dependientes, ya que suponía el reconocimiento de un derecho subjetivo para todos aquellos que hasta el momento estaban recibiendo alguna prestación meramente asistencial, con el consiguiente efecto retroactivo. Y además, junto con las prestaciones ya existentes; residenciales, ayuda a domicilio, teleasistencia,… se configuraba un nuevo panorama constituido por las prestaciones económicas; prestación económica vinculada al servicio, prestación económica para asistencia personal y prestación económica para cuidados familiares. Especialmente ésta última mejoró los ánimos de "las", más que *los*

cuidadores informales, pues son muchas las mujeres que dedican toda su vida al cuidado y atención de un familiar dependiente, sin derecho a haber percibido prestación alguna.

El interés por el correcto desarrollo de la Ley no sólo nace de los potenciales beneficiarios y de sus familias, así como de los sectores profesionales afectados o de los políticos y legisladores. Muchos más ciudadanos están preocupados porque se satisfagan las prestaciones económicas en ella previstas, entre otras cosas, porque todos somos posibles beneficiarios de la misma. Cualquier persona joven y con salud, puede convertirse en dependiente, de manera inmediata, ya sea por una enfermedad sobrevenida o por un accidente, por lo que esta ley importa a todos, tanto a los afectados por la dependencia como a los que pueden llevar a cabo una vida independiente.

El correcto desarrollo de esta Ley, por tanto, con un aporte financiero suficiente y estable en el tiempo, habría supuesto que se hubiera consolidado el llamado "cuarto pilar del Estado de Bienestar", junto con sanidad, educación y pensiones. Sin embargo, la carencia de recursos financieros, una dotación presupuestaria insuficiente y la inexistencia de acuerdo entre las Comunidades Autónomas ha hecho que se frustraran las múltiples expectativas generadas, en el anuncio de su puesta en marcha.

Al comienzo de la implantación de la ley uno de los principales obstáculos que aparecieron fueron las reticencias y la falta de acuerdo entre Comunidades Autónomas, pues la gestión de la dependencia se hacía *ex novo* por parte de las autonomías, sin una experiencia previa estatal, como se hizo con la sanidad o con la educación, por lo que las diferencias territoriales, en lo que a gestión y sobre todo, aportación financiera, se hicieron patentes.

Esto provocó que, junto con la falta de suficiencia financiera una de las principales críticas que desde la óptica económica y sociológica se hacen a la implantación de la ley sea por el incumplimiento del más básico principio de solidaridad interregional que, si bien, no sólo se produce en el ámbito de la dependencia, se debería haber tomado en consideración antes de su entrada en vigor.

Precisamente la falta de acuerdo entre las Comunidades Autónomas para cofinanciar las prestaciones, los recursos administrativos planteados y retirados por algunas de ellas, las diferencias en la dotación de prestaciones previas a la aplicación de la Ley han hecho que se acentúen enormemente las diferencias territoriales, incumpliendo la tan manida solidaridad interregional.

2. SITUACIÓN DE LA DEPENDENCIA EN ESPAÑA ANTES DE LA APROBACIÓN DE LA LEY

2.1. Los modelos de protección en Europa

Las novedades introducidas por la Ley de Autonomía Personal y Atención a las Personas en Situación de Dependencia sólo se entienden con un conocimiento previo de la realidad asistencial previa a 2007, y es precisamente ésta la que está recibiendo gran parte de las críticas, pues no son pocos los que antes eran atendidos por los servicios sociales y ahora quedan fuera del radio de acción de la dependencia.

Y es que la situación española ha sido y es radicalmente distinta a la que presentan otros países de nuestro entorno, por lo que es imprescindible conocer otros modelos de protección social existentes y cómo se desarrollan, ya que los diversos modelos están íntimamente relacionados con los niveles de empleo existentes en el país.

Algunos autores europeos clasifican las formas de protección en el entorno europeo de la siguiente manera:

- Modelo mediterráneo: En el que existe un pronunciado riesgo de pobreza y un nivel de empleo más bien bajo.
- Modelo anglosajón: Donde la protección social es reducida y los niveles de empleo altos.

- Modelo continental: Que protege claramente contra la pobreza y donde el nivel de empleo alcanza cotas bajas.
- Modelo nórdico: Conjuga un alto nivel de empleo y un disminuido riesgo de pobreza. Las prestaciones sociales son altas para el desempleado y todo tipo de servicios sociales para las personas necesitadas y ayudas a la familia. Éste es el modelo que mejor conjuga eficacia y equidad.

España ha encajado siempre perfectamente en un modelo continental de prestaciones y con la aprobación de la Ley de Autonomía Personal, se pretendió que la protección contra la pobreza pasara de ser un reconocimiento asistencial a un derecho subjetivo, lo que supuso un avance muy significativo en lo que a reconocimiento de derechos se refiere.

No obstante, la actual situación de crisis económica puede acentuar la realidad del desempleo y aunque no existe aún un horizonte temporal amplio, si los recursos públicos siguen descendiendo al caer las cotizaciones a la Seguridad Social por el aumento del desempleo y, precisamente por la necesidad de aumentar las prestaciones por desempleo puede que se reduzcan los recursos destinados a las prestaciones de dependencia. Sin embargo, es necesario analizar el estado de la cuestión con una perspectiva temporal más amplia para poder hacer alguna afirmación solvente.

2.2. La discapacidad para la Organización Mundial de la Salud

La dispersión de normas, estadísticas y datos, así como la heterogeneidad en su cálculo ha dificultado enormemente que antes de la puesta en marcha y de la aprobación de la Ley que nos ocupa, se conociera con relativa certeza un número más o menos exacto de los potenciales beneficiarios de la misma.

Siendo conceptos claramente diferenciados dependencia y discapacidad, lo cierto es que las series estadísticas existentes hasta 2007 se referían a cifras y encuestas sobre el número de discapacitados existentes, y eran tan extraordinariamente heterogéneas que dificultaban enormemente el conocimiento exacto del número de discapacitados en el momento. A continuación se exponen las principales fuentes estadísticas empleadas con anterioridad a la aprobación de la Ley 39/2006.

Una de los puntos fuertes de la Ley de Dependencia ha sido la simplificación y agrupación de las categorías de dependencia, como más adelante se verá. Pues, las tipologías existentes hasta el momento de la aprobación de esta ley, eran mucho más dispersas.

- DISCAPACIDAD: Toda limitación grave que afecte o se espere que vaya a afectar durante más de un año a la actividad del que la padece y tenga su origen en una deficiencia.
- DEPENDENCIA FÍSICA LEVE: es aquella en la que para obtener autonomía se necesita una ayuda puntual para algunas actividades de la vida diaria.
- DEPENDENCIA FÍSICA GRAVE: si se necesita ayuda para que una persona supla en algunas actividades de la vida diaria.
- DEPENDENCIA FÍSICA SEVERA: si se necesita ayuda permanente para las actividades de la vida diaria.
- DEPENDENCIA PSÍQUICA LEVE: se presenta en aquellas personas en las que se empieza a tener algunos signos de demencia senil.
- DEPENDENCIA PSÍQUICA GRAVE: se da cuando los signos de demencia son más habituales.
- DEMENCIA PSÍQUICA SEVERA: se presenta cuando existe pérdida de autonomía psíquica.

Como se ve, la tipología de dependencia y discapacidad no facilitaba en absoluto el conocimiento de la realidad sociológica de la dependencia y la discapacidad en España.

Ya la Organización Mundial de la Salud en 1980 establecía tres niveles susceptibles de ayuda, que eran:

- Deficiencias: entendidas como las consecuencias permanentes de las enfermedades y accidentes en el nivel corporal, fisiológico u orgánico.
- Discapacidades: son las restricciones en la actividad del individuo por cualquier deficiencia.
- Minusvalías: Son aquellas situaciones de desventaja derivadas de deficiencias o discapacidades que limitan o impiden participar y desempeñar roles sociales en niveles normales

Uno de los principales problemas con los que la aplicación de la Ley de Dependencia se ha encontrado es el desconocimiento del número exacto de dependientes existentes en España en el momento de su entrada en vigor. No existen cifras certeras, más allá de meras estimaciones, más o menos veraces. A continuación se presentan los principales indicadores sociológicos empleados en España para la cuantificación de las necesidades de los dependientes:

a. ENCUESTA SOBRE DISCAPACIDADES, DEFICIENCIAS Y MINUSVALÍAS; En 1986 se elaboró esta encuesta a 256.000 personas integradas en 74.000 hogares, y mostraba que del total de la población el 15% eran discapacitados, tenían alguna deficiencia o/y minusvalía. Y teniendo en consideración sólo la población entre 16 y 64 años, el 10,5% de este colectivo de la población es dependiente.
b. PANEL DE HOGARES DE LA UNIÓN EUROPEA DE 1996. Con una muestra de 130.000 personas adultas entre 16 y 64 años, integradas en 60.000 hogares, el 9,9% de la población es dependiente.
c. ENCUESTA SOBRE DISCAPACIDADES, DEFICIENCIAS Y ESTADO DE SALUD. Realizada en 1999 sobre la muestra de 70.500 hogares compuestos por 218.195 personas. Según esta encuesta el 9% de la población es dependiente y de esta, si se tiene en consideración la población entre 16 y 64 años, el 5,1% de este colectivo de población es dependiente.
d. MÓDULO SOBRE PERSONAS CON DISCAPACIDAD DE LA EPA. Esta encuesta se realizó con una muestra de 59.000 hogares compuestos por 112.000 personas y según ella el 8,7% de la población entre 16 y 64 años es dependiente, según la terminología actual.
e. ENCUESTA NACIONAL DE SALUD DE 2003. Para la cuantificación de la población dependiente se empleo una muestra de 22.000 viviendas y estima la población dependiente en 3,3% de la población entre 16 y 64 años.

La existencia de esta multiplicidad de encuestas que tratan de conocer la población dependiente en cada una de las fechas referidas no sólo aportan datos diferentes por la divergencia de las magnitudes evaluadas sino que, además, cada una de ellas clasifican las categorías de dependencia de manera distinta[1].

[1] La Encuesta sobre Discapacidades, Deficiencias y Minusvalías de 1986 clasificaba las discapacidades en las catorce categorías siguientes:

1. Ver
2. Oír
3. Hablar
4. Discapacidades de comunicación
5. Cuidado personal
6. Andar
7. Subir escaleras
8. Correr
9. Salir de casa
10. Actividades de la vida cotidiana
11. Dependencia y resistencia
12. Ambiental
13. Conducta de uno mismo
14. Conducta con los demás

Estas categorías fueron clasificadas en las diez categorías siguientes:

1. Ver
2. Oír
3. Comunicarse
4. Aprender, aplicar conocimientos y desarrollar tareas
5. Desplazarse

3. LA ESPAÑA DE LA DEPENDENCIA

El desarrollo y la implantación de la Ley de Promoción de la Autonomía Personal y Atención a las Personas en situación de Dependencia ha sufrido a lo largo de sus dos años de aplicación tantos avatares como críticas, principalmente por las divergencias territoriales surgidas entre las diferentes regiones, con diferentes modelos organizativos, tanto en lo territorial como en la ordenación de los aspectos procedimentales y de gestión, dificultan y han entorpecido enormemente el desarrollo de la ley.

El anclaje territorial de la protección a la dependencia en lo autonómico se observan cuatro realidades[2]:

- MODELOS DESCENTRALIZADOS: Comunidades Autónomas y Diputaciones Forales que incardinan más o menos plenamente la protección a la dependencia en el sistema de Servicios Sociales. Se apoyan en la red de servicios sociales básicos –cuya gestión directa se desarrolla en el ámbito local municipal- como la red que soporta la información, la valoración (conjunta con la Comunidad Autónoma), la elaboración del PIA y el seguimiento. (Castilla y León y Diputaciones forales vascas). Evidentemente, este proceder lleva aparejada la coordinación con las corporaciones locales y un importante esfuerzo por dotarlas de más personal y medios.
- MODELOS DESCONCENTRADOS: Comunidades Autónomas que centralizan la responsabilidad de gran parte de los procesos a través de la creación de unidades y servicios territorializados (provincializados) dependientes del nivel autonómico, especialmente para la valoración y la provisión. Ceden a las corporaciones locales –a los Servicios Sociales Básicos- la gestión/negociación del PIA pero mantienen la responsabilidad total sobre el sistema (Andalucía, Aragón, la Rioja, …)
- MODELOS CENTRALIZADOS: Comunidades Autónomas que asumen de forma directa la gestión de la protección a la dependencia y que no han contado, al menos inicialmente con los servicios sociales básicos, salvo para algunas tareas subordinadas de índole puramente administrativa. (Castilla la Mancha, Asturias, Madrid, …). Además suele coincidir esta opción con una falta de información importante al nivel local municipal. Esto está provocando gran malestar en la "primera línea" de atención a los ciudadanos.
- MISCELÁNA CAÓTICA: Aquí se agrupan Comunidades Autónomas como Extremadura cuya opción ha sido confiar la gestión de la atención a la dependencia a la red del sistema sanitario incorporando una originalísima visión que no estaba contemplada en la Ley de Dependencia y que, en la práctica, está siendo el camino adoptado por Cantabria.

3.1. Valoración de la dependencia en la ley de promoción de la autonomía personal y atención a las personas en situación de dependencia

La Ley objeto de este artículo ha venido a simplificar enormemente las tipologías de dependencia y discapacidad antes referidas. En concreto, sólo se prevén tres grados de dependencias con dos niveles cada una, y, como posteriormente se verá, según el grado y nivel de dependencia está previsto un calendario de reconocimiento de prestaciones. El grado y niveles de dependencia se determinará por el Consejo Territorial del Sistema Nacional de Dependencia.

- GRADO I: DEPENDENCIA MODERADA: La persona necesita ayuda para realizar varias actividades básicas de la vida diaria[3], al menos una vez al día.

6. Utilizar brazos y manos
7. Desplazarse fuera del hogar
8. Cuidar de sí mismo
9. Realizar tareas del hogar
10. Relacionarse con otras personas

[2] Barriga (2009)

- GRADO II: DEPENDENCIA SEVERA: La persona necesita ayuda para realizar varias actividades básicas de la vida diaria dos o tres veces al día, pero no requiere la presencia permanente de un cuidador.
- GRADO III: GRAN DEPENDENCIA: La persona necesita ayuda para realizar varias actividades básicas de la vida diaria varias veces al día y necesita la presencia indispensable y continua de otra persona.[4]

La Ley de Autonomía Personal preveía un calendario de implantación de la misma que no se ha cumplido, fundamentalmente por no conocerse el número exacto de personas en situación de dependencia en el momento de su entrada en vigor.

Particularmente, el calendario de aplicación que la Ley preveía respondía a que se atendiera a las personas en situación de dependencia, de acuerdo a las fechas que se exponen a continuación:

Calendario de implantación de la Ley

Año	Grado y nivel de dependencia determinante para su atención
2007	Todos los grandes dependientes, grado III, nivel 1 y 2
2008	Dependientes grado II, nivel 2 (dependientes severos)
2009[5]	Dependientes grado II, nivel 1
2011	Dependientes grado I, nivel 2 (moderados)
2013	Dependientes grado I, nivel 1
2015	Finaliza el período de aplicación

Fuente: Elaboración propia a partir de Ley 39/2006

En paralelo a este calendario de aplicación, la Ley de Promoción de la Autonomía Personal prevé una serie de prestaciones novedosas y otras que ya existían en los servicios sociales, que se exponen a continuación:

Estos servicios tienen por objeto la prevención del alargamiento de su grado y nivel de dependencia, incluyéndose la teleasistencia, la ayuda a domicilio, centros de día y atención residencial, y se explican a continuación.

1. SERVICIO DE TELEASISTENCIA

Atiende a los beneficiarios mediante el uso de tecnologías de comunicación y apoyo de medios personales, con objeto de dar respuesta inmediata en situaciones de emergencia, inseguridad, soledad y aislamiento. Se prestará de acuerdo al Plan Individual de Asistencia.

Según los datos facilitados por el IMSERSO a 1 de abril de 2009, existen en el conjunto del Estado español 17.929 servicios de teleasistencia, lo que supone un 2,85% del total de prestaciones de dependencia, siendo Andalucía la Comunidad Autónoma que encabeza, tanto en términos absolutos como en valores porcentuales la prestación de este servicio con un total de

[3] ACTIVIDADES BÁSICAS (ABVD)
Ducha, baño, aseo
Vestirse, desvestirse
Peinarse, afeitarse
Andar, levantarse, acostarse
Comer
Quedarse solo de noche

[4] ACTIVIDADES INSTRUMENTALES (AIVD)
Usar el teléfono
Comprar
Preparar comida/cortar una rebanada de pan
Lavar la ropa (a mano/ a máquina)
Coger un bus/taxi
Acordarse de tomar la medicación
Administrar dinero
Subir 10 escalones

[5] Se prevé una revisión en este año para evaluar la puesta en marcha y el desarrollo de la ley.

12.964, seguida muy de lejos por el País Vasco con 3.237 servicios de teleasistencia y mucho más lejos se sitúa Castilla la Mancha con 672 servicios de teleasistencia. Las comunidades que no tienen reconocido ningún servicio de teleasistencia identificados en PIA son; Aragón, Baleares, Canarias, Valencia o Madrid.

2. AYUDA A DOMICILIO

Se prestará según el informe del PIA así como de la intensidad del grado de dependencia, que varía según los siguientes parámetros:

Grado III nivel 2: 70-90 horas mensuales
Grado III nivel 1: 55-70 horas mensuales
Grado II nivel 2: 40-55 horas mensuales
Grado II nivel 1: 30-40 horas mensuales.

En abril de 2009, la Comunidad Autónoma que encabeza la prestación de los servicios de ayuda a domicilio es Andalucía con 36.884, lo que suponen un 21,98% del total de servicios prestados por la región, seguida muy de lejos por el País Vasco (5.422, un 13,42%). A la cola se sitúan; Aragón, Baleares, Canarias, Valencia, Madrid, Murcia y Navarra con cero servicios de ayuda a domicilio identificados en el PIA.

Tanto este servicio como la teleasistencia presentan la dificultad añadido de ser prestados por las Corporaciones locales, por lo que se agravan los problemas para conocer exactamente las cifras y niveles de prestación del servicio, ya que los ayuntamientos no tienen un sistema homogéneo de reporte de datos a su comunidad autónoma respectiva.

3. SERVICIOS RESIDENCIALES

a. Servicio de Atención Residencial

Ofrece atención integral y continuada de carácter personal, social y sanitario, que se prestará en los centros residenciales, con carácter permanente o temporal, con motivo de vacaciones, fines de semana, convalecencia, enfermedades o descansos de los cuidadores no profesionales.

Todos estos servicios son prestados por las Comunidades autónomas, que, a través de su capacidad legislativa tienen obligación de articular un sistema coherente en el que se unifiquen los sistemas de provisión de estos recursos, los criterios de calidad y acreditación, así como la aportación económica del usuario. No obstante, las Entidades locales, están también obligadas a prestar servicios sociales por la Ley Reguladores de Bases del Régimen local, por lo que es imprescindible un correcto sostenimiento de estos servicios mediante conciertos económicos con sus gobiernos autonómicos y el Estado central.

Los datos aportados por el SAAD en lo que a prestaciones residenciales no son nada reveladores del grado de implantación de la ley ya que no distinguen entre plazas residenciales prestadas antes de la entrada en vigor de la ley y plazas que se hayan cubierto al amparo de la ley.

4. PRESTACIONES ECONÓMICAS DE ATENCIÓN A LA DEPENDENCIA

A diferencia de las anteriores, estas prestaciones económicas descritas en este epígrafe son una novedad introducida por la Ley de Dependencia.

Uno de los mayores problemas con los que se encuentra la implantación de la Ley es que algunas Comunidades Autónomas se han limitado a "traspasar" a los beneficiarios de los servicios sociales al SAAD, fundamentalmente en lo que se refiere a plazas residenciales[6]. En algunos territorios se está vaciando de contenido y, por tanto, de presupuesto, al Sistema de Servicios Sociales para atender casi en exclusiva a las situaciones de dependencia. Tal es el caso

[6] El Sistema de Servicios Sociales atiende al 40% de las personas que son beneficiarias de una plaza residencial y que la Ley de Autonomía Personal no los considera por múltiples cuestiones.

del Plan Concertado[7] de Prestaciones Básicas o de otras muchas prestaciones existentes en los Servicios Sociales Básicos para atender a situaciones que no son de dependencia.

No obstante, la ley preveía que se comenzaran a prestar algunas ayudas económicas; en sustitución de otros servicios que por determinadas circunstancias no se pudieran prestar, otras para aliviar la carga económica y la insuficiencia de recursos que el dependiente pudiera tener por sus especiales características y, finalmente, prestaciones económicas que le sirvieran a la persona en situación de dependencia para que le fuera más fácil el acceso al empleo o al estudio.

a. Prestación económica vinculada al servicio.

Es una prestación personal y periódica, sujeta al grado y nivel de dependencia y capacidad económica del beneficiario, destinada a la cobertura de los gastos del servicio previsto en el PIA, cuando no sea posible la atención por un servicio público o concertado.

De acuerdo a los datos facilitados por el IMSERSO, la situación de estas prestaciones a 1 de abril de 2009[8] oscila entre la realidad de Castilla y León donde la prestación vinculada al servicio representa el 13,32% del total de prestaciones reconocidas por la región y la realidad de la Comunidad de Madrid, quienes han reconocido un 0,03% de los servicios de la dependencia a través de esta prestación. Navarra con un 9,92%, Aragón con un 8,37%, Cataluña con un 7,79% o Extremadura con un 6,29% de las prestaciones de la dependencia satisfechas mediante esta prestación se sitúan como las autonomías pioneras en el reconocimiento de la prestación económica vinculada al servicio.

Sería muy revelador saber si el reconocimiento de esta prestación responde a la inexistencia de recursos suficientes con anterioridad a la implantación de la Ley, por lo que esta medida legislativa habría supuesto un avance en la política social o, por el contrario, se reconoce la prestación aun existiendo recursos de otro tipo y son motivos de otra índole las razones por las que se lleva a cabo el pago de esta prestación. No obstante, las dificultades encontradas en el suministro de información tanto por parte de las Comunidades Autónomas como por parte del IMSERSO, impiden dar una respuesta concluyente.

b. Prestación económica para cuidados en el medio familiar y apoyo a cuidadores no profesionales.

Es una prestación de carácter excepcional, siendo su finalidad el mantener al beneficiario en su domicilio atendido por cuidadores no profesionales. Se establecerá a través del PIA y está sujeta al grado y nivel de dependencia.

Esta prestación que se planteó como excepcional está generalizándose como forma de dar cauce a los PIA. Navarra es la región que mayor número de prestaciones económicas para cuidados familiares ha reconocido, de acuerdo a los datos de 1 de abril 2009 es el 55,70% del total de prestaciones, Ceuta y Melilla con un 46,12%, Galicia con un 39,54%, Cantabria con un 37,16% y Andalucía un 32,13% del total de prestaciones reconocidas con prestaciones económicas para cuidados familiares.

Esta prestación da por supuesto que las cantidades percibidas por la persona en situación de dependencia se aplican a tal finalidad, y además, esta prestación supone que el familiar que lo necesite será dado de alta en la Seguridad Social, al ser un cuidador no profesional. Por esto, una asignatura pendiente es la formación que les es exigible a los cuidadores familiares, así como qué apoyos requieren y quien controla los servicios que se prestan. En consonancia con las prestaciones reconocidas, el número de convenios reconocidos, según las cifras aportadas por el IMSERSO a 1 de abril de 2009 son los que se muestran en la siguiente tabla:

[7] El presupuesto estatal para el Plan Concertado en 2009 es de 95.092.360 euros, con un incremento de alrededor de un millón de euros respecto al presupuesto del año anterior.

[8] En el anexo I se recoge la situación de todas las prestaciones, según datos del IMSERSO a 1 de abril de 2009.

Convenios especiales de cuidadores no profesionales de personas en situación de dependencia 1 de abril de 2009

Ámbito territorial	Convenios obligatorios	Convenios voluntarios	P. E. cuidados familiares	%convenios sobre prestaciones
Andalucía	26.137	22	53.910	48,50%
Aragón	1.450	15	7.670	19,10%
Asturias	1.584	30	4.604	35%
Islas Baleares	548	6	3.053	18,14%
Canarias	538	0	1.455	36,97%
Cantabria	1.559	11	4.100	38,29%
Castilla y León	1.693	9	7.503	22,68%
Castilla la Mancha	4.401	13	10.168	43,41%
Cataluña	5.941	51	28.760	20,83%
Valencia	2.564	15	6.120	42,14%
Extremadura	630	0	1.744	36,12%
Galicia	4.106	50	15.356	27,06%
Madrid	338	5	1.227	27,95%
Murcia	2.110	9	5.329	39,76%
Navarra	414	15	6.372	6,73%
País Vasco	4.219	51	10.339	41,30%
La Rioja	508	5	1.492	34,38%
Ceuta	233	0	785 (ambos)	58,34%
Melilla	225	0		58,34%
TOTAL ESTADO	59.198	307	169.987	35%

Fuente: IMSERSO (2009)

Como se ve, son muchos menos los convenios, tanto obligatorios como voluntarios que se han formalizado, respecto al número de prestaciones económicas para cuidados familiares. No existen datos concluyentes que nos permitan afirmar si esta divergencia se debe a que el familiar encargado de prestar atención a la persona en situación de dependencia ya cotiza en algún régimen de la Seguridad Social o, es por algún otro motivo.

Una de las cuestiones que suscitó mayor expectativa la puesta en marcha de la Ley de Autonomía Personal fue la posibilidad de incrementar los puestos de trabajo alrededor de la dependencia, así como se percibió con gran expectativa la posibilidad de percibir una prestación económica y darle de alta en la Seguridad Social a los familiares que se encargaran de algún dependiente..

c. Prestación económica de asistencia personal.

Facilita la promoción de la autonomía de la persona con independencia de su edad. Pretende contribuir a la contratación de una asistencia personal que facilite al beneficiario una vida más autónoma, el acceso a la educación y al trabajo y el ejercicio de las ABVD

Esta prestación representa, hasta el momento, sólo el 0,5% de los servicios del catálogo reconocidos, a pesar de ser una de las prestaciones que más facilitaría la independencia y la autonomía de la persona en situación de dependencia.

Las cuantías se establecerán anualmente por el Gobierno mediante Real Decreto, previo acuerdo del Consejo Territorial, para los grados y niveles con derecho a prestaciones, actualizándose en función del IPC

En esta prestación sólo el País Vasco presenta una situación algo relevante, pues el 0,87% de las prestaciones identificadas en el PIA son prestaciones económicas para asistencia personal. No obstante, no se ha reconocido ni una sola prestación en Aragón, Canarias, Cantabria, Valencia, Extremadura, Madrid, Murcia, La Rioja, Ceuta ni Melilla, según los datos del IMSERSO.

Como se ha visto, del catálogo de servicios que prevé la Ley de Promoción de la Autonomía Personal, existen dos grupos de prestaciones; por una parte, las que ya existían y venían prestándose por los servicios sociales, en las que la principal dificultad es conocer, porque las Comunidades autónomas faciliten el dato, cuáles son prestaciones nuevas al amparo de la Ley, y cuáles ya se prestaban con anterioridad a su entrada en vigor. Y, por otro lado, existen prestaciones económicas que no existían con anterioridad a 2007.

Respecto a éstas últimas, las cuantías máximas de las prestaciones descritas en las páginas anteriores que están vigentes en la actualidad son las siguientes:

GRADOS NIVELES	Prestación vinculada al servicio	Prestación para cuidados familiares	Prestación asistencial personal
Grado III nivel 2	811,98	506,96	811,98
Grado III nivel 1	608,98	405,99	608,98
Grado II nivel 2	450,00	328,36	

INCOMPATIBILIDADES

El catálogo de servicios que plantea la Ley de Promoción de la Autonomía Personal, en ocasiones es incompatible con algunas prestaciones percibidas previamente por las personas en situación de dependencia con anterioridad a la aprobación de la ley y que venían prestándose por los servicios sociales y aún siguen haciéndolo. Por eso, el Real Decreto 272/2007, de 8 de junio, regula la intensidad y protección de los servicios, así como sus incompatibilidades. Es una norma bastante genérica, permitiendo que cada Comunidad Autónoma establezca sus referencias concretas a estos aspectos, con el riesgo de establecer un sistema de protección en el que sean patentes las desigualdades territoriales. No obstante, plantea las siguientes incompatibilidades entre servicios:

El servicio de atención residencial permanente es incompatible con el servicio de teleasistencia, con el de ayuda a domicilio y con el de centro de noche. En los demás servicios se estará a lo que determine la Comunidad Autónoma.

En aquellos casos en los que el beneficiario sea titular de otras prestaciones de análoga naturales se deducirán:

- Complemento de gran invalidez.
- Complemento de asignación económica por hijo a cargo de 18 años con un grado de minusvalía igual o superior al 75.
- Complemento por necesidad de tercera persona de la pensión no contributiva.
- Subsidio de ayuda a tercera persona (LISMI).

El desarrollo reglamentario por parte de las regiones está siendo muy desigual, de forma que los criterios adoptados en cada uno de los territorios difieren notablemente en lo que a incompatibilidades se refiere.

4. DIFERENCIAS TERRITORIALES Y SU FINANCIACIÓN

El proceso de implantación de la Ley ha sido muy arduo y extraordinariamente lento, pues cuando habían transcurrido dieciocho meses de su aprobación sólo habían percibido alguna prestación el 10% de los solicitantes. Las expectativas que generó esta Ley fueron tan altas que ya en el mes de abril de 2007 había 16.200 solicitudes y en el mes de junio se registró el mayor número, 81.271, lo que supone un 12,67% de las peticiones.

Pero quizás uno de los aspectos más polémicos en el desarrollo de la ley ha sido la cofinanciación entre el Estado central y las Comunidades autónomas. Los recursos económicos destinados a la "dependencia" y, sobre todo, saber si este sistema está planteado como un

modelo de financiación sostenible en el tiempo es un debate abierto tras dos años de aplicación de la Ley. Una de las cuestiones más debatidas son las desigualdades territoriales que surgen, así como las fricciones entre los territorios autonómicos y el Estado central. Son patentes las discrepancias entre los criterios de financiación que las comunidades autónomas tienen y la más que suficiente financiación que, a juicio del Estado, aporta para llevar a cabo el desarrollo de la ley. El Estado central es el encargado de establecer un contenido mínimo común en lo que a reconocimiento de derechos se refiere.

Durante los primeros meses de aplicación de la Ley las múltiples dificultades con las que las Comunidades Autónomas se han encontrado ha provocado que se incluyeran como nuevas prestaciones las que ya se estaban reconociendo con anterioridad a la aprobación de la Ley, por lo que se hace imposible realizar un estudio cuantitativo serio sobre la eficiencia de la misma.

Como se dijo al principio, la aplicación de la Ley se ha llevado a cabo por las diferentes Comunidades autónomas, lo que, unido a que las prestaciones sociales que se venían dotando por las regiones presentaban niveles muy distintos, hace que las desigualdades territoriales en el grado de implantación de la ley de dependencia sea muy distinto, como se muestra en el cuadro siguiente;

Solicitudes a 1 de abril de 2009

Ámbito Territorial	Solicitudes registradas (Nº)	Solicitudes registradas (%)	Población (Nº)	Población (%)	S. r. % población
Andalucía	270.914	30,31	8.202.220	17,77	3,30
Aragón	32.392	3,62	1.326.918	2,87	2,44
Asturias	26.645	2,98	1.080.138	2,34	2,47
Baleares	17.648	1,97	1.072.844	2,32	1,64
Canarias	20.687	2,31	2.075.968	4,50	1,00
Cantabria	14.810	1,66	582.138	1,26	2,54
Castilla y León	43.320	4,85	2.557.330	5,54	1,69
Castilla-La Mancha	51.862	5,80	2.043.100	4,43	2,54
Cataluña	125.704	14,07	7.364.078	15,95	1,71
Valencia	38.332	4,29	5.029.601	10,90	0,76
Extremadura	27.428	3,07	1.097.744	2,38	2,50
Galicia	64.920	7,26	2.784.169	6,03	2,33
Madrid	40.455	4,53	6.271.638	13,59	0,65
Murcia	23.240	2,60	1.426.109	3,09	1,63
Navarra	18.398	2,06	620.377	1,34	2,97
País Vasco	65.155	7,29	2.157.112	4,67	3,02
La Rioja	9.127	1,02	317.501	0,69	2,87
Ceuta y Melilla	2.675	0,30	148.837	0,32	1,80
TOTAL	893.712	100,0	46.157.822	100,00	1,94

Los porcentajes de solicitudes se realizan sobre el total de la población y no sobre la población dependiente en la Comunidad Autónoma de referencia, ya que se carece del número exacto de dependientes que existen en España y en cada uno de los territorios. Por esto, los porcentajes de solicitudes registradas no pueden considerarse muy significativos en todas las regiones, pues responden al cociente solicitudes registradas/ total de población y no a solicitudes registradas/total de población dependiente, que sería lo adecuado, especialmente para aquellos territorios donde su población significativamente más joven.

No obstante, el número de solicitudes es muy distinto, según el territorio del que se trate. Así, Comunidades Autónomas que superan la media española, en lo que a solicitudes registradas se refiere (1,94 media española) son; Andalucía, 3,30, Aragón; 2,44 o Asturias, 2,47. Mientras que regiones como Valencia sólo ostenta el 0,76% de las solicitudes presentadas, o Madrid el 0,65% o Canarias el 1% del total.

Por otra parte, hay que decir que durante los primeros meses de aplicación de la ley hubo grandes dificultades en lo que a valoración de las personas en situación de dependencia se refiere, por las razones antes apuntadas. Esto ha hecho que en el momento actual del total de solicitudes haya un 41,78% de ellas pendientes de identificar un PIA, con el consiguiente retraso en la percepción de la prestación que le corresponda. Además, y en consonancia con las desigualdades territoriales que en la aplicación de la ley se están dando, hay Comunidades Autónomas en las que las personas que no tienen aún un PIA identificado está por encima de la media nacional, como son: Canarias (80,09%), Cataluña(55,19%) o Castilla la Mancha (55,04%), y otras como Cantabria (34,52%), Castilla y León (32,40%) o Andalucía (25,50%) en las que las personas pendientes de un PIA están por debajo de la media nacional.

Ámbito territorial	*Pendientes de identificar en PIA*	*%*	*TOTAL*
Andalucía	42.793	25,50	167.810
Aragón	11.776	47,68	24.700
Asturias	7.552	45,12	16.739
Islas Baleares	8.078	65,94	12.251
Canarias	9.128	80,09	11.397
Cantabria	3.809	34,52	11.033
Castilla y León	10.930	32,40	33.734
Castilla la Mancha	20.488	55,04	37.221
Cataluña	57.493	55,19	104.168
Valencia	14.865	45,55	32.631
Extremadura	10.206	58,06	17.577
Galicia	14.838	38,21	38.835
Madrid	18.928	52,90	35.783
Murcia	17.581	75,70	23.224
Navarra	2.478	21,66	11.440
País Vasco	6.792	16,81	40.395
La Rioja	4.224	52,97	7.974
Ceuta y Melilla	648	38,07	1.702
TOTAL	**262.607**	**41,78**	**628.614**

CONCLUSIONES

Las prestaciones reconocidas en la Ley de Promoción de la Autonomía Personal podrían haber constituido el cuarto pilar del Estado del Bienestar, de haberse conocido previamente a la aplicación de la Ley cual era la realidad española, en lo que a dependencia se refiere. Sin embargo, pocos son los territorios los que saben con cierta fiabilidad el número de mayores y discapacitados, en definitiva, el número de dependientes que habitan en su territorio, y de éstos, el número de dependientes con derecho a percibir alguna prestación.

A lo largo de los dos años de implantación de la Ley 39/2006 se han producido aspectos francamente mejorables y desviaciones importantes de lo que inicialmente estaba previsto en la norma, ya que el desarrollo del SAD en los diferentes territorios está siendo muy desigual, tanto en los procedimientos como en las prestaciones percibidas. Estas graves deficiencias son las que están haciendo que la puesta en marcha de la ley haya sido un proceso crítico

La precipitación en la puesta en marcha del SAAD ha hecho que en múltiples ocasiones, y como única salida a muchos PIAS, se ha convertido la excepcionalidad de las prestaciones económicas para cuidados en el entrono familiar en una prestación habitual.

Por otra parte, uno de los mayores problemas con los que se encuentra la implantación de la Ley es que algunas Comunidades Autónomas se han limitado a "traspasar" a los beneficiarios de los servicios sociales al SAAD, fundamentalmente en lo que se refiere a plazas residenciales. En algunos territorios se está vaciando de contenido y, por tanto, de presupuesto, al Sistema de Servicios Sociales para atender casi en exclusiva a las situaciones de dependencia.

Uno de los problemas financieros de la Ley, aún sin resolver es la forma en la que se aplicará el; vulgarmente denominado "copago", pues, como en otros puntos, en éste no existe unanimidad entre las regiones y esto ahonda en la falta de solidaridad interregional en la aplicación de la Ley, pues es potestativo para las Comunidades Autónomas el elevar el mínimo de fondos establecido al efecto. Esta pluralidad regional para la determinación de la capacidad económica del dependiente, de cara a hacerle partícipe de la financiación de los servicios, está generando gran incertidumbre e inseguridad. Pero más llamativo aún es el "abuso" de la prestación que, en principio era una situación de excepcionalidad, como es la prestación económica vinculada al servicio.

Por otra parte, si los derechos reconocidos en la Ley se plantean como derechos subjetivos, entonces, cualquier beneficiario potencial podrá ejercer su derecho por la vía administrativa o la judicial, sin embargo, hasta el momento no se conoce reclamación alguna de ningún potencial beneficiario, incluso en las comunidades donde no se reconoce el efecto retroactivo de las prestaciones no satisfechas por el fallecimiento del potencial beneficiario.

Una de las cuestiones básicas que se han planteado en estos dos años de puesta en marcha de la Ley es la necesidad de no distinguir entre "dependencia" y "no dependencia", ya que la posible incompatibilidad de ayudas y sobre todo la colisión éntrelos servicios del SAAD y los que tradicionalmente venían prestando los servicios sociales, hacen que se dificulte enormemente el seguimiento de la implantación de la ley. Ya que existían personas que eran atendidas en los Servicios Sociales prestados por los niveles municipales y, sin embargo, ahora son excluidos del sistema por no considerarse que debe estar incluidos en el sistema de dependencia.

6. BIBLIOGRAFÍA Y LEGISLACIÓN BÁSICA

Principales normas aplicables previas a la aprobación de la Ley de Autonomía Personal y Atención a las Personas en Situación de Dependencia y algunas fechas significativas se exponen a continuación.

Normativa en los años ´70 y ´80.

- Declaración de los Derechos de las Personas con retraso mental. 1971
- Declaración de los Derechos de las personas con discapacidad 1975
- Clasificación Internacional de Deficiencias, Discapacidades y Minusvalías, OMS, 1980
- Año Internacional de las Personas con Discapacidad. 1981
- Programa de Acción Mundial para las personas con discapacidad Asamblea ONU diciembre 1982
- Ley de Integración Social de Minusválidos. España 1982
- Convenio 159 de la OIT relativo a la rehabilitación profesional y el empleo de las personas con discapacidad 1983
- Decenio de ONU para las personas con discapacidad 1983-1992
- Programa HELIOS a favor de minusválidos 1988-1991

Años ´90.

- Leyes de Estadounidenses con Discapacidad. 1990.
- Nueva definición del retraso mental. Asociación Americana de Retraso Mental, 1992.
- Recomendación R (92)6 sobre una política coherente para la rehabilitación de las personas con discapacidad. Consejo de Europa, 1992
- Normas uniformes para la Igualdad de Oportunidades de las personas con minusvalía. ONU, 1993

- Comunicación de la Comisión y Resolución del Consejo sobre la Igualdad de Oportunidades de las personas con minusvalía. U.E. 1996
- Clasificación Internacional del Funcionamiento, la Discapacidad y la Salud, OMS, 2001
- Ley de Igualdad de Oportunidades, No discriminación y Accesibilidad Universal de las personas con discapacidad. España 2003.
- Decreto 307/2007, de 18 de diciembre, del procedimiento para el reconocimiento de la situación de dependencia y del derecho a la prestaciones del sistema para la autonomía y atención a la dependencia, y de la composición y funciones de los órganos de coordinación, de atención y valoración de la situación de dependencia
- Orden de 9 de mayo de 2008, de la Consejería de Bienestar Social, por la que se modifica la Orden de 24-20-2007, por la que se establecen las intensidades de protección de los servicios y el régimen de compatibilidad de los servicios y prestaciones económicas del sistema para la autonomía y atención a la dependencia en Castilla la Mancha.
- Resolución de 2-11-2007, de la Dirección General de Coordinación de la Dependencia, por la que se delegan competencias en las Delegaciones Provinciales de Bienestar Social en materia de reconocimiento de la situación de dependencia.
- Resolución de 24 de abril de 2007, de la Consejería de Bienestar Social, por la que se aprueban los modelos de solicitudes de reconocimiento de la situación de dependencia del derecho a las prestaciones del sistema y del informe social.

BARRIGA, L. A. y otros (2008a): "I Dictamen del Observatorio. Análisis y evaluación del proceso de implantación de la Ley de Autonomía Personal y Atención a las Personas en Situación de Dependencia". Asociación Estatal de Directores y Gerentes en Servicios Sociales. Ciudad Real.

(2008b): "II Dictamen del Observatorio. Desarrollo e implantación territorial de la Ley de Promoción de la Autonomía Personal y Atención a las Personas en Situación de Dependencia.". Asociación Estatal de Directores y gerentes en Servicios Sociales. Alcorcón (Madrid).

ANEXO

Ámbito territorial	Prevención Dependencia y Promoción A. Personal	Teleasistencia	Ayuda a Domicilio	Centros de Día/Noche	Atención residencial	P.E. Vinculada al servicio	P.E. Cuidados Familiares	P.E. Asistencia Personal	Total
Andalucía	7	12.964	36.884	5.706	13.353	2.168	53.910	25	125.017
Aragón	0	0	0	561	2.625	2.068	7.670	0	12.924
Asturias	0	27	575	468	3.044	468	4.604	1	9.187
Baleares	0	0	0	228	858	31	3.053	3	4.173
Canarias	0	0	0	371	432	11	1.455	0	2.269
Cantabria	0	1	346	607	2.170	0	4.100	0	7.224
Castilla y León	792	243	1.797	1.168	6.775	4.495	7.503	31	22.804
Castilla la Mancha	24	672	124	308	5.137	296	10.168	4	16.733
Cataluña	0	2	3.664	590	5.525	8.119	28.760	15	46.675
Valencia	1	0	0	580	10.448	617	6.120	0	17.766
Extremadura	274	281	1.099	313	2.554	1.106	1.744	0	7.371
Galicia	98	98	2.788	712	3.752	1.169	15.356	24	23.997
Madrid	0	0	0	2.186	13.431	11	1.227	0	16.855
Murcia	0	4	0	0	129	181	5.329	0	5.643
Navarra	0	116	0	177	1.161	1.135	6.372	1	8.962
País Vasco	0	3.237	5.422	4.346	9.072	837	10.339	350	33.603
La Rioja	18	212	600	464	825	139	1.492	0	3.750
Ceuta y Melilla	*0*	*72*	*64*	*16*	*64*	*53*	*785*	*0*	*1.054*
TOTAL	***1.214***	***17.929***	***53.363***	***18.801***	***81.355***	***22.904***	***169.987***	***454***	***366.007***

SOCIAL SCIENCE BIBLIOMETRIC MICROINDEX

MARÍA ISABEL DORTA GONZÁLEZ
E-mail: isadorta@ull.es
Departamento de Estadística, Investigación Operativa y Computación
Universidad de La Laguna
PABLO DORTA GONZÁLEZ
E-mail: pdorta@dmc.ulpgc.es
DOLORES ROSA SANTOS PEÑATE
E-mail: drsantos@dmc.ulpgc.es
RAFAEL SUÁREZ VEGA
E-mail: rsuarez@dmc.ulpgc.es
Departamento de Métodos Cuantitativos en Economía y Gestión
Universidad de Las Palmas de Gran Canaria rsuarez@dmc.ulpgc.es

Resumen

El número h se ha convertido en el indicador bibliométrico más empleado para estimar el éxito del trabajo realizado por un investigador y predecir la evolución del impacto de su producción en el futuro. Esto se debe principalmente a dos razones. En primer lugar, a su simplicidad, dado que se trata de un único indicador que combina producción e impacto, y puede ser determinado fácilmente por cualquier investigador. En segundo lugar, a que elimina los sesgos provocados por las colas de la distribución de citas. Un problema asociado al número h es que discrimina poco entre investigadores con una reducida producción o que trabajan en áreas con bajos niveles de impacto promedio, lo que sucede, por ejemplo, en Ciencias Sociales. En este trabajo se proponen soluciones que consideran aquellos artículos y citas que pueden contribuir con una alta probabilidad a incrementar el valor del número h en el futuro.

Palabras clave: bibliometría, política científica, indicadores, producción, impacto.
Área Temática: Economía Social y Discapacidad.

Abstract

The h index has become the most widely bibliometric indicator used to estimate the success of a researcher. This is mainly due to its simplicity, since it is a single indicator that combines production and impact, and can be easily determined by any researcher. The main problem with this index is the low level of discrimination between researchers with reduced working production in scientific disciplines with low average impact, which is common in social sciences. In this paper, some solutions to consider those publications and citations that may contribute to increase in a high probability its value in the future are proposed.

Thematic Area: Social Economics.

1. INTRODUCCIÓN

La investigación científica tiene gran influencia en el desarrollo económico y social de los países. Sin embargo, los recursos que pueden destinarse a esta actividad son limitados. La demanda de incremento en su financiación ha originado la necesidad de evaluar la actividad científica y sus resultados, siendo uno de los más directos las publicaciones científicas. El objetivo de los *indicadores bibliométricos* es el análisis cuantitativo y cualitativo de dichas publicaciones, y entre sus principales aplicaciones se encuentra el área de Política Científica.

Los indicadores bibliométricos ofrecen diferentes medidas a partir del número de artículos publicados por los investigadores en revistas de reconocido prestigio y el número de citas que estos artículos reciben en estas mismas revistas. Este procedimiento tiene un uso muy extendido en la literatura internacional, y se basa en evidencias que indican que la producción científica de calidad se difunde en este tipo de revistas. En ocasiones, se considera una ponderación de la

producción o del impacto (número de citas), dado que éste es un indicador muy sensible al campo de especialidad.

Los indicadores clásicos como el *número de artículos* N_p, el *número total de citas* N_c y el *promedio de citas por artículo* $n_c = N_c/N_p$, de forma individual no reflejan satisfactoriamente el éxito de la carrera profesional de un investigador, dado que existen evidencias empíricas que indican que no hay correlación entre ellos. Además, no son *robustos* en el sentido que se ven afectados por un inusual número de artículos no citados (o poco citados) y por uno o varios artículos anormalmente muy citados. El par (N_p, N_c) puede dar una medida del éxito en el pasado, pero la pregunta más importante es ¿cómo predecir o estimar el éxito futuro de un investigador?

El número h (Hirsch, 2005) estima el número de trabajos importantes publicados por un investigador, incrementando la exigencia a la vez que aumenta su valor. Existe correlación entre el número h y el éxito de un investigador apreciado por sus pares (Hirsch, 2005) y tiene, además, gran valor predictivo del éxito futuro (Hirsch, 2007).

El número h es un indicador robusto que considera al mismo tiempo aspectos cuantitativos y cualitativos o de visibilidad. Sin embargo, al comparar las carreras profesionales de dos científicos de un mismo campo, surgen dudas en aquellos casos con valores similares de h en los que las curvas de citas se cortan. Algunos investigadores presentan más altos niveles de citas al comienzo de la curva y más bajos al final. En estos casos, es preciso establecer criterios que permitan discriminar que distribuciones de citas corresponden a carreras más consolidadas y constantes. Este problema del número h es especialmente importante en el caso de los investigadores sociales, dado que los promedios de artículos publicados y citas recibidas son significativamente inferiores a otras áreas.

Por tanto, el problema a resolver es el siguiente: ¿Cómo discriminar entre dos investigadores con valores similares de h y curvas de citas que se cortan?. En este trabajo se sugieren soluciones que consideran aquellos artículos y citas que pueden contribuir con una alta probabilidad a incrementar el valor de h en el futuro.

El apartado 2 está dedicado a describir los principales indicadores bibliométricos empleados en la evaluación de la carrera profesional de un investigador, y a analizar aquellos que son robustos. En el apartado 3 se formulan algunos problemas en relación al número h, y se sugieren posibles soluciones. Finalmente, el apartado 4 se dedica a conclusiones y consideraciones finales.

2. MICROINDICADORES

2.1 Indicadores bibliométricos a nivel micro

Existe consenso en admitir que los artículos son el resultado más directo de la investigación, especialmente cuando se consideran aquellas revistas en las que existe un sistema de selección que garantiza la calidad y originalidad de los trabajos. Aunque el juicio de expertos ha demostrado ser el método más apropiado para valorar la calidad de las aportaciones al área de los artículos, el sistema presenta algunas limitaciones como el carácter subjetivo y su elevado coste. En este contexto, los indicadores bibliométricos representan *indicios objetivos de calidad*, que pueden complementar al juicio de expertos.

Los indicadores bibliométricos más empleados se basan en el recuento de las publicaciones y las citas recibidas por los trabajos publicados, así como el impacto de las revistas de publicación. El *número de publicaciones* es un indicador cuantitativo que no valora la calidad del contenido. Es evidente que no todas las publicaciones tienen la misma importancia ni suponen la misma contribución al avance científico. En la búsqueda de indicadores cualitativos, capaces de valorar estos aspectos, se han propuesto el *número de citas* que reciben los trabajos, el *factor de impacto* de la revista de publicación y la *posición relativa* de la revista en los listados temáticos ordenados por factores de impacto del *Journal Citation Reports* (JCR) atendiendo a la base de datos ISI-Web of Science de Thomson Reuters. Sin embargo, estos

indicadores han quedado reducidos, después de una amplia polémica[1], a indicadores del impacto, influencia o visibilidad de la investigación, que es un indicador indirecto de la calidad de las publicaciones.

El factor de impacto de una revista es el número medio de citas de los trabajos publicados en los dos años naturales anteriores. Este indicador fue creado por Garfield para evaluar las revistas de investigación. Quienes defienden su aplicación en la evaluación de investigadores, justifican este procedimiento en base a la creencia que aquellas revistas con mayores factores de impacto exigen mejores manuscritos, lo cual es en parte cierto al comparar trabajos del mismo campo cuando los factores de impacto de las revistas son significativamente diferentes. Sin embargo, la aplicación a investigadores que trabajan en distintos campos no es válida, dado que el factor de impacto depende de parámetros bibliométricos que no tienen que ver con la calidad como, por ejemplo, el número promedio de citas bibliográficas en las revistas del área. Por otra parte, las mejores revistas publican muchos trabajos de escaso éxito y, a la inversa, muchos trabajos de gran éxito se publican en revistas que no están en los primeros puestos de los listados por factor de impacto. Estudios empíricos indican que no existe correlación entre el factor de impacto de una revista y el número de citas de un trabajo concreto en esa revista, por lo que no puede emplearse el factor de impacto para predecir el éxito de un artículo. La regla 80/20 indica que el 80% de las citas de muchas revistas son obtenidas por el 20% de los trabajos (Garfield, 2007). Como consecuencia, existe un rechazo generalizado a evaluar la calidad de un trabajo por el factor de impacto de la revista (*Nature*, 2005), sobre todo cuando determinar las citas reales de los trabajos está hoy en día al alcance de cualquier investigador.

En España, la evaluación basada en el número de publicaciones y en el factor de impacto de las revistas, ha provocado que los investigadores asocien el éxito de la investigación con publicar en revistas con alto factor de impacto, aunque no estén correlacionados (Imperial y Rodríguez-Navarro, 2007; Garfield, 2007). Este hecho estimula, exclusivamente, la producción y no su contenido, favoreciendo la publicación de artículos que no contienen un avance real y significativo del conocimiento.

En la valoración de las publicaciones científicas se suele tener en cuenta, además, algunos aspectos como: el número de páginas publicadas (en ocasiones normalizadas por el promedio de caracteres), el número de autores de cada trabajo y el número de afiliaciones de los autores (en el caso de las evaluaciones institucionales). El número de páginas es significativo sólo cuando se consideran artículos publicados en revistas de gran prestigio (blue ribbon o top journals), en las que la extensión de los artículos refleja el esfuerzo realizado por los investigadores. Además, es conocido que existen culturas distintas en los diferentes campos de especialidad (por ejemplo, los trabajos de naturaleza teórica suelen ser más cortos que los de aplicaciones empíricas), y que algunas revistas establecen límites estrictos al número de páginas de cada artículo, mientras que otras no lo hacen. El período de referencia del análisis es también una variable de decisión importante. Períodos muy cortos (inferiores a tres años) pueden dar valoraciones coyunturales con escaso contenido informativo, mientras que un período de cinco a diez años permite valoraciones mucho más consistentes de la labor de un investigador. Períodos más amplios, entre 25 y 30 años, permiten apreciar la solidez de la investigación realizada a largo plazo, aunque, en estos casos, los investigadores noveles no están en igualdad con los demás, y los valores históricos pueden distorsionar la realidad actual del investigador.

2.2 Indicadores bibliométricos tipo *h*

Analizar el *rendimiento* de los investigadores y evaluar la *calidad* de los trabajos científicos es un tema complejo, dado que los indicadores clásicos no son robustos, y en la evaluación por pares generalmente abundan los conflictos de intereses. Un indicador es *robusto* cuando no se

[1] Se ha observado que algunas publicaciones que reúnen criterios de calidad atendiendo a la opinión de expertos, apenas reciben citas, es decir que producen escaso impacto. La razón es que el número de citas depende de su calidad, pero también de otras variables entre las que figura el prestigio del autor y/o la institución, la actualidad del tema, la lengua empleada y la revista de publicación, que condicionan la mayor o menor difusión del trabajo.

ve afectado ni por un inusual número de artículos no citados (o poco citados) ni por uno o varios artículos anormalmente muy citados.

Aunque el número h han revolucionado el campo de la evaluación científica de pequeñas unidades, es habitual recurrir al factor de impacto de las revistas en aquellas áreas con bajos índices de producción y citas –como sucede en Ciencias Sociales-, lo que incentiva hábitos de publicación no deseados y alejados de la excelencia, como se indicó en el epígrafe anterior. Por tanto, el único procedimiento para valorar la excelencia de un investigador sin recurrir a la evaluación por pares es el análisis de la frecuencia de las citas de sus trabajos. La suma de todas las citas recibidas por los trabajos publicados por un determinado autor refleja, en muchos casos, el éxito de su carrera científica, aunque las grandes colaboraciones introducen sesgos muy importantes, al recoger muchas citas pero derivadas del trabajo de un gran número de investigadores.

En base a la idea anterior, Hirsch propone un sencillo y elegante indicador robusto: "A scientist has index h if h of his or her N_p papers have at least h citations each and the other $(N_p - h)$ papers have $\leq h$ citations each" (Hirsch, 2005, p.1). Este indicador estima el número de trabajos importantes publicados por un investigador, incrementando la exigencia a la vez que aumenta su valor, y proporciona una cota inferior h^2 de N_c. El número total de citas N_c es generalmente mucho mayor que h^2 (entre 3 y 5 veces según ha estimado el propio Hirsch) ya que esta cantidad subestima las citas de los h artículos más citados e ignora los artículos con menos de h citas. Este autor argumenta que N_c puede ser inflado por un pequeño número de importantes éxitos que podrían deberse al efecto de la coautoría. Existe correlación entre el número h y el éxito de un investigador apreciado por sus pares (Hirsch, 2005) y tiene además gran valor predictivo del éxito futuro (Hirsch, 2007). Además, correlaciona fuertemente con el número tal de citas (Cronin y Meho, 2006; Saad, 2006; Scimago, 2007; Hirsch, 2007) y tiene incluso mayor poder predictivo futuro sobre N_p y N_c que las propias variables (Hirsch, 2007).

A la hora de diferenciar entre investigadores activos e inactivos se ha propuesto la tasa de crecimiento $h'(t)$ (Rousseau y Ye, 2008), de tal manera que una función $h(t)$ convexa es preferible a una lineal o cóncava. Hirsch (2007) ha estimado empíricamente que $h(t) = a \cdot t$ es una función lineal del tiempo, mientras que $N_c(t)$ es una función cuadrática. De esta forma, Hirsch propone comparar las tasas de crecimiento $a = h/t$, siendo t los años transcurridos desde la publicación del primer artículo.

Al estimar el éxito de un investigador por el número h se desprecian las colas de la distribución de citas. Las colas de la distribución corresponden a aquellas publicaciones que se alejan del impacto promedio, ya sea porque han sido muy citadas (cola superior en la distribución de citas U), o porque apenas lo han sido (cola inferior L). Una U larga puede corresponder a trabajos con muchos autores en grandes líneas de investigación que generan muchas citas. Una L larga indica un gran número de publicaciones de escaso impacto, lo que podría indicar una sobreproducción.

El número h ha sido extensamente estudiado (ver revisión de Bornmann y Daniel, 2007) y tiene importantes propiedades matemáticas (Glänzel, 2006). Ha sido aplicado en diferentes áreas temáticas, desde la física (Hirsch, 2005) hasta la ciencia de la información (Cronin y Meho, 2006), y también en diferentes áreas geográficas, como Brasil (Batista y otros, 2006). Algunas de sus variantes resuelven muchos de los problemas que se han descrito en la literatura sobre este indicador. El número de citas depende de dos parámetros bibliométricos que son diferentes entre campos y que no tienen que ver con la calidad, como es el número medio de referencias por trabajo y el número de autores. El primero de estos sesgos puede corregirse dado que el h máximo en cada campo está fuertemente correlacionado con los factores de impacto de las primeras revistas del área, lo que permite estimar un número h de referencia en cada especialidad (Imperial y Rodríguez-Navarro, 2007). El segundo puede reducirse con un indicador complementario h_1, que se obtiene al dividir h por el número medio de autores de

esos h artículos (Batista y otros, 2006), e indicaría el número de artículos que un investigador podría haber publicado en solitario a lo largo de su carrera con al menos h_1 citas.

Sin embargo, lo que para Hirsch es una ventaja de su indicador (despreciar la cola superior de la distribución de citas), para otros autores no está justificado. El número g (Egghe, 2006) considera las citas totales de los g artículos más citados, y viene a representar un promedio de citas entre estos g artículos. Ordenados los artículos de forma decreciente del número de citas que reciben, g es el mayor valor de tal forma que los primeros g artículos tienen, en su conjunto, al menos g^2 citas. En este caso se incluye U, por lo que obviamente $g \geq h$.

Finalmente, el número b (Bornmann y otros, 2007) indica el número de artículos incluidos en el 10% de los más citados en su campo, determinados a partir de los percentiles ISI-ESI, por ejemplo.

3. ALGUNOS PROBLEMAS SIN RESOLVER ASOCIADOS AL NÚMERO h Y APORTACIONES

3.1 Formulación matemática de los indicadores tipo h

El procedimiento de obtención de la distribución y la curva de citas consiste en ordenar los trabajos de un autor de forma decreciente en función del número de citas recibidas por cada trabajo. La representación del número de citas frente a la posición en la lista da lugar a la *distribución de citas*, mientras que la línea de aspecto hiperbólico que une estos puntos es la *curva de citas*[2]. El número h aparece cuando el rango de publicaciones (posición en la lista) supera o iguala al número de citas, es decir, la parte entera del punto de intersección de la curva de citas con la recta que representa el rango (bisectriz del primer cuadrante $f(x) = x$). Esto significa que el autor tiene h trabajos con al menos h citas cada uno.

Sea N_p el número total de publicaciones (artículos en revistas científicas) de un autor, p_i, $i = 1,...,N_p$, el conjunto de publicaciones y c_i el número de citas recibidas por la publicación i en cierta base de datos bibliográfica (p.e. ISI-WoS, Scopus, etc.). Sea $c_1 \geq c_2 \geq \cdots \geq c_{N_p}$ las citas recibidas por las publicaciones en orden decreciente del número de citas y $N_c^j = \sum_{i=1}^{j} c_i$ las citas agregadas de las j publicaciones más citadas. Sea además $N_c = \sum_{i=1}^{N_p} c_i$ las citas totales y $n_c = N_c / N_p$ el promedio de citas por artículo.

El número h proporciona una cota inferior de citas $H = h^2$ y no refleja el volumen de citas en las colas de la distribución de citas: *cola superior* U de los artículos más citados y *cola inferior* L de los artículos menos citados. Obsérvese que se tienen las siguientes relaciones:

$N_c = H + U + L,$

$U = N_c^h - H,$

$L = N_c - N_c^h.$

El peso de las colas viene dado por $\frac{N_c}{H}$. Si $\frac{N_c}{H} < 3$ podemos decir que la distribución tiene colas poco pesadas, mientras que si $\frac{N_c}{H} > 5$ la distribución tiene colas muy pesadas. La relación entre las colas viene dada por $\frac{U}{L}$.

El número g es el menor entero tal que $g^2 \geq N_c^g$, mientras que la *mediana* m de la distribución de citas es el menor entero tal que $N_c^m \geq \frac{1}{2} N_c$. El número g garantiza una cota

[2] La función está definida para valores enteros y la curva es una guía que no indica continuidad.

inferior de citas $G = g^2 \geq H$ dado que $g \geq h$ al incluir las citas de U. En este caso, sí se consideran todas las citas de los g artículos más citados. Sin embargo, puede presentar el problema asociado a la coautoría descrito en la sección anterior.

Como se ha indicado en la sección anterior, los indicadores bibliométricos tratan de medir el volumen de publicaciones y el impacto que estas tienen en la comunidad científica, resultando de interés contar con un único indicador a la hora de comparar la trayectoria profesional de dos investigadores. Sin embargo, aunque el número h permite estimar la visibilidad promedio de los trabajos de un investigador, en ciertas ocasiones, dos distribuciones de citas bien diferenciadas pueden dar lugar a un mismo número h. No es posible, entonces, distinguir entre investigadores con índices similares, pero con perfiles muy diferenciados, como por ejemplo entre un investigador con pocos trabajos muy citados y otro con muchos trabajos de escaso impacto (podrían incluso tener citas totales similares). Además, los bajos niveles de producción y citas en algunas áreas (en Ciencias Sociales, por ejemplo) se traducen en números h significativamente inferiores a los de otras áreas, por lo que la cota inferior del volumen de citas garantizadas por el indicador es mucho menor y su capacidad de discriminar se ve reducida.

El Cuadro 1 resume la notación empleada y describe algunos indicadores bibliométricos. El Cuadro 2 y el Gráfico 1 muestran para dos ejemplos reales diferentes medidas que describen la forma de la curva de citas.

3.2 Baja correlación con el juicio de expertos sociales e índice de Gini

En general los números h en Ciencias Sociales son muy inferiores a los de otras ciencias. Así, mientras en Física el h máximo es del orden de 100, y en Biología y Biomedicina del orden de 150 (Hirsch, 2005), en Economía el h máximo o de referencia es del orden de 50. Como se ha indicado en la sección anterior, el h máximo en cada campo está fuertemente correlacionado con los factores de impacto de las primeras revistas del área, lo que permite estimar un h de referencia en cada especialidad $h_R \approx 16 + 11 \cdot f$, donde f es el factor de impacto promedio de las primeras revistas del área (Imperial y Rodríguez-Navarro, 2007). Considerando el área de Economía en su conjunto, la categoría "economics" del JCR 2007 incluye 191 revistas, cuyos factores de impacto máximos varían aproximadamente entre 2 y 4 (aunque el mayor es 4,190 tan sólo doce superan el 2). La única revista española en esta categoría es *Investigaciones Económicas* y su factor de impacto es 0,268. Estos datos indican que el índice de referencia internacional en Economía es del orden de 50 ($h_R = 16 + 11 \cdot 3 = 49$), mientras que en España es del orden de 20 ($h_{RE} = 16 + 11 \cdot 0{,}268 = 18{,}9$).

Aunque algunos economistas de prestigio atendiendo al juicio de expertos, tienen índices en torno al de referencia (por ejemplo Joseph E. Stiglitz, premio Nobel en 2001 y $h = 50$), los índices de otros economistas de prestigio son muy inferiores (por ejemplo Edmund S. Phelps, Nobel en 2006 y $h = 20$). Por este motivo, el número h correlaciona poco con el juicio de expertos y se hace necesaria alguna medida del grado de consolidación y solidez de la carrera profesional de un investigador social. Esta medida de *regularidad en el impacto de las publicaciones* puede reflejarse a través de la curva de Lorenz y también puede ser de interés en otras áreas no sociales, sobre todo en aquellos casos en los que existen dudas razonables entre dos investigadores con similares números h.

La *curva de Lorenz* es una gráfica de concentración acumulada de la distribución del impacto (citas). Su expresión en porcentaje es el *índice de Gini*, que puede ser empleado para medir el grado de desigualdad en la distribución de citas entre artículos. El coeficiente de desigualdad de Gini se calcula como una razón de las áreas que vienen definidas por la curva de Lorenz. Si el área entre la línea de perfecta igualdad (bisectriz del cuadrado) y la curva de Lorenz es A y el área por debajo de la curva de Lorenz es B, entonces el coeficiente de Gini viene dado por $A/(A+B)$. Esta ratio es siempre un número entre 0 y 1. Un coeficiente 0 indica igualdad máxima en la distribución de citas, mientras que un coeficiente 1 indica desigualdad máxima. Además, existe una relación con la mediana m de tal manera que cuanto menor sea ésta mayor será el coeficiente. El índice de Gini se obtiene multiplicando este coeficiente por 100.

La curva de Lorenz es el resultado de unir los puntos (i, N_c^i), $i = 1,...,N_p$, pero esta vez ordenados de forma *creciente del número de citas*. Es estrictamente creciente si consideramos exclusivamente los artículos que han sido citados alguna vez y, en términos de porcentaje (cada eje varía entre 0 y 100), los puntos

$$\left(\frac{i}{N_p} \cdot 100,\ \frac{N_c^i}{N_c} \cdot 100 \right),\ i = 1,...,N_p,$$

proporcionan la distribución porcentual del acumulado de citas. Una forma de ajustar por exceso el índice de Gini es mediante la expresión

$$\left(\frac{1}{2} - \frac{1}{N_p N_c} \sum_{i=1}^{N_p - 1} N_c^i \right) \cdot 100\%.$$

Como puede observarse en el Gráfico 2, diferentes distribuciones de citas pueden dar lugar a una misma curva de Lorenz y, por tanto, a índices de Gini idénticos. Sin embargo, dado un número de publicaciones y citas, a menor índice de Gini mayor igualdad en el reparto de las citas, lo que indica una carrera más consolidada y constante.

3.3 Discriminación para índices similares (no significativamente diferentes)

Al comparar las carreras profesionales de dos científicos de un mismo campo, surgen dudas en aquellos casos en los que las curvas de citas se cortan. Los Gráficos 3 y 4 muestran dos ejemplos en los que esto sucede. En ambos casos, el primer investigador presenta niveles más altos de citas al comienzo de la curva y más bajos al final, pero ¿qué gráfica sugiere una carrera más consolidada y constante?. El número h del primer investigador puede ser mayor o menor que el del segundo, aunque eso no tiene porqué indicar necesariamente que el primero presente una carrera más sólida ni tampoco que el impacto futuro tenga porqué ser mayor a medio plazo. Estas gráficas reflejan el problema de discriminar entre dos distribuciones con similares números h, pero con diferencias importantes en la relación entre las colas de la distribución U/L. Como puede observarse, a mayor ratio de colas mejor promedio de citas por artículo.

La mediana divide la distribución de citas en dos partes similares y cuando la distribución es simétrica respecto a la bisectriz se tiene $m < h$. El caso $m \square h$ corresponde a una distribución donde $U \square L$, mientras que $m \square h$ corresponde a $U \square L$. El número h no discrimina entre perfiles diferentes dado que no tiene en cuenta las colas de la distribución, ni siquiera para m alejados de h (podría darse el caso de dos distribuciones con el mismo h, una con $U \square L$ y otra con $U \square L$.) Parece razonable, en estos casos, medir, de alguna manera, parte de U con una función creciente de mayor pendiente que $f(x) = x$, con algún proceso iterativo sobre las citas no consideradas en U, o complementar el número h con aquellos artículos que incrementarán con mayor probabilidad su valor futuro. Veamos a continuación este último complemento del número h con más detenimiento.

3.3.1 Indicadores generalizados de entorno o confianza tipo h

La zona próxima a H en la distribución corresponde a aquellas citas que con una mayor probabilidad contribuirán a incrementar el valor de h en el futuro. Por ello, parece razonable relajar la condición para incluir también estas citas. De esta forma, se complementa el número h en aquellos casos en los que existen dudas razonables entre investigadores. Esto puede conseguirse considerando las citas de los artículos en cierto intervalo de la forma

$$[h - \alpha, h + \beta],\ 0 \le \alpha \le h,\ 0 \le \beta \le N_p - h.$$

Ambos parámetros son números naturales. Cuando $\alpha = \beta = 0$, se obtiene el propio h; para $\alpha = h,\ \beta = 0$, se obtiene el número g; para $\alpha = h,\ \beta = N_p - h$, se obtiene N_c.

En aquellos casos en los que h es grande se podría tomar por ejemplo $\alpha = \beta = \lceil h/2 \rceil$, siendo $\lceil x \rceil$ la parte entera de x, resultando el intervalo $\left[\frac{h}{2}, h + \frac{h}{2}\right]$ en el caso particular en que h es un número par. Sin embargo, cuando h es pequeño puede parecer más razonable considerar, por ejemplo, $\alpha = \lceil h/2 \rceil$, $\beta = h$, resultando el intervalo $\left[\frac{h}{2}, 2h\right]$ en el caso particular en que h es un número par.

En general, es posible obtener un conjunto de indicadores generalizados de confianza $H_j = h \pm j,\ j = 1,...,h,$ y estimar empíricamente el valor de j que más correlaciona con el H futuro. El Gráfico 5 muestra esta situación.

Previamente a la aplicación empírica, veamos un ejemplo para ilustrar el problema. Consideremos la distribución de citas de los cuatro autores españoles más productivos, a diciembre de 2008, según la base de datos ISI-SSCI. Los Cuadros 3 y 4 muestran el rango de publicaciones, el número h a diciembre de 2003, h_{2003}, y diciembre de 2008, h_{2008}, y el h generalizado para 2003, $H_j = h \pm j,\ j = 1,...,10$. En el Gráfico 6 se observa cómo existe correlación lineal entre el h pasado, año 2003, y el actual, año 2008 $(r = 0,93)$. También puede observarse que existe correlación no lineal entre H_j y H_{2008}, aunque parece que dicha correlación se reduce a medida que aumenta $j,\ j = 1,...4$. En la aplicación empírica se trataría de hacer algo similar pero con un número muy superior de individuos.

4. CONCLUSIONES

En este trabajo se han formulado algunos problemas que presenta el número h en relación a distinguir entre un grupo de investigadores, con valores similares de h, aquellos con una carrera más consolidada y constante. Se sugieren, además, diferentes indicadores tipo h que consideran aquellos artículos y citas que pueden contribuir con mayor probabilidad a incrementar su valor en el futuro.

REFERENCIAS BIBLIOGRÁFICAS

BATISTA, PD, CAMPITELI, MG, KINOUCHI, O, MARTINEZ, AS (2006) "Is it possible to compare researchers with different scientific interests?", *Scientometrics*, v. 68, n. 1, pp. 179–189.

BORNMANN, L, DANIEL, HD (2007) "What do we know about the h Index?", *Journal of the American Society for Information Science and Technology*, v. 58, n. 9, pp. 1381–1385.

BORNMANN, L, MUTZ, R, DANIEL, HD (2007) "The b index as a measure of scientific excellence. A promising supplement to the h index?", *Cybermetrics*, v. 11, n. 1, paper 6.

CRONIN, B, MEHO, LI (2006) "Using the h-index to rank influential information scientists", *Journal of the American Society for Information Science and Technology*, v. 57, n. 9, pp. 1.275-1.278.

EGGHE, L (2006) "Theory and practise of the g-index", Scientometrics, v. 69, n. 1, pp. 131–152.

GARFIELD, E (2007) "The evolution of the Science Citation Index", *International Microbiology*, n. 10, pp. 65-69.

GLÄNZEL, W (2006) "On the h-index. A mathematical approach to a new measure of publication activity and citation impact", *Scientometrics*, v. 67, n. 2, pp. 315-321.

HIRSCH, JE (2005) "An index to quantify an individual's scientific research output", *Proceedings of the National Academy of Sciences*, v. 102, n. 46, pp. 16.569–16.572.

HIRSCH, JE (2007) "Does the h index have predictive power?", *Proceedings of the National Academy of Sciences*, v. 104, n. 49, pp. 19.193–19.198.

IMPERIAL, J, RODRÍGUEZ-NAVARRO, A (2007) "Usefulness of Hirsch's h-index to evaluate scientific research in Spain", *Scientometrics*, v. 71, n. 2, pp. 271–282.

NATURE EDITORIAL (2005) "Ratings games", *Nature*, v. 436, n. 7.053, pp. 889-890.

ROUSSEAU, R, Ye, FY (2008) "A proposal for a dinamic h-type index", *Journal of the American Society for Information Science and Technology*, v. 59, n. 11, pp. 1.853–1.855.

SAAD, G (2006) "Exploring the h-index at the author and journal levels using bibliometric data of productive consumer scholars and business-related journals respectively", *Scientometrics*, v. 69, n. 1, pp. 117-120.

SCIMAGO (2007) "El índice h de Hirsch: su aplicación a algunos de los científicos españoles más destacados", *El profesional de la información*, v. 16, n. 1, pp. 47-49.

Cuadro 1: Notación y descripción de los indicadores bibliométricos.

Indicador	Descripción	Citas representadas	% de citas
N_p	nº publicaciones	-	-
$i = 1,\ldots,N_p$	índice publicación en orden decreciente del nº de citas	-	-
p_i	publicación i	-	-
c_i	citas publicación i	-	-
$N_c^j = \sum_{i=1}^{j} c_i$	citas agregadas hasta j	-	-
$N_c^h = \sum_{i=1}^{h} c_i$	citas de los h más citados	$H+U$	$\frac{H+U}{N_c}\cdot 100$
$N_c = N_c^{N_p}$	total citas	$H+U+L$	100
$n_c = \frac{N_c}{N_p}$	promedio citas por publicación	-	-
H	cuadrado de citas de Hirsch	h^2	$\frac{h^2}{N_c}\cdot 100$
U	cola superior más citados	$N_c^h - H$	$\frac{N_c^h - H}{N_c}\cdot 100$
L	cola inferior menos citados	$N_c - N_c^h = N_c - H - U$	$\left(1 - \frac{N_c^h}{N_c}\right)\cdot 100$
G	cuadrado de citas de Egghe	g^2	$\frac{g^2}{N_c}\cdot 100$
$U+L$	total colas	$N_c - H$	$\left(1 - \frac{H}{N_c}\right)\cdot 100$
h	número h	$H = h^2$	$\frac{H}{N_c}\cdot 100$
g	número g	$G = g^2$	$\frac{G}{N_c}\cdot 100$
m	mediana $\left(N_c^{m-1} < \frac{N_c}{2};\ N_c^m \geq \frac{N_c}{2}\right)$	$\frac{1}{2}N_c$	50
$\frac{U}{L}$	relación de colas	-	-
$\frac{U}{H}, \frac{U}{H+L}$	proporciones cola superior	-	-
$\frac{N_c}{H}$	peso de las colas	-	-
$\frac{g}{h}$	relación de índices	-	-

Cuadro 2: Indicadores bibliométricos de J.E. Stiglitz y E.S. Phelps según WoS (Octubre 2008)

Indicador	J.E. Stiglitz		E.S. Phelps	
	valor	% citas	valor	% citas
N_p	164	-	58	-
N_c	12016	100	2125	100
n_c	73,27	-	36,64	-
h	50	-	20	-
g	108	-	46	-
m	9	50	5	50
N_c^h	10132	84,3	1982	93,3
H	2500	20,8	400	18,8
U	7632	63,5	1582	74,4
L	1884	15,7	143	6,7
G	11664	97,1	2116	99,6
$U+L$	9516	79,2	1725	81,2
$\frac{U}{L}$	4,05	-	11,06	-
$\frac{U}{H}\left(\frac{U}{H+L}\right)$	3,05 (1,74)	-	3,95 (2,91)	-
$\frac{g}{h}$	2,16	-	2,30	-

Gráfico 1: Distribución de citas de J.E. Stiglitz y E.S. Phelps según WoS (Octubre 2008)

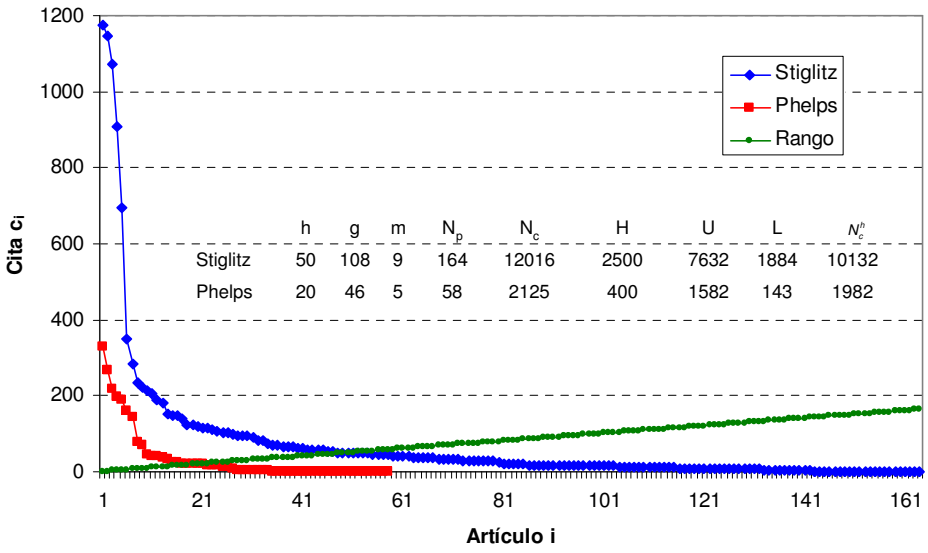


Gráfico 2: Curva de Lorenz e índice de desigualdad de Gini.

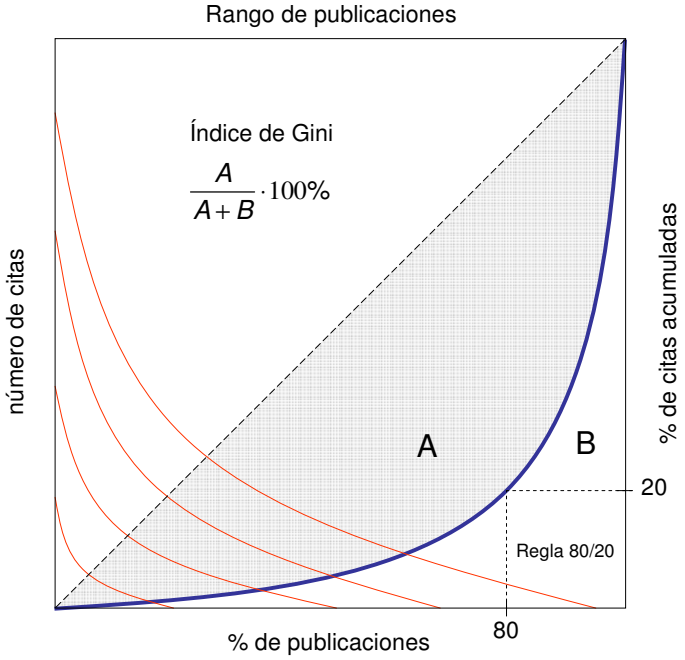


Gráfico 3: Dos curvas de citas que se cortan, con idéntico número *h* y diferente promedio de citas por artículo.

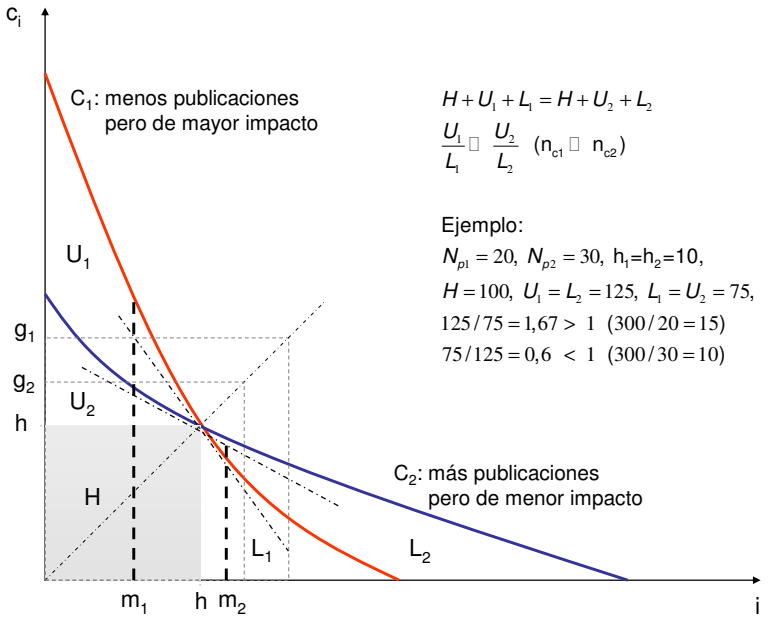


Gráfico 4: Dos curvas de citas que se cortan con similar número y diferente promedio de citas por artículo.

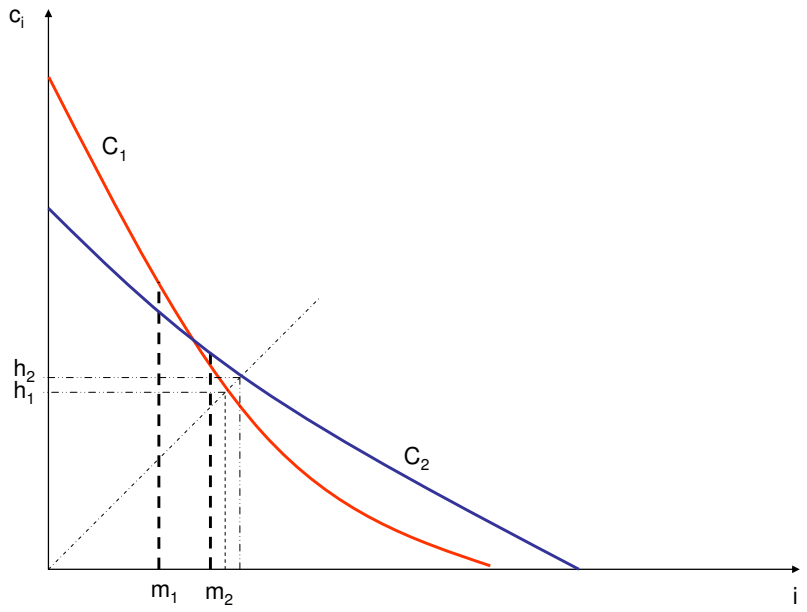
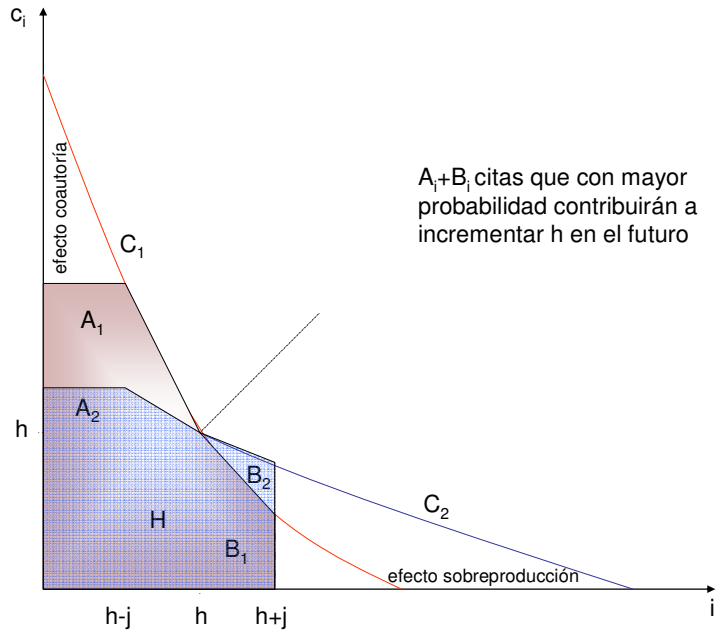


Gráfico 5: Indicadores de confianza tipo *h*.



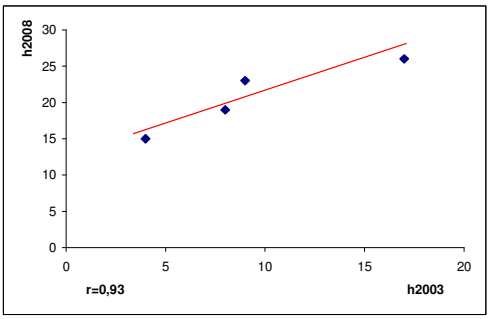
Cuadro 3: Artículos y citas del ejemplo.

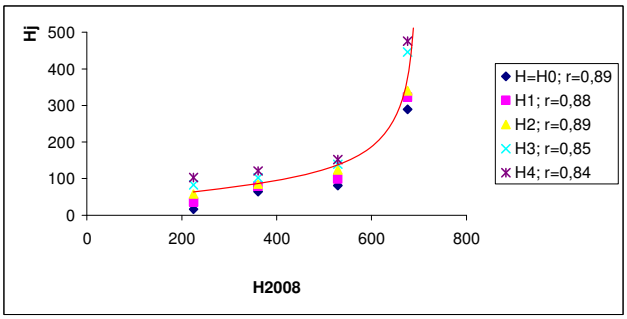
Autores	Rango	Artículos		Citas		Número h		H	
		2003	2008	2003	2008	2003	2008	2003	2008
Alonso, J	[1990-2008]	48	118	925	2708	17	26	289	676
Vieta, E	[1992-2008]	39	88	263	1200	9	23	81	529
Haro, JM	[1992-2008]	18	85	105	733	4	15	16	225
Borrell, C	[1995-2008]	31	78	178	855	8	19	64	361

Cuadro 4: Índice generalizado de entorno para el ejemplo.

Autores	Total citas h generalizado para 2003								
	[h-1,h+1]	[h-2,h+2]	[h-3,h+3]	[h-4,h+4]	[h-5,h+5]	[h-6,h+6]	[h-7,h+7]	[h-8,h+8]	[h-9,h+9]
Alonso, J	323	340	446	476	505	567	592	634	645
Vieta, E	99	124	139	152	216	242	251	256	267
Haro, JM	36	58	83	103	-	-	-	-	-
Borrell, C	79	85	102	121	200	213	217	278	-

Gráfico 6: Correlaciones y diagramas de dispersión.





DISTRIBUCIÓN GEOGRÁFICA DE LA POBLACIÓN INMIGRANTE EN EL PAÍS VASCO Y SU PARTICIPACIÓN EN EL MERCADO DE TRABAJO

LEIRE ALDAZ ODRIOZOLA

E-mail: leire.aldaz@ehu.es

Mª BEGOÑA EGUÍA PEÑA

JUAN RAMÓN MURUA MÚGICA

Universidad del País Vasco

Resumen

This article deals with an analysis of the behavior patterns of the immigrant population residing in the Basque Country in 2008. After describing their main characteristics, it analyzes their distribution throughout the territory, quantifying the degree of concentration in different Basque municipalities. Subsequently, it identifies the main labor areas in which these foreigners are located and provides a comparison with their domestic counterparts. This analysis is completed using a variety of indicators to measure their participation in certain sectors. Finally, we show the evolution of the immigrant population over the last decade, where the presence of immigrants in the Basque population has increased considerably, to know how has been the absorption of this new population in the labor market.

Palabras Clave: Inmigración, mercado de trabajo.

Área Temática: Economía Social y Discapacidad.

1. INTRODUCCIÓN

La última década ha sido partícipe de una serie de cambios fundamentales en lo que al tamaño y composición de la población se refiere, debido a la llegada masiva de inmigrantes extranjeros de diversa procedencia. España ha pasado de ser uno de los países comunitarios con menor porcentaje de población extranjera (1,6% en 1998) a ser uno de los países que ocupan las posiciones de cabeza (con un 11,33% en 2008).[1] La Comunidad Autónoma Vasca (CAV) también ha notado la llegada de estos nuevos efectivos. A finales del siglo pasado ni siquiera suponían un 1% (0,72% en 1998), pero ya en 2008 son un 5,44% de la población total vasca.

[1] En el artículo se utilizan indistintamente el término extranjero e inmigrante para referirse a aquellas personas que residen en territorio español pero no tienen nacionalidad española.

Esta entrada de población procedente de otros países no sólo ha alterado el tamaño de la población total, en progresivo crecimiento, y su composición, por cuanto los nuevos efectivos son relativamente jóvenes, sino también la distribución de la población en el territorio. El mercado de trabajo tampoco está siendo ajeno a estos cambios, pues cada vez existe una mayor segregación laboral no sólo por sexo sino también por nacionalidad.

Por estas razones, las cuestiones relacionadas con la evolución demográfica en general y los flujos migratorios, en particular, son motivo de preocupación e interés en la actualidad, y prueba de ello es la abundancia de estudios que abordan estos temas. Recolons (2005), Torres (2007) e Ikuspegi (2008) son unos ejemplos claros. El papel de la inmigración en el mercado laboral español es uno de los centros de atención de estos estudios (ver, por ejemplo, Pajares (2008)).

La segregación laboral es un concepto que puede ser estudiado siempre que se analicen dos colectivos de personas y se pretenda comparar si se encuentran en igualdad de oportunidades al incorporarse al mercado de trabajo. Hoy en día, este fenómeno se debe a distintos factores. Primero, aunque la presencia de la mujer se ha extendido a todos los sectores económicos y a todas las ocupaciones, diversos estudios empíricos revelan la existencia de una segregación laboral por sexo, de forma que los hombres y las mujeres están ocupando distintos puestos (ver, por ejemplo, De la Rica (2007) y Palacio y Simón (2006)). Segundo, trabajos que analizan la inserción laboral de los inmigrantes en las sociedades desarrolladas también delatan que existe una segregación laboral por nacionalidad (Calderón e Hidalgo (2007) es un ejemplo). Analizando la localización sectorial de los trabajadores extranjeros se observa que se ubican en unas pocas ramas de actividad: construcción, actividades de los hogares, hostelería, comercio y agricultura Cabe suponer, por tanto, que una parte de la población ocupada inmigrante sufre una doble segregación: por sexo y por nacionalidad. Este trabajo pretende reflejar las desigualdades existentes entre los distintos colectivos que participan en el mercado laboral.

El trabajo se estructura de la siguiente forma. La Sección 2 realiza una breve descripción de las fuentes de datos y de los indicadores que se emplean en el análisis. La Sección 3 describe las principales características de la población inmigrante, en cuanto a edad, sexo y origen, y analiza su distribución a largo de la CAV con el objetivo de verificar si existe una concentración de extranjeros en determinados municipios. La Sección 4 realiza un análisis sobre la participación de los trabajadores inmigrantes en el mercado laboral vasco, identifica los espacios laborales prioritarios en los que se concentran y cuantifica su segregación sectorial y ocupacional para el año 2008. Para finalizar, en la Sección 5 se presentan las principales conclusiones.

2. FUENTE DE DATOS Y METODOLOGÍA

El análisis utiliza dos fuentes básicas de datos, que se caracterizan por su carácter censal (no muestral). La primera es la Revisión Continua del Padrón Municipal, elaborada por el Instituto Nacional de Estadística, que proporciona datos oficiales de la población a uno de enero de cada año. Estos datos permiten caracterizar a las personas extranjeras residentes en el País Vasco. Ésta no es una fuente que aporte datos con total grado de precisión, al menos, por dos razones. Por un lado, presenta el inconveniente de que la cifra de población tiende a estar sobrevalorada. La razón es sencilla: mientras que las altas en el Padrón son prácticamente inmediatas tras solicitarlas, las bajas no lo son tanto. Así, existen extranjeros que deciden emigrar y que no gestionan su baja en el padrón y otros que, cambiando de municipio de residencia, mantienen una duplicidad en sus inscripciones municipales. Por otro lado, también existen otros numerosos inmigrantes que viven establemente en España y que no están empadronados. Estos datos publicados por INE obviamente no están exentos de problemas, sin embargo, serán los utilizados en el trabajo, al tener la consideración de oficiales.[2]

[2] El empadronamiento no exige que el inmigrante esté regularizado, por lo que el Padrón municipal incluye tanto a los que se encuentran en situación legal como ilegal. Para obtener datos de los Inmigrantes Extranjeros Regularizados se puede acudir a las cifras de extranjeros con permiso de residencia que ofrece la Dirección General de Policía (Ministerio de Interior).

La segunda fuente es la Seguridad Social, que ofrece datos de afiliados así como de altas y bajas producidas en determinados periodos, información que permite conocer a la población extranjera ocupada en situación regular. Presenta el inconveniente de que al tratarse de un registro de altas, algunos individuos pueden estar contabilizados más de una vez, mientras que otros pueden no estarlo al no estar dados de alta. De hecho, no refleja a los inmigrantes en situación legal irregular, ignorando así a una gran bolsa de trabajadores que ofrece su fuerza laboral en la economía sumergida.

Aunque obviamente ambas fuentes adolecen de problemas, sus datos tienen un carácter oficial y servirán para caracterizar a la población inmigrante y su evolución a lo largo del periodo 1998-2008. Estos años representan más de una década con un crecimiento espectacular de este colectivo e involucran procesos de movilidad, tanto geográfica como ocupacional, por lo que es un periodo suficientemente representativo para establecer diferencias apreciables.

Para el análisis del mercado de trabajo, se utilizan datos de afiliados dados de alta el último día de cada mes, considerando como referencia diciembre de 2008.[3]

En el análisis se hace uso de diversos indicadores. Se recurre principalmente a *indicadores de concentración* para medir el reparto de los inmigrantes de la CAV entre las tres provincias vascas y, dentro de estas, entre los distintos municipios. El coeficiente de Gini servirá para cuantificar este mayor o menor grado de concentración. También se emplean otros indicadores como medida del porcentaje de población residente en la CAV que tiene la condición de inmigrante.

Asimismo, con el objeto de estudiar la segregación laboral, se utilizan tres herramientas complementarias (al igual que en Castaño, Sánchez-Herrero e Iglesias (2002) e Iglesias y Llorente (2008)):

- *Índice de distribución*, como medida de la proporción de inmigrantes en un determinado sector de actividad (u ocupación).
- *Índice de concentración*, usado para conocer dónde trabajan con más frecuencia los inmigrantes, y calculado como el cociente entre el número de inmigrantes en un sector de actividad (u ocupación) y el número total de inmigrantes empleados.
- *Índice de segregación*, que indica en qué medida los inmigrantes presentan pautas de concentración distintas a las de los no inmigrantes. Aunque existen varios indicadores de segregaciÿÿ, en este artículo se opÿÿ por utilizaÿÿel Índice de Disimilitud (IDD) desarrollado por Duncan y Duncan (ver Duncan y Duncan (1955)), construido inicialmente para medir la segregación racial en Estados Unidos. Este indicador se caracteriza por su relativa sencillez de implementación y por ser uno de los más utilizados tradicionalmente para medir la segregación laboral. Adopta valores entre 0 y 100, representando el 0 una distribución igualitaria y el 100 la segregación total y se calcula como:[4]

$$IDD = \left(\frac{1}{2} \sum_{i=1}^{n} \left| \frac{I_i}{I} - \frac{N_i}{N} \right| \right) \times 100$$

donde, I_i y N_i representan el número de inmigrantes y de nacionales en la ocupación i (o régimen de cotización i); I y N constituyen el número total de inmigrantes y de nacionales empleados, respectivamente, y n denota el número de ocupaciones (o regímenes de cotización).[5]

3. LA POBLACIÓN INMIGRANTE EN EL PAÍS VASCO

[3] Desde Enero de 2009 la Seguridad Social ofrece datos sobre los afiliados dados de alta en las distintas actividades clasificadas por el código CNAE tanto para los afiliados al régimen general como para los autónomos, si bien esta información no está disponible hasta esta fecha. En la Sección 4.2., al analizar a los inmigrantes por actividad ocupacional se tomará como referencia este mes.

[4] El inconveniente que presenta y por lo que ha recibido múltiples críticas, es que no es adecuado cuando se comparan distintos períodos de tiempo. Este problema, sin embargo, no afecta a este atrabajo, al calcular este indicador de segregación en un momento dado.

[5] Alternativamente, si se analiza la segregación laboral por sexo, los colectivos recogidos por el índice serán mujeres y hombres, en lugar de inmigrantes y nacionales.

3.1. Distribución de la población inmigrante

La población de la CAV ha aumentado de manera continuada desde comienzos de la nueva década hasta la actualidad, resultando un crecimiento del 2,79% durante el periodo 1998-2008. Este incremento, sin embargo, no ha sido igualitario en los tres territorios vascos: Álava ha incrementado su población un 8,80%, mientras que Vizcaya lo ha hecho sólo en un 0,78%. Esto supone un mayor peso de la población alavesa en el conjunto de la CAV (14,35% en 2008 frente al 13,56% en 1998), contribuyendo así a una distribución más equilibrada de la población en el territorio vasco. Al tratar de determinar la naturaleza de las variaciones demográficas debe considerarse, en primera instancia, las causas naturales como nacimientos y defunciones, y como segunda vía las corrientes migratorias. En la CAV, el saldo vegetativo ha sido negativo hasta 2003, y tras este año, aunque positivo, no ha alcanzado cifras elevadas, por lo que los aportes demográficos han de buscarse en las entradas de población procedente de otras zonas.

Tabla 1. Población total, Saldo Vegetativo y Saldo Migratorio en el País Vasco

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Población total	2.098.628	2.100.441	2.098.596	2.101.478	2.108.281	2.112.204	2.115.279	2.124.846	2.133.684	2.141.860	2.157.112
Saldo vegetativo	-2.038	-1.609	-927	-438	-493	-36	895	281	1.519	1.706	-
Saldo migratorio											
Extranjeros	2.245	3.453	6.362	8.417	9.890	9.700	11.664	14.742	18.134	-	-
Otras CC.AA	-4.457	-4.642	-3.801	-1.338	-3.653	-3.449	-2.640	-1.478	-3.676	-	-
Total	-2.213	-1.188	2.560	7.079	6.237	6.251	9.024	13.264	14.458	17.692	-

Fuente: Movimiento Natural de Población y Saldos Migratorios proceden de Eustat. Datos de Población proceden de INE

Tabla 2. Número de extranjeros en el País Vasco

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
PAIS VASCO	15.198	16.793	21.140	27.438	38.408	49.231	59.166	72.894	85.542	98.524	117.337
Álava	2.460	2.801	3.818	5.462	8.031	10.445	12.058	15.141	16.857	19.392	22.840
Guipúzcoa	5.301	6.359	7.903	8.856	11.716	14.878	18.232	21.536	25.290	29.040	35.935
Vizcaya	7.437	7.633	9.419	13.120	18.661	23.908	28.876	36.217	43.395	50.092	58.562

Fuente: INE

Tabla 3. Porcentaje de extranjeros en el País Vasco dentro de cada demarcación territorial

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
PAIS VASCO	0,72	0,80	1,01	1,31	1,82	2,33	2,80	3,43	4,01	4,60	5,44%
Álava	0,86	0,98	1,33	1,89	2,75	3,55	4,07	5,05	5,58	6,35	7,38%
Guipúzcoa	0,78	0,94	1,16	1,30	1,72	2,17	2,66	3,13	3,66	4,18	5,12%
Vizcaya	0,65	0,67	0,83	1,16	1,65	2,11	2,55	3,19	3,81	4,39	5,11%

Fuente: INE

El siglo XXI ha comenzado con una ruptura de la trayectoria migratoria en el País Vasco, dejando de ser una comunidad exportadora de población, para convertirse en comunidad receptora. La tendencia desfavorable que parecía haberse consolidado en décadas anteriores, cambia de dirección en el año 2000 presentando flujos positivos y crecientes: las entradas netas de población se cifraban en 2.560 en 2000 y en 2007 son de 17.692 (se multiplica casi por 7). Este cambio de comportamiento se justifica por la entrada neta de población procedente de otros países, que compensa con creces la salida de vascos hacia otras comunidades del estado español.

Centrándonos en la población extranjera, se observa que el número de personas residentes en la CAV ha aumentado de forma considerable, pasando de 15.198 habitantes empadronados en 1998 a 117.337 en 2008, una cifra casi ocho veces mayor, lo que supone un salto del 0,72% al 5,44% de la población total en 2008.

Esta evolución positiva queda reflejada en las tres provincias vascas, aunque con mayor intensidad en Álava. Centrándonos en el porcentaje de extranjeros dentro de cada demarcación territorial (ver Tabla 3) se tiene que, partiendo de porcentajes similares en 1998, este territorio alcanza un 7,38% de población inmigrante, frente al 5,1% de Guipúzcoa y Vizcaya. Quizá esta diferencia obedezca a la mayor necesidad de mano de obra no cualificada exigida en algunas zonas alavesas para realizar tareas del campo y otras ligadas a la actividad vitivinícola, que puede ser cubierta por los nuevos efectivos de esta economía.

El peso de la población inmigrante en la capital alavesa asciende al 8,22% y en algunos municipios alaveses supera el 10% de la total (por ejemplo, en Baños de Ebro (14,85%), Samaniego (11,85%) o Leza (11,06%))[6]. La presencia de extranjeros en las otras capitales vascas, aunque menor, también es considerable, pues asciende al 7,06% de la población vizcaína y al 5,82% de la guipuzcoana.

Su distribución geográfica es similar a la de la población autóctona: el análisis del reparto de los inmigrantes de la CAV entre las tres provincias vascas, muestra que la mayor concentración de los inmigrantes se da en Vizcaya (49,91% en 2008) y la menor en Álava (19,47%).

La distribución de la población extranjera en el territorio además parece seguir el mismo patrón que ya existía antes de su llegada: una alta concentración en ámbitos territoriales reducidos, con gran dinamismo económico, normalmente en ciertas áreas urbanas, especialmente de la zona costera, frente a otras grandes áreas interiores despobladas y semi-abandonadas. Su destino principal son las capitales de provincia y algunos municipios limítrofes, que pueden absorber un mayor número de empleos, sobre todo de baja cualificación, característica propia del perfil del inmigrante procedente de otros países. Se ocupan principalmente en el servicio doméstico, el comercio, la hostelería y la construcción.

Con el propósito de analizar el mayor o menor grado de concentración de la población extranjera en los municipios vascos, se calcula el coeficiente de Gini. Este coeficiente es una medida de desigualdad que adopta valores entre 0 y 1, donde 0 se corresponde con la igualdad perfecta (reparto igualitario de la población en el territorio) y 1 con la desigualdad perfecta.

Tabla 4. Índices de Gini. Población Nacional e Inmigrante. 1998 y 2008

	1998		2008	
	Nacional	Inmigrante	Nacional	Inmigrante
Vizcaya	0,8124	0,8744	0,7984	0,8464
Álava	**0,8878**	**0,9365**	**0,8767**	**0,9169**
Guipúzcoa	0,7438	0,8136	0,7352	0,7751
CAV	**0,8166**	**0,8754**	**0,8040**	**0,8488**

Fuente: INE

La proximidad del indicador a la unidad es reflejo claro de la ausencia de uniformidad en el reparto de la población, y especialmente de la inmigrante, a lo largo de la CAV. De hecho, sólo en 2008, las tres capitales vascas albergan al 46,68% de los inmigrantes y por tanto, el 98,8% de los restantes municipios vascos alojan al 53,32% de los extranjeros.

Álava es, entre las provincias de la CAV, la que presenta desequilibrios más acentuados en la distribución territorial de la población: el índice de Gini como expresión del grado de

[6] Evidentemente, también hay municipios vizcaínos (Berriatua (13,88%), Markina-Xemein (11,45%) y Areatza (10,45%)) y guipuzcoanos (Ordizia (12,45%)) con gran proporción de población inmigrante.

desigualdad toma los valores más altos, tanto en la población nacional como en la inmigrante, aunque son mayores en este último colectivo. De hecho, en 4 de los 51 municipios alaveses, los más grandes, se concentra casi el 90% de la población extranjera (y de la nacional). Sólo en la capital reside el 83,64% de los inmigrantes.

Los desequilibrios en Vizcaya y en Guipúzcoa no están tan acentuados, pero distan mucho de la uniformidad: los coeficientes de Gini en 2008 son 0,79 y 0,73 respectivamente, para los nacionales y de 0,84 y 0,77, respectivamente para los inmigrantes. En la capital vizcaína se concentra el 30,19% de la población autóctona y el 42,59% de la extranjera. En la guipuzcoana, estos porcentajes son del 26,09% y del 29,84%, respectivamente.

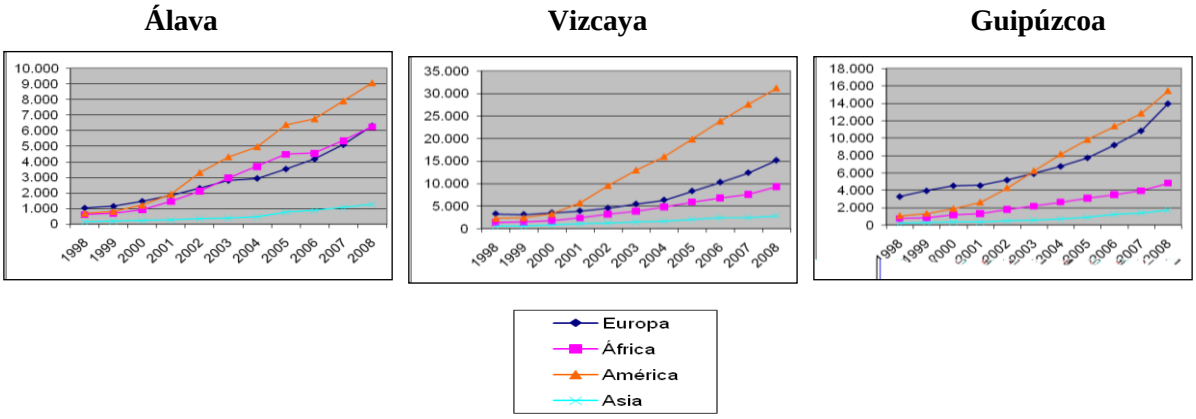
En este contexto, el fenómeno de los flujos migratorios que ha aparecido con fuerza creciente en los últimos años parece que no va a constituir una vía para corregir los desequilibrios territoriales existentes en el conjunto del territorio vasco, sino que más bien puede acentuarlos.

A pesar de que los desequilibrios son importantes, parece que se están atenuando con el tiempo. Prueba de ello son los menores coeficientes obtenidos en 2008 respecto a 1998. La razón puede encontrarse en que, aun persistiendo un claro predominio numérico de los municipios pequeños, está creciendo sensiblemente el número de municipios mayores de 1.000 y 2.000 habitantes, lo cual parece tender a la conformación de nuevas cabeceras comarcales intermedias, que resultan más atractivas para la población.

3.2. La población inmigrante por zona de origen y sexo

La justificación de esta entrada masiva de población inmigrante procedente de otros países se encuentra en las motivaciones económicas. La favorable situación a la que se ha enfrentado la economía vasca desde 1994, con un mercado laboral en expansión, y que incluso en los últimos años ha llegado a demandar nuevos efectivos, ha sido el principal factor de atracción de muchos residentes de otros países con una situación económica relativamente peor. Por continentes, tras América del Sur (41,42%), la Europa comunitaria es la principal abastecedora de efectivos (27,26%) en 2008. A ésta le sigue África (17,30), Asia (4,97%) y Oceanía (0,08%, un valor casi despreciable). Este orden se mantiene en las tres provincias vascas (ver Figura 1).

Figura 1. Evolución de la población inmigrante según zona de origen.



Por países, los rumanos (10,95%), colombianos (10,36%) y bolivianos (10,14%) han sido principalmente los que han respondido a esta llamada en 2008, aunque también ha sido notable la presencia de marroquís (8,75%), portugueses (7,49%) y ecuatorianos (6,44%).

Un análisis de los extranjeros por sexo, a nivel agregado, no revela grandes diferencias ya que, en términos generales, existe un cierto equilibrio entre hombres (51,44% en 2008) y mujeres. Las divergencias aparecen si el análisis se hace, de forma detallada, por zona de origen y lugar de destino.

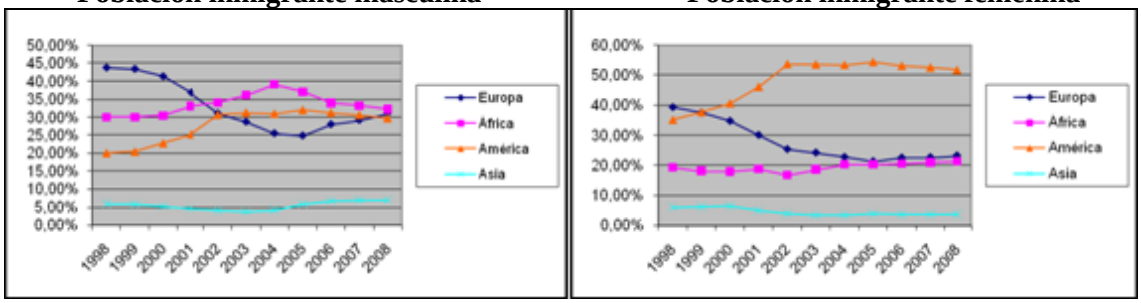
Tabla 5. Evolución del Porcentaje de inmigrantes en el País Vasco según zona de origen (1998-2008)

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
UE(27)	**47,16%**	**46,38%**	**42,37%**	**35,17%**	**28,65%**	**25,95%**	**24,02%**	**23,73%**	**24,77%**	**25,90%**	**27,26%**
Rumania	0,66%	0,70%	1,00%	1,44%	1,82%	3,14%	4,48%	5,83%	7,32%	8,85%	10,95%
Portugal	21,33%	21,01%	18,48%	14,69%	11,24%	9,11%	7,79%	7,28%	7,32%	7,55%	7,49%
EUROPA NO COMUNITARIA	**2,76%**	**2,64%**	**2,64%**	**2,52%**	**2,67%**	**2,90%**	**3,13%**	**3,19%**	**3,00%**	**2,87%**	**2,92%**
Ucrania	0,18%	0,21%	0,25%	0,35%	0,66%	0,97%	1,13%	1,17%	1,15%	1,11%	1,11%
ÁFRICA	**17,86%**	**17,56%**	**18,43%**	**18,82%**	**18,52%**	**18,27%**	**18,85%**	**18,44%**	**17,49%**	**17,07%**	**17,30%**
Marruecos	10,63%	10,27%	10,31%	9,91%	9,74%	9,10%	9,30%	9,27%	8,86%	8,76%	8,75%
Argelia	1,07%	1,26%	1,60%	2,27%	2,63%	3,01%	3,27%	2,98%	2,71%	2,57%	2,63%
Senegal	0,78%	0,77%	0,77%	0,94%	1,04%	1,19%	1,37%	1,31%	1,27%	1,34%	1,45%
AMÉRICA CENTRAL Y CARIBE	**6,57%**	**6,55%**	**7,11%**	**6,38%**	**5,33%**	**4,73%**	**4,36%**	**4,16%**	**3,99%**	**4,18%**	**4,71%**
Cuba	3,48%	3,58%	4,03%	3,59%	2,95%	2,64%	2,31%	2,05%	1,83%	1,60%	1,48%
República Dominicana	2,01%	2,04%	2,21%	1,98%	1,67%	1,46%	1,38%	1,41%	1,28%	1,19%	1,25%
AMÉRICA DEL NORTE	**5,20%**	**5,08%**	**4,35%**	**3,92%**	**3,20%**	**2,81%**	**2,41%**	**2,19%**	**1,89%**	**1,43%**	**1,33%**
AMÉRICA DEL SUR	**14,50%**	**15,60%**	**18,46%**	**27,03%**	**36,11%**	**40,29%**	**42,43%**	**43,10%**	**43,46%**	**43,44%**	**41,42%**
Colombia	1,84%	3,04%	5,00%	10,37%	16,84%	17,21%	15,29%	14,00%	12,96%	11,80%	10,36%
Bolivia	0,27%	0,20%	0,18%	0,24%	0,37%	0,87%	2,54%	4,78%	7,03%	9,71%	10,14%
Ecuador	0,51%	0,62%	1,18%	4,44%	7,34%	9,88%	11,90%	10,45%	9,05%	7,55%	6,44%
Brasil	2,89%	3,55%	4,61%	5,12%	4,83%	4,29%	3,83%	4,18%	4,43%	4,38%	4,37%
ASIA	**5,51%**	**5,73%**	**6,24%**	**5,83%**	**5,24%**	**4,84%**	**4,62%**	**5,02%**	**5,25%**	**5,00%**	**4,97%**
China	2,78%	2,94%	3,50%	3,28%	3,00%	2,72%	2,63%	2,91%	2,96%	2,85%	2,80%
Pakistán	0,18%	0,17%	0,19%	0,23%	0,33%	0,50%	0,62%	0,91%	1,17%	1,20%	1,25%
OCEANÍA	**0,39%**	**0,40%**	**0,38%**	**0,31%**	**0,26%**	**0,21%**	**0,18%**	**0,16%**	**0,13%**	**0,09%**	**0,08%**
APÁTRIDAS	**0,02%**	**0,02%**	**0,01%**	**0,01%**	**0,01%**	**0,01%**	**0,01%**	**0,01%**	**0,01%**	**0,02%**	**0,01%**

Fuente: INE

Figura 2. Evolución de la composición de la población inmigrante según zona de origen. Álava

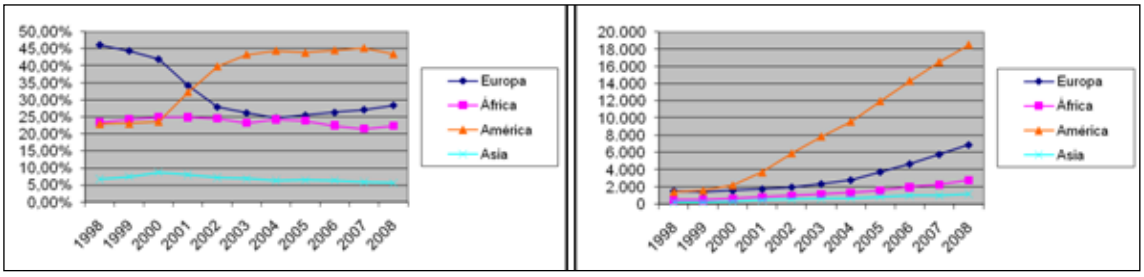
Población inmigrante masculina **Población inmigrante femenina**



De la población extranjera residente en Álava en 2008, el 55,04% son varones y el 44,96% restante mujeres. Centrándonos en el origen de la población masculina, se observa que si bien en 1998 es principalmente europeo (casi el 45%) y africano (alrededor del 30%), once años más tarde, Europa, África y América mantienen porcentajes muy similares (alrededor del 30%). Los cambios más señalados en la procedencia se aprecian en la población femenina. América ha pasado a ser el principal abastecedor de extranjeras de Álava, pues más de la mitad de las residentes en esta provincia tienen este origen. Ha conseguido desbancar, con creces, al continente europeo, que partiendo de un 40% en 1998, sólo conserva poco más de un 20% en 2008.

Figura 3. Evolución de la composición de la población inmigrante según zona de origen. Vizcaya

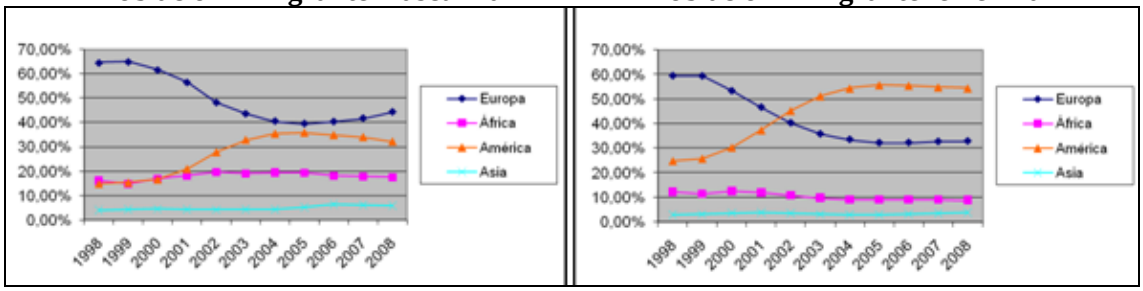
Población inmigrante masculina **Población inmigrante femenina**



En la provincia de Vizcaya, el 49% son inmigrantes varones, por lo que no existen grandes diferencias por sexo. Centrándonos en los hombres extranjeros, en 1998 casi la mitad procede del continente europeo y menos del 25% es americano. En 2008, los extranjeros proceden mayoritariamente de este último continente, al igual que el 63% de las mujeres inmigrantes residentes en Vizcaya.

Figura 4. Evolución de la composición de la población inmigrante según zona de origen. Guipúzcoa

Población inmigrante masculina **Población inmigrante femenina**

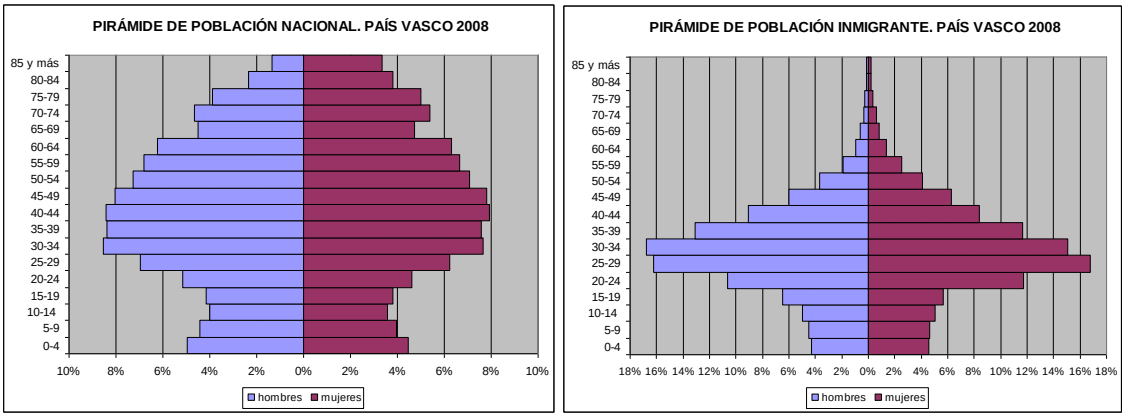


Guipúzcoa es la única provincia en la que el colectivo masculino tiene su origen mayoritariamente en Europa. Aunque se observa un crecimiento considerable de la población extranjera masculina de origen americano, los europeos siguen contando con los mayores porcentajes. Sin embargo, como en las otras provincias vascas, es el continente americano, el principal abastecedor de mujeres inmigrantes en 2008.

3.3. La población inmigrante por edad y sexo

La pirámide demográfica refleja claramente la distribución por edad y sexo de la población, siendo notables las diferencias entre la autóctona y la inmigrante.

Figura 5. Pirámides de población



Centrándonos en los habitantes de nacionalidad española, se aprecia claramente que se trata de una población envejecida, con una elevada esperanza de vida y un reducido peso de su población infantil. Prácticamente un 20% supera la edad legal de jubilación (en su mayoría mujeres, un 58,32%) y sólo alrededor de un 13% tiene menos de 14 años.

La pirámide de población extranjera no mantiene grandes parecidos con la nacional. Los efectivos de mayor edad suponen un porcentaje muy reducido (1,82% de mayores de 65 años) y la gran masa poblacional (un 44,81%) se encuentra entre los 25 y los 40 años, de los que un 52,85% son varones, lo que responde a la edad de entrada del inmigrante en España.

Esta peculiar estructura por edades, sin apenas individuos de edad avanzada, puede actuar de freno al progresivo envejecimiento al que estaba siendo sometida la población vasca, constituyendo así, una vía de rejuvenecimiento de la población total.

Es evidente, por tanto, que uno de los efectos más relevantes de la entrada de estos inmigrantes se está produciendo en la estructura por edades de la población.

4. PARTICIPACIÓN DE LA POBLACIÓN INMIGRANTE EN EL MERCADO DE TRABAJO

La incorporación de la mano de obra inmigrante al mercado de trabajo se produce en determinados sectores de actividad y ocupaciones, catalogados como espacios laborales prioritarios de los trabajadores extranjeros. Esta segmentación del mercado de trabajo es causa de la segregación laboral que sufren determinados colectivos. Con el objetivo de identificar dichos espacios laborales y cuantificar el nivel de desigualdad, se utilizan diversos indicadores.

Primero, el *indicador de distribución* cuantifica la presencia de los trabajadores inmigrantes en el mercado de trabajo. Segundo, el *indicador de concentración* muestra dónde trabajan con más frecuencia y finalmente, el *indicador de segregación* indica en qué medida los inmigrantes presentan pautas de concentración distintas a las de los nativos.

El estudio se divide en dos partes: se trata por un lado, de comparar las pautas de comportamiento de los distintos colectivos de trabajadores en el régimen de cotización elegido a

la hora de afiliarse a la Seguridad Social y por el otro lado, se trata de reflejar la desigualdad que se produce en las actividades que ocupan unos y otros en el mercado de trabajo.

4.1. Afiliados extranjeros por régimen de cotización

La mano de obra extranjera contratada en la CAV asciende a 49.538 personas y supone el 5,18% de la mano de obra total, según datos recogidos por la Seguridad Social para diciembre de 2008.[7] El peso de los trabajadores inmigrantes varía en las tres provincias vascas. Vizcaya, con 21.139 extranjeros dados de alta en la Seguridad Social, es el territorio que absorbe el mayor número de trabajadores inmigrantes en términos absolutos, aunque en términos relativos sólo supone el 4,38% del empleo total. En Guipúzcoa, los 16.983 afiliados extranjeros representan el 5,36% de la afiliación total. Finalmente, Álava empleando 11.416 trabajadores extranjeros, es la que cuenta con mayor presencia de inmigrantes en su mano de obra contratada, un 7,32%.

Analizando el conjunto de trabajadores afiliados a la Seguridad Social por régimen de cotización, se observa que tanto en la CAV como en sus provincias, es el Régimen General (RG) el que cuenta con mayor número de registros, con un 78,33%. A éste le siguen el Régimen Especial de Trabajadores Autónomos (RETA) con un 19,55%, y otros tres regímenes con una participación minoritaria: el Régimen Especial de Empleados del Hogar (REEH) con un 1,45%, el Régimen Especial del Mar (REM) con un 0,38% y el Régimen Especial Agrario (REA) con un 0,29%.

Los trabajadores inmigrantes siguen pautas similares a la de sus homólogos nacionales, aunque las participaciones en cada régimen difieren cuantitativamente: el 68,52% de los afiliados extranjeros cotiza en el general, el 13,8% son autónomos, el 13,19% está dado de alta en el régimen especial de empleadas del hogar, el 4,06% en el agrario y finalmente, el 0,43% en el del mar. Resulta claro que los extranjeros tienen una menor presencia en los regímenes mayoritarios (RG y RETA), y una mayor presencia en los regímenes especiales agrario y de empleadas del hogar.

Indicador de Distribución

Con el objetivo de identificar los espacios laborales prioritarios en los que se ubica cada colectivo de trabajadores, se calculan los indicadores de distribución, que muestran el peso de cada colectivo en un determinado régimen de cotización.

Se analiza si el régimen de cotización refleja rasgos de segregación laboral por sexo y por nacionalidad o si, por el contrario, todos los colectivos se afilian a los distintos regímenes de forma igualitaria. Para ello, se clasifican los regímenes de cotización en función a ambos factores: sexo y nacionalidad, y se sigue a Hakim (1993) para caracterizarlos. Según este autor, se definen los sectores de actividad como femeninos o masculinos en función de que presenten un porcentaje de mujeres superior o inferior al resultado de aumentar o reducir un 10 por ciento la presencia de la mujer en el empleo total. Si el porcentaje de mujeres se encuentra dentro del intervalo definido por ambas condiciones, el sector se considera integrado por género. En nuestro análisis se utiliza esta misma clasificación también para diferenciar los sectores para inmigrantes de los sectores para nativos.

–*clasificación de los regímenes de cotización por sexo*: con una participación de la mujer en el mercado de trabajo vasco del 43,15%, se obtiene que el Régimen General y el Trabajadores Autónomos son regímenes integrados por sexo, mientras que el Agrario y el Mar son regímenes masculinos y el de Empleadas del Hogar femenino.

[7] Entendemos, en lo sucesivo, que la mano de obra contratada coincide con la afiliada a la Seguridad Social.

Tabla 6. Indicador de distribución de la población ocupada por sexo. Diciembre 2008

	RG		RETA		REA		REM		REEH	
	H	M	H	M	H	M	H	M	H	M
Álava	57,51%	42,49%	66,58%	33,42%	84,66%	15,34%	0,00%	0,00%	4,48%	95,52%
Vizcaya	55,02%	44,98%	66,95%	33,05%	85,97%	14,03%	86,77%	13,23%	5,26%	94,74%
Guipúzcoa	55,42%	44,58%	61,71%	38,29%	90,28%	9,72%	90,99%	9,01%	4,54%	95,46%
CAV	**55,66%**	**44,34%**	**64,39%**	**35,61%**	**86,99%**	**13,01%**	**90,09%**	**9,91%**	**4,80%**	**95,20%**

Nota: RG: *Régimen General*; **RETA:** *Régimen Especial de Trabajadores Autónomos*; **REA:** *Régimen Especial Agrario*; **REM:** *Régimen Especial del Mar*; **REEH:** *Régimen Especial de Empleadas del Hogar*
Fuente: Seguridad Social

– *clasificación de los regímenes de cotización por nacionalidad*: la clasificación de los regímenes en función de la nacionalidad muestra que el Régimen Agrario y de Empleadas del Hogar son regímenes en los que mayoritariamente cotizan los trabajadores extranjeros, por lo que pueden ser considerados como regímenes para inmigrantes, mientras que el Régimen General, el de Trabajadores Autónomos y el del Mar se consideran regímenes integrados por nacionalidad.[8]

Tabla 7. Indicador de distribución de la población ocupada inmigrante. Diciembre 2008

	RG	RETA	REA	REM	REEH
Álava	6,38%	5,07%	79,97%	0,00%	59,47%
Vizcaya	4,95%	3,34%	63,75%	9,21%	45,85%
Guipúzcoa	3,64%	3,56%	69,65%	5,00%	45,16%
CAV	**4,53%**	**3,66%**	**72,93%**	**5,90%**	**47,15%**

Nota: RG: *Régimen General*; **RETA:** *Régimen Especial de Trabajadores Autónomos*; **REA:** *Régimen Especial Agrario*; **REM:** *Régimen Especial del Mar*; **REEH:** *Régimen Especial de Empleadas del Hogar*
Fuente: Seguridad Social

– *clasificación de los regímenes de cotización por sexo y nacionalidad.* La distribución de los trabajadores inmigrantes por sexo en los distintos regímenes de cotización, confirma que los hombres que se dan de alta en los regímenes clasificados como regímenes para inmigrantes (REA y REEH) son, de hecho, mayoritariamente extranjeros. En el caso de las mujeres, más de la mitad de las que cotizan en el Régimen Agrario son inmigrantes. Sin embargo, su presencia no destaca tanto en el Régimen de las Empleadas del Hogar por su marcado carácter femenino, es decir, porque también existe un elevado porcentaje de "nativas" empleadas en este sector.

Tabla 8. Indicador de distribución de la población ocupada inmigrante por sexo. Diciembre 2008

	RG		RETA		REA		REM		REEH	
	H	M	H	M	H	M	H	M	H	M
Álava	7,21%	5,27%	5,63%	3,96%	80,59%	76,50%	0,00%	0,00%	85,33%	58,26%
Vizcaya	5,71%	4,03%	3,85%	2,30%	70,55%	22,08%	10,46%	0,98%	83,88%	43,73%
Guipúzcoa	4,29%	2,84%	4,24%	2,47%	71,82%	49,49%	5,46%	0,39%	76,42%	43,68%
CAV	**5,26%**	**3,62%**	**4,25%**	**2,58%**	**75,26%**	**57,38%**	**6,49%**	**0,56%**	**80,48%**	**45,46%**

Nota: RG: *Régimen General*; **RETA:** *Régimen Especial de Trabajadores Autónomos*; **REA:** *Régimen Especial Agrario*; **REM:** *Régimen Especial del Mar*; **REEH:** *Régimen Especial de Empleadas del Hogar*
Fuente: Seguridad Social

8 Para usar la clasificación de Hakim se tiene en cuenta que en la CAV la mano de obra inmigrante contratada es el 5,18% de la mano de obra total. Estos porcentajes son del 4,38% en Vizcaya, 5,36% en Guipúzcoa y 7,32% en Álava.

En el resto de regímenes no es destacable el porcentaje de trabajadores extranjeros, ya que el mayor valor aparece para los hombres en el Régimen Especial del Mar, con un 6,49%.

Los espacios prioritarios en los que cotizan los colectivos de trabajadores inmigrantes son, por tanto, el Régimen Agrario y el de Empleadas del Hogar. Si, además, se considera su sexo, son los hombres inmigrantes los que se ubican prioritariamente en el REA, representando el 89,77% de los extranjeros afiliados, y las mujeres inmigrantes las que mayoritariamente están en el REEH, con un 91,8%.

Indicador de Concentración

El indicador de concentración, como medida del porcentaje de inmigrantes en cada régimen de cotización en relación al número total de inmigrantes empleados, permite saber en qué régimen cotiza con más frecuencia cada colectivo de trabajadores[9]. Se observa que los hombres inmigrantes cotizan con más frecuencia en el régimen general, seguido del régimen de trabajadores autónomos y en menor medida del agrario. Sin embargo, aunque casi dos tercios de las mujeres inmigrantes afiliadas cotizan en el régimen general, casi un tercio lo hace en empleadas del hogar.

Tabla 9. Indicador de concentración de la población ocupada inmigrante por sexo. Diciembre 2008

	RG		RETA		REA		REM		REEH	
	H	M	H	M	H	M	H	M	H	M
Álava	75,62%	67,92%	12,06%	7,09%	11,42%	3,27%	0,00%	0,00%	0,90%	21,72%
Vizcaya	74,49%	61,10%	19,17%	8,02%	3,34%	0,24%	0,70%	0,01%	2,30%	30,63%
Guipúzcoa	72,83%	55,88%	18,82%	9,80%	5,28%	0,57%	1,12%	0,01%	1,95%	33,74%
CAV	**74,06%**	**60,30%**	**17,31%**	**8,59%**	**6,11%**	**1,03%**	**0,71%**	**0,01%**	**1,81%**	**30,06%**

Nota: RG: *Régimen General*; **RETA:** *Régimen Especial de Trabajadores Autónomos*; **REA:** *Régimen Especial Agrario*; **REM:** *Régimen Especial del Mar*; **REEH:** *Régimen Especial de Empleadas del Hogar*
Fuente: Seguridad Social

Los resultados varían si el colectivo de trabajadores inmigrantes se divide por zona de origen. Para los hombres procedentes de países pertenecientes a la Unión Europea (UE) es destacable su presencia en el régimen de trabajadores autónomos. Ahora bien, los hombres procedentes de países de fuera de la UE cotizan mayoritariamente en el régimen general. En el caso de las mujeres inmigrantes, las procedentes de la UE cotizan con más frecuencia en los dos regímenes mayoritarios (RG y RETA) y muestran unos porcentajes muy bajos en el resto de regímenes. Sin embargo, para las procedentes de países de fuera de la UE la diferencia entre el número de afiliadas al régimen general y al de empleadas del hogar es menor que en los casos anteriores.

Tabla 10. Indicador de concentración de la población ocupada inmigrante por sexo y por zona de origen. Población inmigrante UE. Diciembre 2008

	RG		RETA		REA		REM		REEH	
	H	M	H	M	H	M	H	M	H	M
Álava	63,54%	67,97%	14,80%	10,20%	21,52%	15,18%	0,00%	0,00%	0,14%	6,64%
Vizcaya	68,22%	76,32%	29,60%	16,22%	1,11%	0,36%	0,94%	0,00%	0,13%	7,10%
Guipúzcoa	62,32%	70,30%	31,98%	15,40%	4,46%	0,81%	1,14%	0,00%	0,11%	13,49%
CAV	**64,83%**	**72,22%**	**27,09%**	**14,70%**	**7,16%**	**3,47%**	**0,80%**	**0,00%**	**0,12%**	**9,61%**

Nota: RG: *Régimen General*; **RETA:** *Régimen Especial de Trabajadores Autónomos*; **REA:** *Régimen Especial Agrario*; **REM:** *Régimen Especial del Mar*; **REEH:** *Régimen Especial de Empleadas del Hogar*
Fuente: Seguridad Social

[9] Los datos de afiliados extranjeros que ofrece la Seguridad Social distinguen entre afiliados extranjeros de la Unión Europea (UE) y de fuera de la Unión Europea (no UE). Esta misma clasificación se hace en este trabajo, distinguiendo así entre hombres y mujeres nativos y extranjeros, y éstos últimos a su vez desagrupados en trabajadores extranjeros de la Unión Europea y de fuera de la Unión Europea.

Por tanto, entre los hombres extranjeros los procedentes de la UE optan por cotizar en el régimen de trabajadores autónomos y en el régimen agrario en mayor medida que los procedentes de fuera de la UE, que se decantan por el régimen general, y que cotizan en mayor proporción en el régimen de empleadas del hogar. En el caso de las mujeres inmigrantes es destacable la diferencia que muestran las procedentes de la UE y de fuera de la UE al elegir el régimen de empleadas del hogar, con porcentajes relativamente más altos para estas últimas (un 35,62% frente a un 9,61%). Aunque en menor medida, otra diferencia se produce en el régimen agrario, con una presencia prácticamente nula de mujeres de fuera de la UE.

Tabla 11. Indicador de concentración de la población ocupada inmigrante por sexo y por zona de origen. Población inmigrante no UE. Diciembre 2008

	RG		RETA		REA		REM		REEH	
	H	M	H	M	H	M	H	M	H	M
Álava	83,55%	67,91%	10,27%	6,33%	4,79%	0,35%	0,00%	0,00%	1,39%	25,41%
Vizcaya	79,83%	56,27%	10,28%	5,42%	5,24%	0,21%	0,50%	0,02%	4,15%	38,08%
Guipúzcoa	79,29%	52,26%	10,73%	8,40%	5,79%	0,51%	1,11%	0,01%	3,08%	38,81%
CAV	**80,51%**	**57,06%**	**10,48%**	**6,93%**	**5,37%**	**0,37%**	**0,65%**	**0,01%**	**2,99%**	**35,62%**

Nota: RG: *Régimen General*; **RETA:** *Régimen Especial de Trabajadores Autónomos*; **REA:** *Régimen Especial Agrario*; **REM:** *Régimen Especial del Mar*; **REEH:** *Régimen Especial de Empleadas del Hogar*
Fuente: Seguridad Social

La concentración de la población ocupada inmigrante varía para cada territorio. En Álava, aunque la mayoría de los hombres procedentes de países de la UE cotice en el régimen general, son más los que cotizan en el régimen agrario que los autónomos. Los procedentes de fuera de la UE muestran los mayores niveles de concentración en el régimen general (un 83,55%), también por encima de los afiliados nativos. Las mujeres procedentes de la UE siguen las mismas pautas que sus homólogos varones al elegir el régimen de cotización. Sin embargo, las mujeres procedentes de fuera de la UE presentan elevados índices de concentración en el régimen de empleadas del hogar (un 35,62%).

Las pautas de comportamiento de los afiliados extranjeros son similares en Vizcaya y Guipúzcoa. Los hombres y mujeres procedentes de países de la UE presentan mayores niveles en el régimen de autónomos respecto a sus homólogos procedentes de fuera de la UE. En cuanto al colectivo de trabajadores extranjeros procedentes de fuera de la UE presentan un comportamiento similar al de la provincia de Álava.

Indicador de Segregación

Tras calcular los indicadores de distribución y de concentración, por último, se obtiene el indicador de segregación. Este indicador de segregación permite cuantificar la desigualdad en la distribución de los trabajadores extranjeros en los regímenes de cotización o, lo que es lo mismo, el porcentaje de inmigrantes (o de nacionales) que debería cambiar de régimen para igualar la presencia de ambos colectivos en todos los regímenes. Se calcula primero la segregación del mercado de trabajo vasco por nacionalidad y, posteriormente, el análisis se completa con el cálculo de la segregación por nacionalidad y género.

Tabla 12. Índice de Disimilitud por nacionalidad. Diciembre 2008

	IDD
Álava	16,44
Vizcaya	15,25
Guipúzcoa	17,50
CAV	**16,42**

Fuente: Seguridad Social

El índice de Duncan y Duncan, como medida de segregación, muestra una desigual distribución de la población ocupada en los regímenes en los que se dan de alta los afiliados nativos y extranjeros, lo que revela que los extranjeros no tienen las mismas oportunidades que los nativos al elegir el régimen de cotización.

Tabla 13. Índice de Disimilitud por nacionalidad y sexo. Diciembre 2008

	IDD	
	Hombres	Mujeres
Álava	12,07	23,81
Vizcaya	8,71	28,68
Guipúzcoa	7,28	32,43
CAV	**7,90**	**29,22**

Fuente: Seguridad Social

El valor del índice varía al considerar también el sexo del trabajador, observando una mayor similitud entre los colectivos de hombres y una mayor desigualdad entre los colectivos de mujeres.

Si a su vez se desagrupa el colectivo de trabajadores inmigrantes residentes en la CAV por zona de origen y sexo, el índice refleja desigualdades similares entre los trabajadores provenientes de la UE (tanto hombres como mujeres) y los hombres procedentes de fuera de la UE. El colectivo que presenta una mayor desigualdad es el compuesto por mujeres provenientes de países de fuera de la UE. Por tanto, son estas mujeres las que hoy en día sufren los mayores niveles de segregación en el mercado laboral, en relación al régimen en el que cotizan.

Tabla 14. Índice de Disimilitud por zona de origen y sexo. Diciembre 2008

UE

	IDD	
	Hombres	Mujeres
Álava	21,42	20,65
Vizcaya	3,44	5,27
Guipúzcoa	16,09	12,43
CAV	**12,02**	**11,21**

No UE

	IDD	
	Hombres	Mujeres
Álava	6,89	24,59
Vizcaya	17,60	36,09
Guipúzcoa	9,80	37,44
CAV	**11,95**	**34,12**

Fuente: Seguridad Social

Los valores obtenidos con este índice varían para las tres provincias vascas. En Álava el colectivo que presenta una menor desigualdad es el compuesto por hombres de fuera de la UE, mientras que la segregación que sufre el resto es similar. En Vizcaya, los trabajadores extranjeros de la UE muestran un comportamiento similar al de los nativos y entre los trabajadores de fuera de la Unión Europea, son las mujeres las que presentan mayores niveles de desigualdad. Finalmente, en Guipúzcoa, todos los colectivos muestran una distribución desigual, aunque también en este caso son las mujeres de fuera de la UE las que presentan unos mayores valores en este índice.

Se concluye así que los regímenes de cotización prioritarios del colectivo de inmigrantes son el régimen agrario, para los hombres, y el de empleadas del hogar para las mujeres. En cuanto a los hombres extranjeros afiliados al régimen agrario, los procedentes de países de la UE muestran mayores niveles de concentración que los de fuera de la UE. Entre las mujeres afiliadas al régimen de empleadas del hogar las procedentes de países de fuera de la UE presentan un nivel de concentración cuatro veces mayor al de las procedentes de la UE.

Además, la distribución de la población ocupada inmigrante entre los distintos regímenes, en comparación a sus homólogos nacionales, es desigual por lo que no existe una igualdad de oportunidades a la hora de incorporarse al mercado de trabajo para todos los distintos colectivos que conforman la mano de obra extranjera.

4.2. Afiliados extranjeros por actividad ocupacional

Este apartado trata de identificar las actividades prioritarias en las que se ocupan los trabajadores inmigrantes,[10] independientemente del régimen en el que cotizan. Las actividades se clasifican según el código CNAE-2009 y se toma como referencia enero de 2009 (momento

[10] Por falta de datos este apartado agrupa el colectivo de trabajadores por nacionalidad, considerando únicamente el colectivo de trabajadores nativos y el colectivo de trabajadores inmigrantes.

desde el que esta información está disponible). Los indicadores de distribución, de concentración y de segregación son, de nuevo, herramientas útiles para este análisis.

Indicador de Distribución

Primero, se estudia el peso relativo de la mano de obra inmigrante en cada ocupación. Con el indicador de distribución se obtiene información sobre las actividades que emplean mayor número de inmigrantes. Siguiendo el método propuesto por Hakim (1993), se clasifican las distintas actividades ocupacionales y se identifican las ocupaciones que pueden ser consideradas para inmigrantes, obteniendo:

- un elevado peso de la población ocupada inmigrante en las actividades de los hogares, sobre todo en la provincia de Álava, que absorbe a más de la mitad de los ocupados en esta actividad. Se puede considerar, por tanto, que la actividad de los hogares es una ocupación para inmigrantes en las tres provincias vascas.
- agricultura, ganadería, silvicultura y pesca también son ocupaciones para inmigrantes en Álava y Vizcaya, o incluso en la CAV en su conjunto. Aunque el porcentaje de empleo de inmigrantes en estas actividades es alto en Guipúzcoa, en la clasificación se considera como actividad integrada por nacionalidad.
- construcción y hostelería son consideradas como actividades para inmigrantes en Álava. En Vizcaya y en Guipúzcoa, aunque el porcentaje de inmigrantes concentrados en estas ocupaciones sea elevado, no supera en 10 puntos el peso de la mano de obra extranjera ocupada en cada territorio, por lo que se consideran actividades integradas por nacionalidad.

Tabla 15. Indicador de distribución de la población ocupada inmigrante. Enero 2009

CNAE-2009	Álava	Vizcaya	Guipúzcoa	CAV
Agricultura, ganadería, silvicultura y pesca	24,83%	16,62%	13,24%	**18,17%**
Industrias extractivas	6,35%	1,66%	2,43%	**2,90%**
Industria manufacturera	3,41%	2,58%	2,48%	**2,71%**
Suministro de energía eléctrica, gas, vapor y aire acondicionado	0,49%	0,47%	0,48%	**0,47%**
Suministro de agua, actividades de saneamiento, gestión de residuos y descontaminación	13,36%	2,14%	3,01%	**3,74%**
Construcción	19,33%	9,71%	10,89%	**11,46%**
Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas	6,52%	2,98%	4,17%	**3,83%**
Transporte y almacenamiento	6,58%	2,58%	6,63%	**4,45%**
Hostelería	19,12%	11,31%	14,13%	**13,42%**
Información y comunicaciones	4,72%	1,74%	1,81%	**2,06%**
Actividades financieras y de seguros	0,27%	0,47%	0,18%	**0,34%**
Actividades inmobiliarias	4,90%	1,56%	1,96%	**2,04%**
Actividades profesionales, científicas y técnicas	3,12%	2,13%	2,67%	**2,43%**
Actividades administrativas y servicios auxiliares	10,36%	4,30%	6,82%	**5,84%**
Administración Pública y defensa; Seguridad Social obligatoria	1,19%	0,50%	0,56%	**0,69%**
Educación	2,43%	2,68%	2,82%	**2,69%**
Actividades sanitarias y de servicios sociales	4,32%	1,53%	3,09%	**2,54%**
Actividades artísticas, recreativas y de entretenimiento	6,94%	4,18%	5,28%	**4,99%**
Otros servicios	3,31%	2,53%	3,61%	**3,01%**
Actividades de los hogares como empleadores de personal doméstico; actividades de los hogares como productores de bienes y servicios para uso propio	57,71%	39,60%	44,76%	**43,50%**
Actividades de organizaciones y organismos extraterritoriales	0,00%	23,81%	0,00%	**22,73%**

Fuente: Seguridad Social

Aunque según Hakim (1993), el resto de ocupaciones se clasifican como actividades integradas por nacionalidad, existen determinadas actividades en las que la presencia de trabajadores extranjeros es prácticamente nula. Se trata del suministro de energía o de las

actividades financieras y de seguros que, por tanto, podrían ser consideradas como actividades para nativos.

Indicador de Concentración

A continuación, se trata de conocer en qué actividades trabajan con más frecuencia los trabajadores inmigrantes, para lo que se utiliza el indicador de concentración. Este indicador muestra cómo se distribuye la población ocupada entre las distintas actividades en cada territorio. Aunque las actividades en las que se concentra el colectivo inmigrante son las mismas en las tres provincias, varía el orden en el que se clasifican por nivel de concentración:

- En Álava, la construcción es la actividad en la que se concentra el mayor número de inmigrantes, seguida en menor medida de la hostelería y la industria manufacturera. La siguiente actividad con una concentración considerable de trabajadores extranjeros es el comercio al por mayor y al por menor, a la que siguen la agricultura y las actividades de los hogares. Finalmente, destacan las actividades administrativas y servicios auxiliares.
- Vizcaya también concentra a la mayor parte de los trabajadores extranjeros en la construcción, seguida en menor medida de las actividades de los hogares y la hostelería. A continuación destacan el comercio al por mayor y al por menor, la industria manufacturera y las actividades administrativas y servicios auxiliares.
- Finalmente, Guipúzcoa, aunque en menor medida que las provincias anteriores, muestra la mayor concentración de extranjeros en la construcción. A continuación le siguen la hostelería, las actividades de los hogares, la industria manufacturera, el comercio al por mayor y al por menor y, en menor medida, las actividades administrativas y servicios auxiliares.

También existen actividades en las que la concentración de trabajadores inmigrantes es casi nula. Se trata del suministro de energía, actividades financieras y de seguros, actividades inmobiliarias, actividades profesionales, científicas y técnicas, y administración pública, entre otros.

Tabla 16. Indicador de concentración de la población ocupada inmigrante. Enero 2009

CNAE-2009	Álava	Vizcaya	Guipúzcoa	CAV
Agricultura, ganadería, silvicultura y pesca	8,88%	4,97%	2,80%	**5,14%**
Industrias extractivas	0,10%	0,03%	0,04%	**0,05%**
Industria manufacturera	12,72%	10,07%	11,98%	**11,33%**
Suministro de energía eléctrica, gas, vapor y aire acondicionado	0,01%	0,05%	0,01%	**0,03%**
Suministro de agua, actividades de saneamiento, gestión de residuos y descontaminación	0,77%	0,32%	0,29%	**0,41%**
Construcción	20,69%	20,96%	17,24%	**19,64%**
Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas	11,04%	10,93%	10,70%	**10,88%**
Transporte y almacenamiento	3,45%	2,75%	5,42%	**3,81%**
Hostelería	12,94%	14,36%	15,44%	**14,39%**
Información y comunicaciones	0,92%	1,17%	0,60%	**0,92%**
Actividades financieras y de seguros	0,06%	0,26%	0,09%	**0,16%**
Actividades inmobiliarias	0,16%	0,15%	0,13%	**0,15%**
Actividades profesionales, científicas y técnicas	1,73%	2,58%	2,38%	**2,31%**
Actividades administrativas y servicios auxiliares	7,93%	7,15%	6,30%	**7,05%**
Administración Pública y defensa; Seguridad Social obligatoria	1,19%	0,50%	0,45%	**0,64%**
Educación	1,66%	3,44%	3,29%	**2,98%**
Actividades sanitarias y de servicios sociales	4,94%	2,69%	4,88%	**3,95%**
Actividades artísticas, recreativas y de entretenimiento	1,10%	1,13%	1,34%	**1,19%**
Otros servicios	1,18%	1,77%	2,15%	**1,76%**
Actividades de los hogares como empleadores de personal doméstico; actividades de los hogares como productores de bienes y servicios para uso propio	8,53%	14,68%	14,46%	**13,18%**
Actividades de organizaciones y organismos extraterritoriales	0,00%	0,02%	0,00%	**0,01%**

Fuente: Seguridad Social

Indicador de Segregación

Finalmente, el indicador de segregación cuantifica la desigualdad en la distribución de la población ocupada inmigrante en relación a la nacional al considerar las actividades en las que se emplea cada colectivo en mercado de trabajo actual.

El índice de Duncan y Duncan muestra que el mercado laboral vasco está segmentado y existe disimilitud en la distribución de la población ocupada según la nacionalidad del trabajador. Esta disimilitud se interpreta como una medida de la discriminación que sufren determinados colectivos de la población al participar en el mercado de trabajo y representa el porcentaje de trabajadores inmigrantes que debería cambiar de actividad para obtener una distribución igualitaria entre la población ocupada inmigrante y la nacional.

Tabla 17. Índice de Disimilitud por nacionalidad. Enero 2009

	IDD
Álava	38,09
Vizcaya	38,27
Guipúzcoa	37,12
CAV	**37,27**

Fuente: Seguridad Social

Los valores obtenidos muestran mayores niveles de desigualdad en la distribución de los trabajadores extranjeros que los observados al calcular este índice por régimen de cotización. Sin embargo, este resultado no es de extrañar, pues tal y como indican Hidalgo, Calderón y Pérez (2006), a mayor agregación de ocupaciones menor es la segregación que capta este índice.

A diferencia de la sección anterior, en ésta no es posible cuantificar la segregación por sexo ni por zona de origen, debido a que no se dispone de datos con este nivel de detalle. No obstante, se confía en que el comportamiento sea similar al obtenido para los distintos regímenes de cotización. Así, es de esperar que sean las mujeres inmigrantes procedentes de países de fuera de la UE las que obtengan mayores valores en este índice y, por tanto, las que presenten mayores niveles de desigualdad en su incorporación al mercado de trabajo o, lo que es lo mismo, las que hoy en día sufran mayor discriminación laboral.

5. CONSIDERACIONES FINALES

En la última década la población del País Vasco ha experimentado cambios importantes en cuanto a su tamaño, composición y distribución a lo largo del territorio, debido principalmente a la entrada de nuevos efectivos procedentes de otros países. A finales del siglo pasado apenas suponían un 1%, pero ya en 2008 representan un 5,44% de la población de la CAV. El mercado de trabajo no es ajeno a estos cambios y presenta una segregación no sólo por sexo sino también por nacionalidad.

La entrada de estos extranjeros, sin apenas individuos de edad avanzada, puede actuar de freno al progresivo envejecimiento al que estaba siendo sometida la población vasca, constituyendo así, una vía de rejuvenecimiento de su población total. En este sentido, uno de los efectos más relevantes de la entrada de estos inmigrantes se está produciendo en la estructura por edades de la población.

En cuanto a su distribución en el territorio, su destino principal son las capitales de provincia y algunos municipios limítrofes, con gran dinamismo económico y que pueden absorber un mayor número de empleos. Por esta razón, estos flujos migratorios parece que no van a constituir una vía para corregir los desequilibrios territoriales existentes en el conjunto de la CAV, sino que más bien pueden acentuarlos.

Resulta evidente que se trata de una inmigración laboral por lo que la incorporación de estos efectivos repercute sobre las características del mercado de trabajo vasco. Aunque los afiliados extranjeros siguen pautas de cotización similares a las de sus homólogos nacionales, la participación en cada régimen difiere cuantitativamente: los inmigrantes tienen una menor

presencia en los regímenes mayoritarios (general y autónomos) a favor de los regímenes agrario y de empleadas del hogar. Si junto a la nacionalidad, se tiene en cuenta el sexo de los ocupados, se observa que además de ser regímenes para extranjeros, el agrario puede ser considerado régimen masculino y el de empleadas del hogar femenino.

Los inmigrantes muestran asimismo distintos comportamientos en función de su zona de origen. Los hombres procedentes de la UE destacan en el régimen de autónomos mientras que los de fuera de la UE cotizan mayoritariamente en el régimen general. En cuanto a las mujeres, las europeas presentan un comportamiento similar al de sus homólogos varones, mientras que las procedentes de fuera de la UE se concentran altamente en el régimen de empleadas del hogar.

Las actividades ocupacionales en las que se emplean prioritariamente los trabajadores extranjeros son las mismas en las tres provincias vascas: se trata de la construcción, la hostelería, el comercio, las actividades de los hogares y la agricultura.

Finalmente, el indicador de segregación muestra que existe una desigualdad en la distribución de los trabajadores inmigrantes y de los nacionales, siendo el colectivo más desfavorecido el correspondiente a la población ocupada femenina procedente de los países de fuera de la Unión Europea.

Por último, comentar que este análisis puede ser extendido en trabajos futuros a otros indicadores que permitan analizar la evolución en el tiempo de la segregación laboral por sexo y nacionalidad. El índice de Karmel y MacLachlan (1988) y/o el Marginal Maching (Blackburn, Jarman y Siltanen, 1993) pueden constituir ejemplos de este tipo de indicadores.

5.BIBLIOGRAFÍA

BLACKBURN, M., JARMAN, J. y SILTANEN, S. (1993), "The Analysis of Occupational Gender Segregation over Time and Place: Considerations of Measurement and some New Evidence". *Work, Employment and Society*, vol. 3, pág. 335-362.

CALDERÓN, M.J. e HIDALGO, A. (2007), "La Segregación Sectorial y Ocupacional de los Inmigrantes en España: Identificación de Espacios Laborales Prioritarios por Países de Origen". VII Jornadas de Economía Laboral.

CASTAÑO, C., IGLESIAS, C. y SÁNCHEZ-HERRERO, M. (2002), "Tecnología y Empleo en Perspectiva de Género. El Caso Español al Final del Siglo XX". *Economía Industrial*, nº 348, pág. 27-40.

DE LA RICA, S. (2007), "Segregación Ocupacional y Diferencias Salariales por Género en España: 1995-2002". Documento de Trabajo 2007-35. FEDEA.

DUNCAN, O. y DUNCAN B. (1955), "A Methodological Analysis of Segregation Indexes". *American Sociological Review*, nº 20, pág. 210-217.

HAKIM, C. (1993), "Segregated and Integrated Occupations: A New Approach to Analysing Social Change". *European Sociological Review*, vol. 9, nº 3, pág. 289-314.

HIDALGO, A., CALDERÓN, M.J. y PÉREZ, S. (2006), "La Discriminación Salarial y Segregación Laboral de los Inmigrantes: Identificación de Espacios Laborales Prioritarios para la Intervención en Materia de Protección Social". Ministerio de Trabajo y Asuntos Sociales.

IGLESIAS C. y LLORENTE, R. (2008), "Evolución Reciente de la Segregación Laboral por Género en España". Serie Documentos de trabajo. Instituto Universitario de Análisis Económico y Social (IAES). Universidad de Alcalá.

IKUSPEGI (2008), "Población Extranjera en la CAPV 2008", *Panorámica de la Inmigración*, nº 23. Observatorio Vasco de Inmigración.

KARMEL, T. y MACLACHLAN, M. (1988), "Occupational sex segregation. Increasing or decreasing?". *The Economic Record*, nº 64, pág. 187-195.

PAJARES, M. (2008), "Inmigración y Mercado de Trabajo. Informe 2008". Ministerio de Trabajo e Inmigración.

PALACIO, J.I. y SIMÓN, H.J. (2006), "Segregación Laboral y Diferencias Salariales por Razón de Sexo en España". *Estadística Española*, vol. 48, nº 163, pág.493-524.

RECOLONS ARQUER, L (2005), "Las migraciones exteriores en la Evolución de la Población de España, 1950-2004", *Migraciones*, nº 17, pág. 45-89.

TORRES, R. (2007), "Las migraciones internas en el País Vasco durante el periodo 1991-2001. Evidencias de un proceso de contraurbanización". *Boletín de la Asociación de Geógrafos Españoles*, nº 43, pág. 85-106.

LA SOLIDARIDAD INTERTERRITORIAL COMO PRINCIPIO RECTOR EN UNA HACIENDA DESCENTRALIZADA

Mª GABRIELA LAGOS RODRÍGUEZ
E-mail: Gabriela.Lagos@uclm.es
RAQUEL ÁLAMO CERRILLO
E-mail: Raquel.Alamo@uclm.es
Departamento de Economía Política y Hacienda Pública, Política Económica y Estadística
Universidad de Castilla- La Mancha

Resumen

Si podemos afirmar que existe un acuerdo general en considerar que la solidaridad es un principio relevante en la ordenación de las haciendas descentralizadas, debemos reconocer que la controversia acompaña la aplicación efectiva del mismo. Las diferencias en su concepción son importantes –hay quienes la niegan en el ámbito interterritorial y la limitan al interpersonal- y evidencian la dificultad en la precisión del contenido de este principio. Esta indeterminación conlleva la propuesta de fórmulas de aplicación que, en nuestra opinión, son difícilmente compatibles con la esencia de la solidaridad. El propósito de esta comunicación es poner de manifiesto estas contradicciones tras definir el principio en las tres vertientes posibles –económica, jurídica y sociológica- y analizar su trascendencia en un sistema de financiación descentralizado, tomando como referencia el modelo de financiación autonómica español.

Palabras clave: solidaridad, federalismo, autonomía, descentralización.
Área temática: Economía Social y Discapacidad.

1. CONTENIDO DEL PRINCIPIO DE SOLIDARIDAD

Tal vez la tarea más complicada del estudio académico de la solidaridad, estriba en su precisión como concepto antes de analizar su plasmación en otros ámbitos como el económico o el jurídico. Aunque ésta es una cuestión que debería ser tratada como requisito previo en cualquier estudio, excede de los objetivos de éste trabajo. Baste señalar que el valor de solidaridad ha originado una notable profusión de interpretaciones, con frecuencia dispares: ha sido tomado como expresión social de la caridad cristiana, como el deber moral de cada individuo de mejorar la situación de sus semejantes, como el deber legal de contribuir al bienestar social, como elemento indispensable para garantizar la eficacia de la vida social, como expresión del principio de supervivencia de la especie humana, y como comportamiento espontáneo de colaboración que nace de la empatía entre sujetos o grupos afines.

Un elemento que complica aun más el concepto de solidaridad es que se inserta siempre en un contexto de reciprocidad que permite diferenciar varias clases, según los agentes entre los que se entable. Así, la solidaridad se predica tanto de las relaciones interindividuales como intersociales y es también válida como parámetro para evaluar los comportamientos recíprocos de las unidades territoriales en un contexto de descentralización. Por último, la solidaridad puede ser entendida en dos aspectos complementarios: se puede ser sujeto activo de la solidaridad, es decir, actuar de acuerdo a este principio, o sujeto pasivo –o receptor de la solidaridad –.

La multiplicidad de interpretaciones, la variedad de agentes involucrados y la dualidad que presenta su ejercicio, determinan la imprecisión del principio de solidaridad. Esta ambigüedad del concepto es, la causa de la confusión que existe en cuanto a sus fines y a las formulaciones que adopta y tiene su reflejo en la contraposición de posturas que pretenden hacer de su peculiar interpretación de este principio, la referencia a imponer al resto. Con frecuencia, un análisis

somero de estas propuestas revela que la protección de un interés particular y su voluntad de hacerlo prevalecer sobre el común y los intereses del resto sustentan tales interpretaciones. Es decir, la solidaridad, como principio ético, social, económico e incluso jurídico, ha sido tomada como baluarte para el sostenimiento de ideologías, normas y decisiones económicas de la más diversa naturaleza y en las que su concepto se ve alterado o contradicho.

Una mera aproximación histórica permite constatar la afirmación anterior. Sin detenernos en sus antecedentes en la filosofía griega o en la virtud teologal de la caridad, tomaremos como primera referencia el concepto de altruismo, que podemos identificar con la *benevolencia* de HUME o la *beneficencia* de Adam SMITH. Este último autor pone de manifiesto una de las características básicas del concepto de solidaridad: su traducción económica –la *beneficencia* se plasma en la provisión de bienes y servicios esenciales a quienes no tienen los medios propios para lograrlos –. Es precisamente este autor quien muestra ya una de las características de la interpretación de solidaridad que ha originado su multiplicidad de expresiones: el utilitarismo. Así, la razón de la conducta altruista la resume Adam SMITH en su famosa frase *dame tú lo que me hace falta, y yo te daré lo que te falta a ti*. En su obra *La Teoría de los sentimientos morales*, expresa el mismo convencimiento en el concepto de simpatía que mueve al "Homo oeconomicus" a la solidaridad con el afán de maximizar su interés, su utilidad o beneficio propio.

El segundo hito remarcable en el devenir histórico de la solidaridad se liga al principio de fraternidad. No nos quedamos ya en el mundo de las ideas, de los conceptos, sino que la Revolución Francesa propició su aparición en el mundo real, convertido en un principio de ordenación social. A medida que la Revolución avanza, se hace más patente la necesidad de *fraternizar*, lo que será la causa de la radicalización del proceso revolucionario. De esta forma, la lucha y la violencia se convertirán en un compañero constante de la idea de fraternidad no siendo ajena tampoco a la noción de solidaridad contemporánea[1].

Las teorías solidaristas y socialistas aportan otro elemento fundamental en la noción contemporánea del principio de solidaridad: suponen el abandono de la consideración del individuo aislado para centrarse en la proyección social. Tal vez el autor más destacado de estas corrientes será DURKHEIM. En su opinión, la solidaridad es un requisito para la existencia y desenvolvimiento de la vida social, actúa como condición garante de la supervivencia del grupo: sin ella no habría vida social, no habría sociedad[2]. Convierte, pues, a la solidaridad en el motor de la cohesión social. Además, como criterio ético, la solidaridad actúa como elemento de integración y legitimación del poder político[3]. El punto común o nexo de ambas manifestaciones de la solidaridad es hacer máximos y compatibles los niveles de autonomía individual e integración social.

En la concepción social, la solidaridad es una obligación legal impuesta por el poder público, por tanto, un acto desprovisto de toda subjetividad. Bajo el imperio de estas ideas, nacen y se desarrollan los mecanismos de asistencia social de principios del Siglo XX, paradigmas de la solidaridad interindividual. Esta solidaridad asistencial debe ser realizada por el Estado. Las autoridades públicas son las encargadas de satisfacer las necesidades de los más desfavorecidos, lo que se plasmará la doctrina administrativa del servicio público.

Los autores del momento, reaccionan al imperio del individualismo que había originado la supremacía de los valores de igualdad y libertad, predicables de cada sujeto considerado por separado, pero no como ser social. La obra de MARX y ENGELS, *El manifiesto comunista*, publicada a principios de 1848, se convierte en la fuente de inspiración de la organización del Estado. En lugar de instar a los ciudadanos a abandonar un sistema lleno de injusticias por otro mejor, el socialismo científico sostiene que la transformación de la sociedad deriva del conflicto entre los factores productivos y la adjudicación de la propiedad de la producción. La lucha de

[1] La solidaridad se liga a la reivindicación, no siempre pacífica, de mejoras sociales.

[2] Como categoría científica, distingue una solidaridad psicosocial que denomina *solidaridad mecánica* –derivada de la similitud de fines de los individuos de una misma colectividad y propia de las sociedades menos avanzadas –, y una solidaridad funcional a la que llama *solidaridad orgánica* –la que se establece entre individuos con fines contradictorios o antagónicos, que es el caso de las sociedades en las que existe especialización en el trabajo y que requiere un pacto o convenio que asegure el logro de un interés parcial a cada sujeto –.

[3] Aspecto éste muy criticado en cuanto puede servir para sostener regímenes políticos totalitarios que desvirtúen la noción de solidaridad.

clases propia del capitalismo desaparecerá en la sociedad comunista cuyo nexo de unión es la cooperación –solidaridad – entre todos los trabajadores.

La I Guerra Mundial supone una brusca interrupción de las tesis de solidaridad que se extendían por toda Europa. El término solidaridad cayó en desuso y fue sustituido por otros como integración o cohesión. Esta evolución del concepto de solidaridad se produce al tiempo que el Estado social se impone como forma de organización política y económica de las sociedades contemporáneas. Para hacer frente a las responsabilidades que asume, se dota al Estado de amplios poderes que le permitan conseguir los recursos necesarios para proveer bienes y servicios. Uno de los riesgos de este tipo de Estado es que incentive conductas de irresponsabilidad de los ciudadanos, lo que se manifiesta en la falta de iniciativa y esfuerzo en la búsqueda de sus propios medios de subsistencia, amparados en la convicción de que el Estado actuará en su auxilio.

Las críticas al Estado social también afectan a la noción de solidaridad, que, bajo una ideología humanista, no se concibe sino en armonía con la libertad individual. En consecuencia, el Estado ya no es considerado como el único responsable de su realización efectiva sino que los ciudadanos también están obligados a mantener una conducta solidaria. La solidaridad debe abarcar tanto los aspectos públicos –solidaridad obligatoria – como los privados –solidaridad voluntaria –.

El contrapunto de la posición humanista lo ofrece la Teoría de la Justicia de RAWLS. Sin adentrarnos en su obra, merece la pena recordar que para este autor las diferencias en las libertades básicas sólo son aceptables si se producen en la distribución de la renta y la riqueza. La pauta que permitirá saber si una desigualdad debe mantenerse o ser revocada es el *principio de diferencia* que postula que la variedad de posiciones sólo es admisible si las mayores expectativas de los más favorecidos, cuando desempeñan un papel en el funcionamiento del sistema social en su conjunto, mejoran las expectativas de los más desfavorecidos. Como principios complementarios presenta el de *reparación* –cuya pretensión es corregir hacia la igualdad la desviación provocada por las circunstancias personales –y el de *fraternidad* –que postula *no desear tener mayores ventajas salvo que sea en beneficio de los que están peor.*

El concepto rawlsiano de fraternidad se corresponde con la idea de solidaridad imperante en los Estados democráticos contemporáneos, caracterizados por la supremacía de los derechos y libertades individuales y de la preocupación por el bienestar social. Por tanto, la posición ante las desigualdades se presenta como el núcleo de las discusiones actuales sobre la solidaridad, asumido este valor como imprescindible condición de justicia social. La justicia social depende de dos factores: de la igualdad de la distribución -entendida como igualdad en los niveles de bienestar - y del bienestar total -entendido como la suma de las utilidades de todos los individuos -. La primera condición se proyecta en el principio de reparación, mientras que la segunda lo hace en el principio de fraternidad. No es difícil trasladar estos factores y principios a la terminología más común: la primera condición de justicia social encuentra su base en el principio de igualdad y la segunda en el de solidaridad. Ambas son indispensables para que exista justicia social, se complementan y relacionan entre sí, pero son diferentes.

El razonamiento expuesto puede ser trasladado a las relaciones interterritoriales. Cada jurisdicción territorial puede entenderse como un sujeto con características propias que lo separan del resto de los territorios con los que conforma una unidad política, económica y social. El principio de diferencia, con sus colaterales de reparación, reciprocidad y fraternidad, es el modelo de conducta que orientaría sus relaciones hacia la justicia y la maximización del bienestar de cada territorio y del conjunto que forman. Es preciso aceptar, por tanto, que sólo las desigualdades inevitables que redundan a favor de la generalidad son tolerables y que el patrón de medida sólo puede ser el territorio menos favorecido. De tal forma que la distribución del beneficio del trabajo conjunto de los territorios será justa si y sólo si maximiza las posibilidades del territorio más desfavorecido.

En la revisión histórica del concepto de solidaridad queda patente que la solidaridad se justifica y sustenta en la comunidad de intereses y necesidades y en la colaboración en el logro del bien común. En la actualidad, la idea de solidaridad presupone las de igualdad y, además, es esencial para la aplicación conjunta de ambos principios. La conducta solidaria es, pues, un medio para alcanzar la Justicia en las relaciones interindividuales, intersociales o interterritoriales.

Un elemento integrante de la noción actual de solidaridad es que la concibe como un acto voluntario y libre: es un proceso de interacción del conjunto en el que las relaciones de cooperación se imponen sobre las de oposición o competencia, y el fin común sobre los intereses parciales. Sin perjuicio de lo anterior, se mantienen los nexos con la idea de obligatoriedad, lo que explica el papel preponderante del Estado, no sólo en el establecimiento de las condiciones que hacen posible la desaparición de las diferencias sociales injustas, sino también como agente activo obligado a captar de los ciudadanos más favorecidos los recursos que le permitan hacer frente a las necesidades de los que están en desventaja. La función pública de redistribución es un instrumento primordial en el logro de la solidaridad colectiva, tanto de desde el punto de vista interpersonal como interterritorial.

A pesar de la importancia de la acción colectiva en el principio de solidaridad, la configuración de la sociedad en grupos o en individuos aislados exige el respeto de la libertad y voluntariedad individual y social, y explica el retroceso del Estado en la responsabilidad de decidir qué es bueno y cómo conseguirlo. La solidaridad exige hoy individuos, grupos sociales o jurisdicciones responsables de sí mismos y de su bienestar.

Existe la conciencia general de que las disparidades entre individuos, grupos o territorios impiden que la convivencia social resulte satisfactoria para el conjunto y para cada parte involucrada. El único medio eficiente de armonizar los intereses particulares con el interés común de todos es practicar la solidaridad, en suma, aportar cada uno y todos –el Estado – lo necesario para hacer desaparecer las desigualdades que impiden el desarrollo de los menos favorecidos para alcanzar unas condiciones de vida –social, política y económica – acordes a la dignidad humana.

Aunque el concepto es único cabe distinguir dos tipos de solidaridad:

- Una voluntaria, basada en el pacto entre individuos, grupos o jurisdicciones que se apoya en razones políticas de libertad.
- Y otra obligatoria, llevada a cabo por el imperio del poder público y cuyo fin más evidente es la equidad, es decir, la vigencia del principio de igualdad referido a las condiciones específicas de cada sujeto.

De esta manera, queda resuelta la dualidad de interpretación del concepto. La solidaridad voluntaria es racional, no relega el provecho propio y se basa en el egoísmo individual, por el contrario, la solidaridad obligatoria tiene una mayor carga ética porque se dirige al logro del bien común –entendido no como lo que más nos interesa a todos, sino como aquello que es bueno, sin entrar a considerar intereses propios o ajenos –, y por eso, para evitar que el egoísmo consustancial al sujeto impida su consecución, se encarga al Estado que la haga efectiva. Si el peligro de la solidaridad voluntaria es que se deje fuera a quienes, estando en desventaja, no tienen ninguna utilidad que ofrecer al sistema a cambio de la ayuda prestada, el riesgo del segundo es la anulación de la responsabilidad individual. En consecuencia, ambas fórmulas deben darse de forma conjunta para lograr la aplicación efectiva de la solidaridad, facilitando así la convivencia de libertad y equidad.

Para resumir y concretar lo expuesto hasta el momento, partimos de un principio ético máximo: la Justicia. Esta se escinde en dos exigencias: la libertad/ autonomía y la unidad/ equidad. El nexo de unión entre ambas es la solidaridad, que se plasma en un objetivo intermedio que satisface a la vez los dos requisitos anteriores: la corrección de las desigualdades. Si hacemos prevaler la libertad en el modo de corregir los desequilibrios, optaremos por técnicas de cooperación voluntaria, que sólo se pondrán en práctica si los cooperantes obtienen un provecho particular. Si, por el contrario, buscamos garantizar la equidad de trato, el medio idóneo es la ley que obliga a todos por igual. Autonomía y equidad deben estar presentes como condiciones de justicia, el problema es, pues, determinar el grado de libertad y equidad óptimo, que será el que proporcione mayor eficiencia al sistema social, es decir, aquél que asegure la realización de sus objetivos y suponga una menor carga económica.

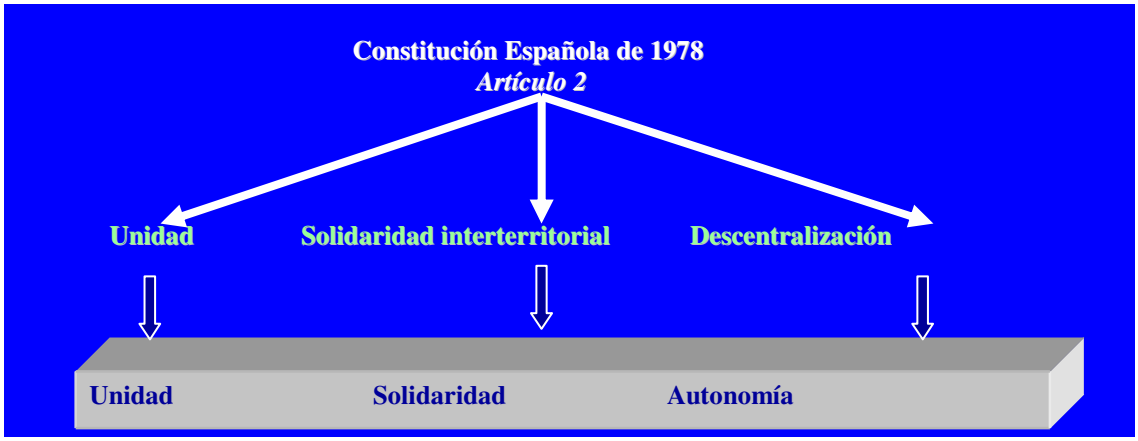
La solidaridad, por tanto, se mueve en el impreciso terreno de los opuestos. En ella conviven la emotividad y la racionalidad, lo bueno y lo eficiente, lo individual y lo social, el altruismo y el egoísmo. Por eso podemos afirmar, sin duda que allí donde haya convivencia –sociedad – será necesaria la solidaridad: una voluntad de actuación común en beneficio de todos.

2. LA SOLIDARIDAD EN EL ESTADO DE LAS AUTONOMÍAS

La Constitución de 1978 supuso la adopción de una fórmula descentralizada en el reparto del poder político, conocida bajo la denominación de Estado de las Autonomías. La decisión del constituyente español se expresa en el artículo 2 de la Norma Fundamental, según el cual *la Constitución se fundamenta en la indisoluble unidad de todos los españoles y reconoce y garantiza el derecho a la autonomía de las nacionalidades y regiones que la integran y la solidaridad entre todas ellas*. Por tanto, podemos afirmar que el nuevo sistema no se limita a una mera descentralización de la organización de la Administración Central, sino que conlleva el reparto efectivo del poder político hacia unidades territoriales de rango inferior, de manera que el Estado Central no ostenta ya el monopolio de la actividad pública. A tenor de lo dispuesto, las Comunidades Autónomas –denominación general de nacionalidades y regiones –se convierten en titulares de un derecho a la autonomía que se traduce en su capacidad de autogobierno, es decir, de dotarse de normas jurídicas propias y poder decidir sobre las cuestiones que les afectan.

Aunque su actuación se desarrolle dentro de los límites que la propia Constitución y las leyes establecen, lo cierto es que el reconocimiento de la autonomía de nacionalidades y regiones implica la existencia de múltiples centros de decisión política. La mera existencia de sujetos distintos con preferencias diferentes puede generar colisiones de actuación y conflictos de intereses que será necesario resolver. En todo caso, de lo dispuesto en el precepto constitucional citado debemos deducir tres elementos fundamentales en la vertebración política de nuestro país: la unidad del Estado español –principio de unidad –, la descentralización del poder político en unidades territoriales subcentrales –principio de autonomía – y la exigencia de un comportamiento solidario entre nacionalidades y regiones –principio de solidaridad –.

Esquema Nº1. Principios derivados del Artículo 2 de la Constitución



Elaboración propia.

A pesar del reconocimiento general de la trascendencia del principio de solidaridad en nuestro Ordenamiento jurídico, no existe acuerdo en cuanto a la precisión de su naturaleza y las obligaciones que conlleva, en especial en el ámbito de las relaciones entre el Estado y las Comunidades Autónomas. Esta situación de ambigüedad y controversia supone un riesgo importante de inaplicación de dicho principio.

La primera y más importante referencia a la solidaridad la encontramos en el Titulo Preliminar de la Constitución, en el ya citado artículo 2. Su posición privilegiada en el Texto Constitucional determina su condición de *fundamento político* del Estado y, como tal, sirve de base a todo el sistema jurídico. A este precepto, se le unen varios más que refuerzan la posición constitucional de principio de solidaridad y concretan su contenido:

⇒ En el Título VIII, *De la Organización Territorial del Estado*, en el Capítulo I, dedicado a los principios generales, el artículo 138 concreta la exigencia de solidaridad como patrón inspirador de la nueva estructura territorial.

→ El apartado primero de este precepto prescribe la realización efectiva de la solidaridad y designa al Estado como responsable de la misma. Es la expresión genérica de la solidaridad interterritorial, si bien, el enunciado del artículo alude sólo a las condiciones económicas de las Comunidades Autónomas.

→ El segundo párrafo, expresa la exigencia de solidaridad en términos negativos, como prohibición de privilegios económicos o sociales entre las diferentes Comunidades Autónomas. En esta segunda formulación se ponen de relieve las conexiones de la solidaridad con el principio de igualdad en sus dos proyecciones: formal porque la ausencia de privilegios se puede expresar también como el derecho a no recibir trato discriminatorio, y material, es decir, en el resultado, en tanto exigencia de que no existan desequilibrios económicos interterritoriales.

⇒ Las otras dos referencias expresas al principio de solidaridad –artículos 156.1 y 158.2– tratan de manera concreta los aspectos económicos y financieros. Ambos preceptos actúan de forma complementaria en la consecución del objetivo que la Constitución impone al Gobierno nacional: la ausencia de desequilibrios económicos entre las Comunidades Autónomas.

→ El artículo 156 configura la solidaridad como una limitación del derecho de autonomía. Es, por tanto, la vertiente negativa del principio que impide que la diversidad que puede originar la libertad autonómica en materia financiera incremente las disparidades existentes.

→ Por su lado, el artículo 158, en su párrafo segundo, obliga a que el Estado actúe contra los desequilibrios económicos y designa para esta función un instrumento financiero concreto –el Fondo de Compensación Interterritorial –, lo que no excluye otras posibilidades que pudieran utilizarse. Según lo dispuesto en este artículo, la corrección de las disparidades económicas se configura como un deber positivo de actuación que, aunque concierne de forma general a todos los poderes públicos, es una carga propia del Estado, en tanto garante del cumplimiento efectivo del principio de solidaridad.

Por tanto, en nuestro Ordenamiento, el principio de solidaridad es una regla jurídica positiva, con expresiones legales concretas, a diferencia de otros países, en los que la noción de solidaridad y su especificación en deberes y derechos es fruto de la actividad doctrinal y jurisprudencial. Así, podemos distinguir:

• Solidaridad Interpersonal.- De la que tratan los artículos 45.2 –*Los poderes públicos velarán por la utilización racional de todos los recursos naturales, con el fin de proteger y mejorar la calidad de vida y defender y restaurar el medio ambiente, apoyándose en la indispensable solidaridad colectiva*– y 156.1, que establece que la autonomía financiera de las Comunidades Autónomas deberá ser ejercida de acuerdo a *los principios de coordinación con la Hacienda estatal y de solidaridad entre todos los españoles*.

• Solidaridad Interterritorial.- Referida en los artículos 2 y 138 –que designa al Estado como garante de la realización efectiva del principio de solidaridad y, para ello le encarga velar *por el establecimiento de un equilibrio económico adecuado y justo entre las diversas partes del territorio español, y atendiendo en particular a las circunstancias del hecho insular*– y 158 .2 –que crea un instrumento financiero destinado a inversión, el Fondo de Compensación Interterritorial, con el fin *corregir desequilibrios económicos interterritoriales y hacer efectivo el principio de solidaridad*–.

Sea cual sea la acepción que tomemos, el contenido del principio de solidaridad no es fácil de determinar. Se trata de un principio sometido a mutaciones dado su carácter relativo y cuya interpretación varía según el momento histórico. Para algunos autores, una minuciosa definición jurídica de su contenido puede ser peligrosa al impedir su adaptación a las circunstancias cambiantes de la sociedad. Sin embargo, como ya apuntamos, una excesiva imprecisión puede convertirlo en una declaración sin valor jurídico. Para evitar este peligro es necesario identificar las exigencias que conlleva la proclamación genérica del principio de solidaridad, en especial, en el ámbito interterritorial.

El enunciado del artículo 2 refleja la convivencia de dos tendencias opuestas en el ordenamiento constitucional: la disgregadora que se acoge al derecho de autonomía y la unitaria que pretende la pervivencia del Estado español como entidad jurídica. En este contexto, la proclamación constitucional del principio de solidaridad es fruto de la nueva estructura territorial que se adopta: sin solidaridad no sería posible la convivencia entre los diversos territorios que integran el Estado español, habida cuenta de los desequilibrios económicos y la diversidad cultural que presenta. Así pues, la solidaridad es el principio de articulación que permite la integración de los intereses particulares –a los que atiende el principio de autonomía – y los intereses generales –contemplados en el principio de unidad del Estado –.

Podríamos diferenciar, diferencia tres clases distintas de la solidaridad interterritorial, en razón de sus ámbitos de proyección:

• Una Solidaridad Política reflejada en el artículo 2 de la Constitución. Según esta interpretación, la solidaridad es un fundamento político de la Nación-Estado y de la propia Constitución. Es el *presupuesto necesario para la integración social de las nacionalidades y regiones en la Nación-Estado, un modo de entender y articular la unidad* y su verdadera garantía política. El riesgo es que se pretendan excluir las medidas impuestas por parte del Estado para velar por el cumplimiento de este principio. Es inherente a la constitución y existencia de un Estado la unidad de los elementos que lo integran, y este nexo y las garantías que lo acompañan, son algo más que meras obligaciones jurídicas, son la expresión de un fundamento político consubstancial a la propia noción del ente estatal, cualquiera que sea su forma jurídica.

• Una Solidaridad Jurídica referida al obligado cumplimiento del principio de solidaridad en las relaciones *entre el Estado y las Comunidades Autónomas y entre éstas entre sí, conforme a la cual se pueden imponer conductas obligatorias para salvaguardar las necesidades generales del sistema autonómico o condicionar actos y comportamientos*. Su mayor virtualidad sería la de actuar como límite positivo y negativo en el ejercicio de las competencias, tanto del Estado como de las Comunidades Autónomas.

• Una Solidaridad Económica que surge de la evidencia de la desigualdad económica existente entre las Comunidades Autónomas y que se configura, sobre todo, como un deber del Estado de asegurar un equilibrio económico justo entre las diversas partes del territorio español –artículo 138.1.CE–. En consonancia con este objetivo, el Fondo de Compensación Interterritorial se instituye como el instrumento más importante de la solidaridad de este tipo, junto con las asignaciones de nivelación de servicios públicos fundamentales.

Esta noción de solidaridad económica orienta también la organización fiscal del Estado, con dos consecuencias importantes:

⇒ Por un lado, actúa como un límite negativo en la descentralización de la potestad tributaria. En consecuencia, la vigencia del principio de solidaridad no contradice la pluralidad fiscal dentro del Estado hasta el límite en que ponga en peligro la utilización de la política fiscal como un medio de garantizar el equilibrio económico justo entre las diversas regiones.

⇒ Y por otro, debe ser la referencia a utilizar al establecer la contribución de una Comunidad Autónoma al sostenimiento de los gastos generales del Estado, lo que supone la

aplicación del principio de capacidad de pago: cuanto mayor sea la riqueza de una región, mayor será su aportación.

Según los sujetos entre los que se establece el vínculo solidario, se distinguen tres dimensiones de la solidaridad interterritorial:

• Entre Comunidades Autónomas y el Estado, en la medida en que es preciso un comportamiento leal por parte de cada una de las entidades que integran en Estado para asegurar la cohesión interna. En consecuencia, la solidaridad de las Comunidades Autónomas para con el Estado cumple la misión de asegurar la prevalencia del interés común y, de manera recíproca, el Estado deberá actuar a favor de la desaparición de los desequilibrios existentes entre las Comunidades Autónomas, en especial en el ámbito económico, y no sólo a través de transferencias sino también a través de acuerdos de colaboración o cooperación.

• Dentro del ámbito intrarregional, lo que obliga a los gobiernos regionales a tomar las medidas necesarias en favor del desarrollo de aquéllas áreas geográficas o colectivos humanos en desventaja respecto al resto de los que integran el territorio y la población de la Comunidad Autónoma.

• Entre Comunidades Autónomas, de manera que se exige de cada Comunidad un comportamiento solidario con el resto de Autonomías, en especial, con las menos desarrolladas. Bajo esta premisa, las relaciones de solidaridad interregional deben ser contempladas desde dos perspectivas diferentes: una general, impuesta por la exigencia constitucional de solidaridad y otra específica, que se expresará en los pactos, convenios o acuerdos de cooperación que cada Comunidad Autónoma concluya con otra u otras.

Debemos tener presente que el respeto al interés común y al ajeno no tiene por qué ser incompatible con la persecución del interés propio. Es legítimo que cada unidad territorial actúe en pro de su bienestar, pretendiendo obtener provecho de sus decisiones. Tal conducta sólo puede ser declarada insolidaria si perjudica a otras jurisdicciones o al conjunto estatal. Si la violación procede del nivel autonómico, el Estado puede actuar a través del mecanismo de ejecución forzosa previsto en el mencionado artículo 155 de la Constitución, según el cual, si una Comunidad Autónoma incumpliese sus obligaciones constitucionales o legales*, o actuare de forma que atente gravemente al interés general de España*, el Gobierno de la Nación podrá *adoptar las medidas necesarias para obligar a aquélla al cumplimiento forzoso de dichas obligaciones o para la protección del mencionado interés general*[4]. En caso de que sea el Estado el que vulnere el interés general o el legítimo derecho de una Comunidad Autónoma a la obtención de un provecho que no ocasiona perjuicio a otra Autonomía o al conjunto nacional, podrán solicitar la tutela de su derecho por vía judicial.

La solidaridad no sólo puede realizarse a través de medidas coercitivas, sino que también se puede actuar a favor del interés común por medio de pactos o acuerdos. Al respecto, el ya citado artículo 145 de la Constitución[5], permite en su párrafo segundo los convenios entre Autonomías *para la gestión y prestación de servicios propios*, remitiendo la regulación de sus condiciones a los respectivos Estatutos. Exige sólo que dichos convenios sean comunicados a las Cortes Generales, por tanto, el Estado no tiene poder de decisión. La causa de esta conformidad estatal se encuentra en la limitación de la materia objeto de convenio: sólo la gestión y prestación de servicios de titularidad autonómica. El razonamiento que subyace en este enunciado jurídico es que estos pactos responderán a criterios técnicos y no políticos, es decir, se concluirán sólo si las

[4] El mismo precepto establece dos condiciones para la actuación estatal, que son el previo requerimiento a la máxima autoridad autonómica, el Presidente de la Comunidad, y la autorización por mayoría absoluta del Senado. Este último requisito, tal como se configura la representación política en la Cámara Territorial, hace muy difícil en la práctica la intervención estatal.

[5] Los antecedentes de la redacción definitiva de este artículo ilustran bien cómo fue el proceso de apertura de esta vía en favor de las Autonomías. Así, de las primeras redacciones en las que se primaba el papel del Estado, éste se reduce al control de los convenios de gestión y prestación de servicios propios si bien se exige autorización de las Cortes para los acuerdos de cooperación..

Comunidades participantes se benefician de su actuación conjunta, lo que en términos económicos supone un incremento del rendimiento a obtener o, lo más frecuente, un ahorro en el gasto de provisión del servicio. Al despreciar el componente político de estos pactos, se entiende que no van a perjudicar el interés del común del Estado. En consecuencia, no sólo no interfieren la ejecución de las competencias estatales sino que, además, se consideran un instrumento adecuado de favorecer la eficiencia de la actuación regional en la prestación de los servicios que tiene encomendados.

El riesgo evidente es que dos o varias Autonomías se comprometan a mantener posturas unánimes frente a ciertos asuntos trascendentales en la dirección y gestión del Estado y actúen así como auténticos grupos de presión. Tal posibilidad se ve favorecida por el mecanismo de representación parlamentaria en nuestro país que permite la presencia de partidos políticos de circunscripción territorial autonómica, por lo general de vocación localista –o nacionalista –, en el Congreso de los Diputados, Cámara que asume el protagonismo en la actividad legislativa estatal. Para evitar que las decisiones que deben tomarse en atención al interés general, es decir, el de todos los ciudadanos españoles con independencia de su Comunidad de residencia, sean perturbadas a favor de los intereses particulares de ciertas Autonomías, se establece que las Cortes Generales analicen con carácter previo el contenido del acuerdo autonómico. Como garantes del interés común sólo darán la autorización al convenio cuando se aseguren de que no se perjudica su consecución.

Lo común en las tres modalidades anteriores de solidaridad, es que se fundamentan en una relación jurídica bilateral. En consecuencia, la solidaridad interterritorial exige la reciprocidad de actuaciones entre los gobiernos regionales, o de estos con el Estado o viceversa, o, incluso, de la autoridad autonómica con las áreas geográficas o colectivos bajo su jurisdicción.

Ahora bien, la regulación constitucional configura la solidaridad como un deber para el Estado, garante de su cumplimiento efectivo, mientras que no hay un precepto expreso que responsabilice a las Comunidades en la realización del principio de solidaridad, sino que se pretende que la conducta solidaria entre Comunidades y con el Estado sea más bien una cooperación voluntaria. Desde el ámbito estatal, la solidaridad es un deber, para las Comunidades Autónomas un derecho, mientras que se carga al Estado con el deber de hacer efectivo el principio de solidaridad, se relaja la responsabilidad de las Comunidades en su consecución.

3. EL SISTEMA DE FINANCIACIÓN AUTONÓMICO Y SOLIDARIDAD

No queremos realizar en este apartado sino un breve repaso de la situación de los instrumentos financieros de corrección de disparidades en el sistema de financiación de las Comunidades Autónomas. Por tanto, nos centraremos en aquellos aspectos más relevantes y de mayor importancia en orden a la argumentación de nuestras conclusiones.

3.1. El fondo de compensación interterritorial

Es sabido que el Fondo de Compensación Interterritorial es el principal instrumento de solidaridad del sistema, es decir, recae sobre él la responsabilidad de corregir de manera efectiva las disparidades económicas interterritoriales. Sin entrar en sus antecedentes más remotos, sí nos interesa remarcar que el FCI no siempre ha servido al fin de la corrección sino que llegó a plantearse como un instrumento de suficiencia, es decir, como un aporte de recursos para hacer frente a los gastos de las CC.AA. La última reforma del FCI, continúa definiéndolo como el instrumento corrector de los desequilibrios económicos interterritoriales, mediante gastos de inversión. Lo que hace la Ley 22/2001, de Financiación de las Comunidades Autónomas de Régimen Común es reconfigurar los criterios que determinan las comunidades beneficiarias -las consideradas por la Unión Europea como objetivo 1, Ceuta y Melilla- y alterar el destino del Fondo para que pueda financiar no sólo gastos de inversión, sino también el gasto corriente asociado a esa inversión. Así, se dispone como mínimo se podrán adscribir para gastos de

inversión el 75% del Fondo, de forma que como máximo el 25% restante podrán destinarse a gastos de funcionamiento asociados a la inversión. Tenemos, pues, dos fondos: uno por importe del 75 por ciento del total, destinado exclusivamente a gastos de inversión, y otro fondo complementario, también dotado en el capítulo de transferencias de capital, que puede ser destinado a cubrir los gastos de funcionamiento asociados a los proyectos de inversión financiados con cargo a ambos fondos.

A tenor de su definición como un instrumento de desarrollo regional destinado exclusivamente a hacer efectivo el principio de solidaridad, su distribución se realiza entre las Comunidades Autónomas menos desarrolladas, considerando como tales aquellas cuya renta por habitante es inferior al 75 por ciento de la media comunitaria. Este criterio coincide con el establecido por la Unión Europea para seleccionar las regiones del Objetivo nº1, beneficiarias de los Fondos Estructurales. Los datos del 2009 dan el siguiente panorama de la financiación por FCI en España:

Cuadro 1

FONDOS DE COMPENSACIÓN INTERTERRITORIAL

	AÑO 2008		AÑO 2009				Variación
	TOTAL		Fondo de Compensación	Fondo Complementario	TOTAL FCI		2009-2008
	Miles de euros	%	(Miles de euros)	(Miles de euros)	Miles de euros	%	(%)
ANDALUCÍA	481.618,05	36,00	358.607,35	119.523,83	478.131,18	35,32	-0,72
ASTURIAS	56.564,87	4,23	42.005,63	14.000,48	56.006,11	4,14	-0,99
CANARIAS	90.564,38	6,77	75.814,07	25.268,83	101.082,90	7,47	11,61
CANTABRIA	10.630,61	0,795	8.077,24	2.692,14	10.769,38	0,796	1,31
CASTILLA Y LEÓN	85.058,06	6,36	62.928,48	20.974,06	83.902,54	6,20	-1,36
CASTILLA-LA MANCHA	114.341,62	8,55	85.470,77	28.487,40	113.958,17	8,42	-0,34
C. VALENCIANA	116.884,02	8,74	100.535,61	33.508,52	134.044,13	9,90	14,68
EXTREMADURA	97.203,57	7,26	71.935,40	23.976,07	95.911,47	7,08	-1,33
GALICIA	203.298,90	15,19	147.919,41	49.301,54	197.220,95	14,57	-2,99
MURCIA	61.135,90	4,57	46.364,14	15.453,17	61.817,31	4,57	1,11
CEUTA	10.340,81	0,77	7.847,32	2.615,51	10.462,83	0,77	1,18
MELILLA	10.340,81	0,77	7.847,32	2.615,51	10.462,83	0,77	1,18
TOTAL	**1.337.981,60**	**100,00**	**1.015.352,74**	**338.417,06**	**1.353.769,80**	**100,00**	**1,18**

Fuente: Presupuestos Generales del Estado de 2008 y 2009.

La base de cálculo del Fondo de Compensación está constituida por los gastos de inversiones reales nuevas de carácter civil incluidas en los Presupuestos del Estado y sus Organismos Autónomos. Esta cantidad se pondera por la población relativa de las Comunidades Autónomas beneficiarias del Fondo respecto de la población total de España y por el índice resultante del cociente entre la renta por habitante media nacional y la renta por habitante media de las Comunidades Autónomas incluidas en el Fondo[6]. A esta cantidad le aplicamos el porcentaje fijado por la Ley General de los Presupuestos[7] y se incrementa en un 1,57% destinándose a las Ciudades con Estatuto de Autonomía de Ceuta y Melilla[8].

La dotación para el Fondo Complementario será 33,33% del Fondo de Compensación de cada Comunidad Autónoma[9]. En el gráfico siguiente podemos ver la evolución de estos Fondos, en el destaca el pronunciamiento de la tendencia creciente que se produce a partir de 2002 y la estabilidad de los dos últimos ejercicios.

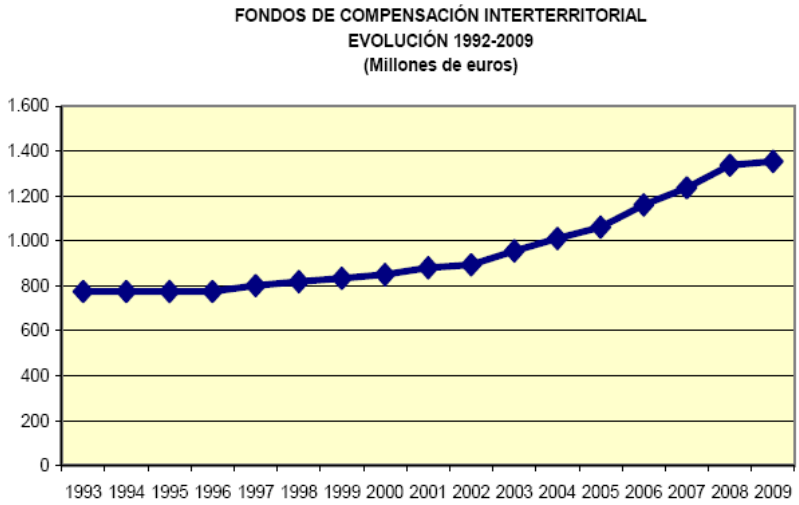
[6] Para 2009, la inversión real civil nueva asciende a 5.302.428,81 miles de euros, y los índices de población relativa y de renta relativa fueron de 0,58163346 y 1,17904647, por tanto, la base de cálculo del Fondo de Compensación es de 3.636.261,850 miles de euros.

[7] Para 2009 es 27,49137% lo que da un valor de 999.658,10 miles de euros.

[8] El Fondo de Compensación en 2009 es de 1.015.352,74 miles de euros.

[9] 338.417,06 miles de euros en 2009.

Gráfico 1



Fuente: Fondos de Compensación Interterritorial (Año 2009). Secretaría de Estado de Cooperación Territorial. MAP

Una vez determinado el volumen del Fondo de Compensación Interterritorial su importe se distribuye entre las Comunidades Autónomas beneficiarias según las siguientes variables: Población (87,5 por ciento), saldo migratorio (1,6 por ciento), paro (1 por ciento), superficie (3 por ciento) y dispersión de la población (6,9 por ciento). Efectuado el reparto del Fondo en función de las variables y ponderaciones citadas, el resultado obtenido se redistribuye según las variables inversas de la renta e insularidad –que supone el incremento del 63,1 % la dotación correspondiente de la Comunidad de las Islas Canarias-.

3.2. Las asignaciones de nivelación de servicios públicos fundamentales

Las asignaciones de nivelación deben garantizar un nivel mínimo en la prestación de los servicios públicos fundamentales en todo el territorio español, que, tras un largo tiempo sin definición, quedaron fijados en educación y sanidad. Su existencia dentro del sistema de financiación se justifica en el interés del Estado en asegurar una determinada prestación de los servicios públicos fundamentales. No son transferencias dirigidas a corregir desequilibrios financieros horizontales, su fin no es garantizar la suficiencia de las Haciendas Autonómicas, sino permitir que todos los ciudadanos tengan acceso a un nivel deseable de prestación pública referido a ciertos servicios. Su propósito es la nivelación de los niveles de provisión pública en determinadas áreas, lo que supondrá un incremento de la financiación transferida a las regiones deficitarias, pero no vinculado al gasto asumido por la Comunidad sino a la corrección de la disparidad existente entre el nivel de prestación que ofrece a sus ciudadanos y el mínimo que debería prestar. Esta diferencia se refleja también en la fórmula elegida para hacerla efectiva, a través de transferencias condicionadas al cumplimiento del fin para el que se dotan.

Cuando el reparto de las fuentes de financiación entre los entes subcentrales produce diferencias en el volumen de recursos que recibe cada uno de ellos, es necesario aplicar mecanismos correctores que permitan restaurar el equilibrio interterritorial y garantizar que todos estarán en condiciones de hacer frente a los gastos que hayan asumido. Pero este objetivo es distinto al de garantizar que todos los ciudadanos del Estado disfruten del nivel de prestación de servicios públicos fundamentales que se estime adecuado. En este sentido puede afirmarse que el nivel de provisión en los servicios fundamentales es considerado como un bien preferente del Estado que se materializa en subvenciones a las Comunidades Autónomas que no pueden alcanzarlo, por sí mismas, en los servicios que son de su competencia.

Las asignaciones de nivelación servirán para mantener la calidad de la enseñanza y de la sanidad pero les han sido impuestas unas restricciones tan severas que hacen difícil su

aplicación práctica. Así, se exige un crecimiento superior al promedio nacional en más de tres puntos porcentuales en el número de estudiantes de enseñanza obligatoria o de la población protegida por el sistema sanitario público[10].

3.3. La solidaridad en la financiación ordinaria del sistema

Nos proponemos en este apartado, hacer una revisión somera de cómo el modelo de financiación autonómico contempla el objetivo de solidaridad en la determinación de los recursos que distribuye a las Comunidades. Es decir, nos centramos en los instrumentos que tendrían que venir a compensar las diferencias que se dan entre las Autonomías en cuanto a la capacidad que tienen para generar los recursos necesarios para hacer frente a sus competencias. Precisemos que ser capaces de financiar sus gastos las hace suficientes pero no garantiza que puedan proveer un nivel similar de servicios públicos a sus ciudadanos. En esencia, lo que tendríamos que encontrar son mecanismos de nivelación que compensen las diferencias de riqueza y su repercusión sobre la financiación que son capaces de generar, sobre todo en un sistema como el actual en el que la potestad tributaria y los recursos directamente provenientes de la recaudación territorializada son la base del mismo.

El análisis debe partir del cálculo de las necesidades de gasto de las Comunidades Autónomas, que se hace a partir de unas variables diferentes en función del destino de la financiación que aportan y con un peso relativo también distinto. Sin ánimo de ser exhaustivos, destacamos que para el bloque de Competencias Comunes se prevé que el 94% corresponda a la población, un 4,2% por Superficie, el 1,2% por Dispersión de la población y un 0,6% por Insularidad. Este reparto da un peso muy considerable a la población como criterio de necesidad. Los fondos que lo complementan son: un fondo de garantía de mínimos[11], un fondo por menor renta relativa que se distribuye entre las Comunidades cuya renta relativa sea inferior a la media y en proporción a las diferencias respecto a la media y un fondo de menor densidad de población para las Autonomías cuya densidad de población sea inferior a 27 habitantes por Km^2 y cuya extensión sea inferior a 50.000 Km^2[12]. El Fondo de Renta Relativa se distribuirá entre las Comunidades que tengan una menor riqueza relativa, calculada a partir de la relación entre su población y la total del conjunto de las Autonomías y el Valor Añadido Bruto a coste de factores que presenta respecto de la media:

$$IRRi = (Pi/PT) - (VABi/VABT)$$

Donde:

Pi = Población de la Comunidad Autónoma i
PT = Población total de las quince Comunidades Autónomas de régimen común
VABi = Valor Añadido Bruto al coste de los factores de la Comunidad Autónoma i
VABT = Valor Añadido Bruto al coste de los factores de las quince Comunidades Autónomas de régimen común

Las Comunidades que son beneficiarias de este fondo adicional son Andalucía, Galicia, Extremadura, Castilla- La Mancha, Castilla y León, Asturias, Cantabria, Canarias, Murcia y La Comunidad Valenciana.

La acción de estas variables así como la disposición de dos Fondos adicionales de carácter redistributivo sirven para que algunos autores sostengan que este Fondo contribuye a la aplicación efectiva del principio de solidaridad. No es ésta nuestra opinión sino que entendemos que, tal como el Acuerdo lo define, es un instrumento financiero de suficiencia, es decir, su objetivo es hacer posible que cada Comunidad perciba los recursos que necesita para ejercer sus competencias. Cuestión distinta es que se calcule a partir de unas variables que indican una

[10] La diferencia se reduce a un punto para las comunidades con más de 90.000 km^2
[11] Que servirá para compensar a aquellas Comunidades que obtengan una financiación inferior a la restricción inicial del sistema (1999).
[12] Son las Comunidades de Extremadura y Aragón.

mayor o menor necesidad pero, ¿quiere ello decir que se asegura que todas las Comunidades estén proveyendo un nivel adecuado de prestaciones públicas? Dicho de otro modo, qué ocurriría si el nivel de prestación del servicio "x" en la comunidad "y" estuviese por debajo del nivel medio de prestación en el conjunto de las comunidades, ¿el Fondo le permitiría obtener la financiación adicional necesaria para incrementar su nivel de prestación?

El Fondo de Suficiencia es el mecanismo de cierre del sistema de financiación autonómico actual. Se calcula a partir de las diferencias entre las necesidades de financiación de las Comunidades Autónomas –tomando como referencia el año base del modelo (1999)- y los recursos de naturaleza tributaria obtenidos por tasas, tributos cedidos, la tarifa autonómica del IRPF (33%) y las participaciones en la recaudación del IVA (35%), Impuestos Especiales (40%) y 100% el Impuesto sobre la Electricidad, del Impuesto Especial sobre Determinados Medios de Transporte y del Impuesto sobre Ventas Minoristas de Determinados Hidrocarburos. Si el resultado es positivo, la Comunidad percibirá la cantidad fijada para el año base incrementada por el índice de evolución que experimenten los Ingresos Tributarios del Estado entre 1999 y el año en el que se calcula el Fondo. Si, por el contrario, el resultado es negativo, el importe es un derecho a favor de la Hacienda Estatal y tendrá que revertirse por parte de la Comunidad Autónoma al nivel central. Al igual que ocurría en el primer caso, el importe negativo que resulte se pondera por el incremento de los ingresos tributarios, ésta vez, de la Comunidad Autónoma, salvo que sea mayor que el índice estatal, en cuyo caso se aplica éste. En la tabla siguiente, extraída de los datos del Ministerio de Hacienda sobre la financiación de las Comunidades Autónomas en 2006, observamos que sólo Madrid y Baleares tienen signo negativo.

4 - LIQUIDACIÓN DEL FONDO DE SUFICIENCIA DE 2006

4.2. CÁLCULO DE LA LIQUIDACIÓN DEL FONDO DE SUFICIENCIA

Miles de euros

Comunidad Autónoma/ Ciudad	Valor Fondo de Suficiencia 1999 aplicable en 1-1-06, excluido Fondo IT	Valor definitivo Fondo de Suficiencia 2006	Entregas a cuenta realizadas en 2006	Liquidación 2006
	(1)	(2)=(1)*ITEn o ITEr	(3)	(4)=(2)-(3)
Cataluña	1.590.059,16	2.692.765,19	2.401.435,01	291.330,18
Galicia	2.133.226,84	3.612.619,65	3.221.856,05	390.763,60
Andalucía	5.335.355,64	9.035.424,78	8.062.553,76	972.871,02
Principado de Asturias	642.529,71	1.088.124,06	971.174,84	116.949,22
Cantabria	405.595,00	686.875,13	612.600,95	74.274,18
La Rioja	221.572,93	375.233,76	334.646,99	40.586,77
Región de Murcia	650.853,83	1.102.220,96	983.056,68	119.164,28
Comunitat Valenciana	1.539.094,60	2.606.456,71	2.324.385,52	282.071,19
Aragón	647.075,67	1.095.822,65	976.992,08	118.830,57
Castilla-La Mancha	1.388.065,15	2.350.688,33	2.096.270,91	254.417,42
Canarias	1.879.952,36	3.183.699,32	2.839.431,19	344.268,13
Extremadura	1.129.347,10	1.912.549,31	1.705.749,10	206.800,21
Illes Balears	-136.879,71	-211.684,47	-206.668,98	-5.015,49
Madrid	-107.654,18	-182.312,35	-158.033,34	-24.279,01
Castilla y León	1.799.862,29	3.048.066,79	2.718.454,65	329.612,14
Total CC.AA.	**19.118.056,39**	**32.396.549,81**	**28.883.905,41**	**3.512.644,40**
Melilla	6.920,05	11.719,10	10.451,88	1.267,22
Ceuta	8.698,45	14.730,83	13.137,93	1.592,90
Total general	**19.133.674,89**	**32.422.999,74**	**28.907.495,22**	**3.515.504,52**

Si comparamos la información anterior con la siguiente, que recoge los recursos tributarios de las Comunidades Autónomas en 2006, podemos constatar que el Fondo de Suficiencia actúa como un instrumento redistribuidor de los recursos tributarios. Aunque lo idóneo es analizar varios ejercicios, nos conformamos en este trabajo con tomar como referencia el ejercicio 2006 –el último para el que disponemos de datos del Ministerio-. Así podemos constatar que mientras los datos de recaudación muestran una mayor dispersión y pueden relacionarse con el nivel de renta de la Comunidad, el resultado final de la financiación de cada Comunidad es más homogéneo y responde más al cálculo de necesidades que acoge el modelo.

Miles de euros

Comunidad Autónoma	Tributos cedidos y Tasas con criterio normativo			IRPF (Tarifa autonómica sin capacidad normativa)	IVA	II.EE.	Rendimiento definitivo en 2006 por el I. sobre Determinados Medios de Transporte (*)	Rendimiento definitivo en 2006 por el I. sobre Ventas Minoristas Hidrocarburos (sin competencias normativas) (*)	Fondo de Suficiencia	TOTAL FINANCIACIÓN
	Valor 99 (a 1-1-06)	ITE aplicable según el art. 15.1 b) de la Ley 21/01	Valor 2006							
	(1)	(2)	(3)=(1)*(2)	(4)	(5)	(6)	(7)	(8)	(9)	(10)=(3)+(4)+(5)+(6)+(7)+(8)+(9)
Cataluña	1.490.339,33	1,6935	2.523.889,66	5.076.239,27	3.688.320,44	1.521.828,68	367.509,64	153.050,38	2.692.765,19	16.023.603,25
Galicia	357.572,15	1,6935	605.548,44	1.112.023,36	1.231.977,51	513.284,51	86.944,35	57.986,28	3.612.619,65	7.220.384,10
Andalucía	978.853,72	1,6935	1.657.688,77	2.914.757,62	3.311.070,21	1.438.885,83	308.264,58	153.880,25	9.035.424,78	18.819.972,04
Principado de Asturias	179.866,09	1,6935	304.603,22	565.117,36	524.508,33	207.273,85	33.635,91	19.986,38	1.088.124,06	2.743.249,11
Cantabria	83.299,08	1,6935	141.066,99	304.144,45	290.369,77	126.086,87	20.854,67	13.791,81	686.875,13	1.583.189,69
La Rioja	48.029,88	1,6935	81.338,60	175.053,07	146.281,22	61.451,81	10.540,85	7.498,38	375.233,76	857.397,70
Región de Murcia	129.882,92	1,6935	219.956,73	549.755,47	543.166,44	300.297,24	63.795,06	35.931,95	1.102.220,96	2.815.123,85
Comunitat Valenciana	839.302,59	1,6935	1.421.358,94	2.343.282,42	2.147.658,02	986.570,18	233.290,37	101.188,91	2.606.456,71	9.839.805,54
Aragón	250.836,61	1,6935	424.791,80	785.829,47	642.484,48	307.696,90	40.302,65	37.077,53	1.096.822,65	3.334.005,48
Castilla-La Mancha	192.337,69	1,6935	325.723,88	726.380,54	743.276,17	466.040,61	66.023,66	61.128,67	2.350.688,33	4.739.261,87
Canarias	288.803,75	1,6935	489.089,15	745.111,83	0,00	59.316,62	0,00	0,00	3.183.699,32	4.477.216,92
Extremadura	99.358,72	1,6935	168.263,99	313.844,29	400.763,09	198.936,64	26.573,80	24.229,38	1.912.549,31	3.045.160,52
Illes Balears	154.023,77	1,5465	238.197,76	552.289,76	905.492,97	275.020,99	48.561,86	20.924,40	-211.684,47	1.828.803,27
Madrid	1.536.408,11	1,6935	2.601.907,13	5.441.659,78	3.418.872,98	954.653,08	536.459,35	94.905,75	-182.312,35	12.866.145,72
Castilla y León	376.996,87	1,6935	638.444,20	1.147.824,13	1.134.038,04	605.775,57	75.494,26	82.286,49	3.048.066,79	6.731.929,48
Total CC.AA.	**7.005.911,27**		**11.841.869,25**	**22.753.312,83**	**19.128.279,68**	**8.023.119,39**	**1.918.251,01**	**863.866,57**	**32.396.549,81**	**96.925.248,53**

(*) Fuente: AEAT

4. CONCLUSIONES

Este trabajo tiene como objetivo realizar un análisis normativo de la recepción del principio de solidaridad en el sistema de financiación de las Comunidades Autónomas vigente en la actualidad. La oportunidad es idónea también, en tanto el Gobierno español está en este momento negociando con las Autonomías un nuevo modelo y por cuanto el encaje del principio de solidaridad se plantea como una de las grandes dificultades a las que se enfrenta.

Retomando la información que hemos ofrecido hasta aquí, cabe señalar que la solidaridad sólo puede establecerse entre entes o sujetos que tengan una voluntad de integración entre sí que se manifieste en compartir un interés común. Siendo éste el presupuesto inicial no puede decirse, en el caso español, que sea un tema menor. Habida cuenta del auge de los nacionalismos y de la voluntad expresa de algunos territorios de separarse del Estado español, cabe preguntarse si existe ese interés común y esa voluntad de convivencia e una única estructura.

Lo que podemos extraer de la experiencia de estos más de treinta años de Autonomías es que la solidaridad se ha ido desvirtuando en su concepción y en su aplicación a la financiación autonómica. Ha sido interpretada como una "imposición estatal", asimilada, por tanto, a la acción interventora del poder central y limitadora de la autonomía. Cada vez que se plantea una nueva demanda de incremento de las competencias tributarias de las Comunidades, la solidaridad se convierte en el límite que se opone. En definitiva, la solidaridad se percibe como una amenaza para el derecho de Autonomía y en cada nueva negociación se plantea la restricción de su influencia en el sistema de financiación autonómico.

En el ámbito opuesto, las Comunidades perceptoras de los fondos de solidaridad procuran que su condición de menos ricas sirva para modular la financiación percibida. Esperan una contribución del conjunto de las Autonomías que les permita no sólo asegurar que van a percibir recursos suficientes para ejercer sus competencias sino recibir la financiación adicional necesaria para proveer un nivel de servicios similar al del conjunto.

Las expectativas de unas y otras están justificadas y amparadas por el principio de solidaridad, entendido correctamente. Pero, ¿por qué hemos llegado a esta situación? En nuestra

opinión la respuesta hay que buscarla en el predominio de la actuación redistributiva del Estado a través de las transferencias generales, condicionadas o no.

Si revisamos nuestro sistema de financiación, constatamos que sólo encontramos instrumentos de solidaridad obligatoria, es decir, aquélla en la que el Estado deriva fondos hacia los territorios que tienen una mayor necesidad, medida ésta en términos de renta per cápita. La potencialidad redistributiva del FCI –que es un recurso externo al sistema-, dada su dotación, no es alta y, sin embargo, los datos de gasto en inversión de las Comunidades perceptoras muestran que hay una gran dependencia de este Fondo y del FEDER. Puesto que exige la dotación correspondiente de la Autonomía para su percepción, no podemos aceptar que se trate de una subvención que desincentive el esfuerzo de la Comunidad perceptora y su mantenimiento y potenciación tendrá que ser correlativo al mantenimiento y crecimiento de las diferencias existentes entre Comunidades. Las Asignaciones de nivelación no se han puesto en marcha, tal vez porque el propio sistema contempla un mecanismo propio para la Sanidad pero, el incremento de la población que ha experimentado España en los últimos años hace que la financiación recibida sea insuficiente para mantener el nivel de prestación que se está ofreciendo. Las dificultades son mayores en las Comunidades que tienen un menor nivel de renta y que, por tanto, obtienen una financiación tributaria menor de ahí que sea previsible que éstas esperen recibir un mayor porcentaje de financiación vía transferencia corriente.

El sistema de financiación actual carece un mecanismo de solidaridad que corrija las diferencias en la financiación que perciben las Comunidades en función de su capacidad de generar recursos. Tenemos que buscar sus manifestaciones en el Fondo de Suficiencia y podemos encontrarla en el mecanismo de crecimiento del mismo, dado que crece con la media nacional de incremento de los tributos cedidos. Las Comunidades que experimentan un menor crecimiento se benefician del que otras han obtenido. Si las Autonomías de menor crecimiento coinciden con las menor renta per cápita –suele ser así, el mecanismo está favoreciendo a los de peor posición relativa. Pero tenemos que admitir que si una Comunidad rica experimenta un retroceso en su crecimiento económico se beneficiará también de este mecanismo, luego no podemos concluir que es un instrumento de solidaridad, en un sentido estricto. En nuestra opinión, el peso fundamental del Fondo de Suficiencia recae en dicho objetivo. Sólo el Fondo de Renta Relativa contribuye de manera efectiva a la solidaridad ya que vincula a la renta por habitante

El núcleo del problema de la solidaridad en el sistema actual está en el cálculo de las necesidades de gasto y así asistimos diariamente a propuestas de las diferentes Comunidades en las que tratan de dar más peso a la variable que más les beneficia. Tal comportamiento es esperable, en la medida en que cada una de ellas trata de hacer prevalecer su propio interés. No es realista exigir una actitud altruista de las Comunidades Autónomas pero sí se pueden disponer los medios adecuados para que la solidaridad actúe como elemento integrador de la voluntad particular –determinada por el beneficio propio- y el interés común que debe asegurarse. Y ese papel debe jugarlo el Estado, debe ser el ente que garantice que el resultado final de la negociación responda a los objetivos nacionales, entendiendo éstos como los comunes. En nuestra opinión, el modelo actual de negociación no puede garantizar este resultado. Se caracteriza por la bilateralidad y la influencia de factores políticos e institucionales, lo que explica la dificultad de llevarlo a cabo de manera satisfactoria y eficiente.

BIBLIOGRAFÍA

CAMPBELL, T. (1992): Siete Teorías de la Sociedad. Durkheim. Cátedra. Madrid.

DE LA FUENTE, A. y GUNDÍN, M. (2008): "El sistema de financiación de las Comunidades Autónomas de Régimen Común: un análisis crítico y algunas propuestas de reforma". En Colección de Estudios Económicos. 18-08. Serie Economía Regional. FEDEA

DUGUIT, L. (1924): Soberanía y Libertad. Librería Española y Extranjera.

LAGOS RODRÍGUEZ, MG (2002): La solidaridad interterritorial en el sistema de financiación de las Comunidades Autónomas. ProQuest Co. UMI Dissertations Publishing. Michigan, USA.

MINISTERIO DE ECONOMÍA Y HACEINDA (2008): Las Haciendas Autonómicas en cifras. 2006. Madrid.

RAWLS, J. (1985): Teoría de la Justicia. Fondo de Cultura Económica. México.

PRESTACIONES DE JUBILACIÓN EN LOS PLANES DE PENSIONES INTEGRADOS CON LA SEGURIDAD SOCIAL: UN ANÁLISIS SEGÚN LAS BASES DE COTIZACIÓN CONSIDERADAS EN EL CÁLCULO DE LA PENSIÓN

PELÁEZ FERMOSO, FRANCISCO JOSÉ
E-mail: ppelaez@eco.uva.es
GARCÍA GONZÁLEZ, ANA
E-mail: anagar@eco.uva.es
Departamento de Economía Aplicada
Universidad de Valladolid

Resumen

Este trabajo trata de analizar la incidencia que tiene la consideración de distintos escenarios para las Bases de Cotización en el cálculo de las prestaciones de jubilación cubiertas por un plan de pensiones de prestación definida integrado con el sistema público de la Seguridad Social. Para este objetivo de integración de ambos sistemas se pueden utilizar el método Offset y el método Excess. Este tipo de planes de pensiones que definen las prestaciones que garantizan teniendo en cuenta las concedidas por el sistema público de pensiones, tratan de subsanar la pérdida de poder adquisitivo que experimenta el trabajador cuando se retira o cesa en su actividad. Son, por tanto, un instrumento financiero bastante adecuado para actuar de forma eficaz ante este problema. Utilizaremos los denominados Ratios de Sustitución para comprobar el nivel de equidad que caracteriza al sistema de la Seguridad Social, siendo éste más reducido cuanto más elevadas sean las rentas salariales de los trabajadores.

Palabras Clave: Bases de cotización, Métodos de integración de planes de pensiones, Prestaciones de jubilación, Sistema de la Seguridad Social
Área Temática: Economía Social y Discapacidad.

Abstract

The aim of this paper is to analyze how the quotation bases used to compute retirement benefits may affect these pension benefits. We focus on a private defined benefit pension plan integrated with the Social Security system. Social Security integration is implemented using both the offset and the Excess methods. For an integrated pension plan, private benefits are calculated taking into account the benefits granted by the public pension system. Integrated plans try to correct for the loss of purchasing power that a worker experiences when he retires or stops in his activity and receives a public pension. This loss in purchasing power is stronger the higher the worker's wage, which makes Social Security integration an adequate financial tool to solve this problem. We will use the so called Replacement Rates to verify the equity level of the Social Security, and the employer-sponsored retirement pensions.

Key words: Quotation bases, Integration methods of pension plans; Retirement benefits, Social Security system.
Thematic Area: Social Economics.

1. INTRODUCCIÓN

El tema sobre el devenir del sistema público de pensiones en España es desde hace más de dos décadas uno de los debates más candentes y controvertidos que existen dentro del campo del Estado del Bienestar. El desafío que plantea el envejecimiento de la población y el consiguiente crecimiento del gasto, no compensado por los ingresos, hace que el Consejo Europeo celebrado en Barcelona (2002) elabore un documento en el que se insta a los Estados Miembros a que se acelere la reforma de los regímenes de pensiones, tanto para asegurar su sostenibilidad como para que puedan seguir cumpliendo sus objetivos sociales. Tras la última aprobación del Programa de Estabilidad Español (2005-2008), la Comisión Europea vuelve a demandar una reforma global del sistema público de pensiones dados los riesgos derivados del gran incremento del gasto que conllevan éstas en el largo plazo.

El problema de las pensiones es, por tanto, un problema técnico. Se trata de cómo organizar los recursos durante la vida laboral activa de cada trabajador con el fin de satisfacer las necesidades que éste demanda a partir del momento de su jubilación. Toda solución política al problema de las pensiones que no tenga en consideración la relación existente entre recursos y necesidades a lo largo del ciclo vital de cada trabajador, será una solución en falso. Como afirma Herce San Miguel (1995), el tema del ahorro a largo plazo se presenta como un problema de instrumentación segura y productiva de la coherencia necesaria entre recursos y necesidades del ciclo vital. En este contexto, las pensiones no son sino una de las muchas vías existentes para lograr dicha coherencia, aunque la instrumentación propiamente dicha corresponde a productos concretos. De este modo, la instrumentación del ahorro-pensión da lugar a la aparición y desarrollo de los denominados "planes y fondos de pensiones", que son productos financieros cada vez más regulados y cuya finalidad debe ser proporcionar seguridad a los partícipes a favor de quienes se crean. Del mismo modo cabe reseñar, que las aportaciones a los planes de pensiones de carácter privado son equivalentes a las contribuciones realizadas al sistema público de la Seguridad Social. Ambas representan un esfuerzo financiero importante de los trabajadores con vistas a la financiación de sus pensiones futuras.

En las últimas décadas principalmente se constata que los planes y fondos de pensiones están adquiriendo en España un protagonismo cada vez más importante como consecuencia de las repercusiones que tienen en el funcionamiento del sistema financiero y de la economía en general. Existen diferentes factores que desde su creación en 1987 han contribuido al creciente desarrollo de los planes y fondos de pensiones en España. Por un lado, la mayor concienciación que adquieren los individuos de la necesidad de complementar el sistema público de la Seguridad Social a raíz de la aprobación del Pacto de Toledo[1] y, por otro, las mejoras que se han ido introduciendo a través de las diferentes normativas tratando de aportar mayores garantías y flexibilidad por medio de la protección de los compromisos de las empresas con los trabajadores, de las medidas de control y las mejoras financiero-fiscales aplicadas a estos instrumentos financieros de carácter privado.

En este contexto, uno de los mayores problemas que se presenta al final de la vida laboral de cualquier trabajador es poder disponer de los recursos monetarios suficientes que le permitan mantener un nivel de vida semejante al que tenía hasta ese momento. Es un hecho fehaciente que en España la pensión de jubilación otorgada por el sistema público de la Seguridad Social suele ser menor que el salario alcanzado por el trabajador a su retiro, hecho tanto más acentuado cuanto más elevadas sean las rentas salariales percibidas por los trabajadores al final de su actividad laboral. Un instrumento financiero bastante adecuado que puede subsanar de forma eficaz este problema son los planes y fondos de pensiones que definen las prestaciones de jubilación que cubren en función de las prestaciones que concede el sistema público de la

[1] El Pacto de Toledo, aprobado el 6 de abril de 1995 por el Pleno del Congreso de los Diputados, Herce San Miguel (2000), surge como resultado del debate generado sobre la sostenibilidad y viabilidad del sistema público de pensiones. Dicho acuerdo recoge un conjunto de medidas mediante las cuales trata de reformar el sistema de la Seguridad Social y realiza varias recomendaciones entre las que se encuentra una relacionada con los planes de pensiones privados y los sistemas colectivos. Estas medidas han sido revisadas en diferentes ocasiones. La última de ellas, el 30 de septiembre de 2003, en la cual se aprueba el Informe de la Comisión no Permanente del Pacto de Toledo del Congreso de los Diputados y en la que se recomienda reforzar la previsión complementaria en España.

Seguridad Social. En consecuencia, los Gobiernos deberían potenciar cada vez más el desarrollo de este tipo de sistemas privados de pensiones.

De esta forma, se han ido desarrollando paulatinamente en algunos países los "planes y fondos de pensiones" integrados con la Seguridad Social, cuyas prestaciones económicas son calculadas teniendo en cuenta las otorgadas por el sistema público de pensiones. Su principal finalidad es proporcionar a los trabajadores ya jubilados un nivel de renta que, junto con el suministrado la Seguridad Social, asegure unos niveles de ingresos similares a los que éstos poseían antes de la jubilación.

Cada vez es más habitual y se hace más aconsejable implantar por parte de las empresas a favor de sus trabajadores un plan de pensiones de estas características, cuyas prestaciones estén definidas en función de las otorgadas por el sistema de la Seguridad Social y en el que la mayoría de las aportaciones al plan de pensiones las realice el propio empresario. Son los denominados planes de pensiones de prestación definida del Sistema de Empleo. Con esta forma de proceder se busca reducir el riesgo inherente que conlleva el sistema público de pensiones en lo que se refiere a la cobertura de las pensiones por él garantizadas.

Una de las principales características de los planes de pensiones integrados con el sistema de la Seguridad Social es que tratan de modificar el riesgo soportado tanto por el promotor del plan como por los trabajadores y el Gobierno en relación con las prestaciones cubiertas por ambos sistemas. Como exponen Merton, Bodie y Marcus (1987), desde esta perspectiva, estos planes de pensiones aseguran a sus partícipes frente a los cambios adversos en las prestaciones de jubilación otorgadas por el sistema público de pensiones. De esta forma y en función de las prestaciones concedidas a los mismos por este último sistema de previsión social, el promotor-empresario se compromete a garantizar prestaciones de carácter aleatorio a los trabajadores que se retiran. La voluntad por parte de los empresarios de contribuir a cubrir en parte las futuras prestaciones de jubilación que corresponden a los trabajadores activos, debería estar sustentada en el deseo de retener e incentivar a dichos empleados durante su vida laboral activa.

De esta forma, los planes de pensiones integrados actúan como una especie de seguro o garantía frente al riesgo que tiene el trabajador respecto de la percepción de un determinado nivel de renta a partir de su jubilación. La existencia de este riesgo es consecuencia de la incertidumbre que conlleva en la actualidad la futura viabilidad financiera de los sistemas públicos de pensiones, en general, y del sistema de la Seguridad Social en España, en particular. A su vez, la viabilidad del plan va a depender del nivel de rentas futuras que sea capaz de suministrar a sus beneficiarios y, como afirma Herce San Miguel (1989), de la estructura del ciclo vital de cada partícipe. En otras palabras, de que la relación entre las contribuciones realizadas y la duración de la vida pasiva de cada beneficiario del plan de pensiones se mantenga dentro de ciertos límites, de modo que la financiación de una pensión razonable no sea excesivamente gravosa. Dada la trascendencia de este hecho, el riesgo compartido de ambos sistemas que se integran puede ser también objeto de estudio desde una perspectiva política. Por tanto, la finalidad última de los planes de pensiones que se integran con el sistema público de pensiones, es conseguir una adecuada combinación de las prestaciones económicas que ambos sistemas de previsión social han de cubrir.

Otra de las metas que busca la integración de los planes de pensiones con el sistema público de la Seguridad Social es tratar de paliar, como argumentan Anderson (1976) y Chang *et al*. (1982), la discriminación que caracteriza al actual sistema público de pensiones respecto de las rentas obtenidas por los trabajadores con más alta cualificación profesional y salarios más elevados frente a aquéllos que obtienen rentas salariales más bajas. Debido a que el sistema público de pensiones en España proporciona, en términos relativos, menores prestaciones a los trabajadores con salarios más altos, la integración de ambos sistemas debería tratar de compensar esta inclinación y, por tanto, que los trabajadores con mayores salarios puedan percibir proporcionalmente prestaciones más altas que los trabajadores que menos ganen. Se busca, en definitiva, lograr que el sistema de la Seguridad Social sea lo más equitativo posible con todos sus afiliados.

Para determinar la cuantía de las prestaciones económicas que garantizan estos sistemas de previsión social, se utilizan diferentes procedimientos o métodos para su integración. Como especifican Schulz y Leavitt (1983), los métodos de integración más habituales usados para este

fin son el método Offset y el método Excess, que tienen en cuenta diferentes formas de cálculo para determinar las prestaciones de jubilación que desee garantizar el plan de pensiones. Por medio de estas fórmulas o expresiones se busca comparar las ganancias obtenidas al final de la vida laboral de cada uno de los partícipes con las prestaciones otorgadas a los mismos por el sistema de la Seguridad Social en el momento de su jubilación.

En este contexto, se puede definir el proceso de integración por un conjunto de técnicas de coordinación que permiten definir las prestaciones del plan de pensiones en función de las prestaciones cubiertas por el sistema de la Seguridad Social. En general, este proceso de integración entre planes de pensiones privados y sistemas públicos de pensiones es regulado teniendo en cuenta un conjunto de normas y disposiciones que a tal efecto tiene la opción de especificar cada país donde estos sistemas pueden desarrollarse, como se constata en Amoroso (1982), Sluscher (1998) y Bender (1999).

Además de esta sección, el trabajo se estructura de esta forma: en la sección segunda se definen los objetivos y la metodología a seguir en el desarrollo del trabajo; en la tercera sección se describen de forma breve los principales métodos utilizados en la integración de los planes de pensiones con el sistema de la Seguridad Social; en la sección cuarta se plantea un modelo representativo de integración de ambos sistemas de previsión social; en la quinta sección se lleva a cabo la simulación del modelo según los diferentes escenarios planteados y los resultados obtenidos; y en la sección sexta se describen las conclusiones obtenidas con la realización de este trabajo.

2. OBJETIVOS Y METODOLOGÍA

Con la realización de este trabajo tratamos de conseguir diferentes objetivos, entre los que caben señalar los siguientes:

- Descripción y análisis de un modelo que recoge el proceso de integración de un plan de pensiones de prestación definida con el sistema público de la Seguridad Social.
- Consideración de distintos escenarios a la hora de calcular la Base Reguladora de la pensión a efectos de la determinación de la prestación de jubilación cubierta por el plan de pensiones.
- A partir de estos escenarios, determinación de los Ratios de Sustitución Total y Parcial, que miden qué parte del salario final bruto de cada trabajador cubren las prestaciones totales (públicas y privadas) y las prestaciones de la Seguridad Social, respectivamente.
- Estudio de cómo se modifica el nivel de discriminación del sistema público de la Seguridad Social cuando se procede a integrar este sistema con un plan de pensiones de prestación definida, teniendo en cuenta los diferentes grupos ocupacionales a los que pertenece cada uno de los trabajadores afiliados al Régimen General de la Seguridad Social.

Para la consecución de estos objetivos, se va a aplicar la metodología utilizada por el método Offset para la integración y análisis conjunto de ambos sistemas de previsión.

3. MÉTODOS DE INTEGRACIÓN

Dos son los métodos más habitualmente utilizados en el proceso de integración de ambos sistemas de previsión social, el método Offset y el método Excess. Tanto los planes de pensiones de prestación definida como los planes de aportación definida pueden ser integrados. Los planes de prestación definida, que garantizan una determinada prestación en el momento de la jubilación, función del número de años de servicio así como del nivel de ganancias de cada partícipe, pueden ser integrados utilizando cualquiera de las dos modalidades anteriormente citadas. Los planes de aportación definida, que fijan las cuantías de las contribuciones que deben realizarse al plan hasta la jubilación de cada partícipe, en cuyo caso, la cuantía de la prestación que percibirá hasta ese momento va a depender de las aportaciones realizadas y debidamente capitalizadas, pueden ser integrados exclusivamente mediante el método Excess.

Siguiendo a Anderson (1976) y a McGill (1984), los planes de prestación definida pueden englobarse en dos grandes categorías: los planes "Final-pay", que relacionan las prestaciones del plan con las ganancias que corresponden a los trabajadores en el momento de su jubilación, o con un promedio de las ganancias consideradas para un determinado número de años previos a la jubilación; y los planes "Career-pay", que definen la prestación del plan en el momento de la jubilación del trabajador como las prestaciones acumuladas que se acreditan por año de servicio activo en la empresa promotora del plan. Por tanto, en este caso, la prestación de jubilación estará en función de las ganancias medias de todos los años de actividad laboral del empleado.

El método Offset. En los planes de pensiones que utilizan este método para su integración con el sistema público de la Seguridad Social, la prestación final garantizada por el plan se calcula detrayendo del salario bruto final a la edad de retiro o de la renta disponible de cada trabajador a esa edad un determinado porcentaje de la prestación de jubilación otorgada por el sistema de la Seguridad Social, que es constante, e independiente de su nivel salarial. El porcentaje de offset utilizado debe ser congelado en el momento de la jubilación, de forma que ningún incremento en las prestaciones de la Seguridad Social posterior al momento de su retiro pueda reducir la prestación del plan de pensiones. Este porcentaje se fija cualquiera que sea el nivel salarial de cada partícipe. Una reducción o porcentaje de Offset deseable sería el porcentaje de la prestación de jubilación cubierta por el sistema de la Seguridad Social equivalente a la cuantía financiada por el empresario a través de sus contribuciones al sistema público de pensiones durante la vida activa de cada trabajador.

En el caso de que se considere la renta disponible como referencia en el cálculo de las prestaciones del plan, en lugar del salario bruto final, como reseñan Chang *et al*. (1982) ésta debería ser el límite superior de la prestación total de la pensión de retiro, es decir, la suma de la prestación del plan más la prestación cubierta por el sistema de la Seguridad Social. Ello es debido a que la política social no fomenta un incremento en el nivel de vida de un trabajador después de su jubilación. Según esta consideración, la renta disponible vendría definida como las ganancias brutas finales reducidas por aquellos conceptos que ya no se aplicarán después de la jubilación, como las contribuciones a la Seguridad Social. Recordemos, llegados a este punto que, como acabamos de señalar, el método Offset se emplea sobre todo en planes "Final Pay".

Teniendo en cuenta este límite de la renta disponible y según sea el objetivo establecido para la misma, pueden considerarse dos casos: un objetivo de sustitución total o parcial de la renta disponible. Para la realización de este trabajo, se considera únicamente el "objetivo de sustitución total de la renta disponible" y, en este caso, ésta es definida como un porcentaje del salario bruto final a través de esta expresión: $B_r^p = p \cdot S_r - q \cdot B_r^s$.

El Método Excess. Los Planes de Pensiones que utilizan métodos de integración Excess se caracterizan principalmente porque establecen previamente un determinado Nivel de Integración para cada partícipe-trabajador con el fin de poder determinar la prestación de jubilación que garantiza el plan de pensiones. Para determinar la prestación de jubilación que cubre el plan a cada partícipe, estos planes deducen de las ganancias salariales finales o medias previas a la jubilación, una cuantía igual al denominado Nivel de Integración, que podrá ser fijado por el propio plan de pensiones. El Nivel de Integración puede permanecer fijo en una determinada cuantía y, en este caso, diremos que el plan de pensiones está estáticamente integrado. Por el contrario, los planes con Niveles de Integración que se ajustan a cambios de las Bases de Cotización máximas establecidas para cada año, se conocen como planes dinámicamente integrados.

En los métodos Excess utilizados para la integración de estos sistemas de previsión social se pueden utilizar diversas fórmulas para el cálculo de la prestación de jubilación de un plan de pensiones de prestación definida Excess "Final-Pay". Según este tipo de plan de pensiones, la prestación de jubilación se define como un porcentaje fijo, λ, de las ganancias finales o promedio de un cierto número de años previos al de la jubilación del trabajador por encima del Nivel de Integración establecido. La fórmula que recoge este cálculo es la siguiente: $B_r^p = \lambda \cdot (S_r - \overline{NI})$. El método Excess es, tal vez, más directo a la hora de pagar prestaciones a ciertos empleados con salarios más altos antes de la jubilación, ya que el Nivel de Integración, es decir, el nivel de ganancias a partir del cual se aplica el porcentaje de Integración

para calcular la prestación cubierta por el plan, es fijado por el empresario, con lo cual está en sus manos que todos o sólo unos pocos empleados reciban prestaciones sólo con elevar o reducir el Nivel de Integración.

4. EL MODELO

Para la realización de este trabajo de investigación se describe un modelo matemático que recoge la estructura de un plan de pensiones de prestación definida del Sistema de Empleo "Final-Pay" integrado con el sistema de la Seguridad Social, como describen McGill (1984) y Herce San Miguel (1989). Esta modalidad de plan de pensiones considera para propósitos de la integración el salario bruto final del trabajador a su retiro. Como sea que uno de los principales objetivos buscados con la realización de este trabajo es determinar y analizar el nivel de cobertura que tanto el plan de pensiones de prestación definida como el sistema de la Seguridad Social garantizan en forma de prestaciones de jubilación (pensiones) a los trabajadores que se retiran, se pasa a definir a continuación las variables y ratios que son relevantes para este análisis:

B_r^p *:* prestación de jubilación que garantiza a cada partícipe el plan de pensiones. En un plan de pensiones integrado va a depender de la prestación concedida por el sistema de la Seguridad Social a cada trabajador en el momento que se jubila y de su salario bruto final o de su renta disponible.

B_r^T : prestación de jubilación total para cada partícipe del plan de pensiones. Es la suma de la prestación otorgada por el sistema de la Seguridad Social y de la prestación cubierta por el plan de pensiones a partir de la edad de jubilación del trabajador. Viene definida por esta expresión: $B_r^T = B_r^s + B_r^p$.

B_r^s *:* prestación de jubilación de la Seguridad Social establecida para cada empleado a su retiro. Se define en función de los salarios computables para un cierto número de años previos a la jubilación, del número de años cotizados y de la edad efectiva de jubilación, teniendo en cuenta la normativa legal que esté establecida en cada momento.

S_r : salario bruto final que corresponde a cada trabajador en el momento de su retiro. En ocasiones, y a efectos de cálculo de la prestación de jubilación del plan de pensiones, se toma el salario promedio de un cierto número de años previos a la jubilación.

p : porcentaje que se aplica al salario final bruto del trabajador en el momento de su jubilación. Puede tomar distintos valores según cuál sea el compromiso que en términos de prestaciones de jubilación desee cubrir el plan a sus partícipes.

q : porcentaje de reducción aplicado a la prestación de jubilación otorgada por el sistema de la Seguridad Social a cada trabajador que se retira. Es el Offset[2] considerado en el proceso de integración y puede tomar diferentes valores.

R^s : porcentaje que sobre el salario final bruto representa la prestación de jubilación concedida a cada trabajador por dicho sistema. Representa el "Ratio de Sustitución Parcial de la Seguridad Social" y se define como: $R^s = B_r^s / S_r$.

R^T : porcentaje del salario final bruto sustituido por la prestación total de jubilación. Denota el "Ratio de Sustitución Total**"** y se define como: $R^T = B_r^T / S_r$.

Una vez descritas las variables y parámetros que caracterizan al modelo representativo de ambos sistemas de previsión social, se pasa a describir distintos escenarios posibles de

[2] El porcentaje de Offset ha ser fijado para permanecer invariable a partir de la edad de jubilación de cada trabajador, como afirman Schulz y Leavitt (1983), de modo que ningún incremento en las prestaciones de la Seguridad Social después de que el empleado se haya jubilado pueda reducir la prestación que éste perciba del plan en el futuro. Este porcentaje suele estar relacionado con el nivel de cotizaciones que el empresario-promotor del plan de pensiones haya aportado al sistema de la Seguridad Social durante la vida activa del trabajador.

realización, con el fin de poder simular el modelo y analizar las variables más relevantes del mismo para el horizonte temporal 1985-2019, es decir, teniendo en cuenta los datos considerados en el pasado y realizando una estimación a futuro de los principales parámetros que en él intervienen. En este contexto**,** se determinarán las prestaciones de jubilación cubiertas por el plan de pensiones teniendo en cuenta las prestaciones de jubilación otorgadas por el sistema público de la Seguridad Social a cada trabajador que se retira.

A efectos prácticos de simulación y análisis del modelo establecido, se considera un trabajador perteneciente a cada uno de los grupos ocupacionales considerados en la Clasificación Nacional de Ocupaciones[3] (CNO-94), a cada uno de los cuales se les ha asignado un valor numérico para mejor identificación, como se recoge a continuación en el Cuadro-I. Sin embargo, como reseña Gala Vallejo (1999), la Ley General de la Seguridad Social de 1994 contempla otra clasificación de los trabajadores por categorías profesionales según la consideración de diferentes Grupos de Cotización[4].

CUADRO-I

Clasificación Nacional de Ocupaciones (CNO-94)	Valores Numéricos
Dirección de las Administraciones Públicas y de empresas de 10 o mas asalariados	1
Profesiones asociadas a Titulaciones de 2º y 3er Ciclo Universitario y afines	2
Profesiones asociadas a una Titulación de 1er Ciclo Universitario y afines	3
Trabajadores cualificados de la metalurgia, la construcción de maquinaria y asimilados	4
Operadores de instalaciones industriales, de maquinaria fija; montadores y ensambladores	5
Técnicos y profesionales de apoyo	6
Conductores y operadores de maquinaria móvil	7
Trabajadores cualificados de la construcción, excepto los operadores de máquina	8
Trabajadores de los servicios de protección y seguridad	9
Empleados de tipo administrativo	10
Trabajadores cualificados de industrias de artes gráficas, textil y de la confección, de la elaboración de alimentos ebanistas, artesanos y otros asimilados	11
Dependientes de comercio y asimilados	12
Trabajadores de los servicios de restauración y de servicios profesionales	13
Peones de la construcción, industrias manufactureras y transportes	14
Trabajadores no cualificados en servicios (excepto transportes)	15

Fuente: Instituto Nacional de Estadística y Elaboración Propia

4.1. Hipótesis de trabajo

Para poder trabajar con este modelo, se describen a continuación un conjunto de hipótesis que es necesario establecer sobre los parámetros y variables que en él intervienen:

- Las prestaciones de jubilación cubiertas por el sistema público de la Seguridad Social son conocidas para cada trabajador retirado, ya que éstas se establecen en función de los salarios computables para un determinado número de años[5].

[3] La Clasificación Nacional de Ocupaciones (CNO-94) utilizada en la realización de este análisis es una adaptación española de la comunitaria ISCO-88 COM de tipo jerárquico en la que se contemplan grupos ocupacionales con un nivel máximo de agregación.

[4] El artículo 26.3 del Reglamento General sobre Cotización establece que la asimilación a las categorías profesionales se efectuará por la Tesorería General de la Seguridad Social (TGSS) previo informe de la Inspección de Trabajo y de la Seguridad Social. Estos Grupos de Cotización están descritos en la Ley General de la Seguridad Social (LGSS), de 20 de junio de 1994, y en el Real Decreto 2064/1995, de 22 de diciembre, por el que se aprueba el Reglamento General de Cotización, Liquidación y otros Derechos de la Seguridad Social**.**

[5] El artículo 163 de la Ley General de la Seguridad Social (LGSS), de 20 de junio de 1994, con las posteriores modificaciones introducidas en la Ley 24/1997, contiene las normas para determinar la cuantía de la pensión (prestación) de jubilación, en su modalidad contributiva, que el Régimen General de la Seguridad Social concede a cada trabajador que se retira.

- Se considera que todos los trabajadores pertenecientes al plan de pensiones han cotizado 35 años al Régimen General de la Seguridad Social y de forma ininterrumpida durante los 15 años previos a la jubilación y que en el año 2019 en el que se retiran de la actividad tienen 65 años. En consecuencia, les corresponde percibir el 100% de la Base Reguladora de la pensión que se utiliza para determinar la cuantía de la prestación de jubilación establecida por este sistema público de pensiones, siempre que ésta sea menor o igual que la cuantía máxima fijada anualmente para la pensión de jubilación en los Presupuestos Generales del Estado.
- Los salarios computables que se consideran en las Bases de Cotización para el cálculo de la Base Reguladora[6] de la pensión de jubilación establecida por el sistema de la Seguridad Social, en su modalidad contributiva, se determinan suponiendo conocidos los valores que van a tomar alguno de los parámetros más relevantes del modelo que intervienen en su cálculo. Para la determinación de los salarios que corresponden a cada año del período 1985-2019, objeto de estudio, se toman como referencia los salarios recogidos en la Encuesta de Estructura Salarial[7], y la tasa media de crecimiento salarial deducida de las series enlazadas de Índices de Costes Laborales (ICL) con la Encuesta de Salarios de la Industria y de los Servicios[8] del INE para el período 1981-2001 y, posteriormente, con la Encuesta Trimestral de Coste Laboral (ETCL).
- Los salarios computables en las Bases de Cotización para el período 2006-2019 se calculan capitalizando a la tasa media de crecimiento salarial previsto y para el período 2002-1985 actualizando a dicha tasa. Durante el intervalo 2006-2019 se supone que los salarios y del IPC crecen a un tanto anual acumulativo constante del 2%. Desde 1985, cada 7 años, se estima un crecimiento de los salarios de un 10%. La referencia son los salarios del año 2002.
- La Base Reguladora de la pensión de cada trabajador, según el año y clasificación de ocupaciones al que pertenece, se determina considerando que las Bases de Cotización máximas crecen a un tanto anual medio estimado desde el año 2006 y que no existen lagunas de cotización, es decir, períodos en los que no ha existido obligación de cotizar.
- El parámetro *p* considerado en el método Offset de integración para determinar el nivel de prestación cubierto por el plan de pensiones a cada partícipe que cesa en su actividad laboral toma el valor del 100%. El parámetro *q* (porcentaje de Offset aplicado) puede tomar distintos valores[9] de referencia: $q = 1$; $q = 0{,}8$; $q = 0{,}75$; $q = 0{,}5$; $q = 0{,}25$.

[6] La Base Reguladora de la pensión, en su modalidad contributiva, se tiene en cuenta para calcular la prestación de jubilación concedida por el sistema de la Seguridad Social a cada trabajador que cesa en su actividad laboral. Según se recoge en el artículo 162 del Texto Refundido de la Ley General de la Seguridad Social, aprobado por Real Decreto legislativo 1/1994, de 2 de junio, junto con las modificaciones introducidas por la Ley 24/1997, de 15 de julio, dicha Base Reguladora viene definida en la actualidad por la expresión $BR = \left(\sum_{j=1}^{24} BCCC_j + \sum_{j=25}^{180} BCCC_j \cdot \frac{IPC_{25}}{IPC_j} \right) / 210$, con BR la Base Reguladora de la pensión de jubilación; $BCCC_j$ la Base de Cotización para contingencias comunes del j-ésimo mes anterior a la jubilación e IPC el Índice de Precios al Consumo correspondiente al j-ésimo mes previo a la jubilación. Como se supone que todos los empleados han cotizado 35 años y se jubilan a los 65 años, la Base Reguladora de la pensión va a conincidir con la prestación de jubilación de la Seguridad Social, siempre que no exceda del tope máximo establecido para esta pensión en los Presupuestos Generales del Estado de cada año.

[7] La Encuesta de Estructura Salarial (EES), realizada por primera vez en el año 1995, es una operación estadística de periodicidad cuatrianual realizada en el marco de la Unión Europea con criterios comunes de metodología y contenidos, a fin de obtener unos resultados comparables sobre la estructura y distribución de los salarios entre los Estados Miembros. La Encuesta de Estructura Salarial 2002 y publicada por el INE en el 2005, investiga la distribución de las ganancias medias por trabajador y año según la Clasificación Nacional de Ocupaciones (CNO-94) realizada en función de distintas variables como el sexo, rama de actividad, antigüedad, tamaño de la empresa y ocupación. Esta última es la que se considera en la realización de este trabajo.

[8] Las series enlazadas de Índices de Costes Laborales (ICL) con la Encuesta de Salarios de la Industria y de los Servicios (ES) realizadas por el INE para el período 1981-2001 tienen como objetivo proporcionar información sobre los niveles y la evolución de las ganancias medias brutas de cada trabajador (antes de la deducción de impuestos y cotizaciones a la Seguridad Social). De los datos recogidos en las mismas obtenemos los tantos de crecimiento salarial para cada uno de los períodos que comprende este análisis. A partir del año 2001, la Encuesta de Salarios de la Industria y de los Servicios ha sido sustituida por la denominada Encuesta Trimestral de Coste Laboral (ETCL).

[9] Se elige el valor $q = 0{,}8$ porque es el porcentaje equivalente al peso que tienen las cotizaciones realizadas por el empresario al sistema de la Seguridad Social durante los años de actividad laboral del trabajador respecto del total realizadas a dicho sistema tanto por el empresario como por cada trabajador.

5. SIMULACIÓN DEL MODELO

Con el fin de poder analizar el efecto que la integración de ambos sistemas de previsión tiene sobre la prestación del plan de pensiones y sobre el nivel de discriminación que conlleva el sistema de la Seguridad social para los trabajadores con salarios más elevados, se procede a simular el modelo planteando distintos escenarios de posible ocurrencia. Esta técnica de los escenarios se utiliza habitualmente cuando en el momento en el que se realiza el estudio y debido a la consideración de un horizonte temporal generalmente amplio para el análisis, se prevé cierta incertidumbre en los valores que pueden tomar algunas de las variables que intervienen en el modelo o en las condiciones y normativa aplicada en el futuro. De este modo, se puede analizar cómo varían los resultados obtenidos de su simulación según el escenario considerado. En este contexto, se plantean tres posibles escenarios según el número de años que se consideren para las Bases de Cotización que se computan en el cálculo de la Base Reguladora de la pensión del sistema de la Seguridad Social.

Escenario I: En este escenario, a efectos de el cálculo de la Base Reguladora de la pensión de la Seguridad Social, además de las hipótesis previamente establecidas, se van a tener en cuenta los salarios computables de los últimos 15 años previos al retiro de cada trabajador, como se describe en la normativa que a este respecto establece la Ley General de la Seguridad Social para España y que ya ha sido reseñada.

TABLA-I

Base Reguladora de la Pensión (15 años)					
Prestación del Plan; p = 100%					
Grupo	*q=1*	*q=0,80*	*q=0,75*	*q=0,50*	*q=0,25*
1	64573	72550	74544	84515	94486
2	25831	33808	35802	45773	55744
3	9169	17146	19140	29111	39082
4	6132	13264	15048	23963	32879
5	6062	13113	14876	23690	32504
6	6055	13098	14859	23663	32467
7	5052	10929	12398	19744	27090
8	4991	10797	12248	19505	26762
9	4555	9853	11177	17799	24422
10	4184	9051	10268	16351	22435
11	4158	8994	10204	16249	22294
12	3956	8558	9709	15461	21214
13	3659	7915	8979	14299	19618
14	3459	7483	8489	13518	18547
15	3037	6569	7451	11866	16281

GRÁFICO-I

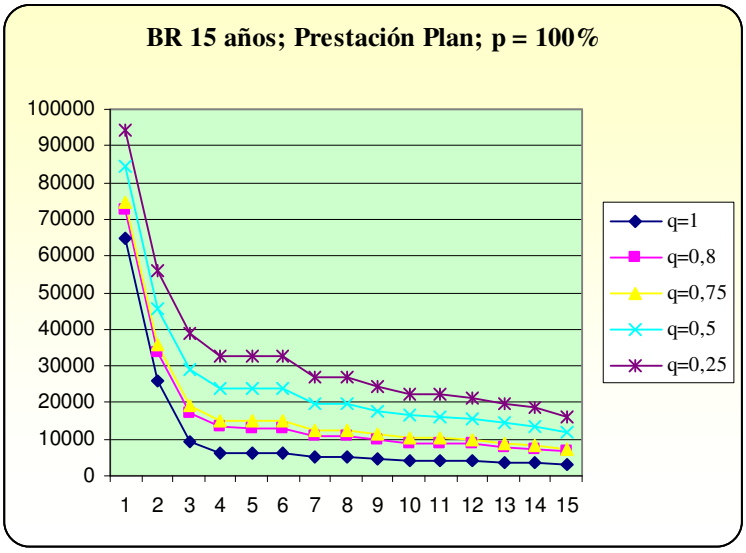
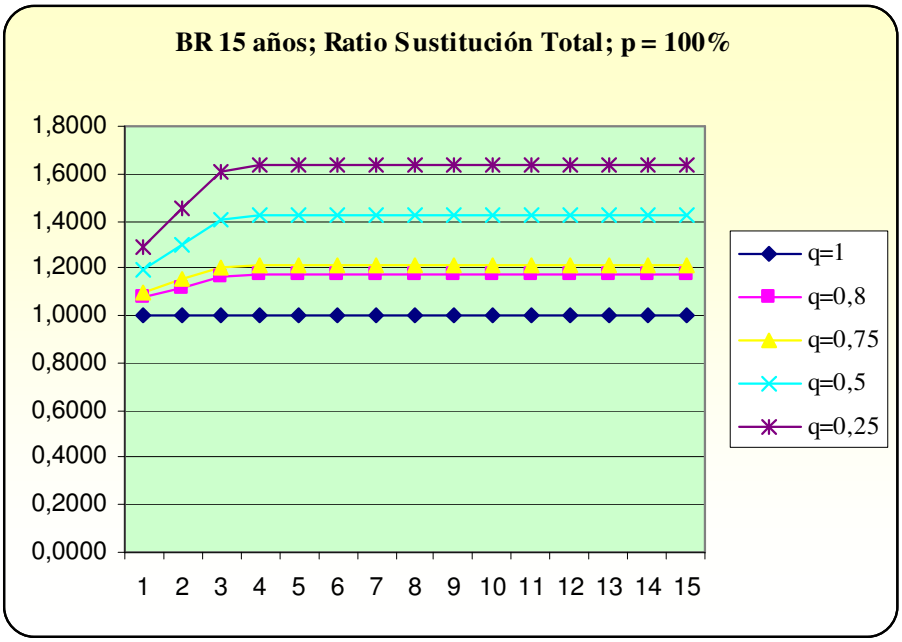


TABLA-II

Base Reguladora de la Pensión (15 años)					
Ratio de Sustitución Total; p = 100%					
Grupo	*q=1*	*q=0,80*	*q=0,75*	*q=0,50*	*q=0,25*
1	1,0000	1,0764	1,0955	1,1909	1,2864
2	1,0000	1,1214	1,1517	1,3035	1,4552
3	1,0000	1,1626	1,2033	1,4065	1,6098
4	1,0000	1,1707	1,2133	1,4266	1,6400
5	1,0000	1,1707	1,2133	1,4266	1,6400
6	1,0000	1,1707	1,2133	1,4266	1,6400
7	1,0000	1,1707	1,2133	1,4266	1,6400
8	1,0000	1,1707	1,2133	1,4266	1,6400
9	1,0000	1,1707	1,2133	1,4266	1,6400
10	1,0000	1,1707	1,2133	1,4266	1,6400
11	1,0000	1,1707	1,2133	1,4266	1,6400
12	1,0000	1,1707	1,2133	1,4266	1,6400
13	1,0000	1,1707	1,2133	1,4266	1,6400
14	1,0000	1,1707	1,2133	1,4266	1,6400
15	1,0000	1,1707	1,2133	1,4266	1,6400

GRÁFICO-II



Escenario II: En este Escenario-II, para realizar el cálculo de la Base Reguladora de la pensión de la Seguridad Social, se consideran los salarios computables (Bases de Cotización) de los últimos 8 años inmediatamente previos a la jubilación, del mismo modo que se procedió en su momento cuando se aplicaron las disposiciones legales establecidas en la Ley 26/1985, de 31 de julio, y hasta su posterior modificación.

TABLA-III

Base Reguladora de la Pensión (25 años) Prestación del Plan; p = 100%					
Grupo	*q=1*	*q=0,80*	*q=0,75*	*q=0,50*	*q=0,25*
1	64573	72550	74544	84515	94486
2	25831	33808	35802	45773	55744
3	9458	17377	19357	29256	39155
4	8058	14805	16492	24926	33360
5	7966	14637	16304	24642	32980
6	7957	14620	16286	24614	32942
7	6639	12199	13589	20538	27487
8	6559	12051	13424	20289	27154
9	5985	10997	12250	18515	24779
10	5498	10102	11253	17008	22763
11	5464	10039	11183	16902	22621
12	5199	9553	10641	16083	21524
13	4808	8834	9841	14873	19906
14	4546	8352	9304	14061	18819
15	3990	7332	8167	12343	16520

GRÁFICO-III

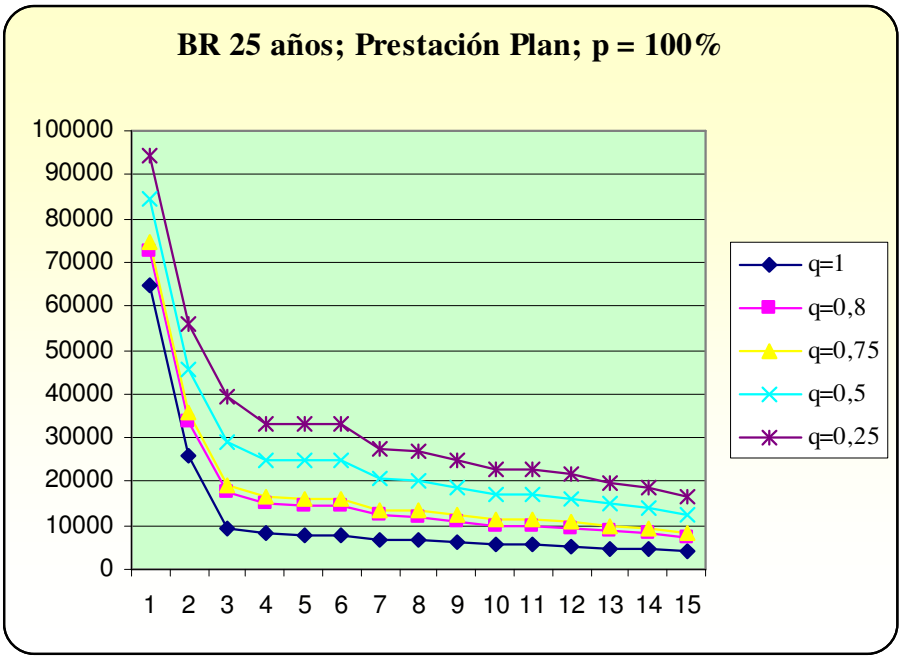
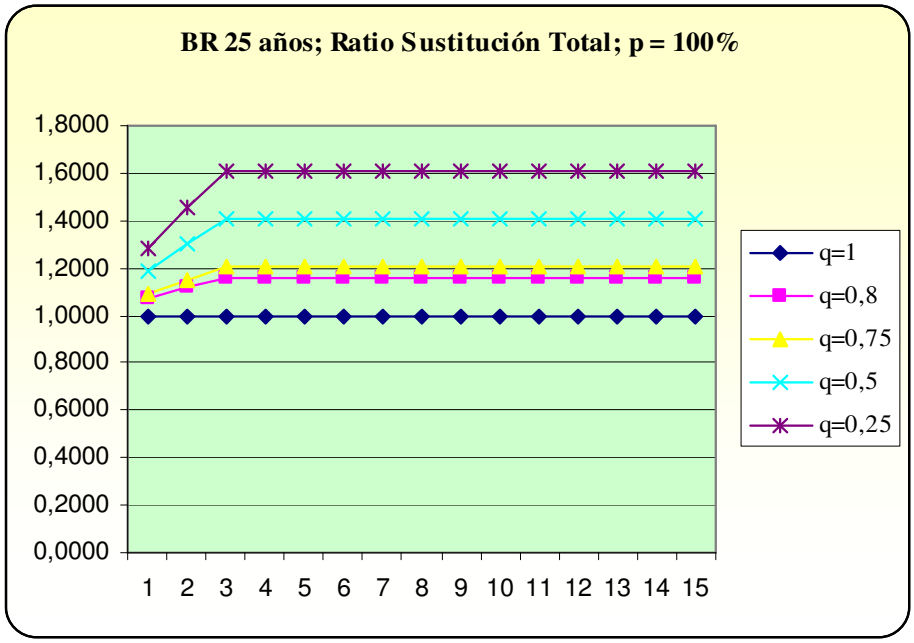


TABLA-IV

Base Reguladora de la Pensión (25 años)					
Ratio de Sustitución Total; p = 100%					
Grupo	*q=1*	*q=0,80*	*q=0,75*	*q=0,50*	*q=0,25*
1	1,0000	1,0764	1,0955	1,1909	1,2864
2	1,0000	1,1214	1,1517	1,3035	1,4552
3	1,0000	1,1614	1,2018	1,4036	1,6054
4	1,0000	1,1614	1,2018	1,4036	1,6054
5	1,0000	1,1614	1,2018	1,4036	1,6054
6	1,0000	1,1614	1,2018	1,4036	1,6054
7	1,0000	1,1614	1,2018	1,4036	1,6054
8	1,0000	1,1614	1,2018	1,4036	1,6054
9	1,0000	1,1614	1,2018	1,4036	1,6054
10	1,0000	1,1614	1,2018	1,4036	1,6054
11	1,0000	1,1614	1,2018	1,4036	1,6054
12	1,0000	1,1614	1,2018	1,4036	1,6054
13	1,0000	1,1614	1,2018	1,4036	1,6054
14	1,0000	1,1614	1,2018	1,4036	1,6054
15	1,0000	1,1614	1,2018	1,4036	1,6054

GRÁFICO-IV



Escenario III: Para el cálculo de la Base Reguladora de la pensión de la Seguridad Social, se tienen en cuenta los salarios computables de los últimos 35 años, es decir, de toda la carrera profesional del trabajador. Este es el escenario más probable en el futuro como consecuencia, de una parte, de un mayor incremento del gasto público ocasionado por el progresivo envejecimiento y longevidad de la población, a la vez que por un descenso de la población activa, que es la que financia dicho sistema, y de otra, por un descenso de la oferta de empleo en el mercado de trabajo y, por tanto, por una reducción de los ingresos que en forma de cotizaciones al sistema se generarían.

TABLA-V

Base Reguladora de la Pensión (35 años)					
Prestación del Plan; p = 100%					
Grupo	*q=1*	*q=0,80*	*q=0,75*	*q=0,50*	*q=0,25*
1	64573	72550	74544	84515	94486
2	25831	33808	35802	45773	55744
3	12122	19508	21355	30588	39821
4	10328	16621	18194	26061	33927
5	10210	16432	17987	25764	33541
6	10198	16413	17967	25735	33503
7	8509	13695	14991	21473	27954
8	8406	13529	14809	21212	27615
9	7671	12346	13514	19358	25201
10	7047	11341	12415	17783	23150
11	7003	11270	12337	17672	23006
12	6664	10724	11739	16815	21890
13	6163	9918	10856	15550	20244
14	5826	9376	10264	14701	19139
15	5114	8231	9010	12905	16801

GRÁFICO-V

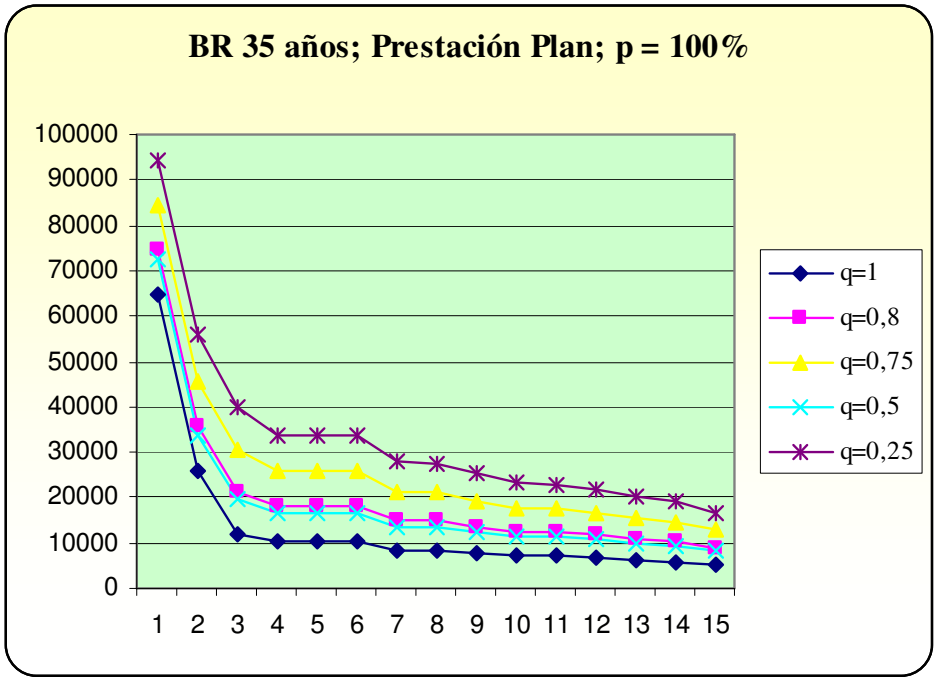
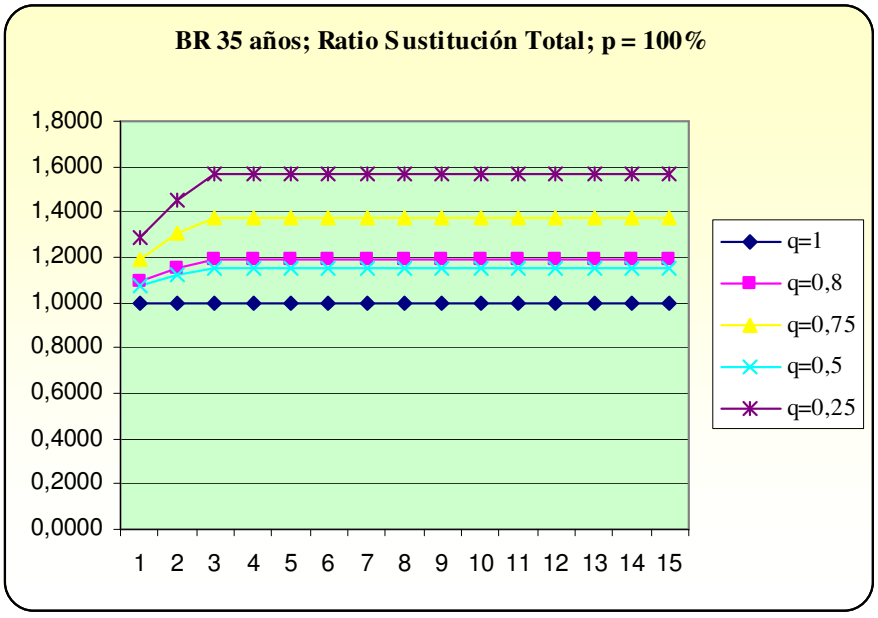


TABLA-VI

Base Reguladora de la Pensión (35 años)					
Ratio de Sustitución Total; p = 100%					
Grupo	*q=1*	*q=0,80*	*q=0,75*	*q=0,50*	*q=0,25*
1	1,00000	1,0764	1,0955	1,1909	1,2864
2	1,00000	1,1214	1,1517	1,3035	1,4552
3	1,00000	1,1506	1,1882	1,3764	1,5647
4	1,00000	1,1506	1,1882	1,3764	1,5647
5	1,00000	1,1506	1,1882	1,3764	1,5647
6	1,00000	1,1506	1,1882	1,3764	1,5647
7	1,00000	1,1506	1,1882	1,3764	1,5647
8	1,00000	1,1506	1,1882	1,3764	1,5647
9	1,00000	1,1506	1,1882	1,3764	1,5647
10	1,00000	1,1506	1,1882	1,3764	1,5647
11	1,00000	1,1506	1,1882	1,3764	1,5647
12	1,00000	1,1506	1,1882	1,3764	1,5647
13	1,00000	1,1506	1,1882	1,3764	1,5647
14	1,00000	1,1506	1,1882	1,3764	1,5647
15	1,00000	1,1506	1,1882	1,3764	1,5647

GRÁFICO-VI



Se comprueba en las Tablas y Gráficos I, III y V que en los tres escenarios, para un valor de *p* establecido, a medida que decrece el parámetro *q* crece la prestación del plan de pensiones, debido a que en su cálculo se detrae de la prestación original cada vez una parte más pequeña.

En las Tablas y Gráficos II, IV y VI, se constata que el Ratio de Sustitución Total crece a medida que disminuye el porcentaje de Offset aplicado. Sin embargo, para los tres primeros grupos ocupacionales este crecimiento es menor que para el resto, que crece en la misma medida para todos los grupos ocupacionales considerados.

Del mismo modo, también se constata que a medida que crece el número de años considerados de los salarios computados en el cálculo de la Base Reguladora de la pensión, la prestación del plan también crece, como consecuencia, tal y como se aprecia en la Tabla VII, de

una disminución de la cuantía de la pensión de jubilación correspondiente al sistema de la Seguridad Social, y por tanto, una mayor necesidad de compensación.

Así, en la Tabla VII se comprueba también que el Ratio de Sustitución Parcial de la Seguridad Social disminuye a medida que se computan más años en el cálculo de la Base Reguladora de la pensión, salvo para los dos primeros grupos (por el efecto del valor de la prestación máxima de jubilación aplicable a ambos, ya que la cuantía de la Base Reguladora sobrepasa dicho valor máximo).

A través de los resultados obtenidos de la simulación del modelo se constata que con la integración de ambos sistemas, el nivel de discriminación de los trabajadores con niveles salariales más altos con respecto a los trabajadores con menores salarios, disminuye, cualquiera que sea el nivel de integración considerado.

Asimismo, se puede comprobar que a medida que crece el número de años considerados en el cálculo de la Base Reguladora de la pensión, el Ratio de Sustitución Total disminuye (excepto para los dos primeros grupos de cotización), debido a que el efecto de la disminución en la prestación de jubilación de la Seguridad Social es más fuerte que el efecto de crecimiento de la prestación del plan privado de pensiones.

TABLA-VII

Clasificación Nacional por Grupos de Ocupación (CON-94)	Salario inicial Bruto (1985)	Salario Final Bruto, S_r (2019)	Prestación Seg. Social, B_r^s (15)	Prestación Seg. Social, B_r^s (25)	Prestación Seg. Social, B_r^s (35)	R^s (15)	R^s (25)	R^s (35)
Dirección de las Administraciones Públicas y de empresas de 10 o mas asalariados	18749	104458	39885	39885	39885	0,3818	0,3818	0,3818
Profesiones asociadas a Titulaciones de 2º y 3er Ciclo Universitario y afines	11795	65716	39885	39885	39885	0,6069	0,6069	0,6069
Profesiones asociadas a una Titulación de 1er Ciclo Universitario y afines	8805	49054	39885	39596	36932	0,8131	0,8072	0,7529
Trabajadores cualificados de la metalurgia, la construcción de maquinaria y asimilados	7502	41794	35662	33736	31466	0,8533	0,8072	0,7529
Operadores de instalaciones industriales, de maquinaria fija; montadores y ensambladores	7416	41318	35255	33351	31108	0,8533	0,8072	0,7529
Técnicos y profesionales de apoyo	7408	41271	35216	33314	31072	0,8533	0,8072	0,7529
Conductores y operadores de maquinaria móvil	6181	34436	29383	27796	25926	0,8533	0,8072	0,7529
Trabajadores cualificados de la construcción, excepto los operadores de máquina	6106	34018	29027	27459	25612	0,8533	0,8072	0,7529
Trabajadores de los servicios de protección y seguridad	5572	31044	26489	25058	23373	0,8533	0,8072	0,7529
Empleados de tipo administrativo	5119	28518	24334	23020	21471	0,8533	0,8072	0,7529
Trabajadores cualificados de industrias de artes gráficas, textil y de la confección, de la elaboración de alimentos ebanistas, artesanos y otros asimilados	5087	28340	24182	22876	21337	0,8533	0,8072	0,7529
Dependientes de comercio y asimilados	4840	26966	23010	21767	20302	0,8533	0,8072	0,7529
Trabajadores de los servicios de restauración y de servicios profesionales	4476	24938	21279	20130	18776	0,8533	0,8072	0,7529
Peones de la construcción, industrias manufactureras y transportes	4232	23577	20118	19031	17751	0,8533	0,8072	0,7529
Trabajadores no cualificados en servicios (excepto transportes)	3715	20696	17660	16706	15582	0,8533	0,8072	0,7529

FUENTE: Instituto Nacional de Estadística y Elaboración Propia

6. CONCLUSIONES

De los resultados obtenidos de la simulación del modelo cuando se considera el método Offset para la integración de ambos sistemas de previsión social, se comprueba para los tres escenarios considerados, que se reduce el grado de discriminación que el sistema de la Seguridad Social ejerce sobre los trabajadores con niveles salariales más elevados, sea cual fuere el nivel de compensación u Offset aplicado en la integración y el número de años considerados para determinar la Base Reguladora de la pensión. Ello es debido a que el Ratio de Sustitución Total crece para todos los niveles salariales de los trabajadores y, además, es mayor cuantos menos años se tengan en cuenta para computar las bases de Cotización que se consideran en el cálculo de la Base Reguladora de la pensión. En consecuencia, se reduce el riesgo que conlleva el pago real de las prestaciones de jubilación garantizadas por la Seguridad Social, lo que implica que el objetivo de equidad buscado con la integración de ambos sistemas de previsión queda satisfecho en todos los casos analizados.

BIBLIOGRAFÍA

AMOROSO, V. (1982), "IRS Limitations on Integration of Defined Benefit Pension Plans with United States Social Security", *Transactions of Society of Actuaries*, 34, pp. 195-213.

ANDERSON, A.W. (1976), "Social Security Integration", *Transactions of Society of Actuaries*, 28, pp. 287-325.

ANUARIO DE ESTADÍSTICAS LABORALES Y DE ASUNTOS SOCIALES (1983-2002), "Bases de Cotización al Régimen General de la Seguridad Social", *Ministerio de Trabajo y Asuntos Sociales*, Madrid.

BENDER, K. A. (1999), "Characteristics of individuals with integrated pensions", *Social Security Bulletin*, 62, 3, pp. 28-40.

CARPIO GARCÍA, M. (2002), "El sistema de Seguridad Social Español: horizonte económico", en "La Seguridad Social: los retos del siglo XXI", UIMP, Santander.

ESTADÍSTICAS LABORALES Y DE ASUNTOS SOCIALES (1983-2002), "Bases de Cotización al Régimen General de la Seguridad Social", *Ministerio de Trabajo y Asuntos Sociales*, Madrid.

CHANG, Y. - FELDTMOSE, J. N. - FURNISH, J. - GULOTTA, M. J. - HODES, D. M. - LHAMON, F. T. - MARGEL, L. N. - MITCHELL, K. - NABLER, S. G. - ROHLFS, A. F. - SANNING, D. E. - SCHNITZER, R. J. (1982), "Integration of Private Pension Plans with Social Security", *Transactions of Society of Actuaries,* 34, pp. 247-276.

GALA VALLEJO, C. (1999), "El Sistema de la Seguridad Social en España", Ministerio de Trabajo y Asuntos Sociales*, Colección Textos Legales*, 31-C, Madrid.

HERCE SAN MIGUEL, J. A. (1989), "Planes y Fondos de Pensiones", Papeles de Economía Española*, Suplementos sobre el Sistema Financiero*, 26, pp. 7-50.

HERCE SAN MIGUEL, J. A. (1995), "Las pensiones públicas en España: perspectivas y posibilidades de reforma", Documentos de Trabajo, *FEDEA*, pp. 95-103, Madrid.

HERCE SAN MIGUEL, J. A. Y ALONSO, J. (2000), "La Reforma de las Pensiones ante la Revisión del Pacto de Toledo", Colección de Estudios Económicos, 19, *FEDEA y Caja de Ahorros y Pensiones de Barcelona*, Barcelona.

INE (1981-2001), Series Enlazadas de Índices de Costes Laborales con la Encuesta de Salarios en la Industria y Servicios, INEbase, www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft22/p131&file=inebase

INE (2001-2008), Encuesta Trimestral de Coste Laboral (ETCL), INEbase. www.ine.es/jaxi/menu.do?type=pcaxis&path=/t22/p187/&file=inebase.

INE (2005), "Encuesta de Estructura Salarial (EES) 2002", Base de datos INEbase, www.ine.es/inebase/cgi/axi.

INE (2005), "Índices de Precios de Consumo (IPC)", Datos históricos bases 2001, 1992 y 1983, INEbase, www.ine.es/jaxiBD/tabla.do.

MCGILL, D. M. (1984), "Fundamentals of Private Pensions" University of Pennsylvania. *Richard D. Irwi,*. Illinois.

MERTON, R. - BODIE, Z. - MARCUS, A.J. (1987), "Pension Plan Integration as Insurance against Social Security Risk", in Bodie, Z., Shoven, J.B. y Wise, D.A. (Eds), *Issues in Pension Economics* (NBER), The University of Chicago Press, pp. 147-172.

MINISTERIO DE TRABAJO Y SEGURIDAD SOCIAL (1995), "Ley General de la Seguridad Social", Instituto Nacional de la Seguridad Social, *Colección Textos Legales*, 54, Madrid.

MINISTERIO DE TRABAJO Y SEGURIDAD SOCIAL (2000-2007), "Bases de Cotización al Régimen General de la Seguridad Social", www.seg-social.es/inicio.

SCHULZ, J. H. - LEAVITT, T. D. (1983), "Pension Integration: Concepts, Issues and Proposals", *Employee Benefit Research Institute (*EBRI), Washington.

SLUSCHER, CH. (1998), "Pension Integration and Social Security Reform", *Social Security Bulletin,* 61, 3, pp. 20-27.

impacto de la política pública SOBRE LA DEMANDA CULTURAL: EL CASO DE CUATRO MUSEOS EN COLOMBIA

GIOVANNI ANDRÉS HERNÁNDEZ SALAZAR
E-mail: giovanni.hernandez@unisabana.edu.co
ANA MARÍA OLAYA PARDO
E-mail: ana.olaya@unisabana.edu.co
Universidad de la Sabana (Colombia)

Resumen

Para incentivar la demanda a los museos, el gobierno colombiano creo la ley 580, la cual estableció la entrada gratuita a los museos entre el 15 de julio y 15 de agosto. La aplicación de esta ley produjo un incremento en el número de visitantes para ese período, repercutiendo negativamente en los ingresos por boletería. Para resolver esta situación, se aprobó la ley 929 que busca establecer la entrada gratuita a los museos el último domingo de cada mes. Este trabajo investiga el impacto de la ley 929 sobre los ingresos por boletería y visitantes a 4 museos de Bogotá. Se construyó una base de datos de visitantes e ingresos para el periodo 2001-2008 y se estimaron modelos univariados ARIMA de intervención. Se concluye que la ley tuvo un impacto significativo mejorando los ingresos por boletería de todos los museos, pero no modificó el comportamiento temporal de los visitantes.

Palabras Clave: Política Pública, ARIMA, Intervención, Museos, Demanda.
Área Temática: Economía Social y Discapacidad.

Abstract

To stimulate the demand to the museums, the Colombian government promulgated the law 580, which established the free admission to the museums between July 15 and August 15. The effect of this law was an increase in the number of visitors for this period, impacting negatively in the museums' income. To solve this situation, the law 929 was approved and established the free admission to the museums only the last sunday of every month. This work investigates the impact of the law 929 on the museums' income and visitors to 4 museums of Bogota. A visitors' database and income were constructed for the period 2001-2008 and ARIMA models of intervention were estimated. The main result is the law had a significant impact improving the museums' income, but did not modify the temporary behavior of the total visitors.

Thematic Area: Social Economics.

INTRODUCCIÓN

Los museos son instituciones culturales que han evolucionado no solo en su infraestructura física y recurso humano, sino también en su concepción, misión y visión institucional, con el objetivo de acoplarse a los acontecimientos sociales de las épocas en las que ellos han estado presentes. En la antigüedad los museos eran considerados lugares donde se albergaban colecciones que representaban la identidad cultural de un pueblo. En principio, los museos eran concebidos como lugares destinados a la admiración y contemplación de sus colecciones.

Actualmente, la definición de la institución museo por parte del Consejo Internacional de Museos (ICOM, por su sigla en ingles) se refiere a una entidad "permanente, sin fines lucrativos, al servicio de la sociedad y su desarrollo, abierta al público, que efectúa investigaciones sobre los testimonios materiales del ser humano y de su medio ambiente, los cuales adquiere, conserva, comunica y exhibe, con propósitos de estudio, educación y deleite".

Los museos colombianos no han sido ajenos a esta evolución y con el ánimo de cumplir cabalmente las tareas que su definición y la sociedad les impone, el Gobierno colombiano a través del Ministerio de Cultura ha desarrollado estrategias de fortalecimiento no solo en infraestructura, recurso humano e ingresos, sino en el marco legal donde se identifican principalmente la promulgación de las leyes 580 de 2000 y 929 de 2004[1].

La ley 580 de 2000 buscaba exaltar los valores, símbolos patrios, manifestaciones autóctonas culturales de Colombia, para lo cual estableció dentro de muchas otras disposiciones, el lapso comprendido entre el 15 de julio y el 15 de agosto de cada año como el mes de la patria, período en el cual los museos, monumentos nacionales y centros culturales del estado, permitían el ingreso gratuito a sus instalaciones (artículo 9°, parágrafo 1).

Como consecuencia de la promulgación de esta ley los museos Quinta de Bolívar, Santa Clara, Arte Colonial y 20 de Julio[2] percibieron un incremento significativo de usuarios durante el mes de la patria, aumento que en promedio superaba en 2.5 veces el nivel medio de visitas en los otros meses del año. Esta situación exacerbo el riesgo para las colecciones, dificultó el servicio eficiente y eficaz de los guías académicos durante las visitas guiadas y, disminuyó los ingresos por venta de boletería.

Por tal motivo, el Gobierno Colombiano intentando suavizar los efectos anteriores, decide promulgar la ley 929 de 2004, por la cual se modifica el parágrafo del artículo 9° de la ley 580 del 2000, e indica que los museos, monumentos nacionales y centros culturales, permitirán el ingreso gratuito a sus instalaciones el último domingo de cada mes, así como el 20 de julio y el 7 de agosto de cada año.

Este documento tiene como objetivo evaluar y estimar el impacto de la ley 929 de 2004 sobre el comportamiento de visitantes a los museos Quinta de Bolívar, Santa Clara, Arte Colonial y 20 de Julio, con el fin de determinar su eficacia en el fortalecimiento de los ingresos por la venta de boletería y la disminución de la congestión en el mes de julio. El documento se desarrolla en cinco partes: la primera realiza una revisión de literatura sobre temas museístico, la segunda describe la base de datos donde se consigna el comportamiento temporal de los visitantes a los museos así como los ingresos por boletería, la tercera describe paso a paso la metodología de modelos ARIMA de intervención aplicada al desarrollo de este documento, la cuarta estima los modelos y resultados y, la quinta presenta las conclusiones y recomendaciones.

1 ESTADO DEL ARTE

Con el desarrollo de los museos públicos en Europa en el siglo XIX, surgen los primeros trabajos teóricos en donde se establecen los principios de la museología, considerada como una disciplina científica que estudia el funcionamiento de los museos, su historia, su papel en la sociedad, los sistemas de organización, conservación y exhibición de las obras artísticas, las relaciones entre el medio físico y las obras[3]. Durante los últimos años esta disciplina se ha venido desarrollando en el ambiente académico y ha sido objeto de estudio de investigadores de diversas disciplinas que han desarrollado metodologías y teorías sobre los mismos.

En los años 70´s surge una nueva rama de estudio denominada “economía de los museos”, la cual se basa en la aplicación de la teoría microeconómica al estudio de variables museísticas tales como: estructura de costos, determinación de precios, economías de escala, fallas del mercado, demanda y oferta, intervención estatal, teoría del principal agente, valor económico, entre otros. Entre los trabajos pioneros en esta línea de investigación se encuentran el de Peacock y Godfrey (1974) y Robbins (1971). El primero demuestra que la teoría y los

[1] Aunque el marco legal en el sector cultural colombiano es amplio, solo se hace referencia a estas dos leyes dado que son fundamentales para el desarrollo de esta investigación.
[2] Esta investigación solo se focalizara en el análisis de estos museos, lo cual no indica que no hayan existido los mismos resultados en otros museos del país.
[3] ICOM, Comité Internacional de Museos, creado en la UNESCO en 1946.

instrumentos que utiliza la economía pueden ser utilizados para el estudio de los museos y, el segundo analiza los beneficios externos que le generan los museos a la sociedad.

El interés académico por el análisis de los museos se debe fundamentalmente a las siguientes razones: a) desde la perspectiva económica, el museo es entendido como una firma que transforma un conjunto de inputs en outputs a través de diversos procesos, lo cual permite que el análisis económico y cuantitativo sea factible en su estudio; b) el debate en torno al impacto económico y social del museo es amplio, va desde la recolección y conserva de patrimonio histórico, hasta la investigación, generación de conocimiento, educación, entretenimiento y diversión, turismo, crecimiento económico, entre otros; c) los museos presentan una estructura institucional variada, resultado del tipo de propiedad pública, privada o mixta y pueden caracterizarse por ser sociedades no lucrativas, lo que resulta interesante para la investigación; d) para Barrio (2007), durante los últimos años, las políticas culturales han tomado gran importancia dentro de los planes de desarrollo económico de los países, ya que la actividad turística es vista como un sector potencial para el crecimiento económico a través de la generación de ingresos y la rehabilitación urbana y, e) la relación entre la cultura y el bienestar individual es un tópico de estudio relevante en la búsqueda del bienestar social.

Un amplio debate se ha desarrollado en torno a la pertinencia del museo público. Algunos autores exponen que el Estado debe proveer los servicios museísticos y velar por la eficiencia y equidad para contribuir al bienestar social, ya que el acceso a la cultura es un derecho de los individuos. Así, las políticas, estrategias y gestión deben provenir del Estado. Otros argumentos para justificar la intervención estatal son basados en las fallas del mercado: en el caso de los museos, los que soportan los costos de la provisión de los servicios museísticos no son los mismos que reciben los beneficios, por lo tanto, sin intervención estatal la producción sería inferior a la óptima socialmente requerida (Frey y Meier 2002).

Uno de los principales problemas objeto de investigación es el financiamiento. El debate teórico se ha centrado en el cuestionamiento acerca de si se debe cobrar la entrada a los museos y por lo tanto, ser una fuente de ingresos adicional a los subsidios otorgados por el Estado. Se encuentran posiciones contrarias con relación a la entrada libre a los museos, donde algunos afirman que el establecimiento del precio reduce el número de visitantes, mientras que otros, demuestran la inelasticidad de la demanda en los museos frente al precio (O´Hagan, 1995; Anderson, 1998; Darnell, 1998; Heilbrun y Gray, 2001; Luksetich y Partridge, 1997; Been et al., 2002). Por otra parte, Robins (1971), señala que las entradas gratis a museos se establecen fundamentalmente como estrategia de promoción para permitir el acceso de personas de bajos ingresos. Otros argumentan que los museos tienen un efecto positivo sobre la población en términos de educación, por lo que deben ser financiados sólo con impuestos (Frey y Meier 2003).

Otro argumento a favor de la gratuidad proviene de la teoría de los costos marginales la cual expone que el costo marginal es cero. De esta manera, algunos autores argumentan que si el costo marginal es cero, el punto óptimo es aquel donde el precio es igual a cero, generando bienestar derivado de la mayor equidad y eficiencia. Sin embargo, Jackson (1988), realiza una estimación de costos en museos y encuentra que el costo marginal es diferente de cero. Así mismo, Maddison y Foster (2003) exponen que en el caso de los museos en los que se presentan altas congestiones, los costos marginales son crecientes y por lo tanto, las tarifas deben ser diferenciales cuando hay eventos especiales, de manera que contribuyan a mejorar los ingresos del museo. Por su parte, Bailey y Falconer (1998) argumentan que hay razones de eficiencia que justifican el precio diferente de cero, siempre y cuando otros sectores de la economía mantengan un precio diferente al costo marginal.

Feldstein (1991), Di Maggio (1991), Dickenson (1992), consideran que la financiación de los museos a través de los impuestos puede ser una política regresiva generando problemas de inequidad en el mercado. Estudios empíricos han demostrado que los visitantes de los museos son individuos con ingresos medios y altos que cuentan con un buen nivel de educación y, los precios de los museos generalmente son muy bajos con relación a los costos, los cuales son financiados con impuestos.

El establecimiento del precio ha sido avalado por algunos autores al ser visto como una fuente adicional de financiamiento para los museos (Creigh-Tyte y Selwood, 1998); aunque se encuentran afirmaciones acerca de que éste reduce el gasto de los visitantes en otras áreas de consumo, Steiner (1997) encuentra que el incremento de este tipo de gastos no compensa la pérdida por establecer un día con entrada gratis.

Otro debate que se presenta es el relacionado con la estructuras de tarifas que se debe implementar en un museo, teniendo en cuenta la equidad y eficiencia. Por un lado, se pueden establecer diferentes precios que permitan el acceso a las personas con bajos recursos, menores, estudiantes y ancianos. También se ha expuesto la necesidad de establecer tarifas diferenciales en los horarios y fechas en las que se presenta mayor congestión y, así lograr mayor eficiencia en el servicio a los visitantes. Por último, se justifican diferentes precios para los diversos servicios o productos ofrecidos. Frey (1994), es uno de los exponentes que justifican la política de precios diferencial.

Con relación a la política pública, se encuentra que el precio de entrada puede ser perjudicial al reducir los niveles de subsidios, ya que el museo cuenta con ingresos adicionales que desmotivan la asignación de recursos hacia éstos (Anderson, 1998). En oposición, las teorías sobre la eficiencia económica en el museo indican que el establecimiento del precio de entrada es necesario porque los costos en que se incurre para obtener ingresos públicos pueden ser mayores que los recursos adquiridos. Con relación a la discusión sobre la justificación de los subsidios a los museos y la cantidad óptima se encuentra una gran variedad de trabajos (Duffy 1992; Martin, 1995; Asworth and Johnson, 1996; Grampp (1996); Barros, 1998; Santagata y Signorelli, 2000; Maddison, 2004).

Así mismo, se han desarrollado diversos estudios acerca de la eficiencia en los museos en los que realizan comparaciones entre actividades específicas de los museos basadas en indicadores de desempeño (Jackson, 1991; Weil, 1995). Otros estudios utilizan las técnicas de frontera para analizar la eficiencia entre museos (Paulus, 1996; Pignataro, 2002; Basso y Funari, 2004). Uno de los indicadores más utilizados para evaluar la eficiencia es el número de visitantes, puesto que, al ser entidades sin ánimo de lucro el interés por la maximización de los beneficios se traslada hacia la maximización en el número de visitantes y la calidad del servicio, por lo que algunos directores se inclinan por establecer un precio bajo o entrada libre para incentivar las visitas al museo.

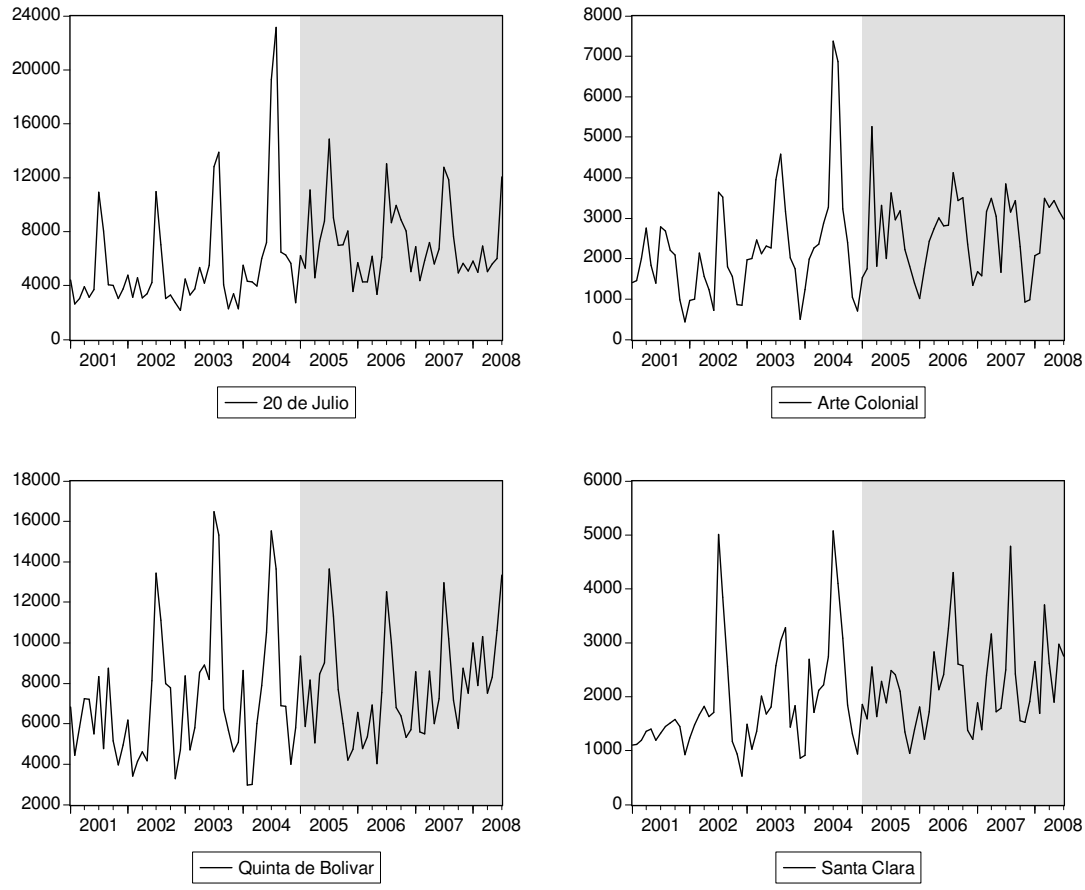
2 DESCRIPCION DE LOS DATOS

La base de datos utilizada para el desarrollo de este trabajo se construyó a partir de los reportes mensuales de gestión que los museos Quinta de Bolívar, 20 de Julio de 1810, Arte Colonial y Santa Clara entregan a la Red Nacional de Museos. Con esta información, se estructuró una base de datos mensual (enero de 2001 a julio de 2008) que da cuenta del número total de visitantes a cada uno de los museos discriminado por gratuitos y no gratuitos, así como los ingresos percibidos por boletería. Esta sección tiene como objetivo describir los patrones temporales de comportamiento de los visitantes a cada uno de los museos antes mencionados.

2.1 Visitantes

Como resultado de la implementación de la ley 929 de 2004, el promedio anual de visitantes totales en La Quinta de Bolívar pasó de 86.502 en el período 2001 – 2004 a 92.589 en el período 2005 – 2008. Para Santa Clara, 20 de Julio y Arte Colonial, los promedios para el primer período fueron 23.472, 63.783 y 26.237, respectivamente. Para el segundo período llegaron a 27.016, 83.061 y 30.910 visitantes. El museo que registró mayor crecimiento del promedio de visitantes totales entre los dos períodos, fue 20 de Julio con 30.2%, seguido por Arte Colonial con 17.8%, Santa Clara 15.1% y Quinta de Bolívar 7%.

Gráfico 1. Total Visitantes



Fuente: Red Nacional de Museos

Sin embargo, el impacto inicial de la medida fue negativo, pues todos los museos registraron una caída importante en el número de visitantes totales al año siguiente de aplicar la ley[4]: en 20 de Julio fue 24.3%, Quinta de Bolívar 15.5%, Arte Colonial 16.1% y Santa Clara 7.7%. Durante el último año analizado, 20 de Julio y Arte Colonial registran nuevamente decrecimiento de 9,3% y 5,7%, respectivamente.

El promedio de visitantes totales en el mes de julio decreció entre los períodos analizados en los cuatro museos. En Arte Colonial y Santa Clara la disminución fue importante, 25.3% y 21.1%, respectivamente. Por su parte, 20 de Julio y Quinta de Bolívar registraron una caída igual de sólo 2.5%. Estos resultados podrían indicar que hubo un impacto positivo de la medida sobre el nivel de congestión en los museos Arte Colonial y Santa Clara, puesto que, el número de visitantes totales durante el mes de julio decreció, aunque registraron un crecimiento anual positivo.

El promedio de visitantes totales en el mes de julio durante 2001 - 2004 en el museo Arte Colonial fue de 4.441, lo que representó el 16.9% de los visitantes totales en el año, en Santa Clara 3.492 (14.8%), en 20 de Julio 13.504 (21.2%) y Quinta de Bolívar 13.449 (15.5%). Para el período comprendido entre 2005 – 2008 los datos son los siguientes: Arte Colonial 3.315 (10.7%), Santa Clara 2.754 (10.2%), 20 de Julio 13.167 (15.8%) y Quinta de Bolívar 13.118 (14.2%).

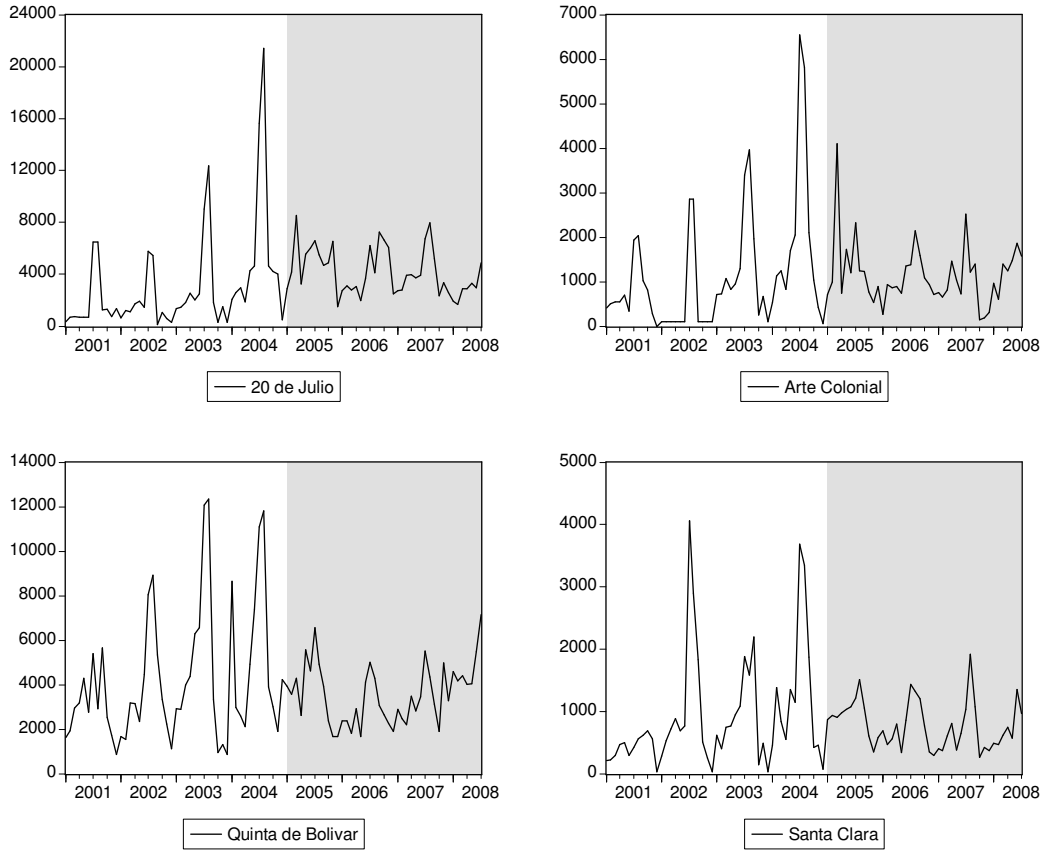
[4] Debido a que las cifras disponibles están hasta julio de 2008, los crecimientos anuales se calcularon con relación a este mes.

2.2 Visitantes gratuitos

El promedio anual de visitantes gratuitos presentó un crecimiento negativo entre los dos períodos analizados en todos los museos, excepto en 20 de julio. Santa Clara pasó de 12.046 en el período 2001 – 2004 a 8.865 en el período 2005 – 2008, con un decrecimiento de 26.4%; Quinta de Bolívar pasó de un promedio de 52.152 a 41.066, con una caída de 21.3%; Arte Colonial se redujo de 13.585 a 12.616, decreciendo en 7.1% y 20 de julio aumento de 34.429 a 47.623, con un crecimiento de 38.3%.

Todos los museos registraron una caída sustancial en los visitantes gratuitos al año siguiente de aplicar la ley: Arte Colonial 47.8%, Quinta de Bolívar 37.8%, 20 de Julio 35% y el museo Santa Clara 29.9%. Aunque al año siguiente 20 de Julio y Arte Colonial registran crecimiento, nuevamente caen entre 2007-2008 en 22.8% y 14.1, respectivamente. Por su parte, Quinta Bolívar crece 39.6% y Santa Clara 13.2% durante el último año analizado.

Gráfico 2. Visitantes Gratuitos



Fuente: Red Nacional de Museos

Como es de esperar, la participación de los visitantes gratuitos y los no gratuitos dentro del total cambió en todos los museos, después de la intervención. Durante el período 2001-2004, los gratuitos participaban con el 60.3% en La Quinta de Bolívar, 51.3% en Santa Clara, 54% en 20 de Julio y 51,8% en Arte Colonial. Para el período 2005-2008, su participación cayó en Quinta de Bolívar 44.4%, Santa Clara 32.8% y Arte Colonial 40.8%, mientras que, en el 20 de Julio aumentó a 57,3%. El promedio de visitantes gratuitos en el mes de julio decreció entre los períodos analizados en los cuatro museos. El que mayor caída registró fue Santa Clara con (53.9%), seguido por Arte Colonial (47%), Quinta de Bolívar (33.8%) y 20 de Julio (33.8%).

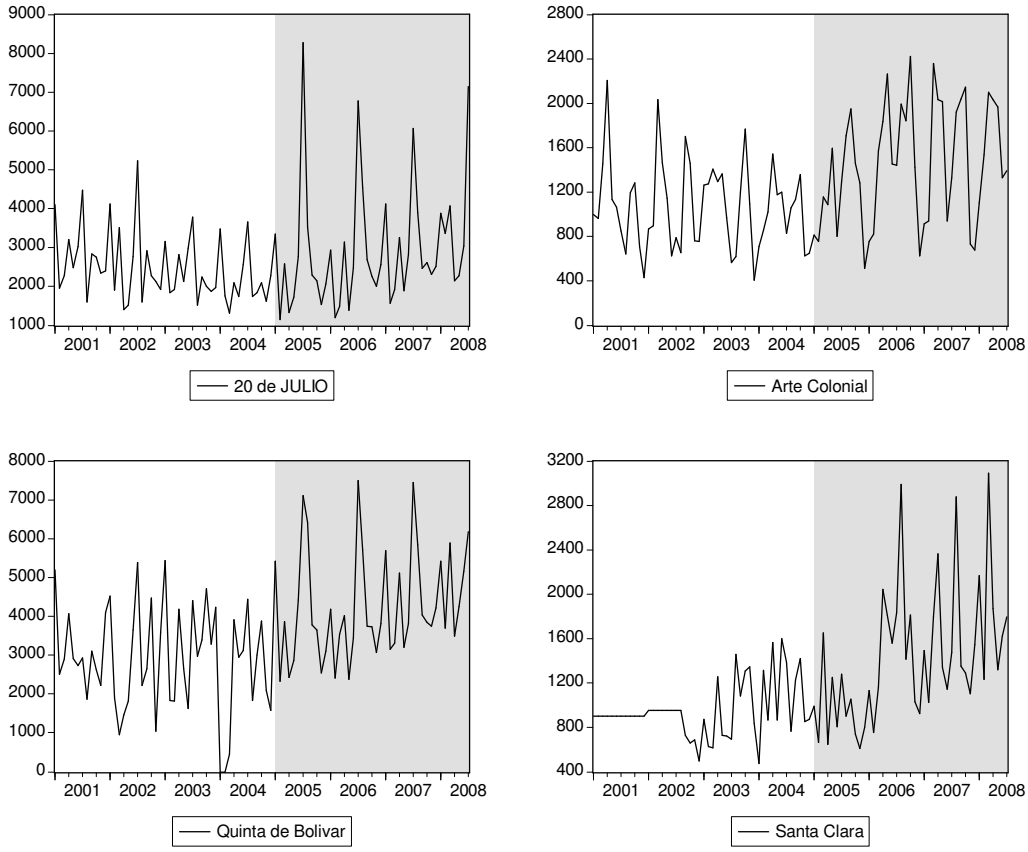
El promedio de visitantes gratuitos en el mes de julio durante 2001 - 2004 en Santa Clara fue de 2.511, lo que representó el 20.9% de los visitantes totales en el año. En Arte Colonial

3.648 (27.1%), en 20 de Julio 9.217 (26.8%) y Quinta de Bolívar 9.153 (17.6%). Para el período comprendido entre 2005 – 2008 los datos son los siguientes: Santa Clara 1.159 (13.1%), Arte Colonial 1.951 (15.5%), 20 de Julio 6.100 (12.8%) y Quinta de Bolívar 6.058 (14.8%).

2.3 Visitantes no gratuitos

Después de la implementación de la nueva política, el promedio anual de visitantes no gratuitos se incrementó al comparar el período comprendido entre 2001- 2004 y el período 2005-2008, en los cuatro museos. El que mayor crecimiento registró fue Santa Clara (58.9%) al pasar de 11.425 a 18.151, seguido de Quinta de Bolívar (50%) al pasar de 34.349 a 51.523, Arte Colonial (44.6%) al aumentar de 12.651 a 18.293 y, 20 de julio (20.7%) al aumentar de 29.354 a 35.437 visitantes no gratuitos.

Gráfico 3. Visitantes No Gratuitos



Fuente: Red Nacional de Museos

El promedio de visitantes no gratuitos en el mes de julio se incrementó entre los períodos analizados en los cuatro museos. En Arte Colonial registró un crecimiento de 80.2%, 20 de Julio 64.8%, Quinta de Bolívar 64.4% y Santa Clara 62.6%. En el período comprendido entre 2001 - 2004, el promedio de visitantes no gratuitos en el mes de julio en el museo Arte Colonial fue de 757, lo que representó el 6% de los visitantes totales en el año, en Santa Clara 981 (8.6%), en 20 de Julio 4.287 (14.6%) y Quinta de Bolívar 4.295 (12.5%). Para el período comprendido entre 2005 – 2008 los datos son los siguientes: Arte Colonial 1.595 (7.5%), Santa Clara 1.595 (8.8%), 20 de Julio 7.067 (19.9%) y Quinta de Bolívar 7.060 (13.7%).

2.4 Ingresos por boletería

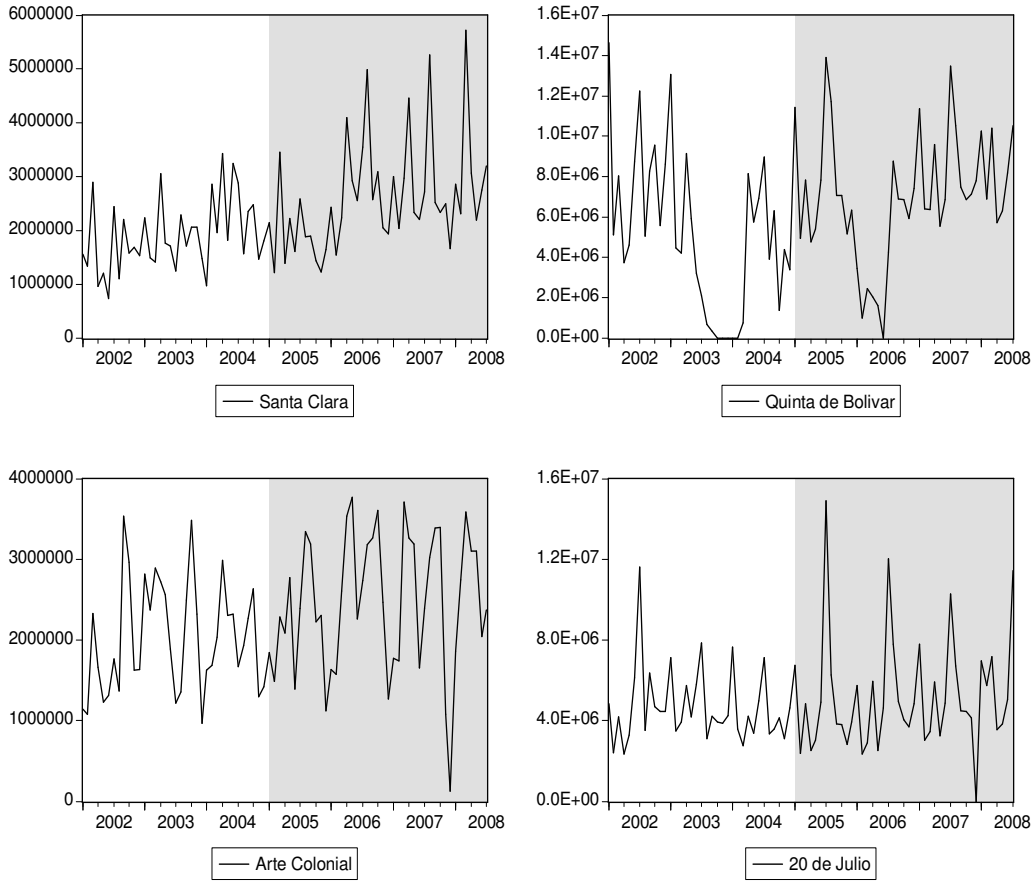
El promedio de los ingresos por boletería registraron crecimiento en los cuatro museos al comparar el período comprendido 2002 - 2004 y 2005 - 2007: el que mayor crecimiento presentó fue Santa Clara (36.2%) al pasar de $24 millones en el primer período a $32,7 millones

en el segundo período, seguido por Quinta de Bolívar (32%) al pasar de $62 millones a $81,8 millones, Arte Colonial (19.6%) que aumentó de $25,5 millones a $30,5 millones y 20 de Julio (6.6%) al crecer de $57,6 millones a $61,4 millones.

El promedio de los ingresos por boletería en el mes de Julio registró un crecimiento en los cuatro museos al comparar los períodos comprendidos entre 2002 - 2004 y 2005 – 2007. En orden de crecimiento, Arte Colonial (59.1%) al pasar de $1,5 millones en el primer período a $2,4 millones en el segundo período, seguido por Santa Clara (37.5%), al pasar de $2,1 millones a $3 millones; 20 de Julio (37.3%), al crecer de $8,8 millones a $12,1 millones; Quinta de Bolívar (35.7%) al pasar de $7,7 millones a $10,5 millones.

Comparando los dos períodos, la participación de los ingresos por boletería en julio dentro de los ingresos por boletería en el año aumentó en todos los museos. En 20 de Julio pasó de 15.4% a 19.8%, Quinta de Bolívar de 12.5% a 12.9%, Santa Clara de 9.1 % a 9.2% y Arte Colonial de 6.1% a 8.1%.

Gráfico 4. Ingresos por Boletería



Fuente: Red Nacional de Museos

3 METODOLOGIA

El análisis de intervención es una técnica econométrica que busca estimar específicamente la influencia de ciertos eventos externos en el comportamiento histórico de una serie temporal. Estos eventos externos se denominan intervenciones y están asociados, entre otras acciones, a cambios específicos en la política pública. Según Guerrero (2003), medir el impacto de esta intervención es de suma importancia dado que puede ayudar a explicar el comportamiento de la serie y mejorar la estimación de los parámetros y los resultados del modelo.

La técnica de análisis de intervención esta basada en la metodología ARIMA desarrollada por Box y Jenkins en 1976, por cuanto una serie temporal que involucre intervenciones externas puede ser expresada como la suma de un modelo ARIMA estacionario invertible y una función que represente los efectos de la intervención. La expresión $T(Z_t) = N_t + E_{I,t}$ incorpora a una serie temporal el efecto de una intervención, donde N_t se asocia con el modelo temporal estacionario ARIMA $\phi(B)\nabla^d N_t = \phi_0 + \theta(B)a_t$, y $E_{I,t}$ es la función de intervención. Por tanto, para estimar el efecto de una intervención sobre el comportamiento de una serie de tiempo es indispensable estimar un modelo ARIMA estacionario e invertible e identificar la forma funcional de la intervención.

3.1 Modelos para series de tiempo estacionarios (ARIMA)

En 1976 Box y Jenkins desarrollaron una metodología que permite ajustar y estimar modelos paramétricos finitos a series temporales de la forma $Y_t = a_0 + a_1 Y_{t-1} + ... + a_p Y_{t-p} + e_t + ... + \beta_q e_{t-q}$ llamados modelos autoregresivos integrados de media móvil o ARIMA (P,D,Q) por sus siglas en ingles. Esta técnica se desarrolla en 5 pasos a saber: estacionariedad, identificación, estimación, verificación de supuestos y finalmente el uso del modelo (pronóstico, simulación o explicación del fenómeno de estudio).

Una variable aleatoria Z_t es un ***proceso estacionario*** de segundo orden cuando sus dos primeros momentos, estos son, la esperanza matemática y varianza son invariantes en el tiempo y su covarianza es constante para todo tiempo y rezago. Formalmente, la estacionariedad de segundo orden se expresa con las siguientes expresiones:

$$E(Y_t) = \nu; \gamma_0 = E\left[(Y_t - \nu)^2\right] = \sigma_a{}^2 \sum_{i=0}^{\infty} a_i{}^2; \gamma_k = E\left[(Y_t - \nu)(Y_{t+k} - \nu)\right] = \sigma_a{}^2 \left(\sum_{i=1}^{\infty} a_i a_{k+i} - a_k \right)$$

$$\sum_{i=0}^{\infty} a_i{}^2 \rightarrow Convergente; \left(\sum_{i=1}^{\infty} a_i a_{k+i} - a_k \right) \rightarrow Convergente$$

Para identificar la estacionariedad o no de una serie de tiempo se realizan las pruebas de raíz unitaria[5]. El algoritmo estadístico de la prueba aumentada de Dickey – Fuller requiere la estimación por mínimos cuadrados ordinarios de tres ecuaciones, las cuales se diferencian unas de otras por la inclusión de elementos determinísticos identificados como parámetros constantes y tendencias. Las igualdades a estimar para el desarrollo de este algoritmo son las siguientes:

$$\Delta Y_t = \rho Y_{t-1} + \sum_{t=1}^{p} b_j \Delta Y_{t-1} + e_t; \Delta Y_t = a_0 + \rho Y_{t-1} + \sum_{t=1}^{p} b_j \Delta Y_{t-1} + e_t; \Delta Y_t = a_0 + a_2 t + \rho Y_{t-1} + \sum_{t=1}^{p} b_j \Delta Y_{t-1} + e_t$$

El parámetro de interés en cada una de las ecuaciones anteriores es ρ. La hipótesis nula a ser contrastada es la presencia de raíz unitaria en la serie, lo que equivale a decir que ρ es estadísticamente igual a cero. Una vez estimado ρ para cada una de las ecuaciones, se calcula su error estándar y se compara su estadístico t con el valor apropiado reportado en las tablas de Dickey-Fuller. De esta forma, es posible rechazar o no la hipótesis nula y así determinar la estacionariedad o no de la serie analizada.

Una vez se constate que la serie es estacionaria, se pasa a la etapa de ***identificación*** del modelo. Esta etapa tiene como objetivo equiparar aproximadamente los patrones de autocorrelación que gobiernan los diferentes procesos. Estos procesos son conocidos como autorregresivos de orden P (AR(P)), de media móvil de orden Q (MA(Q)) y la combinación de los dos anteriores ARMA(P,Q). Si fuera necesario aplicar una diferencia a la serie en niveles para conseguir la estacionariedad, el proceso se denominaría ARIMA(P,D,Q).

Los procesos autorregresivos de orden P se caracterizan porque la realización actual de la variable aleatoria se relaciona con sus valores anteriores mediante la siguiente representación:

[5] En la literatura estadística existe una gran variedad de pruebas de raíz unitaria entre las que se encuentran las Dickey y Fuller, Phillips Perron, KPSS, entre otras. En este trabajo se analizaran las pruebas Dickey y Fuller.

$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p}$, los cuales presentan como característica fundamental funciones de autocorrelación infinita y funciones de autocorrelación parcial "truncadas" en el rezago P.

Por su parte, los procesos de promedio móviles de orden Q representan un proceso estocástico como una suma finita ponderada de choques aleatorios independientes $Y_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + ... + \theta_q a_{t-q}$, que presentan como característica fundamental funciones de autocorrelación "truncadas" en el rezago Q y funciones de autocorrelación parcial infinitas.

Finalmente, la combinación de los dos procesos anteriores da como resultado los modelos ARMA(P,Q), los cuales se conforman de una parte autoregresiva y la otra de una suma de promedios móviles los cuales se configuran matemáticamente de la forma $Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - ..... - \phi_p Y_{t-p} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - ..... - \theta_q a_{t-q}$. Tienen como característica fundamental funciones de autocorrelación y de autocorrelación parcial infinitas.

En conclusión, en la etapa de identificación se intenta reconocer por medio de instrumentos gráficos y estadísticos (funciones de autocorrelación y autocorrelación parcial) los ordenes de los polinomios autorregresivos y de promedios móviles y las pautas que caracterizan los distintos procesos.

Una vez se han identificado y formulado el proceso que gobierna la variable estocástica se continua con el proceso de ***estimación***, el cual, según Guerreo (2003) busca encontrar los mejores valores de los parámetros para que dicho modelo represente apropiadamente la serie en consideración. En esta etapa se utiliza el método de máxima verosimilitud bajo el supuesto de que a_t es un proceso de ruido blanco con distribución normal, media cero y varianza σ^2_a.

Posterior a la etapa de estimación se realiza la ***verificación de los supuestos*** anteriores y pruebas estadísticas adicionales que permiten seleccionar dentro de los posibles modelos el que mejor represente la serie estudiada. El primer test de verificación se hace sobre los residuos los cuales deben tener media cero, varianza constante, mutuamente independientes y distribuidos normalmente. Posteriormente se debe comprobar que el modelo considerado sea parsimonioso, admisible económicamente y estable en los parámetros.

Para calcular si los residuos tienen media cero, y siguiendo a Guerrero (2003), se computa la media aritmética y la desviación estándar muestral de los residuos de la siguiente forma $m(\hat{a}) = \sum \frac{\hat{a}_t}{(N-d-p)}$ y la desviación estándar $\hat{\sigma}_a = \sqrt{\frac{\sum[\hat{a}_t - m(\hat{a})]^2}{(N-d-p-q)}}$.

Una vez calculado las ecuaciones anteriores, se construye el cociente $\frac{\sqrt{N-d-p-m(\hat{a})}}{\hat{\sigma}_a}$. Si su valor absoluto es menor a 2 no se rechaza el supuesto de media cero en los residuos. En caso contrario, si este cociente es mayor o igual a 2, la media de los residuos es significativamente diferente de cero.

Respecto al supuesto de varianza constante en el tiempo, basta con realizar una gráfica de los residuos obtenidos de la estimación contra el tiempo y observar su comportamiento. Si los residuos pueden ser definidos por un límite constante superior e inferior en el tiempo, la varianza es constante. En caso contrario se rechaza la hipótesis de varianza invariante. Referido a la normalidad de los residuales, es posible aplicar test estadísticos como el Jarque-Bera o realizar histogramas que permitan visualizar la forma de la distribución con el fin de detectar posible asimetrías en la misma.

Referido a la no correlación en los residuales se debe calcular primero la función de autocorrelación muestral de los residuos y su respectiva desviación estándar. Acto seguido, construir el estadístico Q de Box y Pierce para realizar la prueba de significación conjunta de K autocorrelaciones simultáneamente y comparar el resultado con los valores de la tabla ji-cuadrada con los grados de libertad adecuados. Si la hipótesis nula de no correlación en los residuales no se rechaza, el supuesto es adecuado. El estadístico Q de Box y Pierce se define como $Q = (N-d-p)\sum_{K=1}^{K} r^2{}_k(\hat{a}); r^2{}_k(\hat{a}) = Autocorrelación$

3.2 Función de intervención

La función de intervención es una modelación matemática que tiene como objetivo detectar y estructurar las modificaciones ocurridas en la serie, no solo en el momento en que se presenta una intervención, sino en los cambios posteriores que son atribuibles a ella. En este sentido, Guerrero (2003) muestra la necesidad de concebir la intervención como un modelo dinámico. Formalmente, un modelo dinámico general de intervención puede representarse con la siguiente expresión:

$$\delta(B)\nabla^b \xi_{I,t} = W(B)P_{I,t}; \delta(B) = 1 - \delta_1 B - .... - \delta_r B^r; W(B) = w_0 - w_1 B - .. - w_s B^s$$

$$P_{I,t} = 1_si_t = I; P_{I,t} = 0_si_t \neq I; I = Intervención; r, s = \text{Re}\, zagos$$

La estructura de los efectos esperados de la intervención dependerá principalmente de los coeficientes de los polinomios *δ(B)* y *W(B)*, así como de los rezagos óptimos representados por *r* y *s*. Por ejemplo, si el efecto esperado de la intervención sobre la serie es un cambio en el nivel de la misma después de un periodo de retraso, el modelo de intervención tomo la forma de $\Delta\xi_{I,t} = (-w_1 B)P_{I,t}$.

En conclusión, Guerrero (2003) muestra que la metodología del análisis de intervención consiste en determinar el momento en que se introdujo algún mecanismo o alguna política que se piense pudo haber influido de manera exógena sobre el comportamiento de la variable de interés. Acto seguido, se construye un modelo que represente la serie desde la observación inicial hasta el periodo antes de la intervención. Posteriormente se procede a postular un modelo dinámico para representar el efecto de la intervención, para luego estimar un modelo completo para todo el periodo muestral.

4 ESTIMACION DE MODELOS Y RESULTADOS

Esta sección tiene como objetivo describir, estimar y presentar los resultados de los modelos ARIMA de intervención para las series total de visitantes (TOTAL), visitantes gratuitos (GRATUITOS) y no gratuitos (BOLETA) de los museos 20 de Julio, Quinta de Bolívar, Arte Colonial y Santa Clara.

4.1 Total visitantes

El análisis de las series temporales de los visitantes totales a los museos muestra una homogeneidad en el proceso generador de datos. Es así que, el mejor modelo ARIMA estimado para todos los museos esta constituido por el primer rezago autoregresivo regular y el primer rezago autoregresivo estacional. El cambio de la política cultural (intervención) se modela con una dummy de paso la cual recoge los cambios en el nivel medio de la serie (INTER_PASO) y una dummy de pulso instantánea que busca estimar el impacto sobre el componente estacional (INTER_PULSO).

La Ecuación 1 formaliza matemáticamente la estructura del modelo estimado y la Tabla 1 muestra los resultados.

Ecuación 1. Total Visitantes

$$MT_i = c + \beta_0 INTER_PASO + \beta_1 INTER_PULSO + \beta_2 MT_{i,t-1} + \beta_3 MT_i^{S}{}_{t-1} + E_{i,t}$$

i = Museo
MT_i = Logaritmo del total de visitantes[6].
INTER_PASO = Toma el valor de 0 antes de 2005 y 1 en otro caso
INTER_PULSO = Toma el valor de 1 en las estacionalidades y 0 en otro caso
MT_{it-1} = Primer rezago regular
$MT^{s}_{i\,t-1}$ = Primero rezago estacional

[6] El logaritmo del total de visitantes se aplica a todos los museos excepto al museo Quinta de Bolívar.

Tabla 1. Resultados Total Visitantes

	20 DE JULIO	QUINTA DE BOLIVAR	ARTE COLONIAL	SANTA CLARA
INTER_PASO	0,2 (-0,14)	704 (1084)	0,19 (0,18)	0,19 (0,16)
INTER_PULSO	0,08 (0,29)	-734 (3144)	-0,21 (0,30)	-0,06 (0,34)
AR(1)	0,37* (0,09)	0,39* (0,08)	0,41* (0,09)	0,469* (0,10)
ARs(1)	0,64* (0,06)	0,66* (0,07)	0,53* (0,11)	0,40* (0,11)
Log-likelihood	-22,38	-817,07	-39,81	-26,72

*** Significativas al 1%**

Errores estandar entre paréntesis

Fuente: Cálculos del autor

Los resultados mostrados en la tabla anterior indican que para todos los museos excepto Santa Clara la estacionalidad es el factor que más se pondera al explicar el número de visitantes actuales. Este resultado no sorprende dado que estas se presentan en la temporada de vacaciones estudiantiles de mitad de año, periodo en el cual los estudiantes aprovechan su receso escolar para realizar labores escolares muchas de ellas relacionadas con las colecciones que estos museos exhiben razón por la cual son de obligatoria visita. Referido al rezago regular, no se presentan grandes diferencias en los museos dado que se exhibe un comportamiento alrededor de 0.4.

Por su parte, los modelos indican que la intervención gubernamental debido al cambio de política cultural no modificó sustancialmente el comportamiento de los visitantes a los respectivos museos. Aunque los coeficientes de la dummy INTER_PASO muestran crecimientos en el nivel promedio de los visitantes para todos los museos. Es importante notar que el coeficiente del 20 de Julio es superior a los otros museos (0.2) y la similitud de los coeficientes de Arte Colonial y Santa Clara (0.19). No obstante el comportamiento anterior estos crecimientos no son significativos a ningún nivel de confianza.

Respecto al comportamiento estacional se encuentra que existió una disminución en todos los museos excepto en el 20 de julio el cual muestra un leve crecimiento (0.08). Los coeficientes de la dummy INTER_PULSO muestran que en los museos Arte Colonial (-0.21) y Santa Clara (-0.06) existieron decrecimientos importantes, mientras que en Quinta de Bolívar y 20 de Julio no fue así. Si bien se presenta este comportamiento, los coeficientes no se estiman significativos a ningún nivel de confianza.

4.2 Visitantes gratuitos

Respecto a los modelos estimados para el grupo de visitantes gratuitos existen pequeñas diferencias para cada uno de los museos. Si bien todos se especifican adecuadamente con el primer rezago autorregresivo regular y el primer rezago autorregresivo estacional, en el museo Santa Clara se encontró como significativo el séptimo rezago regular, mientras que en el museo Quinta de Bolívar se adiciona el décimo primer rezago de media móvil. La Ecuación 2 muestra la especificación matemática para cada museo y la Tabla 2 presenta los resultados.

Ecuación 2. Visitantes Gratuitos

$$VG_i = c + \beta_0 INTER_PASO + \beta_1 INTER_PULSO + \beta_2 VG_{i,t-1} + \beta_3 VG_i{}^{S}{}_{t-1} + E_{i,t}$$

$$VG_k = c + \beta_0 INTER_PASO + \beta_1 INTER_PULSO + \beta_2 VG_{k,t-1} + \beta_3 VG_k{}^{S}{}_{t-1} + \beta_4 E_{k,t-11} + E_{k,t}$$

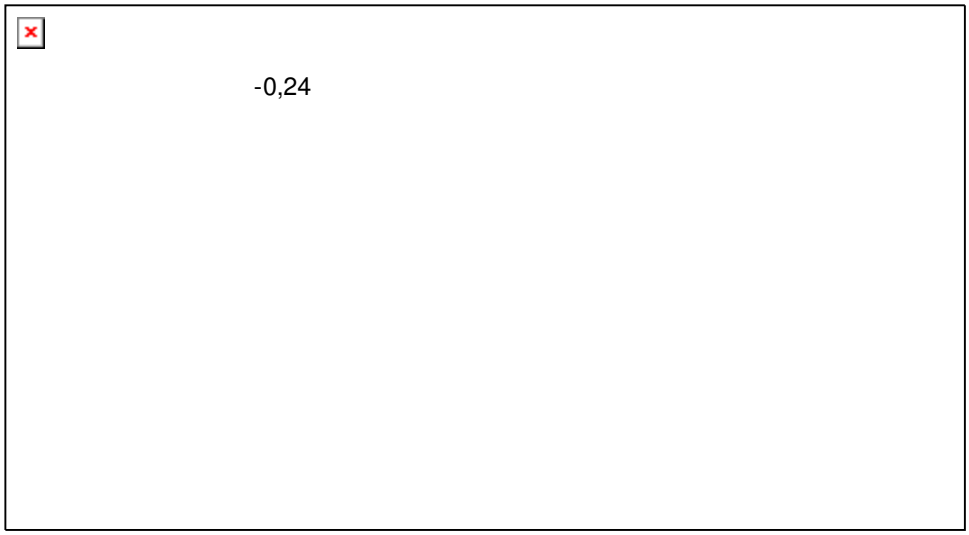
$$VG_m = c + \beta_0 INTER_PASO + \beta_1 INTER_PULSO + \beta_2 VG_{m,t-1} + \beta_3 VG_{m,t-7} + \beta_4 VG_m{}^{S}{}_{t-1} + E_{m,t}$$ i =

Museo 20 de Julio / Museo de Arte Colonial

k = Museo Quinta de Bolívar

m = Museo Santa Clara

Tabla 2. Resultados Visitantes Gratuitos



Fuente: Cálculos del autor

La especificación de los modelos y sus respectivos resultados para los visitantes gratuitos revelan una hetorogeneidad para cada uno de ellos. Por ejemplo, los patrones estacionales son más importantes en el Museo del 20 de Julio (0.55), Quinta de Bolívar (0.51) y Santa Clara (0.68) por dos razones: la primera se relaciona a la temporada de vacaciones estudiantiles que incentiva la visita a los mismos para el desarrollo de las labores escolares, mientras que la segunda se asocia a la estrecha relación de las colecciones con las fiestas patrias que se celebran en esa época. Por ejemplo, el Museo del 20 de Julio se asocia con el grito de independencia llevado a cabo el 20 de julio de 1810 y la batalla del puente de Boyacá se asocia más con la vida y obra del libertador Simón Bolívar.

Respecto a la intervención, los resultados indican que ésta no modificó sustancialmente el comportamiento de visitas gratuitas a los respectivos museos, excepto en Quinta de Bolívar donde se redujo el número de visitantes estacionales. Aunque los patrones estacionales son los mismos, es importante notar que los signos negativos en los coeficientes de los modelos indican una disminución en el número de visitas, aunque éstas no sean significativas estadísticamente. Por ejemplo, los coeficientes de la variable INTER_PULSO fueron de -0.24 para 20 de Julio, -1306 para Quinta de Bolívar, -0.53 para Santa Clara y -0.12 para Arte Colonial. Es importante notar que el Museo Santa Clara tuvo la mayor caída respecto de los demás.

4.3 Visitantes no gratuitos

Los modelos para los visitantes no gratuitos muestran gran heterogeneidad para cada uno de los museos. La mejor especificación se logra estimando solo rezagos autorregresivos y de media móvil regulares sin ser significativos los patrones estacionales. La Ecuación 3 muestra la especificación matemática para cada museo y la Tabla 3 muestra los resultados:

Ecuación 3. Visitantes No Gratuitos

$$VNG_i = c + \beta_0 INTER_PASO + \beta_1 INTER_PULSO + \beta_2 VNG_{i,t-1} + E_{i,t}$$

$$VNG_k = c + \beta_0 INTER_PASO + \beta_1 INTER_PULSO + \beta_2 VNG_{k,t-12} + \beta_3 E_{k,t-2} + E_{k,t}$$

$$VNG_m = c + \beta_0 INTER_PASO + \beta_1 INTER_PULSO + \beta_2 VNG_{m,t-1} + \beta_3 VNG_{m,t-12} + \beta_4 E_m{}^{S}{}_{t-3} + E_{m,t}$$

$$VNG_z = c + \beta_0 INTER_PASO + \beta_1 INTER_PULSO + \beta_2 E_{z,t-16} + E_{z,t}$$

i = Museo del 20 de Julio
k = Museo Quinta de Bolívar
m = Museo de Arte Colonial
z = Museo Santa Clara

Tabla 3. Resultados Visitantes No Gratuitos

	20 DE JULIO	QUINTA DE BOLIVAR	ARTE COLONIAL	SANTA CLARA
INTER_PASO	0,003 (0,09)	966* (253)	0,2* (0,07)	0,41* (0,07)
INTER_PULSO	0,851* (0,18)	2666*** (1698)	0,11*** (0,52)	0,55** (0,21)
AR(1)	0,45* (0,10)		0,21** (0,10)	
AR(12)		0,32* (0,11)	0,54* (0,11)	
MA(2)		-0,15*** (0,11)		
MA(3)			-0,24** (0,11)	
MA(16)				-0,37** (0,15)
Log-likelihood	-17,30	-760,94	-16,17	-13,0

*** Significativas al 1%; ** Significativa al 5%; *** Significativa al 10%**

Errores estandar entre paréntesis

Fuente: Cálculos del autor

Los resultados de la tabla anterior muestran la ausencia de patrones estacionales en la serie de visitantes no gratuitos y la heterogeneidad de cada uno de los modelos respecto a los rezagos autorregresivos y de media móvil. Referido a la intervención se encuentran dos resultados: el primero relacionado con el incremento promedio de visitantes para todos los museos, aunque en el 20 de Julio su efecto no sea significativo, y el segundo, enfocado al surgimiento de un comportamiento estacional en todos los museos debido al cambio en la política.

El incremento promedio se observa por el signo positivo de los coeficientes de la variable INTER_PASO la cual se estima para Quinta de Bolívar en 966, Arte Colonial en 0.2 y Santa Clara en 0.41 donde en este último se presenta el mayor impacto. Por su parte, el comportamiento estacional se ve reflejado en la variable INTER_PULSO donde los coeficientes son significativos y positivos, los cuales se estiman 0.8 para el 20 de Julio, 2666 para Quinta de Bolívar, 0.11 para Arte Colonial y 0.55 para Santa Clara. Esta conducta se explica porque al ser el periodo de vacaciones de mitad de año el único tiempo que se destina para la visita a estos museos (y el tiempo en el cual los padres pueden compartir el tiempo libre con sus hijos), no es posible esperar el día en que éste es gratuito y es preferible pagar para visitarlo que dejar de hacerlo.

4.4 Ingresos por boletería

Los modelos estimados por los ingresos por boletería son heterogéneos para cada uno de los museos mostrando, en general, comportamientos autorregresivos y de media móvil. La Ecuación 4 muestra la especificación matemática para cada museo y la Tabla 4 indica los resultados.

Ecuación 4. Ingresos por Boletería

$$IB_i = c + \beta_0 INTER_PASO + \beta_1 INTER_PULSO + \beta_2 E_{i,t-18} + \beta_3 E^S_{i,t-1} + E_{i,t}$$

$$IB_k = c + \beta_0 INTER_PASO + \beta_1 INTER_PULSO + \beta_2 IB_{k,t-1} + E_{k,t}$$

$$IB_m = c + \beta_0 INTER_PASO + \beta_1 INTER_PULSO + \beta_2 IB_{m,t-1} + \beta_3 IB_{m,t-3} + E_{m,t}$$

$$IB_z = c + \beta_0 INTER_PASO + \beta_1 INTER_PULSO + \beta_2 IB_{z,t-2} + E_{z,t}$$

i = Museo del 20 de Julio
k = Museo Quinta de Bolívar
m = Museo de Arte Colonial
z = Museo Santa Clara

Tabla 4. Resultados Ingresos por Boletería

	20 DE JULIO	QUINTA DE BOLIVAR	ARTE COLONIAL	SANTA CLARA
INTER_PASO	0,054 (0,08)	2187820* (839272)	640079* (139887)	0,43* (0,112)
INTER_PULSO	0,784* (0,20)	2758329* (835193)	368577 (322913)	0,33*** (0,125)
AR(1)	-	0,473* (0,08)	0,375* (0,09)	-
AR(2)	-	-	-	0,31* (0,114)
AR(3)	-	-	-0,373* (0,10)	-
MA(18)	0,22*** (0,12)	-	-	-
MAs(1)	0,3* (0,11)	-	-	-
Log-likelihood	-21,09	-1459,19	-1157,4	-30,90

*** Significativas al 1%; ** Significativa al 5%; *** Significativa al 10%**

Errores estandar entre paréntesis

Fuente: Cálculos del autor

Los resultados de la tabla anterior indican que en los museos existió un impacto positivo y significativo en los ingreso por venta de boletería. Por ejemplo, los museos 20 de julio, Quinta de Bolívar y Santa Clara presentan incrementos positivos en el periodo de receso escolar de mitad de año por cuanto los coeficientes de la variable INTER_PULSO resultaron significativos, positivos e iguales a 0.7, 0.33 y 2.7 respectivamente. Por su parte el promedio de ingreso de los museos tuvo un incremento estadísticamente significativo e importante en el museo Santa Clara (0.43) y Quinta de Bolívar (2.1).

5 CONCLUSIONES

Los principales resultados encontrados en este documento muestran que la modificación de la legislación produjo por un lado, un impacto positivo pero no significativo sobre el crecimiento anual promedio del total de visitantes para todos los museos, y por el otro una disminución del número de visitantes en el mes de julio en todos los museos, disminución que se percibe con más fuerza en Arte Colonial y Santa Clara, y menos importante en Quinta de Bolívar. Sin embargo, aunque esta disminución está asociada a una descongestión en el mes de julio, su resultado no es estadísticamente significativo.

Respecto a los visitantes gratuitos la modificación de la legislación no tuvo un impacto estadísticamente significativo, aunque los modelos arrojaron decrecimientos en el número de visitantes gratuitos en los meses de julio lo que refleja descongestión para este mes. Esta disminución de visitantes gratuitos es más importante en Santa Clara.

Para los visitantes no gratuitos se encuentra que el cambio legislativo tuvo un impacto positivo y significativo en el nivel de visitas destacándose el museo Santa Clara como aquel que tuvo el mayor impacto. Es importante anotar que para todos los museos se presenta un comportamiento estacional en los visitantes no gratuitos después del cambio legislativo, mostrando que las personas destinan sus vacaciones de mitad de año al turismo cultural y desarrollo de labores escolares.

Por último, se encuentra un impacto positivo y significativo sobre los ingresos por boletería para todos los museos. Es importante anotar que el crecimiento promedio fue menor en el museo 20 de Julio y mayor en Santa Clara. Por otro lado, surgen patrones estacionales después de la política debido al incremento de los visitantes no gratuitos en el periodo de vacaciones.

En conclusión, la modificación de la ley trajo por un lado una recomposición de los visitantes entre gratuitos y no gratuitos que generó una descongestión en el mes de julio dado que la disminución de visitantes gratuitos no fue compensada por el incremento en el número de visitantes no gratuitos, y por el otro, un incremento de los ingresos por boletería tanto en el nivel medio como en los meses de julio. Estos comportamientos se explican porque después de la ley las personas están más dispuestas a visitar el museo en otros meses aunque siguen prefiriendo los días de vacaciones de mitad de año este o no el museo gratuito. Aunque la descongestión no es estadísticamente significativa, el incremento de los ingresos por boletería si lo es.

6 RECOMENDACIONES

Las recomendaciones más importantes se refieren al fortalecimiento de los sistemas de captura de información no solo de los visitantes y sus características, sino de las variables propias de cada museo (infraestructura, costos, ingresos, capital humano) que permitan continuar con el desarrollo de estudios cuantitativos enfocados a ser el sustento técnico para la generación de estrategias y modificaciones legislativas culturales que permitan fortalecer el sector museístico en Colombia.

BIBLIOGRAFIA

ALBI, E. (2003), "Economía de las Artes y Política Cultural", Instituto de Estudios Fiscales, Madrid.

ANDERSON, R. G. W. (1998), "Is Charging Economic?", Journal of Cultural Economics, vol. 22, nº 2-3, pp. 179-187.

ASWORTH, J. and JOHNSON, P. (1996), "Sources of Value for Money for Museum Visitors: Some Survey Evidence". Journal of Cultural Economics 20, 67-83.

BAILEY, S. y FALCONER, P. (1998), "Charging for Admission to Museums and Galleries", Journal of Cultural Economics, vol. 22, nº 2-3, pp. 167-177.

BARRIO M., HERRERO L. y SANZ J., (2007), "Medición de Eficiencia de Instituciones Culturales. Aplicación a un Sistema Regional de Museos", Universidad de Valladolid, Ponencia en Seminario Internacional sobre Economía Inducida por los Museos, Valencia.

BEEN, I., VISSCHER, K. y GOUDRIAAN, R. (2002), "Fee or Free?", Working Paper Series (http://flyaquis.mov.vic.gov.au/been/Fee_or_free.doc).

BARROS, C. (1998), "Museum Funding: Hyphotetical Philanthropy and Fair Price". Paper presented in the Xth International Conference on Cultural Economics, Barcelona.

BASSO, A. and FUNARI, S., (2004), "A Quantitative Approach to Evaluate the Relative Efficiency of Museums", Journal of Cultural Economics, Vol. 28, pp. 195-216.

CREIGH-TYTE, S. y SELWOOD, S. (1998), "Museums in the U.K.: Some Evidence on Scale and Activities", Journal of Cultural Economics, vol. 22, nº 2-3, pp. 151-165.

DARNELL, A. C. (1998), "Some Simple Analytics of Access and Revenue Targets", Journal of Cultural Economics, vol. 22, nº 2-3, pp. 189-196.

DICKENSON, V. (1992), "Museum Visitor Survey: An Overview, 1930-1990" en Towse, R. (ed.) (1997), Cultural Economics: The Arts, The Heritage and the Media Industries, vol. 1, Edward Elgar, Cheltenham, pp. 272-281.

DiMAGGIO, P. (1991), "The Museum and the Public", en Feldstein, M. (ed.), The Economics of Museums, University of Chicago Press, Chicago, pp. 39-50.

DUFFY, C. T. (1992), "The Rationale for Public Funding of a National Museum". In: Towse R. and Khakee, A., (Eds.), Cultural Economics, Springer, Berlin, 37-48.

FELDSTEIN, M. (1991), "The Economics of Museums", University of Chicago Press, Chicago.

FREY, B. (1994), "Cultural Economics and Museum Behavior", Scottish Journal of Political Economy, 39 (3/4): 325-335.

FREY, B. and MEIER, S. (2003), "The Economics of Museums", Working Paper, University of Zurich Zurich.

FREY, B. y MEIER, S. (2002), "Museums between private and public: the case of the Beyeler Museum in Basile Working Paper Nº 116

GUERRERO, V. (2003), Análisis estadístico de series de tiempo económicas, International Thomson, 2ª ed, México.

HEILBRUN, J. y GRAY, Ch. (2001), "The Economics of Art and Culture". An American Perspective, Cambridge University Press,Cambridge.

JACKSON, J.M.(1991), "Performance Indicators: Promises and Pitfalls", en S. PEARCE, Museums Economics and the Community, Athlone Press, Londres, pp. 156-172.

JACKSON, R. (1988), "A Museum Cost Function", Journal of Cultural Economics, vol. 12, nº 1, p. 41-50.

JOHNSON P. y THOMAS B. (1998), "The Economics of Museums: A Research Perspective", Journal of Cultural Economics, Vol. 22 (2-3), pp. 75-85.

LUKSETICH, W. y PARTRIDGE, M. (1997), "Demand Functions for Museum Services", Applied Economics, vol. 29, nº 11, pp. 1553-1559.

MADDISON, D. y FOSTER, T. (2003), "Valuing Congestion Costs in the British Museum", Oxford Economic Papers, vol. 55, nº 1, pp.173-190.

MADDISON, D. (2004), "Causality and Museum Subsidies", Journal of Cultural Economics 28: 89–108.

MARTIN F. (1995), "Determining the Size of Museum Subsidies", Journal of Cultural Economics, Vol.18 (4), pp. 255-270.

O'HAGAN J.W. (1995), "National Museums: To Charge or Not to Charge", Journal of Cultural Economics, Vol. 19 (1), pp. 33-47.

PAULUS, O. (1996), "Museums Efficiency, Fourth European Workshop on Efficiency and Productivity Analysis", Université Catholique du Louvain.

PEACOCK, A. y GODFREY, C. (1974), "The Economics of Museums and Galleries", en Towse, R. (ed.), Cultural Economics: The Arts, The Heritage and the Media Industries, vol. 1, Edward Elgar, Cheltenham, pp. 364-375.

PIGNATARO, G.,(2002), Measuring the Efficiency of Museums: A Case Study in Sicily, en I. Rizzo y R. Towse (eds.), The Economics of Heritage: a Study in the Political Economy of Culture in Sicily, Capítulo 5, Edward Elgar, Cheltenham, pp. 65-78.

ROBBINS, L. (1971), "Unsettled Questions in the Political Economy of the Arts", en Towse, R. (ed.), Cultural Economics: The Arts, The Heritage and the Media Industries, vol. 1, Edward Elgar, Cheltenham, 1997, pp. 347-363.

RODRÍGUEZ J. y BLANCO V. (2002), "Optimal Pricing and Grant Policies for Museums", Instituto de Estudios Fiscales, Madrid.

SANTAGATA, W., and SIGNORELLO, G. (2000): "Contingent Valuation and Cultural Policy Design: The Case of Napoli Musei Aperti". Journal of Cultural Economics 24, 181-204.

STEINER, F. (1997), "Optimal Pricing of Museum Admission", Journal of Cultural Economics, vol. 21, nº 4, pp. 307-333.

TRIMARCHI, M. (2005). "El Análisis Principal-Agente" en Towse R. Manual de Economía de la Cultura. Fundación Autor. Madrid

ANEXO

Anexo 1: Pruebas de Especificación de los modelos

Tabla 5. Pruebas de Correlación y Normalidad. Visitantes Totales

	20 DE JULIO	QUINTA DE BOLIVAR	ARTE COLONIAL	SANTA CLARA
Portmanteau	37,72	33,41	55,08	36,34
P-value	0,57	0,75	0,06	0,63
Shapiro-Wilk	0,98 0,809	0,98 0,82	0,97 1,72	0,97 1,84
P-value	0,68	0,66	0,11	0,08

Fuente: Autor

Tabla 6. Pruebas de Correlación y Normalidad. Visitantes Gratuitos

	20 DE JULIO	QUINTA DE BOLIVAR	ARTE COLONIAL	SANTA CLARA
Portmanteau	20,53	44,08	31,69	35,37
P-value	0,99	0,3	0,82	0,67
Shapiro-Wilk	0,91 6,23	0,97 2,06	0,97 2,15	0,93 4,84
P-value	0,00	0,06	0,04	0,00

Fuente: Autor

Tabla 7. Pruebas de Correlación y Normalidad. No Gratuitos

	20 DE JULIO	QUINTA DE BOLIVAR	ARTE COLONIAL	SANTA CLARA
Portmanteau	40,95	37,55	32,59	23,52
P-value	0,42	0,58	0,79	0,98
Shapiro-Wilk	0,99 0,48	0,99 0,67	0,95 3,35	0,98 1,36
P-value	0,94	0,8	0,003	0,24

Fuente: Autor

Tabla 8. Pruebas de Correlación y Normalidad. Boletería

	20 DE JULIO	QUINTA DE BOLIVAR	ARTE COLONIAL	SANTA CLARA
Portmanteau	42,04	51,19	43,15	27,2
P-value	0,38	0,11	0,22	0,88
Shapiro-Wilk	0,98 0,76	2,02 1,55	0,98 1,26	0,98 1,13
P-value	0,723	0,06	0,3	0,39

Fuente: Autor

INDICADORES SINTÉTICOS DE BIENESTAR: UNA REVISIÓN.

MÓNICA DOMÍNGUEZ SERRANO
E-mail: mdomser@upo.es
FLOR Mª GUERRERO CASAS
E-mail: fguecas@upo.es
Universidad Pablo de Olavide

Resumen

La construcción de indicadores sintéticos para el análisis de aspectos económicos y sociales, así como para la gestión y evaluación de políticas públicas, se ha convertido en una herramienta muy habitual en el ámbito internacional. El estudio del bienestar dentro de este contexto es especialmente relevante, pues son muchos los indicadores de este tipo que han surgido en los últimos años, entre los que destaca el Índice de Desarrollo Humano (IDH) propuesto por el Programa de las Naciones Unidas para el Desarrollo (PNUD).
El cálculo de indicadores de bienestar pasa por el diseño y/o adopción de un sistema de variables de partida pero, además, exige la definición de una metodología de agregación que se ajuste al fenómeno objeto de estudio.
En el presente trabajo se pretende recopilar, a modo de revisión del estado del arte, un compendio de los principales indicadores sintéticos de bienestar, haciendo especial hincapié en las metodologías de agregación utilizadas en su cálculo.

Palabras Clave: Indicadores sintéticos, metodología de agregación, bienestar.
Área Temática: Economía Social y Discapacidad.

Abstract

Composite indicators have become a very common tool in the analysis of economic and social aspects, and in the management and evaluation of public policies. In this context, the study of welfare is particularly relevant, because there are many indicators that has emerged in recent years, including the Human Development Index (HDI) proposed by the United Nations Program for Development Program (UNDP). The construction of well-being indicators requires the design and/or the adoption of a system of variables and the definition of an aggregation methodology that fits the phenomenon under study.
This work aims to collect, as a review of the state of the art, a compendium of key summary indicators of wellbeing, with particular emphasis on methods of aggregation used in its calculation.

Key Words: Composite indicators, aggregation methodology, well-being.
Thematic Area: Social Economics.

1. INTRODUCCIÓN.

Los intentos de medir el bienestar han constituido en los últimos años un aspecto central en el análisis socioeconómico, principalmente por la necesidad de contar con indicadores fiables y útiles para el diseño y la evaluación de políticas públicas. Una de las principales corrientes que se ha ocupado de este aspecto es la de los Indicadores Sociales, tanto desde un punto de vista amplio basado en la búsqueda de sistemas de indicadores capaces de recoger la diversidad y detalles que son propios en un ámbito complejo como es el bienestar, como desde una perspectiva agregativa que se ha centrado en el diseño de indicadores sintéticos que hacen posible resumir en un solo índice la mayor información posible.

En este trabajo se presenta, en el apartado 2, el índice de bienestar más utilizado en la última década en el ámbito internacional, y en el apartado 3 se recogen cuáles son las principales

metodologías de agregación utilizadas en la construcción de indicadores sintéticos de este tipo, haciendo referencia a algunos de los trabajos más relevantes.

2. LA MEDICIÓN DEL BIENESTAR EN EL PNUD.

El indicador más utilizado en el ámbito internacional para medir bienestar es el sugerido en 1990 por el Programa de las Naciones Unidas para el Desarrollo (PNUD): el Índice de Desarrollo Humano (IDH).

El IDH trata de combinar los tres aspectos que considera esenciales para que cualquier individuo tenga una vida digna: salud, educación y nivel de ingresos. Su construcción se basa en un enfoque de reducción de brechas, es decir, se observa la reducción de la distancia entre cada variable y el valor máximo que puede alcanzar. De esta forma se consigue asignar una mayor importancia a los logros más difíciles de lograr. Su cálculo se realiza en tres pasos (Informe PNUD: 1990, notas técnicas, pp. 232):

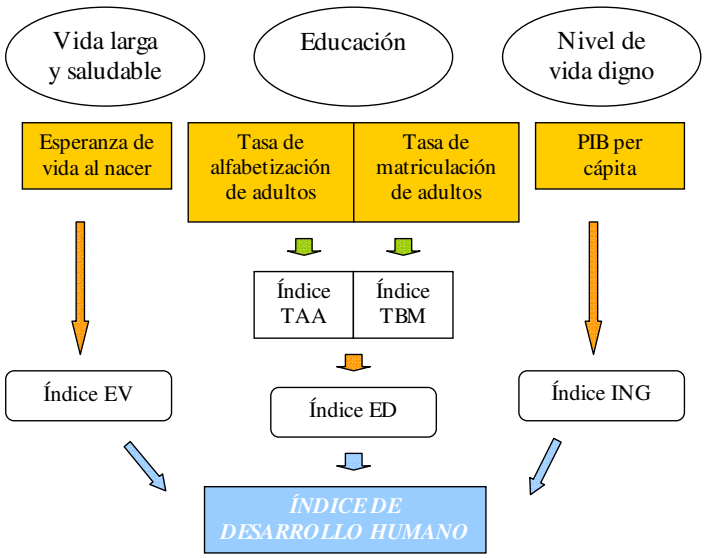
I. Definición de indicador de privación (I_{ij}) para cada país (j) en cada una de sus tres variables básicas: esperanza de vida (X_1), alfabetismo (X_2) y logaritmo del PIB per cápita real (X_3). Se determina un valor máximo y uno mínimo para cada una de las tres variables basándose en los valores reales. Luego, se establece la relación entre el máximo y el valor real y la diferencia entre el máximo y el mínimo, y se clasifica el país en un rango de cero a uno.

$$I_{ij} = \frac{(\max_j x_{ij} - x_{ij})}{(\max_j x_{ij} - \min_j x_{ij})}$$

II. Definición de un indicador promedio (I_j) entre los tres indicadores.

$$I_j = \frac{\sum_{i=1}^{3} I_{ij}}{3}$$

Figura 1 Esquema cálculo IDH



Fuente: Elaboración propia basada en Informe PNUD 2007-08 (2008, pp. 357).

Como ejemplo, en el Informe de Desarrollo Humano de 2005 se calcula el IDH de Sudáfrica para 2003:

I. Cálculo de los índices de cada componente:

a. Índice de esperanza de vida (IEV): mide los logros alcanzados en esperanza de vida. En 2003 el dato de esperanza de vida en el país es de 48,4 años. Los valores máximo y mínimo respectivos quedan establecidos desde 1994 en 85 y 25 años, por lo que el índice quedaría como sigue:

$$\text{IEV} = \frac{48{,}4 - 25}{85 - 25} = 0{,}391$$

b. Índice de educación (IE): mide el progreso relativo en materia de alfabetización de adultos y matriculación bruta combinada de educación primaria, secundaria y terciaria. En primer lugar se calcula el índice de alfabetización de adultos y el índice de la tasa bruta de matriculación. A continuación, se combinan ambos índices ponderando respectivamente como dos tercios y un tercio. Para Sudáfrica, donde la alfabetización de adultos era en 2003 de 82,4% y la tasa bruta combinada de matriculación era de 78%, y una vez considerados como valores extremos el intervalo [0, 100] establecido en 1995, el índice quedaría:

$$\text{Índice de alfabetización de adultos (IAA)} = \frac{82{,}4 - 0}{100 - 0} = 0{,}824$$

$$\text{Índice de TBM (ITBM)} = \frac{78 - 0}{100 - 0} = 0{,}780$$

$$\text{IE} = \frac{2}{3} IAA + \frac{1}{3} ITBM = \frac{2}{3} 0{,}824 + \frac{1}{3} 0{,}780 = 0{,}809$$

c. Índice de nivel de ingresos (II): mide los logros alcanzados en el resto de componentes relativas al desarrollo no contenidas en las anteriores. Dado que una mejora en el desarrollo no requiere un nivel de ingresos ilimitados, el nivel de ingresos medido en términos de PIB per cápita se ajusta utilizando la función logarítmica. Para Sudáfrica, con un PIB per cápita en 10.346 $US (PPA en $US) para 2003 y unos valores extremos fijados desde 1999 en 100 $US y 40.000 $US, el índice sería:

$$\text{II} = \frac{\log 10.346 - \log 100}{\log 40.000 - \log 100} = 0{,}774$$

II. Cálculo del IDH:

Una vez que se han calculado los índices para cada componente, como simple promedio se calcula el IDH.

$$\text{IDH} = \frac{1}{3} IEV + \frac{1}{3} IE + \frac{1}{3} II = \frac{1}{3} 0{,}391 + \frac{1}{3} 0{,}809 + \frac{1}{3} 0{,}774 = 0{,}658$$

Pese a ser el más utilizado en la actualidad, en la medición del bienestar se han empleado multitud de índices además del IDH. En el siguiente apartado se hace referencia a ellos.

3. METODOLOGÍAS DE AGREGACIÓN EN LA CONSTRUCCIÓN DE INDICADORES SINTÉTICOS DE BIENESTAR.

En el contexto de los indicadores sociales, el denominado enfoque agregativo busca la forma de construir una medida única y global que resuma toda la información suministrada por los sistemas de indicadores sociales preexistentes o definidos para su cálculo (Sharpe, 2004). Los estudios a los que nos referiremos en este trabajo que adoptan esta perspectiva tienen así como finalidad la construcción de un indicador sintético útil para la medición del bienestar. En la definición del mismo serán importantes tanto las componentes que lo forman como la metodología que se utilizará para agregarlas. Ambos elementos permitirán diferenciar unos índices de otros.

La definición de indicadores sintéticos exige un proceso complejo que comprende diversas etapas que van desde la delimitación del fenómeno a medir, el establecimiento del marco teórico, la selección de variables a estudiar, etc.; pero sin lugar a dudas, una de las fases más

complicadas la constituye la elección, aplicación y análisis de la metodología de agregación a utilizar.

A continuación, se aborda esta cuestión sin vocación de hacer una lista exhaustiva de trabajos en los que se utilice una u otra técnica sino, más bien, presentando un esquema estructurado de algunas de las metodologías aplicadas más frecuentemente y de forma más consistente en la construcción de indicadores sintéticos de bienestar.

I. Análisis multivariante.

En primer lugar nos centraremos en algunas técnicas de análisis multivariante, que tratan de "*medir, explicar y predecir todas las relaciones existentes entre los elementos que conforman una tabla de datos, proporcionando un resultado que debe ser analizado minuciosamente por el analista*" (Lévy & Varela, 2003, pp. 6).

I.a. Regresión Lineal Múltiple:

El modelo de Regresión Lineal Múltiple plantea la relación existente entre una variable dependiente (y_i) que se desea explicar y un conjunto de variables independientes (x_i) que explicarán el comportamiento de la anterior. La relación entre las componentes no es determinista e incluye cierto grado de aleatoriedad, lo cual se recoge en el modelo a través de la introducción de un término de perturbación ó error (u_i):

$$y_i = \alpha_i + \beta_i x_i + u_i$$

La principal ventaja de esta metodología radica en su sencillez computacional. Por este motivo se trata de uno de los modelos econométricos más utilizados en el ámbito del análisis socioeconómico, fundamentalmente para realizar aproximaciones descriptivas.

En el ámbito del bienestar se pueden encontrar multitud de trabajos, entre ellos podemos destacar los de Ranis et al. (2000), Merz (2002) y Grzeskowizk et al. (2006). En los tres se tienen en cuenta el bienestar , si bien con esta metodología no se busca la propuesta de indicadores sintéticos, como es el caso de la mayor parte de las técnicas siguientes a las que nos referiremos.

El primero analiza las dependencias recíprocas entre crecimiento económico y desarrollo humano en países pertenecientes a distintos territorios (Sudeste asiático, Oriente Medio, África, Latinoamérica y Caribe). Las regresiones entre países muestran una relación significativa en ambas direcciones con el gasto público, salud y educación, así como la tasa de inversión y la distribución de ingresos.

El segundo trabajo considera la incidencia en el bienestar individual de la componente ingresos y otras variables de uso del tiempo, fundamentalmente las horas trabajadas, en el contexto alemán. Es especialmente interesante porque llama la atención sobre la necesidad de incorporar variables de este tipo en el análisis del bienestar que, pese a tratarse de componentes de gran relevancia en el mismo, han sido tradicionalmente olvidadas en su medición (Domínguez, 2007).

Por último, el trabajo de Grzeskowizk et al. plantea el bienestar desde la perspectiva del hogar, basándose en la satisfacción obtenida en función de una serie de variables explicativas como consumo, propiedad, mantenimiento, etc. comparativamente entre EEUU y Corea.

I.b. Análisis de Componentes Principales.

El Análisis de Componentes Principales (ACP)[1] es una técnica de estadística multivariante de la interdependencia, pues en ella todas sus variables tienen una importancia equivalente, desarrollada por Hotelling en 1933[2], que se utiliza tradicionalmente para reducir el número de variables inicial de un análisis, tratando de explicar el mayor porcentaje posible de variabilidad de la muestra con un menor número de variables, que se denominarán *componentes principales* y que serán combinaciones lineales de los datos de origen.

El ACP ha sido una metodología bastante utilizada en la construcción de indicadores sintéticos. Según Krishnakumar y Nagar (2007) la principal razón por la que se ha utilizado esta

[1] Para profundizar en la técnica pueden consultarse Lèvy y Varela (2003) y López (2005), entre otros.
[2] Aunque sus orígenes podemos encontrarlos en los análisis de Pearson (1901) sobre mínimos cuadrados ortogonales.

metodología en relación a la medición del bienestar es por su sencillez computacional. Su ventaja respecto a otras técnicas es que facilita la interpretación del comportamiento de un elevado número de casos si presentan una gran cantidad de características a tener en cuenta. En la medida en que se reduce el número de variables a considerar, la interpretación es más operativa. Asimismo la primera componente recogerá aquellos subindicadores que muestren la máxima variabilidad entre las unidades a comparar mientras que los subindicadores que presentan características similares entre las distintas unidades estarán relegados a componentes menos importantes (Nardo et al., 2005). Por otra parte, el método presenta la ventaja de que siempre proporcionará una solución, sin embargo, esta solución puede generar estimadores muy sesgados a la matriz de cargas factoriales, especialmente si existen variables con comunalidades bajas. En sentido contrario, como cualquier método de reducción de datos, el ACP presenta la desventaja de la pérdida de información. Así, si existiese una dependencia perfecta entre las variables originales se generarían componentes despreciables, puesto que no aportarían nada nuevo a la explicación de la diversidad observada. Sin embargo, en la medida en que no se produzca esta dependencia perfecta, todas las componentes aportarían algo de forma incorrelada con las otras, por tanto, al eliminar alguna, se está perdiendo información que suministran. Algunos otros inconvenientes señalados por Nardo et al. (2005) en relación a esta metodología son: a) La existencia de una fuerte correlación no necesariamente representa una fuerte influencia del subindicador sobre el fenómeno que se estudia; b) Es sensible a las modificaciones en las bases de datos; c) Es sensible a la presencia de valores extremos; d) Es sensible a los problemas que presentan las muestras pequeñas.

El primero en utilizar el ACP para la construcción de indicadores sintéticos relacionados con el bienestar fue Ram (1982) que combinó tres componentes básicas (esperanza de vida, mortalidad infantil y alfabetismo de adultos) con el PIB per cápita para obtener un indicador único e hizo un análisis comparativo entre 187 países. El ACP en el ámbito de los indicadores sintéticos normalmente utiliza un conjunto más amplio de variables de partida, como ocurre en el trabajo de Slottje (1991), que sigue el mismo procedimiento seleccionando 20 atributos de partida en los ámbitos de la dinámica demográfica, educación, nutrición y salud, infraestructuras, estabilidad social y política, instituciones políticas y civiles, ingresos y crecimiento económico y trabajo.

Esta metodología ha sido de las más utilizadas tanto en la medición del bienestar o la pobreza, con trabajos como los de McGillivray (1995), Klasen (2000), Nagar and Basu (2001), Biswas and Caliendo (2002), Rahman et al. (2003), Noorbaksh (2003), en los que se parte de un conjunto amplio de variables entre las que se busca una interrelación que permita establecer medidas adecuadas en diferentes contextos; como en otros ámbitos relacionados como la integración social, donde se puede señalar el trabajo de Klitgaard and Fedderke (1995) que analiza cómo diversas dimensiones como la inestabilidad y conflictos en derechos civiles y políticos, crímenes y violencia afectan de forma desigual a personas ricas y pobres en un conjunto de 116 países; el análisis de la desigualdad regional con estudios como los de Quadrado et al. (1996, 2001) que tratan de analizar la sensibilidad del bienestar a determinados aspectos de carácter económico como el nivel de ingresos y crecimiento económico utilizando variables de salud, educación, ocio, etc. Combinando el ACP con el Análisis Cluster y aplicándolo al contexto español (1996) y húngaro (2001); etc.

I.c. Análisis Factorial.

El análisis factorial es una técnica estadística multivariante, inicialmente propuesta por Pearson a finales del siglo XIX, que pretende reducir el elevado número de variables que describen un fenómeno, de difícil interpretación, a un número de variables menor con el fin de acercarnos a una observación más sencilla del problema. Para ello, trata de comprobar si las variables originales tienen un pequeño número de factores en común que expliquen básicamente lo mismo que las variables de partida.

Algunos indicadores de bienestar que utilizan esta metodología son: el propuesto por Adelman & Morris (1971, 1982) que desde un punto de vista teórico señalan la utilidad de esta metodología en la medición del desarrollo, y en Mishra et al. (2001) que la aplican para estudiar

las relaciones entre el estatus socioeconómico (empleo, vivienda, migración, familia y educación) y el nivel sanitario según sexo y edad en Australia.

I.d. Análisis de Conglomerados o Cluster.

El Análisis de Conglomerados o Análisis Cluster es una técnica cuyo propósito es agrupar un conjunto de observaciones en un número de clusters o grupos, basándose en la idea de distancia o similitud entre las observaciones, en función de unas variables preestablecidas a tal efecto. El objetivo es pues segmentar el conjunto de observaciones originario de manera que se obtengan grupos constituidos por elementos homogéneos entre sí pero con gran heterogeneidad respecto al resto de grupos[3].

La principal ventaja de la técnica es que ofrece una perspectiva distinta y sencilla de la estructura de una base de datos, aunque tiene el inconveniente de que es puramente descriptiva y su proceso de construcción puede resultar confuso.

Algunos trabajos que han utilizado esta de análisis son Hirschberg, Maasoumi & Slottje (1991) en relación a la medición del bienestar, combinando atributos como los derechos políticos, libertades civiles, esperanza de vida, etc.; Pedregal, Torres y Zoido (2006) para la observación de las desigualdades territoriales y el desarrollo en el contexto andaluz; Yang & Hu (2007) que comparan niveles de desarrollo humano en China desde una perspectiva histórica; ó Hernández (2008) que, combinando el Análisis Cluster con otras técnicas, construye un indicador sintético para la comparación del bienestar social de las regiones españolas basándose en variables de salud, educación, renta y exclusión social.

I.e. Análisis Conjunto.

El Análisis Conjunto[4] (Green & Wind, 1975; López, 2005) es una metodología que trata de estudiar el comportamiento de los consumidores según la valoración que éstos realizan de las características (o atributos) de un determinado bien o servicio.

Entre las ventajas de esta metodología encontramos que es posible captar la preferencia a un nivel individual, sin embargo, presenta limitaciones como que el número de características a analizar debe ser muy limitado porque las combinaciones entre ellas aumentan a medida que lo hacen las alternativas y hacen más complejo el análisis. Además, la recogida de información es difícil y costosa.

Podemos encontrar algunos ejemplos de utilización del Análisis Conjunto en el ámbito del bienestar en los trabajos de Ülengin, Ülengin & Güvenç (2001) que estudian la calidad de vida urbana en Estambul a través de las preferencias de sus ciudadanos ó Peterson (2006) que utiliza esta metodología para analizar cómo los individuos asignan pesos a 7 dimensiones referentes a la calidad de vida en EEUU, Francia y Turkía.

II. Análisis multicriterio.

II.a. Análisis Envolvente de Datos.

La programación multiobjetivo constituye un enfoque multicriterio que, ante la imposibilidad de optimizar simultáneamente un conjunto de objetivos, pretende establecer un conjunto de soluciones eficientes.

El Análisis Envolvente de Datos (DEA) fue ideado (Charnes, Cooper & Rhodes, 1978) como herramienta útil para medir la eficiencia técnica en empresas. Los principios básicos de la teoría microeconómica establecen que éstas buscan maximizar sus beneficios. La comparativa entre las empresas pasa por confrontar, de forma conjunta, los productos resultantes del proceso productivo (output) con los insumos utilizados en el mismo (input).

Dado el vector formado por p input (x_1, x_2 , …,x_p) utilizados para la producción de q output (y_1, y_2 , …,y_q) en n unidades productivas, la eficiencia técnica de una determinada unidad 0 podría estimarse:

[3] Para la descripción detallada de la técnica se pueden consultar Hair et al. (1998), Lèvy & Varela (2003) y López (2005).

[4] Tiene sus orígenes en el campo de la psicología (Luce y Tukey, 1964) pero su máximo desarrollo en el ámbito del marketing (Green y Rao; 1971), aunque puede utilizarse, en general, para el análisis del valor asignado por los sujetos en procesos de elección entre varias alternativas.

$$\max_{u_t v_i} \frac{\sum_{f=1}^{q} u_f y_{f0}}{\sum_{i=1}^{p} v_i x_{i0}} \qquad (1)$$

s.a.:

$$\frac{\sum_{f=1}^{q} u_f y_{fj}}{\sum_{i=1}^{p} v_i x_{ij}} \leq 1 \qquad j=1,\ldots, n \qquad (2)$$

donde u_f, $v_i \geq \varepsilon$ (siendo ε un parámetro mayor que cero aunque muy próximo a él) $\forall$ f=1,…,q; i=1,…,p, y (u_1,…, u_q , v_1,…, v_p) es el conjunto de ponderaciones que se desea determinar de manera que maximice la ratio para cada unidad y al tiempo consiga que las ratios de las demás unidades (incluida la unidad evaluada) no superen el valor 1. Este problema se resuelve para cada una de las n unidades.

La utilización de esta técnica para la construcción de indicadores sintéticos -que en este contexto se conoce como *enfoque del beneficio de la duda*-, tiene su máxima representación en los trabajos de Cherchye (2002, 2003, 2006). Esta nueva utilidad parte de la utilización de una serie de subindicadores para la construcción de un indicador que recoja de forma conjunta la aportación de cada uno de estos subindicadores de partida.

El indicador consistiría en la definición de la ratio entre el valor de determinados indicadores de las unidades analizadas y el valor de indicadores establecidos como valor óptimo (benchmarking). La unidad que maximice el valor total del indicador, bajo los pesos específicos de la unidad analizada, será establecida como benchmarking. Asimismo, se seleccionan para cada unidad aquellos pesos que maximizan el valor del indicador, sujeto a un conjunto de restricciones comunes.

Esta descripción es únicamente válida cuando todos los subindicadores tienen un carácter positivo, es decir, cuanto mayor es su valor, mejor es la situación de la unidad analizada. Los trabajos anteriores a los de Cherchye presentan una alternativa distinta, en caso de que exista la necesidad de considerar indicadores tanto positivos como negativos. La formulación, en este caso, se aproxima a la formulación tradicional de DEA, siendo los output los subindicadores positivos y los input los subindicadores negativos. En estos casos el indicador sintético se definiría como la ratio entre una serie de indicadores de carácter "positivo" (que actuarían como output o resultados del proceso a analizar) y una serie de indicadores de carácter "negativo" (que serían los que en los modelos de eficiencia de empresas se consideraban input). Los primeros tendrían un sentido positivo desde la perspectiva de que cuanto mayor fuese el subindicador mejor sería el valor obtenido para el indicador sintético, y lo contrario ocurriría para los negativos, cuanto mayor fuesen, peor sería el indicador sintético obtenido.

La función objetivo (1) representará el indicador sintético cuyo valor es el óptimo para cada unidad (en este supuesto, cada una de las unidades suelen ser unidades territoriales: países, regiones, etc.). Así, al maximizar la función se busca el conjunto de ponderaciones que otorgan el mayor valor posible del indicador sintético a esa región. Las restricciones (2) permiten comparar cada región con el resto de regiones que presentan los mejores valores, que obtienen el valor 1 para el indicador mientras las demás obtienen un valor inferior. Será necesario repetir la operación tantas veces como indicadores se deseen construir. El objetivo final es el establecimiento de un ranking en función de los resultados obtenidos para el indicador.

En el ámbito de análisis del bienestar esta técnica ha tenido una creciente incorporación. Algunos indicadores propuestos son los de Hashimoto & Ishikawa (1993) y Hashimoto & Kodama (1997) en el campo de la medición de la situación social en Japón; Storrie & Bjurek (2000) que construyen un indicador sintético y lo aplican al mercado de trabajo de la UE; Mahlberg & Obersteiner (2001) que construyen un indicador sintético alternativo al IDH utilizando sus mismas componentes; Zhu (2001) que aplica la metodología a la estimación de un

indicador de nivel de calidad de vida; Cherchye et al. (2003, 2006) que propone un indicador relativo al ámbito tecnológico; Martínez, de Miguel y Murías (2005) que aplican al ámbito español la metodología en relación a determinados aspectos socioeconómicos y Murias, Martínez y de Miguel (2006) que tratan de construir un índice sintético de bienestar basándose en los niveles de ingreso y desempleo en el territorio español; Smith & Street (2005) que desarrollan una medida para la eficiencia de los servicios públicos, etc.

II.b. Programación por metas.

El enfoque anterior plantea el problema de que proporciona una serie de indicadores que permite establecer un ranking, pero éste no es comparable con los subindicadores originales, dado que cada unidad adopta sus propias ponderaciones.

Surge así la necesidad de desarrollar una nueva metodología que permitiera el establecimiento de unos pesos comunes de forma que el valor del indicador global obtenido se aproximara lo máximo posible a la frontera eficiente. Es decir, las unidades eficientes serían aquellas que mantuviesen su eficiencia intacta al considerar pesos comunes. De esta manera, se desarrollaría un método para discriminar entre unidades eficientes, y no únicamente para separar las eficientes de las que no lo son (Despotis, 2004).

Destaca en el ámbito que nos ocupa el trabajo de Despotis (2004) que compara entre el IDH y un nuevo indicador de medición de desarrollo humano creado según la metodología descrita aplicada a los mismos subindicadores de partida.

II.c. Proceso Analítico Jerárquico.

El enfoque del Proceso Analítico Jerárquico (AHP) es una técnica de decisión multicriterio discreta, a diferencia de los dos procedimientos anteriores que son continuos, aunque pueden ser utilizados igualmente en problemas discretos. Fue creada por Saaty (1980) y representa la escuela americana dentro del enfoque de modelos de decisión multicriterio discretos[5].

Algunos trabajos en el ámbito de los indicadores sintéticos son el de Blancas, Domínguez y Guerrero (2006) en el terreno de la medición del bienestar con perspectiva de género en las regiones españolas; Liu et al. (2007) que analizan comparativamente la sostenibilidad del desarrollo en Canadá y Australia ó Bonini (2007) en el ámbito de la satisfacción en relación a la salud, desarrollo humano y condiciones ambientales en un conjunto de países de distintos continentes.

II.d. Teoría de la Utilidad Multiatributo.

Se trata de un enfoque discreto, al igual que el anterior, desarrollado por Keeney y Raiffa (1976), que busca expresar las preferencias del centro decisor en términos de la utilidad que le reporta, dentro de un contexto de la teoría de la decisión en condiciones de incertidumbre. Esta propuesta pasa por la definición de una función de utilidad multiatributo agregada según una serie de funciones de utilidad individuales definidas previamente.

Algunos autores consideran que la utilidad multiatributo es un enfoque con un gran fundamento teórico (Ballestero y Romero, 1998), sin embargo, presenta el inconveniente de la dificultad del cálculo de las funciones de utilidad, en caso de que el número de alternativas sea elevado, así como la presunción de que el centro decisor es racional a la hora de atribuir probabilidades a los diferentes estados del entorno, asignar un valor de utilidad a cada estado y comparar alternativas (Bell et al., 1988), lo cual es cuestionado por diferentes autores (March, 1978, Thaler y Russel, 1988).

Entre estos trabajos podemos encontrar el de Barrera-Roldán y Saldívar-Valdés (2002), que utilizan 21 indicadores de carácter económico, ambiental y social en la región mejicana de Coatzacoalcos para diseñar un índice de desarrollo sostenible; y Munda (2005) que, en el mismo sentido, justifica teóricamente la utilidad de este enfoque en la agregación de indicadores de índole económica, social y ambiental para la construcción de un índice de desarrollo sostenible. En cuanto a bienestar en términos exclusivamente sociales, es interesante el proyecto de Pérez

[5] La escuela europea está representada en cuanto a modelos de decisión multicriterio discretos por Roy, B. (1968) que desarrolla los denominados métodos ELECTRE. Pese a la importancia de dichos modelos, no serán desarrollados en este trabajo por la escasa utilización de los mismos en el contexto de construcción de indicadores sintéticos en que nos desenvolvemos.

(2006) cuyo propósito es evaluar la marcha de distintos países en relación al cumplimiento de los Objetivos del Milenio mediante la construcción de indicadores al efecto.

III. Agregaciones simples basadas en normalización de subindicadores.

Una parte muy importante de los trabajos sobre construcción de indicadores sintéticos utilizan métodos simples de agregación, diferenciándose únicamente en la metodología utilizada para normalizar los subindicadores que los componen. La característica común a todos ellos es que parten de la base de que las unidades de medida de los subindicadores deben ser las mismas, por lo que estos subindicadores deben ser normalizados como paso previo. Las diferencias en las técnicas de normalización, así como la manera de ponderar los subindicadores, dan como resultado un conjunto de métodos entre los que se pueden diferenciar dos grandes grupos.

III.a. Igualdad en la ponderación.

Consiste básicamente en asignar a cada subindicador el mismo peso que a los demás, agregando la información mediante una suma. La ponderación y agregación suele hacerse en niveles sucesivos de manera que previamente se ponderan y agregan una serie de variables para construir los subindicadores relativos a una determinada dimensión y, posteriormente, se agregan éstos para construir el indicador sintético.

La principal desventaja que presenta este método es que ignora las relaciones causales existentes entre los subindicadores, lo que provocaría problemas en caso de que se incorporasen variables con un alto grado de correlación, pues se estaría contabilizando doblemente el efecto.

Algunos de los índices más empleados a nivel internacional en la medición del desarrollo, como el IDH y el IDG, utilizan esta metodología, principalmente por su sencillez en el cálculo.

III.b. Otras ponderaciones.

Entre los procedimientos que asignan distinta ponderación podemos encontrar los citados por Saisana & Tarantola (2002):

- Método 1:

Consiste en el establecimiento de ranking entre las alternativas para cada subindicador utilizando la regla de Borda. La forma de agregación es una suma de los ranking obtenidos.

$$I_a^t = \sum_{i=1}^{N} \text{Rank}_{ia}^t$$

donde I_a^t representa el valor del indicador para la alternativa *a* en el momento *t*, e *i* hace referencia a cada uno de los subindicadores.

Aunque no en el ámbito puramente social, utilizan este método por ejemplo Fagerberg (2001) en la construcción del *Índice de Información y Tecnologías de la Comunicación* donde combina 5 subindicadores (número de teléfonos móviles, usuarios de Internet, ect.) calculando el ranking de cada país para cada uno de los subindicadores y sumando a continuación este ranking; y Jencks et al. (2003) para estudiar la evolución experimentada entre 1998-1999 y 2000-2001 en el ámbito del cuidado médico para personas mayores de 65 años sobre la base de 22 indicadores.

- Método 2:

El procedimiento de agregación es la suma de los subindicadores una vez han sido normalizados, utilizando para ello la diferencia entre el número de indicadores que están por encima y por debajo de un umbral arbitrario (p) definido por el investigador alrededor de la media.

$$I_a^t = \sum_{i=1}^{N} \text{sgn} \left(\frac{x_{ia}^t}{x_{i\bar{a}}^t} - (1+p) \right)$$

donde I_a^t representa el valor del indicador para la alternativa *a* en el momento *t*, x_{ia}^t es el valor del subindicador *i-ésimo* para la alternativa *a* en el momento *t*, $x_{i\bar{a}}^t$ es el valor del subindicador *i-ésimo* para la media de los subindicadores en el momento *t*.

Es usado en la construcción del *Summary Innovation Index* (DG Enterprise, 2001) que mide convergencia europea en materia de innovación mediante la utilización de 17 subindicadores

pertenecientes a 4 áreas (fuentes humanas, creación de conocimiento, transmisión y aplicación de nuevo conocimiento e innovación financiera y en los mercados). Así, para cada país el índice será igual al número de subindicadores que se sitúen por encima del 20% de la media del total de países menos el número de subindicadores que supere el 20% por debajo de la media.

- Método 3:

La normalización en este método consiste en la construcción de una ratio entre el valor del subindicador para cada unidad (x_{ia}^t) con respecto al valor del subindicador para otra unidad que se toma como referencia (x_{iR}^t). Esta alternativa que se toma como referencia puede ser la que se sitúa en la media del conjunto o alguna otra que cumpla una característica determinada, como que sea aquella que presenta los valores para el subindicador a los que todas las demás alternativas aspiran. En cualquier caso, se deja al criterio del analista.

$$y_{ia}^t = x_{ia}^t / x_{iR}^t$$

La construcción del indicador sintético sería la agregación ponderada de estos subindicadores una vez normalizados.

$$I_a^t = \frac{\sum_{i=1}^{N} w_i y_{ia}^t}{\sum_{i=1}^{N} w_i} \qquad (1)$$

donde I_a^t representa el valor del indicador para la alternativa *a* en el momento *t*, y w_i es el peso dado al subindicador *i-ésimo*.

En caso de que los pesos asignados a los indicadores sean unitarios, se trataría de una media aritmética y estaríamos en la situación de igualdad en la ponderación referida anteriormente, donde se sitúan gran parte de los índices de bienestar más comunes como, por ejemplo, el *Índice de Capacidades Básicas* (Social Watch, 2000) que agrega porcentaje de niños en quinto nivel de primaria, supervivencia hasta los 5 años de edad y porcentaje de partos asistidos por personal especializado. Así encontramos también el IDH o el *Índice de Bienestar Económico* (Osberg & Sharpe, 2002). No obstante, pese a que la ponderación para la construcción del índice final es la misma, no ocurre lo mismo para la formación de los subindicadores que constituyen cada índice global. Así, por ejemplo, la componente educativa del IDH se compone de tasa de alfabetización de adultos y tasa bruta de matriculación combinada, ponderadas, respectivamente, con 2/3 y 1/3. Lo mismo ocurre con el propuesto por Osberg & Sharpe, que ponderan de forma distinta sus componentes internas.

- Método 4:

La construcción del índice es similar al anterior (1) pero los valores de los subindicadores son sustituidos por la diferencia de valores entre el año analizado y el anterior y dividido por el valor del año actual:

$$y_{ia}^t = \frac{(x_{ia}^t - x_{ia}^{t-1})}{x_{ia}^t}$$

donde x_{ia}^t es el valor del subindicador *i-ésimo* para la alternativa *a* en el momento *t*.

Un ejemplo lo encontramos en el *Índice de Mercado Interno* (Tarantola et al., 2002) que trata de medir el grado de cumplimiento de beneficios (tales como incremento de ingresos, bajada de precios, posibilidades de trabajo, etc.) en función de las estrategias de mercado utilizando la metodología citada para construir 19 variables que, posteriormente, son tratadas mediante PCA.

- Método 5:

La construcción del indicador, al igual que los anteriores, sigue la estructura de la ecuación (1) pero, en este caso se basa en las puntuaciones estandarizadas para cada subindicador como la diferencia entre el valor de la alternativa y el valor medio del conjunto, dividida por el error estándar.

$$y_{ia}^t = \frac{x_{ia}^t - x_{\bar{a}i}^t}{\sigma_{\bar{a}}},$$

Donde x_{ia}^t es el valor del subindicador *i-ésimo* para la alternativa *a* en el momento *t* y el subíndice $\bar{a}$ representa la media aritmética.

Es uno de los más usados en la construcción de indicadores. Entre ellos podemos encontrar los trabajos del World International Forum (2001) para la construcción del *Environmental Sustainability Index* que es utilizado por 122 países para medir su grado de sostenibilidad medioambiental. El índice se compone de un conjunto de 22 subindicadores cada uno de los cuales incluye entre 2 y 6 variables (67 en total), que son normalizados según la metodología descrita y agregados mediante el sistema de igualdad en la ponderación.

- Método 6:

El último de los métodos, que también construye su índice con la ecuación (1), es similar al anterior pero usa un cambio de escala en los valores de los subindicadores. El resultado es que las puntuaciones estandarizadas de todos los subindicadores presentan el mismo rango de variación (0,1).

$$y_{ia}^t = \frac{x_{ia}^t - \min_a x_i^t}{\max_a x_i^t - \min_a x_i^t}$$

donde x_i^t es el valor del subindicador *i-ésimo* en el momento *t*.

Este cambio de escala consigue añadir robustez al método, sin embargo, esta característica introduce el problema contrario, es decir, el rango de los subindicadores con pequeñas variaciones se ve incrementado.

IV. Otros.

Finalmente, nos referiremos a otras técnicas que, si bien no tienen como finalidad principal la construcción de indicadores sintéticos, también han sido utilizadas con este propósito. Entre ellas destacaremos tres: Conjuntos Difusos (Fuzzy Sets), Opinión Pública y Budget Allocation y métodos basados en la medición de distancias.

IV.a. Fuzzy sets.

El análisis del fenómeno del bienestar es complejo pues la observación de los comportamientos humanos y el entorno social que los rodea exige la consideración de múltiples variables que no siempre son cuantificables o que, de serlo, admiten gradación. Es decir, en un sistema en que la incertidumbre está presente, la medición en términos absolutos que nos permite la lógica clásica se presenta como insuficiente. Esta idea es la que llevó a Zadeh (1965) a desarrollar los conjuntos difusos (fuzzy sets) ó la lógica difusa (fuzzy logic)[6], que permite cuantificar esta incertidumbre.

En el contexto en que nos movemos, el proceso trata de inferir una solución a un problema que no puede ser observado directamente a partir de los aspectos del problema que pueden ser observados directamente (subindicadores).

La mayor ventaja de esta metodología es la precisión y rapidez en la implementación puesto que utiliza método heurístico, así como la posibilidad que brinda de analizar datos que por su propia naturaleza son escasos o imprecisos y marcadamente no lineales. Su principal inconveniente su complejidad computacional.

La consideración de estas técnicas en la construcción de indicadores sintéticos es bastante reciente. Algunos de los trabajos más interesantes a este respecto son los desarrollados por Cornelissen, Van den Berg, Koops, Grossman, & Udo (2001); Phillips and Adriantiatsaholiniaina (2001) y Andriantiatsaholiniaina, Kouikoglou & Phillis (2004) en relación a la medición del desarrollo sostenible. En relación a la medición del bienestar Baliamoune & McGillivray (2006) aplican esta metodología a las mismas componentes del IDH y establecen una comparación entre ambos indicadores para 14 países asiáticos; Bérenguer & Verdier-Chouchane (2007, 2008) que, respectivamente en ambos trabajos, elaboran índices de calidad de vida utilizando diversas variables en las dimensiones de salud, educación, nivel económico y medioambiente en África, y construyen índices de calidad de vida en las mujeres atendiendo a su faceta de madres. Asimismo, existen trabajos como el de Tauchert (2002) que

[6] Ver Carlsson & Fullér (2002).

aborda desde un punto de vista teórico las posibilidades de la metodología fuzzy en el estudio de bienestar desde una perspectiva de género y el de Chiappero (2006) y Addabbo et al. (2004) que señala sus posibilidades en cuanto al enfoque de las capacidades.

IV.b. Opinión Pública y Budget Allocation.

Este tipo de metodología se basa en la aportación subjetiva realizada por expertos o por la opinión pública sobre la temática de los problemas que se trata de resolver.

Según Nardo (2005), en el caso del Budget Allocation el primer paso consiste en seleccionar un grupo de expertos para realizar la evaluación pertinente. El grupo de expertos, a través de técnicas de discusión participativa, establecen sus opiniones acerca de determinados aspectos. En concreto, acerca de determinados subindicadores relativos a los fenómenos a los que se refiere, a los que además asignan pesos en función de su idoneidad, importancia, etc. En este caso los expertos deben repartir un "presupuesto" de N puntos entre los subindicadores, de manera que tendrán que "pagar" más por aquellos indicadores cuya importancia quieran destacar. Para identificar los indicadores más adecuados y calcular sus pesos se utiliza el conocimiento de expertos aplicando la técnica Delphi, pudiendo establecerse pasos sucesivos hasta que se alcanza la convergencia en caso de discrepancia inicial. Finalmente, resta realizar una agregación para la obtención de la medida sintética, que suele seguir un esquema multiplicativo.

La participación de la Opinión Pública, por su parte, exige que las personas que intervienen manifiesten su grado de conformidad o no en relación a determinados problemas propuestos medidos a través de una serie de subindicadores.

No tenemos constancia de indicadores que midan estrictamente bienestar, aunque sí se aplican en aspectos relacionados con el mismo.

IV.c. Distancias.

Finalmente, la última de las metodologías a la que haremos referencia es el cálculo de indicadores a través de la medición de las distancias entre la situación de partida y los objetivos que se pretende conseguir. De esta manera se detecta la mayor o menor urgencia de actuación en un determinado aspecto en función de la mayor o menor distancia entre el estado en que se encuentra un fenómeno concreto y la situación a la que se desea llegar.

En esta metodología los pesos de cada subindicador se obtienen como la ratio (o como la diferencia) entre el valor del subindicador y el valor del correspondiente objetivo, expresados en la misma unidad. Posteriormente estos parámetros pueden ser agregados mediante una simple media aritmética para generar el indicador sintético.

Existen multitud de medidas de distancia (Pena Trapero, 1977, pp. 65-71 y 91 y siguientes): distancia CRL de Pearson, distancia de Frechet, distancia Generalizada de Mahalanobis, distancia de Stone, distancia-I de Ivanovic, etc. En el ámbito de los indicadores sintéticos de bienestar, la más utilizada es la desarrollada por Pena Trapero (1977) que consiste básicamente en una modificación del indicador de distancia de Ivanovic, aunque con un sistema de ponderaciones distinto basado en el coeficiente de determinación. También los trabajos de Zarzosa (1996, 2005) aportan una interesante perspectiva al respecto. A nivel internacional, algunos ejemplos de indicadores sintéticos basados en medidas de distancia son los desarrollados por el PNUD.

4. CONCLUSIONES.

El análisis del bienestar desde la óptica de los indicadores sintéticos ha experimentado un gran auge en los últimos años a nivel internacional. La utilización de distintas metodologías de agregación y de distintas componentes como variables de partida son los principales elementos que diferencian unos indicadores de otros.

Tras la revisión de las distintas técnicas queda pendiente el problema de cuál de ellas se adapta mejor a los datos de partida de cada análisis. Este es un problema sin resolver, la elección entre los distintos métodos de agregación es difícil y depende del objetivo del índice que se pretende construir, de las características de los subindicadores e incluso del coste computacional que el investigador pueda asumir.

REFERENCIAS BIBLIOGRÁFICAS.

ADDABBO, T., DI TOMMASO, L. & FACCHINETTI, G. (2004), "To what extend fuzzy set theory and structural equation modelling can measure functionings? An application to child well-being", Child Working Papers WP 03-04.

ADELMAN, I. & TAFT MORRIS, C. (1971), "Analysis-of-variance techniques for the study of economic development", Journal of Development Studies, Vol. 8, Issue 1, pp. 90-105.

ADELMAN, I. & MORRIS, CT (1982), "Factor analysis and development: a reply", Journal of Development Economics, Vol. 11, pp. 109-112 and pp. 113-123.

BALIAMOUNE-LUTZ, M. & MCGILLIVRAY, M. (2006): "Fuzzy well-being achievement in Pacific Asia", Journal of the Asia Pacific Economy, Vol. 11, Issue 2, pp. 168-177.

BALLESTERO, E. Y ROMERO, C. (1998), "Multiple criteria decisión making and its applications to economic problems", Kluwe Academic Publishers, Holanda.

BELL, DE: ET AL. (1988), "Descriptive, normative, and prescriptive interactions in decision making", in Bell et al. (Eds.), Decision making, Cambridge University Press, Cambridge, pp. 9-30.

BÉRENGUER, V. & VERDIER-CHOUCHANE, A. (2007): "Multidimensional measures of well-being: standard of living and quality of life acrros countries", World Development, Vol. 35, Issue 7, pp. 1259-1276.

???——. (2008): "From gender equality to Women´s quality-of-life index", Communication for the Annual Conference of Human Development and Capabity Association, New Delhi 10-13 September 2008.

BISWAS, B. & CALIENDO, F. (2002), "A multivariate analysis of the human development index", Indian Economic Journal, Vol. 49, Issue 4, pp. 96–100.

BLANCAS, FJ., DOMÍNGUEZ, M. Y GUERRERO, F. (2006), "Desigualdad regional de la mujer: un método de análisis alternativo", Anales de Economía Aplicada, Vol. 2008, pp. 164-165.

BONINI, A. (2007), "Cross-national variation in individual life satisfaction: effects of nacional wealth, human development and environmental conditions", Social Indicators Research, Accepted: 7 July.

CHARNES, A., COOPER, WW. & RHODES, E. (1978), "Measuring the efficiency on Decision Making Units", European Journal of Operational Research, Vol. 2, pp. 429-444.

CHERCHYE, L. & KUOSMANEN, T. (2002), "Benchmarking Sustainable Development: A synthetic meta-index approach", Working Paper, World Institute for Development Economic Research, Helsinki.

CHERCHYE, L., MOESEN, W. & PUYENBROECK, T.V. (2003), "Legitimately diverse, yet comparable: On synthesising social inclusion performance in the EU", Working Paper, World Institute for Development Economic Research, Helsinki.

CHERCHYE, L., MOESEN, W., ROGGE, N., PUYENBROECK, T.V., SAISANA, M., SALTELLI, A., LISKA, R. &TARANTOLA, S. (2006), "Creating Composite Indicators with DEA and Robustness Analysis: the case of the Technology Achievement Index", Public Economics Working Paper Serie, N. ces0613, Centrum voor Economische Studiën, Katholieke Universiteit Leuven.

CHIAPPERO, E. (2006): "Capability Approach and fuzzy set theory: description, aggregation and inference issues", in Springer (Eds.): Fuzzy Set Approach to Multidimensional Poverty Meassurement, Springer, US, pp. 93-113.

CORNELISSEN, AMG.; VAN DEN BERG, J.; KOOPS, WJ.; GROSSMAN, M. & UDO, HMJ. (2001): "Assessment of the contribution of sustainability indicators to sustainable development: a novel approach using fuzzy set theory", Agriculture, Ecosystems & Environment, Vol. 86, pp. 173-185.

DESPOTIS, DK. (2004), "A reassessment of the human development index via Data Envelopment Analysis", Journal of the Operational Research Society, pp. 314-323.

DG ENTERPRISE (2001), Summary Innovation Index.

DOMÍNGUEZ, M. (2007), "La variable tiempo: análisis de las diferencias de género en la estructura socioeconómica de Andalucía", Anales de Economía Aplicada, Vol. 2007, pp66-82.

FAGERBERG, J. (2001), "Europe at the crossroads: the challenge from innovation-based growth" in Archibugi, D. & Lundvall, B. Å. (Eds.): The Globalizing Learning Economy, Oxford University Press, New York, pp. 45-60.

GREEN, P. & WIND, Y. (1975), "New way to measure consumers´ judgments", Harvard Bussines Review, July-August, pp. 107-117.

GRZESKOWIZK ET AL. (2006), "Housing well-being: developing and validating a measure", Vol. 79, pp. 503-541.

HAIR ET AL. (1998), "Multivariate Data Analysis", Prentince-Hall International, London.

HASHIMOTO, A. & ISHIKAWA, M. (1993), "Using DEA to evaluate the state of society as measured by multiple social indicators", Socio-Economic Planning Sciences, Vol. 27, pp. 257-268.

HASHIMOTO, A. & KODAMA, M. (1997), "Has livability of Japan gotten better for 1956-1990?: A DEA approach", Social Indicators Research, Vol. 40, pp. 359-373.

HERNÁNDEZ, A. (2008), "El mapa de bienestar social: propuesta metodológica para comparaciones regionales. Una aplicación a las comunidades autónomas españolas", Contribuciones a la Economía, Vol. mayo.

HIRSCHBERG, J., MAASOUMI, E. & SLOTTJE, D. (1991), "Cluster analysis for measuring welfare and quality of life across countries", Journal of Econometrics, Vol. 50, pp. 131-150.

JENCKS, SF.; HUFF, ED. & CUERDON, T. (2003), "Change in the quality of care delivered to Medicare beneficiaries, 1998-1999 to 2000-2001", Journal of the American Medical Association, Vol. 289, Issue 3, pp. 12.

KLASEN, S. (2000), "Measuring poverty and deprivation in South Africa", Review of Income and Wealth, Vol. 46, pp.33–58.

KLITGAARD, R. & FEDDERKE, J. (1995), "Social integration and disintegration: an exploratory analysis of cross-country data", World Development, Vol. 23, Issue 3, pp. 357-369.

KRISHNAKUMAR, J. & NAGAR, AL. (2007), "On exact statistical properties of multidimensional indices based on Principal Components , Factor Analysis, MIMIC and Estructural Equations Models", Social Indicators Research, Vol. 86, Issue 3, pp. 481-496.

LÈVY, JP & VARELA, J (2003), "Análisis multivariable para las ciencias sociales", Pearson: Prentice Hall, Madrid.

LIU, P. & TSAI, C. (2007), "A comparative analysis of Sustainable Fishery Development Indicator System in Australia and Canada", Sustainable Development, Vol. 15, pp. 28-40.

LÓPEZ, C. (2005), "Métodos estadísticos avanzados con SPSS", Thomson, Madrid.

MAHLBERG, B. & OBERSTEINER (2001), "Remeasuring the HDI by Data Envelopment Analysis", Report 01-069, International Institute for Applied Systems Analysis Interim. Austria.

MARTÍNEZ, F. MIGUEL DE, J.C. Y MURIAS, P. (2005), "El análisis envolvente de datos en la construcción de indicadores sintéticos. Una aplicación a las provincias españolas", Estudios de Economía Aplicada, Vol. 23, Issue 3, pp. 753-771.

MCGILLIVRAY, M. (1995), "Measuring non-economic well-being achievement", Review of Income and Wealth, Vol. 51, Issue 2, pp. 337–364.

MERZ, J. (2002), "Time and economic well-being –a panel analysis of desired versus actual working hours", Review of Income and Wealth, Vol. 48, Issue 3, pp. 317-346.

MISHRA, G. ET AL. (2001), "The measurement of socio-economic status: investigation of gender-and age-specific indicators in Australian National Health Survey", Social Indicators Research, Vol. 56, pp. 73-89.

MUNDA, G. (2005), "Measuring sustainability: a multi-criterion framework", Environment, Development and Sustainability, Vol. 7, Issue 1, pp. 117-134.

MURIAS, P.; MARTÍNEZ, F. Y DE MIGUEL, JC. (2006), "An economic wellbeing index for the Spanish provinces: a Data Envelopment Analysis approach", Social Indicators Research, Vol. 77, Issue 3, 395-417.

NAGAR, A. L., & BASU, S. (2001), "Weighting socio-economic indicators of human development (a latent variable approach)", National Institute of Public Finance and Policy, New Delhi.

NARDO, M.; SAISANA, M. SALTELLI, A., TARANTOLA, S., HOFFMAN, A. & GIOVANNINI, E. (2005), "Handbook of constructing composite indicators: methodology and user guide", Study Documents 2005-3. OCDE.

NOORBAKSH, F. (2003), "Human development and regional disparities in India", Discussion Paper, UN-WIDER, Helsinki.

OSBERG, L. & SHARPE, A. (2002), "An Index of Economic Well-being for selected OECD countries", Review of Income and Wealth, Vol. 48, Issue 3, pp. 291-316.

PEDREGAL, B.; TORRES, FJ & ZOIDO, F. (2006), "Propuesta metodológica para la medición del desarrollo y las desigualdades territoriales. Aplicación al territorio andaluz", Revista Electrónica de Geografía y Ciencias Sociales, Vol. X, Nº 220.

PENA, B. (1977): Problemas en la medición del bienestar y conceptos afines. Una aplicación al caso español. INE. Madrid.

PÉREZ, CO. (2006), "Equidad, Pobreza y Bienestar Social: construcción de indicadores con un enfoque multiatributo", Proyecto de investigación y desarrollo tecnológico, Universidad Nacional de Córdoba, Argentina.

PETERSON, M. (2006), "Identifying quality-of-life priorities for societal development: using a market orientation to benefit citizens", Journal of Macromarketing, Vol. 26, Issue 1, pp. 45-58.

PNUD: "Informe sobre el Desarrollo Humano". Oxford University Press. New York. (Informes 1990, 2005 y 2007-2008).

PHILLIS, A. ANDRIANTIATSAHOLINIAINA, L. (2001): "Sustainability: an ill-defined concept and its assessment using fuzzy logic", Ecological Economics, Vol. 37, pp. 435-456.

QUADRADO, L. Y LÓPEZ-BAZO (1996), "Macroeconomic variables and regional wellfare. An application to the Spanish regions", European Regional Science Association, 36th European Congress, 26-30 August.

QUADRADO, L.; HEIJMAN, W & FOLMER, H. (2001), "Multidimensional análisis of regional inequality: the case of Hungary", Social Indicators Research, Vol. 56, pp. 21-42.

RAHMAN, T. ET AL. (2003), ·"Measuring the quality of life across countries: A sensitivity analysis of well-being indices", WIDER International Conference on Inequality, Poverty and Human Well-being, Helsinki, Finland.

RAM, R. (1982), "Composite indices of physical quality of life, basic needs fulfilment and income", Journal of Development Economics, Vol. 11, pp. 227-247.

RANIS, G.; STEWART, F. & RAMIREZ, A. (2000): "Economic growth and human development", World Development, Vol. 28, Issue 2, pp- 197-219.

ROY, B. (1968), "Classement et choix en presence de points de vue multiples (la methode Electre), Rev. Fr. Inf. Rech. Oper., Vol. 2, Issue 8, pp. 57-75.

SAATY, T. (1980), "Multicriteria Decision Making: The analytic Hierarchy Process.",McGraw-Hill, New York.

SAISANA, M. & TARANTOLA, S. (2002), "State of the Report on Current Methodologies and Practices for Composite Indicator Development", Joint Research Center, European Commission.

SHARPE, A (2004), "Literature review of frameworks macro-indicators", Centre for the Study of Living Standards, Ottawa, Ontario.

SLOTTJE, D. (1991), "Measuring the quality of life across countries", The Review of Economics and Statistics, Vol. 73, Issue 4, pp. 684-693.

SMITH, P-. & STREET, A. (2005), "Measuring the efficiency of public services: the limits of analysis", Journal of Research Statistic, Vol. 168, Part 2, pp. 401-417.

SOCIAL WATCH (2000): Social Watch Annual Report 1999. (http://www.socialwatch.org/en/informeImpreso/TablaDeContenidos1999.htm) (03/02/2008).

STORRIE, D. & BJUREK, H. (2000), "European Lobour Market Performance with Efficiency Frontier Techniques", Discussion Paper Nº FS I 00-211, Wissenschaftszentrum Berlin für Sozialforschung.

TARANTOLA, S.; SAISANA, M. & SALTELLI, A. (2002), "Internal Market Index 2002: technical details of the methodology", Joint Research Center, EU. Italy.

THALER, R. & RUSSEL, T. (1988,: "The relevance of quasi rationality in competitive markets", en Bell et al. (Eds.): Decision making, Cambridge University Press, Cambridge, pp. 508-524.

TAUCHERT, A. (2002): "Fuzzy gender: between female-embodiment and intersex", General of Gender Studies, Vol. 11, Issue 1, pp. 29-38.

ÜLENGIN, B.; ÜLENGIN, F. & GÜVENÇ, U. (2001), "A multidimensional approach to urban quality of life: the case of Istanbul", European Journal of Operational Research, Vol. 130, pp. 361-374.

WORLD INTERNATIONAL FORUM (2001): Environmental Sustainability Index (http://sedac.ciesin.columnbia.edu/es/esi/rank_01.html) (03/05/2007).

YANG, Y. & HU, A. (2007), "Investigatin regional disparities of China´s human development with Cluster Analysis: a historial perspective", Social Indicators Research, Accepted 13 August.

ZADEH, L. (1965), "Fuzzy sets", Information and Control, Vol. 8, pp. 338-353.

ZARZOSA, P. (1996): Aproximación a la medición del bienestar social. Secretariado de publicaciones e intercambio científico de la universidad de Valladolid. Valladolid.

—'??—. (2005): La calidad de vida en los municipios de la provincia de Valladolid. Diputación de Valladolid. Valladolid.

ZHU (2001), "Multidimensional quality-of-life measure with an application to Fortune´s best cities", Socio-Economic Planning Sciences, Vol 35, pp. 263-284.

PROYECCIÓN DE LA DISCAPACIDAD EN ESPAÑA A PARTIR DE UN MODELO DE ESTADOS MÚLTIPLES

RAFAEL HERRERÍAS PLEGUEZUELO
JOSÉ CALLEJÓN CÉSPEDES
FRANCISCO JAVIER BLANCO ENCOMIENDA
E-mail: jble@ugr.es
Universidad de Granada

Resumen

En España, el incremento continuado en el número de personas con alguna discapacidad ha motivado un gran interés desde una perspectiva tanto social como económica.
Una de las cuestiones más importantes es estimar cuántas personas con discapacidad puede haber a corto o medio plazo. A este respecto, en el presente trabajo se muestra un método para proyectar el número de personas en dicha situación, en base a un modelo de estados múltiples y al avance de resultados de la recientemente publicada *Encuesta de Discapacidad, Autonomía Personal y Situaciones de Dependencia* (2008).
La proyección de las tasas de discapacidad puede constituir el punto de partida para diseñar iniciativas en favor de este colectivo.

Palabras clave: discapacidad, proyección, modelo de estados múltiples.
Área temática: Economía Social y Discapacidad.

Abstract

In Spain, the continued increase in the number of people with some disability has motivated a great interest from a social and economic perspective.
One of the most important issues is to estimate how many people will have a disability at short or medium term. In this respect, in the present paper a method appears to project the number of people who will be in that situation, in according to a multiple state model and to the advance of results of the recently published *Survey of Disability, Personal Autonomy and Situations of Dependency* (2008).
The projection of the disability rates can constitute the starting point to design initiatives in favour of this group.

Key words: disability, projection, multiple state model.
Thematic Area: Social Economics.

1. INTRODUCCIÓN

Las personas con algún tipo de discapacidad forman un sector cada vez más numeroso en nuestra sociedad, de ahí que el interés por conocer la situación de este colectivo esté creciendo de manera tan notable desde hace algunos años. Este interés queda patente en numerosas iniciativas de política social y económica emprendidas, las cuales se están convirtiendo en uno de los grandes retos de las sociedades modernas.

Han sido varias las medidas adoptadas en torno al fenómeno de la discapacidad, sobre todo, desde la declaración del 2003 como "*Año Europeo de las Personas con Discapacidad"*. En el caso de España, se pueden citar por ejemplo el II Plan Estatal de Acción para las Personas con Discapacidad 2003-2007 (de cuyas funciones de impulso, coordinación y evaluación se encargó el entonces Ministerio de Trabajo y Asuntos Sociales, a través del Instituto de Migraciones y Servicios Sociales, en colaboración con el Consejo Nacional de la Discapacidad y la

Conferencia Sectorial de Asuntos Sociales) y la Estrategia Global de Acción para el Empleo de Personas con Discapacidad 2008-2012, presentada en 2008 por el actual Ministerio de Trabajo e Inmigración.

No obstante, pese a que nos encontramos en una fase de desarrollo dinámico en materia de discapacidad, aún se necesita llevar a cabo más actuaciones al respecto, con el fin de conseguir la integración, la igualdad de oportunidades y la mejora en la calidad de vida de este colectivo.

2. LAS PERSONAS CON DISCAPACIDAD EN ESPAÑA

En nuestro país, la información más amplia que sobre discapacidad se dispone es la procedente del Instituto Nacional de Estadística, el cual ha prestado -desde hace varias décadas- una especial atención a las personas que cuentan con ella a través de tres importantes encuestas.

En 1986 realizó, en colaboración con el por entonces INSERSO, la Encuesta sobre Discapacidades, Deficiencias y Minusvalías (EDDM), donde por primera vez se estimaba el número de personas con discapacidad y los tipos de discapacidades.

En 1999, el INE, el Instituto de Migraciones y Servicios Sociales (IMSERSO) y la Fundación ONCE llevaron a cabo la Encuesta sobre Discapacidades, Deficiencias y Estado de Salud (EDDES), donde se introdujeron importantes novedades respecto a la anterior, como por ejemplo, la consideración de la severidad de la discapacidad y la necesidad de ayudas.

Y la tercera, publicada muy recientemente bajo el nombre de Encuesta sobre Discapacidad, Autonomía personal y situaciones de Dependencia (EDAD, 2008), incorpora como nota distintiva el dirigirse a personas residentes tanto en viviendas familiares (EDAD-hogares) como en determinados establecimientos colectivos (EDAD-centros).

De acuerdo con la EDAD-2008, el número de personas de 6 años y más con alguna discapacidad en España es el que aparece en la siguiente tabla:

Tabla 1: Personas de 6 y más años con alguna discapacidad por grupos de edad y sexo
(Unidades: miles de personas)

	Ambos sexos	Varones	Mujeres
De 6 a 44 años	**608,2**	**345,2**	**263,1**
De 6 a 16 años	85,5	55,6	30
De 17 a 24 años	67,8	41,6	26,2
De 25 a 34 años	168,7	98,8	69,8
De 35 a 44 años	286,1	149,1	137
De 45 a 64 años	**951,9**	**409**	**542,8**
De 45 a 54 años	406	181,9	224,1
De 55 a 64 años	545,8	227,1	318,7
De 65 a 79 años	**1201,7**	**454,8**	**746,8**
De 65 a 69 años	292,8	124,2	168,6
De 70 a 74 años	405,1	147,5	257,6
De 75 a 79 años	503,8	183,2	320,6
De 80 y más años	**1025,8**	**301,9**	**723,9**
De 80 a 84 años	482,6	148,6	333,9
De 85 a 89 años	339,8	103,2	236,6
De 90 y más años	203,4	50	153,4
Total	**3787,4**	**1510,9**	**2276,5**

Fuente: EDAD-2008

Analizando la tabla anterior se observa que el número de españoles mayores de 6 años con discapacidad asciende a más de 3 millones y medio, de los cuales prácticamente el 60% son mujeres. Las edades a las que hay más personas con discapacidad son a partir de los 65 años, aunque el dato entre los 45 y 64 años también es significativo.

3. EL MODELO DE ESTADOS MÚLTIPLES

Los modelos de estados múltiples se han convertido en una valiosa herramienta estadística para el análisis de datos longitudinales. De ahí que, dado los notables avances en estadística, se utilicen en diversos campos como la medicina, la biología, la física, las ciencias sociales, etc.

Este tipo de modelos se emplea en un proceso estocástico en el que, en cualquier momento, se ocupa uno de los estados discretos de un conjunto. En medicina, los estados pueden describir condiciones como la salud, la enfermedad o la muerte, denominándose "transición" al cambio entre estados. En cuanto a la estructura del modelo, éste debe especificar tanto los distintos estados como las transiciones que pueden darse entre los mismos. Así, el modelo estadístico completo ha de especificar la estructura entre estados, además de la forma de la función de riesgo (función de la intensidad) para cada posible transición (Hougaard, 1999).

Entre las distintas estructuras con múltiples estados que se pueden presentar está el modelo de la discapacidad, que cobra especial importancia ante enfermedades irreversibles, sobre todo, cuando éstas aumentan el riesgo de muerte. En concreto, la mayor utilidad de este modelo radica en la posibilidad de proyectar el número de personas que estarán discapacitadas, en base a las probabilidades o tasas de transición entre los estados, esto es, que una persona sin discapacidad pase a tenerla, que una persona con un grado de discapacidad pase a otro nivel de severidad o que fallezca.

Además de las tasas de transición, a la hora de estimar el número de personas discapacitadas que habrá en el futuro, se necesita conocer las tasas de prevalencia y determinar las tendencias que se puedan asumir. Las primeras posibilitarán el cálculo de las tasas de transición y las segundas permitirán establecer distintos escenarios para la proyección.

Así pues, el modelo de proyección a emplear requiere tres pasos fundamentales:

- Estimar el nivel de discapacidad de la población actual (apartado 3.1).
- Determinar las probabilidades de transición entre los distintos estados (apartado 3.2).
- Establecer diferentes escenarios sobre discapacidad y mortalidad bajo los cuales poder llevar a cabo la proyección (apartado 3.3).

3.1. Las tasas de prevalencia de la discapacidad

Las tasas de prevalencia, que constituyen el punto de partida del modelo, indican la proporción de personas de cada edad que tiene algún tipo de discapacidad en un momento determinado respecto al total de la población.

Teniendo en cuenta los datos recogidos en la Tabla 1 en cuanto al número de personas de 6 y más años con alguna discapacidad en España (EDAD-2008), así como las estimaciones de la población actual calculadas por el INE a partir del censo de 2001 y tomando como fecha de referencia el 1 de enero de 2008, se obtienen las siguientes tasas de prevalencia por rangos de edades:

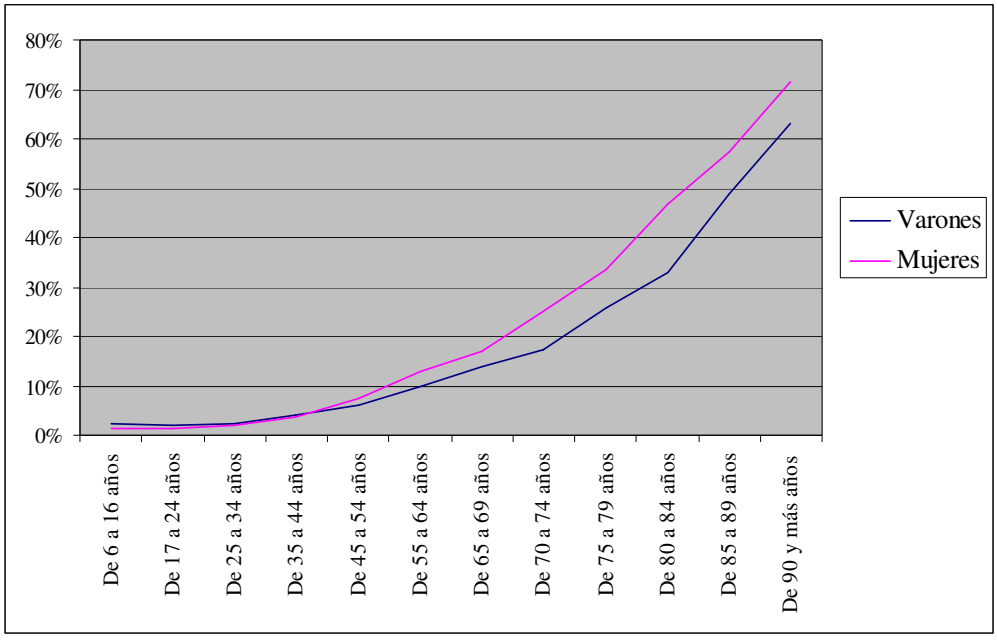
Tabla 2: Tasas de prevalencia de la discapacidad en España por grupos de edad y sexo

	Ambos sexos	Varones	Mujeres
De 6 a 44 años	**2,53%**	**2,80%**	**2,25%**
De 6 a 16 años	1,82%	2,31%	1,31%
De 17 a 24 años	1,62%	1,94%	1,28%
De 25 a 34 años	2,19%	2,48%	1,88%
De 35 a 44 años	3,85%	3,92%	3,77%
De 45 a 64 años	**8,69%**	**7,59%**	**9,76%**
De 45 a 54 años	6,65%	5,99%	7,31%
De 55 a 64 años	11,27%	9,67%	12,79%
De 65 a 79 años	**22,10%**	**18,57%**	**24,98%**
De 65 a 69 años	15,43%	13,89%	16,80%
De 70 a 74 años	21,57%	17,40%	24,99%
De 75 a 79 años	30,32%	25,92%	33,57%
De 80 y más años	**49,27%**	**40,73%**	**54,00%**
De 80 a 84 años	41,43%	32,97%	46,75%
De 85 a 89 años	54,50%	48,90%	57,36%
De 90 y más años	69,31%	62,89%	71,69%
Total	**8,92%**	**7,23%**	**10,56%**

Fuente: Elaboración propia a partir de la EDAD-2008 y de las estimaciones de la población actual de España calculadas por el INE a partir del censo de 2001

Al representar gráficamente los valores anteriores se pone de manifiesto, de forma más clara si cabe, cómo aumentan las tasas de prevalencia con la edad, siendo mayor la proporción de varones con discapacidad hasta los 45 años, edad a partir de la cual pasa a ser superior la proporción de mujeres.

Figura 1: Comportamiento de las tasas de prevalencia de la discapacidad en España por grupos de edad y sexo



Fuente: Elaboración propia

3.2. El modelo de transición

Rickayzen y Walsh (2002) proponen un modelo con el que proyectar el número de personas de Reino Unido que estarán discapacitadas hasta el año 2036, con el fin de determinar la asistencia a precisar.

Este modelo ha sido ampliamente utilizado en numerosos trabajos relativos a los múltiples estados de salud que puede presentar un individuo, como por ejemplo Leung (2004) para Australia o Del Giudice (2006) para Italia.

El modelo requiere conocer las tasas de prevalencia, las tasas de transición y las tendencias de estas últimas. Para ello, parten de los datos arrojados por la encuesta de discapacidad OPCS (Martin, Meltzer y Elliot, 1988) y establecen distintos supuestos acerca del comportamiento de las tasas de transición. En este sentido, aunque se siga el mencionado modelo, a la hora de estimar las probabilidades de transición se podrían descartar algunos de sus componentes y plantear modificaciones en varias expresiones y parámetros asociados, dadas las características de la población española.

3.2.1. La relación de dependencia entre la mortalidad y la discapacidad

No cabe duda de que la mortalidad depende, en parte, del grado de discapacidad que presente el individuo. Así, es de esperar que las tasas de mortalidad sean superiores para personas con grados de discapacidad severos.

Rickayzen y Walsh (2002) incorporan el término de *Mortalidad-Extra* para representar la dependencia entre la mortalidad y el grado de discapacidad en función de la edad. De este modo, la fórmula con la que expresar la *Mortalidad-Extra* para una persona de edad *x* con un grado de discapacidad *n* (correspondiendo *n*=1 a una persona con discapacidad moderada, *n*=2 a una con discapacidad severa y *n*=3 a una con discapacidad total) es:

$$Mortalidad\text{-}Extra\left(x,n\right)=\frac{0,2}{1+1,1^{50-x}}\cdot\frac{Max\left(n-1,0\right)}{5}$$

donde 50 es la edad pivote y el incluir $n-1$ conlleva a que sólo se considere una mortalidad adicional para los individuos con discapacidad severa o total. La Tabla 3 muestra los valores ilustrativos para distintas edades y para dichos grados de discapacidad:

Tabla 3: Mortalidad-Extra para discapacidad severa y total

Edad	Grado de la discapacidad	
	Discapacidad severa	Discapacidad total
10	0,0009	0,0017
20	0,0022	0,0043
30	0,0052	0,0104
40	0,0111	0,0223
50	0,0200	0,0400
60	0,0289	0,0577
70	0,0348	0,0696
80	0,0378	0,0757
90	0,0391	0,0783
100	0,0397	0,0793
110	0,0399	0,0797

Fuente: Rickayzen y Walsh (2002)

Una vez obtenidos estos datos, se puede determinar $Mortalidad\left(x,n\right)$, que representa la probabilidad de que una persona de edad *x*, que se encuentra en un estado de discapacidad *n*, muera a lo largo del próximo año. La expresión empleada es la siguiente:

$$Mortalidad\left(x,n\right)=Mortalidad\left(x,0\right)+Mortalidad\text{-}Extra\left(x,n\right)$$

3.2.2. La probabilidad de pasar a estar discapacitado

La expresión utilizada por Rickayzen y Walsh (2002) tiene en cuenta cuatro parámetros para las mujeres y uno adicional para los hombres.

Para las mujeres, la fórmula es:

$$Nueva_Discapacitada\left(x\right)=A+\frac{D-A}{1+B^{C-x}}$$

en la que los cuatro parámetros son *A*, *B*, *C* y *D* y $Nueva_Discapacitada\left(x\right)$ es la probabilidad de que una mujer de edad *x* pase a estar discapacitada en un año.

Para los hombres, se emplea la siguiente expresión:

$$Nuevo_Discapacitado\left(x\right)=\left(A+\frac{D-A}{1+B^{C-x}}\right)\times\left(1-\frac{1}{3}\cdot\exp\left[-\left(\frac{x-E}{4}\right)^{2}\right]\right)$$

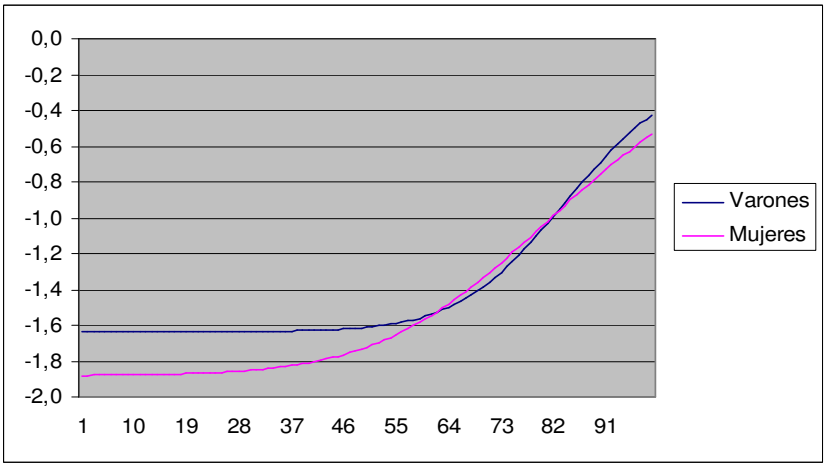
donde:

- El parámetro *A* es el límite de la probabilidad de pasar a estar discapacitado a edades muy jóvenes.
- El parámetro *D* es el límite de la probabilidad de pasar a estar discapacitado a edades muy avanzadas.
- El parámetro *B* indica la velocidad en la que cambian las probabilidades entre los dos valores extremos.
- El parámetro *C* es la edad máxima.
- El parámetro adicional *E* representa la edad a la que hay un cambio de pendiente en la función de $Nuevo_Discapacitado\left(x\right)$.

Dado que aún no se dispone de los datos pormenorizados de la EDAD-2008, las tasas de prevalencia para cada edad no se conocen, por lo que no es posible determinar la edad exacta en la que hay un cambio de pendiente. De ahí que, en un principio, no se tenga en cuenta el parámetro *E*, utilizando la misma expresión para determinar la probabilidad de que ambos sexos pasen a estar discapacitados.

La siguiente figura representa el logaritmo (base 10) de la probabilidad anual de pasar a tener una discapacidad para ambos sexos.

Figura 2: Log probabilidad anual de pasar al estado de discapacidad para varones y mujeres



Fuente: Elaboración propia a partir de la metodología propuesta por Rickayzen y Walsh (2002)

Se observa que la probabilidad crece con la edad y es superior para el caso de los hombres hasta los 62 años, a partir de los cuales la probabilidad que tienen las mujeres de pasar a discapacitadas es mayor. Por último, desde los 83 años en adelante la probabilidad de discapacitarse vuelve a ser superior para los hombres, aunque la diferencia no es tan acusada como en el primer tramo.

3.3. Tendencias de la mortalidad y la discapacidad

Las tendencias de la mortalidad y la discapacidad se determinan a través de la mortalidad extra y de la probabilidad de pasar a un estado de discapacidad.

Una disminución en la mortalidad extra, quizás como consecuencia de los avances médicos o de la mejora de los cuidados prestados, conlleva a que las personas vivan más tiempo una vez que pasan a estar discapacitadas. Por el contrario, un aumento en la misma indica que la esperanza de vida de las personas con discapacidad se acorta.

Siguiendo a Rickayzen y Walsh (2002), la manera en que se establecen distintos escenarios en base a las tendencias de la mortalidad es sustituyendo en la expresión de la *Mortalidad-Extra* la cifra 0,2 por:

$\left(0,2+\Delta\cdot\frac{t-2008}{10}\right)$ donde $\Delta=\pm 0,02$ y t representa un año entre el 2009 y el último año para el que se quiera realizar la proyección.

Por lo que respecta a las tendencias de la discapacidad, el aumento o la disminución interanual a considerar se incorporarían en la función de $Nueva_Discapacidad$.

En la tabla siguiente se presentan algunos de los escenarios de proyección que podrían plantearse teniendo en cuenta las hipótesis anteriores.

Tabla 4: Análisis de la sensibilidad

Escenario	Mortalidad-Extra	Nueva_Discapacidad
A	↓	↓
B	↓	↑
C	↑	↑
D	↑	↓

Fuente: Elaboración propia

4. CONCLUSIONES

El método de proyección propuesto en este trabajo parte de la estimación del nivel de discapacidad en España. Aunque ya ha sido publicado un avance de resultados de la *Encuesta de Discapacidad, Autonomía Personal y Situaciones de Dependencia* (2008), se precisa de los datos detallados de la misma para poder obtener las tasas de prevalencia para cada una de las edades y poder desarrollar el modelo de transición entre estados de salud de manera exhaustiva.

Así pues, en base a la prevalencia de la discapacidad se podría determinar una de las dos componentes principales del modelo: la probabilidad de que una persona no discapacitada pasase a estarlo. La otra componente, la mortalidad-extra de aquellas personas que presentasen alguna discapacidad, combinada con la mortalidad global de la población española, permitiría conocer las probabilidades de fallecimiento/supervivencia, a partir de las cuales se realizaría la proyección.

Por último, sería en torno a ambas componentes como habría que llevar a cabo el análisis de la sensibilidad, estableciendo distintos escenarios en función de que experimentasen un incremento, una disminución o se mantuviesen constantes.

6. BIBLIOGRAFÍA

DEL GIUDICE, A. (2006). Una proiezione dinamica del numero delle persone confinate e implicazioni per i futuri bisogni assistenziali di lungo periodo. *Politiche Sanitarie*, *7* (1), 15-29.

HOUGAARD, P. (1999). Multi-state Models: A Review. *Lifetime Data Analysis*, *5*, 239-264.

INE (1987). *Encuesta sobre Discapacidades, Deficiencias y Minusvalías 1986*. Madrid: INE e INSERSO.

INE (1999). *Encuesta sobre Discapacidades, Deficiencias y Estado de Salud*. Madrid: INE, IMSERSO y Fundación ONCE.

INE (2008). *Encuesta sobre Discapacidad, Autonomía personal y situaciones de Dependencia*. Madrid: INE, Ministerio de Educación, Política Social y Deporte, Fundación ONCE, CERMI y FEAPS.

LEUNG, E. (2004). Projecting the Needs and Costs of Long Term Care in Australia. *Australian Actuarial Journal*, *10* (2), 343-385.

MARTIN, J., MELTZER, H. & ELLIOT, D. (1988). *OPCS surveys of disability in Great Britain, Report 1, The prevalence of disability among adults*. London: HMSO.

RICKAYZEN, B. D. & WALSH, D. E. P. (2002). A multi-state model of disability for the United Kingdom: Implications for future need for Long-Term Care for the elderly. *British Actuarial Journal*, *8* (2), 341-393.

EL DESARROLLO LOCAL Y LAS EMPRESAS DE ECONOMÍA SOCIAL. UN ANÁLISIS PARA LA C. A. DE CASTILLA Y LEÓN

JUAN JOSÉ JUSTE CARRIÓN
E-mail: juste@eco.uva.es
JESÚS MARÍA GÓMEZ GARCÍA
E-mail: jmgomez@eco.uva.es
JOSEFA EUGENIA FERNÁNDEZ ARUFE
E-mail: jfarufe@eco.uva.es
Departamento de Economía Aplicada
Facultad de Ciencias Económicas y Empresariales
Universidad de Valladolid

Resumen

Los procesos de desarrollo local requieren la concurrencia en el territorio de factores muy diversos (económicos, sociales, culturales, etc.), así como de un considerable esfuerzo de cohesión económica y social en el que deben implicarse las fuerzas vivas del territorio, entre las cuales se encuentran las empresas de la Economía Social. El objetivo de este trabajo es analizar la relevancia del subsector empresarial de la Economía Social en el proceso de desarrollo local en Castilla y León, con especial referencia al mundo rural. De acuerdo con ello, partiendo de la identificación de los aspectos fundamentales que caracterizan el desarrollo local como elemento de reestructuración productiva (origen, factores condicionantes y espacios productivos típicos), en el trabajo se describen sus vínculos con el sector de la Economía Social. A continuación, se analiza, desde una perspectiva aplicada, la contribución de las empresas de la Economía Social al desarrollo endógeno local/rural de Castilla y León en términos de empleo y producción. Los datos requeridos para llevar a cabo este análisis proceden, además de las fuentes estadísticas habituales, de los resultados de una encuesta propia realizada a las empresas de la Economía Social en Castilla y León.

Palabras clave: Economía social, Desarrollo local/rural, Empresas de economía social.
Área Temática: Economía Social y Discapacidad.

Abstract

Local development processes require the presence in the territory of many different factors (economic, social, cultural, etc.) as well as a considerable economic and social cohesion in which should be engaged the forces of the territory, among which are the enterprises of the Social Economy. The aim of this paper is to analyze the relevance of the business sector of the Social Economy in the process of local development in Castilla y León, with special reference to rural areas. Accordingly, based on the identification of the fundamental aspects that characterize the development as part of economic restructuring (origin, production factors and spaces typical), the paper describes its links with the Social Economy sector. Later, it analyzes, from an applied perspective, the contribution of the Social Economy enterprises to local/rural endogenous development of Castilla y León in terms of employment and production. The data to conduct this analysis is based, in addition to those obtained from the usual statistical sources, on the results of a own survey conducted to the Social Economy enterprises in Castilla y León.

Key words: Social economy, Local/rural development, Social economy enterprises.
Thematic Area: Social Economics.

1. DESARROLLO LOCAL Y REESTRUCTURACIÓN PRODUCTIVA.

La profunda reestructuración productiva subsiguiente a la crisis económica de los años 70 se ha traducido, a lo largo de los últimos treinta años, en una compleja reorganización de los espacios productivos en las economías occidentales, de la que ha resultado un variopinto puzzle de regiones y localidades, interpretable en términos de ganadores (+) y perdedores (-) y cuyas piezas fundamentales serían (Benko y Lipietz, 1992; Vázquez Barquero, 2005),:

1.*Zonas de antigua industrialización* (-). Densamente pobladas y con predominio de sectores maduros sensibles a los costes laborales y energéticos (textil, siderurgia, minería…). Duramente castigadas por la crisis y pasto de un acelerado proceso de ajuste empresarial, dominado por una lógica de desindustrialización y reconversión.

2.*Áreas rurales deprimidas* (-). Eminentemente agrarias, en riesgo de descapitalización, y poco atractivas para la actividad industrial y terciaria. Al paro y al éxodo y envejecimiento demográficos, se añaden la concentración de las explotaciones; la creciente dependencia de las grandes firmas industriales y de distribución; el bajo nivel de vida; la escasez de infraestructuras y servicios; el abandono y degradación del territorio; o la falta de oportunidades.

3.*Áreas urbanas de tejido económico reconvertido* (+). Protagonistas de una regeneración urbana basada en la revitalización económica y en la reutilización de terrenos degradados, resultantes del cierre de empresas y de la congestión y deterioro del entorno físico y del tejido económico. Son ciudades que han sabido reaccionar frente a la desconcentración industrial hacia la periferia, impulsando un nuevo papel metropolitano en los negocios, con los servicios como punta de lanza.

4.*Espacios ligados a actividades estratégicas y a la innovación* (+). Resultado de la aparición, en áreas específicas, de concentraciones tecnológicas planificadas, de la mano de la Administración Pública, y/o de otras instituciones (universidades, asociaciones empresariales…), con objeto de promover el desarrollo regional, a través de la I+D+i, la reindustrialización y la instalación de nuevas industrias en su seno. Figuran aquí las llamadas tecnópolis y los parques científicos y tecnológicos.

5.*Espacios productivos emergentes* (+). Territorios con una trayectoria tecnológica de desarrollo difuso, consistente con la lógica de la acumulación flexible donde niveles retributivos más bajos, escasa conflictividad laboral, accesibilidad a los mercados, entorno agradable, receptividad social, etc., constituyen fuentes de ventaja competitiva. Es el caso de numerosas ciudades modestas favorecidas por la descentralización productiva y la menor fricción de la distancia, y de otros espacios periféricos (ciudades medias y cabeceras de comarca rurales) sustentados autónomamente gracias a una amplia dotación de recursos locales.

Esta heterogénea geografía productiva parece reflejar, en los inicios del siglo XXI, más que la implantación de un nuevo y único patrón de desarrollo territorial postfordista, la coexistencia de una pluralidad limitada de modelos para la resolución de los problemas económicos. Una de sus posibles sendas, típica de los *espacios emergentes*, es la del *desarrollo local endógeno*[1], patrón organizativo flexible y descentralizado basado en la valorización y aprovechamiento de los recursos propios y en la versatilidad de las PYMES, y donde el territorio juega un papel activo como agente de desarrollo.

Las figuras en que se materializa dicho modelo de *desarrollo desde abajo - distritos industriales* y, más en general*, sistemas productivos locales* (SPL)[2] – están detrás del éxito de

[1] Vázquez Barquero (1988) define el desarrollo local endógeno en los siguientes términos: "un proceso de crecimiento económico y cambio estructural que conduce a una mejora del nivel de vida de la población local, en el que se pueden identificar, al menos, dos dimensiones: Una económica, en la que los empresarios locales usan su capacidad para organizar los factores productivos locales con niveles de productividad suficiente para ser competitivos en los mercados; otra socio-cultural, en la que los valores y las instituciones locales sirven de base al proceso de desarrollo. La estrategia de desarrollo local, no obstante, debe considerar una dimensión adicional, la político-administrativa, en la que las políticas territoriales permiten crear un entorno económico local favorable, protegerlo de las interferencias externas y favorecer e impulsar el desarrollo del potencial local".

[2] El distrito industrial constituye una categoría sui generis de SPL, basado en la presencia de un amplio contingente de PYMES locales especializadas. En general, entre los elementos económicos que suelen concurrir en todo SPL, como patrón de organización industrial, figuran: 1) Una elevada división del trabajo entre empresas (PYMES y/o grandes) radicadas en un mismo espacio geográfico, a escala intra e intersectorial, cuyo último eslabón sería el trabajo a domicilio y, en ocasiones, la economía sumergida, como exponente de flexibilidad laboral. 2) Una especialización productiva en empresas e instalaciones que facilita su modernización tecnológica y su relativa autonomía económica. 3) El juego combinado de mercado y reciprocidad (entendida ésta como intercambio

numerosos territorios por la vía de la industrialización, apareciendo como una válvula de escape a la marginación para muchas áreas rurales y periféricas.

Los elementos que concurren en dichos espacios productivos – explicativos de su singular dinamismo e integrantes de su *potencial endógeno* - son de muy diversa índole, destacando en particular los siguientes (Castillo et al. 1994):

a)Recursos humanos: un nivel de población suficiente, una amplia oferta de trabajo, flexible y con un notable grado de cualificación, y la presencia de emprendedores capaces de generar empleo y diversificar el tejido económico.

b)Recursos materiales: una amplia dotación de recursos naturales, una suficiente red de equipamientos e infraestructuras técnicas y sociales y un importante caudal de recursos económicos financieros: mecanismos de ahorro local, servicios financieros y un importante número de establecimientos productivos (especialmente PYMES).

c)Recursos técnicos: el conocimiento contextual y la receptividad al cambio tecnológico como fuente de creatividad local.

d)Recursos culturales: una escala de valores compartidos y una fuerte identidad local.

e)El poder y capacidad de organización a escala local, merced a la implicación de las fuerzas vivas del área: empresas y otros agentes privados (sindicatos, patronal, cámaras de comercio, bancos…), poderes públicos, formas sociales y económicas de solidaridad e integración (familias, cooperativismo, redes de ayuda, asociaciones ciudadanas…), e incluso estructuras de animación e información (partidos políticos, grupos de acción local, agentes de desarrollo local,…). Entre esa amplia diversidad de figuras se inscriben las que configuran la Economía Social.

2. LA ECONOMÍA SOCIAL Y SUS VÍNCULOS CON EL DESARROLLO LOCAL.

La noción de Economía Social presenta acepciones diversas que le confieren un carácter esencialmente abierto, a la vez que una cierta dosis de ambigüedad. En el ámbito científico, y en aras de reflejar su comportamiento y creciente importancia económica, la definición más operativa es la propuesta por la CIRIEC-España (Chaves, 2006; Chaves y Monzón, 2001, 2003; Barea y Monzón, 2002)[3]:

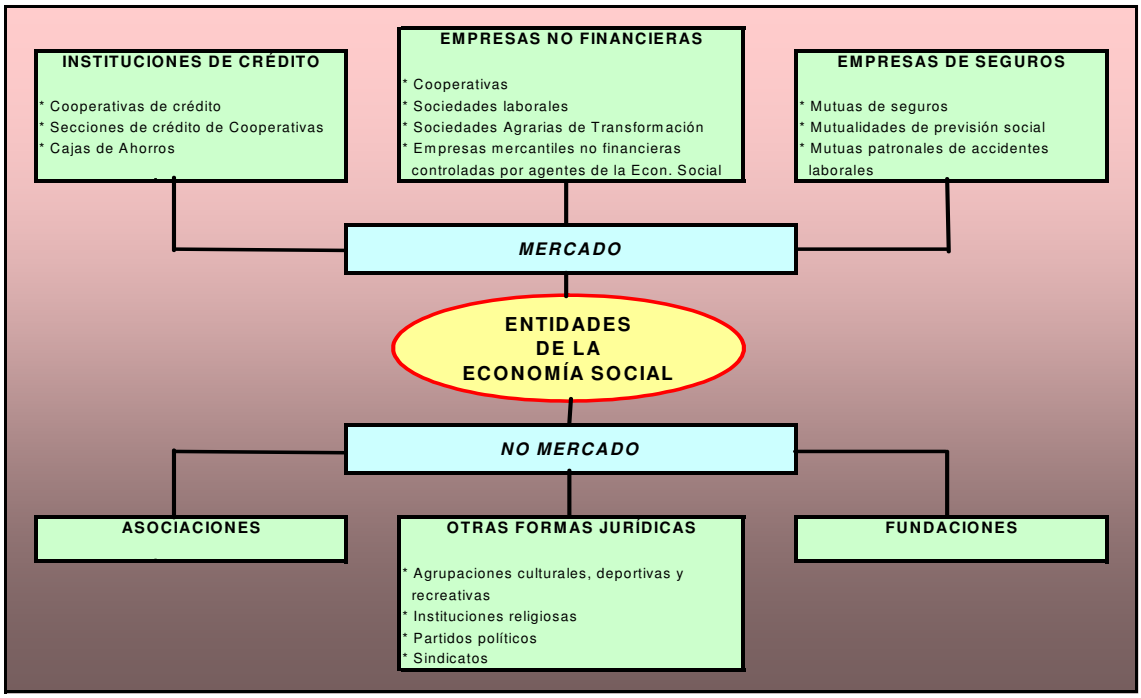
"aquel conjunto de empresas privadas creadas para satisfacer las necesidades de sus socios a través del mercado, produciendo bienes y servicios, asegurando o financiando y en las que la distribución del beneficio y la toma de decisiones no están ligadas directamente con el capital aportado por cada socio, correspondiendo un voto a cada uno de ellos. La Economía Social también incluye a las instituciones sin fines de lucro que son productores no de mercado privados, no controlados por las Administraciones Públicas y que producen servicios no destinados a la venta para determinados grupos de hogares, procediendo sus recursos principales de contribuciones voluntarias efectuadas por los hogares en su calidad de consumidores, de pagos de las Administraciones Públicas y de rentas de la propiedad".

A la luz de dicha definición, cabe distinguir dos subsectores dentro de la Economía Social: el de *mercado* o empresarial y el de *no mercado* (sin ánimo de lucro) cuyas diferentes formas jurídicas se recogen en la Figura 1:

gratuito y directo de información y servicios entre empresarios y proveedores, y vinculada a la fidelidad e identidad locales), que, por un lado, favorece la adquisición de conocimientos profesionales y experiencia y el acceso a mejoras tecnológicas y organizativas, y, por otro, genera el ambiente propicio para la asunción local de riesgo en los negocios, reforzando el potencial sinérgico de la comunidad. 4) La atmósfera de la industrialización difusa como trasfondo, ya que la fuerte interacción entre economía y sociedad y las iniciativas locales de desarrollo constituyen la base de su nacimiento y consolidación. Para mayor información sobre el concepto, la tipología y los rasgos de los SPL, véase Garofoli y Mazzoni (1994).

[3] Esta definición de la asociación científica CIRIEC-España (Centro Internacional de Investigación e Información sobre la Economía Pública Social y Cooperativa), se aproxima a las vigentes en Francia o Bélgica, países también pioneros - junto con España y Portugal – en el reconocimiento, a finales de los 70, de la creciente importancia de las formas sociales privadas cuyos parámetros no encajan con la lógica de acumulación capitalista. Chaves (2006) subraya la existencia de otras visiones del concepto de Economía Social colaterales. Así, en el terreno político y legislativo, los poderes públicos suelen identificar el concepto de Economía Social con una parte de su Subsector de mercado: las cooperativas y las sociedades laborales. En el ámbito profesional-empresarial, la Economía Social se sitúa a caballo entre las visiones académica y político-institucional, incorporando otras formas empresariales orientadas a la cohesión social: mutualidades, fundaciones, centros especiales de empleo y empresas de inserción.

Figura 1: Entidades integrantes de la economía social por sectores institucionales



FUENTE: Adaptado de Barea y Monzón (2002).

En cualquier caso, de acuerdo con Toscano (2000), se trata de iniciativas muy diversas que, en esencia, pretenden cubrir aquellas lagunas que ni el mercado en su función asignativa, ni la intervención pública correctora de las desigualdades pueden colmar. Ya que tal pretensión se identifica, en primera instancia, con una finalidad de servicio a los miembros componentes o a una colectividad dada – con el consiguiente fomento del bienestar social -, el verdadero germen de la Economía Social radica en la solidaridad.

Con todo, el papel de la Economía Social no se limita a su contribución genérica a una sociedad más plural, participativa, democrática, solidaria y justa, en definitiva; de hecho, el entramado institucional que la integra cuenta con una reconocida capacidad para desempeñar eficazmente múltiples funciones de índole económica, entre las que destacan las siguientes (Chaves y Monzón, 2003; Gómez y Román, 2004):

1.La distribución más igualitaria de la renta y la riqueza, a través de mecanismos como la preferencia otorgada a personas, servicios y factor trabajo respecto al factor capital en el reparto de los beneficios de las cooperativas, o el voluntariado y las donaciones privadas en el ámbito de las organizaciones sin ánimo de lucro.

2.Corrección de desequilibrios en el mercado laboral, tanto en épocas de recesión como de auge económico. La lucha contra la inestabilidad en el empleo, el paro y la exclusión laboral, o la generación de nuevos puestos de trabajo a menudo asociados a los llamados *nuevos yacimientos de empleo*, así lo confirman.

3.Oferta de servicios de bienestar social en campos en continua expansión: servicios a la tercera edad, infancia y minusválidos, servicios educativos, sanitarios y culturales, ayuda a refugiados y a marginados.

4.Estabilidad económica. El carácter participativo de la Economía Social permite afrontar las crisis con fórmulas flexibles distintas de las tradicionales (ajustes de plantilla, expedientes de regulación, cierres de empresas,…), como son la reestructuración de las remuneraciones de los socios o la reactivación de empresas en dificultades a través de su adquisición en propiedad por los propios trabajadores.

5.Desarrollo sostenible. Los valores inherentes a la Economía Social (prácticas democráticas y solidarias, tejido asociativo, formación, voluntariado, cooperación al desarrollo,…) configuran un sustrato favorable al cambio cultural requerido por un estilo de vida más respetuoso con el medio ambiente y la biodiversidad.

6.Autonomía territorial. Al promover la democracia y la participación ciudadana, las entidades de la Economía Social confieren un gran protagonismo a la sociedad civil del territorio en que asientan a la hora de determinar el modelo de desarrollo de la zona y de controlar la evolución del proceso de crecimiento y cambio estructural.

7.Desarrollo local. La capacidad de la Economía Social para estimular iniciativas de desarrollo endógeno donde el sector privado lucrativo y las autoridades públicas no han logrado hacerlo satisfactoriamente tiene su raíz en fuertes elementos de convergencia entre las funciones anteriores y los rasgos del modelo de *desarrollo desde abajo* (Montero, 1997; Toscano, 2000; Coque, 2003; Mozas y Bernal, 2006):

- El carácter del **desarrollo local como estrategia integral**, supone que en su diseño concurren los diferentes ámbitos que determinan el pulso vital de todo territorio: económico, social, institucional, cultural, histórico, medioambiental, etc., a la vez que el esfuerzo permanente de todos los agentes implicados, incluida la **Economía Social, como parte activa del sector empresarial y de las fuerzas vivas locales.**

- El **desarrollo local es un modelo descentralizado y participativo**. Este espíritu cooperativo encuentra un caldo de cultivo propicio en las entidades de **la Economía Social**, **caracterizadas por la democratización en la toma de decisiones**.

- **El desarrollo local es un proceso reactivador de la economía y dinamizador de la sociedad local**, siendo esenciales en este contexto la flexibilidad, el dinamismo y la cultura de la pequeña y mediana empresa, como principal ingrediente del tejido productivo de mayoría de las iniciativas de desarrollo local. **La Economía Social ha venido jugando un importante papel en áreas afectadas por** la lógica de la desinversión y **la reconversión industrial**, a menudo mediante la creación de PYMES (cooperativas, sociedades laborales) a instancia de los propios empleados.

- **El desarrollo local se fundamenta en la movilización de los recursos de la zona**, hecho preceptivo en las experiencias de desarrollo endógeno. **Las entidades de la Economía Social no son intensivas en capital**, **debiendo compensar la escasez de medios mediante el empleo adecuado de los recursos del entorno** (naturales, humanos, técnicos,…) que configuran el potencial endógeno local.

- **El desarrollo local se halla íntimamente ligado a la vocación emprendedora**. Si algo caracteriza el surgimiento de iniciativas locales es la existencia de empresarios y/o el empeño de los trabajadores de llegar a serlo. **Ese espíritu es consustancial a la Economía Social, dado el recurso al autoempleo** y el riesgo asumido por muchas personas al vincular su futuro y recursos a actividades por cuenta propia.

- **El desarrollo local se fundamenta en el capital humano del área:** en los trabajadores y su cualificación, crucial para adecuarse a los requerimientos de las empresas locales. El **principio de educación, formación e información es esencial para la competitividad** de las empresas **de Economía Social,** que suelen destinar parte de sus propios fondos a la formación continua de sus socios y trabajadores.

- **El desarrollo local persigue la fijación de la población en el territorio**, **mediante** la oferta de **puestos de trabajo** y calidad de vida en la zona. Aunque las entidades de **la Economía Social** no se guían por el ánimo de lucro, su presencia **implica la creación de riqueza –** a menudo orientada a suplir carencias sociales **- y de empleo estable**, anteponiéndose frecuentemente éste a la propia actividad empresarial.

- **Desarrollo local y Economía Social presentan un enfoque de largo plazo**. El desarrollo local pretende crear las condiciones para un desarrollo sostenido del territorio, mediante una estrategia de largo alcance. La Economía Social propende hacia la obtención de resultados duraderos, en forma de beneficios sociales, empleo y bienestar social, y no tanto hacia el logro de efímeros beneficios a corto plazo.

- **El desarrollo local** promueve la especialización productiva en actividades en las que el área goza de ventaja competitiva, pero a la vez **fomenta sectores emergentes** como factor de diversificación. **Muchos de estos nuevos yacimientos de empleo** tienen marcado carácter

local, al estar relacionados con los servicios a la comunidad, la provisión de bienes preferentes o el cuidado del medio ambiente. Tales rasgos **entroncan con valores como la ausencia de ánimo de lucro, la participación democrática o el objetivo de bienestar social propios de la Economía Social**, lo que la sitúa en una posición favorable al desarrollo de nuevos nichos de mercado.

➢ **El desarrollo local está abierto a la iniciativa pública**. Aunque en el pasado las autoridades no han sido cruciales en el nacimiento de experiencias de desarrollo endógeno, hoy pueden fomentar la consolidación de las existentes, así como la aparición de otras nuevas, mediante una planificación estratégica atenta a las especificidades de cada zona. **Dicho complemento no es ajeno a la Economía Social, mediante normas, estructuras de acompañamiento y políticas –** como el impulso del *partenariado* entre autoridades locales y organizaciones de la Economía Social **- orientadas a la creación de entidades y mejora de su financiación**.

➢ **Desarrollo local y Economía Social se identifican con el desarrollo sostenible**. La supervivencia de una economía local depende de la permanencia de la actividad productiva. Ello implica la necesidad de considerar su impacto ambiental, así como la integración de las estrategias de desarrollo y de conservación. El principio de interés por la comunidad inherente a la Economía Social o su inclinación al sostenimiento de actividades en trance de desaparición (artesanía, industria tradicional) la hacen particularmente compatible con la idea de sostenibilidad.

➢ **Las experiencias de desarrollo local presentan una marcada identidad local**. **Las entidades de Economía social**, con o sin ánimo de lucro, **se identifican estrechamente con el territorio en que operan**. Se suelen constituir allí donde surge una iniciativa o la fuerte convicción personal de poder salir adelante de un conjunto de emprendedores locales. Buscando mejorar su calidad de vida, tratarán de aprovechar los recursos a su alcance, vinculando a menudo su actividad a las necesidades locales y tendiendo a reinvertir beneficios en la zona, lo que conjura el riesgo de traslado o cierre selectivo, típico de los procesos de deslocalización.

➢ El **desarrollo local y** la **Economía Social son trascendentales para el entorno rural**. El afán de reactivación económica, el aprovechamiento de los recursos del área, la vinculación de los emprendedores al territorio, la fijación de la población, son elementos de confluencia entre lo "local" y lo "social" con efectos "balsámicos" sobre el mundo rural, considerando sus problemas y las oportunidades de inversión inherentes al apoyo público en los últimos años, con vistas a su diversificación productiva. En este sentido, resulta crucial el papel de las empresas y el de figuras como las Asociaciones de agentes de empleo y desarrollo local y los Grupos de Acción Local (GAL), forma representativa de partenariado a escala microterritorial.

3. ECONOMÍA SOCIAL Y DESARROLLO TERRITORIAL EN CASTILLA Y LEÓN

Admitidas las sinergias entre desarrollo local y Economía Social, cabe preguntarse cómo se plasman en la práctica en un ámbito territorial concreto como es la Comunidad Autónoma de Castilla y León. Teniendo presente que los procesos de desarrollo local/regional se enmarcan en contextos más amplios, antes de abordar dicha cuestión, resulta conveniente realizar algunos comentarios previos a nivel nacional.

3.1 La Economía Social y el desarrollo local en España: aspectos generales.

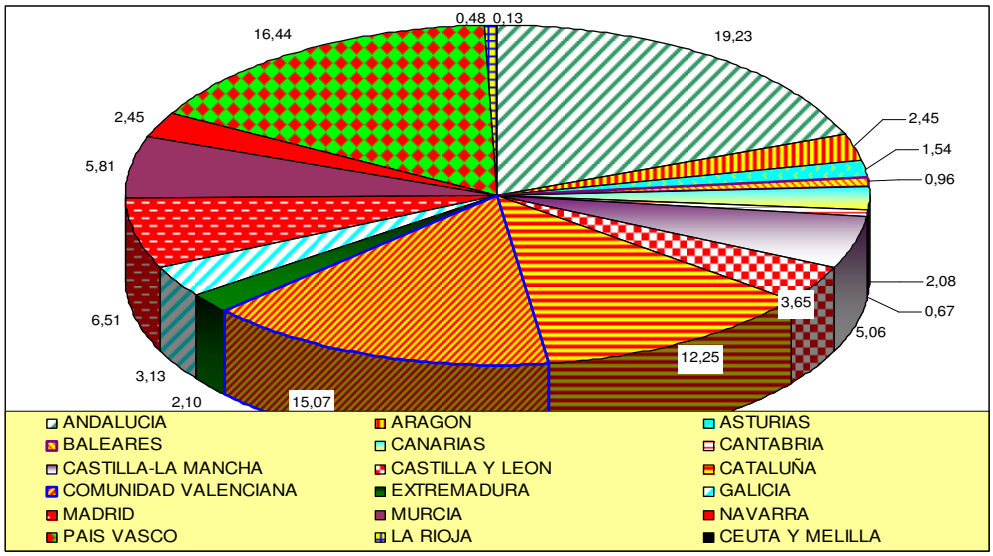
El análisis de la literatura sobre Economía Social y desarrollo local en España permite realizar las siguientes consideraciones de orden general:

- Los procesos de desarrollo local cobran especial importancia como respuesta, a menudo autónoma, a la crisis industrial de los 70. Las experiencias más representativas, muchas con una dilatada tradición, giran en torno al dinamismo industrial, cristalizando en la configuración de Sistemas Productivos Locales.
- Los SPL españoles, por lo común con base en la PYME, suelen localizarse en ciudades medias o cabeceras de comarca, a escala municipal, y, preferentemente, en las CCAA del Arco

Mediterráneo y el Valle del Ebro, a escala regional[4], en claro contraste con la realidad de regiones del interior como Castilla y León.

- España cuenta en la actualidad con uno de los más altos porcentajes de empleo a escala europea en el sector de la Economía Social, constituyendo las cooperativas y las sociedades laborales las entidades más representativas[5]. Dichas entidades han venido mostrando una elevada capacidad para afrontar las crisis, merced a la adopción de mecanismos flexibles de difícil aceptación en otro tipo de firmas, tales como el ajuste de la jornada laboral o de los salarios.
- Tradicionalmente se ha tratado de entidades localizadas en zonas en declive o deprimidas, impulsando sus inversiones procesos de revitalización económica local. En los últimos años, no obstante, se viene asistiendo al surgimiento de entidades sociales – PYMES en su inmensa mayoría - en múltiples actividades intensivas en factor trabajo, como es el caso de los servicios de ayuda a personas dependientes, enseñanza, servicios de limpieza, agricultura ecológica, turismo rural o los ámbitos de inserción laboral y de lucha contra la exclusión.
- Esa expansión no ha sido homogénea desde la óptica territorial, como demuestran las cifras de trabajadores y empresas en las diferentes CCAA. La Figura 1 refleja el mayor peso de Andalucía, País Vasco, Comunidad Valenciana y Cataluña, con porcentajes de participación en la Economía Social por encima del 12%, en contraste con Castilla y León, cuyo menor dinamismo en cuanto a la creación de empleo se traduce en una contribución muy inferior (del 3,65%). Hecho que llama la atención a tenor de su tamaño geográfico y demográfico, especialmente si se compara con regiones menos extensas y pobladas como Murcia, Navarra o Canarias. Este comportamiento es resultado de una elevada y heterogénea natalidad empresarial, que conduce a una desigual distribución territorial.

Figura 1: Distribución regional del número de trabajadores en cooperativas y sociedades laborales a 31-12-2008 (%)



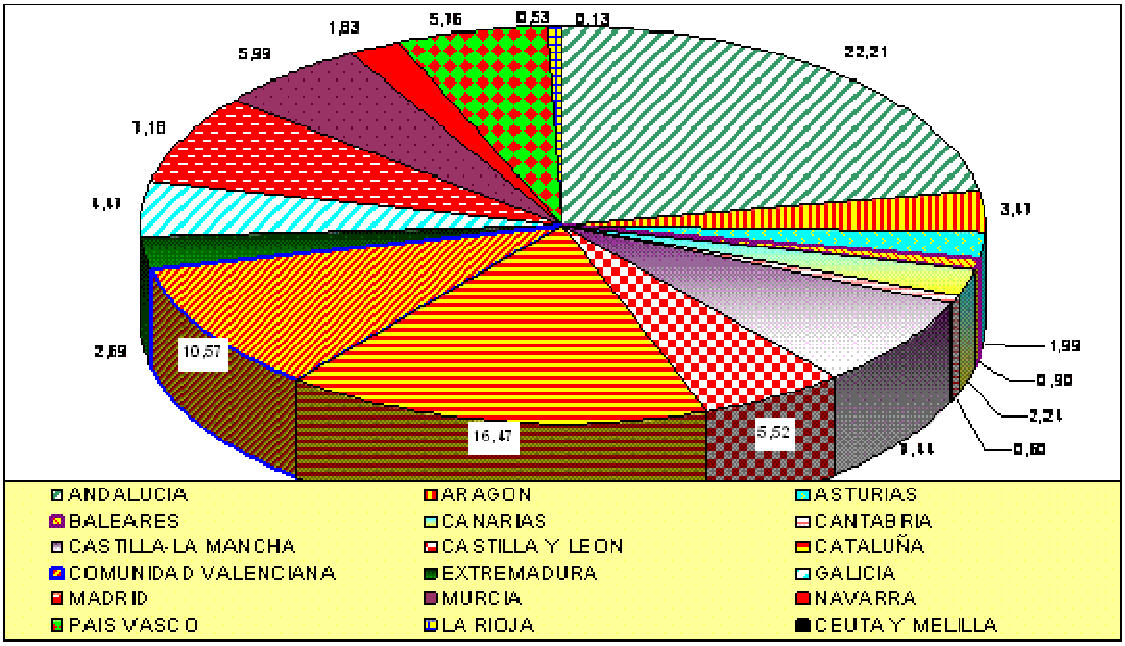
FUENTE: Elaboración propia a partir de los Datos de la Dirección General de la Economía Social, del trabajo autónomo y del Fondo Social Europeo. Ministerio de Trabajo y Asuntos Sociales.

[4] Muchas de esas experiencias presentan una marcada base cooperativa, merced a la configuración de grandes grupos de Economía Social, como es el caso de Mondragón en Guipúzcoa o de Guissona en Lérida. En Castilla y León la industrialización de raíz local se limita a un número reducido de experiencias- no más de una veintena -, entre las que destacan Aranda de Duero, Tierra de Pinares, Ponferrada, Guijuelo, Cantimpalos, Íscar, la Ribera del Duero y Medina del Campo (Juste, 2001; Juste et al. 2007).

[5] De acuerdo con Chaves (2006), el impulso acelerado de la Economía Social en España se debe, en esencia, a la conjunción de una serie de factores: 1) El advenimiento del régimen democrático que cristaliza en el reconocimiento y consolidación de las libertades de asociación y expresión, en las que se enmarca la filosofía del Estado benefactor. 2) Los dolorosos ajustes frente a las crisis económicas, y la necesaria recuperación industrial y tecnológica. 3) El agravamiento del desempleo de larga duración. 4) La aparición de nuevas demandas sociales, en particular aquellas asociadas al proceso de envejecimiento demográfico. 5) La progresiva desertización social y económica de numerosas zonas rurales. 6) La persistencia de problemas como la pobreza y la exclusión. Tomando como referencia las citadas entidades, la expansión de la Economía Social a lo largo de los últimos años resulta ciertamente evidente, hasta el punto de duplicar en 2007 (con una cifra de 442.300 trabajadores), el volumen de empleo existente en 1990 (224.000) y el número de entidades (unas 20000 en 1995 y 45000 en 2007).

• La Figura 2 refleja el fuerte peso en el panorama nacional de la Economía Social de regiones como Andalucía, Cataluña y la Comunidad Valenciana, que aglutinan el 50,4% del número total de entidades. El porcentaje correspondiente a Castilla y León resulta mucho más modesto (5,52%), registrando la región la menor dimensión media por entidad (6,4 trabajadores, frente a los 9,7 de España).

Figura 2: Distribución regional del número de cooperativas y sociedades laborales a 31-12-2008 (%)



FUENTE: Elaboración propia a partir de los Datos de la Dirección General de la Economía Social, del trabajo autónomo y del Fondo Social Europeo. Ministerio de Trabajo y Asuntos Sociales.

• Este menor peso específico de Castilla y León se halla en sintonía con el menor protagonismo de la Economía Social en su tejido productivo, como lo demuestra el reducido porcentaje que sus entidades suponen sobre la población ocupada total de la región (1,4%). Dicha cifra, sensiblemente inferior a la media nacional (2,1%), queda, como se recoge en el Cuadro 1, muy por debajo de las registradas por regiones como el País Vasco (7%), Murcia (3,9%), Navarra (3,6%), Comunidad Valenciana (2,8%) o Andalucía (2,6%).

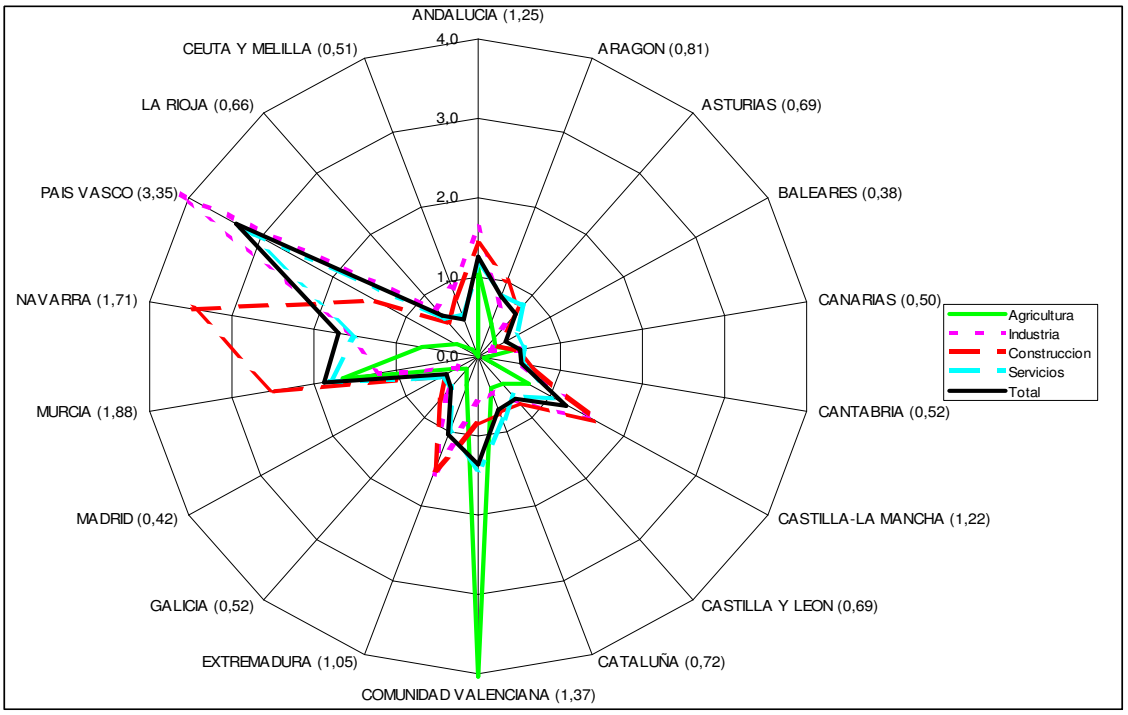
Cuadro 1: Población ocupada en Economía Social (cooperativas y sociedades laborales) respecto a población ocupada total, por sectores productivos a 31-12-2008 (%)

Regiones	*Agricultura*	*Industria*	*Construccion*	*Servicios*	***Total***
ANDALUCIA	5,4	5,0	2,4	2,0	***2,6***
ARAGON	2,2	2,2	1,8	1,5	***1,7***
ASTURIAS	1,4	1,5	1,3	1,4	***1,4***
BALEARES	1,2	0,5	0,5	0,9	***0,8***
CANARIAS	2,2	1,5	0,8	1,0	***1,0***
CANTABRIA	0,3	1,7	1,2	0,9	***1,1***
CASTILLA-LA MANCHA	3,4	4,7	2,8	1,8	***2,5***
CASTILLA Y LEON	2,2	2,2	1,3	1,1	***1,4***
CATALUÑA	2,1	1,5	1,3	1,5	***1,5***
COMUNIDAD VALENCIANA	19,8	1,8	1,5	2,5	***2,8***
EXTREMADURA	1,9	4,8	2,6	1,7	***2,2***
GALICIA	1,0	1,8	1,3	0,9	***1,1***
MADRID	1,5	0,8	0,8	0,9	***0,9***
MURCIA	8,0	3,7	4,3	3,2	***3,9***
NAVARRA	3,4	4,8	6,0	2,6	***3,6***
PAIS VASCO	1,4	12,5	2,4	5,7	***7,0***
LA RIOJA	0,6	2,4	1,0	1,1	***1,4***
CEUTA Y MELILLA	0,0	2,8	1,4	1,0	***1,1***

FUENTE: Elaboración propia a partir de los Datos de la Dirección General de la Economía Social, del trabajo autónomo y del Fondo Social Europeo. Ministerio de Trabajo y Asuntos Sociales.

• Por sectores productivos, destaca la creación de puestos de trabajo en los sectores secundario (el 22,6% del empleo total en Economía Social) y, sobre todo, terciario (el 58,3%). La positiva evolución registrada en los últimos años en todos los sectores se ha truncado en 2008, a consecuencia de la actual crisis, destacando el impacto en el sector de la construcción (con una caída del 34% respecto al año anterior, que le ha llevado a pasar del 12,6% al 8,9% del empleo total generado por cooperativas y sociedades laborales a escala nacional. Con todo, como muestra el citado Cuadro 1, a escala intrasectorial es la agricultura el sector con mayor grado de penetración de la Economía Social (especialmente visible en la Comunidad Valenciana), al suponer ésta el 4,9% del empleo agrario nacional. Le sigue la industria, donde sociedades laborales y cooperativas generan el 3,1% del empleo manufacturero nacional (con penetración máxima en el País Vasco). Los menores porcentajes corresponden a la construcción y al terciario, ambos con un 1,7% (donde destacan, respectivamente, los cifras de Navarra y el País Vasco). La incidencia de la Economía Social en Castilla y León es inferior a la media nacional en todos los sectores, resultando un menor grado de especialización general de la región en este ámbito respecto a España (0,69), en contraste con el País Vasco, Navarra, Murcia, la Comunidad Valenciana o Andalucía (Figura, 3).

Figura 3: Grado de especialización en el empleo en Economía Social (cooperativas y sociedades laborales) de las regiones españolas por sectores productivos a 31-12-2008 (%)



FUENTE: Elaboración propia a partir de los Datos de la Dirección General de la Economía Social, del trabajo autónomo y del Fondo Social Europeo. Ministerio de Trabajo y Asuntos Sociales.

3.1 La Economía Social en Castilla y León: rasgos fundamentales.

El análisis anterior pone de manifiesto que, en términos generales, la Economía Social en Castilla y León, en su segmento mercado, no alcanza la dimensión ni el grado de penetración en el tejido productivo de otras regiones de España. Debido a ello, no obstante, cabe albergar expectativas respecto a un amplio potencial de expansión futura de este sector en la región (Fernández y Gómez, 2007; Gómez y Román, 2004).

Aun admitiendo, en consecuencia, que su impacto territorial va a ser más limitado que en otras Comunidades Autónomas, la aproximación a la realidad de la Economía Social como

vehículo de desarrollo local en Castilla y León requiere un análisis más profundo; un examen, a nuestro juicio, que permita verificar de algún modo la plasmación de los puntos de convergencia identificados previamente (al menos, de buena parte de ellos).

Para efectuar dicha labor partimos de los resultados obtenidos en un estudio sobre la situación actual y rasgos de la Economía Social en Castilla y León, realizado con base en una encuesta lanzada a un universo de 2600 empresas6: Cooperativas (el 41,5% del total), Sociedades Laborales (52,9%) y Centros Especiales de Empleo (C.E.E.) (5,6%).

Una población empresarial que, como se recoge el Cuadro 2, se ubica preferentemente en las provincias de Valladolid y León - ambas con el 19,1% del total regional -, cuyo peso es más notorio en los sectores no agrarios de la región y en el número total de sociedades laborales y C.E.E.. Les siguen Salamanca y Burgos (16,8% y 14,6% del total, respectivamente), con la más amplia presencia de cooperativas, destacando en la agricultura e industria regionales; finalmente, con la menor participación, figuran, en general, las cinco provincias menos pobladas.

Cuadro 2: Distribución provincial de las empresas de Economía Social por sectores

	% s/total regional					% s/total provincial				
Provincias	Agricultura	Industria	Construcc.	Servicios	Total	Agricultura	Industria	Construcc.	Servicios	Total
AVILA	5,68	4,70	5,37	5,03	5,12	22,56	20,30	9,77	47,37	100,00
BURGOS	17,99	15,13	12,81	13,33	14,63	25,00	22,89	8,16	43,95	100,00
LEÓN	10,23	15,65	27,27	22,83	19,09	10,89	18,15	13,31	57,66	100,00
PALENCIA	10,23	8,52	6,20	6,30	7,58	27,41	24,87	7,61	40,10	100,00
SALAMANCA	24,05	16,52	15,70	14,13	16,82	29,06	21,74	8,70	40,50	100,00
SEGOVIA	6,06	4,52	2,07	3,91	4,31	28,57	23,21	4,46	43,75	100,00
SORIA	3,41	8,52	5,79	3,75	4,93	14,06	38,28	10,94	36,72	100,00
VALLADOLID	11,55	18,26	18,18	22,83	19,09	12,30	21,17	8,87	57,66	100,00
ZAMORA	10,80	8,17	6,61	7,90	8,43	26,03	21,46	7,31	45,21	100,00
CASTILLA Y LEÓN	100,00	100,00	100,00	100,00	100,00	20,32	22,13	9,31	48,23	100,00

FUENTE: Elaboración propia a partir de los datos de la encuesta.

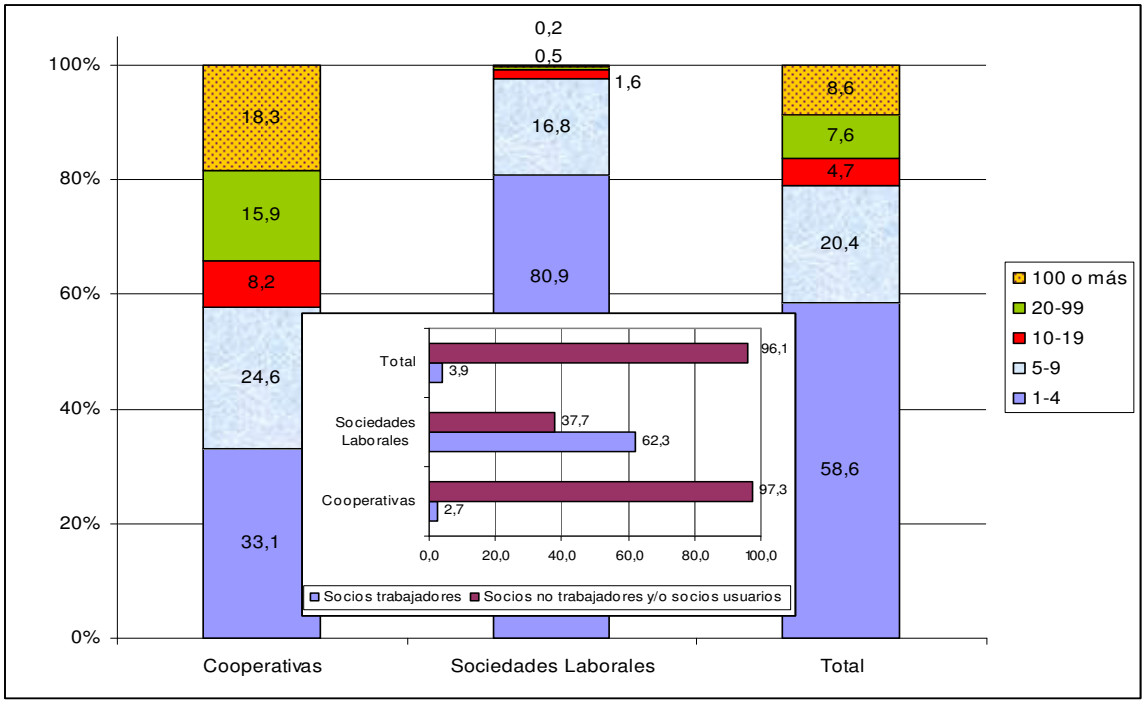
A escala interprovincial se detecta un importante grado de terciarización de la Economía Social, al situarse en el sector servicios el 48,2% de las empresas, cifra que se eleva al 56% si se tiene presente que hasta un 39% de las empresas que cabe calificar de agrarias tienen como actividad dominante la comercialización de productos agrarios. Dicha terciarización es especialmente intensa en Valladolid y León (57,7% de la Economía Social provincial en ambos casos), registrando el valor mínimo en Soria (36,7%), donde destaca el sector manufacturero, con una cifra - 38,3% - muy superior a la media regional (22,1%). El sector agrario sobresale en las provincias de Salamanca, Segovia, Palencia y Zamora, y la construcción registra su más alto porcentaje en León (13,3%).

Los datos de la muestra objeto de análisis, revelan una composición societaria muy diversa (Figura 4). Mientras el 81% de las sociedades laborales cuenta con un número de socios inferior a 5, dicha cifra se reduce al 33,1% para las cooperativas, pues algunas de sus categorías (agrarias, de vivienda, de crédito) cuentan con un elevado número de socios que conduce a un promedio de 202,6 socios por cooperativa (frente a los 4,7 socios de las sociedades laborales). No obstante, en las cooperativas los socios son, por lo común, no trabajadores y/o usuarios

[6] El proyecto, financiado por la Junta de Castilla y León, lleva por título "La Economía Social en Castilla y León. Estudio del sector empresarial. La encuesta fue efectuada de manera telefónica entre el 19 de junio y el 6 de julio de 2006. Respondieron un total de 1191 empresas, con una tasa de respuesta del 45% y un margen de error del ±2,25% para datos globales en condiciones normales de muestreo (p=q=0,5, 2σ), con un nivel de confianza del 95,5%). Se plantearon 4 tipos de cuestionarios: uno para Centros Especiales de Empleo, otro para Cooperativas de Vivienda, otro para Cooperativas de crédito y un cuarto para las demás clases de Cooperativas y las Sociedades Laborales. En cualquier caso, la encuesta constaba de 89 preguntas organizadas en 10 secciones: 1) Identificación; 2) Sector de actividad y dimensión productiva; 3) Estructura societaria y laboral; 4) Cualificación y formación de los recursos humanos; 5) Origen y características organizativas; 6) Gestión empresarial y nuevas tecnologías; 7) Mercado y estrategia empresarial; 8) Proyección externa y asociacionismo; 9) Relaciones con la Administración Pública; 10) Aspectos económicos y financieros.

(97,3%), en tanto que en las sociedades laborales el porcentaje de socios trabajadores alcanza el 62,3%, siendo socios capitalistas el resto.

Figura 4: Composición societaria de las cooperativas y sociedades laborales (%)



FUENTE: Elaboración propia a partir de los datos de la encuesta.

En todo caso, como refleja el Cuadro 3, el tamaño medio de estas entidades es muy reducido: 9,3 puestos de trabajo (inflado por la mayor dimensión de los C.E.E.), y con un claro predominio de las empresas que dan ocupación a menos de 4 personas (el 58,6% del total), sobre todo en el caso de las sociedades laborales. Se trata, pues, de PYMES - o, mejor, de microempresas -, de acuerdo con panorama productivo dominante en los procesos de desarrollo local, cuya impacto económico resulta más notable en las cooperativas (que generan el 52,5% del empleo y el 69% de la producción en Economía Social); en consonancia con este último dato, dicha categoría registra los mayores niveles de facturación por empresa y de productividad aparente del trabajo (2,1 millones de € y 200.000 €, respectivamente).

Cuadro 3: Contribución al empleo y la producción de la Economía Social (%)

Población ocupada	Cooperativas	Sociedades Laborales	Centros Esp. Empleo	Total
2 o menos	34,6	34,3	17,3	**32,9**
3-4	22,6	31,5	14,4	**25,7**
5-9	21,7	23,0	28,8	**22,9**
10-19	12,0	8,5	14,4	**10,7**
20-99	7,4	2,8	23,1	**6,8**
100 o más	1,6	0,0	1,9	**1,0**
Total	100,0	100,0	100,0	**100,0**
Tamaño medio	10,5	5,2	22,1	**9,3**
Contribución a empleo	52,5	32,7	14,8	**100,0**
Contribución a producción	*69,0*	*24,3*	*6,7*	***100,0***
Facturación/empresa (€)	*2116450,0*	*503988,4*	*1305729,2*	***1364805,1***
Productividad media tbjo.	*201566,7*	*96920,8*	*59082,8*	***146753,2***

FUENTE: Elaboración propia a partir de los datos de la encuesta.

Considerando el empleo asalariado, el Cuadro 4 corrobora lo anterior: así, casi un tercio de las cooperativas y sociedades laborales no emplean ningún trabajador por cuenta ajena, claro indicio del recurso al autoempleo, presentando casi el 75% menos de 4 trabajadores (cifra que roza las cuatro quintas partes en las sociedades laborales).

Cuadro 4: Estructura laboral por número de asalariados y tipos de contrato (%)

Total empleados	Cooperativas	Sociedades Laborales	Total
0	31,7	31,3	**31,5**
1-2	20,4	27,6	**23,8**
3-4	16,2	20,1	**18,0**
5-9	15,7	12,9	**14,4**
10-19	7,7	5,5	**6,7**
20-99	6,5	2,6	**4,7**
100 o más	1,7	0,0	**0,9**
Total	100,0	100,0	**100,0**
Contratados indefinidos	75,8	53,8	**70,5**
Contratados temporales	24,2	46,2	**29,5**

FUENTE: Elaboración propia a partir de los datos de la encuesta.

Respecto a la modalidad de contratación, algo más del 70% son contratos indefinidos, especialmente en las cooperativas, en cuyo seno la contratación temporal representa algo menos del 25%. Al **autoempleo** y relativa **estabilidad laboral** – propios del fenómeno del desarrollo local – les acompaña un apreciable grado de **democratización en el ámbito decisorio**, jugando la Asamblea General de socios un papel relevante en más del 35% de los casos en la adopción de decisiones de gestión, laborales y, sobre todo, de carácter estratégico (39,8%), seguida del Consejo Rector en las cooperativas y del Consejo de Administración en las sociedades laborales (Cuadro 5).

Cuadro 5: Responsabilidad de los órganos de la empresa en la toma de decisiones

Órganos de la empresa	De gestión	Laborales	Estratégicas
Director o gerente	22,1	20,2	16,8
Presidente o administrador	13,6	13,4	12,4
Consejo Rector o Consejo de Administración	29,0	30,5	31,0
Asamblea o Junta de Socios	35,3	36,0	39,8
Total	**100,0**	**100,0**	**100,0**

FUENTE: Elaboración propia a partir de los datos de la encuesta.

El nivel de formación dominante entre el personal de estas entidades son los estudios primarios (40,9%), siendo escaso el personal sin estudios (8,7%), sobre todo en las sociedades laborales (1,7%), donde destacan los estudios secundarios (43,8%). En los Consejos Rectores o de Administración predominan los miembros con estudios primarios o secundarios, así como los directivos de cooperativas y sociedades laborales; no así en el caso de los C.E.E., donde adquieren relevancia los estudios universitarios (al igual que en ciertas clases de cooperativas, como las de enseñanza y las de crédito).

Cuadro 6: Niveles formativos por tipos de personal y empresas (%)

Tipo de empresa	Nivel de estudios	Total por tipo de empresa (%)	Consejo Rector o de Admón.	Personal de dirección	Resto plantilla	Total
Cooperativas	Sin estudios	10,0	12,2	6,2	81,7	100,0
	Primarios	41,4	49,4	15,3	35,3	100,0
	Secundarios	25,1	39,4	18,5	42,1	100,0
	Universitarios	23,5	24,6	8,3	67,1	100,0
Sociedades Laborales	Sin estudios	1,7	40,6	25,0	34,4	100,0
	Primarios	37,3	29,7	17,2	53,1	100,0
	Secundarios	43,8	34,9	24,0	41,1	100,0
	Universitarios	17,3	37,4	30,8	31,7	100,0
Centros Especiales de Empleo	Sin estudios	12,1	1,3	0,0	98,7	100,0
	Primarios	43,7	3,7	1,5	94,8	100,0
	Secundarios	28,9	12,3	8,6	79,1	100,0
	Universitarios	15,4	27,7	25,1	47,1	100,0
TOTAL	Sin estudios	8,7				
	Primarios	40,9				
	Secundarios	29,0				
	Universitarios	21,4				

FUENTE: Elaboración propia a partir de los datos de la encuesta.

Respecto a la formación no reglada, el personal de más de la mitad de las empresas de Economía Social ha asistido a cursos de formación y/o perfeccionamiento profesional en los últimos dos años. La participación media es del 62% (de hecho, en casi el 50% de los casos ha asistido más del 60% de la plantilla, y más especialmente en los C.E.E. (Cuadro 7). Estos datos traslucen, no obstante, la existencia de importantes necesidades formativas[7], cara al fomento del **capital humano** como recurso local primordial.

Cuadro 7: Porcentaje de la plantilla participante en cursos de formación

Porcentaje de la plantilla participante en cursos	Cooperativas	Sociedades Laborales	Centros Esp. Empleo	Total
0-20	16,9	14,7	11,3	**15,4**
21-41	17,2	19,3	9,7	**17,3**
41-60	17,6	23,2	24,2	**20,6**
61-80	11,0	7,3	14,5	**9,8**
81-100	37,2	35,5	40,3	**36,8**
Total	**100,0**	**100,0**	**100,0**	**100,0**
% medio de participación	62,1	60,3	68,7	**62,0**

FUENTE: Elaboración propia a partir de los datos de la encuesta.

La antigüedad media de las empresas analizadas es de unos 12,5 años, siendo las cooperativas las más veteranas. La gran mayoría de cooperativas y sociedades laborales surgieron por iniciativa de los socios, en situación de paro en casi el 50% de los casos. Este **carácter reactivo** frente a la adversidad, consustancial al inicio de todo proceso de desarrollo local, sigue marcando el comportamiento estratégico de más de un tercio de estas entidades (sobre todo cooperativas), que tienen como prioridad la supervivencia de la empresa, afrontando la presión del mercado mediante los recursos disponibles.

Con todo, conviene señalar el comportamiento prospectivo, abierto a innovaciones y nuevos mercados, como elemento distintivo en más del 25% de los casos (sociedades laborales y C.E.E.). De hecho, en relación con la innovación (Cuadro 8), una notable proporción de empresas (50,6%) declara haber innovado recientemente (hasta el 68,3% en los C.E.E.), con un claro predominio de la innovación de proceso sobre las de producto y organizativas (ambas, de nuevo, con un peso llamativo en los C.E.E.).

[7] El contenido de los cursos, en general, viene girando en torno a cuestiones técnicas en materia de producción y comercialización, a aspectos organizativos, de gestión y de administración y a aspectos relacionados con las nuevas tecnologías (informática y comunicaciones), siendo en su mayoría gestionados por una Asociación u organización representativa del sector (en cooperativas y sociedades laborales) y, en menor medida, por la propia empresa mediante un Plan de Formación (en el caso de los C.E.E.). Llama la atención el hecho de que casi la mitad de las cooperativas de la región no haya realizado aportación alguna al Fondo de Educación y Promoción Cooperativa en los tres últimos años. Al menos, entre las que sí han contribuido el destino principal fue la formación de los socios (con un 42% del total de recursos financieros del citado Fondo).

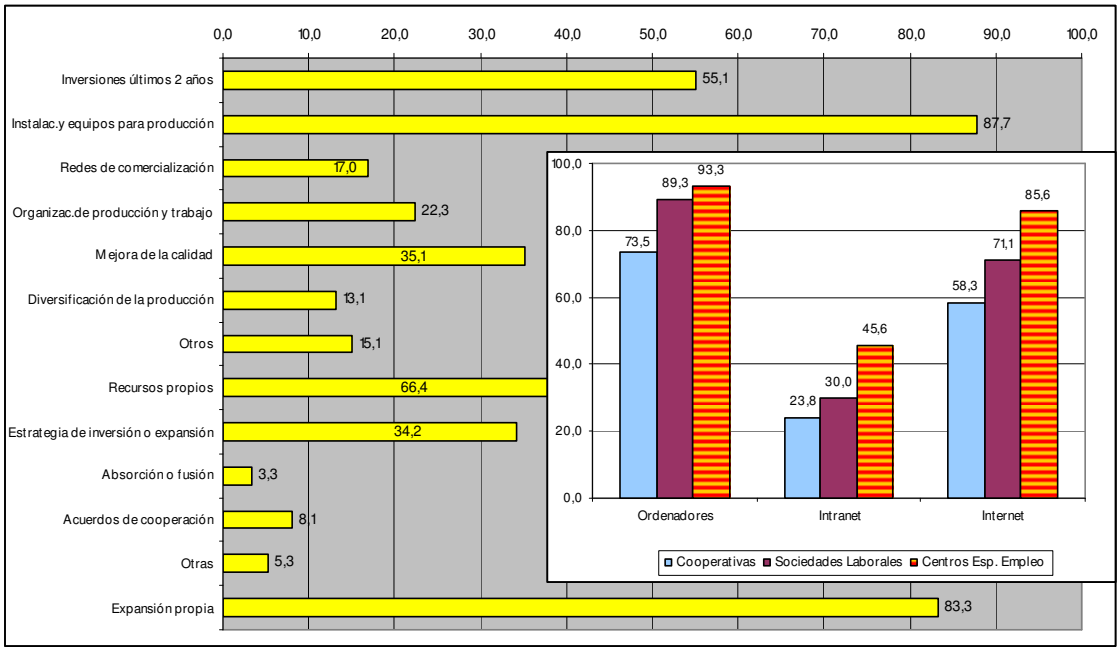
Menos frecuente resulta la realización de actividades de I+D, presentes sobre todo en los C.E.E. (52,9%), aunque son las cooperativas quienes registran el mayor porcentaje de empresas que, realizando tales actividades, protegen sus resultados (22,7%). Esta apertura a la innovación concuerda con una alta penetración de las TICs (Figura 5): las empresas cuentan con buen equipamiento informático (sobre todo C.E.E. y sociedades laborales), fruto de su política inversora, notablemente activa en los últimos dos años en el 55% de las firmas, y con las instalaciones y equipos de producción como destino primordial en el 87,7% de los casos, seguido de la mejora en la calidad (35,1%). Los recursos propios constituyen la principal fuente de financiación de las inversiones (en el 66,4% de los casos).

Cuadro 8: Actividad empresarial innovadora en los últimos dos años (%)

Concepto	Cooperativas	Sociedades Laborales	Centros Esp. Empleo	Total
Alguna forma de innovación	51,4	46,0	68,3	**50,6**
Innovación organizativa	17,6	17,8	23,1	**19,3**
Innovación de producto	18,2	17,8	35,6	**18,8**
Innovación de proceso	42,8	33,4	26,9	**39,6**
Actividades de I+D (*)	14,3	16,4	52,9	**16,0**
Empresas (*) que protegen su I+D	22,7	17,7	18,4	**20,2**

FUENTE: Elaboración propia a partir de los datos de la encuesta.

Figura 5: Política inversora y equipamiento en materia de TICs (%)



FUENTE: Elaboración propia a partir de los datos de la encuesta.

Las previsiones acerca de una expansión futura no son muy frecuentes (34,2% de los casos), apuntando a la expansión propia como modalidad preferida.

Innovación e inversión reflejan, de cualquier modo, la presencia de una cierta **visión de largo plazo** – inherente al desarrollo local – que sería más sólida de potenciarse una estrategia competitiva apoyada en la diferenciación del producto y no tanto vía costes.

En esa labor de fomento podría jugar un papel importante, sin duda, el apoyo público. Como refleja el Cuadro 9, un 56% de las empresas (el 92% de los C.E.E.) reconoce haber recibido subvenciones en los últimos tres años, procedentes, en su mayoría, de la Administración Regional (hasta el 87,9% de las firmas que declaran contar con ayudas).

El destino principal de las ayudas lo constituyen la inversión en maquinaria y/o instalaciones y el fomento del empleo. Esa **permeabilidad respecto a la iniciativa pública** – palanca de desarrollo local – queda patente si se tiene en cuenta que el apoyo obtenido ha sido

determinante en la elección de la forma jurídica de la empresa (sobre todo en las sociedades laborales); en todo caso, las principales demandas de servicios públicos por parte de las empresas giran en torno a la mejora de los sistemas de financiación y de los canales de información.

Cuadro 9: Ayudas de las Administraciones Públicas (%)

Concepto	Cooperativas	Sociedades Laborales	Centros Esp. Empleo	Total
Reciben ayuda pública	***56,8***	***47,8***	***92,3***	***56,1***
De entidades locales	7,9	7,4	7,4	7,6
De Junta de Castilla y León	85,5	86,8	99,0	87,9
De alguna entidad nacional	10,3	9,9	7,3	9,7
De entidades europeas	30,6	7,9	8,3	19,2
De otras entidades	4,8	4,1	5,3	4,6
Servicios demandados a las A.A.P.P.				
Información	60,5	58,7	54,8	59,2
Seguimiento de proyecto empresa	29,6	32,0	29,4	30,6
Ayuda a obtención de financiación	63,6	68,4	68,0	66,0
Ayuda integral	25,3	27,5	38,8	27,4
Otros	11,5	9,7	12,7	10,8

FUENTE: Elaboración propia a partir de los datos de la encuesta.

Un importante número de empresas que operan en el sector terciario tiene como función principal la realización de actividades enmarcadas en los llamados **nuevos yacimientos de empleo**, generando oportunidades para una mayor **diversificación económica a escala local**. Destacan, en este sentido, los *servicios a la vida cotidiana* (servicios a domicilio, atención a la infancia, inserción laboral de jóvenes en dificultades (Cuadro 10) y, en menor medida, los *servicios de mejora de la calidad de la vida* (seguridad y vigilancia, transportes colectivos locales, revalorización de espacios urbanos,…) y los *servicios orientados al ocio*(turismo rural, valorización del patrimonio, deporte,…).

Cuadro 10: Nuevos yacimientos de empleo por tipos de empresas (%)

Ámbitos ocupacionales (categorías de servicios)	Cooperativas	Sociedades Laborales	Centros Esp. Empleo	Total
Servicios a la vida cotidiana	17,4	21,0	88,9	**57,1**
De mejora de la calidad de vida	52,2	40,3	11,1	**26,1**
De ocio	30,4	33,9	0,0	**15,2**
Medioambientales	0,0	4,8	0,0	**1,6**
Total	**100,0**	**100,0**	**100,0**	**100,0**

FUENTE: Elaboración propia a partir de los datos de la encuesta.

La presencia de servicios medioambientales es muy inferior; con todo, la valoración que estas entidades realizan de su **compromiso con el desarrollo sostenible** es notable (Cuadro 11), y muy similar a la referida a su respeto por los valores democráticos y participativos y a su **contribución al desarrollo y cohesión social a escala local**.

Cuadro 11: Valoración de las ventajas de la Economía Social (Escala de 1 a 5)

Empresas	*Compromiso con el medio ambiente*	*Responsabilidad social de la empresa*	*Desarrollo y cohesión social a nivel local*	*Valores solidarios democráticos y participativos*	*Aplicación del princ. de igualdad de oportunidades*	*Integración sociolaboral de colect.desfavorecidos*
Cooperativas	4,1	4,1	3,9	3,9	4,0	3,6
Soc. laborales	4,2	4,3	4,0	3,9	4,2	3,9
C. E. Empleo	3,9	4,4	4,0	4,0	4,4	4,7
TOTAL	4,1	4,2	3,9	3,9	4,1	3,8

FUENTE: Elaboración propia a partir de los datos de la encuesta.

A tales valores no resulta ajena la **política de reinversión del excedente**: el 72,3% tiene como destino la autofinanciación de la empresa, frente a un 17% de retorno a los socios.

Cuadro 12: Destino del beneficio neto (porcentaje medio por empresa)

Destino de excedentes (%)	Cooperativas	Sociedades Laborales	Total
Retorno a socios	24,6	9,0	**17,2**
Autofinanciación de la empresa	63,9	81,5	**72,3**
Otros destinos	11,6	9,5	**10,6**
Total	100,0	100,0	**100,0**

FUENTE: Elaboración propia a partir de los datos de la encuesta.

Dicha reinversión recogida en el Cuadro 12, fortalece el **vínculo de estas entidades con el territorio**, elevado ya de por sí, merced al origen local de la mayoría de empresarios y trabajadores y al carácter de la zona de asentamiento como principal punto de venta para casi un tercio de las empresas, cifra que alcanza casi el 58,6% si se incorpora la provincia (Cuadro 13), concentrando ambos ámbitos de mercado el 63% de las ventas. La mayor amplitud de mercado se registra en los C.E.E. En cualquier caso, el porcentaje que supone la actividad exportadora sobre la facturación total es muy reducido.

Cuadro 13: Ámbitos de mercado (%)

Ámbito de mercado	Cooperativas	Sociedades Laborales	Centros Esp. Empleo	Total
Local	34,6	34,3	17,3	**32,9**
Provincial	22,6	31,5	14,4	**25,7**
Regional	21,7	23,0	28,8	**22,9**
Nacional	12,0	8,5	14,4	**10,7**
Unión Europea	7,4	2,8	23,1	**6,8**
Resto del mundo	1,6	0,0	1,9	**1,0**
Total	100,0	100,0	100,0	**100,0**
% Exportación/volumen de ventas	3,3	1,2	0,5	**2,1**

FUENTE: Elaboración propia a partir de los datos de la encuesta.

La incidencia sobre el territorio resulta, en cualquier caso, ampliamente variada si se considera la distribución municipal de la población total de empresas de la Economía Social de Castilla y León. Éstas se hallan ubicadas en poco más de 600 municipios, que si bien aglutinan el 84% de la población y el 87% de la ocupación regionales, tan suponen el 27% de los 2248 municipios de la región. En esta línea, del examen del Cuadro 14 se deduce lo siguiente:

a) Las localidades de más de 30000 habitantes – las capitales de provincia más Ponferrada, Miranda de Ebro y Aranda de Duero (el 2% de los municipios con presencia de las empresas de Economía Social, y casi el 60% de la población y de la ocupación del universo considerado) – reúnen conjuntamente el 46,2% del número total de empresas. Les sigue en importancia el colectivo formado por las localidades menores de 500 habitantes (el 41,5% del total), – donde se ubica el 12,4% de las entidades, cifra que asciende al 20,9% si se consideran los municipios menores de 1000 habitantes.

b) El primer grupo domina todos los sectores (y en especial, los servicios), salvo la agricultura, donde sobresale el segundo bloque (tanto en empresas agrarias de explotación común (*Cód. 1*) como en las dedicadas a la comercialización al por mayor de productos agrarios (*511*). En la actividad forestal (*2*) despunta el grupo de entre 1000 y 5000 habitantes.

c) Por subsectores, fuera del primer bloque municipal mencionado, destacan la industria agroalimentaria (*15*), dominante en los municipios menores de 500 habitantes, los sectores textil-confección (*17/18*), química (*24*), material de transporte (*34-35*) telecomunicaciones (*64*)

y alquileres (*74*) en los municipios de 2001 a 5000, la industria maderera en el grupo de 1001 a 2000 y otros servicios como las actividades asociativas (*91*) (501-1000), transporte terrestre (*60*) (10001-30000) y actividades recreativas, culturales y deportivas (*92*) (5001-10000).

Cuadro 14: Distribución intermunicipal de las empresas de Economía Social CNAE-93

Nº habitantes	<500	501-1000	1001-2000	2001-5000	5001-10000	10001-30000	30001-50000	capitales	Total
Población	2,98	4,49	6,82	9,52	9,21	9,00	6,47	51,49	100,00
Ocupados	2,85	4,20	6,42	11,60	8,54	8,26	6,15	51,99	100,00
Localidad	41,53	22,15	16,78	10,75	4,89	1,95	0,49	1,47	100,00
511	28,64	15,53	15,53	12,14	6,80	2,43	3,40	15,53	100,00
1	45,90	23,93	12,46	6,89	4,26	0,98	0,66	4,92	100,00
2	11,76	11,76	35,29	23,53	0,00	0,00	0,00	17,65	100,00
Agricultura	***38,07***	***20,27***	***14,39***	***9,47***	***5,11***	***1,52***	***1,70***	***9,47***	***100,00***
15	24,85	13,02	13,61	14,79	10,65	4,14	3,55	15,38	100,00
17/18	11,76	16,47	15,29	22,35	5,88	9,41	1,18	17,65	100,00
19	0,00	0,00	0,00	0,00	100,00	0,00	0,00	0,00	100,00
20	10,53	15,79	21,05	13,16	13,16	0,00	5,26	21,05	100,00
21-22	0,00	0,00	12,77	10,64	10,64	2,13	4,26	59,57	100,00
24	0,00	0,00	0,00	60,00	0,00	0,00	40,00	0,00	100,00
25	0,00	0,00	0,00	0,00	0,00	0,00	33,33	66,67	100,00
26	5,26	21,05	15,79	15,79	5,26	10,53	5,26	21,05	100,00
27/28	5,49	9,89	6,59	14,29	12,09	8,79	17,58	25,27	100,00
29	0,00	17,65	5,88	5,88	11,76	0,00	5,88	52,94	100,00
31-33	0,00	0,00	7,69	7,69	15,38	0,00	0,00	69,23	100,00
34-35	14,29	14,29	0,00	28,57	14,29	14,29	0,00	14,29	100,00
36	6,49	10,39	16,88	11,69	16,88	16,88	2,60	18,18	100,00
Industria	***11,83***	***11,65***	***12,87***	***14,96***	***11,13***	***6,96***	***6,09***	***24,52***	***100,00***
Construcción	***4,96***	***5,79***	***8,68***	***17,36***	***7,85***	***10,33***	***10,74***	***34,30***	***100,00***
50	3,92	2,94	7,84	9,80	7,84	9,80	12,75	45,10	100,00
51	3,95	3,39	8,47	11,86	7,91	5,65	7,34	51,41	100,00
52	2,52	0,72	3,24	6,83	5,04	9,71	5,76	66,19	100,00
55	7,58	6,82	9,85	7,58	7,58	6,82	3,79	50,00	100,00
60	0,00	13,51	10,81	5,41	5,41	16,22	5,41	43,24	100,00
63	0,00	0,00	0,00	4,55	9,09	0,00	0,00	86,36	100,00
64	0,00	0,00	0,00	20,00	10,00	0,00	0,00	70,00	100,00
65-67	3,23	0,00	3,23	0,00	9,68	0,00	0,00	83,87	100,00
70	0,00	2,67	2,67	1,33	1,33	0,00	4,00	88,00	100,00
71	0,00	0,00	0,00	25,00	0,00	0,00	0,00	75,00	100,00
72	2,50	0,00	5,00	7,50	2,50	5,00	10,00	67,50	100,00
74	3,00	1,50	4,00	4,50	6,00	7,00	7,50	66,50	100,00
80	4,00	0,00	0,00	4,00	12,00	0,00	0,00	80,00	100,00
85	1,72	1,72	3,45	13,79	3,45	5,17	8,62	62,07	100,00
91	0,00	50,00	0,00	0,00	0,00	0,00	0,00	50,00	100,00
92	4,76	4,76	4,76	14,29	23,81	14,29	9,52	23,81	100,00
93	2,56	0,00	0,00	2,56	0,00	12,82	12,82	69,23	100,00
Servicios	***3,19***	***2,63***	***5,19***	***7,34***	***6,23***	***7,10***	***6,62***	***61,69***	***100,00***
TOTAL	**12,36**	**8,51**	**9,08**	**10,39**	**7,24**	**6,24**	**5,89**	**40,30**	**100,00**

FUENTE: Elaboración propia a partir de los datos de la encuesta.

El Cuadro 15 viene a reforzar buena parte de las afirmaciones anteriores. Así, la actividad agraria viene a caracterizar el tejido productivo en Economía Social de los municipios eminentemente rurales, los menores de 2000 habitantes (particularmente en los de menos de 500, donde un 62,6% de sus empresas están vinculadas al sector primario). Algo similar ocurre con actividades agroindustriales como la industria agroalimentaria y, en menor medida, la maderera (ésta última en los municipios de entre 501 y 2000). Otras ramas manufactureras con peso en las economías locales son la de textil-confección (501-5000), la del mueble (*36*) (5001-30000) y la de transformados metálicos (*27/28*) que, junto a la construcción alcanza las cifras más relevantes en la Economía Social de los municipios intermedios (de más de 2000 habitantes no capitales de provincia). Las actividades terciarias destacan con diferencia en los municipios mayores de 10000 habitantes, figurando con los porcentajes más destacados en todas las categoría municipales el comercio (*51/52*), especialmente al por menor (*52*), otros servicios (limpieza, servicios profesionales,…) (*74*) y la hostelería (*55*).

Cuadro 15: Distribución intramunicipal de las empresas de Economía Social CNAE-93

Nº habitantes	<500	501-1000	1001-2000	2001-5000	5001-10000	10001-30000	30001-50000	capitales	Total
Empr/pobl.	0,51	0,23	0,16	0,13	0,10	0,09	0,11	0,10	0,12
Empr/ocup.	1,40	0,66	0,46	0,29	0,27	0,24	0,31	0,25	0,32
Empr/munic.	1,26	1,63	2,29	4,09	6,27	13,50	51,00	116,33	4,23
511	18,38	14,48	13,56	9,26	7,45	3,09	4,58	3,06	7,93
1	43,61	33,03	16,10	7,78	6,91	1,85	1,31	1,43	11,74
2	0,62	0,90	2,54	1,48	0,00	0,00	0,00	0,29	0,65
Agricultura	***62,62***	***48,42***	***32,20***	***18,52***	***14,36***	***4,94***	***5,88***	***4,78***	***20,32***
15	13,08	9,95	9,75	9,26	9,57	4,32	3,92	2,48	6,51
17/18	3,12	6,33	5,51	7,04	2,66	4,94	0,65	1,43	3,27
19	0,00	0,00	0,00	0,00	0,53	0,00	0,00	0,00	0,04
20	1,25	2,71	3,39	1,85	2,66	0,00	1,31	0,76	1,46
21-22	0,00	0,00	2,54	1,85	2,66	0,62	1,31	2,67	1,81
24	0,00	0,00	0,00	1,11	0,00	0,00	1,31	0,00	0,19
25	0,00	0,00	0,00	0,00	0,00	0,00	1,31	0,38	0,23
26	0,31	1,81	1,27	1,11	0,53	1,23	0,65	0,38	0,73
27/28	1,56	4,07	2,54	4,81	5,85	4,94	10,46	2,20	3,50
29	0,00	1,36	0,42	0,37	1,06	0,00	0,65	0,86	0,65
31-33	0,00	0,00	0,42	0,37	1,06	0,00	0,00	0,86	0,50
34-35	0,31	0,45	0,00	0,74	0,53	0,62	0,00	0,10	0,27
36	1,56	3,62	5,51	3,33	6,91	8,02	1,31	1,34	2,96
Industria	***21,18***	***30,32***	***31,36***	***31,85***	***34,04***	***24,69***	***22,88***	***13,47***	***22,13***
Construcción	***3,74***	***6,33***	***8,90***	***15,56***	***10,11***	***15,43***	***16,99***	***7,93***	***9,31***
50	1,25	1,36	3,39	3,70	4,26	6,17	8,50	4,39	3,93
51	2,18	2,71	6,36	7,78	7,45	6,17	8,50	8,69	6,81
52	2,18	0,90	3,81	7,04	7,45	16,67	10,46	17,57	10,70
55	3,12	4,07	5,51	3,70	5,32	5,56	3,27	6,30	5,08
60	0,00	2,26	1,69	0,74	1,06	3,70	1,31	1,53	1,42
63	0,00	0,00	0,00	0,37	1,06	0,00	0,00	1,81	0,85
64	0,00	0,00	0,00	0,74	0,53	0,00	0,00	0,67	0,38
65-67	0,31	0,00	0,42	0,00	1,60	0,00	0,00	2,48	1,19
70	0,00	0,90	0,85	0,37	0,53	0,00	1,96	6,30	2,89
71	0,00	0,00	0,00	0,37	0,00	0,00	0,00	0,29	0,15
72	0,31	0,00	0,85	1,11	0,53	1,23	2,61	2,58	1,54
74	1,87	1,36	3,39	3,33	6,38	8,64	9,80	12,70	7,70
80	0,31	0,00	0,00	0,37	1,60	0,00	0,00	1,91	0,96
85	0,31	0,45	0,85	2,96	1,06	1,85	3,27	3,44	2,23
91	0,00	0,45	0,00	0,00	0,00	0,00	0,00	0,10	0,08
92	0,31	0,45	0,42	1,11	2,66	1,85	1,31	0,48	0,81
93	0,31	0,00	0,00	0,37	0,00	3,09	3,27	2,58	1,50
Servicios	***12,46***	***14,93***	***27,54***	***34,07***	***41,49***	***54,94***	***54,25***	***73,83***	***48,23***
TOTAL	**100,00**	**100,00**	**100,00**	**100,00**	**100,00**	**100,00**	**100,00**	**100,00**	**100,00**

FUENTE: Elaboración propia a partir de los datos de la encuesta.

Aunque la ratio empresas de Economía Social por municipio resulta muy superior en las capitales de provincia y localidades mayores de 30000 habitantes, su importancia se ve obviamente diluida en su más amplio tejido económico. De hecho, es el paisaje productivo de los municipios rurales el que se ve más intensamente afectado por la Economía Social, como revelan los mayores coeficientes de empresas por habitante y por ocupado en los municipios con menos de 2000 habitantes.

Algunas localidades mayores de 2000 también muestran ratios dignos de atención, como Vitigudino, Paredes de Nava, Villares de la Reina, Peñaranda de Bracamonte, Ágreda, Almazán e incluso Aranda de Duero; todas ellas presentan una cierta entidad como plataformas para la industrailización local, especialmente Aranda, en su calidad de sistema local. Con todo, donde se detecta la más peculiar correlación entre el fenómeno del desarrollo local y la Economía Social es en los SPL de Medina del Campo, Tierra de Pinares (Con Covaleda y Duruelo de la Sierra en cabeza), Íscar o la Ribera del Duero burgalesa, con un significativo número de empresas de cooperativas y sociedades laborales centradas en la especialización productiva típica a escala local.

4. CONCLUSIONES

A la luz del análisis efectuado, cabe extraer las siguientes conclusiones:

1.El desarrollo local, modelo organizativo de carácter eminentemente participativo (fundamentado en el papel activo del territorio, en la utilización de los recursos del área y en la

adaptabilidad de las PYMES, y focalizado, principalmente, en pequeños núcleos urbanos y ciudades medias) se ha ido configurando como una forma de acumulación y de regulación flexible generadora de prosperidad, y convirtiendo en una alternativa al tradicional patrón de concentración/difusión urbano/industrial.

2.La continuidad del éxito económico de los llamados SPL (espacios típicos de ese modelo descentralizado de desarrollo), así como su supervivencia, se halla actualmente supeditada a su capacidad para superar los retos de la globalización. Cobran así una especial relevancia como fuentes de competitividad territorial el saber contextual generador de ideas a escala local, la receptividad al cambio y la capacidad de captar conocimiento codificado allende el entorno inmediato, a la vez que el diseño de estrategias institucionales de planificación estratégica y empresariales, encaminadas a superar problemas como las debilidades de las PYMES, el insuficiente apoyo y descoordinación institucionales, el abandono del medio rural, las crecientes exigencias de respeto al medio ambiente, etc.

3.Las entidades que configuran el sector de la Economía Social desempeñan una serie de funciones que - más allá de su genérico impulso hacia una sociedad más plural, participativa, democrática y solidaria - muestran su gran trascendencia económica. Así, cabe destacar su contribución a una mayor equidad en el reparto de la renta y la riqueza, su acción correctora de los desequilibrios en el mercado laboral, su vinculación a actividades productivas ligadas a las necesidades sociales, su fomento de la estabilidad macroeconómica y el desarrollo sostenible, así como su capacidad para estimular procesos de desarrollo local/rural, fenómeno con el cual presenta considerables sinergias.

4.La proyección territorial del desarrollo local en España (focalizada mayormente en torno a la industria) es mucho más intensa que en Castilla y León, donde los SPL constituyen un reducido número de experiencias. Dicha debilidad guarda, al menos de manera indirecta, una cierta correlación con el menor peso de la Economía Social – en particular de su subsector de mercado - en el tejido productivo de la región, dando como resultado un impacto territorial más limitado y modesto.

5.Esto no significa que la estrategia del desarrollo local no revista interés para Castilla y León, en especial para su mundo rural; ni tampoco que la Economía Social no pueda jugar un papel más importante en el futuro, dado el amplio margen existente para su expansión hasta alcanzar la media nacional. De hecho, el análisis del sector de mercado de la Economía Social revela la presencia de importantes elementos de convergencia con el desarrollo local: base en la PYME, autoempleo y estabilidad laboral; uso del potencial endógeno; valores democráticos; orientación hacia el largo plazo y hacia la sostenibilidad; nuevos yacimientos de empleo intensivos en trabajo (asistencia a personas dependientes, enseñanza, servicios de limpieza, agricultura ecológica, turismo rural y agroturismo, inserción laboral y lucha contra la exclusión social); vinculación al territorio; apertura al complemento de la iniciativa pública,…

6.Dicho sector de mercado presenta, no obstante, debilidades que limitan su potencial como vehículo para el desarrollo local, como son el minifundismo empresarial, el modesto nivel de formación de la plantilla de las empresas analizadas, del que se derivan necesidades formativas en materia técnica en directivos y trabajadores, y la infrecuente tendencia a la cooperación (ligada a una mentalidad marcadamente individualista), que se traduce en una escasa importancia del asociacionismo.

7.El esfuerzo en los frentes formativo, tecnológico y cooperativo resulta, pues, crucial al objeto de fortalecer la dimensión y el papel de la Economía Social de mercado en la región cara al futuro. Esta labor no sólo compete a las entidades analizadas, sino también a los poderes públicos, a través de actuaciones orientadas a: potenciar su constitución; a mejorar sus mecanismos de financiación y demás servicios de apoyo; fomentar la apuesta por las nuevas tecnologías y la implantación de una cultura de la innovación y de la calidad y excelencia empresarial; aumentar su nivel de capital humano; impulsar la cooperación e integración interempresarial

8.En este contexto pueden desempeñar funciones relevantes a escala local figuras como los agentes de empleo y desarrollo local y, sobre todo, los llamados Grupos de Acción Local (GAL), los cuales, como Asociaciones sin ánimo de lucro, constituyen una pieza fundamental de la Economía Social en su sector de no mercado en España. El examen de la actuación de

estos GAL – 44 en Castilla y León – en tanto que gestores (en el escenario de la actual política territorial europea) de un amplio volumen de recursos financieros con fines de desarrollo local/comarcal y como exponentes de partenariado público-privado – resultaría, indudablemente, enriquecedor como complemento del presente análisis centrado en el sector de mercado. En todo caso, cualquier consideración a realizar en futuras investigaciones tendrá que incorporar necesariamente el impacto de la actual crisis económico-financiera sobre la situación y perspectivas de la Economía Social castellana y leonesa, con inevitables repercusiones sobre su potencial como factor al servicio del desarrollo local en la región.

BIBLIOGRAFÍA

BAREA, J. Y MONZÓN, J.L., (dir.) (2002): *La Economía Social en España en el año 2000.* CIRIEC-España, Valencia.

BECATTINI, G., COSTA, M.T. Y TRULLÉN, J. (coord.) (2002): *Desarollo local: teorías y estrategias*. Civitas, Madrid.

BENKO, G. Y LIPIETZ, A. (eds.) (1992): *Les régions qui gagnent. Districts et réseaux: les nouveaux paradigmes de la géographie économique*. PUF, París.

BUENDÍA, I. Y GARCÍA, J.V. (coord.) (2003): *Cooperativismo y desarrollo local.* Centro de Estudios Ramón Areces-Fundación Luis Vives, Madrid.

CASTILLO, J. DEL ET AL. (1994), Barroeta, B., Bayón, M.M. y Cordero, E. (1994): *Manual de desarrollo local*. Gobierno Vasco, Departamento de Economía y Hacienda, Bilbao.

CEP (1993): *EXCEL. Cooperación entre empresas y Sistemas productivos locales*. IMPI-Ministerio de Industria Comercio y Turismo, Madrid.

CHAVES, R. (2006): "La Economía Social en España: concepto, agentes y políticas públicas". En Pérez de Uralde, J.Mª (coord.): *La Economía Social en Iberoamérica*. Fundación Iberoamericana de la Economía Social, Madrid, pp. 215-269.

CHAVES, R. Y MONZÓN, J.L. (2001): "Economía Social y sector no lucrativo: Actualidad científica y perspectivas". CIRIEC-España, *Revista de Economía Pública, Social y Cooperativa*, 37, pp. 7-33.

CHAVES, R. Y MONZÓN, J.L. (2003): "La Economía Social y la Política Económica". En Jordán Galduf, J.Mª et al. (coord.): *Política Económica: fundamentos, objetivos e instrumentos*, Tirant lo Blanch, Valencia, pp. 363-389.

COQUE, J. (2003): "El desarrollo local sobre bases cooperativas. Valorización de los recursos locales y creación de redes". En Buendía, I. y García, J.V. (coord.) (2003): *Cooperativismo y desarrollo local*. C.E. R. Areces-F. L. Vives, Madrid, pp. 261-299.

FERNÁNDEZ, J.E. Y GÓMEZ, J. Mª (2007): *La Economía Social en Castilla y León. Estudio del sector empresarial*. Ed. Sever-Cuesta. Junta de Castilla y León. Valladolid.

GAROFOLI, G. Y MAZZONI, R. (1994): "I sistemi produttivi locali: un'introduzione". En Garofoli, G. y Mazzoni, R. (coords.): *Sistemi produttivi locali: struttura e trasformazione*, Franco Angeli, Milán, pp. 7-24.

GÓMEZ, J:Mª Y ROMÁN, A. (2004): "La Economía Social en Castilla y León: un sector clave para generar empleo y luchar contra la exclusión laboral*". IV Congreso de Castilla y León sobre Mercado de Trabajo y Relaciones Laborales*, Palencia.

JUSTE, J.J. (2001): *Desarrollo Local y Mercado Global: los sistemas productivos locales y la industria agroalimentaria en Castilla y León*. Tesis Doctoral, Departamento de Economía Aplicada. Universidad de Valladolid.

JUSTE, J.J. ET AL. (2007): "Los sistemas productivos locales industriales en Castilla y León". En *Anales de Economía Aplicada 2007*, ASEPELT-Delta publicaciones, Madrid, pp. 83-104.

MONTERO, A. (1997): "Cooperativismo y desarrollo rural". *REVESCO*, 63, pp. 63-77.

MOZAS, A. Y BERNAL, E. (2006): "Desarrollo territorial y economía social". *CIRIEC-España, Revista de Economía Pública, Social y Cooperativa*, 55, agosto, pp. 125-140.

PÉREZ RAMÍREZ, B. Y CARRILLO BENITO, E. (2000): *Desarrollo local: Manual de uso*. Esic Editorial/Federación Andaluza de Municipios y Provincias, Madrid.

SANCHIS, J.R. Y CAMPOS, V. (2005): "Inserción sociolaboral, economía social y desarrollo local. Estudio empírico sobre la realidad actual del agente de empleo y desarrollo local en España". *CIRIEC-España, Revista de Economía Pública, Social y Cooperativa*, 52, agosto, pp. 279-306.

TOSCANO, F. (2000): "Desarrollo local y Economía Social". En Pérez Ramírez, B. y Carrillo Benito, E.: *Desarrollo local: Manual de uso*, Esic Editorial/Federación Andaluza de Municipios y Provincias, Madrid, pp. 445-470.

VÁZQUEZ BARQUERO, A. (1988): *Desarrollo local: Una estrategia de creación de empleo.* Pirámide, Madrid.

VÁZQUEZ BARQUERO, A. (2005): *Las nuevas fuerzas del desarrollo*. Pirámide, Madrid.

PAUTAS DE CONSUMO EN LOS HOGARES DE LA TERCERA EDAD: UN ANÁLISIS DESDE LA PERSPECTIVA DEL GASTO EN SERVICIOS

MANUEL HERNÁNDEZ PEINADO
E-mail: mhpeinad@ugr.es
JOSÉ A. CAMACHO BALLESTA
E-mail: jcamacho@ugr.es
Instituto de Desarrollo Regional y Departamento de Economía Internacional y de España
Universidad de Granada

Resumen

Las necesidades específicas de las personas mayores requieren diferentes líneas de consumo, como puede ser el consumo que sirve para preservar la salud o para restaurarla, los bienes y servicios que permiten seguir atendiendo el propio hogar, o los bienes y servicios que contribuyan a ocupar el tiempo libre.
El acelerado incremento de las personas mayores que se está produciendo en Europa y en particular en España, obliga a dedicar una atención especial al estudio del gasto realizado por los mayores, haciendo hincapié en el consumo privado de servicios, por tratarse de un aspecto no suficientemente analizado en la literatura económica.
En el contexto anterior se trata de llevar a cabo un análisis de los patrones de consumo de los hogares de la tercera edad, mediante una explotación de los datos contenidos en la Encuesta de Presupuestos Familiares de 1998 y 2005, mediante la utilización de técnicas estadísticas multivariantes.

Palabras Clave: Consumo, Hogares, Servicios, Tercera Edad, Presupuestos Familiares.
Área Temática: Economía Social y Discapacidad.

Abstract

The specific needs of older people require different lines of consumption, such as the consumption that serves to preserve or restore the health, goods and services that continue to allow the home, or goods and services to fill the free time.
The rapid increase of older people is occurring in Europe and particularly in Spain, forced to pay special attention to the study of the expenditure incurred by the elderly, with emphasis on private consumption of services, as a matter not sufficiently analyzed in economic literature.
In the above context it is carrying out an analysis of consumption patterns of households in the elderly, by exploiting the information contained in the Family Budget Survey 1998 and 2005, using multivariate statistical techniques.

Thematic Area: Social Economics.

1. INTRODUCCIÓN

Durante las últimas décadas el aumento de la esperanza de vida, unido al descenso de la natalidad, ha producido un aumento de las personas mayores. Este envejecimiento de la población está teniendo sus previsibles consecuencias, como es el incremento de una parte cada vez mayor del gasto, tanto público como privado, encaminado a satisfacer la demanda de las personas mayores. De ahí que el conocimiento de cómo está evolucionando el colectivo anterior tenga una importancia económica y social.

El aumento del colectivo anterior supone un crecimiento en la demanda de bienes y servicios que puedan atender las necesidades específicas de las personas mayores, se trata de líneas de consumo orientadas a conservar la salud, a atender el propio hogar y a disfrutar del tiempo libre.

Pero también el incremento de las personas mayores obliga a una reorientación de las políticas sociales. Las políticas públicas, tanto en la Unión Europea como en España, han de ir dirigiéndose cada vez más hacia los mayores, sus hogares o familiares, por lo que una buena parte de los recursos públicos tendrán que ir destinados a cubrir gran parte de sus necesidades.

Estos cambios que van produciéndose en la estructura demográfica se reflejan en la información estadística de los países europeos. Así, según datos de Eurostat, recogidos en la tabla 1, en una década, el intervalo temporal comprendido entre 1998 y 2007, la Unión Europea de los quince ha experimentado una subida del 2 por ciento en la proporción de personas de 65 años y más con respecto a la población total, siendo Italia y Alemania los dos países más envejecidos en 2007, con casi un 20 por ciento de su población con 65 años y más. España, no es ninguna excepción, y aunque se ha ralentizado el dinamismo de crecimiento de las personas mayores en los últimos años[1], en 2007 sigue presentando una proporción elevada. El resto de países presentan todos porcentajes superiores al 10 por ciento, a excepción de Turquía.

Tabla 1. Proporción de población en la Unión Europea con 65 años y más (1998-2007)
(porcentaje sobre la población total)

	1998	2005	2007
UE-27	**15.3**	**16.6**	**16.9**
UE-25	15.4	16.7	17.0
UE-15	**15.9**	**17.4**	**17.9**
Bélgica	16.5	17.2	17.1
Bulgaria	15.6	17.1	17.3
República Checa	13.6	14.0	14.4
Dinamarca	14.9	15.0	15.3
Alemania	15.8	18.6	19.8
Estonia	14.5	16.5	17.1
Irlanda	11.4	11.2	10.9
Grecia	15.9	18.1	18.6
España	**16.2**	**16.8**	**16.7**
Francia	15.5	16.2	16.2
Italia	17.5	19.5	19.9
Chipre	11.1	11.9	12.3
Letonia	14.4	16.5	17.1
Lituania	13.2	15.1	15.6
Luxemburgo	14.3	14.1	14.0
Hungría	14.7	15.6	15.9
Malta	11.8	13.3	13.8
Holanda	13.5	14.0	14.5
Austria	15.4	16.0	17.2
Polonia	11.7	13.1	13.4
Portugal	15.6	17.0	17.3
Rumania	12.9	14.7	14.9
Eslovenia	13.2	15.3	15.9
Eslovaquia	11.2	11.6	11.9
Finlandia	14.6	15.9	16.5
Suecia	17.4	17.2	17.4

[1] A principios del siglo XXI la entrada de inmigrantes en España ha contribuido en parte a un cierto rejuvenecimiento de la población, si bien, en los tiempos actuales, se está produciendo la salida del país de importantes contingentes de ellos, derivada de los efectos de la crisis económica.

Tabla 1 (cont.)

	1998	2005	2007
Reino Unido	15.9	16.0	16.0
Croacia	12.3	16.7	17.0
Macedonia	9.2	10.9	11.2
Turquía	5.2	5.8	6.6
Islandia	11.6	11.8	11.6
Liechtenstein	10.2	11.1	11.9
Noruega	15.7	14.7	14.6
Suiza	15.0	15.8	16.2

Fuente: Eurostat. Elaboración propia

En cuanto a las características de los hogares europeos teniendo en cuenta a la persona de referencia o sustentador principal con sesenta años y más, la tabla 2 presenta la proporción de dichos hogares con respecto al total en el horizonte temporal 1999-2005, para aquellos países en los que se disponían de datos en Eurostat. Como se puede observar, la Unión Europea de los quince arroja una evolución positiva, llegando a alcanzar en 2005 algo más del 35 por ciento los hogares cuya persona de referencia es un mayor. Los países que presentan los porcentajes más altos de este tipo de hogares son Italia y Bulgaria. Por su parte, España también refleja una de las cifras más altas de la Unión Europea, llegando casi al cuarenta por ciento. Se puede comprobar también que las sucesivas ampliaciones de países en el bloque europeo han disminuido el porcentaje medio de estos hogares.

Tabla 2. Proporción de hogares de la Unión Europea cuyo sustentador principal tiene 60 años y más (1999-2005)

	1999	2005
Austria	30,61	31,15
Bélgica	30,97	28,71
Bulgaria	45,50	42,33
Chipre	31,50	30,43
República Checa	29,80	31,87
Alemania	33,40	35,04
Dinamarca	27,74	29,71
España	**36,98**	**39,83**
UE-15	**33,56**	**35,28**
UE-25		34,48
UE-27		33,45
Finlandia	28,93	31,81
Francia	30,79	33,17
Grecia	40,67	38,36
Hungría	34,00	34,82
Islandia	27,17	29,89
Italia	44,14	42,45
Lituania	31,70	32,17
Luxemburgo	31,07	27,27
Letonia	33,30	30,69
Macedonia	31,60	26,57
Holanda	26,45	26,63
Noruega		22,90
Polonia	25,00	28,97
Portugal	38,51	33,75
Rumania	37,00	8,35
Estonia	22,26	29,48
Eslovenia	27,40	29,65
Eslovaquia	16,00	26,57
Turquía		16,48
Reino Unido	29,81	33,46

Fuente: Eurostat. Elaboración propia.

A pesar de que el impacto del envejecimiento demográfico sobre la composición del consumo privado es un tema de amplio debate y análisis, hasta estos momentos no existe un pleno consenso. Es por lo que, teniendo en cuenta el contexto anterior, se hace preciso realizar un estudio del gasto llevado a cabo por los mayores en nuestro país, haciendo especial referencia al consumo privado de servicios. Para ello, se utilizará la Encuesta Continua de Presupuestos Familiares, considerando el periodo temporal 1998-2005.

En concreto, se trata de realizar un estudio de las características, patrones y factores determinantes del consumo privado de los mayores en España, con especial referencia al gasto en servicios.

La hipótesis de trabajo consiste en que la explicación del gasto en consumo de servicios por las personas mayores no se debe solamente al nivel de renta sino que además se dan otros tipos de factores.

En el sentido anterior, los objetivos que se pretenden alcanzar con el análisis son dos, primero, estudiar el efecto de las características de los hogares de la tercera edad sobre el consumo de servicios en España entre 1998 y 2005; y segundo, detectar y confirmar los factores explicativos del consumo de servicios en dichos hogares.

Los resultados del trabajo se podrían tomar como referencia para profundizar en estudios de carácter regional, o bien atendiendo a la diferenciación entre el mundo rural y el urbano.

De acuerdo con lo anteriormente expuesto, estructuramos el estudio en los apartados siguientes. En el primer epígrafe se realiza una revisión de la literatura sobre los patrones de consumo de las personas mayores, en el segundo apartado se recoge la metodología y fuentes estadísticas empleadas, en el tercer bloque se muestran los resultados empíricos y, en un último apartado, se presentan las conclusiones.

2. REVISIÓN DE LA LITERATURA

A la hora de abordar el estudio del comportamiento de los consumidores en cualquier edad, y de los consumidores mayores en concreto, el referente inicial viene marcado por las hipótesis de la Renta Permanente y del Ciclo Vital.

La hipótesis de la Renta Permanente (Milton Friedman, 1957), se basa en la idea de que el consumo y el ahorro no son función del ingreso corriente, sino de dos tipos de ingresos: el ingreso permanente, entendido como el ingreso futuro esperado; y, el ingreso transitorio o no esperado, de naturaleza aleatoria. Además, señala que parte de la renta permanente que se dedica al consumo está influenciada por factores como la edad, el tamaño y composición de la familia, etc., que son factores que difieren de unas familias a otras.

La hipótesis del Ciclo Vital (Modigliani y Brumberg, 1954; Ando y Modigliani 1957, 1963; Brumberg, 1966; Modigliani 1985) argumenta que el individuo típico tiene una corriente de ingresos que se distribuye de forma distinta según el período de vida en el que se encuentre, es decir, es relativamente baja al principio y al final de su vida, y alcanza sus máximos en los años de madurez. Por otra parte, el nivel de consumo será más o menos constante a lo largo de su vida.

Los trabajos de Ando y Modigliani basándose en datos de presupuestos familiares indican que la renta de la unidad familiar primero aumenta y después decrece al ir aumentando la edad de ésta. El tamaño de la familia, dado por el número de miembros de la unidad familiar, sigue el mismo esquema general de comportamiento, aunque más agudizado que el de la función de renta.

Los autores de la hipótesis del Ciclo Vital señalan que a medida que aumente el tamaño familiar los gastos en consumo necesarios para mantener un cierto nivel de vida lo harán en menor proporción (Ando y Modigliani, 1957), llegando a la conclusión de que las unidades de consumo pueden basar sus expectativas en la experiencia y el comportamiento reales de las unidades familiares que se hallan en un punto más avanzado del ciclo vital. De esta forma, el modelo variará con la edad.

A medida que va transcurriendo el ciclo vital familiar se van franqueando una serie de etapas en las que se modifican la renta de la familia y la composición del hogar, lo que lleva a originar distintas necesidades. De este modo, parece evidente pensar que los diferentes tipos de familia u hogares originan comportamientos de consumo diferenciados.

Sin embargo, la literatura revisada nos indica que los estudios sobre el comportamiento de los consumidores mayores no empezaron a enfocarse con profundidad desde la óptica microeconómica hasta un poco antes de 1970, anteriormente la investigación se centraba más en su carácter macroeconómico.

Los estudios revisados muestran que no ha habido un alto consenso a la hora de delimitar el concepto de consumidores mayores. Así, en el plano internacional, se han empleado edades de 60, 62 y 65 años (Axelson and Penfield, 1983; Moehrle, 1990; Schwenk, 1995); y también se ha llegado a establecer entre dos y cuatro agrupaciones de edades (Harrison, 1986). Incluso se ha considerado un rango de edades comprendido entre 45 y 65 años (Gantz et al., 1980; Greco, 1989; Peterson, 1992; Ursic et al., 1986). Otros autores (Andel-Ghany y Sharpe, 1997) consideran dos grupos de consumidores mayores, aquellas con edades entre 6 y 74 años (consumidor joven-viejo) y los que tienen 75 años y más (consumidor viejo-viejo). En esa misma línea se encuadran otros autores (Hitschler, 1993; Paulin, 2000), que comparaban familias cuyo sustentador principal tenía entre 65 y 74 años con aquellas otras unidades de consumo en las que la persona de referencia tenía 75 y más años. Por su parte, hubo quien examino los patrones de gasto de familias cuya persona de referencia tenía entre 62 y 74 años (Moehrle, 1990).

Otros trabajos revisados consideran la edad de 50 años como un indicador para un consumidor mayor (Beck, 1996; Buró y Gabbot, 1995; Cardigan et al., 2004; Kang y Ridgway, 1996; Mitchell, 1996; Semon, 1995; Weaver, 1996).

En España, distintos autores consideran personas mayores aquellas que tienen 65 y más años (Mañas et al., 2002; Grande, 2002; Marchante et al., 2002; Hernández (2005); Camacho y Hernández (2006).

Así pues, de acuerdo con los estudios españoles y para los fines de este trabajo, se ha considerado persona mayor a partir de los 65 años de edad.

3. METODOLOGÍA Y FUENTES ESTADÍSTICAS

Los sujetos de la investigación se corresponden con hogares españoles cuyo sustentador principal tiene 65 y más años. Por tanto, quedan excluidos del análisis aquellos hogares en los que convive alguna persona de 65 años o más pero en los que no son ellos mismos los sustentadores principales.

Las fuentes estadísticas empleadas son las Encuestas Continuas de Presupuestos Familiares, información trimestral, elaboradas por el Instituto Nacional de Estadística, que permiten disponer de información fiable y actualizada para conocer el comportamiento de consumo de los hogares y la asignación de su presupuesto a los distintos conceptos de gasto. El ámbito temporal analizado abarca desde el cuarto trimestre de 1998 hasta el cuarto trimestre de 2005.

La elevada cantidad de variables a manejar en las Encuestas nos ha llevado a tener que emplear técnicas estadísticas que afrontaran adecuadamente los objetivos de la investigación, así pues intentando tener una visión conjunta e integrada de los aspectos implicados se han utilizado técnicas multivariantes, caracterizadas por resumir y sintetizar grandes conjuntos de datos con el fin de mejor el conocimiento que se tiene de la realidad.

4. RESULTADOS EMPÍRICOS

4.1. ANÁLISIS DESCRIPTIVO

Con el propósito de analizar la evolución del comportamiento de los consumidores mayores se ha extraído una muestra compuesta por aquellos hogares en los que la edad del sustentador

principal es de 65 y más años. Las muestras seleccionadas se corresponden con el cuarto trimestre de los años 1998 y 2005, siendo el número de hogares extraídos de 2.526 y 2.384, respectivamente.[2]

El conjunto de hogares pertenecientes a la tercera edad puede diferenciarse del resto tanto en el origen de los ingresos[3] que percibe como en su comportamiento respecto al consumo.

Tabla 3. Consumo medio en los hogares en los que el sustentador principal tiene 65 o mas años, 1998-2005

Grupos de gastos en bienes y servicios	Importe medio trimestral por hogar (€)		% sobre total gastos		% sobre total servicios	
	4T98	4T05	4T98	4T05	4T98	4T05
Alimentos y bebidas no alcohólicas	686,58	863,80	21,56	19,47		
Bebidas alcohólic., tabaco y narcóticos	67,32	72,94	2,11	1,64		
Artículos de vestir y calzado	201,16	239,21	6,32	5,39		
Vivienda, agua, elec., gas y otr. comb.	1.216,58	1.864,59	38,20	42,02		
Mobiliario, equip. hogar y gastos corr.	150,23	201,46	4,72	4,54		
Salud	92,37	97,51	2,90	2,20		
Transportes	208,63	245,39	6,55	5,53		
Comunicaciones	64,92	104,22	2,04	2,35		
Ocio, espectáculos y cultura	153,48	226,64	4,82	5,11		
Enseñanza	13,09	10,36	0,41	0,23		
Hoteles, cafés y restaurantes	186,73	255,08	5,86	5,75		
Otros bienes y servicios	143,40	256,20	4,50	5,77		
Suma bienes y servicios	3.184,49	4.437,41	100,00	100,00		
Grupos de gastos en servicios						
Servicios y reparaciones para el hogar	136,40	173,33	4,28	3,91	15,63	13,99
Servicios de la vivienda	204,72	260,50	6,43	5,87	23,46	21,03
Servicio doméstico	27,68	58,11	0,87	1,31	3,17	4,69
Servicios médicos	50,86	51,91	1,60	1,17	5,83	4,19
Transporte personal	26,03	37,58	0,82	0,85	2,98	3,03
Transporte público	26,75	29,08	0,84	0,66	3,07	2,35
Correos y comunicaciones	5,00	32,78	0,16	0,74	0,57	2,65
Servicios de ocio y cultura	88,20	136,40	2,77	3,07	10,11	11,01
Servicios de guardería	0,47	0,69	0,01	0,02	0,05	0,06
Servicios de enseñanza	12,82	10,05	0,40	0,23	1,47	0,81
Servicios para el cuidado personal	28,94	62,30	0,91	1,40	3,32	5,03
Restaurantes	65,89	104,04	2,07	2,34	7,55	8,40
Cafés y bares	111,39	137,63	3,50	3,10	12,77	11,11
Hoteles	9,45	13,41	0,30	0,30	1,08	1,08
Otros servicios y seguros	77,93	130,80	2,45	2,95	8,93	10,56
Suma Servicios	872,54	1.238,61	27,40	27,91	100,00	100,00

Fuente: ECPF, 1998-2005. INE. Elaboración propia

En general, los cambios en la distribución del gasto en el intervalo temporal analizado muestran que se ha producido un desplazamiento desde los gastos en alimentación, bebidas, vestido, mobiliario, salud, transportes, enseñanza, y hoteles, hacia vivienda, comunicaciones, ocio, y otros bienes y servicios.

El gasto en servicios correspondiente a aquellos hogares en los que el sustentador principal tiene 65 y más años muestra una evolución creciente tanto en términos absolutos como relativos. Dicha evolución guarda una estrecha relación con el aumento de pensiones y prestaciones destinadas a las personas mayores, propio de los países altamente desarrollados que siguen la senda del estado del bienestar. Por lo que resulta previsible que continúe dicho aumento en los próximos años.

[2] Se ha optado por el último trimestre para contar con la última información disponible de las Encuestas Continuas de Presupuestos Familiares, base 1997, en su carácter trimestral. No se ha optado por los ficheros longitudinales anuales porque no se disponían de los microdatos de 2005.

[3] La renta de la tercera edad tiene su procedencia más de los sistemas públicos de transferencias y del propio ahorro acumulado que del propio mercado laboral.

Haciendo hincapié en los grupos de servicios, se produce una desviación del gasto hacia servicio doméstico, transporte personal, comunicaciones, ocio y cultura, cuidado personal, restaurantes y otros servicios; en detrimento de servicios y reparaciones para el hogar, servicios de la vivienda, médicos, transporte público, enseñanza, y cafés y bares.

En el año 2005, servicios y reparaciones para el hogar, servicios de la vivienda, ocio y cultura, cafés y bares, y otros servicios y seguros, suponen cerca de un setenta por ciento del total del gasto en servicios.

En un intento de avanzar un paso más en el estudio, resulta interesante llevar a cabo un análisis del consumo diferencial en servicios entre aquellos hogares cuyo sustentador principal tiene más de 64 años y aquellos en los que el sustentador principal tiene menos de 65 años. Para ello, se selecciona una muestra formada por aquellos hogares en los que el sustentador principal tiene menos de 65 años.

En este caso también se observa que se ha producido una desviación del gasto en bienes y servicios a lo largo del período considerado, tomando una mayor importancia relativa los gastos en vivienda, transportes, comunicaciones, ocio, y otros bienes y servicios. Al igual que sucedía para el caso de los hogares con sustentador principal de 65 años y más, hay una evolución creciente en el consumo de servicios, pero además en el caso que nos ocupa este avance se intensifica más. Así, para el año 2005, el consumo de servicios se sitúa en una cifra superior al 33,5 por ciento con respecto al gasto total. De este modo, se puede hablar de una cierta contracción del gasto en servicios para el caso de los hogares cuyo sustentador principal es un mayor de 65 y más años, en el que los servicios representan el 27,91 por ciento del presupuesto total del hogar.

Tabla 4. Consumo medio en los hogares en los que el sustentador principal tiene menos de 65 años, 1998-2005

Grupos de gastos en bienes y servicios	Importe medio trimestral por hogar (€)		% sobre total gastos		% sobre total servicios	
	4T98	4T05	4T98	4T05	4T98	4T05
Alimentos y bebidas no alcohólicas	906,52	1.144,89	17,87	15,81		
Bebidas alcohólic., tabaco y narcóticos	139,88	164,90	2,76	2,28		
Artículos de vestir y calzado	399,06	497,01	7,87	6,86		
Vivienda, agua, elec., gas y otr. comb.	1.418,03	2.177,55	27,96	30,06		
Mobiliario, equip. hogar y gastos corr.	227,29	307,57	4,48	4,25		
Salud	112,42	167,65	2,22	2,31		
Transportes	582,81	844,36	11,49	11,66		
Comunicaciones	93,04	205,98	1,83	2,84		
Ocio, espectáculos y cultura	364,37	544,59	7,18	7,52		
Enseñanza	110,60	113,01	2,18	1,56		
Hoteles, cafés y restaurantes	475,58	680,51	9,38	9,40		
Otros bienes y servicios	242,27	394,83	4,78	5,45		
Suma bienes y servicios	5.071,88	7.242,85	100,00	100,00		
Grupos de gastos en servicios						
Servicios y reparaciones para el hogar	168,57	213,82	3,32	2,95	10,67	8,80
Servicios de la vivienda	249,62	390,46	4,92	5,39	15,79	16,07
Servicio doméstico	23,74	49,04	0,47	0,68	1,50	2,02
Servicios médicos	46,42	77,97	0,92	1,08	2,94	3,21
Transporte personal	76,73	107,55	1,51	1,48	4,85	4,43
Transporte público	47,97	61,66	0,95	0,85	3,03	2,54
Correos y comunicaciones	14,13	121,87	0,28	1,68	0,89	5,02
Servicios de ocio y cultura	183,20	301,30	3,61	4,16	11,59	12,40
Servicios de guardería	19,84	20,37	0,39	0,28	1,26	0,84
Servicios de enseñanza	101,52	105,12	2,00	1,45	6,42	4,33
Servicios para el cuidado personal	33,04	68,15	0,65	0,94	2,09	2,81
Restaurantes	186,15	265,23	3,67	3,66	11,78	10,92
Cafés y bares	256,22	358,16	5,05	4,94	16,21	14,74
Hoteles	33,22	57,12	0,65	0,79	2,10	2,35
Otros servicios y seguros	140,14	231,42	2,76	3,20	8,87	9,53
Suma Servicios	1.580,49	2.429,22	31,16	33,54	100,00	100,00

Fuente: ECPF, 1998-2005. INE. Elaboración propia

4.2. Análisis Factorial de componentes principales

Para detectar los posibles factores explicativos del consumo de servicios en los hogares de la tercera edad, partimos de una serie de variables,[4] como son: tamaño del municipio, capital de provincia, densidad, número de miembros del hogar, tipo de hogar, número de adultos en el hogar y número de mujeres en el hogar.

Las muestras seleccionadas se corresponden con el cuarto periodo trimestral de 1998 y 2005. Y se trata de hogares de la tercera edad que presentan gasto no nulo en consumo de servicios.

Los cuadros 1 y 2 muestran los coeficientes de las matrices de componentes rotados por cada factor detectado para cada trimestre considerado: *composición del hogar* y *aspectos geográficos y demográficos*.[5]

CUADRO 1
COEFICIENTES DE LA MATRIZ DE COMPONENTES ROTADOS
FACTOR *COMPOSICIÓN DEL HOGAR*
(4T-1998 A 4T-2005)

	Nº miembros del hogar	Tipo de hogar	Nº adultos	Nº mujeres
4T-1998	0,953	0,803	0,931	0,821
4T-2005	0,956	0,820	0,940	0,817

Fuente: Elaboración propia

CUADRO 2
COEFICIENTES DE LA MATRIZ DE COMPONENTES ROTADOS
FACTOR *ASPECTOS GEOGRÁFICOS Y DEMOGRÁFICOS*
(4T-1998 A 4T-2005)

	Tamaño del municipio	Capital de provincia	Densidad
4T-1998	0,971	0,911	0,882
4T-2005	0,966	0,904	0,856

Fuente: Elaboración propia

Como se puede comprobar las variables que forman el factor *composición del hogar* vienen dadas por: número de miembros del hogar, tipo de hogar, número de adultos y número de mujeres; mientras que las variables que configuran el factor *aspectos geográficos y demográficos* son: tamaño del municipio, capital de provincia y densidad.

El cuadro 3 recoge el orden y varianza explicada por cada factor así como el porcentaje total de varianza explicada en cada uno de los modelos factoriales obtenidos para cada trimestre considerado.

CUADRO 3
ORDEN Y VARIANZA EXPLICADA POR CADA FACTOR (SOLUCIÓN ROTADA)
(4T-1998 A 4T-2005)

	Factor 1 (Composición del hogar)	**Factor 2 (Aspectos geográficos y demográficos)**	**Porcentaje total de Var. explicada**
4T-1998	44,19%	36,50%	80,69%
4T-2005	44,83%	35,49%	80,33%

Fuente: Elaboración propia

[4] Teniendo en cuenta el marco teórico y los trabajos revisados, dichas variables definen características de los hogares españoles pertenecientes a la tercera edad y se intuye que en algunos casos puedan estar fuertemente correlacionadas.

[5] La denominación de los factores se ha realizado en base a los resultados obtenidos en las matrices de componentes rotados, y de acuerdo con la teoría y la literatura revisada.

4.3. Análisis de Regresión Lineal Múltiple

Se va a trabajar con las puntuaciones factoriales en los factores extraídos a través del ACP, siempre teniendo en cuenta que se trata de puntuaciones tipificadas. Se trata de realizar un ARLM de los importes de gasto en servicios con las componentes extraídas y adicionando la variable intervalo de ingresos, como variable que recoge la presencia de la renta.

Variable dependiente: Logaritmo neperiano de Suma de servicios del 4º trimestre.

Variables independientes: Intervalo de ingresos, Factor Composición del hogar, Factor Aspectos geográficos y demográficos.

El cuadro 4 refleja un resumen de los resultados de la aplicación del ARLM[6] a cada uno de los trimestres estimados.

CUADRO 4
RESUMEN RESULTADOS APLICACIÓN ARLM
(4T-1998 A 4T-2005)

				Coef. Estandarizados (nivel significación)		
Servicios	N	R^2	Signif. Anova	Intervalo de ingresos	Fact. Comp. hogar	Fact. Asp geog.. y demo.
4T-1998	2.476	18.8%	0.000	0.319 (0.000)	0.134 (0.000)	-0.129 (0.000)
4T-2005	2.383	15.4%	0.000	0.321 (0.000)	0.048 (0.023)	-0.142 (0.000)

Fuente: elaboración propia

5. CONCLUSIONES

En este trabajo hemos analizado el consumo en el segmento de población de más de 65 años, pudiendo decir que si bien este grupo de población tiene necesidades comunes al resto de los grupos de edad, también precisa otra serie de bienes y servicios que los diferencian del resto de consumidores. De hecho, el envejecimiento de la población supone una alteración de los patrones de consumo ya que el crecimiento de la población de mayor edad produce una mayor demanda de productos sanitarios o de bienes públicos y, por el contrario, se reducirá la demanda de otros bienes y servicios que son más propios de otros grupos de edad.

Desde el punto de vista de la oferta, el incremento progresivo de la tercera edad, con una tendencia de la población hacia el envejecimiento, crea la necesidad de un mercado destinado a los consumidores mayores, que implica una modificación de las estrategias empresariales.

Con el transcurso de los años, de acuerdo con el periodo analizado y según hemos podido comprobar, adquiere una mayor importancia el gasto llevado a cabo por los hogares de la tercera edad, pero también muestra un mayor auge el gasto en servicios con respecto al gasto total. En este contexto, se hace preciso analizar cuáles son las pautas y patrones de consumo que siguen los hogares de la tercera edad, así como determinar cuáles son sus factores explicativos en cuanto al volumen de gasto en servicios.

El marco teórico de nuestro estudio nos lleva a establecer que:

[6] La utilización en esta investigación de datos microeconómicos y de sección cruzada (cada hogar es un caso que contempla distintos atributos en una fecha determinada) implica, siguiendo a Schroeder, Sjoquist y Stephan (1986), que la valoración de la bondad de los modelos ajustados, tomando como referencia al R^2, no debe ser la misma que cuando se emplean datos macroeconómicos y datos longitudinales, ya que para éstos últimos se esperan valores de R^2 más elevados que para los primeros, en los que, con frecuencia, son muy bajos debido a que la aleatoriedad que presentan suele ser muy superior. En esta misma línea, los coeficientes de determinación correspondientes a los modelos estimados en esta investigación no deben de tenerse en cuenta como referentes principales de la calidad de los resultados sino como una información más, prestando una mayor atención a los test que contrastan la significatividad de los coeficientes correspondientes a las variables.

- El consumo de los hogares de la tercera edad está influenciado por variables como la edad, el tamaño de la familia, la composición del hogar, entre otras, tal y como expuso Friedman en su Teoría de la Renta Permanente.
- De acuerdo con la Teoría del Ciclo Vital, el modelo de consumo de la familia varía con la edad, la renta de la familia y la composición del hogar, ya que las necesidades son distintas, por lo que el ciclo vital es un factor discriminante del consumo.

El análisis de las pautas y patrones de consumo de los hogares de la tercera edad, de acuerdo con el análisis descriptivo realizado, nos lleva a plantear, en el intervalo temporal 1998-2005, como puntos más destacados que:

- El consumo de los hogares de la tercera edad ha crecido tanto en términos absolutos como relativos, incremento que está relacionado con el aumento de las pensiones y prestaciones dirigidas a las personas mayores. Tendencia que es muy probable que continúe en los próximos años.

- No obstante, cuando se compara el consumo de estos hogares con el del resto de la población, se muestra una contracción del gasto. Y ello es así por la propia disminución de ingresos que se produce en los hogares de la tercera edad con la llegada de la jubilación.

La aplicación del Análisis Factorial de Componentes Principales y del Análisis de Regresión Lineal Múltiple nos ha llevado a determinar, para el periodo 1998-2005, los factores explicativos del consumo de los hogares de la tercera edad en cuanto al volumen de gasto en servicios. Y, además de la *renta*, como variable explicativa de obligada presencia en los estudios sobre consumo, se han determinado otros dos factores como son: la *composición del hogar* y los *aspectos geográficos y demográficos*:

- La *renta*, identificada por la variable *intervalo de ingresos*, manifiesta una relación positiva con el gasto en servicios, es decir, a mayor nivel de renta del hogar mayor es el gasto en servicios.

- La *composición del hogar*, factor formado por las variables: número de miembros del hogar, tipo de hogar, número de adultos y número de mujeres, muestra también una relación positiva con el volumen de gasto en servicios. Cuánto más amplia es la composición del hogar mayor es el volumen de gasto en servicios.

- Los *aspectos geográficos y demográficos*, factor formado por: tamaño de municipio, capital de provincia y densidad, evidencia una relación negativa con la intensidad de gasto en servicios. Es decir, conforme aumenta la zona de menor afluencia de habitantes menor es el gasto en servicios.

En definitiva, los resultados indican que la *renta* es la variable principal para explicar las diferencias de gasto en servicios de los hogares de la tercera edad, y le sigue en orden de importancia, la *composición del hogar* y los *aspectos geográficos y demográficos*. Resultados que coinciden en parte con los obtenidos por Bigné et al. (2002), los cuales demostraron, utilizando los microdatos de la Encuesta Continua de Presupuestos Familiares de 1998 y para los gastos en bienes y servicios, que es el nivel de renta la variable más relevante para explicar las diferencias en dichos gastos de los hogares de la tercera edad, seguidos en importancia por el nivel de estudios y el hábitat.

Por otra parte, desde el punto de vista de las políticas sociales, sería interesante instrumentar mecanismos que faciliten la satisfacción de las necesidades de los mayores, estimulando la oferta de bienes y servicios dirigidos a este colectivo.

Por último, la determinación en nuestro estudio de un factor espacial como los aspectos *geográficos y demográficos* nos sugiere seguir trabajando en un futuro en el análisis del consumo de los hogares de la tercera edad desde una perspectiva regional, y poder plantearnos también el análisis del consumo de servicios en los entornos rural y urbano.

BIBLIOGRAFÍA

ABDEL-GHANY, M. y SHARPE, D. (1997): "Consumption Patterns Among the Young-Old and Old-Old". *Journal of Consumer Affairs*, vol. 31, n.º 1, págs. 90-112.

ARNALTE, E. (1998): "La problemática rural en el sur de Europa: líneas de análisis y temas de debate". Paris, Centre d'Études et de Recherches Internationales de la Fondation Nationale des Sciences Politiques.

ANDO, A. y MODIGLIANI, F. (1957): "Tests of the Life-Cycle Hypothesis of savings: comments and suggestions". *Bulletin of the Oxford University Institute of Statistics*, vol. 19.

ANDO, A. y MODIGLIANI, F. (1963): "The Life-Cycle Hypothesis of saving: aggregate implications and tests". *American Economic Review*, vol. 53, págs. 55-84.

AXELSON, M. y MARJORIE, P. (1983): "Factors Associated with Food Expenditure of Elderly Persons Living Alone". *Home Economics Research Journal*, vol. 12, n.º 2, págs. 228-236.

BECK, B. (1996): "The luxury of longer life". *Economist*, 27, Enero, págs.. 7-9.

BELLANTE, D. y FOSTER (1984): "Working Wives and expenditures on Services". *Journal of Consumer Research*, 11, Septiembre, págs. 700-707.

BIGNÉ, J., ALDÁS, J., KÜSTER, I. y VILA, N. (2002): "Patrones y determinantes del consumo de los hogares de tercera edad en España". *Revista Española de Investigación en Marketing ESIC*, págs. 107-128.

BURT, S. y GABBOT, M. (1995): "The elderly consumer and non-food purchase behaviour". *European Journal of Marketing*, Vol. 29, N.º 2, págs. 15-43.

CAMACHO, J. A. y HERNÁNDEZ, M. (2006): "Consumo de servicios y ciclo vital: un análisis de los hogares españoles con datos de corte transversal". *Servilab, Serie Documentos de Trabajo,* 04/2006.

CAPEL, H. (1975): "La definición de lo Urbano". *Estudios Geográficos*, 138-139, págs. 265-301.

CARRIGAN, M. (1998): "Segmenting the Grey Market: The Case of Fitty-Plus 'Lifegroups'". *Journal of Marketing Practice*, vol. 4, n.º 2, págs. 43-56.

CARRIGAN, M. (1999): "Old spice-Developing successful relationships with the grey market". Long Range Planning, vol. 32, n.º 2, págs. 253-262.

CARRIGAN, M., SZMIGIN, I. y WRIGHT, J. (2004): "Shopping for a better world? An interpretive study of the potential for ethical consumption within the older market". *Journal of Consumer Marketing*, vol. 21, n.º 6, págs. 401-417.

COLE, C. y BALASUBRAMANIAN, S. (1993): "Age Differences in Consumers' Search for Information: Public Policy Implications". *Journal of Consumer Research*, vol. 20, Junio, págs. 157-169.

DEATON, A. S., RUIZ-CASTILLO, J. y DUNDAN, T. (1989): "The Influence of Household Composition on Household Expenditure Patterns: Theory and Spanish Evidence". *Journal of Political Economy*, vol. 97, n.º 1, págs. 179-200.

DOUTHITT, R. A. y FEDYK, J. M. (1988): "The influence of children on family life cycle spending behavior: theory and applications". *Journal of Consumer Affairs*, n.º 22, págs. 220-248.

DOUTHITT, R. A. y FEDYK, J. M. (1990): "Family composition, parental time and market goods: Life cycle trade-offs". *Journal of Consumer Affairs*, n.º 24, págs. 110-133.

FRIEDMAN, M. (1957): "A Theory of the Consumption Function". Princeton University Press, Princeton, N. J. Traducción: "Una Teoría de la Función de Consumo". Alianza Editorial. Madrid, 1973.

GANTZ, W., GARTENBERG, H. y RAINBOW, C. (1980): "Approaching invisibility: the portrayal of the elderly in magazine advertisements". *Journal of Comunication*, págs. 56-60.

GARCÍA BARTOLOMÉ, J. (1996): *Los procesos rurales en el ámbito de la Unión Europea*. En Teresa, A. y Cortés, C. (coords.): *La nueva relación campo-ciudad y la pobreza rural*. México, UNAM, INAH, UAM y Plaza y Valdés Editores.

GARCÍA SANZ, B. (1996). "La sociedad rural ante el siglo XXI". Madrid, MAPA, Serie Estudios, 125.

GARCÍA SANZ, B. (2000): *La diversificación económica de la sociedad rural*. Madrid, mimeo.

GILLY, M. C. y ENIS, B. M. (1982): "Recicling the family life cycle: a proposal for redefinition". *Advances in Consumer Research*, vol. 9, págs. 271-276.

GRACIA, A., GIL, J. y ANGULO, A. (1998). "El consumo de alimentos en España: el consumidor rural versus urbano". *Revista de Estudios Regionales*, vol. 50, págs. 111-130.

GRANDE, I. (2002): *El Consumo de la Tercera Edad*. Madrid, ESIC Editorial.

GRECO, A. (1989): "Representation of the elderly in advertising: crisis or inconsequence?". *Journal of Consumer Marketing*, vol. 19, n.º 1, págs. 37-44.

HAIR, J. F.; ANDERSON, R. E.; TATHAM, R. L. y BLACK, W. C. (1999): *Análisis Multivariante*. Madrid, Prentice Hall.

HARRISON, B. (1986): "Spending Patterns of Older Persons Reveled in Expenditure Survey". *Monthly Labor Review*, vol. 109, n.º 10, págs. 15-17.

HERNÁNDEZ, M. (2005): *Servicios y consumo familiar: un análisis regional de los patrones de consumo*. Tesis doctoral. Universidad de Granada.

HITSCHLER, P. (1993): "Spending by older consumers: 1980 and 1990 compared". *Monthly Labor Review*, Mayo, págs. 3-13.

INE (1991): "Censo de Población y Viviendas". Instituto Nacional de Estadística, Madrid.

INE (1998a): "Encuesta Continua de Presupuestos Familiares 1998. Base 1997. Ficheros de microdatos trimestrales. Instituto Nacional de Estadística, Madrid.

INE (1998b): "Encuesta Continua de Presupuestos Familiares. Renovada en el 2º trimestre de 1997. Metodología". Instituto Nacional de Estadística, Madrid.

INE (1999): "Encuesta Continua de Presupuestos Familiares 1998. Base 1997. Ficheros de microdatos trimestrales. Instituto Nacional de Estadística, Madrid.

INE (2000): "Encuesta Continua de Presupuestos Familiares 1998. Base 1997. Ficheros de microdatos trimestrales. Instituto Nacional de Estadística, Madrid.

INE (2001a): "Encuesta Continua de Presupuestos Familiares 1998. Base 1997. Ficheros de microdatos trimestrales. Instituto Nacional de Estadística, Madrid.

INE (2001b): "Censo de Población y Viviendas". Instituto Nacional de Estadística, Madrid.

INE (2002): "Encuesta Continua de Presupuestos Familiares 1998. Base 1997. Ficheros de microdatos trimestrales. Instituto Nacional de Estadística, Madrid.

INE (2003): "Encuesta Continua de Presupuestos Familiares 1998. Base 1997. Ficheros de microdatos trimestrales. Instituto Nacional de Estadística, Madrid.

INE (2004): "Encuesta Continua de Presupuestos Familiares 1998. Base 1997. Ficheros de microdatos trimestrales. Instituto Nacional de Estadística, Madrid.

INE (2006): "Boletín informativo del Instituto Nacional de Estadística. Cifras INE". N.º 4/2006. Octubre. INEbase.

INSTITUTO NACIONAL DE CONSUMO (2000): *La Tercera Edad y el Consumo*. Madrid, Ministerio de Sanidad y Consumo.

KANG, Y. y RIDGWAY, N. (1996): "The importance of consumer market interactions as a form of social support for elderly consumers". *Journal of Public Policy and Marketing*, vol. 15, n.º 1, págs. 108-118.

LE JEANNIC, T. (1996). "Une nouvelle approche territoriale de la ville". Économie et Statistique, 294-295, págs. 25-45.

LINO, M. (1990a): "Factors affecting expenditures of single-parent households". *Home Economics Research Journal*, vol. 18, n.º 3, Marzo, págs. 191-201.

LINO, M. (1990b): "Expenditures on a child by husband-wife families". *Family Economics Review*, vol. 3, n.º 3, págs. 2-18.

MAÑAS, E. (1999): *La demanda final de servicios por parte de las familias*. En Cuadrado, J. R. (dir. y coord.): *El sector servicios y el empleo en España. Evolución reciente y perspectivas de futuro*. Bilbao, Fundación BBV, 1999, págs. 377-400.

MAÑAS, E., GABALDÓN, P. y CUADRADO, J. R. (2002): *El consumo de servicios en los hogares españoles. Análisis de su relación con el trabajo femenino y el tiempo*. Madrid, Instituto de Estudios Económicos.

MARCHANTE, A., ORTEGA, B., SORIA, T. y COLOMER, L. (2002): "Renta, Consumo y Ahorro de los Mayores: un análisis aplicado". *Estudios de Economía Aplicada*, vol. 20, n.º 1, págs. 197-215.

McLEOD, P. B. y ELLIS, J. R. (1982): "Housing consumption over the family life cycle: An Empirical Análisis". *Urban Studies*, n.º 19, Mayo, págs. 177-185.

MILLER, N. y KIM, S. (1999): "The Importante of Older Consumers To Small Business Survival: Evidence From Rural Iowa". Journal of Small Bussiness Management, vol. 37, n.º 4, págs. 1-15.

MITCHELL, A. (1996): "Why grey means gold". *Management Today*, Enero, págs. 76-77.

MODIGLIANI, F. (1986): "Life Cycle, Individual Thift, and the Wealth of Nations". *The American Economics Review*, June, vol. 76, n.º 3.

MODIGLIANI, F. y BRUMBERG, R. (1954): "Utility Analysis and the consumption function: an interpretation of cross-section data". *Post Keynesian Economics*, Rutgers University Press.

MOEHRLE, T. (1990): "Expenditure Patterns of the Elderly: Workers and Nonworkers". *Monthly Labor Review*, vol 113, Mayo, págs. 34-41.

MURPHY, P. E. y STAPLES, W. A. (1979): "A modernized family life cycle". *Journal of Consumer Research*, vol. 6, Junio, págs. 12-22.

PAULIN, G. (2000): "Expenditure patterns of older Americans, 1984-97". *Monthly Labor Review*, vol. 123, n.º 5, págs. 3-28.

PETERSON, R. (1992): "The depiction of senior citizens in magazine advertisement: a content analysis". *Journal of Busines Ethics*, vol. 11, págs. 701-706.

PEÑA, D. (2002): *Análisis de Datos Multivariantes*. Madrid, McGraw-Hill/Interamericana de España, S.A.U.

RAMA, R. y CALATRAVA, A. (2001). "Industrialización no tradicional en municipios rurales españoles". *Economía Agraria y Recursos Naturales*, vol. 1, págs. 29-54.

SCHANINGER, C. M. y DANKO, W. D. (1993): "A conceptual and empirical comparison of alternative household life cycle models". *Journal of Consumer Research*, vol. 19, Marzo, págs. 580-594.

SCHROEDER, L.; SJOQUIST, D. L. y STEPHAN, P. E. (1986): "Understanding Regression Analysis". Sage University Paper serie on *Quantitative Applications in the Social Sciences*, vol. 57, Beverly Hills: Sage Publications.

SCHWENK, F. (1995): "Assets of Elderly Households". *Family Economics Review*, vol. 8, n.º 1, págs. 13-19.

SEMON, T. (1995): "Gerontophobia in marketing research". *Marketing News*, vol. 29, n.º 1, págs. 23-24.

SOBERON-FERRER, H. y DARDIS, R. (1991): "Determinants of Household Expeÿÿitures for Services". *Journal ofÿÿonsumer Research*, vol. 17, Marzo, págs. 385-397.

TONGREN, H. (1988): "Determinant Behavior Characteristics of Older Consumers". *Journal of Consumer Affair*, vol. 22, n.º 1, págs. 136-157.

URSIC, A., URSIC, M. y URSIC, V. (1986): "A longitudinal study of the use of elderly in magazine advertising". *Journal of Consumer Research*, vol. 13, págs. 131-133.

WAGNER, J. y HANNA, S. (1983): "The effectiveness of family life cycle variables in consumer expenditure research". *Journal of Consumer Research*, vol. 10, Diciembre, págs. 281-291.

WARD, S. y WACKMAN, D. B. (1972): "Children's purchase influence attemps and parental yielding". *Journal of Marketing Research*, n.º 9, págs. 316-319.

WEAVER, P. (1996): "Opportunity knocks as America ages". Nation's Business, Vol. 84, n.º 8, págs. 30-33.

WELLS, W. y GUBAR, G. (1966): "Life cycle concept in marketing research". *Journal of Marketing Research*, vol. III, Noviembre, págs. 355-363.

WILKES, R. E. (1995): "Household life-cycle stages, transitions and product expenditures". *Journal of Consumer Research*, vol. 22, Junio, págs. 27-42.

ANEXO: VARIABLES UTILIZADAS EN LOS ANÁLISIS ESTADÍSTICOS

Fuentes: Encuestas Continuas de Presupuestos Familiares (Instituto Nacional de Estadística)
Periodo de estudio: 1998-2005 (4º trimestres)
Variables de partida:

- Tamaño del municipio.
  Etiquetas:
  1= "Municipio capital de provincia"
  2= "Municipio no capital con más de 100.000 habitantes"
  3= "Municipio no capital con más de más de 50.000 y menos de 100.000"
  4= "Municipio no capital con más de 20.000 y menos de 50.000"
  5= "Municipio más de 10.000 y menos de 20.000"
  6= "Municipio con menos de 10.000"
- Capital de provincia
  Etiquetas:
  1= "Municipio capital de provincia"
  2= "Municipio no capital de provincia"
- Densidad
  Etiquetas:
  1= "Zona densamente poblada"
  2= "Zona intermedia"
  3= "Zona diseminada"
- Número de miembros del hogar
- Tipo de hogar
  Etiquetas:
  1="Persona sola de menos de 65" (no considerada)
  2="Persona sola de 65 o más años"
  3="Pareja sin hijos"
  4="Pareja con un hijo"
  5="Pareja con dos hijos"
  6="Pareja con tres hijos"
  7="Un adulto con hijos"
  8="Otro tipo de hogares"
- Número de adultos en el hogar
- Número de mujeres en el hogar

Otras variables:

- Intervalo de ingresos
  Etiquetas:
  1= "Hasta 499 euros"
  2= "De 500 a 999 euros"
  3= "De 1.000 a 1.499 euros"
  4= "De 1.500 a 1.999 euros"
  5= "De 2.000 a 2.499 euros"
  6= "De 2.500 a 2.999 euros"
  7= "De 3.000 a 4.999 euros"
  8= "Más de 5.000 euros"

- Suma de servicios del 2º trimestre (en euros)

Factores creados

- **Composición del hogar*:***
  - Número de miembros del hogar
  - Tipo de hogar
  - Número de adultos
  - Número de mujeres
- **Aspectos geográficos y demográficos:**
  - Tamaño del municipio
  - Capital de provincia
  - Densidad

LA CONTRATACIÓN DE LAS PERSONAS CON DISCAPACIDAD EN ESPAÑA: SITUACIÓN ACTUAL Y TENDENCIAS DE FUTURO(*)

MIGUEL ÁNGEL FAJARDO CALDERA
E-mail: fajardo@unex.es
Departamento de Economía Aplicada
Universidad de Extremadura
PABLO ALONSO GONZÁLEZ
E-mail: pablo.alonsog@uah.es
Departamento de Estadística, Estructura Económica y O.E.I.
Universidad de Alcalá
IRENE ALBARRÁN LOZANO
E-mail: ialbarra@est-econ.uc3m.es
Departamento de Estadística
Universidad Carlos III de Madrid

Resumen

Uno de los principales aspectos de la integración de las personas con discapacidad es el referido a su incorporación al mercado de trabajo. En los últimos años, se ha registrado una mayor contratación de personas con estas características. Sin embargo, la situación actual parece sugerir que el número de discapacitados con un contrato laboral puede reducirse.
En este trabajo se hace una breve descripción de la evolución reciente de la contratación de personas con discapacidad en España y se esbozan las principales líneas de tendencia futura, distinguiéndose entre hombres y mujeres. Para ello, se presentan diferentes escenarios de evolución del sistema económico general y su repercusión sobre el empleo de este grupo de personas.

Palabras clave: discapacidad, empleo, género, tendencias futuras
Área temática: Economía Social y Discapacidad.

Abstract

Participation in Labour Market is one of the most important aspects concerned to social integration in handicapped people. During the last years, an increasing number of them have been contracted in Spain. However, the present economic context may suggest that the number of workers belonging to this group may be reduced.
This paper tries to do a brief description of the recent evolution and future trends referred to employment for this group. Gender will be a key factor in the analysis. In order to do this task, different future scenarios for the whole economic system will be shown and the impact of each on the employment will be estimated.

Keywords: disability, employment, gender, future trends
Thematic Area: Social Economics.

1.- INTRODUCCIÓN

La tarea de buscar, obtener y conservar un empleo estable se ha convertido, actualmente, en uno de los objetivos más perseguidos y, se puede decir que, no es siempre fácil. Las encuestas sociológicas de opinión demuestran que una de las cuestiones que más preocupan a la población

(*) Este trabajo ha sido financiado por el Ministerio de Educación y Ciencia y el Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica, proyecto SEJ2005-08070/ECON, cofinanciado con fondos FEDER.

española es el trabajo. El acceso y mantenimiento en el mundo laboral proporciona a la persona (con y sin discapacidad) independencia económica, favorece el sentimiento de utilidad social y la autoestima por la vía del reconocimiento social. Además, el trabajo desde una dimensión ética, ayuda a que la persona se desarrolle como ser humano. De esta forma se convierte en un instrumento fundamental en la vida y, especialmente en la persona con discapacidad, imprescindible para lograr la integración social, la incorporación plena a la sociedad y un desarrollo personal al aumentar la satisfacción personal del individuo y las relaciones sociales. En definitiva, se convierte en un instrumento para lograr una imagen positiva de sí mismo.

Tener un trabajo adecuado a sus habilidades y capacidades es uno de los derechos de los individuos. En el caso de las personas con discapacidad es fundamental que esté adaptado a sus capacidades y habilidades puesto que así se logra una mayor integración en la sociedad mejorando, por tanto, su calidad de vida y alcanzando así una vida más independiente.

El análisis de la situación laboral de las personas con discapacidad no es nuevo. La discapacidad se ha convertido en uno de los ámbitos más potenciados de la protección social en la Unión Europea. Durante los últimos años se han producido avances considerables en las políticas sociales que contribuyen a la integración de las personas con discapacidad. Uno de los pilares de estas políticas es la integración laboral, aspecto básico para un desarrollo social de este colectivo

En los países occidentales la discapacidad, según Giménez y Ramos (2003), se configura como una de las áreas neurálgicas de actuación en la configuración de la protección social de los ciudadanos. A partir de la década de los años noventa, se ha ido abandonando el modelo médico de la discapacidad (Verdugo, 2003) y han surgido otros modelos entre los modelos que cabe destacar el modelo o teoría social de la discapacidad que plantea la concepción de la discapacidad como una construcción social impuesta que conforma una clase social deprimida (Oliver, 1998; Verdugo, 2003) y el modelo constructivista (Schalock, 1997) donde la discapacidad se conceptualiza de acuerdo con las actitudes, prácticas y estructuras institucionales más que por las propias deficiencias. El objetivo fundamental, en este segundo modelo, radica en reducir al máximo tanto las barreras físicas como las sociales (Imrie, 1997).

En relación a la literatura sobre la participación laboral de las personas con discapacidad, en la década de los años 70 aparecen trabajos empíricos centrados en el mercado laboral en Estados Unidos debido, en gran medida, al interés por el análisis de los efectos de sus leyes Antidiscriminación. Posteriormente en los años 90 surgen estudios empíricos relacionados con la participación laboral de los discapacitados en Europa (O'Donnell, 1998; Kidd, Sloane y Ferko, 2000; Malo, 2000; García Serrano y Malo, 2002; Zwinkels, 2001; Malo, 2003; Malo y Dávila, 2006; Dávila, 2006). Siguiendo a Dávila (2006), la creación en 2000 del European Expert Group on Employment for Disabled People ha contribuido de manera significativa al desarrollo de trabajos empíricos en Europa, cuyo interés se centra principalmente en la participación de los discapacitados en el mercado de trabajo.

Además, existen trabajos como el de Loprest et al (1995) que comprueban cómo la participación en el mercado de trabajo de este grupo de personas es superior en los varones que en las mujeres.

Es necesario señalar que, tal y como propone Livermore et al (2000), el análisis de la participación en el mercado de trabajo de las personas con discapacidad en el mercado de trabajo ha de contemplar aspectos que afectan tanto a la oferta como a la demanda de trabajo. Dentro de los aspectos relacionados con la oferta de trabajo de los discapacitados debe tenerse en cuenta que éstos incurren en costes adicionales (transporte, rehabilitación, asistencia técnica o personal, etc). Adicionalmente, existen costes directos de búsqueda de empleo así como otras circunstancias (por ejemplo su menor esperanza de vida en algunos casos) que reducen su participación en el trabajo. Hay que considerar, igualmente, que la discapacidad resta tiempo al trabajo para destinarlo al cuidado médico y personal. Algunos aspectos que pueden afectar a la reducida demanda de trabajo de este colectivo son la discapacidad percibida por el empresario como una señal de baja productividad así como el temor a incurrir en costes adicionales (por ejemplo la adaptación al medio de trabajo). Por desgracia, no se dispone de información a nivel nacional para analizar estos aspectos que, sin duda son importantes y, por tanto, en este estudio no pueden ser abordados.

Es importante señalar que una de las áreas de discriminación más importantes que presenta la población con discapacidad se refleja en el ámbito laboral. Tal y como se detallará en los siguientes apartados, las personas con alguna discapacidad pueden incluirse en el mercado laboral ordinario, en el que se integra toda la población activa española, o en puestos de trabajo protegidos. Concretamente, dos son las principales modalidades de este tipo de trabajo en España: Los Centros Ocupacionales y los Centros Especiales de Empleo (Giménez y Ramos, 2003).

Según el II Plan de Acción para las Personas con Discapacidad 2003-2007 (Ministerio de Trabajo y Asuntos Sociales, 2003), a pesar de estas iniciativas, de las personas con discapacidad sólo el 33,7% (787.800) son personas activas. Sin embargo, en la población sin discapacidad éstas representan una tasa de actividad del 70%. Si realizamos una comparación en función del género, las diferencias también son notables. El 39,3% de los hombres con discapacidad son activos frente al 27,1% de las mujeres. Estos datos son muy inferiores a los de la población sin discapacidad, cuya tasa de actividad tanto en hombres como en mujeres es más del doble que la de la población con discapacidad (84,4% en los hombres y 55,8% en las mujeres).

Teniendo en cuenta las consideraciones mencionadas anteriormente, este trabajo profundiza en el análisis de la situación y previsible evolución futura de la contratación de personas con discapacidad en España describiendo su tendencia global así como la diferenciación en función del sexo.

En el siguiente apartado se describe y analiza la contratación de este colectivo durante los últimos años. Posteriormente, en el apartado 3 se estima la evolución futura de dicha contratación considerando para ello diversos escenarios económicos y analizando la influencia que éstos pueden tener sobre el empleo de las personas con discapacidad. Finalmente, se exponen las principales consecuencias extraídas tras el análisis de los resultados obtenidos comentándose los mismos.

2.- SITUACIÓN LABORAL DE LAS PERSONAS CON DISCAPACIDAD EN ESPAÑA

Antes de pasar a estimar el impacto de la adversa situación económica presente, se procede a realizar una descripción de la población con minusvalía, las personas con discapacidad, que se encuentra actualmente trabajando.

La realidad laboral de este conjunto de la población es heterogénea. Tal y como se ha mencionado, básicamente los puestos de trabajo ocupados por ellos pueden englobarse en dos grandes bloques. Por un lado, están los que se encuentran trabajando en los llamados Centros Especiales de Empleo (CEE). Es lo que se conoce como empleo protegido. Y, en el otro extremo se encuentra el resto del colectivo, que se conoce como empleo no protegido.

Por lo que se refiere al empleo protegido, la información estadística disponible está muy fragmentada. De ella, cabe destacar la contenida en la base Discapnet, creada a partir de un proyecto financiado conjuntamente por el FEDER y la Fundación ONCE. Su acceso se realiza a través de la página web www.discapnet.es. En ella, se clasifican los Centros especiales de empleo en 51 sectores económicos, recogiéndose de cada centro la siguiente información:

- nombre completo del centro
- sectores en los que trabaja
- productos que elabora
- ámbito territorial
- dirección, teléfono, fax y correo electrónico de contacto
- accionistas
- clientes
- ventas
- plantilla total y de personas con discapacidad

- desglose de estas últimas según el tipo de discapacidad, considerándose las siguientes: visual, física, sensorial, psíquica, auditiva, enfermo mental y con parálisis cerebral
- persona de contacto

De acuerdo con la información contenida en esta base, hay 14.644 personas con discapacidad trabajando en los citados Centros. Su desglose por sector de actividad (según CNA-93) se recoge en la tabla 1:

Tabla 1: Distribución por actividad de las personas ocupadas en los CEE

	A	D	F	G	H	I	J	K	M	N	O
Nº	368	4.750	101	3.775	804	51	221	1.524	78	20	2.952
%	2,5%	32,4%	0,7%	25,8%	5,5%	0,3%	1,5%	10,4%	0,5%	0,1%	20,2%

Nota: A = Agricultura, ganadería, caza y silvicultura; D= Industria manufacturera; F = Construcción; G = Comercio, Reparación de vehículos de motor, motocicletas y ciclomotores y artículos personales y de uso doméstico; H = Hostelería; I = Transporte, almacenamiento y comunicaciones; J = Intermediación financiera; K = Actividades inmobiliarias; M = Educación; N = Actividades sanitarias y veterinarias. Servicios Sociales; O = Otras actividades sociales y de servicios prestados a la Comunidad. Servicios personales

Fuente: elaboración propia a partir de Discapnet

A pesar de la evidente utilidad de esta base de datos, es necesario señalar que adolece de una serie de defectos que obliga a cuestionarse los datos que en ella aparecen. En concreto:

1) no se sabe a qué fecha hacen referencia los datos. Únicamente se indica que son de 2008, pero no de qué momento del año, y que pueden ser actualizados a través de la página web, lo cual no quiere decir que lo hayan sido.
2) como consecuencia de lo anterior, no hay una serie histórica que permita hacer el seguimiento del empleo.
3) no se sabe si Discapnet recoge todos los CEE que están abiertos en España. Por tanto, no hay una idea precisa acerca de la representatividad de la base
4) dentro de cada CEE no hay un desglose entre empleados masculinos y femeninos
5) por lo que se refiere al desglose de la plantilla, pueden encontrarse las siguientes situaciones erróneas:
   a) que el centro aparezca con plantilla nula
   b) que el total de la plantilla sea inferior al de personas con discapacidad empleadas
   c) que el total de personas con discapacidad sea inferior a la suma de los siete tipos de discapacidad considerados

Por lo que se refiere a las estadísticas referidas al empleo no protegido, la situación es algo mejor. Existen registros elaborados mensualmente por el Servicio Público de Empleo Estatal (SPEE, antiguo INEM). Los datos se encuentran disponibles con esta periodicidad desde enero de 2002. Sin embargo, hay que destacar que en mayo de 2005 se produjo un cambio metodológico en la elaboración de las series, truncándose la continuidad de las series, salvo en el total anual de contratos. Además, existe otro problema adicional y es que sólo hay estadísticas de número de contratos firmados pero no de número de personas trabajando. Es decir, es posible conocer únicamente el flujo de entrada, pero no el de salida ni el volumen total de personas con discapacidad trabajando.

Los contratos firmados por las personas con discapacidad -a partir de ahora PCD- recogidos en las estadísticas del SPEE pueden ser:

- indefinido minusválidos: contrato de duración indefinida dirigido a la integración laboral de trabajadores minusválidos y que puede ser de jornada completa o a tiempo parcial. Se regulan por el RD 145/183, modificado por RD 4/99 y la Ley 12/2001.
- temporal para minusválidos: contrato temporal de seis meses a tres años, con jornada completa o parcial, para trabajadores minusválidos. Se regulan por la Ley 42/94, RDL 13-96 en su Disposición Adicional Sexta, y por la Ley 12/2001.

Yendo ya a una descripción de la situación laboral de las PCD, es necesario decir que, en líneas generales, su participación en el mercado de trabajo español es muy reducida. En concreto, para el periodo comprendido entre enero de 2006 y diciembre de 2008, solamente el 0,15% de los contratos registrados en el SPEE estaban asociados a personas de este colectivo. No obstante, es cierto que esta participación está creciendo en el tiempo, pues ha pasado del 0,14% en 2006 al 0,16% en 2008. Esta aceleración ha sido más intensa entre los hombres que entre las mujeres, tal y como se recoge en la tabla 2:

Tabla 2: Proporción de contratos de minusválidos por sexos

	Total	Hombres	Mujeres
2006	0,14%	0,17%	0,11%
2007	0,16%	0,20%	0,12%
2008	0,16%	0,20%	0,12%
Media global	0,15%	0,19%	0,12%

Fuente: SPEE y elaboración propia

Por lo que se refiere a su desglose por sectores de actividad, las estadísticas disponibles permiten su agregación tanto por actividades como por sectores económicos. Dentro de las primeras, las cifras disponibles se han agrupado en 10 grandes grupos -GG-, que se identifican por un dígito[1]. En este trabajo se ha considerado la agregación de subgrupos con el primer dígito, que es coincidente con el GG. La tabla 3 recoge, por hombres y mujeres, los pesos anuales según GG para todos los contratos registrados en el SPEE:

Tabla 3: Proporción de contratos por GG y sexos: total del mercado

	Hombres			Mujeres		
GG	**2006**	**2007**	**2008**	**2006**	**2007**	**2008**
1	0,8%	0,7%	0,7%	0,4%	0,3%	0,3%
2	2,0%	1,8%	2,0%	2,6%	2,4%	2,7%
3	7,4%	7,7%	8,3%	6,6%	6,5%	6,9%
4	14,2%	13,6%	14,0%	29,1%	29,8%	27,6%
5	10,2%	9,8%	9,6%	19,3%	20,2%	20,2%
6	1,6%	1,6%	1,5%	0,3%	0,8%	0,2%
7	12,0%	12,2%	10,4%	2,1%	1,6%	1,7%
8	10,7%	12,1%	10,9%	2,2%	2,3%	2,0%
9	41,2%	40,5%	42,7%	37,5%	36,0%	38,4%
0	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

Nota: 1 = Dirección de Empresas y de las Administraciones Públicas; 2 = Técnicos y Profesionales científicos e intelectuales; 3 = Técnicos y profesionales de apoyo; 4 = Empleados de tipo administrativo; 5 = Trabajadores de servicios de restauración, personales, protección y vendedores de comercio; 6 = Trabajadores cualificados en la agricultura y en la pesca; 7 = Artesanos y trabajadores cualificados en actividades manufactureras, la construcción y la minería, excepto operadores de instalaciones y maquinaria; 8 = Operadores de maquinaria e instalaciones, montadores; 9 = Trabajadores no cualificados; 0 = Fuerzas armadas

Fuente: SPEE y elaboración propia

Como puede apreciarse, las mayores concentraciones de empleo se dan en los GG 5, 7, 8 y 9 en hombres (casi un 80% del total, aunque decrece ligeramente en el tiempo) y en 4, 5 y 9 en mujeres (algo más de un 77% del total).

Por sectores económicos, aun cuando mayoritariamente el empleo se centra en el sector servicios, el comportamiento es diferente entre sexos. Así, en la población masculina se ha registrado una reorientación desde la construcción y la industria hacia los servicios y, en menor medida, hacia la agricultura. Por su parte, en el caso de las mujeres, el trasvase ha sido íntegramente hacia los servicios. El resumen de los últimos tres años y los dos primeros meses de 2009 se recoge en la tabla 4:

[1] Este sistema de clasificación está recogido en el Real Decreto 917/84, de 6 de mayo, (CNO-94). Esta clasificación se estructura en 10 grandes grupos -GG-, que se identifican por un dígito. Los grandes grupos a su vez se dividen en 19 grupos principales -GP- identificados con letras. Cada GG puede subdividirse en uno o más GP. El siguiente nivel está constituido por 65 subgrupos principales identificados con dos dígitos.

Tabla 4: Empleo de PCD por sectores económicos

	Hombres				Mujeres			
	2006	2007	2008	2009	2006	2007	2008	2009
Agricultura	1,6%	1,8%	2,0%	2,2%	0,6%	0,6%	0,7%	0,5%
Industria	16,1%	15,4%	14,1%	10,4%	10,5%	9,7%	7,7%	6,5%
Construcción	12,3%	12,7%	10,1%	7,5%	3,4%	3,3%	2,6%	1,9%
Servicios	70,0%	70,1%	73,8%	79,9%	85,6%	86,5%	89,1%	91,2%

Fuente: SPEE y elaboración propia

3.- ESTIMACIÓN DEL NÚMERO FUTURO DE CONTRATOS

El objetivo final de este trabajo es el de evaluar el número de contratos que se podrán firmar en el futuro. Se ha tomado como horizonte el año 2012, ejercicio para el cual, los principales servicios de estudios tanto nacionales como extranjeros, estiman que la actual fase recesiva puede haber concluido.

Para poder estimar el impacto que la desfavorable situación económica actual tendrá sobre el empleo de las personas con discapacidad, es necesario partir de la relación que debería existir entre la evolución de algún agregado económico básico y la citada variable. Como indicativo de la actividad se ha tomado el PIB en términos reales.

Los datos utilizados proceden de diversas fuentes. Así, los referidos a empleo se han tomado de las Estadísticas de contratos, elaboradas mensualmente por el SPEE. Como se ha indicado anteriormente, los datos se encuentran disponibles con esta periodicidad desde enero de 2002. Sin embargo, hay que destacar que en mayo de 2005 se produjo un cambio metodológico en la elaboración de las series que hace que su tratamiento mensual sea imposible. Por ello, se ha trabajado con agregados anuales desde 2002 hasta 2008 ambos inclusive.

Por lo que se refiere a la información sobre PIB y precios, se ha tomado del INE. De esta última fuente también se han recogido las previsiones de población futura española, que se utilizarán para estimar la posible senda de contratos en años venideros para el colectivo analizado.

Tal y como se ha apuntado anteriormente, existe un grado de integración escaso en el mercado laboral. Como se señala en la tabla 5, el número de contratos visados por el SPEE es muy bajo en relación con el del total de contratos, si bien se ha registrado un ligera tendencia al alza en la proporción, siendo más acusada en las mujeres.

Tabla 5: Evolución del empleo en PCD

	Hombres						
	2002	2003	2004	2005	2006	2007	2008
Total	8.042.501	8.247.360	9.086.448	9.495.229	10.217.310	10.056.196	8.784.437
PCD	12.498	12.459	14.068	15.308	17.680	19.660	16.388
%	0,16%	0,15%	0,15%	0,16%	0,17%	0,20%	0,19%
Δ anual							
Total		2,55%	10,17%	4,50%	7,60%	-1,58%	-12,65%
PCD		-0,31%	12,91%	8,81%	15,50%	11,20%	-16,64%
	Mujeres						
	2002	2003	2004	2005	2006	2007	2008
Total	6.136.747	6.420.703	7.264.336	7.669.736	8.489.069	8.565.912	7.816.800
PCD	5.035	5.387	6.522	7.359	9.480	10.175	9.370
%	0,08%	0,08%	0,09%	0,10%	0,11%	0,12%	0,12%
Δ anual							
Total		4,63%	13,14%	5,58%	10,68%	0,91%	-8,75%
PCD		6,99%	21,07%	12,83%	28,82%	7,33%	-7,91%
	Mujeres/Total						
	2002	2003	2004	2005	2006	2007	2008
Total	43,3%	43,8%	44,4%	44,7%	45,4%	46,0%	47,1%
PCD	28,7%	30,2%	31,7%	32,5%	34,9%	34,1%	36,4%

Nota: M/TOTAL = mujeres sobre total
Fuente: SPEE y elaboración propia

Además, es necesario señalar el menor grado de integración de la población femenina, si bien se ha avanzado casi ocho puntos en los siete años analizados, estando muy lejos de la práctica equiparación que se da en el conjunto del mercado laboral español.

Para realizar la estimación del posible número de contratos que pueden ser firmados por PCD en el futuro (2009-2012) se ha distinguido entre hombres y mujeres. Se ha procedido en varias etapas, de la siguiente manera:

Paso 1: estimación de la relación entre el crecimiento del total de contratos firmados y el de la actividad global.

Dado el escaso volumen de datos disponibles se ha estimado un modelo del tipo:

$$CC_t = \alpha + \beta CPIB_t + \varepsilon_t \quad / \quad \varepsilon_t \approx N(0, \sigma_\varepsilon)$$

siendo CC_t la tasa de crecimiento de los contratos en el año t y $CPIB_t$ la del crecimiento del PIB real en ese año. En este caso, el modelo estimado se recoge en la tabla 6.

Tabla 6: relación entre CC y CPIB

Estimador	Hombres	Mujeres
□	-0,13	-0,09
	(0,05)	(0,05)
□	3,92	3,62
	(1,12)	(1,17)
□	4,50%	4,72%
R^2	0,76	0,71

Fuente: elaboración propia

Paso 2: Estimación de la relación entre el crecimiento del total de contratos firmados por PCD y el total de contratos.

Se ha utilizado una regresión del tipo:

$$CPCD_t = a + bCC_t + \gamma_t \quad / \quad \gamma_t \approx N(0, \sigma_\gamma)$$

siendo $CPCD_t$ la tasa de crecimiento de los contratos para *PCD* en el año *t*. En este caso, el modelo estimado se recoge en la tabla 7:

Tabla 7: relación entre CC y CPIB

Estimador	Hombres	Mujeres
a	-0,13	-0,09
	(0,05)	(0,05)
b	3,92	3,62
	(1,12)	(1,17)
□	4,50%	4,72%
R^2	0,76	0,71

Fuente: elaboración propia

Paso 3: Predicción del número de contratos futuros

En esta fase se realizan proyecciones a partir de los modelos anteriormente estimados pero incorporando la información extraída de las proyecciones de población del INE.

Para llevar a cabo las predicciones se hace necesario suponer una determinada senda de evolución del PIB real. Dado el contexto actual en el que se ha producido una abrupta caída de la actividad, se hace difícil poder estimar la evolución que seguirá el PIB. Por ello se ha optado por presentar tres posibles sendas para el periodo analizado (2009-2012), que son las que se recogen en la Tabla 8:

Tabla 8: Posibles sendas de evolución del crecimiento del PIB real

Escenarios	2009	2010	2011	2012
Pesimista	-2,50%	-0,50%	0,50%	0,90%
Neutro	-3,00%	-1,00%	0,00%	0,30%
Optimista	-3,50%	-1,50%	-0,50%	-0,25%

Fuente: elaboración propia

A partir de estas sendas, la proyección de contratos masculinos se logra sin más que proyectar hacia adelante las dos ecuaciones anteriormente estimadas. Usando la estimación del número de contratos totales como variable independiente en la estimación de los contratos para PCD, se obtiene la cifra asociada a la senda de crecimiento del PIB utilizada en cada caso. Por lo que se refiere al empleo femenino, es necesario tener en cuenta que la ecuación que recoge el total de contratos viene afectada por un crecimiento más acusado que en la población masculina.

De extrapolar directamente la ecuación se llegaría a unas cifras de contratos muy superiores a las de los hombres. Ahora bien, en los últimos años se había alcanzado prácticamente la igualdad entre ambos sexos en lo que a número de contratos se refiere, por lo que se parece razonable pensar que esta situación no se alterará en el futuro. Por ello, se ha tomado como tasa de evolución del empleo femenino la correspondiente a la menor de entre la proporción que representan las mujeres en edad de trabajar -obtenida a partir de las proyecciones del INE- y la que se obtiene de aplicar el modelo estimado. A partir de esta tasa, se estima el número de contratos correspondiente a mujeres. Con se hiciera en el caso anterior, se utiliza esa tasa como variable independiente en la ecuación que recoge el número de contratos para mujeres con discapacidad, obteniéndose finalmente la tasa estimada y a partir de ella, el número de contratos asociado. Los resultados son los recogidos en la Tabla 9:

Tabla 9. Número estimado de contratos y % de empleo femenino

Sexo	Escenario	2009	2010	2011	2012	Acumulado
Hombre	Optimista	12.128	10.194	9.081	8.272	39.675
	Neutro	11.716	9.554	8.270	7.284	36.824
	Pesimista	11.305	8.934	7.509	6.406	34.154
Mujer	Optimista	7.226	6.325	5.572	5.043	24.166
	Neutro	6.966	5.850	4.978	4.327	22.121
	Pesimista	6.707	5.393	4.428	3.702	20.231
% empleo femenino	Optimista	37,3%	38,3%	38,0%	37,9%	
	Neutro	37,3%	38,0%	37,6%	37,3%	
	Pesimista	37,2%	37,6%	37,1%	36,6%	

Fuente: elaboración propia

Finalmente, cabría pensar en estimar el coste en términos de empleo de la actual situación. Si suponemos un crecimiento anual del PIB del 3%, el número de contratos perdidos por año y acumulados, por escenario y sexo, es el que se recoge en la tabla 10:

Tabla 10: Número estimado de contratos y % empleo femenino perdidos

Sexo	Escenario	2009	2010	2011	2012	Acumulado
Hombre	Optimista	4.529	6.736	8.127	9.219	28.611
	Neutro	4.941	7.377	8.938	10.206	31.461
	Pesimista	5.352	7.996	9.699	11.084	34.131
Mujer	Optimista	2.852	4.514	6.085	7.168	20.620
	Neutro	3.111	4.989	6.679	7.885	22.665
	Pesimista	3.371	5.446	7.229	8.509	24.555
% empleo femenino	Optimista	0,4%	0,7%	2,4%	3,2%	
	Neutro	0,4%	1,1%	2,8%	3,8%	
	Pesimista	0,5%	1,4%	3,3%	4,5%	

Fuente: elaboración propia

4.- CONCLUSIONES

La evaluación del nivel de integración laboral de un determinado colectivo impone la existencia de estadísticas veraces y completas que permitan seguir la evolución de las variables que se quieren medir. En el caso de la población española con discapacidad, las estadísticas disponibles son escasas y reflejan únicamente determinados aspectos de su participación en el mercado de trabajo. Así, las cifras referidas al empleo protegido son muy escasas y de reducido valor, mientras que las del resto del empleo ofrecen una visión más completa, especialmente en lo que se refiere al número de contratos firmados. Sin embargo, no es posible conocer ni el flujo de reducciones de contratos ni el número total de personas con estas características que se encuentra actualmente empleadas.

Partiendo de estas carencias en la información de base se ha pretendido estimar el impacto que sobre el nivel de ocupación de este colectivo tiene y puede tener en el futuro la actual situación económica, caracterizada por un hundimiento generalizado y rápido de la actividad económica, la cual, ha incidido duramente sobre el empleo, en especial el de las PCD. Las estimaciones sugieren que el número de contratos va a ir disminuyendo a lo largo del periodo considerado, al principio de forma más intensa en hombres que en mujeres.

Otra de las consecuencias de esta crisis es que la convergencia en el empleo femenino se va a ver truncada manteniéndose en tasas similares a las actuales o incluso comenzando a decrecer a partir de 2011, según el escenario considerado. Si se utiliza un escenario hipotético de no crisis se habrían firmado entre 28 y 34.000 nuevos contratos en hombres y entre 20 y 24.000 en mujeres y la tasa de participación femenina podría haber superado el 41% en 2012. Es decir, el número de contratos potencialmente perdidos sería semejante e incluso superior, según el escenario considerado, al de los que pueden firmarse durante el periodo considerado.

5.- BIBLIOGRAFÍA

CES (2003). La Situación de las Personas con Discapacidad en España, Colección Informes, Madrid: Consejo Económico y Social.

COMITÉ ESPAÑOL DE REPRESENTANTES DE MINUSVÁLIDOS -CERMI- (2001). Un Plan de Empleo para las Personas con Discapacidad en el siglo XXI, Madrid.

CONSEJO DE EUROPA (1998). Recomendación nº (98)9 del Comité de Ministros a los Estados miembros relativa a la dependencia.

DÁVILA, C.D. (2006). "Discapacidad y género: un estudio de participación en el mercado de trabajo español". Moneda y crédito, 223, pp. 127-158.

GANNON, B. Y NOLAN, B. (2003). Disability and Labour Market Participation, HRB Documento de trabajo, Economic and Social Research Institute, Ireland.

GARCÍA ALONSO, J. V. (Coord.) (2003). El movimiento de Vida Independiente. Experiencias internacionales, Madrid: Fundación Luis Vives.

GARCÍA SERRANO, C. Y MALO, M.A. (2002). "Discapacidad y mercado de trabajo en la Unión Europea". Cuadernos Aragoneses de Economía, 12, 2, pp. 237-255.

GIMÉNEZ, D. Y RAMOS, M.M. (2003). "La discriminación de las mujeres con discapacidad en España". Revista del Ministerio de Trabajo y Asuntos Sociales, 45, pp. 61-76.

IMRIE, R. (1997). "Rethinking the relationships between disability, rehabilitation and society", Disability and Rehabilitation, 19, pp. 263-271.

IMSERSO (2003). La Discapacidad en cifras, Madrid.

INE (2002a). Encuesta sobre Discapacidades, Deficiencias y Estado de la Salud 1999. Resultados detallados, Madrid.

INE (2002b). Encuesta sobre Discapacidades, Deficiencias y Estado de la Salud 1999. Metodología, Madrid.

INEM (2000). Información sobre mercado de trabajo. Resumen anual de datos del Observatorio Ocupacional 1999, Madrid: Observatorio Ocupacional.

INEM (2002). Información sobre mercado de trabajo. Resumen anual de datos del Observatorio Ocupacional 2000, Madrid: Observatorio Ocupacional.

INSS (1994). Clasificación Internacional de Deficiencias, Discapacidades y Minusvalías, Madrid: Ministerio de Trabajo y Asuntos Sociales.

JIMÉNEZ, A. Y HUETE, A. (2002). Epidemiología de la Discapacidad, Madrid: Real Patronato sobre discapacidad.

KIDD, M., SLOANE, P. Y FERKO, I. (2000). "Disability and the Labour Market: An Analysis of British Males". Journal of Health Economics, 19, pp. 961-981.

LEY 39/2006 de Promoción de la Autonomía Personal y Atención a personas en situación de dependencia, BOE de 15 de diciembre de 2006.

LIVERMORE, G.A., STAPLETON, D.C, NOWAK, M.W., WITTENBURG, D.C., Y EISEMAN, E.D. (2000). The Economics of Policies and Programs Affecting the Employment of People with Disabilities, Cornell University.

LOPREST, P., RUPP, K., Y SANDELL, S.H. (1995). "Gender, Disabilities, and Employment in the Health and Retirement Study". The Journal of Human Resources, XXX. Suplemento, S293-S318.

MADDALA, G.S. (1990): Limited dependent and qualitative variables in Econometrics, Cambridge University Press.

MALO, M.A. (2000) "Discapacidad y participación en el mercado de trabajo: Una comparación de la situación española con la del conjunto de países europeos del panel de Hogares" en L. Garrido y L. Toharia, Condiciones de vida en España y en Europa, Madrid: INE.

MALO, M.A (2003). "Las Personas con Discapacidad en el Mercado de Trabajo Español". Revista del Ministerio de Trabajo y Asuntos Sociales, 46, pp. 99-126.

MALO, M.A. Y DÁVILA, C.D. (2006). "Género, discapacidad y posición familiar: la participación laboral de las mujeres con discapacidad". Cuadernos aragoneses de economía, 16, 1, pp. 61-82.

MINISTERIO DE TRABAJO Y ASUNTOS SOCIALES (2001). Plan de Acción para el Empleo del Reino de España 2001, Madrid.

MINISTERIO DE TRABAJO Y ASUNTOS SOCIALES (2003). Plan Nacional de Acción para el Empleo 2003, Madrid.

O'DONNELL, O. (1998). The Effect of Disability on Employment Allowing for Work Incapacity, University of Kent, Department of Economics, Discussion Paper 98/13.

OLIVER, M. (1998). "¿Una sociología de la discapacidad o una sociología discapacitada?", en L. Barton (ed.) Discapacidad y sociedad, pp. 34-58 (Madrid, Morata/Fundación Paideia).

OMS (1997). Clasificación Internacional de Deficiencias, Discapacidades y Minusvalías (CIDDM), Madrid: INSERSO.

PUGA, M.D. Y ABELLÁN A. (2004). El proceso de discapacidad. Un análisis de la Encuesta de Discapacidades, Deficiencias, y Estado de Salud, Madrid: Fundación Pfeizer.

REAL DECRETO 504/2007, DE 20 DE ABRIL, por el que se aprueba el baremo de valoración de la situación de dependencia establecido por la Ley 39/2006, de 14 de diciembre, de promoción de la autonomía personal y atención a las personas en situación de dependencia.

RIVAS, F. (1999). "Conducta, asesoramiento vocacional y minusvalía", en F. Rivas y M. L. López (coords.) Asesoramiento vocacional de estudiantes con minusvalías físicas y sensoriales. Valencia. Servei de Publicacions de la Universitat de València.

SCHALOCK, R. L. (1997). "Can the concept of quality of life make a difference?", en R. L. Schalock (ed.) Quality of life: Vol. II. Application to persons with disabilities, pp. 245-267 (Washington, D. C., American Association of Mental Retardation).

STERN, S. (1989). "Measuring the Effect of Disability on Labor Force Participation". The Journal of Human Resources, XXIV, 3, pp. 361-395.

VERDUGO, M.A. (2003). "La concepción de la discapacidad en los modelos sociales", en: M.A. Verdugo Alonso y F.B. Jordán de Urríes Vega (coord.) Investigación, innovación y cambio. Salamanca. Amaru.

WILKINS, R. (2003). Labor Markets Outcomes and Welfare Dependence of Persons with Disabilities in Australia, Melbourne Institute, Documento de trabajo nº 2/03.

ZWINKELS, W. (2001). The Employment Situation of People with Disabilities in the European Union, Research paper, EIM Business and Policy Research.

UNEMPLOYMENT, FINANCIAL FACTORS AND WELL-BEING

MARIA DA CONCEIÇÃO COSTA PEREIRA
E-mail: mpereira@fe.uc.pt
FILIPE JORGE FERNANDES COELHO
E-mail: fcoelho@fe.uc.pt
Faculdade de Economia da Universidade de Coimbra

Resumo

Este estudo contribui para a economia do bem-estar explorando algumas questões que não foram previamente investigadas ou que receberam pouca atenção na literatura, usando dados do European Social Survey. Em primeiro lugar, este estudo é aparentemente um dos primeiros a investigar a forma como as restrições ao crédito, sob a forma de dificuldades sentidas em obter crédito, se relacionam com o bem-estar subjectivo, na presença de outras variáveis relacionadas com o rendimento. Em segundo lugar, também consideramos o grau com que as pessoas conseguem satisfazer as suas necessidades de consumo com o rendimento corrente, investigando o seu impacto no bem-estar subjectivo. Em terceiro lugar, investigamos o papel de experiências de desemprego passadas no bem-estar presente.

Palabras Clave: bem-estar subjectivo, restrições ao crédito, desemprego.
Área Temática: Economia Social.

Abstract

This paper contributes to the literature on the economics of happiness by exploring some issues that have not been investigated before or scarcely so, drawing upon data from the European Social Survey. Firstly, this study is apparently one of the first to investigate how borrowing constraints, in the form of perceived difficulties in obtaining credit, relate to subjective well-being (SWB) in the context of other income related variables. Secondly, we also consider the extent to which people make ends meet, investigating the impact of how people are coping with their present income on SWB. Thirdly, we further investigate the role of past unemployment experiences in shaping current SWB.

Keywords: subjective well-being, life satisfaction, borrowing constraints, unemployment.
Thematic Area: Social Economics.

1. INTRODUCTION

Recently, there has been a surge of research in the economic literature, namely in mainstream economics journals, investigating the determinants of subjective well-being. Improving the welfare of individuals is an important goal of economic policy, and this provides an important motivation for such research (e.g. Van Praag Frijters and Ferrer-i-Carbonell, 2003; Hayo and Seifert, 2003). The base for this research involves self-reports of people about their happiness or life satisfaction. However, being inherently subjective, reports on general life satisfaction raise serious doubts to many researchers, and this explains why economists have long avoided dealing with such types of information, preferring instead to study the actual choices of people (Kahneman and Krueger, 2006).

Although common in the social sciences, the collection and analysis of people's beliefs, values and attitudes has been largely excluded from economics due to the positivist paradigm (Easterlin, 2001). However, considering the abundant evidence of individuals' choices departing from rationality, Kahneman and Krueger (2006: 3) conclude that "if people display bounded rationality when it comes to maximizing utility, then their choices do not necessarily reflect

their "true" preferences, and an exclusive reliance on choices to infer what people desire loses some of its appeal."

SWB is an individual's global evaluation of how he/she thinks and feels about his/her life (Dolan, Peasgood, and White, 2008). It has been established that such personal reports contain many imperfections, being affected, for example, by the interaction with the interviewer, question order effects, social desirability biases, respondent's mood state, consistency predisposition and memory effects (e.g., Kahneman and Krueger, 2006; Bertrand and Mullainathan, 2001; Podsakoff et al., 2003). Notwithstanding, evidence has been found that global satisfaction or happiness questions correlate in expected ways with a number of measures. Kahneman and Krueger (2006) and Blanchflower and Oswald (2004), for example, summarize a number of studies on neural science and psychology that establish robust correlations between self-reported general life satisfaction or happiness levels and physiological measures (such as frequent smiling, smiling with the eyes, one's happiness evaluated by friends and family), health factors (e.g. heart rate and blood-pressure measures in reaction to stress, reports of headache and digestive disorders, recovery from illness), and objective life circumstances (like unemployment, marriage, sudden illness, sudden increase in income). These and other studies have given support to the accuracy of SWB data, and this has been a major reason why research on the economics of happiness has been flourishing in these last 10 years or so (Kahneman and Krueger, 2006; Kahneman Krueger and Schkade, 2004; Senik, 2005). The fact that Daniel Kahneman has been laureated with the economics Nobel Prize in 2002 has probably also contributed to unwind the prejudice that persisted towards the idea of economists meddling with such types of data.

Given the recent surge in the SWB literature, a number of factors have been identified as determinants. Besides personal characteristics (e.g., age, gender, race, and marital status), some of the most commonly tested determinants of happiness include absolute and relative income, education, employment status and health. More recently, the variety of determinants that received researchers' attention has been enlarged to include religion (e.g., Clark and Lelkes, 2009), voluntary work (Meier and Stutzer, 2006), environment (e.g., Welsch, 2006), and democracy (e.g., Frey and Stutzer, 2000), amongst others. Some extended surveys on the economics literature concerning the determinants of SWB are provided by Frey and Stutzer (2002) and Dolan Peasgood and White (2008), Clark, Frijters and Shields (2007).

In this paper we look at two important and interrelated determinants of SWB: unemployment experiences and financial issues. These are two top issues affecting individuals' happiness. For example, Easterlin (1974) commented that financial issues are amongst individuals' top worries. The literature on SWB also contends that humans are highly vulnerable to joblessness, and that unemployment constitutes an important economic determinant of unhappiness (Clark and Oswald, 1994; Oswald, 1997). In terms of unemployment, our concern lies essentially on the interactions between the effects of past and present individual unemployment experiences. As to the income-related determinants we study the impact of perceived income (how people are coping with their income) and the role of borrowing constraints, in regressions in which we control for a number of other variables, thus providing a more robust test of the effects analysed.

Thus, this study enables us to make a number of contributions. Firstly, we further investigate the role of past unemployment experiences in shaping current SWB, an issue that has not received that much attention in past work (Clark, Georgellis, and Sanfey, 2001). In addition, we consider this in the context of current employment status and also of financial issues. Secondly, the study considers the extent to which people make ends meet, investigating how people are coping with their present income, and the corresponding impact on SWB. This gauges the extent to which individuals' needs and consumption standards are being met, an issue that has not been much investigated in SWB economic studies. Thirdly, as far as we are aware, this study also innovates in investigating how borrowing constraints, in the form of perceived difficulties in obtaining credit, relate to SWB in the context of other income related variables.

The paper is organized as follows. In section 2 we survey the literature focusing on unemployment and financial determinants of SWB. In section 3 we present the methodology, section 4 reports the estimation results, and section 5 concludes.

2. RESEARCH BACKGROUND

2.1. Subjective well-being and unemployment

The literature reveals a robust negative relationship between unemployment and SWB (e.g., Clark, Georgellis, and Sanfey, 2001; Clark and Oswald, 1994; Hayo and Seifert, 2003). Explanations for this negative effect have been associated with financial losses and concerns about future financial situation, as well as with other non-pecuniary issues arising from the unemployment state. Winkelman and Winkelman (1998), for example, mentions the non-pecuniary costs of unemployment, which are related with its impact on the individual's self-esteem, social relationships, and identity in society.

The literature further provides some evidence that past unemployment experiences leave a scar on individuals (e.g., Clark, Georgellis, and Sanfey, 2001; Lucas et al., 2004), affecting their current SWB beyond their current occupation status. Lucas et al. (2004), for example, observe that once the unemployment situation ends and individuals return to work, they do not return to previous SWB levels. Against this, set-point theories advance that individuals' SWB reacts to life events but soon return to previous baseline levels (Suh, Diener, and Fujita, 1996).

Moreover, there is some evidence in the literature that the impact of current unemployment on SWB is reduced for those with more extensive unemployment experiences in the past, implying an adaptation effect. Clark, Georgellis, and Sanfey (2001), for example, observed a lower coefficient for current unemployment once past unemployment history was taken into account. However, Winkelman and Winkelman (1998) observed for the unemployed that satisfaction is not affected by unemployment duration. In addition, Lucas et al. (2004) found that the effect of current unemployment was not affected by previous unemployment experience.

2.2. Subjective well-being and financial issues

Income and related issues have been the focus of much research (Dolan, Peasgood, and White, 2008), with the literature pointing to a positive impact on SWB (e.g., Di Tella, MacCulloch, and Oswald, 2001, Frey and Stutzer, 2002). However, past research indicates that the robustness of this effect is not that strong, and that there are diminishing returns from increasing income (Clark and Oswald, 1994; Dolan, Peasgood, and White, 2008; Frey and Stutzer, 2002). In this respect we refer readers to the surveys Clark, Frijters and Shields (2007), Frey and Stutzer (2002).

In this vein, Diener (1984), amongst others, notes that what matters is not simply income, but the extent to which individuals are satisfied with their income. One way of looking at satisfaction with income is by assuming that people care about other people's income (McBride, 2001; Senik, 2004). This involves the notion that utility is relative (Senik, 2004), with an individual's utility being influenced by comparison with others in the reference group. Easterlin (1995) suggested that when the income of everyone increases, the happiness of people does not increase, because of the general increase in living norms. More specifically, the well-being of an individual is influenced by how his/her income compares with that of other people around him/her. This differs from mainstream economics, which considers that an individual's utility, for example, results from his/her own consumption (Luttmer, 2005). In this context, a significant number of studies have been developed to investigate the role of interdependent preferences on individual's behaviour (for a review see, for example, Ferrer-i-Carbonell, 2005). In particular, a number of studies have considered the role of reference income in shaping an individual's SWB, indicating that it reacts as much or even more to reference income than to absolute income levels (e.g., Ferrer-i-Carbonell, 2005; Luttmer, 2005; McBride, 2001). This has

led some economists to question whether a major goal of economic policy should be the equality of incomes rather than economic growth (e.g., McBride, 2001).

Ferrer-i-Carbonell (2005) adds that an additional explanation for the relative lower importance of absolute income is that what matters to people is their subjective evaluations of the extent to which their income satisfy their needs. She further adds that people adapt their expectations, so that an increase in income produce a short-lived increase in well-being, as increases in income are accompanied by raising expectations.

Given the inconsistent findings in the literature concerning the effects of past unemployment experiences on reported levels of SWB, we investigate how unemployment periods of different lengths and the recency of that experience affects an individual SWB, and whether these effects influence different groups of people equally. Regarding financial issues, this study also considers the role of relative income, but adds two additional factors that have not received much attention from the literature on SWB, including the extent to which individuals are coping with their present income, and the role of borrowing constraints.

3. METHODOLOGY

The analysis draws upon data collected recently, in the three rounds of the European Social Survey (ESS)[1], which is funded by the European Commission, the European Science Foundation, and scientific funding bodies in each participating country. The first round was fielded in 2002/2003, the second in 2004/2005 and the third in 2006/2007. The data covers 24 countries[2], but some are totally or partially excluded on account of missing values[3], apart from the fact that some of the countries did not participate in one round.

We will adopt the typical model used for estimating subjective satisfaction responses, which assumes the form:

$$IW^*_i = \alpha X_i + \beta_i + \delta_t + e_i$$

where IW* is the well-being level reported by individual i, which acts as a proxy for the true individual well-being IW, X is a vector of explanatory variables, β is a country fixed effect, δ is a time effect (dummy variable for each round) and e_i captures the unobservable personality traits / disposition and other non available factors (e.g. environment degradation, amount of leisure time), embracing as well the error term associated with the subjectivity of the responses (Blanchflower and Oswald, 2004). Country dummies capture cultural and other institutional factors and round dummies capture time effects, like those related to prices. This model is estimated as ordered probit in which, as usual, the true individuals' well-being is the latent variable.

The dependent variable in this study (IW*), life satisfaction, is measured by the question "All things considered, how satisfied are you with your life as a whole nowadays?". Answers to this question vary on a 0 to 10 scale, where 0 means extremely dissatisfied and 10 means extremely satisfied. The explanatory variables include gender[4], age, marital status[5], main

[1] The European Social Survey (ESS) aims at measuring and explaining trends in attitudes, beliefs and values across countries in Europe and its close neighbours. The questionnaire includes two main sections, each consisting of approximately 120 items; a 'core' module which will remain relatively constant from round to round, plus two or more 'rotating' modules, repeated at intervals. For more details and data access please consult the ESS home site at www.europeansocialsurvey.org.

[2] The cumulative file contains data from countries that have fielded two or three rounds: Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland, the United Kingdom and Ukraine.

[3] Estonia and Ukraine were eliminated from this study on the basis of missing data on income in the rounds they fielded (2nd and 3rd). Other countries were excluded from some rounds: France in the first round, Hungary in the first and third rounds, and Ireland in the third round, again because of missing data on income; Austria was eliminated from the first round and the United Kingdom in the second round, because of missing data on education level; France was eliminated from the second round due to missing data on the variable feelings about one's income.

[4] Gender is a dummy with zero for male and one for female.

[5] Marital status considers individuals living with a partner in the married category, with the single people constituting the omitted category. In the third round of the ESS a variant to the legal marital status question (F62) has been introduced in countries where

activity/employment status[6], education level[7], domicile (a description of the area where one lives)[8], income, past unemployment experiences (and this includes whether individuals have ever been unemployed for a period between 3 and 12 months, over 12 months, and whether any of these periods occurred in the last 5 years), reference income, perceived income and borrowing constraints. Before the estimation process, we have deleted those observations with missing data on any of these variables, which led to a final sample with 74,478 individuals.

4. RESULTS

Conforming to previous empirical studies, we first estimate an equation with a set of traditional explanatory variables. These include gender, age, marital status, main activity/employment status, education level, domicile and income. Income is considered in equivalent terms, meaning that it is adjusted for differences in household size, in a way that the existence of economies of scale in consumption is admitted. It was obtained by dividing the midpoint of the monthly income interval bracket by the square-root of the number of household members. We also tested the OECD-modified equivalence scale to construct equivalent income. This method involves assigning a value of 1 to the first adult household member, of 0.5 to each additional member and of 0.3 to each child under 14. However, the results did not show to be sensitive to the equivalence scale used. Hence, we will only report results for the former method.[9]

The results of this first estimation, reported in equation (1), generally conform to the evidence provided in the literature. More specifically, woman report higher satisfaction levels than man; life satisfaction is U-shaped in age; living with a partner/spouse contributes positively to SWB, whereas separation, divorce and widowed contributes negatively to it; unemployment is negatively associated with SWB; higher education levels are associated with higher levels of SWB; income is positively associated with SWB; finally, the estimated coefficients for domicile reveal that the place where one lives matters significantly, with living in smaller, more peaceful agglomerations contributing positively to well being.

legal partnerships other than marriage (legal recognition of a union between two people other than marriage whether of the opposite or same sex) exist. Respondents in these countries replied to version A of question F62 with the following categories: 1. Married; 2. In a civil partnership; 3. Separated (still legally married); 4. Separated (still in a civil partnership); 5. Divorced; 6. Widowed; 7. Formerly in a civil partnership, now dissolved; 8. Formerly in a civil partnership, partner died; 9. Never married and never in a civil partnership. Whereas in countries with no legal partnerships other than marriage, respondents can only choose from categories 1, 3, 5, 6 and never married in version B. Hence bridging between the rounds in those countries is not straightforward. Furthermore, it has been noticed to us (Orten, undated) that in some countries the word "married" is used in a colloquial rather than in a legal sense. Some respondents choose such an answer, in describing their relationship to their de facto partner, when their legal marital status is something else, thus the answer married tends to be over-reported compared to Eurostat statistics. Given these constraints, we have adopted the following procedure to bridge between rounds. We have created a category for those who are married and live with a partner, whether the relationship is legally recognized or not. These include the married in round 1 and 2, the married and in a civil partnership in round 3 and those who answer yes in question "are you currently living with a partner?" in all rounds. In round 1 and 2 and in countries where version B in round 3 is used, those who are separated, divorced, widowed or never married and answer no in question "are you currently living with a partner?", see their category preserved. For countries where version A is used in round 3, we aggregate categories 3 and 4, categories 5 and 7 and categories 6 and 8. Again, those who are in these categories and in category 9 and answer no in question "are you currently living with a partner?" remain in their reported category; Although we recognize that this bridging between rounds is far from perfect, note that respondents in categories 2, 4, 7 e 8 is quite residual.

[6] Employment status considers a number of categories, with the omitted variable being "in paid work (or away temporarily)(employee, self-employed, working for your family business)". Unemployed category aggregates both who are actively looking for a job and who want a job but are not actively looking for one. The other categories are retired, home care (doing housework, looking after children or other), in school (in education (not paid for by employer) even if on vacation), permanently sick or disabled, and other activity, which incorporates "in community or military service".

[7] Education level omitted combines "not completed primary education" and "primary or first stage of basic". The other levels are: "lower secondary or second stage of basic"," upper secondary", "post secondary, non-tertiary", "first stage of tertiary" and "second stage of tertiary".

[8] Specifically: big city, suburbs or outskirts of big city, town or small city, country village, farm or home in the country side. The omitted category is a big city.

[9] Another experiment concerning income consisted in computing it in purchasing power parity, using comparative price level indices per country and year the interview took place. This did not change the results either. Note that country and year dummies capture, amongst many others, the differences in purchasing power across countries.

4.1. The effects of unemployment experiences

We now pay a closer look at past unemployment experiences, which has been shown to affect the SWB of individuals apart from current employment status (e.g., Clark, Georgellis, and Sanfey, 2001; Winkelmann and Winkelmann, 1998). The specific questions that we use from the ESS are as follows: "*Have you ever been unemployed and seeking work for a period of more than 3 months?*"; "*Have any of these periods lasted 12 months or more?*"; "*Have any of these periods been within the past 5 years?*" Regarding the two initial questions, we created a variable with the value 1 for those individuals that have been unemployed for a period between 3 and 12 months (and zero otherwise), which we called "short", and another variable with the value 1 for those who have been unemployed for a period of over 12 months (and zero otherwise), which we designated "long". Therefore, we investigate whether individuals' current SWB is affected by short and/or long unemployment periods, and by whether any of these unemployment experiences occurred in the past 5 years.

Equation (2) adds to equation (1) whether people have ever been unemployed for a period from 3 to 12 months. The results indicate that such past experience detrimentally affects SWB, and that this effect does not substantively alter the coefficient for current unemployment status. In equation (3) we add to the estimation whether individuals have ever had a long unemployment experience in the past. We observe that such a large unemployment period has a much stronger negative impact on SWB than shorter unemployment periods, and that it reduces the coefficient for current unemployment. Overall, these results are consistent with the scarring hypothesis of Clark, Georgellis, and Sanfey (2001), according to which experiences of past unemployment leave a psychological mark on individuals. In addition, our results indicate that longer unemployment experiences produce a much stronger scar on SWB than shorter past unemployment experiences.

There is also some support in the literature to the adaptation hypothesis, according to which the way individuals feel their current unemployment status is tempered by past unemployment experiences. Accordingly, being unemployed in the past should mitigate an individual's disutility derived from the current unemployment situation. Our results are also consistent with this adaptation hypothesis, as introducing unemployment periods lasting more than 12 months, in equation (3), substantially reduces the negative effect of current unemployment on SWB. However, our results further add in that the adaptation hypothesis is more likely to take place for longer past unemployment situations, as consideration in equation (2) of being unemployed from 3 to 12 months did not change in a meaningful way the coefficient for current unemployment. This suggests that there may exist a threshold above which the adaptation hypothesis takes place. This may account, to some extent, for the contradictory results in the literature, as Clark, Georgellis, and Sanfey (2001) obtained evidence to support the adaptation hypothesis, whereas Winkelman and Winkelman (1998) and Lucas et al. (2004) have failed to do so.

Table 1: Life satisfaction equations and past unemployment experiences

	(1)		(2)		(3)		(4)	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Gender	0.0830	10.33	0.0812	10.11	0.0825	10.27	0.0827	10.29
Age	-0.0352	23.72	-0.0353	-23.79	-0.0320	21.44	-0.0328	21.91
Age squared×1000	0.3593	23.81	0.3575	23.69	0.3195	21.04	0.3254	21.40
Marital status								
Married/l.w. partner	0.2682	22.70	0.2674	22.63	0.2614	22.11	0.2603	22.02
Separated	-0.2750	-7.91	-0.2742	-7.89	-0.2693	-7.74	-0.2688	-7.73
Divorced	-0.0947	-4.98	-0.0927	-4.87	-0.0836	-4.39	-0.0823	-4.33
Widowed	-0.0775	-4.05	-0.0778	-4.07	0.0824	-4.31	-0.0830	-4.34
Employment status								
Unemployed	-0.5174	-27.98	-0.5118	-27.66	-0.3827	-19.75	-0.3442	-17.10
Retired	0.0002	0.02	-0.0023	-0.16	0.0014	0.10	0.0014	0.09
Homecare	-0.0149	-1.06	-0.0180	-1.28	-0.0018	-0.13	0.0027	0.19
In education	0.1354	7.17	0.1209	6.38	0.1041	5.49	0.1000	5.27
Sick	-0.5649	-21.72	-0.5661	-21.76	-0.5448	20.93	-0.5442	-20.90
Other	-0.0990	-2.73	-0.1041	-2.87	-0.0863	-2.38	-0.0819	-2.26
Education								
Lower secondary	0.1041	7.78	0.1047	7.83	0.1066	7.96	0.1070	7.99
Upper secondary	0.1414	10.68	0.1439	10.87	0.1407	10.63	0.1409	10.64
Post sec. (non tert.)	0 .1937	9.94	0.1957	10.04	0.1886	9.67	0.1891	9.69
1st stage tertiary	0.1711	11.38	0.1731	11.51	0.1628	10.82	0.1628	10.83
2nd stage tertiary	0.1794	7.71	0.1808	7.77	0.1674	7.20	0.1679	7.22
Domicile								
Suburbs	0.0190	1.37	0.0187	1.34	0.0157	1.13	0.0142	1.02
Town	0.0305	2.61	0.0301	2.58	0.0291	2.49	0.0280	2.40
Village	0.0966	8.19	0.0953	8.07	0.0904	7.66	0.0896	7.59
Farm	0.1605	8.99	0.1589	8.90	0.1489	8.34	0.1480	8.29
Income×10^3	0.0807	22.70	0.0795	22.32	0.0736	20.64	0.0731	20.49
Past unemployment								
Short			-0.0957	-8.56	-0.1482	-12.97	-0.1065	-8.28
Long					-0.2892	-22.50	-0.2313	-15.18
Last 5 years							-0.1133	-7.07
Number Observations	74478		74478		74478		74478	
LR	15280.93		15354.19		15860.18		15910.16	
LL	-143741.40		-143704.77		-143451.78		-143426.78	
Pseudo R^2	0.0505		0.0507		0.0524		0.0525	

Estimations include round and country dummies, which are not reported. Cutpoints are also omitted, for parsimony reasons.

In equation (4) we add whether individuals had a past unemployment situation lasting more than 3 months in the last 5 years. Interestingly, the results indicate that a recent (in the last 5 years) unemployment situation (lasting more than three months) produced an additional detrimental effect on SWB. This indicates that the more recent a lengthy experience of unemployment, the stronger the impact on current well-being. Recent unemployment experiences are fresher in people's minds, thus leaving a stronger scar, thus more adversely affecting SWB. In some way, this is consistent with set-point theories of SWB and with findings that life events such as marriage and bereavement have mainly temporary effects on SWB (e.g., Clark, Diener, Georgellis, and Lucas; 2008; Kahneman and Krueger, 2006; Oswald and Powdthavee, 2005).

4.2. The effects of economics factors

4.2.1 Reference income

We now investigate the role of financial factors in shaping individuals' SWB (Table 2). In the previous block of equations we observed that equivalent income was positively related to SWB, a result that is consistent with previous research. We now start by considering in Table 2 the role of reference income. For comparison reasons, equation (4) in Table 2 has been reproduced in Table 1. Equation (5) adds to equation (4) the variable reference income, which was computed as the mean equivalent income of the individual's reference group defined by country, round, and education level[10]. The results of this estimation indicate that reference income has a significant negative impact on SWB, i.e., as the income of people in the reference group increases, the SWB of the focal individual decreases. This conforms to previous findings regarding the effect of reference income on SWB (e.g., Blanchflower and Oswald, 2004; Ferrer-i-Carbonell, 2005; Luttmer, 2005; McBride, 2001). Moreover, it is important to notice that the coefficient for income did not change meaningfully with the introduction of reference income. This indicates that reference income has an impact on SWB that is distinct from the absolute equivalent income level influence. Therefore, people derive some utility from their own absolute income level as well as from the extent to which their income level compares with the income of the reference group.

4.2.2. Income and consumption needs: Perceived income

Another way we can look at satisfaction with income is by realizing that there is a huge heterogeneity in consumption standards across households. This may result from inherited consumption habits (what McBride (2001) designates by internal or psychological norms), but also because of other consumption determinants, like health issues, living in an expensive city or children's private schooling, amongst others. While household income does not capture the adequateness of income to consumption standards, which we designate as perceived income, the ESS enables us to approach such problematic with the following question: "*Which of the descriptions on this card comes closest to how you feel about your household income nowadays?*" Answers are: "*1. living comfortably on present income, 2. coping on present income, 3. finding it difficult on present income, 4. finding it very difficult on present income*". The results in equation (6) show that the more people live well on their income, the greater their SWB. This is in line with Ferrer-i-Carbonell's (2005) contention that an explanation for the lower explanatory role of absolute income is that individuals consider the extent to which they are able to satisfy their consumption needs with their income. Somewhat surprisingly, however, we observe that this variable has quite strong effects on SWB, having the highest t-statistics across estimations.

Importantly, inclusion in the equation of the perceived income had a strong impact on the coefficients of many of the variables in the equation, namely marital status, employment status,

[10] Age and employment status were also used, as alternatives to education level. But the reference income measures obtained these ways were not statistically significant.

education, equivalent income, and past unemployment experiences. Regarding marital status, we observe that separated now have a smaller but still negative and significant coefficient, whereas divorced and widowed, which were previously negatively related to SWB, are now non-significant. This does not indicate that these marital statuses are less important or unimportant at all. On the contrary, the results suggest that the effects of these statuses are partially or totally mediated through consumption levels enabled by current income. That is, the effects of being separated, divorced, or widowed are now smaller and even insignificant because consumption standards reflect a possible drop in incomes following the taking place of such marital events. This result is consistent with set-point theories, according to which most life events only affect SWB temporarily (Suh, Diener, and Fujita, 1996). People will then tend to return to previous SWB levels, with some suggesting that long-term SWB levels depend to a great extent from personality traits (Lykken and Tellegen, 1996). Thus, our results suggest, to some extent, that the remaining negative effects of widowing and divorcing on SWB are mostly related to the negative impact of such events on living standards, rather than psychological issues (e.g., the loss of someone important) related to such events. The results are also coherent with news that credit defaults in mortgage loans are frequently related with life events such as divorce.

Relative to employment status, we observe that the negative coefficient for unemployment and sickness are now substantially lower but still highly significant. As above, a possible explanation for these attenuated effects relates to the mediating role exercised by consumption levels associated with present income. These levels are likely to become lower for individuals in such situations, given the ensuing drop in incomes, and this possibly explains the loss of significance for such occupational statuses. Surprisingly, the coefficients for retired and homecare, which were initially insignificant, now become significant with a positive effect. This means that, after controlling for income adequacy to consumption needs, retiring and homecare contribute positively to SWB. A possible explanation relates to the calmer life people in these occupations live, away from work and its related stress.

In addition, consideration of perceived income produces a marked decrease in the education coefficients, which still are significant. Once again, these effects should result from the mediating role exercised by the extent to which individuals are satisfied with consumption levels enabled by current income. This same effect should explain the sharp decrease in significance associated with the variable equivalent income. Thus, these results indicate that it is not so much the absolute income per si that it is important, but the extent to which it enables individuals to achieve desired consumption levels. This result is generally in line with the literature, which points to a relatively weak effect of absolute income on SWB.

Table 2: Life satisfaction equations and income related variables

	(4)		(5)		(6)		(7)	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Gender	0.0827	10.29	0.0818	10.17	0.0904	11.23	0.0995	12.34
Age	-0.0328	21.91	-0.0325	-21.76	-0.0282	-18.79	-0.0286	-19.06
Age squared×1000	0.3254	21.40	0.3223	21.18	0.2673	17.52	0.2713	17.78
Marital status								
Married/l.w. partner	0.2603	22.02	0.2598	21.97	0.2464	20.82	0.2429	20.51
Separated	-0.2688	-7.73	-0.2678	-7.70	-0.1706	-4.90	-0.1722	-4.94
Divorced	-0.0823	-4.33	-0.0822	-4.32	0.0070	0.37	0.0062	0.32
Widowed	-0.0830	-4.34	-0.0839	-4.38	-0.0212	-1.10	-0.0225	-1.17
Employment status								
Unemployed	-0.3442	-17.10	-0.3432	-17.05	-0.1987	-9.79	-0.1887	-9.29
Retired	0.0014	0.09	0.0040	0.28	0.0447	3.03	0.0515	3.49
Homecare	0.0027	0.19	0.0022	0.16	0.0504	3.56	0.0528	3.73
In education	0.1000	5.27	0.1029	5.42	0.1149	6.05	0.1132	5.95
Sick	-0.5442	-20.90	-0.5447	-20.92	-0.3894	-14.88	-0.3742	-14.29
Other	-0.0819	-2.26	-0.0810	-2.23	-0.0136	-0.37	-0.0050	-0.14
Education								
Lower secondary	0.1070	7.99	0.1248	9.11	0.0457	3.33	0.0013	2.86
Upper secondary	0.1409	10.64	0.1859	12.28	0.0724	4.87	0.0607	3.98
Post sec. (non tert.)	0.1891	9.69	0.2659	11.48	0.1361	5.85	0.1183	5.08
1st stage tertiary	0.1628	10.83	0.2786	11.56	0.1234	5.09	0.1078	4.44
2nd stage tertiary	0.1679	7.22	0.3202	9.42	0.1601	4.69	0.1442	4.23
Domicile								
Suburbs	0.0142	1.02	0.0155	1.11	0.0055	0.40	0.0013	0.09
Town	0.0280	2.40	0.0288	2.47	0.0176	1.51	0.0149	1.27
Village	0.0896	7.59	0.0915	7.75	0.0756	6.39	0.0733	6.20
Farm	0.1480	8.29	0.1496	8.37	0.1316	7.36	0.1217	6.80
Income$\times 10^3$	0.0731	20.49	0.0770	21.24	0.0102	2.67	0.0032	0.83
Past unemployment								
Short	-0.1065	-8.28	-0.1063	-8.27	-0.0850	-6.60	-0.0783	-6.07
Long	-0.2313	-15.18	-0.2305	-15.12	-0.1641	-10.73	-0.1509	-9.86
Last 5 years	-0.1133	-7.07	-0.1130	-7.05	-0.0576	-3.58	-0.0548	-3.41
Reference income$\times 10^3$			0.1245	-6.14	-0.1116	-5.50	-0.1156	-5.69
Perceived income								
Difficult on pres. inc.					0.3975	19.85	0.3856	19.11
Coping on pres. inc.					0.8602	44.24	0.8213	41.45
Comfort. on pres. inc.					1.2062	57.51	1.1322	52.87
Borrowing								
Quite difficulty							0.0344	2.81
Neither easy nor diff.							0.0967	7.43
Quite easy							0.1923	14.78
Very easy							0.3156	19.73
Number observations	74478		74478		74478		74478	
LR	15910.16		15947.92		20919.26		21443.59	
LL	-143426.78		143407.79		-140922.23		-140660.07	
Pseudo R^2	0.0525		0.0527		0.0691		0.0708	

Estimations include round and country dummies, which are not reported. Cutpoints are also omitted, for parsimony reasons.

In line with the above findings, the inclusion of perceived income also decreases the coefficients regarding previous unemployment experiences. This is particularly so for unemployment periods over 12 months and for unemployment periods lasting more than three months that have taken place in the last five years. A reason for the more pronounced reduction in the coefficient of these two variables is that longer unemployment periods, and unemployment periods that are more recent should have a much higher impact on the current economic situation and, therefore, on satisfaction with income suitability to finance consumption needs.

Somewhat surprisingly, the coefficient for reference income did not change meaningfully. This indicates that, and against the above findings, the effect of reference income is not mediated by perceived income. Thus, individuals obtain utility from how their income compares with that of others, utility which is not much affected by their income adequacy to meet consumption needs.

4.2.3. Borrowing constraints

Still on the vein of exploring new ways of capturing the financial determinants of SWB, it must be recognized that borrowing constraints may also matter, because by borrowing against future income people can alleviate the difficulties of balancing current low incomes against expenditures and smooth their living standards. Another way that can be used to motivate this suspicion is by realising that people not only care about income but about what they can do with it as well. Hsee et al (2008) distinguish three types of happiness: from money, from acquiring an item and from consuming an item. If acquiring and consuming in themselves provide well-being, then anything that interferes in that process is likely to affect SWB. Indeed, that is the case of borrowing constraints that affect consumption, as supported by a large volume of evidence suggesting the existence of borrowing constraints on consumers (e.g., Bayoumi, 1993; Campbell and Mankiw, 1990; Sarno and Taylor, 1998; Jappelli, Pischke and Souleles, 1998). This line of thought indicates that when agents are not entirely free to borrow against future income, their well-being is penalized. However, no formal studies have yet related borrowing constraints with SWB.

In equation (7) we have thus included a variable that attempts to measure borrowing constraints. This variable is based on the following ESS question. "*If for any reason you were in serious financial difficulties and had to borrow money to make ends meet, how difficult or easy would that be?*" The possible answers are: *1-very difficulty, 2-quite difficulty, 3-neither easy nor difficulty, 4-quite easy, 5-very easy*. The results indicate that perceived access to borrowing is positively related with SWB, with such effects evolving monotonically. The results further indicate that the introduction of this variable did not change meaningfully the coefficients of the remaining variables, except income, which looses its significance. This indicates that the effect of absolute income is also mediated by borrowing constraints. Thus, absolute income matters to the extent to which it contributes to meet desired consumption standards and ease borrowing constraints.

4.3. Further results

In order to enlighten the effects of the unemployment experiences and financial effects considered in this paper, this section provides further results by disaggregating the sample according to whether individuals are currently employed or unemployed, as well as by perceived income and level of borrowing constraints. This enables the investigation of whether people with different employment status and financial situations respond to unemployment and financial issues differentially.

Employment status and SWB. Equations (8) and (9) break the sample according to whether individuals are unemployed or employed, respectively. The results denote that past unemployment experiences do not affect equally the SWB of these two groups of people. More specifically, for employed individuals, having been unemployed in the past for more than three or 12 months, and the recency of that unemployment spell, have a negative impact on SWB.

Moreover, the results of equation (8) indicate that for the currently unemployed, that suffer from such occupational status in the present, past unemployment experiences do not impact on SWB. Thus, the results suggest that the scarring hypothesis applies in particular for the currently employed individuals, for whom the difficulties associated with the present situation seem to prevail over past experiences. This may contribute to enlighten the contradictory findings in the literature concerning the effect of past unemployment experiences on the disutility derived from current employment status.

Moreover, we observe that although reactions to equivalent income are insignificant for both groups, reference income is highly significant for employed, but not significant for the unemployed (although the coefficient for the latter is significant at a 10% level). This conforms to evidence that reference income matters mostly for higher income individuals (e.g., McBride, 2001), and this is more likely to characterize the employed. In addition, we observe that the unemployed tend to present stronger (and positive) coefficients for borrowing possibilities. The unemployed are likely to suffer from a reduction in income, which is likely to affect living standards. In these circumstances, the availability of credit should contribute to attenuate the negative consequences of unemployment on income. For this reason, they should react in a much stronger way to borrowing constraints. The employed, on the contrary, because they have higher incomes, do not need so much to borrow to make ends meet, thus reacting less strongly to borrowing facilities.

We also disaggregated the sample according to past unemployment experiences (as to whether individuals have never been unemployed in the past for more than 3 months, have been unemployed from 3 to 12 months, or for more than 12 months and as to whether individuals have been unemployed for more than 3 months in the last 5 years or not). The results obtained, which are not reported, do not add new insights. Basically, we obtained that those with more negative past unemployment situations, when compared with those with less negative ones, tended to behave similarly to the currently unemployed versus employed individuals. The most notable result is that the more prolonged, and the more recent, the past unemployment experience, the less sensitive individuals are to current unemployment, which is in accordance with the adaptation effect observed in Table 1.

Table 3: Life satisfaction equations desegregations

	(8) Unemployed		(9) Employed		(10) Very diff/Diff. on present inc.		(11) Coping/Comfort. on present inc.		(12) Difficult to borrow		(13) Easy to borrow	
	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio	Coeff.	t-ratio
Gender	0.1241	3.45	0.0984	9.27	0.1028	5.64	0.0977	10.84	0.1101	8.65	0.0810	6.24
Age	-0.0422	-4.40	-0.0397	-12.70	-0.0354	-11.43	-0.0283	-16.41	-0.0280	-12.42	-0.0262	-10.19
Age squared×1000	0.4482	3.86	0.3904	10.85	0.3237	10.50	0.2738	15.52	0.2603	11.61	0.2578	9.70
Marital status												
Married/l.w. partner	0.2372	5.38	0.2873	19.19	0.1945	7.71	0.2616	19.47	0.2132	11.54	0.2944	15.14
Separated	-0.0682	-0.56	-0.1586	-3.61	-0.2115	-3.45	-0.2002	-4.70	-0.1911	-3.77	-0.1646	-2.79
Divorced	0.0192	0.29	0.0338	1.35	-0.0611	-1.71	0.0054	0.24	-0.0237	-0.84	0.0569	1.75
Widowed	0.1998	1.47	-0.0672	-1.57	-0.0763	-2.09	-0.0115	-0.51	-0.0418	-1.52	0.0405	1.16
Employment status	n/a	n/a	n/a	n/a								
Unemployed					-0.1828	5.80	-0.2276	-8.26	-0.1667	-6.17	-0.2775	-6.53
Retired					-0.0069	-0.22	0.0605	3.58	0.0315	1.43	0.0707	2.82
Homecare					-0.0485	-1.69	0.0717	4.38	0.0055	0.26	0.0667	2.69
In education					0.1719	4.04	0.1006	4.71	0.1407	4.55	0.0781	2.58
Sick					-0.3633	-8.63	-0.4332	-12.67	-0.3846	-10.77	-0.4396	-8.90
Other					-0.0762	-1.05	0.0113	0.27	-0.0696	-1.30	0.0149	0.24
Education												
Lower secondary	-0.0121	-0.20	0.0360	1.53	0.0469	1.74	0.0443	2.72	0.0271	1.38	0.0372	1.45
Upper secondary	0.0078	0.11	0.0245	1.01	0.0284	0.86	0.0820	4.65	0.0533	2.31	0.0327	1.22
Post sec. (non tert.)	0.1360	1.15	0.0874	2.67	0.0826	1.46	0.1477	5.68	0.1182	3.12	0.0530	1.40
1st stage tertiary	0.0590	0.48	0.0989	2.89	0.0789	1.31	0.1433	5.30	0.1088	2.76	0.0576	1.46
2nd stage tertiary	0.3537	1.96	0.1313	2.89	0.0565	0.62	0.1801	4.79	0.1595	2.73	0.0880	1.67
Domicile												
Suburbs	-0.1560	-2.41	-0.0075	-0.40	-0.0143	-0.46	0.0068	0.43	-0.0009	-0.04	-0.0014	-0.06
Town	-0.0848	-1.64	0.0198	1.23	-0.0020	-0.08	0.0233	1.74	0.0076	0.42	0.0278	1.42
Village	0.0806	1.52	0.0823	5.02	0.0945	3.93	0.0712	5.22	0.0847	4.77	0.0783	3.85
Farm	-0.0082	-0.09	0.1437	5.99	0.0928	2.01	0.1295	6.61	0.1402	4.54	0.1045	3.91
Income×10^3	-0.0095	-0.39	0.0053	1.09	0.0097	0.77	0.0038	9.78	0.0025	0.35	0.0091	1.71
Past unemployment												
Short	0.1044	1.34	-0.0752	-4.69	-0.0774	-2.72	-0.0973	-6.69	-0.0718	-3.54	-0.1090	-5.20
Long	0.0094	0.12	-0.1198	-5.76	-0.1830	-6.10	-0.1478	-8.21	-0.1580	-7.14	-0.1215	-4.42
Last 5 years	-0.0509	-0.77	-0.0821	-3.96	-0.0972	-3.18	-0.0483	-2.53	-0.0861	-3.68	-0.0300	-1.05
Reference income×10^3	-0.2182	-1.92	-0.1340	-5.02	0.0100	0.19	-0.1205	-5.41	-0.0678	-1.99	-0.1202	-3.83
Perceived income					n/a	n/a	n/a	n/a				
Difficult on pres. inc.	0.4559	8.70	0.3339	8.60					0.4094	17.92	0.2906	4.87
Coping on pres. inc.	0.8265	14.69	0.8026	21.32					0.8156	35.41	0.7399	13.13
Comfort. on pres. inc.	1.1095	13.63	1.1367	29.12					1.1542	41.68	1.0432	18.29
Borrowing									n/a	n/a	n/a	n/a
Quite difficulty	0.0833	1.81	0.0058	0.30	0.1929	9.65	-0.0377	-2.43				
Neither easy nor diff.	0.2636	4.75	0.0537	2.75	0.2552	9.75	0.0412	2.66				
Quite easy	0.1222	2.02	0.1861	9.68	0.2875	9.81	0.1810	11.85				
Very easy	0.4215	4.71	0.2945	13.05	0.3463	6.85	0.3524	19.74				
Number observations	3484		39406		15305		59173		30318		28615	
LR	1099.01		10061.24		2527.42		9271.37		8598.26		4949.97	
LL	-7451.64		-71511.81		-33410.736		-107206.86		-61393.065		-49769.412	
Pseudo R^2	0.0687		0.0657		0.0364		0.0414		0.0654		0.0474	

Estimations include round and country dummies, which are not reported. Cutpoints are also omitted, for parsimony reasons.

Perceived income and SWB. Given the four response options associated with the question regarding *feelings about present income*, for parsimony reasons we have combined the two lowest categories (finding it difficult and very difficult to live on present income), in equation (10), and the two highest (coping and living comfortable on present income), in equation (11). The results indicate that these two groups differ strongly on their reaction to current equivalent income, whose coefficient is non-significant for those who find it difficult to live with their current income, and highly significant for the remaining group. Thus, after controlling for the satisfaction with consumption levels enabled by current income, absolute/equivalent income provides further utility. This result contrasts, to some extent, with past evidence indicating that for higher income levels, income becomes less significant. Based on past evidence, Dolan, Peasgood, and White (2008:98), for example, conclude that "additional income for those who are not at low levels of income is unlikely to increase SWB in the long run if the additional income serves to increase expectations of necessary income." However, our results indicate that, after controlling for people's consumption aspirations, which *feelings about present income* capture to some extent, absolute income does indeed contribute to happiness. Thus, for some people, money does seem to buy happiness. As to reference income, only those that are coping or living comfortably on present income seem to care about it, which conforms to the results obtained by McBride (2001), which show that relative income effects are stronger at high income levels.

Regarding past unemployment experiences, we observe meaningless differences between the two subsamples as to the duration, but significant ones as to the recency of the past unemployment spell. More specifically, those reporting to finding it difficult or very difficult to live on present income react more negatively to the fact that their past unemployment experience has occurred in the last 5 years. This probably results from recent unemployment experiences producing a higher impact on current income and consumption levels.

As to borrowing constraints, we observe that those reporting to live difficulty or very difficulty on present income tend to react more strongly to borrowing capacity. As with the unemployed, this suggests that those living with more difficulties may look at borrowing facilities as a way of overcoming current and future income limitations.

Borrowing constraints and SWB. Given the 5 response options associated with borrowing constraints, again for parsimony reasons we have aggregated the two lowest (difficult and very difficult) and the two highest (quite easy and very easy) categories. The results indicate that those finding it easy to obtain credit react more adversely to current unemployment. Individuals with more resources, be it income or borrowing facilities, probably are the ones who have most to loose from being currently unemployed, a status that could damage, for example, their future borrowing capacity. In addition, there are also non-pecuniary costs, which are likely to be higher for those with access to credit.

As to past unemployment experiences, we do not find meaningful differences between the two groups of people for unemployment spells over three months. However, we do find that unemployment spells occurring in the last five years do leave a significant negative mark on those reporting to finding it difficult to borrow money. This contrasts with those finding it easy to borrow, for whom the recency of such unemployment spells are not significant. An explanation for these findings is that those with recent unemployment experiences have probably suffered economically in a way that limited their borrowing capacity. This possibly increased their sensitivity to recent unemployment spells.

Although at a marginal significance level of 8.7%, absolute income influences the well-being of those finding it easier to borrow money, but it does not affect those with borrowing difficulties. Reference income enters negatively and significantly in the equations for both subsamples. However, those pertaining to the highest categories of the borrowing facilities care about it more than the others. If we believe that our borrowing variable is a good proxy for financial endowment, we conclude that the results reinforces those obtained by McBride (2001), that people with higher financial resources are more likely to compare their own income with their reference group's income.

Finally, we look at the perceived income. Interestingly, we observe that the well-being of the borrowing constrained group reacts more to such a definition of income than those that have

more freedom to borrow. We sustained that one channel of transmission from borrowing facilities to well-being is through consumption. Therefore, there is an analogy between our result and that obtained by the extended literature that searches for differences in consumption behaviour between borrowing constrained and unconstrained individuals (e.g. Campbell and Mankiw, 1990; Bayoumi, 1993; Sarno and Taylor, 1998; Jappelli, Pischke and Souleles, 1998). If individuals face sudden financial difficulties and are not able to borrow in order to smooth consumption, then their consumption will react to current income, contrary to what the permanent income hypothesis contends. Accordingly, this literature has found an augmented sensitivity of consumption to current income for those constrained to borrow.

5. CONCLUSION

Using data from the European Social Survey, this paper investigated the impact on SWB of two important interrelated factors: past unemployment experiences and financial issues. Regarding unemployment, the main findings support previous contentions of an adaptation effect. Accordingly, past unemployment experiences mitigate the disutility individuals obtain from current unemployment. However, the results obtained in this study suggest that there may be a threshold above which the adaptation takes place. Moreover, the results support the scarring hypothesis associated with past unemployment (e.g., Clark, Georgellis, and Sanfey, 2001). Notwithstanding, we observed that the scarring effect only affected the employed individuals. In addition, the results indicate that individuals are not indifferent to the duration of past unemployment experiences. More specifically, longer unemployment spells in the past produce a much stronger scar on individuals than shorter ones. Finally, the results indicate that SWB, apart from being influenced by the duration of past unemployment experiences, is also affected by the recency of the past unemployment spell.

Taken together, the results suggest that to improve the well being of individuals it is particularly important to minimize long term unemployment, as shorter unemployment experiences leave a smaller scar and do not produce the adaptation effect, which could reduce an individual's future efforts to return to the labour market. This follows contentions that unemployment's psychological effects may alter an individual's search strategies and preferences for work (Darity and Goldsmith, 1996).

Regarding financial issues, the major findings concern the relevance of borrowing constraints in shaping SWB. More specifically, the results indicate that SWB depends positively on perceived access to credit. In addition, we observed that perceived income, i.e., the extent to which people are coping with their income, had a quite strong impact on SWB. Together, borrowing constraints and perceived income contributed to the non significance of equivalent income. Thus, the results suggest that the relatively low importance past studies have found for absolute income is due, at least in part, to the mediating effect of other variables, in this case perceived borrowing constraints and income.

REFERENCES

BERTRAND, M. AND MULLAINATHAN, S. (2001), "Do people mean what they say? Implications for subjective survey data", *American Economic Review, AEA Papers and Proceedings*, pp. 67-72.

BAYOUMI, T. (1993), "Financial deregulation and Consumption in the United Kingdom", *The Review of Economics and Statistics*, Vol. 75(3), pp. 536-539.

BLACHFLOWER, D. G. AND OSWALD, A. (2004), "Well-being over Time in Britain and the USA", *Journal of Public Economics*, Vol. 88, pp. 1359-1386.

CAMPBELL J. AND MANKIW G. (1990), "Permanent Income, Current Income and Consumption", *Journal of Business and Economic Statistics*, Vol. 8, pp.265-279.

CLARK, A.E., DIENER, E. GEORGELLIS, Y. AND LUCAS, R.E. (2008), "Lags and Leads in Life Satisfaction: A Test of the Baseline Hypothesis", *The Economic Journal*, Vol. 118, June, pp. 222-243.

CLARK, A.E., FRIJTERS, P. AND SHIELDS, M. (2007), "A survey of the income happiness gradient", *Journal of Economic Literature*, Vol. 46 (1), pp. 95-144.

CLARK, A.E., GEORGELLIS, Y. AND SANFEY, P. (2001), "Scarring: The Psychological Impact of Past Unemployment", *Economica*, Vol.68, pp. 221-241.

CLARK. A.E. AND OSWALD, A. (1994), "Unhappiness and Unemployment", *The Economic Journal*, Vol. 104, May, pp. 648-659.
CLARK, A.E. AND LELKES, O. (2009), "Let us pray: religious interactions in life satisfaction", Working Paper Nº 2009-01, Paris School of Economics.
DARITY, W. JR AND GOLDSMITH, A. H. (1996), "Social psychology, unemployment, and macro-economics", *Journal of Economic Perspectives*, Vol. 10, pp. 121-40.
DOLAN, P., PEASGOOD, T. WHITE, M., (2008), "Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being", *Journal of Economic Psychology*, Vol. 29, pp. 94-122.
EASTERLIN, R. (1974), "Does economic growth improve the human lot? Some empirical evidence" in David, P.A., Reder, M.W. (Eds.), Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz. Academic Press, New York, pp. 89-125.
EASTERLIN, R.A. (1995), "Will raising the incomes of all increase the happiness of all?", *Journal of Economic Behavior and Organization*, Vol. 27, pp. 35-47.
FERRER-I-CARBONELL, A. (2005), "Income and well-being: an empirical analysis of the comparison income effect", *Journal of Public Economics*, Vol. 89, pp. 997-1019.
FREY, B. S. AND STUTZER, A. (2002), "What Can Economists Learn from Happiness Research?", *Journal of Economic Literature*, Vol. XL, June, pp. 402-435.
FREY, B. S. AND STUTZER, A. (2000), "Happiness prospers in democracy", *Journal of Happiness Studies*, Vol. 1(1), pp. 918-938.
HAYO, B. AND SEIFERT, W. (2003), "Subjective Economic Well-Being In Eastern Europe", *Journal of Economic Psychology*, Vol. 24, pp. 329-348.
HSEE, C.K., YANG, Y., LI, N. AND SHEN, L. (2008), "Wealth, Warmth and Well-Being: Whether Happiness is relative or Absolute Depends on Whether it is about Money, acquisition or Consumption", *Journal of Marketing Research*, Vol. 45, forthcoming.
JAPPELLI, T., PISCHKE, J-S. AND SOULELES, N.S. (1998), 'Testing for liquidity constraints in Euler equations with complementary data sources', *Review of Economics and Statistics*, Vol. 80(2), pp. 251-62.
KAHNEMAN, D, AND KRUEGER, A. B. (2006), Developments in the Measurement of Subjective Well-Being", *Journal of Economic Perspectives*, Vol.20(1), Winter, pp. 3-24.
KAHNEMAN, D, AND KRUEGER, A. B., SCHKADE, D. (2006), "Toward National Well-Being Accounts", *American Economic Review, Papers and Proceedings*, Vol. 94(2), pp.429-434.
LUCAS, R. E., CLARK, A. E. GEORGELLIS, Y. AND DIENER, E. (2003), "Re-examining Adaptation and Setpoint Model of Happiness: Reactions to Changes in Marital Status", *Journal of Personality and Social Psychology*, Vol. 84(3), March, pp.527-39.
LUCAS, R. E., CLARK, A. E. GEORGELLIS, Y. AND DIENER, E. (2004), "Unemployment alters the set-point for life satisfaction", *Psychological Science*, Vol. 15, pp. 8-13.
LUTTMER, E.F.P. (2005), "Neighbors as Negatives: Relative Earnings and Well-Being", *The Quarterly Journal of Economics*, Vol. 20(3), August, pp. 963-1002.
LYKKEN, D. AND TELLEGEN, A. (1996), "Happiness is a stochastic phenomenon", *Psychological Science*, Vol. 7, pp.186-189.
OSWALD, A. J. (1997), "Happiness and Economic Performance", *Economic Journal*, Vol. 107, November, pp. 1815-1831.

OSWALD, A.J. AND POWDTHAVEE, N. (2008), "Does Happiness Adapt? A Longitudinal Study of Disability with Implications for Economists and Judges", *Journal of Public Economics*, Vol. 92(5-6), pp. 1061-1077.
MEIER, S. AND STUTZER, A. (2008), "Is volunteering rewarding in itself?", *Economica*, Vol. 75(297), pp. 39-59.
MCBRIDE, M. (2001), "Relative-income effects on subjective well-being in the cross-section", *Journal of Economic Behavior and Organization*, Vol. 45, pp. 251-278.
PODSAKOFF, PHILIP, M., MACKENZIE, SCOTT B., LEE, JEONG-YEON, AND PODSAKOFF, NATHAN P. (2003), "Common Method Biases in behavioural research: A critical Review of the literature and recommended remedies", *Journal of Applied Psychology*, Vol. 88 (5), 879-903.
SARNO L. AND TAYLOR M. P. (1998), "Real Interest Rates, Liquidity Constraints and Financial Deregulation: Private Consumption Behaviour in the UK", *Journal of Macroeconomics*, Vol. 20 (2), pp. 221-242.
SENIK, C. (2005), "Income Distributions and Well-being: what can we learn from subjective data?", *Journal of Economic Surveys*, Vol. 19(1), pp. 43-63.
SUH, E., DIENER, E. AND FUJITA, F. (1996), "Events and subjective well-being: Only recent events matter", *Journal of Personality and Social Psychology*, Vol.70, pp.1091-1102.
VAN PRAAG, BM.S., FRIJTERS, P. AND FERRER-I-CARBONELL, A. (2003), "The anatomy of subjective well-being", *Journal of Economic Behaviour & Organization*, Vol. 51, pp. 29-49.
WELSCH, (2006), Environment and happiness: Valuation of air pollution using life satisfaction data", *Ecological Economics*, Vol. 58, pp. 801-813.
WINKELMANN, L. AND WINKELMANN, R. (1998), "Why Are the Unemployed So Unhappy? Evidence from Panel Data", *Economica*, Vol. 65, pp.1-15.

GRADUATE OVEREDUCATION IN SPAIN: A PROBLEM OR A SOLUTION?

DANIELA CASTAÑEDA
E-mail: daniela.castaneda@unican.es
NOELIA GONZÁLEZ-PRIETO
MARTA PASCUAL
Department of Economics
University of Cantabria

Resumen

Spain is suffering an important economic slowdown which affects significantly to unemployment rises. For this reason, many individuals are considering to increase their educational level. However, this decision can lead to get worse labour market indicators. In this paper, we analyse the problem of overeducation and its effect on labour market using the data contained in the European Community Household Panel (ECHP) and other educational indicators. In Spain, many workers are accepting jobs for which the required level of education is lower than the level they have attained. Thus, less-educated workers have many problems to find a job. As a result, these workers suffer a high and persistent unemployment status. Econometrically, we use different regression models.

Key Words: Overeducation, ECHP, panel data, probit models.
Área temática: Economía Social y Discapacidad.

1. INTRODUCTION

According to the Spanish Labour Force Survey (EPA) 2004, over 25% of the illiterate population of working age is unemployed compared to only 8% of graduates in higher education. This may be a reason why the number of university students in Spain is higher than in other countries. Moreover, in times of crisis like now, where unemployment in Spain has risen from 8.6% in January 2007 to 13.91% at the end of 2008, a settlement arising from the unemployed to improve their future work is to resume its academic training. Mostly all Spanish people study to work on a civil serVant as this job is secured forever if the exam is passed. The most striking is that, in general, those with more formal education prepare the exam for jobs that require a lower educational level, probably due to fact that they have a better chance of achieving the exam. Another striking fact is that the number of students enrolled in the UNED (Distance University Education) in the last year has increased by 12%. This may have two consequences, firstly to increase competitiveness to keep or find a new job and on the other hand, increases the risk that graduates perform a job they are overeducated.

The first question we ask is what is meant by overeducation. Taking into account previous studies it is assumed that overeducation is the fact of having a level that exceeds the educational requirements necessary to perform a job. Something more concrete, overeducation occurs when the educational effort does not receive sufficient social or economic compensation in the labour market.

The analysis of overeducation based on the premise that the labour market and education systems are designed to meet each other, according to theoretical models (García-Espejo (1998), Rubinson and Browne (1994), Gobernado (2003)). For political, social, economic, ... this adjustment is carried out in different countries.

Today most of the countries are involved in a serious economic and financial crisis. As a result of this crisis, there exists an important increase of unemployment rates, especially in

Spain. Unemployed workers try to improve their training in order to access to a new and better job. Another possibility, given the economic problems even if it is of a lower category is to accept a job, thus resulting in overeducation. Besides this problem that occurs with workers overeducated, this also has an impact on other with lower education as they will have more difficulties to access jobs.

Businessmen, also suffer the consequences of overeducation. Firstly, they have a greater number of candidates to fill vacancies so they increase the requirements to the candidates in order to access the job. Secondly, overeducation may have effects on productivity, there are studies that indicate that productivity increases and others conclude that the productivity decreases.

The structure of the paper is as follow. Section 2 reviews the empirical studies related with overeducation and labour market. In Section 3, the methodological decisions adopted are defined. Section 4 and 5 describe, respectively, the data we have used and the empirical results and finally, Section 6 presents the main conclusions.

2. OVEREDUCATION AND LABOUR MARKET: A REVIEW

The study of overeducation in the United States began when Jolger and Nam (1964) studied the increase in educational requirements that were required to fill jobs in the decades of the 40s and 50s which was due mainly to increasing the educational level of the employed population and not the need for such studies to do the job. Although indepth study of the subject was from the seventies. Freeman (1976) analyzed the overeducation and infraeducation using microdata (United States Panel Study of Income Dynamics). Following his example, Duncan and Hoffman (1981) showed that overeducation is a widespread problem in the labour market of the United States.

Büchel and Van Ham (2003) and Büchel and Battu (2003) highlighted the importance of spatial mobility in the labour market by region as a explanatory factor of global education. They obtained empirical evidence that variables related to regional spatial distribution of jobs and the size of the local labour market affect the overeducation. The possibility of access to wider geographical areas decreases the probability of overeducation for all individuals.

Simpson (1992), states that an individual looking for a job in a local labour market has three options when they don´t obtain appropriate jobs with their characteristics. The first option is choosing not to work and keep looking, this means, to become an unemployed person. A second option is to accept a job which the requirements are lower than their level of education, which result is overeducation. The third option is to accept a job beyond the local labour market, travelling daily with longer distances than desired.

Büchel and Van Ham (2003) discuss the extent to which employment opportunities in the regional labour market (unemployment rate) and personal mobility (availability of vehicle and length of travel to work) affect the probability of overeducation. Frank (1978) was the first author in studying the differential overeducation, which is the difference in overeducation between men and women, as a consequence of geographical restrictions on married women when it comes to finding work. In his study, he concludes that when a couple tries to maximize revenue while minimizing the overeducation of the one who has higher level of study (usually the husband) and for that we will go where needed. The other member of the couple looks for a job where they are established and this limits their options to find job according to her level of education. Therefore, as Frank states, the, overeducation and the size of local markets are the cause of the difference between the income levels between men and women. In both cases, geographical restrictions played a key aspect in the conclusions, but the main differences are that Büchel and Van Ham extend the effect to all workers without distinction between men and women, while Frank highlights this difference. Besides the possibility of expanding the space of searching for a job through commuting. The results obtained for the German labour market show the variables related to the regional spatial distribution of jobs explaining the phenomenon of overeducation. Also, greater mobility of individuals (ie: having a car or travelling long

distances) allows expanding the effective size of the labor market and decreases the probability of overeducation.

For the Spanish case, Sanromá and Ramos (2004) concluded that the higher the educational level the greater the probability of overeducation. While this probability is not affected by the experience of an individual in the labour market, as there are evidences that in the Spanish labour market an individual could have followed a different strategy such as not accepting the job and continue searching for a better job.

This study along with work due to the current crisis in which we are involved this can cause social change.

The high unemployment rate plays a crucial role in overeducation, since they lack of ability to access to a job of your level of education and they can consciously accept a job for which is they are overeducated. High unemployment rate leads to fierce competition for scarce jobs available, leading to young people who seek to improve their education by extending the years of study and businesses to tighten their criteria.

Morgan and Scott (1987) and Fields (1995) concluded that any newly-qualified initially presents greater difficulty in finding a job in its class education, with the same conditions, a person engaged in the labor market. This may cause difficulties in finding a job acording with their level of studies if accepting other in which the level of education needed is lower or facing unemployment at the end of the studies.

Verdugo and Verdugo (1989) studied the relationship between overeducation and wages, and concluded that having a high educational level is not matched by high wages, relative wages of university graduates in the year of study was lower than in those with the same degree of a previous generation.

In Portugal, Kiker, Santos and Oliveira (1997) concluded that overeducation workers earn more than their peers with exactly the years of study appropriate to the job. However, these overeducated workers earn less than persons of the same level of education but to work at their level of education.

Alba-Ramirez and San Segundo (1995) conducted a study on the economic benefits that accrue to workers according to their educational level. They concluded that high school graduates get the best compensation in the private sector while the college graduates earn in the public sector. On the other hand the returns to college education are higher among women than among men.

Thurow presents the argument that employers hire workers with better education levels so that lower costs in training. In the same way, Forgeot and Gautier (1997) argue that recruit overeducated people with the hope that they have a higher productivity. Büchel (2000) concludes that in Germany, the overeducated workers tend to be more productive than those with a job according with their educational level.

However Tsang (1987) and Franz (1991) conclude that a overeducated person has a labour dissatisfaction implies a reduction in productivity, increased absenteeism and health problems, and he will try to find a new job suited to their educational level. As a result, companies will appreciate the possibility of employing or not hiring these people.

The reason for the inconsistency in the conclusions of these authors may result in a difference in the approach of overeducation.

3. MEASURING OVEREDUCATION

There are three different methods to measure overeducation in particular and the mismatch of education in general. The presentation of different methods can be found in Hartog (2000), White (1997) and Madrigal (2002).

The objective method is based on the comparison of the level of formal education that the worker has, with the one is deemed necessary in his job. Some authors who applied this method are Rumberger (1987), Hartog and Oosterbeek (1988), Kiker and Santos (1991), Kiker et al. (1997) and Garcia-Montalvo (1995).

The subjective method (direct and indirect), which is based on the workers themselves who reveal whether their level of formal education is adapted to use; authors who have used this method in their work are Duncan and Hoffman (1981) and Sicherman (1991) for United States, and Hartog Oosteerbeck (1988) for Holland and Alba (1993) and García Serrano and Malo (1996) for Spain.

The statistical method has two versions, the first of them is the method of interval on the average worker feels that this has overeducation when a number of years of formal education increased by more than one standard deviation of the average years of education workers in their occupation. Authors who have used this method in their work are Verdugo and Verdugo (1989), Cohn and Kahn (1995), Garcia-Montalvo (1995). The second version is the modal method which assumes that the level of education that is considered appropriate for a particular occupation is the mode criterion of the educational levels of individuals who develop this occupation, i.e., level of education prevailing among workers this occupation. When the educational attainment of worker is equal the educational level of the mode workers, it is considered that the worker is adequately educated. Authors who have used this method are Kiker et al. (1997), Mendes de Olivera et al. (2000).

4. DATA DESCRIPTION: THE EUROPEAN COMMUNITY HOUSEHOLD PANEL

Participation in higher education in the European Union has increased considerably for the last years. The proportion of population attending tertiary education has varied considerably among countries (see Table 1). Thus, in 2005, Finland has the highest value (35%) followed by Denmark (34%), Belgium (31%), Netherlands (30%), Sweden (30%), United Kingdom (30%) and Spain (28%). On the other side, the lowest rates are in Italy (12%), Portugal (13%) and Austria (18%).

However, in Spain the proportion of population with education attainment less than upper secondary level is one of the highest levels in the European Unión countries (see Table 3). The highest level is found in Portugal (78%) follows by Spain (51%) and Italy (50%). As consequence, Spain has a low percentage of population with attainment upper secondary level (see Table 2). Also, Spain is one of the European Union countries with higher level of unemployment as percentage of labour force.

TABLE 1. Attainment tertiary level (% population w/ education level)

	1997	1998	1999	2000	2001	2002	2003	2004	2005
Austria	11	11	11	14	14	14	15	18	18
Belgium	25	25	27	27	28	28	29	30	31
Denmark		25	27	26	28	30	32	33	34
Finland	29	30	31	32	32	33	33	34	35
France	20	21	21	22	23	24	24	24	25
Germany	23	23	23	23	23	23	24	25	25
Greece	16	17	17	18	18	19	19	21	21
Ireland	23	21	20	19	24	25	26	28	29
Italy		9	9	9	10	10	10	12	12
Luxembourg			18	18	18	19	14	24	27
Netherlands		24	23	23	23	25	28	30	30
Portugal		8	9	9	9	9	11	13	13
Spain	19	20	21	23	24	24	25	26	28
Sweden	28	28	29	30	32	33	33	35	30
United Kingdom	23	24	25	26	26	27	28	29	30

Source: OECD Health Data (2008).

TABLE 2. Attainment upper secondary level (% population w/ education level)

	1997	1998	1999	2000	2001	2002	2003	2004	2005
Austria	63	63	64	62	63	63	64	62	63
Belgium	30	31	31	31	32	33	33	34	35
Denmark		53	53	52	52	52	49	48	47
Finland	39	39	40	41	42	42	43	43	44
France	39	40	40	41	41	41	41	41	41
Germany	61	61	58	58	59	60	59	59	59
Greece	29	29	31	31	32	33	34	35	36
Ireland	27	30	35	28	32	35	35	35	35
Italy		32	33	33	33	34	38	37	38
Luxembourg			38	38	35	43	45	40	39
Netherlands		40	32	41	42	43	42	41	42
Portugal		10	10	11	11	11	12	13	14
Spain	13	13	14	16	16	17	18	19	21
Sweden	48	48	48	47	49	49	49	48	54
United Kingdom	36	36	37	37	37	37	37	36	37

Source: OECD Health Data (2008).

TABLE 3. Attainment less than upper secondary level (% population w/ education level)

	1997	1998	1999	2000	2001	2002	2003	2004	2005
Austria	27	26	25	24	23	22	21	20	19
Belgium	45	43	43	41	41	39	38	36	34
Denmark		21	20	21	19	19	19	19	19
Finland	32	31	28	27	26	25	24	22	21
France	41	39	38	37	36	35	35	34	33
Germany	17	16	19	18	17	17	17	16	17
Greece	56	54	52	51	50	48	47	44	43
Ireland	50	49	45	54	45	40	38	37	35
Italy		59	58	58	57	56	52	51	50
Luxembourg			44	44	47	38	41	37	34
Netherlands		36	45	35	35	32	31	29	28
Portugal		82	81	81	80	80	77	75	74
Spain	69	67	65	62	60	59	57	55	51
Sweden	25	24	23	22	19	18	18	17	16
United Kingdom	41	40	38	37	37	36	35	35	33

Source: OECD Health Data (2008).

TABLE 4. Total unemployment (% of labour force)

	1997	1998	1999	2000	2001	2002	2003	2004	2005
Austria	4,2	4,2	3,8	3,5	3,6	4	4,3	5	5,2
Belgium	12,4	11,6	8,5						
Denmark	6,1	5,4	5,5	4,6	4,8	4,7	5,5	5,7	5
Finland	12,5	11,3	10,1	9,7	9	9	9	8,7	8,3
France	12,1	11,5	10,8	9,5	8,7	9	9,8	10	9,9
Germany	9,8	9,2	8,4	7,8	7,8	8,6	9,3	10,3	11,1
Greece	10,2	10,8	11,9	11,2	10,4	9,9	9,3	10,2	9,6
Ireland	10,3	7,8	5,7	4,3	3,7	4,2	4,4	4,4	4,2
Italy	11,6	11,7	11,3	10,5	9,5	9	8,7	8	7,7
Luxembourg	2,7	2,3	2,1	1,8	1,7	2	2,5	2,8	3,1
Netherlands	5,5	4,3	3,5	2,9	2,7	3,2			
Portugal	6,7	4,9	4,4	3,9	4	5,1	6,4	6,7	7,6
Spain	20,4	18,5	15,5	13,8	10,5	11,5	11,5	11	9,2
Sweden	10,2	8,5	7,2	5,9	5,1	5,2	5,8	6,6	7,8
United Kingdom	7	6,1	6	5,5	4,7	5,1	4,8	4,6	4,6

Source: OECD Health Data (2008).

Our focus is the impact of overeducation on the earnings of individuals who worked in 2001 as full-time workers in Spain. Thus, we excluded from the sample those employed part-time, unpaid family workers and unemployed individuals. We follow the model proposed by Verdugo and Verdugo (1989). Thus, a worker is overeducated if his/her education is more than one standard deviation above the average for his census occupation code in 2001. On the other hand, a worker is undereducated if his/her education is more than one standard deviation below the average for his/her census occupation code. So, adequately workers are those within ± 1 standard deviation of the average education for their census occupation code.

The data used in this study come from the European Community Household Panel (ECHP). This survey contains data on individuals and households for the European Union countries with eight waves available (1994 to 2001). The main advantage is that information is homogeneous among countries since the questionnaire is similar across them. This source of data is coordinated by the Statistical Office of the European Communities (EUROSTAT). Also, this survey includes rich new information about income, education, employment, health, etc.

This representative survey of households of different European Union countries was carried out for the first time in 1994 and 60500 households were interviewed (approximately 170000 individuals). The income measure is disposable (after tax) individual income. However the reference period of income is the year prior to interview. The interviews corresponding to the first eight waves of the ECHP were performed from 1994 to 2001, meaning that the corresponding incomes refer, respectively, from 1993 to 2000 (eight years). All monetary amounts in the data are expressed in national currency units. However, comparisons among countries can be made in equivalent units taking into account differences in the national currency purchasing power.

5. THE IMPACT OF OVEREDUCATION ON EARNINGS: EMPIRICAL RESULTS

This section empirically studies socio-demographic variables which could explain earnings differences in Spain. Our analysis of the determinants of income inequality consists of estimation the following model:

$$\log y_i = \beta_0 + \beta_1 PC_i + \beta_2 EL_i + \beta_3 LS_i + \beta_4 OS_i + \beta_0 HS_i + \varepsilon_i , \quad (1)$$

where *i* refers to the individual, *y* denotes individual's earnings, *PC* is a set of personal characteristics (gender, age and marital status), *EL* denotes education level, *SE* and *HS* denote, respectively, sector of employment and health status and ε is an error term. The definition of each variable used in the estimates is given in Table 5.

TABLE 5. Variable definitions.

Name	Definition
Income	
LOGWAGE	Natural logarithm of the individual's wage
Personal Characteristics	
MALE	1 if male, 0 otherwise
AGE	Individual's age
AGE^2	Square of the individual's age
Marital Status	
MARRRIED	1 if married, 0 otherwise
SEPARATED	1 if separated, 0 otherwise
DIVORCED	1 if divorced, 0 otherwise
WIDOW	1 if widowed, 0 otherwise
Education Level	
Lower Education (LOWEDUC)	1 if highest academic qualification is first level (ISCED 0-2), 0 otherwise
Overeducation (OVEREDUC)	1 if an individual is overeducated, 0 otherwise
Undereducation (UNDEREDUC)	1 if an individual is undereducated, 0 otherwise
Sector of employment	
PRIVATE	1 if current activity is in private sector, 0 otherwise (government sector)
OCCUPATION	A set of dummy variables representing occupational classifications
$OCCU_1$	Legislators, seniors officials
$OCCU_2$	Professionals
$OCCU_3$	Technicians and associate professionals
$OCCU_4$	Clerks
$OCCU_5$	5.- Service workers and shop and market sales workers
$OCCU_6$	Skilled agricultural and fishery workers
$OCCU_7$	Craft and related trades workers
$OCCU_8$	Plant and machine operators and assemblers
$OCCU_9$	Elementary occupations
$OCCU_0$	Armed forces
Health Status	
GHEALTH	1 if individual's self assessed health is good or very good, 0 otherwise

Source: Author´s elaboration from ECHP.

Our income measure is disposable (after-tax) individual income. It includes all the net monetary income received by the individual during the reference year (from 1993 to 2000). It includes income from work (employment and self-employment), private income, pensions and other direct social transfers received. Thus, the dependent variable is the logarithm of individual's earnings (measured in real terms).

The explanatory variables are divided into five groups: personal characteristics, education level, marital status, occupational status and health status. As personal characteristics we have included individual's age, gender and household size. Creedy (1977) shows that the variation of median real earnings with age is quadratic in form. Thus, to allow for a flexible relationship between earnings differences and age, a quadratic polynomial function of this variable is included (AGE, AGE^2). Also, the gender of individuals has been taken into consideration and a dummy variable which takes value of 1 if individual is male has been built. In fact, differential earnings constitute an important part of gender inequality in Spain.

Secondly, representing marital status, we have considered four variables (single, separated, divorced and widow) with married as the reference category. Thirdly, we will focus on the relationship between education and income distribution. Becker and Chiswick (1966) in their seminal paper show that income inequality is positively correlated with inequality in schooling and negatively correlated with the average level of schooling. The group of variables used in this analysis are referred to the maximum level of education completed. In the ECHP, education is classified into three categories based on ISCED classification: less than secondary level (ISCED 0-2), second stage of secondary level (ISCED 3) and third level (ISCED 5-7). Thus, another dummy variable has been included: less than secondary level (LOWEDUC). In this sense, many studies have shown that education is an important socio-economic characteristic in determining income inequality, so the attainment of higher educational levels can be reflecting important changes in income distribution. In addition, we know for each individual the highest level of general or higher education completed which is classified into the following categories: Level 0.- Pre-primary education; Level 1.- Primary education or first stage of basic education; Level 2.- Lower secondary or second stage of basic education; Level 3.- Lower vocational training; Level 4.- Upper vocational training; Level 5.- (Upper) secondary education; Level 6.- Post-secondary non-tertiary education; Level 7.- First stage of tertiary education; Level 8.- Second stage of tertiary education.

Also, we have include in the model if the business of organization where the individual works, belongs to the private sector or to the public sector. In addition, we have included the occupation in current job based on the International Standard Classification of Occupation – ISCO-88. Thus, a series of dummy variables representing occupational classification have been defined. The different categories are the following ones: 1) Legislators, seniors officials; 2) Professionals; 3) Technicians and associate professionals; 4) Clerks; 5) Service Workers and shop and market sales workers; 6) Skilled agricultural and fishery workers; 7) Craft and related trades workers; 8) Plant and machine operators and assemblers; 9) Elementary occupations; 10) Armed forces.

Finally, we are concerned with the influence of health status on income inequality. There exists evidence on the relationship between income inequality and health in the European Union (see Cantarero, Pascual and Sarabia, 2005). In this way, we have built a dummy variable which takes value of 1 if individual's self assessed health is good or very good and 0 otherwise.

The results of the estimation using STATA 10.0 are given in Tables 6-8. Table 6 provides the proportions of individuals defined as overeducated, adequately educated and undereducated within each of the 10 broad occupational categories. In our sample, about 16% percent of individuals are classified ass overeducated and about 19% were found to be undereducated. Obviously the proportion of over and under education employees varies among occupations. Thus, overeducated workers are most likely to be found among "Clerks" and "Service workers and shop and market sales workers". Also, the majority of individuals working as "Professionals" are adequately educated.

Regression results from our model are presented in Tables 7 and 8. We can test that overeducated workers earn about 11 percent more than workers who are adequately educated. On the other hand, undereducated workers earn about 21% less than those who are adequately educated. As can be noticed, the most important remaining variables of the model are statistically significant (sex, age, health, etc.).

Finally, we have divided the sample into two sub-samples (males and females). The results obtained are similar and as can be test, females obtained less earnings than males, as expected.

6. CONCLUSIONS

Spain's unemployment rate has increased to 17.4% in 2009. The employment problems appear as a result of the mismatch produced between offer and demand of the working force. The search of solutions for this problem depends, among other factors, on education training. Nowadays, people leave the educational system willing to find a job adequate to their level of education. However, this is not always possible. In this paper, we analyse the problem of overeducation (and underedecuation) and its effect on labour market using the data contained in the European Community Household Panel (ECHP).

In our study, about 16% percent of individuals are classified ass overeducated and about 19% were found to be undereducated. However, overeducation workers earn more than their peers with an adequate level of education.

In European Union countries, and in particular in Spain, it is necessary to create more and better jobs, improving the adaptability, mobility and competitiveness of the working force ("New beginning of the Lisbon Strategy, 2005"). In Spain, there exists a problem of overeducation which varies among occupations. Overeducated workers are most likely to be found among "Clerks" and "Service workers and shop and market sales workers". Also, the majority of individuals working as "Professionals" are adequately educated. The results obtained in this paper add more empirical evidence supporting the effect of education on earnings which needs to be considered together with occupational placement.

TABLE 6. The incidence of Overeducation by Occupational Category (percentage)

Occupation	Over-educated	Adequately Educated	Under-educated
All occupations	16,01	64,99	19,01
1.- Legislators, seniors officials	19,89	56,23	23,87
2.- Professionals	0,00	96,68	3,32
3.- Technicians and associate professionals	14,81	68,91	16,28
4.- Clerks	23,27	52,69	24,04
5.- Service workers and shop and market sales workers	22,34	60,44	17,22
6.- Skilled agricultural and fishery workers	13,65	70,85	15,50
7.- Craft and related trades workers	18,33	50,00	31,67
8.- Plant and machine operators and assemblers	19,60	49,90	30,50
9.- Elementary occupations	14,27	75,82	9,92
10.- Armed forces	10,64	85,11	4,26

Source: Author´s elaboration from ECHP.

Table 7. Regression estimates. Dependent variable: Natural log of the individual's age.

Name	**Coef.**	**Std. Err.**	**t**	**P>t**
MALE	0,3131	0,0383	8,17	0,0000
AGE	0,0925	0,0108	8,58	0,0000
AGE^2	-0,0009	0,0001	-7,01	0,0000
SINGLE	-0,2808	0,0430	-6,53	0,0000
SEPARATED	-0,1163	0,1149	-1,01	0,3120
DIVORCED	-0,3533	0,1494	-2,37	0,0180
WIDOW	0,0752	0,1422	0,53	0,5970
LOWEDUC	-0,0188	0,1061	-0,18	0,8590
OVEREDUC	0,1107	0,0503	2,20	0,0280
UNDEREDUC	-0,2172	0,0507	-4,28	0,0000
PRIVATE	-0,2439	0,0467	-5,23	0,0000
OCCUPATION				
$OCCU_1$	0,3767	0,0788	4,78	0,0000
$OCCU_2$	0,6776	0,0697	9,73	0,0000
$OCCU_3$	0,3242	0,0675	4,80	0,0000
$OCCU_4$	0,3106	0,0797	3,90	0,0000
$OCCU_5$	0,0872	0,0654	1,33	0,1830
$OCCU_6$	-0,7251	0,0904	-8,02	0,0000
$OCCU_7$	0,2415	0,0628	3,84	0,0000
$OCCU_8$	0,3039	0,0720	4,22	0,0000
$OCCU_0$	(dropped)			
GHEALTH	0,1806	0,0451	4,01	0,0000
Constant	11,8359	0,2324	50,93	0,0000
Adj. R^2	0,1606			
Numb. Observ.	4774			

Source: Author´s elaboration from ECHP.

Table 8. Regression estimates. Dependent variable: Natural log of the individual's age. Individuals classified by sex

	Males				Females			
Name	**Coef.**	**Std. Err.**	***t***	**P>*t***	**Coef.**	**Std. Err.**	***t***	**P>*t***
AGE	0,0798	0,0134	5,96	0,0000	0,1149	0,0183	6,27	0,0000
AGE^2	-0,0007	0,0002	-4,70	0,0000	-0,0012	0,0002	-5,42	0,0000
SINGLE	-0,3979	0,0562	-7,08	0,0000	-0,1102	0,0666	-1,66	0,0980
SEPARATED	-0,1785	0,1702	-1,05	0,2950	-0,0199	0,1538	-0,13	0,8970
DIVORCED	-0,3373	0,2568	-1,31	0,1890	-0,3209	0,1807	-1,78	0,0760
WIDOW	0,1917	0,2571	0,75	0,4560	0,2084	0,1706	1,22	0,2220
OCCUPATION								
$OCCU_1$	0,3232	0,0962	3,36	0,0010	0,4036	0,1407	2,87	0,0040
$OCCU_2$	0,6190	0,0940	6,59	0,0000	0,7214	0,1041	6,93	0,0000
$OCCU_3$	0,1285	0,0863	1,49	0,1370	0,5934	0,1080	5,50	0,0000
$OCCU_4$	0,2893	0,1177	2,46	0,0140	0,3351	0,1113	3,01	0,0030
$OCCU_5$	0,0319	0,0898	0,36	0,7230	0,1404	0,0965	1,46	0,1460
$OCCU_6$	-0,6840	0,1067	-6,41	0,0000	-1,0343	0,1796	-5,76	0,0000
$OCCU_7$	0,1663	0,0746	2,23	0,0260	0,3402	0,1452	2,34	0,0190
$OCCU_8$	0,2333	0,0840	2,78	0,0050	0,3209	0,1726	1,86	0,0630
$OCCU_0$	(dropped)				(dropped)			
OVEREDUC	0,1479	0,0619	2,39	0,0170	0,0429	0,0862	0,50	0,6180
UNDEREDUC	-0,2145	0,0584	-3,67	0,0000	-0,2481	0,1045	-2,37	0,0180
PRIVATE	-0,2202	0,0617	-3,57	0,0000	-0,2601	0,0712	-3,66	0,0000
GHEALTH	0,1685	0,0558	3,02	0,0030	0,1735	0,0763	2,27	0,0230
LOWEDUC	0,0602	0,1201	0,50	0,6160	-0,4743	0,2307	-2,06	0,0400
Constant	12,4687				11,3653			
Adj. R^2	0,1525				0,1792			
Numb. Observ.	3197				1577			

Source: Author´s elaboration from ECHP.

7. REFERENCES

ALBA-RAMÍREZ, A. (1993). "Mismatch in the Spanish labor market: overeducation?" Journal of Human Resources, Volume 28, Spring pp 259–278.

ALBA-RAMÍREZ, A., SAN SEGUNDO, M.J. (1995), "The Returns to Education in Spain", Economics of Education Review, Volume 14, April pp 155-166.

ASHWORTH, J.,LYNNE, J. (2001): "Student Subject Choice at Secondary and Tertiary Level: A Cross-Section Study" The Journal of Economic Education, Volume 32, Autumn, pp. 311-320.

BECKER, G.S.; CHISWICK, B.R. (1966). "Education and the Distribution of Earnings", American Economic Review, Papers and Proceedings, 56, 358-69.

BÚCHEL, F. (2007): "The effects of overeducation on productivity in Germany -the firms' viewpoint", Economics of Education Review, Volume 21,June pp 263-275.

CANTARERO, D., M. PASCUAL SARABIA, J.M. (2005). "Effects of Income Inequality on Population Health: New Evidence from the European Community Household Panel", Applied Economics, 37, 89-91.

CREEDY, J. (1977). "The Distribution of Lifetime Earning"s, Oxford Economic Papers, 29, 412-29.

DOLTON, P., VIGNOLES, A. (2000). "The Incidence and Effects of Overeducation in the UK Graduate Labour Market" Economics of Education Review, Volume 19, 179-198.

FIELDS, G. (1995), "Educational Expansion and labor markets", Pergamon Press, 2ª ed, Oxford.

FOLGER JK, NAM CB (1964), "Educational Trends from Census Data." Demog 1 2, pp

FOLGER JK, NAM CB,(1967) "Education of the American Population". Washington: USGPO, LA209.2.F66

FRANK, R. (1978), "Why women earn less: the Theory and Estimation of Differential Overqualification", American Economic Review, Volume 68, June pp 360-372.

GREEN, C., PARVINDER, K., GARETH, L. (2007), "Inmigrant overeducation: Evidence from Recent Arrivals to Australia", Economics of Education Review, Volume 26, August, pp 420-432.

HARTOG, J.(2000) "Over-education and earnings: where are we, where should we go?" Economics of Education Review, Volume 19, April pp 131–147.

HUNG, C.Y. (2008) "Overeducation and undereducation in Taiwan" Journal of Asian Economics , Volume 19, April pp 125–137.

KIKER, B.F., SANTOS, M.C., MENDES DE OLIVEIRA, M. (1997) "Overeducation and Undereducation: Evidence for Portugal" Economics of Education Review, Volume 16, April, pp. 111-125.

KIKER, B.F., SANTOS, M.C., MENDES DE OLIVEIRA,, M. (2000) "The role of human capital and technological change in overeducation" Economics of Education Review, Volume 19, April pp 199–206.

LI, F., MORGAN, J., DING, X. (2008) "The expansion of higher education, employment and over-education in China" International Journal of Educational Development, Volume 28, pp 687–697.

LINDLEY, J. (2009): "The over-education of UK immigrants and minority ethnic groups: Evidence from the Labour Force Survey", Economics of Education Review, Volume 28,February pp.80-89.

MCGUINNESS, S., BENNETT, J. (2007) "Overeducation in the graduate labour market: A quantile regression approach", Volume. Economics of Education Review, Volume 26, October pp 521–531.

SANROMÁ, E., RAMOS, R. (2004) "Overeducation and local labour markets in Spain" . 44th Congress of the European Regional Science Association, Porto, August 2004.

SIGNE, J (2006), "Regional differences in Overeducation", 46th Congress of the European Regional Science Association, Volos 2006.

VERDUGO, R., VERDUGO, N. (1989), "The Impact of Surplus Schooling on Earnings", The Journal of Human Resources, Volume 24, Autumn pp 629-643.

LA ECONOMÍA SOCIAL EN ANDALUCÍA: EL MODELO COOPERATIVO DE LA PROVINCIA DE ALMERÍA

JAIME DE PABLO VALENCIANO
E-mail: jdepablo@ual.es
JUAN URIBE TORIL
E-mail: juribe@ual.es
Departamento de Economía Aplicada
Universidad de Almería
PATRICIO JIMÉNEZ MARTÍNEZ
E-mail: pjimenez@andaluciaemprende.es
ROSA MARÍA MARTÍNEZ VÁZQUEZ
E-mail: azalea_85@hotmail.com

Resumen

La Economía Social juega un papel muy importante dentro del desarrollo económico de Andalucía, siendo la primera Comunidad Autónoma de España por el número de este tipo de empresas y por el de cooperativistas y trabajadores en sociedades laborales, aportando el sector casi el 14% del P.I.B. andaluz. La Economía social está presente en todos los sectores económicos y en más del 82% de los municipios andaluces. Es un importante actor del desarrollo local, pues en el compromiso con lo local, con las solución de sus problemas, es donde nacen y se desarrollan estas empresas, mediante la explotación de los recursos endógenos del territorio, fijando a la población en sus lugares de origen y actuando allí donde surge la necesidad y genera actividad económica donde no existía antes, siendo especialmente significativo el caso de Almería. Este es su auténtico valor añadido.

Palabras Clave: Economía Social, Cooperativas, recursos endógenos, empleo.
Área Temática: Economía Social y Discapacidad.

Abstract

Social Economics plays an important role in the economic development of Andalusia, being the first region of Spain by the number of such companies and the employees of cooperative societies and industrial sector contributing nearly 14% of GDP. Social economy deals in all economic sectors and in over 82% of the Andalusian municipalities. It is an important actor in local development, as in the commitment to the local, the solution of their problems, which are born and grow these businesses through the exploitation of endogenous resources of the territory, setting the people in their places of origin and act where the need arises and generates economic activity where there was before, being especially significant for Almería. This is the real added value.

Keywords: Cooperatives, endogenous resources, employment.
Thematic Area: Social Economics.

1. QUÉ ES LA ECONOMÍA SOCIAL

Una primera aproximación al término "Economía Social", también denominada "Tercer Sector" nos obliga, aunque en el campo teórico no exista un consenso generalizado en torno a una concepción universal de Economía Social y a su definición, a empezar por definir qué entendemos por ella:

Para el profesor de Estructura Económica de la Universidad de San Pablo-CEU, D. Javier Morillas, "podemos entender como Economía Social toda actividad económica, de carácter privado, basada en la asociación de personas en entidades de tipo democrático y participativo, con primacía de las aportaciones personales y de trabajo sobre las de capital".

Para la Confederación Empresarial Española de la Economía Social (CEPES): "Es toda actividad económica basada en la asociación de personas en entidades que apuestan por un funcionamiento democrático y participativo, tanto en la gestión como en el reparto de los excelentes".

El Profesor Barea Tejeiro conceptúa a la economía social como "la economía integrada por empresas que actúan en el mercado, produciendo bienes y servicios, y cuya distribución del beneficio y toma de decisiones no están directamente ligadas con el capital aportado por cada socio".

Para superar el problema de indefinición de la referida Economía Social se está aludiendo a un conjunto de actividades integradas en alguna de estas seis formas jurídicas:

- Cooperativas.
- Sociedades Laborales.
- Mutualidades.
- Asociaciones y Fundaciones.
- Empresas de Inserción.
- Centros Especiales de Empleo

El sector pretende dar respuesta a determinados objetivos de consenso en nuestra sociedad, entre otros:

- Al fomento de la capacidad emprendedora y empresarial, ofreciendo fórmulas empresariales adecuadas para el desarrollo del espíritu emprendedor y de la iniciativa empresarial.
- Al problema del empleo, generando puestos de trabajo de mayor estabilidad y calidad, y con más alto crecimiento que el del sistema económico tradicional, contribuyendo a fijar la población en áreas geográficas donde la economía convencional está ausente o en crisis.
- A la cohesión e inserción social, mediante el trabajo asociativo o cooperativo se favorece la integración laboral y social de personas y grupos en riesgo de exclusión social.

La Economía Social es un modelo empresarial integrado por empresas de carácter privado que actúan en el mercado produciendo bienes y servicios. Se caracterizan porque en la distribución del beneficio y el proceso de toma de decisiones interno predomina el factor humano frente al capital.

Las sociedades cooperativas y laborales son las principales componentes de la economía social, cuyos fines son la eficiencia económica y la redistribución de la riqueza generada con carácter social.

Las cooperativas son sociedades participativas que asocian a personas físicas o jurídicas con intereses y necesidades socio-económicas comunes y para cuya satisfacción y en interés de la comunidad realizan su actividad empresarial. En las cooperativas de trabajo asociado, las más extendidas, los socios son también trabajadores y funcionan bajo los principios de democracia, participación y reparto equitativo de las rentas.

Por su parte, en las sociedades laborales la mayoría del capital es propiedad de los trabajadores, que prestan en ella su trabajo retribuido de forma personal y directa a través de una relación laboral por tiempo indefinido. Estas sociedades pueden ser de responsabilidad limitada o anónimas.

2. EVOLUCIÓN DE LA ECONOMÍA SOCIAL

No cabe duda que el mayor hito del último siglo respecto a la historia de la Economía Social en España ha sido el reconocimiento constitucional de 1978. En el artículo 129.2 de la Constitución española se incluyó un mandato a los poderes públicos de fomento y promoción eficaz de las diversas formas de participación en la empresa, del cooperativismo y otras formas de economía social, y del acceso de los trabajadores a los medios de producción

A partir de la promulgación de la Constitución Española de 1978, el cooperativismo tiene un marco de desarrollo diseñado fundamentalmente por el artículo 129.2. Pero también se ve afectado por otras dos cuestiones que son: tanto el modelo socioeconómico participativo y solidario constitucional, como el diseño de la organización territorial del Estado y el reparto de competencias. Estos han sido, hasta nuestros días, los tres pilares sobre los que se ha asentado el desarrollo de la Economía Social.

Un año más tarde de la promulgación de la Carta Magna, y estando aún vigente, de forma transitoria, la Ley General de Cooperativas de 1974, se iniciaron los trabajos de redacción de un Anteproyecto de *Ley General de Cooperativas*, comprometiéndose el gobierno de la UCD a la presentación del Proyecto antes de Abril de 1980.

Pero, ante la falta de promulgación de la anunciada Ley, y en virtud de la autorización constitucional, las Comunidades Autónomas comenzaron a asumir las competencias legislativas en la materia y acabar con el monopolio legislativo del Estado al respecto.

Así lo hicieron, en un primer momento, en sus respectivos Estatutos de Autonomía, el País Vasco en 1982, seguido de Cataluña, Andalucía, Valencia y Navarra, que procedieron a la sucesiva aprobación de leyes propias sobre cooperativas, hasta que en 1987 se aprobó la Ley estatal General de Cooperativas.

La década de los 80 fue importante en la regulación y expansión de la Economía Social. Por una parte el ingreso de España en la Comunidad Europea, que mostraba un creciente interés por estas entidades; y por otro lado, la sensibilidad de los sucesivos gobiernos desde 1982, fueron determinantes en el desarrollo y extensión de este "nuevo fenómeno" de la Economía Social.

En 1985 el Parlamento andaluz aprobó la primera ley de cooperativas, que dio origen al actual Consejo Andaluz de Cooperación como máximo órgano consultivo del gobierno en materia de cooperativas. Y en 1986 el gobierno central dio el visto bueno a la Ley de Sociedades Anónimas Laborales. Estos fueron importantes instrumentos para potenciar las políticas de creación de empleo y los referentes legislativos de la Economía Social durante más de una década.

En Andalucía a partir de los 90 se experimenta una nueva fase de expansión y consolidación del sector de la Economía Social. En 1992 nace la Federación de Cooperativas de Trabajo Asociado, FAECTA, como resultado de la fusión de 3 organizaciones sectoriales: FUCA, FECOOPTA y FECAE. El proceso de unión culminaría tras la fusión con FECOAN en 1995. Un año después, en 1993 se crea CEPES- Andalucía, aglutinando con representación única a todo el movimiento asociativo de la Economía Social, que hasta ese momento estaba disperso en diversas federaciones sectoriales.

Desde esta fecha tuvieron lugar espectaculares cambios en el sector. En 1997 se aprobó la nueva legislación sobre Sociedades Laborales, con el nacimiento de la figura de la Limitada Laboral. Y en 1999 el gobierno andaluz promulgó la nueva Ley de Sociedades Cooperativas Andaluzas. Fue esta una ley destacada, por ser consensuada con el sector y por responder a los requerimientos socioeconómicos, legislativos y políticos del momento.

La segunda oleada de legislación cooperativa en la esfera autonómica, fue también a principios de los 90, (1992) cuando se transfieren las competencias del Estado en la materia, a las Comunidades que accedieron a su autonomía por la vía del artículo 143. De este modo, las Comunidades Autónomas afectadas modificaron sus estatutos y en la actualidad la totalidad tienen asumidas las competencias legislativas en materia de cooperativas.

Una vez completado el traspaso de competencias sobre cooperativas a las Comunidades Autónomas, tuvo lugar una fase de modernización y actualización de la legislación promulgada hasta la fecha. La misma ha obedecido, no solo a razones de índole técnica, sino también al afán de modernizar las cooperativas para adaptarlas a un mercado cada vez más competitivo, y dotarlas de una verdadera estructura empresarial que les permita contar con las mismas ventajas que el resto de sociedades.

En el plano orgánico también hay que resaltar, con la creación de la Consejería de Empleo y Desarrollo Tecnológico en esta legislatura, la actualización de la denominación de la Dirección General de Cooperativas, adecuándola a sus competencias y a la terminología usual, es decir, la actual Dirección General de Economía Social y Emprendedores.

En la esfera internacional se ha de destacar el reconocimiento de la Economía Social por la Comisión Europea, cuando en 1989 incluyó estas políticas en el ámbito de la Dirección General 23 como una unidad específica en la materia. Posteriormente, en 1992, se creó el Comité Consultivo Europeo de las Cooperativas, Mutuas, Asociaciones y Fundaciones (CMAF). Y recientemente con el nuevo Reglamento sobre la Sociedad Cooperativa Europea, que la ha convertido en la primera fórmula empresarial que permite a los ciudadanos europeos establecerse comercialmente en más de un país integrante, realizando un único registro y contando con una única estructura jurídica y legal.

3. LAS BONDADES DE LA ECONOMÍA SOCIAL

La actividad de este sector en los últimos 25 años ha sido pues, una palpable demostración de vitalidad y un instrumento de humanización y equilibrio, que se ha inspirado en aquellos principios de nuestra Constitución que aluden a una sociedad cohesionada, vertebrada y participativa.

Pero, a pesar de este reconocimiento, la Economía Social no siempre ha sido bien tratada ni valorada. Durante muchos años las empresas de Economía Social han sido desprestigiadas, bajo la excusa de ser peores empresas que el resto de sociedades, desde una visión puramente mercantilista. Los calificativos de empresas de subsistencia, menos competitivas, subsidiadas, o poco rentables han sido lanzados por quienes han estado interesados en su descrédito, con el fin de excluir a la Economía Social como un Agente Económico interviniente en la sociedad, o por quienes no concebían un mundo empresarial dirigido por la clase trabajadora.

En la actualidad esto parece haber cambiado. El esfuerzo hecho por las empresas de economía social, por sus organizaciones representativas y el gobierno andaluz por un gobierno socialista, han conseguido mostrar a un tipo de empresa tan competitiva como cualquier otra, tan capacitada como las demás y tan rentable como es cualquier buena empresa

La Economía Social está presente en todos los sectores económicos, y que se está introduciendo en los nuevos yacimientos de empleo como la asistencia sociosanitaria, y las nuevas tecnologías de la información y la comunicación entre otros, gracias en gran medida a su flexibilidad y capacidad de adaptación.

Es cierto que en Andalucía el sector de la Economía Social está compuesto fundamentalmente por empresas de pequeña dimensión. Pero no es menos cierto que ésta es una característica general de nuestro tejido productivo, en el que las microempresas y pequeñas empresas, con menos de 50 trabajadores, suponen el 99,3 % del total de empresas.

Por tanto, tendremos que responder a la duda: las cooperativas y las sociedades laborales, ¿Son empresas peores que las que operan con otras fórmulas jurídicas?, ¿O son exactamente iguales?

Ni lo uno, ni lo otro. No son peores, ni tampoco son iguales. Son diferentes, distintas. Y lo más importante, son necesarias para alcanzar los objetivos de una Europa cohesionada y equilibrada.

La Economía Social tiene su valor, no tanto en la figura jurídica, sino en el objeto social de la empresa y en su modo de actuación en el mercado y en la sociedad. Su éxito no se mide únicamente en función de los resultados económicos, -que son una condición indispensable para llevar a cabo su finalidad- sino, sobre todo, por su contribución en términos de solidaridad, cohesión social y actuación en el territorio.

No cabe duda de que las empresas de Economía Social están sometidas a los mismos retos y deben responder con la misma agilidad que el resto de sectores económicos que quieren estar en primera línea. Pero en estas circunstancias sociales y de mercado, en la era de la globalización, los comportamientos que conllevan una inyección de solidaridad, de democracia, de compromiso por el empleo y por la cohesión social, deben ser valorados y distinguidos.

Las empresas de economía social incorporan estos valores, pero no como moda, sino como expresión de su forma tradicional de trabajar en la sociedad,.... de su ser de empresa.

Cooperativas y sociedades laborales son, pues, empresas comprometidas con la cohesión social y con la solidaridad. Actúan allí donde surge la necesidad y generan actividad económica donde no existía antes. Son empresas que se apoyan en la capacidad emprendedora de las personas que deciden empeñarse en un proyecto empresarial colectivo, porque entre todos se tiene más fuerza que uno a uno. Donde nacen se desarrollan y de allí no se van. Empresas que integran a todas las personas y que no excluyen. Son empresas de personas porque éstas son su eje vertebrador y no la composición de capital, como ocurre en la empresa tradicional.

Esta es la contribución que se deriva del sistema de organización de las empresas de Economía Social. Un sistema original, específico, distintivo y exclusivo, que ningún otro esquema de organización empresarial puede aportar.

Pero abundando aún más en las diferencias, que son sus valores, es importante subrayar también la aportación de la Economía Social a la articulación económica y territorial de nuestra Comunidad, ya que estas empresas actúan en todos los sectores de actividad y están presentes en más del 82% de los municipios andaluces. Son así, importantes actores del desarrollo local, pues en el compromiso con lo local, con la solución de sus problemas, es donde nacen y se desarrollan estas empresas.

Esta forma de hacer empresa ha sido, y es, una de las mejores vías para ayudar a enraizar a la población en sus lugares de origen. Y ha sido una de las mejores formas de enriquecer los pueblos andaluces como lugar de vida y habitabilidad. Probablemente, de no ser por la Economía Social muchos andaluces se hubiesen visto obligados al infortunio de la emigración y muchos de nuestros pueblos estarían hoy abocados a la desertización.

El modelo económico representado por cooperativas y sociedades laborales ha demostrado su capacidad para canalizar la actitud emprendedora y promover el autoempleo.

La flexibilidad y la capacidad de adaptación e innovación permanente que caracteriza a estas empresas, son un efectivo estímulo para poner en práctica ideas y convertirlas en empresas con cabida en todos los sectores.

Además, por sus principios, ésta es la fórmula empresarial que mejor puede asumir la responsabilidad para desarrollar servicios públicos sobre una base privada, convirtiéndose en importantes entidades colaboradoras de las Administraciones.

Todas las empresas compiten por igual en el mercado, pero las de Economía Social, además de crear riqueza y empleo como el resto de las sociedades, se caracterizan por acciones de solidaridad y de apreciación de las personas por encima de otros intereses. Este es un valor añadido que tiene que ser reconocido por los gobiernos y que requiere de instrumentos reguladores y normativas en los ámbitos económicos y sociales.

La importancia de este colectivo ha sido reconocida, asimismo, por la Unión Europea a través del denominado Reglamento de la Sociedad Cooperativa, consensuado por los Estados Miembros y por la Comisión. No en vano la Economía Social representa el 8% de las empresas y el 10% del empleo total de la U. Europea.

Creemos que se deben estimular y apoyar unas iniciativas empresariales que apuestan por los valores sociales y humanos, que representan una forma alternativa de hacer empresa, de crear empleo estable y de calidad y que han demostrado esa capacidad en momentos de crecimiento y expansión de la economía y en situaciones menos favorables.

Andalucía es hoy la primera Comunidad Autónoma de España por el número de este tipo de empresas y por el de cooperativistas y trabajadores en sociedades laborales, con un ritmo de creación superior a las 2.000 empresas por año. En la actualidad el 12% de las nuevas empresas que se constituyen cada año son de Economía Social.

Este sistema socio-económico ha contribuido de forma muy importante a dinamizar e impulsar el medio rural, aprovechando los recursos endógenos y creando un tejido productivo en localidades y zonas por las que otros modelos empresariales no se han sentido atraídos o motivados para estar.

Los proyectos e iniciativas de carácter cooperativo han demostrado, además, importantes ventajas comparativas, como la flexibilidad y la capacidad de adaptación e innovación permanente, que le han conferido la notoriedad de ser uno de los instrumentos más eficaces para que los emprendedores pongan en práctica sus ideas y un buen medio de acceder al autoempleo o de crear empleos estables y de calidad.

Las cooperativas son buenos instrumentos para la ejecución de políticas cada vez más demandadas por los ciudadanos, como la atención a los mayores, a personas dependientes y a aquellas que tienen problemas de inserción social o laboral (Nuevos Yacimientos de Empleo)

4. MAGNITUDES DE LA ECONOMÍA SOCIAL

En España, a cierre de año 2007, hay 43.665 Empresas de Economía Social, 20.780 Cooperativas y 22.885 Sociedades Laborales.

Andalucía se sitúa en primer lugar en el Ranking nacional con 11.315 Empresas de Economía Social, lo que supone el 25,91% del total español. Le sigue Cataluña (11,38%) y La Comunidad Valenciana (9,93%). Entre las tres Comunidades aglutinan al 47,22% del total de empresas de Economía Social a nivel nacional.

De las 11.315 Empresas de Economía Social andaluzas, 5.596 son Cooperativas y 5.719 Sociedades Laborales. Sevilla representa el 23%, seguida de Málaga con el

14% y Córdoba del 12%.

El Sector Servicios es el más representativo en Andalucía según el número de Empresas (50%), seguido de la Industria (21%), la Construcción (20%) y la Agricultura con el 9%.

En la actualidad, el sector aporta casi el 14% del P.I.B. andaluz 442.326 son los empleos generados por la Economía Social Española. De ellos, 317.542 los generan cooperativas y 124.784 Sociedades Laborales.

Andalucía genera el 20,35% del total nacional en cuanto a empleo en la Economía Social. Esto hace que sea la primera en el Ranking nacional, seguida del País Vasco (15,71%) y la Comunidad Valenciana (14,60%). Las tres comunidades generan el 50,66% de todo el empleo de la Economía Social lo que supone más de la mitad a nivel nacional.

Las empresas de Economía Social andaluzas emplean a 90.016 trabajadores, 60.982 en cooperativas y 29.034 en Sociedades Laborales

Sevilla es la provincia que más empleo genera en Andalucía con 19.194 trabajadores (12.263 en Cooperativas y 6.931 en Sociedades Laborales), lo que supone el 21% del total andaluz. Le sigue Almería con 14.193 (16%) y Málaga con 10.194 (12%).

El Sector Servicios representa el 50% del empleo en la Economía Social andaluza, la Industria el 20%, la Construcción el 16% y la agricultura el 14%.

Según sexo, el 40,8% de los trabajadores de empresas de economía social a nivel nacional son mujeres. A nivel andaluz, este porcentaje es de 38,5%. Almería es la provincia en la que este porcentaje es mayor con un 52,2%.

Según Edad, 3 de cada 5 de los trabajadores de la Economía Social Andaluza (62,5%) tienen menos de 40 años, por lo que podemos concluir que los jóvenes son los principales generadores de empleo en el Sector. Este porcentaje supera en más de 8 puntos a la media nacional.

Según tamaño de la Empresas de Economía Social, casi el 30% de los empleados tienen su centro de trabajo en empresas con más de 100 trabajadores. En Almería este porcentaje aumenta hasta el 57,3%.

Estos datos ponen de manifiesto el carácter emergente de este sector, su pujanza, su dinamismo y su aportación a la generación de empleo y riqueza, en lo que algunos expertos ya han calificado de revolución silenciosa, de milagro empresarial y fenómeno social sin precedentes

5. EL COOPERATIVISMO EN LA PROVINCIA DE ALMERÍA

Gráfico 1

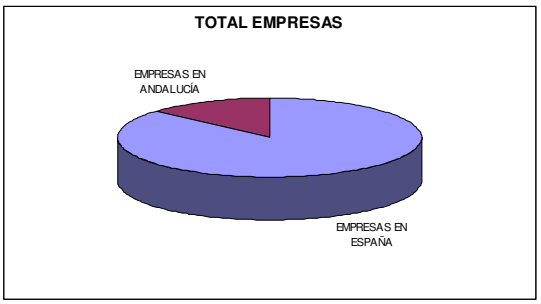


Gráfico 2

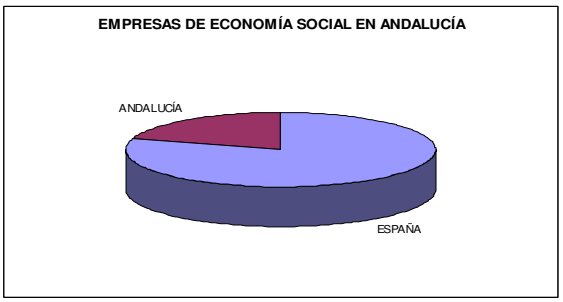


Gráfico 3

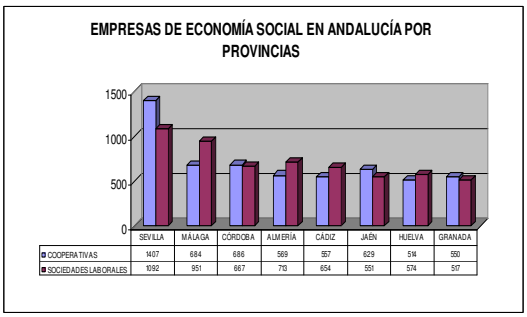
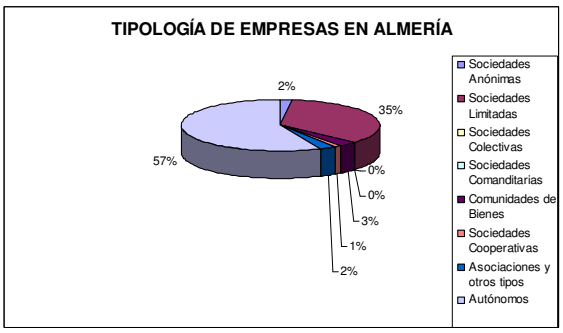


Gráfico 4



Según los datos del Gráfico 1, (Fuente: Directorio Central de Empresas (DIRCE). Año 2.007) enn España, según el Directorio Central de Empresas, existen 3.336.657 empresas de las cuales 511.728 son andaluzas, lo que representa el 15,34% del total.

Según los datos del Gráfico 2 (Fuente: Ministerio de trabajo e Inmigración. Año 2007) del total de empresas existente en España, la economía social representa un 1,31 del total con 43.665 empresas. De éstas un 47,59 son cooperativas y un 52.41% son sociedades laborales

En Andalucía, del total de empresas existente, la economía social representa el 2,25%, con un total de 11.315 empresas. De éstas un 49,46% son cooperativas y un 50,54% son sociedades laborales.

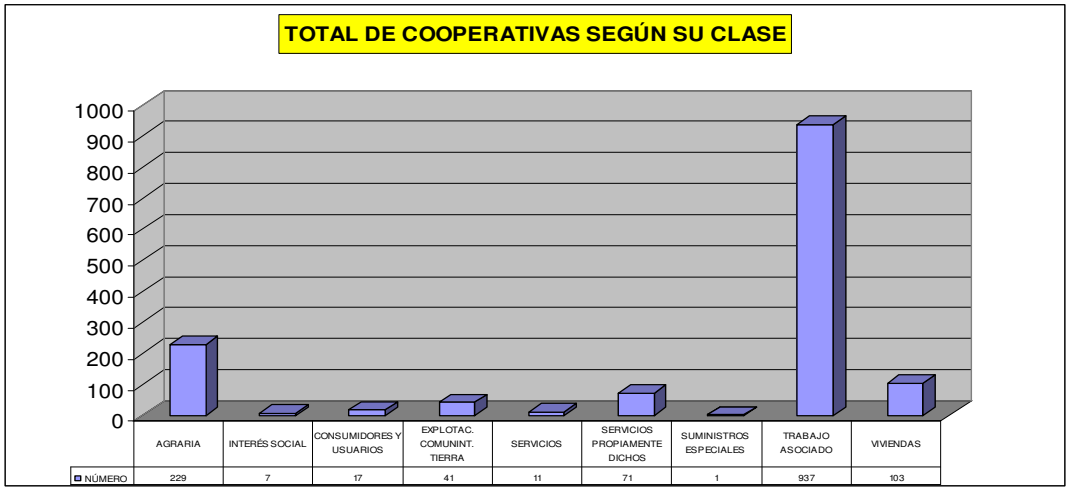
Andalucía ocupa el primer puesto dentro de ránking nacional de empresas de economía social con un 25,91% del total. La segunda Comunidad Autónoma es Cataluña con 4.970 empresas (11,38), seguida de la Comunidad Valenciana con 4.335 (9,93%). Entre las tres comunidades engloban al 47,22% de las empresas españolas de este sector.

Según los datos del Gráfico 3 (Fuente: Ministerio de trabajo e Inmigración. Año 2007), la distribución provincial muestra que el mayor número de empresas está en Sevilla, que concentra el 22,09% de toda Andalucía, seguida de Málaga con el 14,45% y Córdoba con el 11,96%. Las provincias de Almería y Cádiz concentran un 11% cada una de ellas, mientras que en Huelva y Jaén se encuentran respectivamente el 10%. En Granada se concentra el 9%. Esto da idea del equilibrio en la distribución territorial de estas empresas que se encuentran presentes en más del 82% de los municipios andaluces.

Almería ocupa el 4º puesto del ránking andaluz en cuanto al número de empresas de economía social con 1.282 empresas (11,33% del total). Específicamente, en cuanto al número de cooperativas pierde un puesto situándose en 5ª lugar con 569 (10,17%) y en cuanto a sociedades laborales gana un puesto colocándose en la tercera posición andaluza con 713 empresas (12,47% del total)

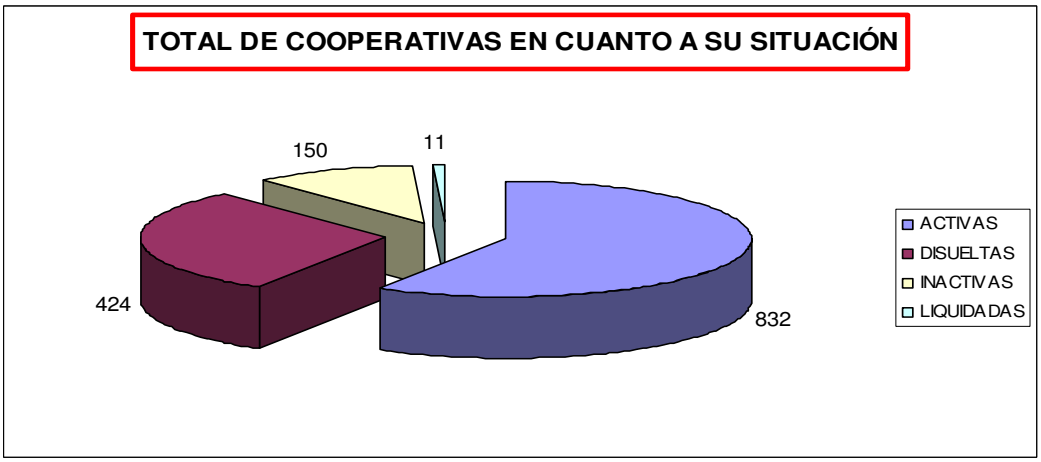
Según los datos del Gráfico 4, (Fuente: Instituto Nacional de Estadística (INE) Directorio Central de Empresas (DIRCE). Datos 2.008), Almería cuenta con un total de 45.130 empresas, en su mayoría autónomos (56,69%) y sociedades de responsabilidad limitada (34,79%), representado entre ambos más del 90% del total de empresas de la provincia, en la que las empresas cooperativas tan sólo representan el 1,32% del total.

Sociedades Cooperativas en Almería

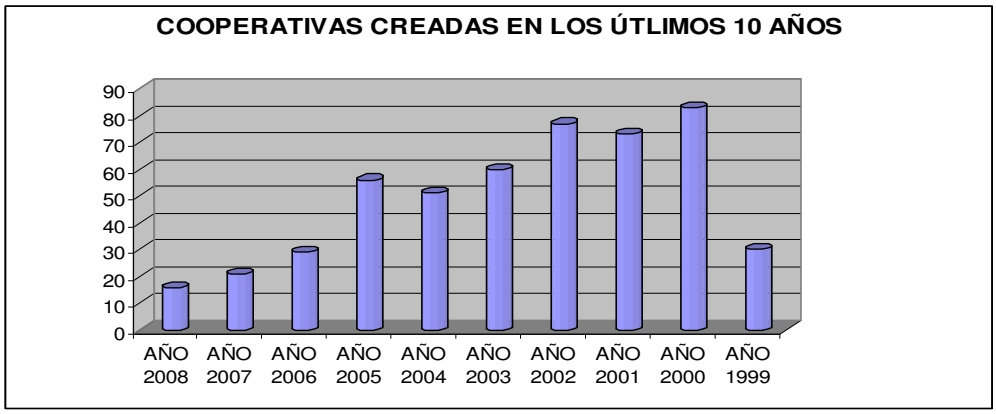


La provincia de Almería, con gran tradición cooperativa, cuenta con un total de 1.417 sociedades cooperativas andaluzas, distribuídas según el tipo de cooperativa de la siguiente forma.

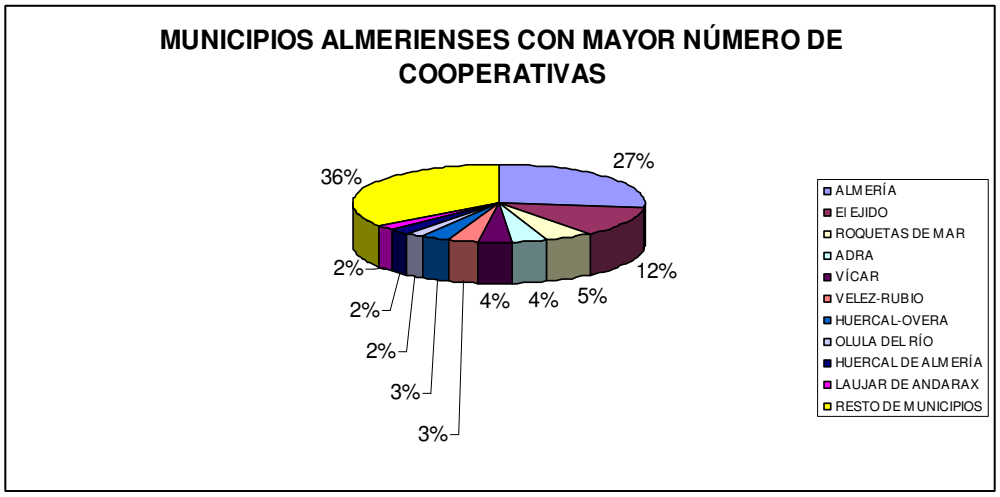
CLASE DE COOPERATIVA	NÚMERO	%
Agraria	229	16,16
Interés social	7	0,49
Consumidores y usuarios	17	1,2
Explotación comunitaria de la tierra	41	2,89
Servicios	11	0,78
Servicios propiamente dichos	71	5,01
Suministros especiales	1	0,07
Trabajo asociado	937	66,13
Viviendas	103	7,27



Del total de cooperativas existente en la provincia, tan sólo 832 están activas, representando el 58,72% del total. El resto o están inactivas (150) o en estado de disolución o liquidadas (435).



En los últimos diez años se han creado en Almería 496 sociedades cooperativas, lo que supone el 35% del total de cooperativas existentes en la provincia desde 1944. En el año 2000 se produce la mayor creación de cooperativas en la provincia, apreciándose a partir de entonces un descenso considerable acrecentándose a partir de 2.005. Esto es debido, probablemente, a que durante este período 2000-2005 estaba en vigor el Plan de Fomento de la Economía Social en Andalucía, que contemplaba importantes subvenciones para este tipo de empresa.



De entre los municipios almerienses, Almería capital cuenta con el mayor número de sociedades cooperativas con un total de 383 (27% del total), seguida de El Ejido 173 (12,21%) y Roquetas de Mar 76 (5,36%). Estas 3 localidades suman casi el 45% del total provincial

6. BIBLIOGRAFÍA

BAREA, J. (1990), "Conceptos y agentes de la Economía Social", Ciriec-España, *Revista de Debate sobre Economía Pública, Social y Cooperativa,* Octubre 1990, Nº 8, Madrid.

BAREA, J y MONZÓN, J.L. (1992) "Libro Blanco de la Economía Social en España". Ministerio de Trabajo y Seguridad Social, Madrid.

CEPES ANDALUCÍA (2007) "Informe Anual de la Economía Social Andaluza"

CEPES ANDALUCÍA (2007) " Informe Estadístico de Cepes Andalucía"

CEPES ANDALUCÍA (2007) "Central de Balances de la Economía Social Andaluza"

CEPES ANDALUCÍA (2007) "Economía Social y desarrollo Rural"

MONZON, J.L. y otros (1987), *Economía Social. Entre Economía Capitalista y Economía Pública*, Valencia. CIRIEC.

MORALES GUTIÉRREZ, A.C. (2002), "La construcción de capital social a través de la economía social: El caso Andaluz ", *REVESCO, Revista de Estudios Cooperativo*s, nº 78.

POBREZA Y PRIVACIONES SOCIALES COMO OBSTÁCULOS PARA EL DESARROLLO ECONÓMICO DE LOS PAÍSES DEL SUR [1]

ANTONIO GARCÍA LIZANA
E-mail: aglizana@uma.es
SALVADOR PÉREZ MORENO
E-mail: sperezmoreno@uma.es
Dpto. de Economía Aplicada
Universidad de Málaga

Resumen

En las últimas décadas se ha ido produciendo una paulatina aceptación de la compatibilidad entre crecimiento económico y equidad en la distribución de la renta, si bien las diferentes propuestas no se han preocupado, en general, de su forma funcional. En este trabajo se estudia la existencia de una relación específica ("Hipótesis L"), la cual se contrasta mediante la utilización de indicadores de pobreza, para un conjunto de países de América, África y Asia. Asimismo, con el propósito de buscar posibles fundamentos para tal relación, siguiendo a Sen y Myrdal, se analizan las relaciones entre pobreza y privaciones sociales en salud y educación, observando que las mismas, en cuanto representativas de la capacidad para funcionar económicamente de las amplias capas de población empobrecida, constituyen un obstáculo para el desarrollo económico de los países del Sur, de donde se sigue que una reducción de las tasas de pobreza resulta fundamental para el mismo.

Palabras Clave: Distribución de la renta, crecimiento económico, pobreza, desarrollo.
Área Temática: Economía Social y Discapacidad.

Abstract

In recent decades the relationship between economic growth and equal income distribution has gradually been accepted, such that different proposals have been overlooked, in general, in their functional form. This paper studies the existence of a specific relationship ("Hypothesis L"), which is contrasted through the utilization of poverty indicators, for a group of countries from America, Africa and Asia. In order to look for possible foundations for such a relation, following Sen and Myrdal, the relations between poverty and social deprivations in health and education are analyzed. These deprivations influence the capacity to function economically of the amply-layered poor population, constituting an obstacle for the economic development of the countries of the South. In consequence, reduction of the rates of poverty is fundamental to improve economic situation of these countries.

Thematic Area: Social Economics.

1. INTRODUCCIÓN

La relación entre equidad en la distribución de la renta y crecimiento económico constituye un campo de investigación que ha despertado mucho interés, debido, entre otras cuestiones, a su importancia intrínseca, al tratarse de dos objetivos que deben ser alcanzados si adoptamos una perspectiva completa del desarrollo. De hecho, son muchos los autores que se han ocupado del tema, aunque no siempre con iguales planteamientos.

[1] Este trabajo esta vinculado a la línea de investigación desarrollada por el Grupo de Investigación SEJ 124 de la Junta de Andalucía.
Los autores agradecen a Guillermina Martín Reyes los comentarios realizados.

Una idea muy difundida presenta la distribución equitativa de la renta y el crecimiento económico como objetivos incompatibles, bien porque se considera que la desigualdad distributiva es un requisito para el crecimiento, como hacen Mill o Kaldor, o bien porque se entiende que es la propia dinámica de crecimiento la que genera desigualdad, en la línea, por ejemplo, de Marx.

La mayor parte de los esfuerzos investigadores en este campo se ha dirigido, no obstante, y hasta fecha reciente, a examinar el cumplimiento de la conocida "hipótesis U" de Kuznets (1955), según la cual el crecimiento económico, efectivamente, incrementa la desigualdad en la distribución de la renta, pero sólo en una primera fase, favoreciendo posteriormente la igualdad distributiva. Sin duda, son numerosos los estudios aportados en esta dirección, tanto desde un punto de vista especulativo como empírico (ver Pérez Moreno, 2002). Sin embargo, la acumulación de pruebas y la presentación de argumentos contrarios a la hipótesis han ido planteando, con el paso del tiempo, serias dudas sobre la validez de la misma, al menos con carácter universal, hasta el punto de que a comienzos de los años ochenta podía afirmarse que no había razones económicas para presentar el modelo de Kuznets como algo inevitable o necesario; aun cuando la hipótesis pudiera ser útil para prevenirnos de la existencia de elementos que, en una economía en crecimiento, operan hacia la desigualdad. En consecuencia, no debía ser aceptada "como una ley económica, sino para introducir las correcciones de política necesarias a la forma de crecimiento, en términos de un objetivo de distribución" (ver Ruiz, 1982, p. 122). Con posterioridad, la opinión de los autores se ha ido endureciendo más, hasta el punto de que cabe afirmar, de acuerdo con Fields (2001) o Deiniger y Squire (1998), entre otros, que no parece plausible continuar apoyando una propuesta que está sustentada en endebles apoyos empíricos.

Mientras tanto, cada vez más investigadores se muestran confiados en la mejora del bienestar colectivo en términos de crecimiento económico y equidad simultáneamente, admitiendo una relación de compatibilidad entre ambos, aun cuando se haga bajo aproximaciones y planteamientos diferentes. Así, cabe hablar de *compatibilidad unidireccional*, *bidireccional* o *indirecta*, según se mantenga una línea de influencia que va exclusivamente de uno a otro objetivo; se suponga que, de algún modo, la influencia es mutua y recíproca; o, finalmente, la compatibilidad sea el reflejo de factores terceros que inciden de manera análoga y simultánea sobre el crecimiento y la distribución. Dentro de la primera, pueden señalarse, a su vez, dos posiciones diferentes, según se considere que son los avances en términos de crecimiento económico los que posibilitarían una sociedad más equitativa, al mantenerse, p. ej., la idea de que los beneficios del crecimiento se distribuyen ampliamente, en línea con la interpretación neoclásica del desarrollo; o bien que la equidad favorece el crecimiento económico, como se plantea dentro de la vasta literatura del crecimiento endógeno (Aghion *et al*., 1999, p. 1617).

Ahora bien, resulta llamativo que las numerosas y variadas propuestas presentadas en favor de la relación de compatibilidad no se han preocupado, en general, de la forma funcional de la misma. Ante esta situación, nuestro Equipo viene sosteniendo desde hace algunos años, además de la existencia de una relación de compatibilidad entre equidad en la distribución de la renta y crecimiento económico, la presencia de una relación específica, que responde a una configuración determinada, tal como fue presentado en la Reunión Internacional sobre *Personal Income Distribution, Inequality and Poverty*, celebrada en la Universidad de Siena en 1991 (ver Martín Reyes *et al*., 1995a).

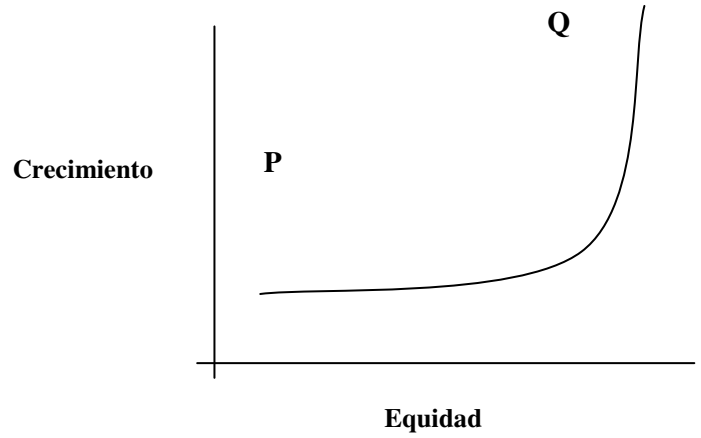
En este trabajo pretendemos continuar con el estudio de dicha relación particular, aportando nuevos contrastes empíricos y nuevos elementos de reflexión sobre el contenido económico de la misma, en el camino para mejorar nuestra comprensión del fenómeno analizado. Para ello, se tienen en cuenta las sugerencias realizadas por Gunnar Myrdal y Amartya. Sen[2], aportando algunos avances provisionales dentro de la línea de investigación que venimos desarrollando, y, por tanto, sin dar por cerrado el tema.

[2] En trabajos anteriores se han tenido en cuenta las ideas de autores como Keynes o Furtado, a partir, obviamente, de marcos explicativos distintos de los aquí utilizados (ver, p. ej., Pérez Moreno, 2002 y Zayas, 2000).

2. LA RELACION ENTRE CRECIMIENTO ECONOMICO Y EQUIDAD DISTRIBUTIVA

Desde su constitución en 1988, nuestro Equipo se alineó con los autores que defienden la compatibilidad entre crecimiento económico y distribución de la renta, entendiendo que los avances en términos de equidad contribuyen a incrementar la renta por habitante. Ahora bien, los estudios empíricos realizados nos han ido convenciendo de que tal relación responde a una forma definida y característica, creciente, no lineal y convexa[3], semejante a la curva PQ que aparece en la Figura 1 (ver Martín Reyes *et al*., 1995b).

FIGURA 1. Relación entre crecimiento y equidad



Si observamos tal relación, podremos advertir que no existe una correspondencia estricta entre los incrementos en términos de equidad y de crecimiento, como ocurriría si la función fuese lineal. Aun cuando en el punto P coinciden los menores niveles de ambas variables, en una situación dada, y en el punto Q los más elevados, la trayectoria seguida entre ambos pone de relieve un comportamiento muy desigual.

Obsérvese que la consideración detallada de esta relación específica permite apreciar la existencia de sendos tramos en los que el crecimiento económico es cuasi independiente de la distribución de la renta y viceversa, al tiempo que se advierte la presencia de una relación de compatibilidad más estricta en la parte central de la forma funcional. Este hecho nos induce a pensar que debemos tener en cuenta toda la relación en su conjunto a la hora de analizar las características de la misma, si no queremos extraer conclusiones erróneas como consecuencia de una observación parcial. Tal vez podría encontrarse aquí una explicación de algunas de las discrepancias detectadas en diferentes autores.

Por otra parte, es importante subrayar que las cosas serían igualmente muy distintas si la relación fuese cóncava, pues mientras que en el caso anterior los valores reducidos en términos de equidad nunca se corresponden con incrementos significativos en términos de crecimiento, en este segundo sí pueden coincidir. Por el contrario, si en presencia de una relación convexa los valores elevados de equidad pueden corresponderse con valores relativamente bajos de crecimiento, no ocurrirá esto en el caso de la forma funcional cóncava.

Por lo tanto, puede decirse, a modo de resumen, que la lectura económica inmediata que se desprende de la referida relación reflejada en la figura 1 es la siguiente: a) La equidad y el

[3] Conviene precisar en este punto que la definición de convexidad que empleamos en este trabajo, de acuerdo con Caballero *et al*. (1992, pp. 313-326), es la siguiente: sea K un subconjunto convexo de R^n y f una función real definida sobre K, diremos que f es convexa si, y sólo si,

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2), \quad \forall x_1, x_2 \in K \text{ y } 0 \leq \lambda \leq 1$$

La idea gráfica de esta definición es que el segmento que une dos puntos de la gráfica de la función está por encima de los puntos intermedios de la gráfica correspondiente.

crecimiento están correlacionados positivamente; b) La falta de equidad coincide con fracasos en términos de crecimiento; c) No ocurre lo mismo en el caso contrario, toda vez que niveles reducidos de crecimiento pueden coincidir con diferentes niveles de equidad; d) Un nivel de crecimiento económico elevado sólo se encuentra en condiciones de elevada equidad; e) Niveles elevados de equidad pueden coincidir con una amplia gama de niveles de crecimiento.

La admisión de estas reglas de juego entre ambas variables es independiente de que exista o no una relación de causalidad entre estos objetivos y del sentido de esa relación. Cabe la posibilidad, como hemos apuntado más arriba, de que un elemento influya sobre el otro; que ambos elementos se refuercen mutuamente, siendo cada uno de ellos causa y efecto al mismo tiempo; o que sus variaciones sean reflejo de una causa común o subyacente a ambos. Pero lo que no ofrece duda es que la frustración de la equidad puede interpretarse como una limitación para el crecimiento, por lo que al comportamiento de la primera podríamos considerarlo "dominante"[4]. Lo que no ocurriría, aun cuando pudiera hablarse de compatibilidad, en presencia de una relación cóncava, manteniendo la misma posición de ambas variables en los ejes de ordenadas. En este caso, el elemento dominante sería el crecimiento, ya que actuaría con respecto a la equidad de la manera indicada para ésta en la relación que nos ocupa.

En tales condiciones, definiremos que existe una relación de "semicomplementariedad con la equidad como elemento dominante". Entenderemos, a su vez, por *semicomplementariedad* toda relación entre dos variables caracterizada por ser positiva y no lineal, y definiremos como *variable dominante* a la situada en el eje de abscisas, cuando la relación entre ambas es convexa; o bien a la situada en el eje de ordenadas, cuando la relación es cóncava[5].

3. LA "HIPOTESIS L"

Desde un punto de vista práctico, el concepto de equidad, sin ninguna matización adicional, puede resultar excesivamente genérico como soporte de un tratamiento analítico riguroso del tema que nos ocupa, y, sobre todo, político, al no precisar exactamente la incidencia de la distribución sobre los diferentes segmentos sociales. Es decir, desde las perspectivas de la convivencia, así como del propio funcionamiento social, no sólo parece relevante el hecho de la equidad en sí, desde un punto de vista agregado; sino que además interesa conocer el modo a través del cual se manifiesta la mayor o menor equidad existente, entre los diversos grupos sociales. Esto puede resultar más claro si descendemos al terreno de la propia medición de la equidad en la distribución. Como es conocido, la distribución de la renta se suele medir utilizando distintos indicadores, en función de los objetivos y juicios de valor del propio investigador. Sin duda, la elección de un indicador u otro puede suponer la obtención de resultados divergentes sobre el comportamiento de la distribución de la renta, pues, no en vano,

[4] Si los fracasos en los niveles de equidad coinciden con los fracasos en los niveles de crecimiento, pero no al contrario, lo razonable es pensar que la equidad condiciona de algún modo el comportamiento del crecimiento. Con independencia de la existencia de relaciones de causalidad, en sentido estricto, lo que se desprendería es que la falta de equidad inhibe y obstaculiza el crecimiento, segando, al menos en parte, las posibilidades de éste. Sin embargo, el que haya o no crecimiento es irrelevante para que los niveles de equidad sean más o menos elevados. Asimismo, conseguir resolver el problema de la equidad es propiciar que el crecimiento sea una realidad, puesto que una situación más equitativa parece una condición necesaria para el crecimiento, aunque no suficiente. Por el contrario, un nivel de crecimiento económico elevado no es una condición para tener éxito en la equidad, ya que puede haber éxito en ésta sin necesidad de que existan altas cotas de crecimiento.

[5] Conviene advertir que en la literatura económica se suele hablar, en general, de compatibilidad o incompatibilidad entre los objetivos. Sin embargo, como puede deducirse a partir de las anteriores consideraciones en torno a la reseñada relación específica entre distribución de la renta y crecimiento económico, entendemos que las consecuencias son muy diferentes dependiendo de que la relación de compatibilidad sea de tipo lineal o no lineal y, en general, de la forma funcional que ésta adopta. Por tanto, se hace preciso matizar la relación funcional que liga a los objetivos que nos atañen. Cuando se habla de compatibilidad debe diferenciarse, por tanto, entre compatibilidad absoluta o simplemente complementariedad, en la medida en que la relación funcional muestre una forma lineal, y compatibilidad semicomplementaria (o semicomplementariedad) con X o Y como elemento dominante, según la relación responda a una función no lineal convexa o cóncava.
De esta forma, la expresión compatibilidad semicomplementaria con la distribución como elemento dominante, que empleamos en esta investigación, está haciendo referencia a un tipo de relación de compatibilidad en el que la distribución de la renta y el crecimiento económico no presentan una correspondencia lineal, sino convexa, cuando presentamos la variable que representa a la distribución en el eje de abcisas, de modo que su fracaso va acompañado del fracaso del crecimiento, pero no a la inversa.

cada indicador posee unas características determinadas. En concreto, una de las primeras cuestiones a la que hay que hacer frente a la hora de trabajar con indicadores de desigualdad es la relativa al empleo de juicios sobre las preferencias sociales. En torno a éstos, de acuerdo con los criterios de Sen ([1973] 1979, pp. 14-16), suelen diferenciarse dos grandes grupos de medidas de desigualdad, según empleen de manera explícita o no alguna función de bienestar. En el primer caso estaríamos hablando de medidas normativas, y, en el segundo, de medidas positivas.

No obstante, hay que tener en cuenta que en las medidas positivas los juicios sobre las preferencias sociales aparecen implícitos en la propia expresión matemática, pudiéndose derivar a raíz de ella la función de bienestar asociada y, por ende, evaluar estas medidas, al igual que en el caso de los indicadores normativos, en términos de bienestar. A partir de tales consideraciones, se puede apreciar que las distintas medidas de desigualdad, en consonancia con las preferencias sociales que incorpora su formulación, enfatizan unos aspectos u otros, ponderan de forma diferente la presencia de ciertos grupos de población, etc.

En este sentido, si consideramos por determinadas razones políticas, económicas o sociales que la atención a los estratos con menores rentas debe convertirse en una elección prioritaria, puede estar justificado utilizar directamente indicadores de pobreza para medir la desigualdad en la distribución de la renta, ya que permitirán, precisamente, reflejar la situación particular de tales grupos de población dentro de la sociedad[6].

En este estudio, nos decantamos por la utilización de los indicadores de pobreza para evaluar la distribución de la renta por diversos motivos. Por un lado, a partir de las distintas aportaciones teóricas en que podemos apoyarnos para defender la relación de compatibilidad, en particular las que interpretan la equidad como elemento que favorece el crecimiento (por ejemplo, en Myrdal, Prebisch, Keynes, etc.), podemos destacar la especial incidencia que cabe atribuir a las capas más desfavorecidas económicamente, de tal manera que su mejora juega un papel muy importante para impulsar el crecimiento económico. Por otro, desde un punto de vista axiológico, consideramos que si las grandes desigualdades sociales son deplorables *per se* para la mayor parte de la sociedad, el rechazo es aún mayor cuando lleva consigo unas deficientes condiciones de vida para los estratos más desfavorecidos de la sociedad. Así, reducir el número de personas calificadas como pobres parece más deseable que disminuir la desigualdad de una sociedad en general o, dicho de otra manera, anteponemos el principio de equidad categórica al de equidad vertical (ver Cuadrado Roura *et al*., 2000, pp. 290-291). Por otro lado, las investigaciones realizadas (ver Martín Reyes *et al.*, 1989, 1995a y 1995b; ECB, 1998; García Lizana y Pérez Moreno, 2000, 2003; Pérez Moreno, 2002) ponen de manifiesto la existencia de una relación más estrecha con el crecimiento económico cuando se utilizan indicadores de pobreza como medida de la distribución de la renta, que cuando se emplean otros indicadores convencionales de distribución[7].

Y además de ello, desde un punto de vista político la preocupación fundamental es la transferencia de rentas desde los más favorecidos a los menos, con el fin de mejorar la situación de estos.

De acuerdo con tales ideas, es más adecuado fijarnos en la situación de las capas más desfavorecidas para llevar a cabo nuestro análisis, tomando como indicador, por tanto, alguna medida de pobreza. Si situamos, pues, un indicador de pobreza como medida de la falta de equidad en el eje de abscisas en lugar de considerar a la equidad, la relación específica de compatibilidad a que nos venimos refiriendo vendría reflejada por una función decreciente, no lineal y convexa simétrica a la mostrada en la Figura 1, tal como la que se muestra en la Figura 2, en la que se presenta al PIB por habitante en ordenadas, como expresión del nivel de

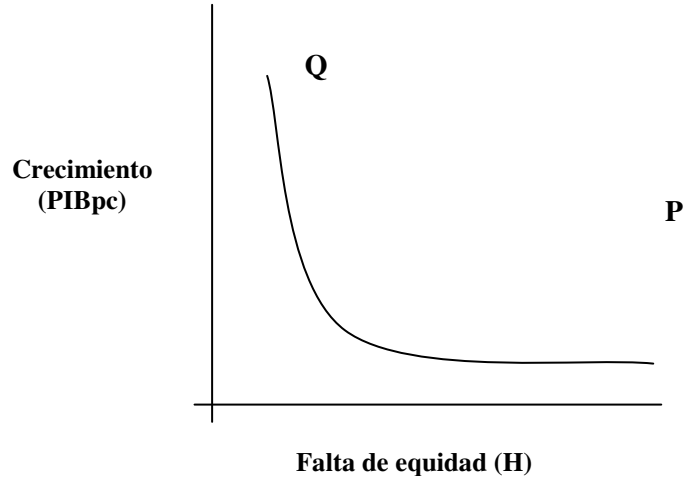
[6] En esta línea, Spicker (1993, p. 47) señala que la relación entre desigualdad y pobreza es de naturaleza lógica, ya que «si definimos la pobreza como una forma de desventaja en un contexto social, y desigualdad es el término con el que nos referimos a la desventaja en un contexto social, entonces la pobreza es una forma de desigualdad».

[7] Entre otras razones, podría argumentarse que la mejora de la equidad del conjunto puede ser debida a que la renta se reparta mejor entre las capas altas y medias de la población (sin que se altere la situación relativa de las bajas), lo cual, desde el punto de vista económico, puede ser poco relevante, dado que los comportamientos de cada grupo social (en cuanto grupo de referencia) suelen ser bastante similares.

crecimiento, y un indicador de pobreza, H, en abscisas, como medida de la falta de equidad en la distribución de la renta.

Dada la forma funcional que adopta la representación gráfica empleada para expresar la relación entre crecimiento y falta de equidad, en este artículo denominamos "hipótesis L" a la referida relación de compatibilidad semicomplementaria con la equidad como elemento dominante, utilizando como elemento de referencia la magnitud que representa la pobreza en el cuerpo social.

FIGURA 2. Relación entre crecimiento y falta de equidad



4. CONTRASTACIÓN DE LA "HIPÓTESIS L" EN LOS PAÍSES EN DESARROLLO

En trabajos anteriores nos hemos ocupado de contrastar la hipótesis indicada en distintos entornos geográficos, con datos nacionales e internacionales, y para situaciones caracterizadas por diferentes situaciones de desarrollo. Nos proponemos, en este caso, fijar nuestra atención en un conjunto de países en desarrollo (adoptando, por tanto, una perspectiva internacional) de Asia, África y América.

No obstante, antes de adentrarnos en esta cuestión, nos parece oportuno dejar sentados ciertos aspectos referentes a la medición de la pobreza, no sólo por la incidencia que puede ejercer sobre los propios resultados finales, sino igualmente por el amplio debate mantenido al respecto, lo que nos aconseja una cierta precisión.

Por un lado, encontramos a los defensores de la consideración de la pobreza desde una perspectiva absoluta, como la incapacidad para acceder a un determinado paquete básico de bienes y servicios definidos exógenamente a partir de unos criterios objetivos y relativamente estables a lo largo del tiempo y en cualquier lugar del espacio, relacionados con la propia subsistencia. Por otro lado, se sitúan quienes optan por una definición relativa de la pobreza, entendida como una situación carencial en relación con el nivel de vida medio de la población de referencia. De esta forma, este criterio relativo obvia la definición de una cesta de bienes y servicios fija y aplica una interpretación relativa de la privación, según la cual el concepto de necesidad es básicamente una construcción social y, por lo tanto, cambiante, por lo que la pobreza, desde esta perspectiva, se debe entender fundamentalmente en términos de privación relativa con respecto a lo que es normal en una sociedad.

Ahora bien, parece evidente que, como subrayan Muñoz Bustillo y Esteve (1998, p. 36), "la presentación de estos dos criterios como alternativos, sin embargo, no debe llevarnos a la opción a favor de uno de ellos y al abandono del otro por irrelevante, ya que en cierta medida, como tantas veces ocurre en las ciencias sociales, tal disyuntiva es falsa, en el sentido de que la debilidad del criterio absoluto de pobreza: el ser incapaz de reflejar aquellos componentes sociales y por lo tanto relativos de la pobreza, es la fuerza del concepto relativo, mientras que, simultáneamente, éste fracasa a la hora de hacer énfasis en los aspectos de pobreza severa

recogidos por el criterio absoluto, pudiendo ayudar a trivializar los efectos de la pobreza sobre determinados colectivos". Asimismo, cabe cuestionarse el carácter absoluto del criterio así denominado, toda vez que, a pesar de lo apuntado con anterioridad, las necesidades que son calificadas como básicas pueden considerarse propias de cada escenario temporal y espacial determinado, lo cual se corresponde con la relativa estabilidad señalada más arriba. Lo que no impide que al hablar de pobreza absoluta estemos llamando la atención sobre condiciones mínimas requeridas para una adecuada vida humana, aun cuando sea en un marco social determinado. O, siguiendo el planteamiento de Sen (1985), existe un "núcleo irreductible de pobreza", el cual se encuentra implícito en el mismo concepto de pobreza, y que está formado por aquellos elementos que dotan al individuo de "capacidad para funcionar" en el medio social.

Ante esta tesitura entendemos plausible sostener que, especialmente en los países en desarrollo, los datos sobre pobreza absoluta debieran presentar una conexión más acusada con la actividad económica general, en tanto en cuanto dicho criterio hace referencia al nivel de subsistencia que dota a los individuos de "capacidad para funcionar"; es decir, que permite a los individuos su participación, en condiciones aceptables, en los distintos procesos socioeconómicos de cada sociedad (producción, consumo, participación política, etc.), debiendo ser mayor dicha conexión en la medida en que consideramos la privación de las necesidades más básicas y primarias y, por tanto, más restrictivas para dicha participación. Sólo en la medida en que tales restricciones se superen de manera sensible en un determinado medio social (es decir, se reduzca la pobreza absoluta y se logre con ello la expansión de la capacidad para funcionar de un número suficiente de individuos, las posibilidades de funcionamiento efectivo de la sociedad, y con ello su dinámica económica, se verán desbloqueadas. Y esto es así porque, entre otras razones, la pobreza absoluta implica niveles de vida reducidos (en términos de nutrición, salud, educación, etc.), lo que, como señala Myrdal (1970), afecta al nivel de productividad, a través de sus efectos sobre el deseo y la capacidad de trabajar y la eficiencia en el trabajo[8]. En la medida en que las tasas de pobreza absoluta son mayores, los niveles de vida reducidos afectan a más individuos, con las correspondientes consecuencias sobre la productividad, y, en consecuencia, sobre la producción. De ahí la insistencia de Myrdal (1970) en la idea "de que elevar el nivel de vida de las masas puede tener importantes efectos sobre la productividad" (p. 36).

Otro aspecto de singular importancia en la medición de la pobreza es la determinación del universo de referencia con respecto al que se define el umbral o línea de pobreza, que permite dividir a la sociedad en dos segmentos diferenciados en función de su posición económica. Y ésta es una cuestión que no sólo está presente en las aproximaciones relativas de la pobreza, sino también en las medidas absolutas. No en vano, en la fijación de los estándares absolutos que definen los umbrales de pobreza pueden considerarse alternativamente el nivel de desarrollo y las normas y valores sociales de una región, de un país o de un conjunto de países; o bien, adoptar un umbral de pobreza universal para todos los países que permita calibrar la situación de la pobreza en las distintas latitudes del planeta atendiendo a referencias comunes que se estiman significativas.

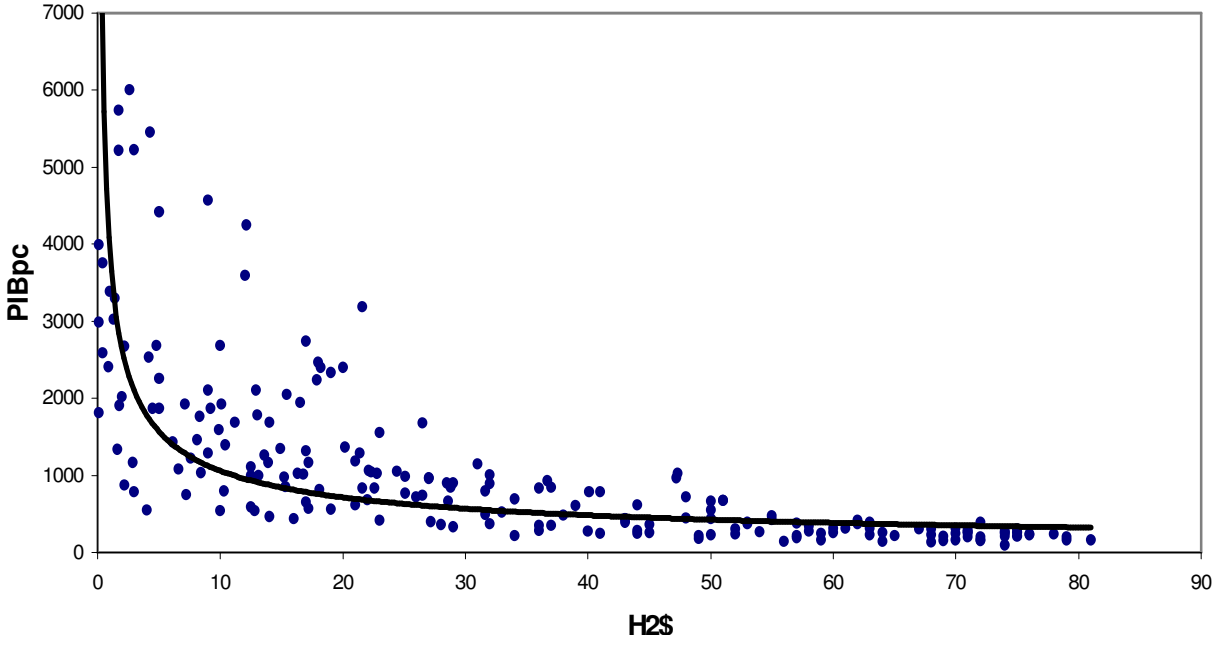
Teniendo en cuenta las apreciaciones precedentes, en este trabajo tomamos como base de nuestra contrastación empírica los indicadores de pobreza estimados por Sala-i-Martín (2002b) a partir de los datos estadísticos sobre distribución de la renta que dispone el Banco Mundial. En concreto, consideramos el porcentaje de población que vivía con menos de 2 dólares diarios –a los precios internacionales de 1985, ajustado en función de la paridad del poder adquisitivo– en 1970, 1980, 1990 y 1998 en 52 países en desarrollo de Asia, África y América –que en 1998 suponían alrededor del 63% de la población mundial–, adoptando, pues, un umbral de pobreza

[8] Gunnar Myrdal (1970) establece a este respecto una diferenciación muy precisa entre los países desarrollados y subdesarrollados: "los altos niveles de renta alcanzados en los países desarrollados, así como su sistema de seguridad social, hacen posible que, para estos países, los niveles de nutrición y, en general, el nivel de vida puedan considerarse sólo desde el punto de vista del bienestar, sin necesidad de referirse al deseo y a la capacidad de trabajar y a la eficiencia del trabajo" (p. 29). De donde concluye que, a diferencia de los países subdesarrollados, en el estudio de la dinámica del desarrollo puede prescindirse en los primeros de la consideración de los factores apuntados; pero no en los segundos. Lo que resulta congruente con la hipótesis L, toda vez que superados unos determinados niveles de pobreza, la mejora de las condiciones de vida alcanzadas permite avanzar en términos de crecimiento, rompiéndose la dependencia con las tasas de pobreza, y pudiendo buscar las explicaciones en los factores económicos habituales.

internacional común y de carácter absoluto[9]. Por su parte, los datos relativos al PIB *per capita* (ajustados en PPA, en dólares USA de 1987), correspondientes a 1970, 1980, 1990 y 1997 han sido tomados del Programa de las Naciones Unidas para el Desarrollo (1998 y 1999).

A partir de dicha información estadística, en la Figura 3 hemos representado las nubes de puntos que se obtienen al relacionar el PIB *per capita* y la mencionada tasa de pobreza.

FIGURA 3. PIB *per capita* y tasa de pobreza (ajuste potencial)



Como puede comprobarse, esta nube de puntos presenta una forma congruente con nuestra tesis, toda vez que la relación manifestada entre PIB *per capita* y tasa de pobreza no sólo es decreciente, sino que además muestra una forma no lineal y convexa. No obstante, para verificar empíricamente la no linealidad de esta relación funcional hemos aplicado el test Box-Cox, pudiéndose concluir de acuerdo con los resultados obtenidos como la especificación potencial se ajusta mejor que la lineal[10].

Ante los problemas detectados inicialmente de heteroscedasticidad y autocorrelación residual, hemos optado por utilizar la matriz de varianzas y covarianzas propuesta por Newy y West (1987) para la estimación de los errores estándar de los coeficientes, considerando un parámetro de truncamiento igual a 3. En este sentido, los resultados obtenidos en la estimación del mencionado modelo por mínimos cuadrados ordinarios pueden resumirse como sigue:

[9] Como es sabido, la medida de pobreza absoluta más usual es la fracción de la población que vive con menos de un dólar diario. Sala-i-Martin (2002a) recuerda que la definición original de este umbral de pobreza proviene de Ravallion *et al*. (1991). Estos investigadores tratan de establecer cuál es el nivel mínimo de vida en los países menos desarrollados por debajo del cual una persona puede ser considerada *pobre*, situando dicha línea de pobreza entre 30.42 y 31 US$ mensuales; a partir de estas estimaciones, se extendió el redondeo de esta cantidad a 1 US$ diario –a los precios internacionales de 1985, ajustado en función de la paridad del poder adquisitivo–. Muy pronto esta línea de pobreza fue adoptada por el Banco Mundial como la definición «oficial» de pobreza absoluta. Poco después, por alguna razón, apareció en la literatura una segunda línea de pobreza que dobló la cifra original a 2 US$ por día. En este trabajo consideramos únicamente esta segunda línea de pobreza, toda vez que los datos manejados muestran que un número considerable de los países contemplados presenta para algunos años porcentajes de población con menos de 1 US$ diario nulos.

[10] La elección del modelo potencial entre los distintos ajustes no lineales radica, además de su adecuación a la relación existente entre los datos de las variables que estamos examinando, a la sencillez y fácil linealización e interpretación de los mismos; asimismo, este modelo muestra directamente el valor de la elasticidad (parámetro β).

CUADRO 2. Modelo potencial $(PIBpc = \alpha \cdot H2\$^{\beta})$

N	$Ln(\alpha)$	β	F	R^2	$\overline{R^2}$
208	8.2710 (44.293)	-0.5675 (-9.628)	295.591	0.589	0.587

Fuente: Elaboración propia

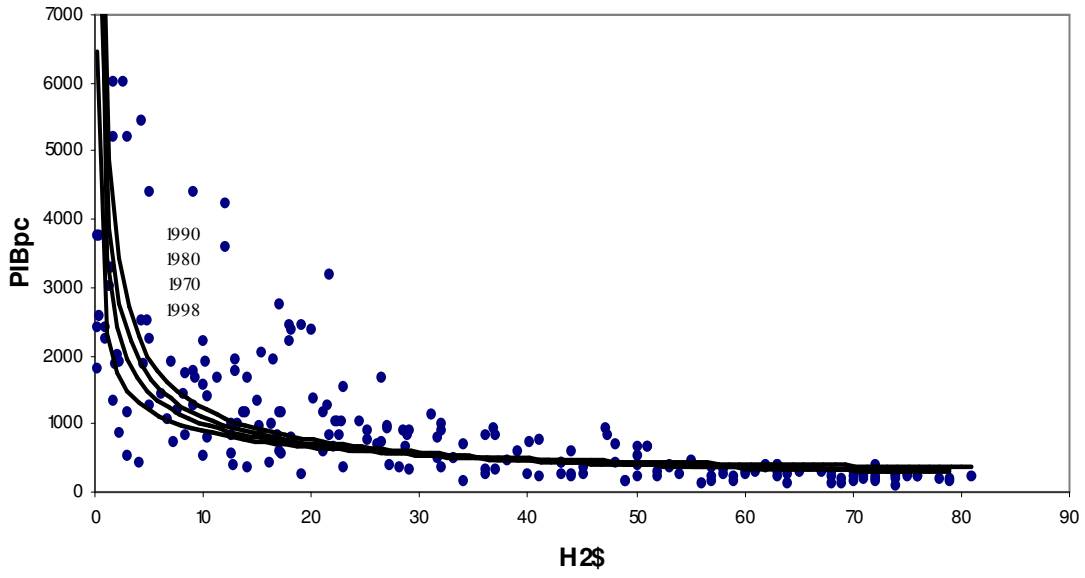
Como puede constatarse en el cuadro 2, la estimación de dicho modelo proporciona resultados bastante satisfactorios. Por una parte, los valores observados de los estadísticos t de Student (entre paréntesis) y F de Snedecor nos lleva a rechazar que los parámetros, individual y conjuntamente considerados, sean significativamente iguales a cero, al tiempo que los coeficientes de determinación explican en torno al 60% de las variaciones de PIB *per capita* (tomado en logaritmos).

En suma, cabe afirmar, que la nube de puntos anterior muestra la presencia de una relación bastante definida, que se corresponde con la forma funcional propuesta. No obstante, conviene subrayar la presencia de anchas bandas de fluctuación, que nos pone sobre aviso de las significativas peculiaridades que presentan dichos territorios en sus perfiles socioeconómicos, y que en los ajustes anteriores estamos obviando en la medida en que la variabilidad de la tasa de pobreza no recoge suficientemente tales circunstancias.

Con el fin de contemplar por separado las distintas curvas que ajustan la relación entre PIB *per capita* y tasa de pobreza para las distintas fechas y continentes considerados, en las figuras 4 y 5 hemos representado en el plano de coordenadas las respectivas formas funcionales que se desprenden de los datos pertenecientes, por un lado, a cada fecha y, por otro, a cada una de las áreas continentales estudiadas.

Asimismo, hemos procedido a estimar los ajustes econométricos que se derivarían de las distintas divisiones efectuadas en la base de datos, obteniendo unos resultados, en todos los casos, igualmente satisfactorios y plenamente conformes con una forma geométrica que caracteriza a la "hipótesis L" (ver cuadros 3 y 4).

FIGURA 4. PIB *per capita* y tasa de pobreza por fecha (ajustes potenciales)

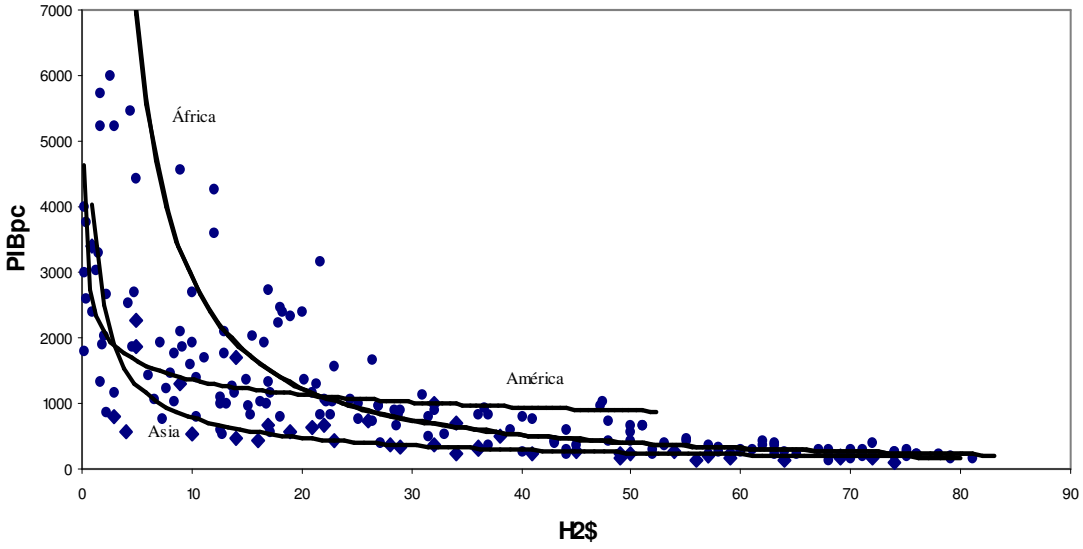


CUADRO 3. Modelo potencial $(PIBpc = \alpha \cdot H2\$^{\beta})$

	N	$Ln(\alpha)$	β	F	R^2	$\overline{R^2}$
1970	52	8.1706 (14.059)	-0.5479 (-3.243)	42.756	0.461	0.450
1980	52	8.4467 (38.915)	-0.6300 (-9.376)	87.918	0.637	0.630
1990	52	8.6730 (42.089)	-0.678 (-10.809)	116.835	0.700	0.694
1997-98	52	8.0081 (44.649)	-0.4782 (-8.505)	72.336	0.591	0.583

Fuente: Elaboración propia

FIGURA 5. PIB *per capita* y tasa de pobreza por continente (ajustes potenciales)



.

CUADRO 4. Modelo potencial $(PIBpc = \alpha \cdot H2\$^{\beta})$

	N	$Ln(\alpha)$	β	F	R^2	$\overline{R^2}$
Asia	40	8.3012 (24.091)	-0.7133 (-7.750)	92.679	0.709	0.702
América	84	7.8304 (70.026)	-0.2659 (-6.289)	48.268	0.371	0.363
África	84	10.8121 (30.215)	-1.2330 (-14.053)	455.468	0.847	0.846

Fuente: Elaboración propia

En la Figura 4 puede observarse la enorme similitud existente entre los ajustes funcionales relativos a cada una de las fechas analizadas, pudiéndose comprobar que, si bien los extremos de las respectivas curvas convergen en buena medida entre sí, se aprecia un incremento paulatino del valor de la elasticidad entre 1970 y 1990; esta tendencia se ve abortada en 1997-98, mostrándose en esta fecha el parámetro representativo de la elasticidad como el menor de todos los estimados. Al hilo de estos datos, puede afirmarse que a finales de la década de los 90 una reducción del nivel de pobreza en un determinado grado porcentual suponía un menor efecto favorable sobre el PIB *per capita* que en los inicios de las décadas de los 70, 80 y 90, periodo a lo largo del cual la magnitud del impacto había sido creciente.

Por su parte, en la Figura 5 conviene poner de relieve que la familia de curvas que se deriva de la consideración por separado de los grupos de países relativos a cada una de las áreas continentales muestra algo más que un simple desplazamiento entre los ajustes funcionales. Por

un lado, puede apreciarse que los países americanos con mayores cotas de pobreza se encuentran en el plano de coordenadas en un nivel superior a los países homólogos de los otros continentes, lo que sugiere una mayor capacidad de crecimiento para análogos niveles de pobreza; pero, sin embargo, no ocurre igual a medida que la tasa de pobreza se reduce, ofreciendo la menor elasticidad de los tres continentes. Asimismo, llama la atención que el ajuste funcional realizado a partir de los datos que presentan los países africanos, si bien parte de posiciones con elevadas tasas de pobreza similares a las de los países considerados de Asia, logra despegar con niveles de pobreza superiores a la de los países de los otros continentes, alcanzando, además, en algunos casos puntuales (como Gabón, Sudáfrica o el propio Botswana –el país del mundo que ha experimentado un mayor crecimiento en los últimos 30 años–) cifras considerables de PIB por habitante; en este sentido, la elevada elasticidad que muestra el modelo africano puede interpretarse como un síntoma más del enorme potencial de crecimiento que detenta el continente negro y que, desgraciadamente, aún hoy continúa sin rentabilizar, en detrimento de millones de vidas humanas.

Sin duda, en la determinación de las distintas situaciones observadas que presentan los países en desarrollo influyen muchas circunstancias. Con carácter general, podría afirmarse que aspectos tales como la dotación de recursos naturales, el ordenamiento político y social, el sistema económico, el nivel y distribución del capital físico y humano, la estructura productiva, el nivel tecnológico, la composición de las exportaciones, el nivel de desarrollo del sistema financiero, la estructura del empleo, los movimientos financieros, etc. constituyen factores que podrían contribuir a explicar de manera significativa esas diferencias observadas entre los países de los tres continentes.

5. ALGUNOS FACTORES EXPLICATIVOS DE LA HIPÓTESIS L

Aun cuando las repetidas observaciones empíricas efectuadas coinciden en confirmar la forma funcional de la relación entre PIB por habitante y tasa de pobreza, parece pertinente preguntarnos por los posibles factores que subyacen detrás de dicha relación. En el complejo escenario existente resulta difícil discernir cuáles pueden ser los principales aspectos que en mayor medida influyen en la misma. Aunque *a priori* puede resultar muy aventurado decantarnos por algún grupo de factores determinado, a la luz de las explicaciones aducidas por G. Myrdal (1968 y 1970) y A. Sen (1985 y 1995), queremos centrar nuestra atención en dos de las dimensiones básicas que conforman la *pobreza humana*[11]; esto es, la privación de una vida larga y saludable y la privación de conocimientos, en la medida en que ambas reflejan de manera inequívoca el deficiente nivel de vida de la población, condicionando la que Sen denomina *capacidad para funcionar*, o Myrdal identifica como *deseo y capacidad de trabajar y eficiencia en el trabajo*.

Recordemos que entre las razones que Myrdal argumenta para explicar que la pobreza inhibe y obstaculiza el crecimiento económico, el autor sueco subraya que en los países subdesarrollados una parte importante de la población está mal alimentada y sufre unas deficientes condiciones de salud, vivienda, higiene y educación; "todo esto reduce la capacidad y los deseos de trabajar intensamente, lo que es causa de que se alcance una producción inferior". Y añade, "esto implica que la aplicación de medidas para aumentar el nivel de sus rentas aumentará la productividad" (Myrdal, [1970], 1973, p. 77). Obviamente, en la medida en que las rentas reducidas sean compartidas por un mayor porcentaje de la población, las deficiencias en términos de salud (y que podemos apreciar mediante la reducida esperanza de vida) y educación (medida en términos de tasas de analfabetismo) estarán tanto más extendidas, de donde se sigue, a través del deseo y la capacidad y eficiencia para trabajar, una menor productividad de la mano de obra del país correspondiente y, por consiguiente, una producción inferior. Por tanto, la mejora de las condiciones de vida de la población desfavorecida supone

[11] El denominado índice de pobreza humana mide la privación con respecto a las tres dimensiones empleadas para definir el desarrollo humano básico, utilizando como variables para mostrar estas privaciones el porcentaje de población que morirá antes de los 40 años, el porcentaje de adultos analfabetos, y alguna medida que permita captar el acceso a los servicios básicos por parte de la población (como el porcentaje de población sin acceso a servicios de salud o de agua potable, o el porcentaje de niños con peso insuficiente).

casi un requerimiento para lograr altas cotas de producción; lo cual requiere, en los términos planteados por Myrdal, la mejora de las rentas de esas numerosas masas de la población o, lo que es lo mismo, reducir las tasas de pobreza[12].

Como puede comprenderse, tanto la privación de una vida larga y saludable como la privación de conocimientos son dos fenómenos que están fuertemente interconectados con la pobreza. De hecho, si consideramos el porcentaje de población que se estima que no sobrevivirá hasta los 40 años de edad (SOB) y la tasa de analfabetismo de adultos (ANA) como variables representativas de las privaciones apuntadas, a partir de los datos de 1998 para 52 países en desarrollo de los tres continentes con los que venimos trabajando, se observa una fuerte correlación con las tasas correspondientes de pobreza, lo que corrobora las opiniones mantenidas por Myrdal y aporta elementos que contribuyen a dotar de contenido teórico a la hipótesis L, desde una perspectiva de oferta[13].

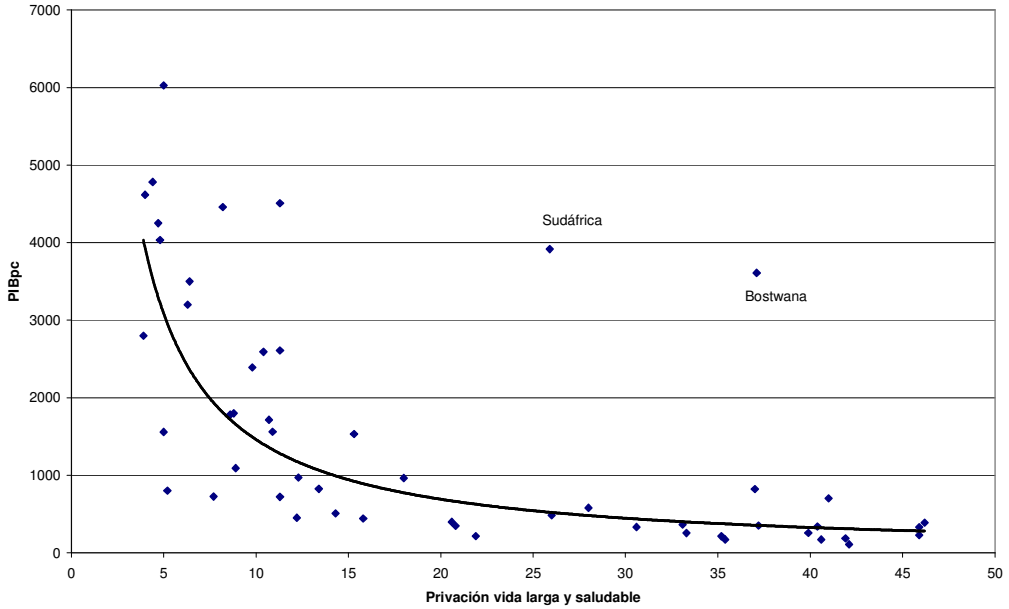
En efecto, en el cuadro 5 se recoge un sencillo ajuste econométrico en el que incorporamos como variables explicativas del crecimiento económico tanto el porcentaje de población que se estima que no sobrevivirá hasta los 40 años de edad (SOB), como la tasa de analfabetismo de adultos (ANA). Los resultados de la estimación –aplicando nuevamente la matriz de varianzas y covarianzas propuesta por Newy y West (1987)– son altamente satisfactorios, presentando claramente mejores ajustes la especificación potencial que la lineal, de acuerdo con el test Box-Cox.

CUADRO 5. Modelo potencial $(PIBpc = \chi \cdot SOB^{\delta} \cdot ANA^{\varepsilon})$

N	$Ln(\chi)$	δ	ε	F	R^2	$\overline{R^2}$
52	10,032 (32,686)	-0,705 (-3,652)	-0,433 (-2,643)	55,301	0,652	0,638

Fuente: Elaboración propia

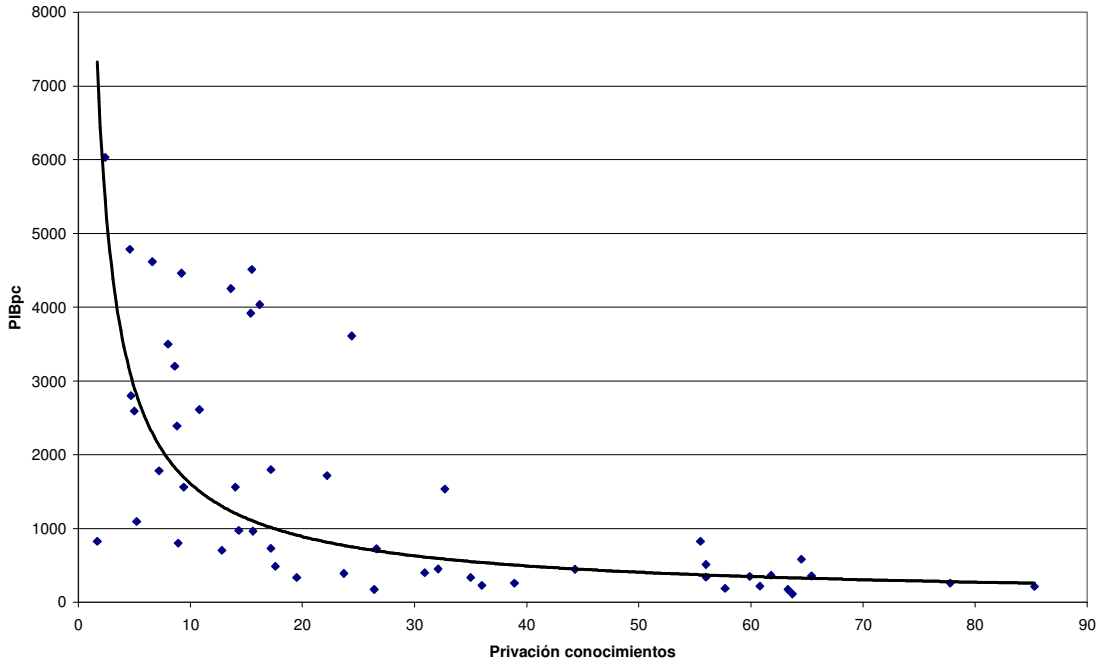
FIGURA 6. PIB *per capita* y privación de una vida larga y saludable (ajuste potencial)



[12] Conviene advertir que los argumentos expuestos por Myrdal en contra de la tesis de la conveniencia de la desigualdad para el crecimiento no se limitan a los señalados. Así, por ej., alude a la poca relevancia del argumento que defiende la desigualdad porque favorece el ahorro, ya que "los grandes terratenientes y capitalistas suelen gastar sus rentas en consumo e inversiones conspicuas y, sobre todo en América Latina, enviando el dinero al extranjero" (Myrdal, 1970, p. 77). La referencia a América Latina podría ayudar a entender la menor elasticidad observada en la función correspondiente con respecto a las de África y Asia, a que se ha aludido más arriba.

[13] En trabajos anteriores (ver, p. ej., Pérez Moreno, 2002) se han expuesto argumentos en defensa de la relación desde el lado de la demanda.

FIGURA 7. PIB *per capita* y privación de conocimientos (ajuste potencial)



En las figuras 6 y 7 se pueden visualizar por separado las nubes de puntos que relacionan las respectivas las dimensiones básicas de la pobreza humana con el PIB per capita, pudiéndose apreciar claramente la relación no lineal y convexa presente en ambos casos, similar a la descrita anteriormente bajo el rubro de 'hipótesis L'.

De esta forma, la lectura que más arriba hacíamos sobre la relación específica existente entre crecimiento económico y equidad –utilizando como indicadores el PIBpc y la tasa de pobreza, respectivamente– encuentra su cabal interpretación en la relación del crecimiento con respecto a la privación de una vida larga y saludable, por un lado, y con respecto a la privación de conocimientos, por otro. Podrían afirmarse, pues, también ahora, las conclusiones fundamentales siguientes: a) La capacidad de disponer de una vida larga y saludable y de conocimientos están correlacionadas positivamente con el crecimiento; b) La privación de una vida larga y saludable y de conocimientos coincide con fracasos en términos de crecimiento; c) No ocurre lo mismo en el caso contrario, toda vez que niveles reducidos de crecimiento pueden coincidir con distintos niveles de tenencia de conocimientos y de una vida larga y saludable; d) Un nivel de crecimiento económico elevado sólo se encuentra en condiciones de reducida privación de conocimientos y de una vida larga y saludable; e) Reducidos niveles de privación de conocimientos y de una vida larga y saludable pueden coincidir con una amplia gama de niveles de crecimiento.

Así las cosas, cabría inferir que el hecho de que las altas cotas de pobreza sean una rémora para el crecimiento económico y la reducción de la pobreza una condición necesaria pero no suficiente para el mismo, puede venir explicado, al menos en parte, por la conexiones existentes entre el crecimiento económico y la privación de conocimientos y de una vida larga y saludable. Dichas conexiones podrían ser interpretadas como justificación de la relación existente entre el PIB per capita y la tasa de pobreza, si tenemos presente, además, que cuando empleamos otros indicadores de la distribución de la renta menos correlacionados con la privación de conocimientos y de una vida larga y saludable –como pueden ser los índices de desigualdad de Gini, Theil, etc.–, la relación de éstos con el crecimiento económico se muestra bastante menos definida.

En definitiva, de acuerdo con el análisis efectuado, cabe concluir que detrás de la 'hipótesis L' parece subyacer la influencia en la determinación del crecimiento económico de la privación de capital humano por parte de la población con menos recursos, factor que puede interpretase como un particular determinante del crecimiento, en la medida en que actúa como elemento dominante en el sentido descrito más arriba; pero, igualmente, en la medida en que influye en la capacidad laboral de la población y, por tanto, en su productividad.

6. CONCLUSIONES

El estudio de las relaciones entre crecimiento económico y distribución de la renta constituye uno de los capítulos más atractivos vinculados con la economía del desarrollo. No obstante, a pesar de los esfuerzos realizados, aún no se ha llegado a un acuerdo entre los economistas. Sin embargo, poco a poco se va afianzando la idea de que existe una relación de complementariedad entre la equidad en la distribución y el crecimiento económico, debiendo anotar en la lista de defensores de tal aproximación teórica a autores tan preclaros como Keynes, Myrdal o Prebisch. Las recientes teorías del crecimiento endógeno también han permitido el aporte de algunas consideraciones teóricas al respecto.

Sin embargo, no basta con afirmar que ambos (crecimiento y equidad) se mueven en la misma dirección, al ser compatibles, sino que necesitamos conocer la forma de la relación de compatibilidad de que estamos hablando, pues las implicaciones políticas que se siguen son muy distintas según las diversas aproximaciones concebibles.

En concreto, la forma que viene siendo defendida por nuestro equipo de investigación, desde 1988, constituye una función no lineal, convexa y creciente, cuando se representa la distribución de la renta en el eje de abscisas. A tal relación la denominamos relación de semicomplementariedad con la equidad como elemento dominante, por entender que cualquier fracaso en términos de equidad se corresponde con un fracaso en términos de crecimiento, pero no al contrario, pues distintos niveles de crecimiento pueden coincidir con el nivel más elevado de equidad en la distribución de la renta.

La consideración de la equidad en términos de reducción de las tasas de pobreza, medidas por indicadores diversos, ofrece una amplia capacidad explicativa en relación con los niveles alcanzados en el desarrollo económico (medidos en términos de PIB por habitante). La forma funcional resultante es decreciente, no lineal y convexa, por lo que la hemos denominado "Hipótesis L", por sugerir tal consonante su representación gráfica.

La estimación de distintas relaciones del tipo indicado, a partir de diversas colecciones de datos ha sido abordada por nuestro equipo en los últimos años. Ofrecemos en este trabajo estimaciones relativas a diversos países del Hemisferio Sur, confirmándose una vez más la relación señalada.

Ahora bien, para una adecuada valoración de la situación es importante identificar los posibles factores explicativos de la relación observada. En este sentido, puede ser de interés tener en cuenta las opiniones postuladas por A. Sen y, en especial, por G. Myrdal, sobre la incidencia de la pobreza extrema sobre la productividad, y, por tanto, sobre el crecimiento económico, a través de su influencia sobre las condiciones de vida de la población empobrecida, y, en concreto, de su salud y conocimientos. En este sentido, hemos constatado la influencia de dos dimensiones básicas de la pobreza humana, a saber, la privación de una vida larga y saludable y la privación de conocimientos, con respecto al PIB por habitante. De acuerdo con todo ello, sería preciso superar un determinado nivel de pobreza para mejorar las condiciones de vida de la población y lograr que las inhibiciones y obstáculos al crecimiento que ello supone puedan reducirse hasta el punto de poder encontrar niveles de PIB por habitante significativamente más elevados. De este modo, se podrá avanzar conjuntamente en el crecimiento económico y la distribución de la renta en los países menos desarrollados, con todo lo que puede suponer para interpretar el comportamiento económico real y para diseñar, subsiguientemente, los tipos de intervención que los agentes políticos deben asumir.

BIBLIOGRAFÍA

AGHION, P., CAROLI, E. y GARCÍA-PEÑALOSA, C. (1999): "Inequality and Economic Growth: The Perspective of the New Growth Theories", *Journal of Economic Literature*, vol. 37, num. 4, pp. 1615-1660.

BANCO MUNDIAL (2001): *Informe sobre el desarrollo Mundial 2000-2001. Lucha contra la pobreza*. Banco Mundial, Washington.

BANCO MUNDIAL (2002): *Informe sobre el desarrollo Mundial 2000-2001. Instituciones para os mercados*. Banco Mundial, Washington.

CABALLERO, R., GONZÁLEZ PAREJA, A. y TRIGUERO, F. (1992): *Métodos matemáticos para la Economía*. McGraw-Hill, Madrid.

CUADRADO ROURA, J. R., MANCHA, T. VILLENA, J. E., CASARES, J., GONZÁLEZ, M., MARÍN, J. M. y PEINADO, M. L. (2000): *Política Económica. Objetivos e instrumentos*. McGraw-Hill, Madrid.

DEININGER, K. y SQUIRE, L. (1998): "New ways of looking al old issues: inequality and growth". *Journal of Development Economics*, vol. 57, núm. 2, pp. 259-287.

ECB (1998): "Pobreza y desarrollo en España", en EDIS et al.: *Las condiciones de vida de la población pobre en España*. FOESSA, Madrid.

FIELDS, G. S. (2001): *Distribution and Development. A new look at the Developing World*. The MIT Press, Cambridge.

GARCÍA LIZANA, A. y PEREZ MORENO, S. (2000): "Una relación específica entre distribución de la renta y crecimiento económico", en *XIV Reunión ASEPELT ESPAÑA*. ASEPELT España, Oviedo.

GARCÍA LIZANA, A. y PEREZ MORENO, S. (2003): "Una relación específica entre distribución de la renta y crecimiento económico: 'la Hipótesis L'". *Estadística. Journal of the Interamerican Statistical Institute*, núm. 164-165, pp. 153-180.

KUZNETS, S. (1955): "Economic growth and income inequality". *American Economic Growth*, núm. 45, pp. 1-28.

KUZNETS, S. (1966): *Modern Economic Growth*. Yale University Press, Yale.

MARTÍN REYES, G.; GARCÍA LIZANA, A. y FERNÁNDEZ MORALES, A. (1989): "La distribución territorial de la pobreza en España", en CEBS: *La pobreza en la España de los 80*. Editorial ACEBO, Madrid.

MARTÍN REYES, G.; GARCÍA LIZANA, A. y FERNÁNDEZ MORALES, A. (1995a*)*: "Poverty and development: An Econometric Approach", *Research on Economic Inequality*, 6, pp. 365-377.

MARTÍN REYES, G.; GARCÍA LIZANA, A. y FERNÁNDEZ MORALES, A. (1995b*): Distribución de la renta y crecimiento económico: el caso de Andalucía*. Instituto de Desarrollo Regional de la Universidad de Sevilla, Sevilla.

MUÑOZ DE BUSTILLO, R. y ESTEVE, F. (1998): "Marco General: Pobreza y Economía de mercado". En EDIS ET AL. (1998): *Las Condiciones de vida de la población pobre en España. Informe General*. Fundación FOESSA, Madrid.

MYRDAL, G. (1968): *Asian Drama: An Inquiry into the Poverty of Nations*. Random House, New York.

MYRDAL, G. [1970] (1973): *Reto a la pobreza.* Ariel, Espluges de Llobregat. Obra original: *The Callenge of World Poverty.* Patheon Books.

NEWEY, W. y WEST, K. (1987): "A simple Positive Semi-definite, Heroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, núm. 55, 703-708.

PÉREZ MORENO, S. (2002): *El Papel de la Distribución de la Renta en el Crecimiento Económico*. Servicio de Publicaciones de la Universidad de Málaga, Málaga.

PNUD (1998): *Informe sobre Desarrollo Humano 2001*. Mundi-Prensa, Madrid.

PNUD (1999): *Informe sobre Desarrollo Humano 2001*. Mundi-Prensa, Madrid.

RAVALLION, M., DATT, G. y VAN DE VALLE, D. (1991): "Qualifying Absolute Poverty in the Developing World". Review of Income and Wealth, núm. 37, pp. 345-361.

RUIZ BRAVO DE MANSILLA, G. (1982): *Igualdad humana y realidad económica*. Pirámide, Madrid.

SALA-I-MARTIN, X. (2002a): The Distrurbing "Rise" of Global Income Inequality, NBER Working Paper 8904 (April).

SALA-I-MARTIN, X. (2002b): The World Distribution of Income (estimated from Individual Country Distributions). NBER Working Paper 8933 (May).

SEN, A. [1973] (1979): *Sobre la desigualdad económica*. Crítica, Barcelona. Obra original: *On Economic Inequality*. Clarendon, London.

SEN, A. [1985]: *Commodities and Capabilities*. Noth Holland, Amsterdam.

SEN, A. [1995]: *Nuevo examen de la desigualdad*. Alianza Editorial, Madrid.

SPICKER, P. (1993): *Poverty and Social Security*. Routledge, London.

ZAYAS FUENTES, S.J. (2000): "Vigencia actual del pensamiento de Raúl Prebisch. Un marco de análisis de la distribución de la renta en la periferia, *Cuadernos de Ciencias Económicas y Empresariales*, nº 38, pp. 61-85.

ÁREA XII
MODELIZACIÓN ECONÓMICA Y FINANCIERA
MODELAÇÃO ECONÓMICA E FINANCEIRA

NUMERICAL ANALYSIS OF THE MILSTEIN METHODS APPLIED IN ASSETS PRICING MODEL AND OTHER MAIN FINANCE STOCHASTIC MODELS

JULIO CEZAR ALVES THOMAZ
jcthomaz@statistiken.mat.br
Faculty of Economics
University of Coimbra

Resumo

Cada vez mais, as Equações Diferencias Estocásticas (SDE) tem diversas aplicações na modelagem de vários problemas em economia e finanças. Sendo que algumas dessas equações, não tendo solução análitica, faz-se necessário o uso de métodos numéricos para encontrar uma solução aproximada. Um exemplo clássico de SDE são modelagens dos ativos subjacentes dos contratos de opções, que em geral são modelos por movimento Browniano Geométrico, uma SDE. Neste artigo é desenvolvido a investigação da análise numérica da solução aproximada, na discretização temporal pelo método de Milstein no movimento Browniano Geométrico, sendo esse um processo de precificação dos ativos. Bem como, a aplicação do método Milstein aos principais modelos estocásticos financeiros. Assim, o estudo fica concentrado especialmente na apresentação da modelagem e solução numérica da SDE, como também no estudo da convergência para a solução análitica, com ilustração de resultados numéricos.

Palavras-chave: Precificação de Ativos, Análise Numérica, Método de Diferenças Finitas, Modelagem de Financeira.
Área temática: Modelação Económica e Financeira.

Abstract

The Stochastic differential equations (SDE) find application in the modelling of diverse problems in economics and finance. Since some of these equations having no analytical solution, making it necessary the use of numerical methods to find an approximate solution. A classic example of SDE, is modelling the underlying assets of the contracts of options, in general, are models for Geometric Brownian motion, which is an SDE. This paper are presents the numerical analysis of the approximate solution in the time discretization by the scheme Milstein in Geometric Brownian motion, and this is a process of assets pricing. And, applying the scheme Milstein to the main finance stochastic models. Once the study is concentrated especially in the presentation of the modelling and numerical solution of the SDE in finance, as well as the study of convergence to the analytical solution with numerical results for illustration.

Key words: Assets Pricing, Numerical Analysis, Finite Differences Methods, Financial Modelling.
Thematic Area: Economic and Financial Modelling.

1 - INTRODUCTION

Stochastic Differential Equations (SDE) find application in modelling of diverse problems in Economics and Finance. We note that the behavior of a assets (or some other underlying) can be described by a SDE. We know that the options are a financial instrument that depend on the underlying assets, has been modeled by a Brownian geometric motion, that is stochastic differential equations given by

$$\begin{aligned} dS_t &= \mu\, S_t dt + \sigma\, S_t dW_t \,, \\ S(t) &= S_0 \,, \end{aligned} \tag{1}$$

where the constant μ is the expected return of the assets, σ^2 is the variance of the return. The undefined integral,

$$W_t = \int_0^t \xi\, dv \,, \tag{2}$$

can be identified as the Wiener's process. This is a Gaussian stochastic process with continuous but no differentiable, samples function, with mean $EW_t = 0$ and covariance $EW_t W_s = \min(t;s)$, but (2) is symbolically presented as $dW_t = \xi_t\, dt$. In general, (1) can be written in general differential form,

$$\begin{aligned} dS_t &= f(t,S_t)\, dt + G(t,S_t)\, dW_t \,, \\ S(t) &= S_0 \,, \end{aligned} \tag{3}$$

which is the Ito's SDE for the process S_t, ie underlying assets pricing. As $f(t,S_t) = \mu\, S_t$ and $G(t,S_t) = \sigma\, S_t$ for the Brownian geometric motion. It should be understood as an abbreviation for the integral equation,

$$S_t = S_0 + \int_0^t f(v,S_v)\, dv + \int_0^t G(v,S_v)\, dW_v \tag{4}$$

The samples functions of W_t are, with probability 1, continuous but not to limited variation in any interval. The second integral in (4) can not be considered as a Riemann-Stieltjes ordinary integral with respect to sample functions of W_t, even for G uniform, because the depends value of the intermediate points in the approximate sums. In 1951, Ito set integrals of the form,

$$Y_t = \int_0^t G(v)\, dW_v \,, \tag{5}$$

for a large class of functional not anticipate G of Wiener's process, and thus puts the theory of stochastic differential equations on a solid basis of calculations.

The solution of SDE (3) is a Markov's process with continuous sample functions, ie a diffusive process. Conversely, each diffusive process is the solution of SDE (3) where f and G^2 are, respectively, the coefficients of drift and diffusion.

The calculation of SDE belongs to the probabilistic methods or direct since it is interested in the random variable S_t and its variation. An Eq. (3) or (4) is a rule of construction, with which

to construct a sample path of S_t given the trajectory of a Wiener process W_t and an initial value S_0.

In case of (1), we know the analytical solution, that was presented in Arnold (1974) and Thomaz and Loula (2007),

$$S_t = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + W_t\right\} . \tag{6}$$

This equation was studied by Black and Scholes (1973) and Merton (1973) in the development of the modeling of pricing options. But in other situation as in "*Mean Reverting Process*",

$$dS_t = \lambda(\mu - S_t)dt + \sigma S_t dW_t ,$$

which is used to model of the dynamics of the risk-free interest rate, we do not know the analytical solution, among others cases which are presented below. Thus the need for application of numerical methods.

2 - CONVERGENCE

In assessing the quality of a numerical scheme is necessary to have some kind of measure of how well the method approximates the analytical solution, ie how it is that the numerical solution converges to the analytical solution. In the stochastic case, unlike the deterministic case, there more than one method of measuring convergence. We consider the two main types of convergence, such as *Weak* and *Strong Convergence*.

2 . 1 - Strong convergence

The concept of a strong convergence using the concept of absolute error, which is just the expectation of the absolute value of the difference between the numerical approximation and analytical solution in time "t", ie $E\left|X_t - Y_{\tau_n}\right|$. We say that a discrete approximation Y_{τ_n}, in time, of exact solution X_t of a SDE, converging in the strong sense in order $\gamma > 0$ if there is constant $C < \infty$ such that

$$E\left|X_t - Y_{\tau_n}\right| \leq C\ \Delta^{\gamma} , \tag{7}$$

for all the fixed steps $\Delta \in (0,1)$. The strong convergence provides a approximation of trajectories of individual sample way of the solution of a SDE.

2 . 2 - Weak convergence

Strong convergence is computationally expensive and not a necessary condition for the expected value, only information on the type of solution is of interest. For example, if and only if request to calculate of the moments of the solution X_t, we are not obliged the individual approach of the paths of X_t which leads to the concept of weak convergence. We say that a discrete approximation in time, Y_{τ_n}, of a solution X_t of SDE converges in a weak sense with order $\gamma > 0$ if, for any polynomial g, there is a constant $C < \infty$ such that

$$\left|E\left(g\left(X_t\right)\right)-E\left(g\left(Y(\tau_n)\right)\right)\right| \le C\ \Delta^{\gamma}\ ,$$
(8)

for all fixed steps $\Delta \in (0,1)$, provided that they are functional. This criterion provides an error of the mean, variance or any moments that are necessary calculate. It is usually easier and faster to implement numerical methods subject to the condition of weak convergence. It is important when approaching a problem of identification (if possible) is a good approximation of the path is required or is an approximation of a functional solution.

3 – MILSTEIN NUMERICAL METHODS

In this work we introduce the numerical analysis of solution of the discrete time approximation of stochastic differential equations in finance. Thus here we present the discrete time approximation called the numerical methods of Milstein has the order of convergence greater than the numerical method of the Euler and Maruyama, which is the simplest of numerical methods.

The *Euler-Maruyama approximation* is one of the time discrete approximations of an Ito's process $S=\{S_t, t\ge 0\}$ satisfying (3) for $t\in[0,T]$ with an initial value $S_0 \in (-\infty,\infty)$. For a given discretization $t_0=\tau_0<\cdots<\tau_N=T$ of the $[t_0=0,T]$, an *Euler-Maruyama Numerical Methods* is a discretized time stochastic process Y satisfying,

$$Y_{n+1}=Y_n+f(\tau_n,Y_n)\left(\tau_{n+1}-\tau_n\right)+G(\tau_n,Y_n)\ \Delta W_n\ ,$$
(9)

where $\Delta W_n = W_{\tau_{n+1}}-W_{\tau_n}$ for $n=0,1,\cdots,N-1$, with initial value $Y_0=S_0$. Already the *Milstein approximation*, in the case one-dimensional, is given for only a sum of term of the expansion of *Ito-Taylor*,

$$G'G\ I_{(1,1)}=\frac{1}{2}G'G\left\{\left(\Delta W_n\right)^2-\left(\tau_{n+1}-\tau_n\right)\right\},$$

in the Euler-Maruyama approximation, then we get the *Milstein Numerical Methods*

$$Y_{n+1}=Y_n+f(\tau_n,Y_n)\left(\tau_{n+1}-\tau_n\right)+G(\tau_n,Y_n)\ \Delta W_n+\frac{1}{2}G'G\left\{\left(\Delta W_n\right)^2-\left(\tau_{n+1}-\tau_n\right)\right\}$$
(10)

The increments ΔW_n in (9) and (10), of the Wiener process may be generated by a generator of pseudo-random numbers of Gaussian independent. This work used the Polar Marsaglia generator presented in Kloeden and Platen (1995).

The numerical methods (9) or (10) determines values of the approximating process at the discretization times only. If required, values can then be determined at the intermediate instants by an appropriate interpolation method. The simplest is the *piecewise constant interpolation* with $Y(t)=Y_{n_t}$, $t\in[0,\infty)$, where n_t is the integer defined by $n_t=Max(0,1,\cdots,N:\tau_n\le t)$ that is the largest "n" for which τ_n does not exceed "t". However,

$$Y(t)=Y_{n_t}+\frac{t-\tau_{n_t}}{\tau_{n_t+1}-\tau_{n_t}}\left(Y_{n_t+1}-Y_{n_t}\right),$$
(11)

linear interpolation, is often used because it is continuous and also simple.

We shall concentrate on values of a discrete time approximation at the given discretization instants. It will not be possible to reproduce the finer structure of sample paths of an Ito's process as they inherit the irregularity of the sample paths of the driving Wiener's process, in particular their no differentiability.

We verify that the expected strong convergence rates theoretically for Euler-Maruyama approximation is of $\gamma = 1/2$, already for Milstein approximation is of $\gamma = 1$. Thus, the Milstein approximation is better for discretization of the problem of finance stochastic model, because of not requiring a high refinement of the mesh for the numerical solution, which reduces computational cost, and leads us to accurate response in less time.

4 - MAIN FINANCE STHOCHASTIC MODELS

It presents some cases of stochastic differential equations that have applications in finance.

4 . 1 – Linear with constant coefficients

We consider the simple case of stochastic differential equation where the coefficients of drift and diffusion are independent of information received in time, ie

$$dS_t = \mu\, dt + \sigma\, dW_t \ , \tag{12}$$

where W_t a Wiener's process. This stochastic differential equation, the coefficients are $f(t, S_t) = \mu$ and $G(t, S_t) = \sigma$. So they do not depend on the set of information of the time I_t. The average of ΔS_t for a small interval of time of length $\tau_{n+1} - \tau_n = h$ is given by $E[\Delta S_t] = \mu\, h$. The expected variation in ΔS_t is given $\text{var}(\Delta S_t) = \sigma^2\, h$

4 . 2 – Geometric stochastic differential equation

The stochastic differential equation are used for modelling the price of underlying assets is not a linear stochastic differential equation of constant coefficients, but the equation of the Geometric Brownian motion, as presented in (1). This equation was studied by Black and Scholes (1973) in developing the modelling of pricing options.

This model implies that in general the formal notation, $f(t, S_t) = \mu\, S_t$ and $G(t, S_t) = \sigma\, S_t$. Therefore, the coefficient of drift and diffusion depend on the information that becomes available in time "t". However, this dependence is direct preference. The drift and standard deviation change proportionally with S_t. In fact, the dividing eq. (1) for S_t, we obtain

$$\frac{dS_t}{S_t} = \mu\, dt + \sigma\, dW_t .$$

This means that although the coefficients of drift and diffusion of the geometric stochastic differential equation has variation with time "t", these in relation the S_t still are parameters invariant in the time.

Geometric Stochastic differential equation has analytical solution as presented in (6) in the introduction of this work. Thus, can be made a study of convergence of the numerical method presented here.

4 . 3 – Square root process

A model of square root process is,

$$dS_t = \mu\, S_t dt + \sigma\sqrt{S_t}\; dW_t \ . \tag{13}$$

Here the drift coefficient follows in the same way that of geometric SDE, however the diffusion coefficient is a function of the square root of S_t. If the volatileness of the price of the asset not to increase "*over*" when S_t increases, this model can be more appropriate.

4 . 4 – Mean reverting process

A stochastic differential equation that has been useful in modelling price of assets is the mean reverting model, frequently used to model the dynamics of interest rate:

$$dS_t = \lambda(\mu - S_t)dt + \sigma\, S_t dW_t \ , \tag{14}$$

as S_t is generally under some "average value" μ, the term in brackets, $(\mu - S_t)$, become positive. This makes with that dS_t either most likely positive. Thus, S_t eventually will move in the direction to revert to the value μ.

Another stochastic differential equation related to this, is when the drift is of the mean reverting type, but the distribution is dependent of the square root of S_t,

$$dS_t = \lambda(\mu - S_t)dt + \sigma\sqrt{S_t}dW_t \ . \tag{15}$$

There is a significant difference between two models, stochastic differential equations, for mean reverting process.

The mean reverting process has a tendency, but the derivations on this tendency are not completely random. Prices of assets may have some predictions with periodicities. This usually makes the model inconsistent with market efficiency.

5 – NUMERICAL RESULTS

It presents results of numerical experiments with the objective to illustrate some important aspects concerning the application of numerical methods in stochastic differential equation in the financial market. The first part of the experiments is presented to study the convergence of computational implementation, and the second part is given the various situations presented previous stochastic differential equation.

5 . 1 – Study of convergence

The simulations presented have as main objective to verify the convergence of the numerical solution of the stochastic differential equation (1) for the analytical solution (6). The Euler-Maruyama numerical methods in (9) for Brownian geometric motion is,

$$Y_{n+1} = Y_n + \mu\, Y_n(\tau_{n+1} - \tau_n) + \sigma\, Y_n\, \Delta W_n \ , \tag{16}$$

and *Milstein Numerical Methods in* (10) is

$$Y_{n+1} = Y_n + \mu\, Y_n (\tau_{n+1} - \tau_n) + \sigma\, Y_n\, \Delta W_n + \frac{1}{2}\sigma^2 Y_n \left\{ (\Delta W_n)^2 - (\tau_{n+1} - \tau_n) \right\}. \quad (17)$$

This computational simulations using the parameters: h = 0.001 , $\mu = 0.15$, $\sigma = 0.30$ and with initial value $S_0 = 100$. Also adopted was a relationship of Δ^γ, for time steps equal 16, 32, 64, 128 and 256.

Figure 1 - Analysis of Error for Euler-Maruyama and Milstein

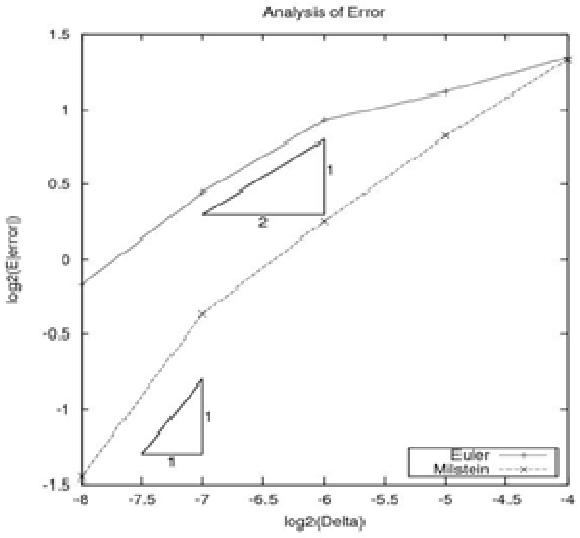
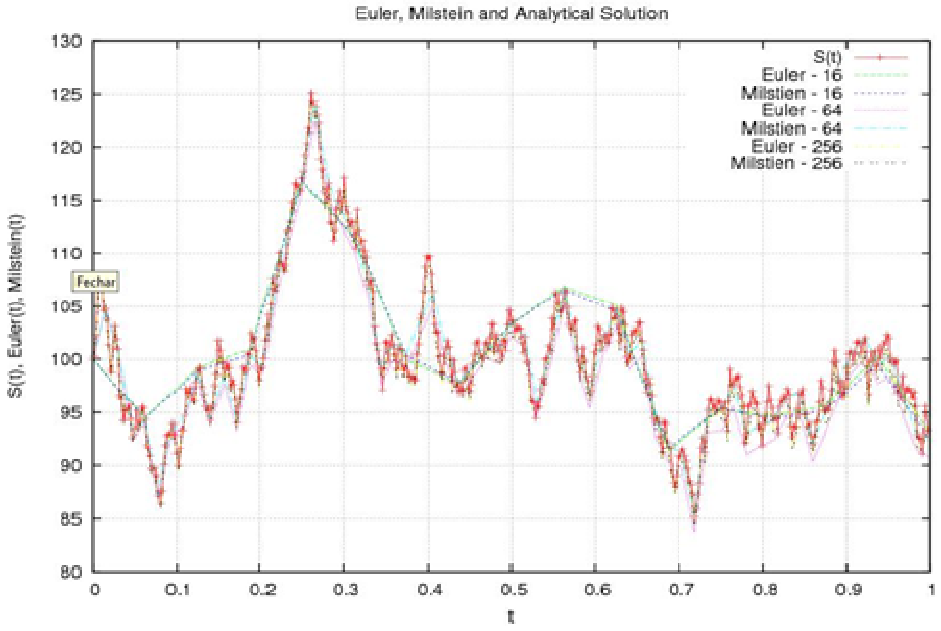


Figure 2 - Analytical solution and Euler-Maruyama and Milstein approximations



A new example also comparing the methods presented, but using the parameters: h = 0.001, $\mu = 0.30$, $\sigma = 0.50$ and $S_0 = 150$. Also adopted was a relationship of Δ^γ, for time steps equal 16, 32, 64, 128 and 256.

Figure 3 - Analysis of Error for Euler-Maruyama and Milstein

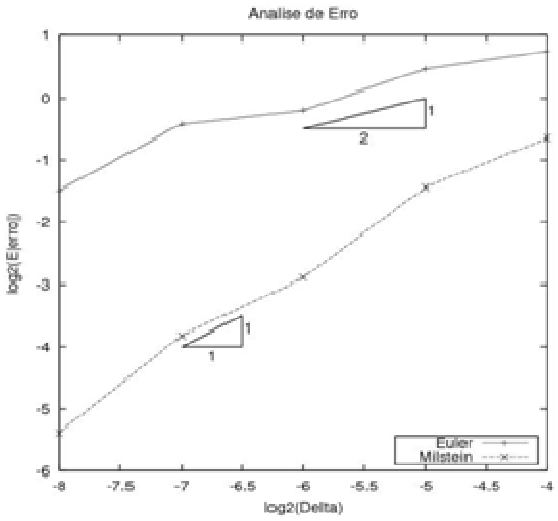
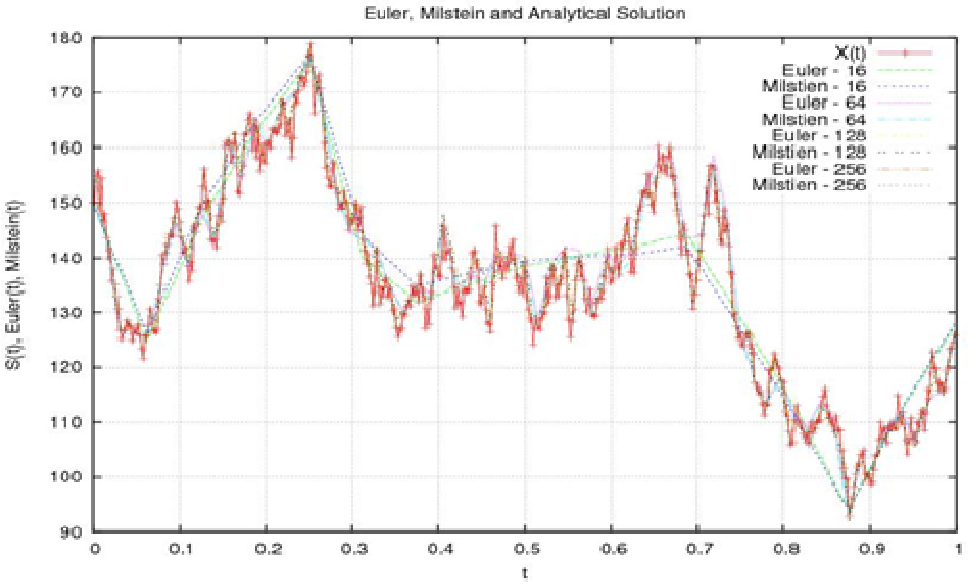


Figure 4 - Analytical solution and Euler-Maruyama and Milstein approximations



The figures (1) and (3) shows the graph of convergence for approximation of the Brownian geometric motion, geometric stochastic differential equation. It confirmed the convergence with the rate of 1/2 for Euler-Maruyama approximation and 1 for Milstein approximation, as expected in theory, as can be seen by the slope of line corresponding to the graph $\log_2(\Delta)\times\log_2(E|error|)$ presented. Already in figures (2) and (4), we have samples trajectories of each simulation.

5 . 2 – Simulation main finance sthochastic models

First example of simulation of sample paths is linear SDE with constant coefficients that is shown in Figure (5). In computer simulation we use the values: $\mu = 0.01$, $\sigma = 0.03$ and h = 0.001. The initial condition as $S_0 = 100$.

From the random number generator of the 1000 values were obtained independently. Since the random variables normally distributed with zero mean and variance 0.001. The fact that W_t in (12) is a Martingale allows the use of independent random variables.

A discrete approximations of (12) is used to obtain the S_t are shown in Figure (5), which were calculated to K = 1,2, … , 1000,

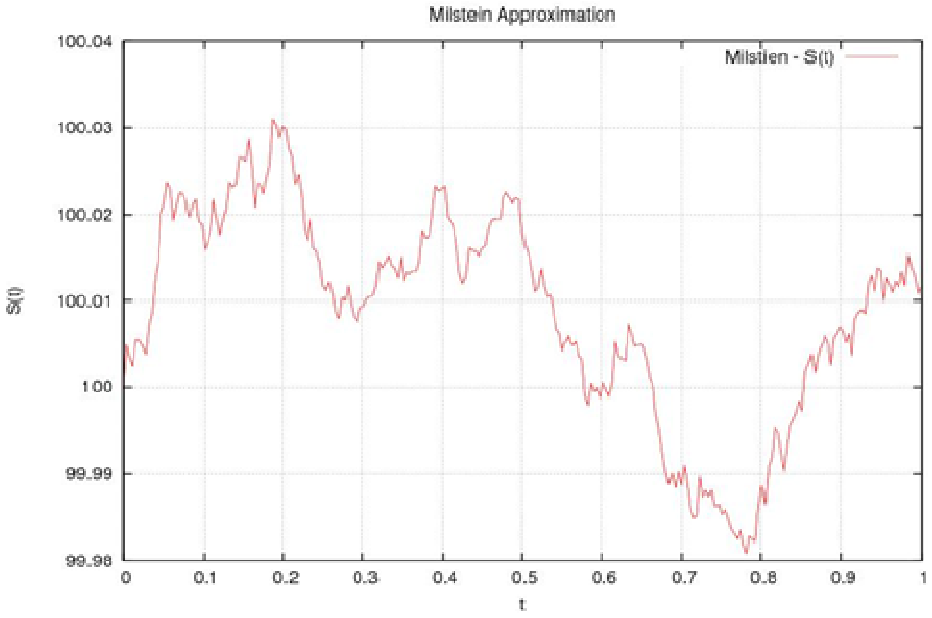
$$S_K = S_{K+1} + 0.01\,(0.001) + 0.03\,\Delta W_k + \frac{1}{2}\times 0.0\times 0.03\left\{(\Delta W_k)^2 - (0.001)\right\}$$

$$S_K = S_{K+1} + 0.01\,(0.001) + 0.03\,\Delta W_k$$

We find that the Milstein approximation in SDE is equal to the Euler-Maruyama approximation, as $G(t,S_t)=\sigma$ soon $G'(t,S_t)=0.0$ canceling the additional term of the Milstein approximation.

Given an initial condition S_0, replaced ΔW_t from normally distributed random numbers getting the S_t successively.

Figure 5 - Milstein approximation for linear SDE with constant coefficients



As can be seen, S_t seems to fluctuate around a straight line with slope μ. The value of σ determine the extent of fluctuations around this line. Note that these fluctuations do not become larger as time passes. This suggests that SDE is appropriate in practice. In particular, the SDE is a good approximation if the "variation" in prices of assets is stable over time and if not appear systematic jumps in the prices of assets.

5 . 2 . 1 – Geometric stochastic differential equation

Figure (6) shows a realization of S_t obtained for finite difference approximation of

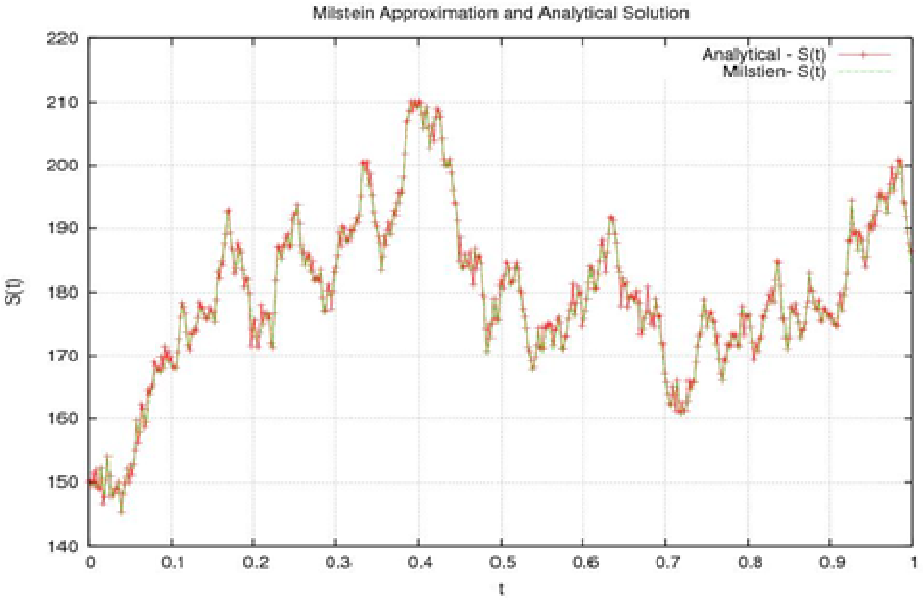
$$dS_t = 0.15\,S_t dt + 0.30\,S_t dW_t,$$

formulated as

$$Y_{n+1} = Y_n + 0.15\left(\tau_{n+1} - \tau_n\right) + 0.30 Y_n \Delta W_n + \frac{1}{2} \times 0.30^2 Y_n \left\{\left(\Delta W_n\right)^2 - \left(\tau_{n+1} - \tau_n\right)\right\}$$

with initial condition $S_0 = 150$. As can be seen in the graph, the S_t is made of two components. First, there is a exponential trend of growth of 15%. The second, a random fluctuation around this trend. These changes increase with over time because of higher prices.

Figure 6 - Milstein approximations and analytical solution for underlying assets pricing.



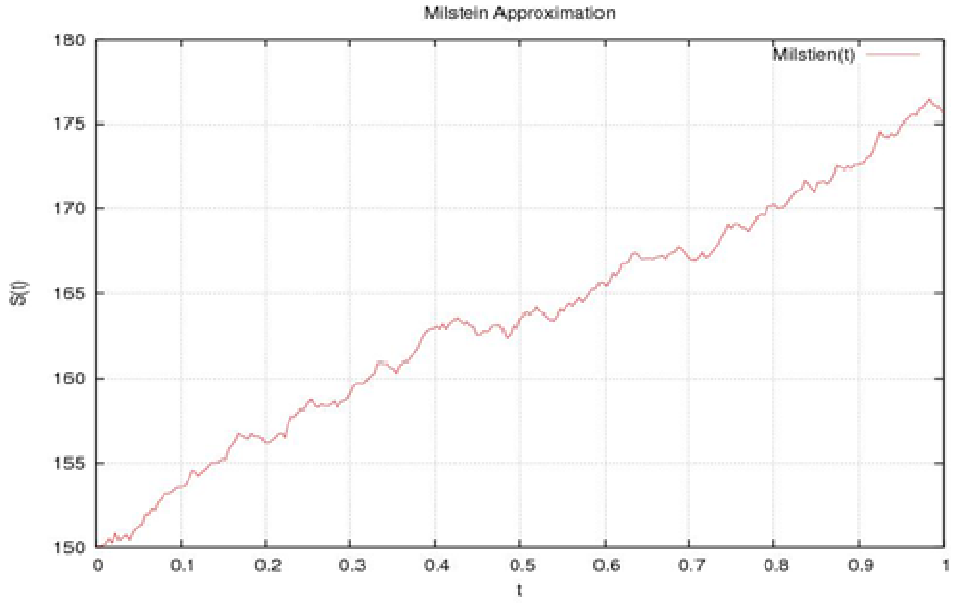
5 . 2 . 2 – Square root process

For example, in the Figure (7) present the sample trajectory to the same terms of dW_t used to generate the Figure (6). Considering the equation

$$dS_t = 0.15 S_t dt + 0.30\sqrt{S_t}\, dW_t,$$

where the coefficients of drift and diffusion are as in Figure (6), but now the diffusion is proportional to $\sqrt{S_t}$ instead of being proportional to S_t. Select is the initial condition as $S_0 = 150$.

Figure 7 - Milstein approximations for square root process



Clearly, the fluctuations in Figure (7) are more dominated than in Figure (6), however the sample trajectory shows "similar" trends.

Another characteristic of this process is the significance of the parameter σ. It is the specification of the component of diffusion, the σ can not be interpreted as the volatility of S_t. But the markets, on another line, have as the convention σ as the volatility of percent of the underlying assets.

5 . 2 . 2 – Mean reverting process

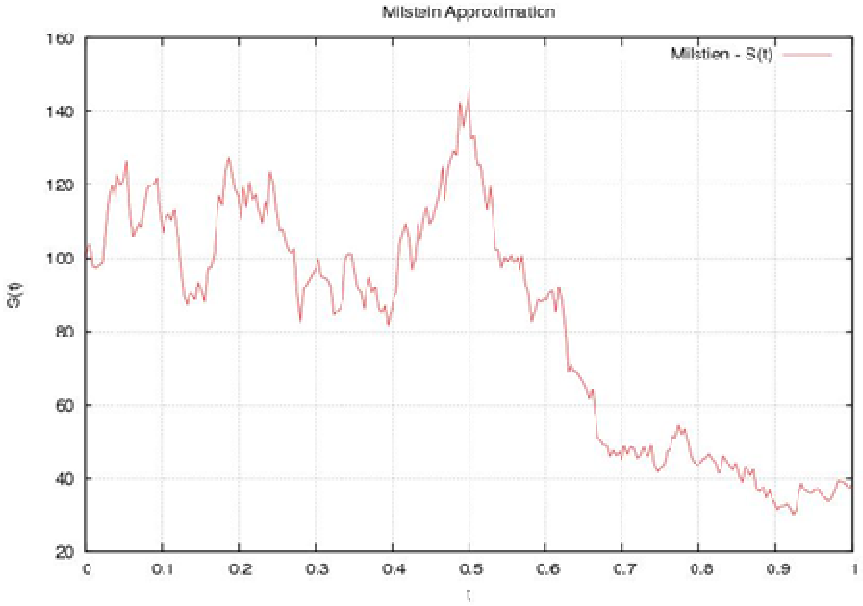
An example of the sample trajectory of a mean reverting process is shown in Figure (8). Selecting $\mu = 0.05$, $\lambda = 0.5$, $\sigma = 0.8$. This implies an average of 5% and a volatility of 80% over a time interval of length 1, and the λ implies an adjustment of 50%.

We adopt the length of the finite intervals as h = 0.001. According to this, during a time interval of length 1, observed the 1000 values of S_t.

The random numbers with zero mean and variance 0.001 was obtained, and the sample path was generated using the increments for k = 1, …, 1000, the initial condition used was $S_0 = 100$, for (14) we have the following approximation

$$\Delta S_K = 0.5\,(0.05 - S_{K-1})(0.001) + 0.8\,\Delta W_k + \frac{1}{2}\times 0.8^2 \times S_K\,\left\{(\Delta W_k)^2 - (0.001)\right\},$$

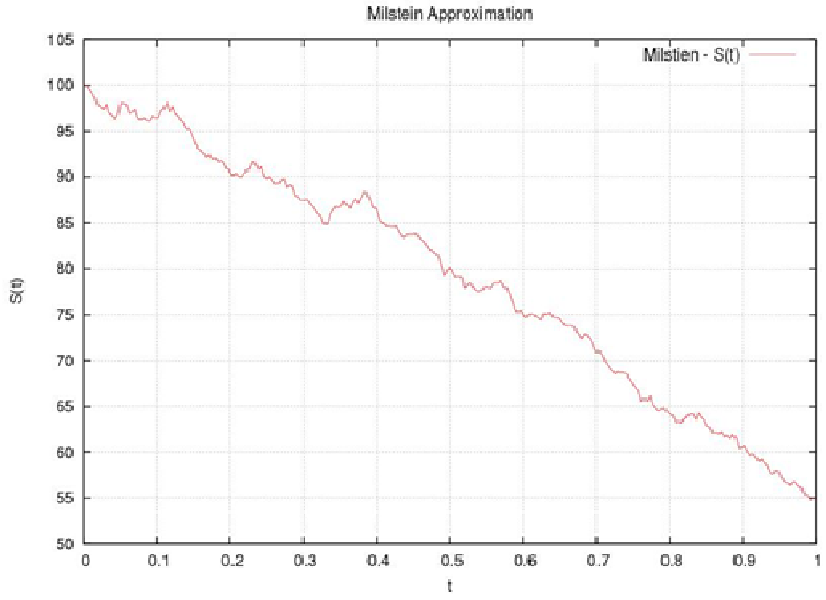
Figure 8 - Milstein approximations for mean reverting process in eq. (14).



For (15), with the same values in the parameters, we have the following approximation

$$\Delta S_K = 0.5\,(0.05 - S_{K-1})(0.001) + 0.8\,\Delta W_k + \frac{1}{2}\times 0.8^2\,\left\{(\Delta W_k)^2 - (0.001)\right\}$$

Figure 9 - Milstein approximations for mean reverting process in eq. (15).



6 – CONCLUSION

This paper presents an application example of the numerical method, finite differences method, a problem of stochastic differential equations applied to financial engineering.

It presented a study of stochastic mathematical modeling of stochastic differential equations and its numerical solution in the discrete time approximation, by finite differences and also was presented the study of convergence for stochastic differential equations in two types of methods, which confirmed the theoretical rates of convergence presented in the literature, for simulation of one assets in financial market, modeled by Brownian geometric motion.

As was also presented the mathematical stochastic models, stochastic differential equations, and their numerical solution in the discrete time approximation, by finite differences, Milstein numerical methods for the main types of stochastic differential equations used in modeling financial market, as an example of application direct stochastic differential equations in the problem of economic origin.

Studies of various problems modeled by stochastic differential equations applied to Financial Engineering, and application of other numerical methods with higher rates of convergence presented here, are motivated for future work.

ACKNOWLEDGMENTS

The author thanks the financial support received by *USA Contractors, INC* for implementation of this paper. The use of facilities of Group for Monetary and Financial Studies, GEMF, of Faculty of Economics of the University of Coimbra. And also to all teachers of Faculty of Economics of the University of Coimbra.

REFERENCES

ARNOLD, L. (1974), "Stochastic Differential Equations: Theory and Applications", John Wiley & Sons.

BLACK, F., and SCHOLES, M. (1973), "The Pricing of Options and Corporate Liabilities", Journal of Political Economy, Vol. IV, pp. 323-338.

ITÔ, K. (1951), "On stochastic differential equations", American Math. Soc., nº 4, New York.

KLOEDEN, P.E., and PLATEN, E. (1995), "Numerical Solution of Stochastic Differential Equations", Number 23 in Applications of Mathematics, Stochastic Modelling and Applied Probability, Springer.

MERTON, R.C. (1973), "Theory of rational option pricing". The Bell Journal of Economics and Management Science, Vol. IV, nº 1, pp. 141-183.

THOMAZ, J.C.A. and LOULA, A.F.D. (2007), "Numerical solution of stochastic differential equation in finance", In Third Brazilian Conference on Statistical Modelling in Insurance and Finance, pp. 344-347.

ESTRUCTURA DE TIPOS EN ESPAÑA

PEDRO A PÉREZ PASCUAL
E-mail: pperez@cee.uned.es
BASILIO SANZ CARNERO
E-mail: bsanz@cee.uned.es
Departamento de Economía Aplicada Cuantitativa I
Universidad Nacional de Educación a Distancia

Resumen

En este trabajo utilizamos el análisis de cointegración de Campbell y Shiller (1987) para tratar de contrastar la teoría de las expectativas de los tipos de interés en España en el periodo 1994 – 2005. El análisis se lleva a cabo siguiendo tanto un enfoque basado en una sola ecuación, como un análisis basado en el empleo de vectores autorregresivos. En ambos casos se ha obtendo evidencia empírica favorable.

Palabras clave: Estructura temporal de tipos, Cointegración, VAR
Area temática: Modelación Económica y Financiera.

Abstract

In this paper we use the cointegration analysis framework proposed by Campbell and Shiller (1987) in order to test the Expectation Theory of Term Structure in Spain during the period 1994-2005. We do this in a univariate as well as a multivariate VAR framework, finding some evidence that supports the theoretical hypothesis.

Keywords: Term Structure, Cointegration, VAR
Thematic Area: Economic and Financial Modelling.

1 - INTRODUCCIÓN

Los modelos de valor presente, constituyen una importante herramienta en la econometría financiera. Estos modelos relacionan el precio de un activo con sus flujos de caja esperados descontados. Desde un punto de vista económico esta forma de valorar un activo es más apropiada que otras, si bien al implicar agentes racionales y expectativas de futuro, resulta difícil de implementar en la práctica.

Campbell y Shiller (1987) han propuesto utilizar el análisis de cointegración para tratar de solventar en parte este problema. El modelo desarrollado por estos dos autores puede utilizarse para contrastar la teoría de las expectativas de la estructura temporal de tipos lo que ha dado lugar a una instrumento útil para el estudio de la teoría de las expectativas y la racionalidad en los mercados financieros.

Diversos autores han intentado contrastar empíricamente la validez de la teoría utilizando datos españoles. Por ejemplo, Prats y Beyaert (1998) sometieron a contraste la teoría de las expectativas de los tipos de interés en el periodo 1986-1992, hallando evidencia positiva. Massot y Nave (2003) proporcionan también evidencia favorable para el periodo 1993-1998, mientras que Esteve (2006) utiliza una aproximación no lineal para alcanzar resultados parecidos.

En este trabajo utilizamos una muestra más amplia de datos que se extiende como hemos dicho desde enero de 1994 hasta diciembre de 2005.

Suponiendo que el tipo de descuento sea constante a través del tiempo, el valor presente de un activo vendrá dado por,

$$Y_t = \sum_{i=0}^{\infty} \frac{1}{(1+r)^i} E_t(y_{t+i}) \qquad (1)$$

donde Y_t representa el valor presente del activo, y_{t+1} es la renta derivada de la posesión de dicho activo en el periodo comprendido entre (t, t+1), E_t es el operador expectativas (esperanza) y r el tipo de descuento.

En este modelo Y_t puede representar el precio de un activo e y_t su dividendo; o Y_t el rendimiento a largo plazo e y_t el correspondiente a corto (estructura de tipos). Incluso con ligeras modificaciones, el modelo anterior puede utilizarse para medir la teoría de la renta permanente (Campbell y Shiller, 1987).

Restando en ambos lados el término $Y_t / (1+r)$ se llega a,

$$Y_t - \frac{y_t}{r} = \frac{1+r}{r} \sum_{i=0}^{\infty} \frac{1}{(1+r)^i} E_t(\Delta y_{t+i}) \qquad (2)$$

El término $Y_t - (y_t / r)$ recibe el nombre de *spread* y expresaría la diferencia entre el valor del activo y su renta descontada. De esta ecuación se deduce que, si las series Y e y son integradas de orden uno, el *spread* debe ser estacionario, lo que significa que ambas deben estar cointegradas, siendo el vector de cointegración (1 – 1/r).

Este modelo se ha modificado (ver p.e., Wang, 2003) para tener en cuenta la posibilidad de que las variables experimenten una tasa de crecimiento g, dando lugar a,

$$Y_t - \frac{(1+g)y_t}{r-g} = \frac{1+r}{r-g} \sum_{i=1}^{\infty} \frac{(1+g)^i (E_t \Delta u_{t+i})}{(1+r)^i} \qquad (3)$$

donde $\Delta u_{t+i} = \Delta \ln y_{t+i} - g$. El término $Y_t - \frac{(1+g)y_t}{r-g}$ recibe ahora el nombre de *spread completo*. En el supuesto de existiese una tasa de crecimiento constante en la renta, el modelo (2) no sería estacionario, pero sí lo sería el (3): es decir el vector de cointegración sería ahora $\left(1 - \frac{1+g}{r-g}\right)$[1].

La ecuación (2) puede escribirse también de la forma,

$$S_t = Y_t - \frac{y_t}{r} = \frac{1+r}{r} E_t \Delta Y_{t+1} \qquad (4)$$

Si hay una burbuja racional (*rational bubble*) en el mercado b_t, tal que $b_t = \frac{1}{1+r} E_t b_{t+1}$, es decir, $b_{t+1} = (1+r)b_t + e_t$, donde e_t es i.i.d.(0, σ^2), en la ecuación (1), entonces aparecerá en el segundo miembro de (2), pero no en el de (4). El término b_t tiene una raíz fuera del círculo unidad y es por tanto explosiva (no estacionaria). En consecuencia incluso si Δy_t es estacionario, el spread será no estacionario si hay una burbuja racional en (1) que induce la no estacionaridad de ΔY_t a través de (5). Por tanto contrastar la racionalidad equivale a contrastar la cointegración

[1] De hecho (3) es el modelo de valoración de Gordon para crecimientos constantes perpetuos.

entre Y_t e y_t. La cointegración implica la estacionaridad del *spread* y la inexistencia de burbujas racionales en el mercado.

Las ecuaciones (2) y (4) sugieren también una forma alternativa de análisis utilizando los VAR. Sea $\mathbf{z}_t = [S_t \ldots S_{t-p+1} \ \Delta y_t \ \ldots \ \Delta y_{t-p+1}]'$. Entonces puede formularse un modelo VAR de la forma,

$$\begin{bmatrix} \mathbf{S}_t \\ \Delta \mathbf{y}_t \end{bmatrix} = \begin{bmatrix} \mathbf{a}(L) & \mathbf{b}(L) \\ \mathbf{c}(L) & \mathbf{d}(L) \end{bmatrix} \begin{bmatrix} \mathbf{S}_{t-1} \\ \Delta \mathbf{y}_{t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \qquad (5)$$

donde $u_{1t} = [v_{1t}\ 0 \ldots 0]'$, $u_{2t} = [v_{2t}\ 0 \ldots 0]'$, y los vectores **a, b, c** y **d,** están definidos de forma que el VAR tuviese sólo las ecuaciones correspondientes a S_t y Δy_t. Es decir,

$$\begin{bmatrix} S_t \\ \vdots \\ S_{t-p+1} \\ \Delta y_t \\ \vdots \\ \Delta y_{t-p+1} \end{bmatrix} = \begin{bmatrix} a_1 & \ldots & a_p & b_1 & \ldots & b_p \\ 1 & & & & & \\ & \ddots & & & & \\ & & & & & \\ c_1 & \ldots & c_p & d_1 & \ldots & d_p \\ & & & 1 & & \\ & & & & \ddots & \end{bmatrix} \begin{bmatrix} S_{t-1} \\ \vdots \\ S_{t-p} \\ \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p} \end{bmatrix} + \begin{bmatrix} v_{1t} \\ 0 \\ \vdots \\ v_{2t} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad (6)$$

donde los espacios en blanco son ceros. En forma compacta podemos escribir la ecuación como,

$$\mathbf{z}_t = \mathbf{A}\mathbf{z}_{t-1} + \mathbf{u}_t \qquad (7)$$

La implicación de las ecuaciones anteriores es que el spread, S_t, debe causar en el sentido de Granger a Δy_t a menos que S_t sea una combinación lineal exacta de valores corrientes y retardados de Δy_t. Por tanto S_t debería tener la capacidad de mejorar la predicción de Δy_t. Además, sean $\mathbf{e}_1$' y $\mathbf{e}_2$' dos vectores fila (1x2p) con ceros en todas las celdas excepto 1 en el primer elemento para $\mathbf{e}_1$' y en el p+1 elemento para $\mathbf{e}_2$', es decir,

$$S_t = \mathbf{e}_1' \mathbf{z}_t \qquad (8)$$

$$\Delta y_t = \mathbf{e}_2' \mathbf{z}_t \qquad (9)$$

Obsérvese que,

$$E[\mathbf{z}_{t+k} / H_t] = \mathbf{A}^k \mathbf{z}_t \qquad (10)$$

donde H es el conjunto de información disponible sobre S_t y Δy_t. Sustituyendo las ecuaciones (8) y (9) en (2), se tiene,

$$\mathbf{e}_1' \mathbf{z}_t = \frac{1}{r} \sum_{i=1}^{\infty} \frac{1}{(1+r)^i} \mathbf{e}_2' \mathbf{A}^i \mathbf{z}_t = \mathbf{e}_2' \frac{1}{r(1+r)} \mathbf{A} \left[\mathbf{I} - \frac{1}{1+r} \mathbf{A} \right]^{-1} \mathbf{z}_t \qquad (11)$$

La ecuación (11) impone restricciones sobre los parámetros del VAR, necesarias para que la racionalidad se mantenga,

$$\mathbf{e}_1' \left[\mathbf{I} - \frac{1}{1+r} \mathbf{A} \right] = \mathbf{e}_2' \frac{1}{r(1+r)} \mathbf{A} \qquad (12)$$

En consecuencia el *spread* "teórico" debe ser,

$$S_t^* = \mathbf{e}_2' \frac{1}{r(1+r)} \mathbf{A} \left[\mathbf{I} - \frac{1}{1+r} \mathbf{A} \right]^{-1} \mathbf{z}_t \qquad (13)$$

Puede demostrarse que la diferencia entre el spread corriente y el teórico es,

$$S_t - S_t^* = \sum_{i=0}^{\infty} \frac{1}{(1+r)^i} E(\xi_{t+i} / H_t) \qquad (14)$$

donde $\xi_t = Y_t - [(1+r)Y_{t-1} - y_t] = Y_t - E_{t-1}Y_t$ representa la innovación en la predicción de Y_t. Contrastar las restricciones (12) es equivalente a contrastar que la parte derecha de la ecuación (14) es ruido blanco con media cero. Es decir $\xi_t = S_t - (1/\delta)S_{t-1} + \theta \Delta y_t$ ha de ser ruido blanco.

Asimismo, usando un test de volatilidad, el ratio var(S_t)/var(S_t*) no debería ser significativamente mayor que la unidad, si se ha de mantenerse el modelo de valor presente. Este test puede implementarse también con la innovación ξ_t y su esperanza matemática donde,

$$\xi_t^* = S_t^* - (1+r)S_{t-1}^* + \frac{1}{r}\Delta y_t \qquad (16)$$

y,

$$\xi_t = S_t - (1+r)S_{t-1} + \frac{1}{r}\Delta y_t \qquad (17)$$

El ratio var(ξ)/var(ξ*) puede contemplarse como la "innovación en el ratio de varianza" mientras que var(S_t)/var(S_t*) sería el ratio de varianza en niveles.

Las implicaciones del análisis anterior pueden resumirse de la siguiente manera. Si el mercado de un activo es racional, entonces las variables valor/precio y renta deberían estar cointegradas y por tanto el *spread* debería ser estacionario. Sin esa relación de cointegración el spread será no estacionario, y existirá en ese mercado una "burbuja racional" que por definición es explosiva. Si el mercado es eficiente y se mantiene el modelo de valor presente, entonces el spread teórico no debería diferir sistemáticamente del corriente, y el ratio de sus varianzas no debería ser significativamente diferente de la unidad. La potencia predictiva del spread sobre Δy_t está condicionada al conjunto de información de los agentes. Si los agentes no disponen de información útil para predecir Δy_t, entonces S_t será una combinación lineal de valores presentes y retardados de Δy_t sin capacidad predictiva. La predicción puede o puede que no resulte mejorada por el hecho de que las variables precio y renta estén cointegradas. Por tanto se usa la cointegración entre esas variables como criterio para evaluar la racionalidad del mercado contra la posibilidad de que existan burbujas en el mismo. Además utilizamos la representación VAR y los ratios de varianza derivados del sistema, para examinar si es posible mantener el modelo de valor presente y en qué medida el mercado está siendo eficiente.

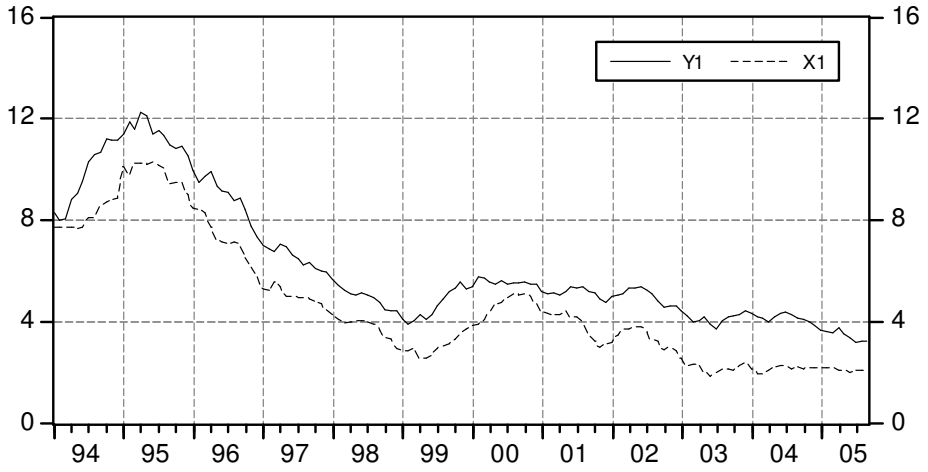
2 - ANÁLISIS EMPÍRICO

Utilizaremos el modelo expuesto en el apartado anterior para contrastar la *teoría de las expectativas en la estructura de tipos* en España. Emplearemos datos mensuales entre enero de 1994 y septiembre de 2005. Los datos de Y_t representan los rendimientos a largo plazo de la deuda pública, en tanto que los de y_t (que a partir de ahora será x_t) son los rendimientos de las

Letras del Tesoro a un año. Ambas series han sido obtenidas de la base de datos que el Banco de España tiene en su página web (http://www.bde.es).

La figura 1 ofrece la representación gráfica de ambas variables,

Figura 1. Representación gráfica de *Y* (tipos a largo) y *X* (tipos a corto)



Puesto que el análisis está basado en el supuesto de que las series originales son integradas de orden 1, la tabla 1 muestra los resultados de someter a contraste este supuesto,

Tabla 1. Contrastes de estacionaridad

	ADF		PP		KPSS	
	Sin tend. (-2.88)	Con tend. (-3.44)	Sin tend. (-2.88)	Con tend. (-3.44)	Sin tend. (0.46)	Con tend. (0.15)
Y	-0.93	-2.35	-0.77	-1.83	1.03	0.22
X	-0.91	-1.41	-1.10	-1.76	1.07	0.19

Con objeto de asegurar todo posible la conclusión respecto a la hipótesis de raíz unitaria en niveles, se han utilizado tres procedimientos diferentes. El primero es el habitual test de Dickey y Fuller aumentado (ADF), el segundo es el contraste alternativo propuesto por Phillips y Perron (PP), y el tercero es el contraste de Kwiatkowski, Phillips, Schmitdt y Shin. En este último la hipótesis nula es la estacionaridad en lugar de la no estacionaridad, como sucede en los dos anteriores. Los tres contrastes se han implementado sin y con tendencia, figurando entre paréntesis los valores críticos para un nivel del confianza del 95%. Los resultados no dejan lugar a dudas: en los dos primeros test, el resultado del estadístico de contraste es siempre mayor (menos negativo) que los valores críticos, tanto cuando se ha considerado una tendencia como cuando se ha excluido ésta, de manera que la hipótesis nula (raíz unitaria) no puede ser rechazada. Por el contrario en el caso del test KPSS, la hipótesis nula (la estacionaridad) sí queda claramente rechazada, al ser el estadístico de contraste superior en todos los casos, al valor crítico. En resumen, tanto *Y* como *x* son series no estacionarias.

La tabla 2 muestra este mismo análisis para las primeras diferencias de ambas series,

Tabla 2. Contraste de raíces unitarias para ΔY y ΔX.

	ADF		PP		KPSS	
	Sin tend. (-2.88)	Con tend. (-3.44)	Sin tend. (-2.88)	Con tend. (-3.44)	Sin tend. (0.46)	Con tend. (0.15)
ΔY	-8.52	-8.50	-9.10	-9.09	0.10	0.11
ΔX	-8.98	-8.95	-9.68	-9.65	0.09	0.07

Claramente la conclusión es ahora la contraria, es decir las primeras diferencias de ambas series son I(0), o sea, estacionarias.

La *teoría de las expectativas de la estructura de tipos* implica que las series Y_t y X_t deben estar cointegradas siendo además conocido el vector de cointegración. Si consideramos las variables en desviaciones con respecto a sus medias, entonces el vector de cointegración entre y_t y x_t será $(1\ -1)$[2]. Por tanto el denominado *spread* representado por $s_t = y_t - x_t$ debe ser estacionario.

La estacionaridad queda confirmada por el análisis estadístico de la tabla 3, donde se muestran los resultados de los contrastes de raíces unitarias aplicados ahora al *spread* (sin tendencia pero con término independiente),

Tabla 3. Contrastes de raíces unitarias de S_t

	Spread, S_t
ADF (-2.88)	- 3.79**
PP (-2.88)	- 3.90**
KPSS (0.46)	0.15**

La cointegración puede comprobarse empíricamente, obteniéndose unos resultados muy próximos a los postulados por la teoría. Así la ecuación de cointegración ahora en niveles, proporciona,

$$\underset{(18.53)}{\hat{Y}_t} = \underset{(60.04)}{1.602 + 0.984X_t}, \quad R^2 = 0.96,\ DW = 0.32,\ SSR = 31.87,\ N = 141 \qquad (18)$$

es decir que el parámetro correspondiente a *X* está muy cercano a la unidad. Un contraste basado en el ratio *t* de que el verdadero valor del parámetro es la unidad, arroja un resultado de *t* = -0.947, de manera que no podría rechazarse esa hipótesis a ninguno de los niveles de confianza habituales. El trazado de los residuos de la ecuación (18) es idéntico al de la figura 2.

El contrate de cointegración a partir de un modelo de corrección de error, proporciona,

$$\underset{(1.60)}{\Delta\hat{Y}_t} = \underset{(1.84)}{0.105} + \underset{(8.56)}{0.131\Delta Y_{t-1}} + \underset{(-1.85)}{0.624\Delta X_{t-1}} - \underset{(1.78)}{0.064Y_{t-1}} + 0.061X_{t-1}, \quad R^2 = 0.45$$

La hipótesis de que el parámetro correspondiente a Y_{t-1} sea nula puede ser rechazada al 6,6 %, de manera que la cointegración puede aceptarse a ese nivel de confianza. El vector de cointegración implícito es muy similar: (1 – 0.953). Una conclusión idéntica se deriva de la aplicación del más sofisticado procedimiento de contrastación de Johansen para la cointegración. Los valores de los estadísticos de la traza y máximo, son respectivamente, λ_{traza} = 13.5 y λ_{max} = 12.46. Estos valores no son significativos al 5% pero sí lo serían al 10 %, siendo el vector de cointegración (1 – 1.008).

En resumen, hay evidencias de que las series en cuestión están cointegradas y de que el vector de cointegración tiene un valor muy cercano al postulado por la teoría de las expectativas. Este resultado descarta además la posibilidad de una burbuja racional en el mercado, dado que la existencia de dicho bucle causaría la no estacionaridad de S_t y ΔX_t, circunstancia podemos descartar basándonos en el análisis empírico. Sin embargo las evidencias anteriores no son concluyentes.

La formulación VAR con dos retardos, queda (estadísticos *t* entre paréntesis),

[2] En otro caso el vector de cointegración incluirá una constante que representa la prima por la espera derivada de invertir en bonos a largo en lugar de hacerlo a corto.

$$S_t = \underset{(4.24)}{0.3} + \underset{(10.3)}{1.15}S_{t-1} - \underset{(-2.99)}{0.336}S_{t-2} + \underset{(5.56)}{0.693}\Delta X_{t-1} - \underset{(-3.59)}{0.401}\Delta X_{t-2},\ R^2 = 0.75,\ SSR = 7.3966$$

$$\Delta X_t = \underset{(-3.11)}{-0.206} + \underset{(0.612)}{0.064}S_{t-1} + \underset{(0.512)}{0.054}S_{t-2} + \underset{(1.87)}{0.219}\Delta X_{t-1} + \underset{(1.42)}{0.149}\Delta X_{t-2},\ R^2 = 0.15,\ SSR = 6.5089$$

La capacidad explicativa de las variables retardadas sobre ΔX es aceptable (15%) y a juzgar por los valores de los estadísticos *t*, parece que ΔX tiene poder explicativo sobre *S* pero no a la inversa. Sin embargo el contraste de causalidad de Granger no avala esta impresión, toda vez que,

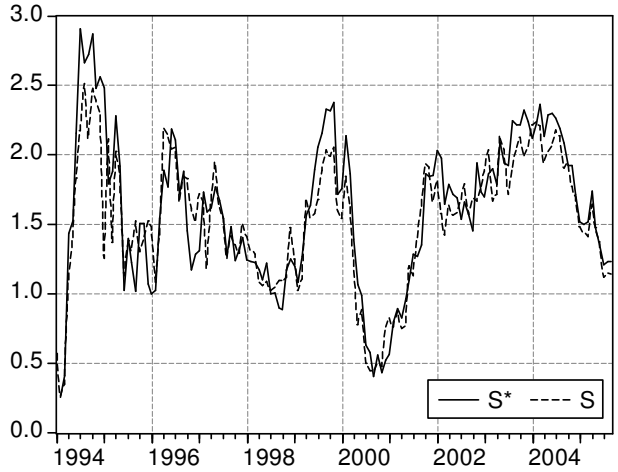
Tabla 4. Contraste de causalidad de Granger

Pairwise Granger Causality Tests			
Sample: 1994:01 2005:09			
Lags: 2			
Null Hypothesis:	Obs	F-Statistic	Probability
ΔX does not Granger Cause *S*	139	15.5632	8.4E-07
S does not Granger Cause ΔX		4.09708	0.01875

de manera que habría causalidad bidireccional, y el spread tendría poder de predicción sobre el cambio de los tipos a corto.

Finalmente, por lo que se refiere a las restricciones de la ecuación (12), calculamos *S** a partir de (13) y posteriormente se lleva a cabo un contraste de igualdad de varianzas[3]. La figura 2 muestra la representación gráfica de *S* y *S**,

Figura 2. Representación gráfica del spread (*S*) y el spread teórico (*S**)



Aunque hay algunas discrepancias, aparentemente el movimiento es muy parecido, lo que implicaría una evidencia empírica favorable a la teoría. Esta impresión resulta confirmada por diversos test de igualdad de varianzas, como se muestra en la tabla siguiente,

[3] El valor de 1/1+*r* se ha calculado a partir de la ecuación (4) en Campbell y Shiller (1987, p. 1066).

Tabla 5. Test de igualdad de varianzas

	(Y_t, X_{1t})		(Y_t, X_{2t})	
F-test	1.38	(0.061)*	1.21	(0.26)**
Siegel-Tuckey	2.26	(0.026)	1.44	(0.15)**
Barlett	3.50	(0.061)*	1.24	(0.26)**
Levene	3.66	(0.057)*	0.06	(0.81)**
Brown-Forsythe	4.24	(0.051)*	0.07	(0.80)**

3 - CONCLUSIÓN

El objeto de de este trabajo ha sido contrastar la teoría la teoría de las expectativas de los tipos de interés en España, siguiendo la aproximación sugerida por Campbell y Schiller. Hemos llevado a cabo la contrastación empleando tanto un enfoque univariante como utilizando un VAR y, en ambos casos, hemos obtenido evidencia que apoya la teoría en el periodo considerado, lo que excluye la existencia de burbujas racionales en el mercado.

BIBLIOGRAFÍA

AYUSO, J., NOVALES, A. AND DE LA TORRE, M. L. (1992), ¿Incorporan los tipos del interbancario una evaluación del riesgo? *Revista Española de Economía,* 9, 343-379.

ALONSO, F., AYUSO, J. AND MARTÍNEZ J. (2001), How informative are financial asset prices in Spain? *Revista de Economía Aplicada*, 25, 5-38.

BEAULIEU, J. J. AND MIRON, J. A. (1993), Seasonal Unit Roots in Aggregated US Data. *Journal of Econometrics*, 55, pp. 305-328.

BEYAERT, A., GARCÍA, J. G. AND PRATS, M. A. (2001), La estructura temporal y las expectativas de tipos en el Mercado interbancario español, *Moneda y Crédito*, 213, 71-96.

CAMPBELL, J. Y. AND SHILLER, R. J. (1987), Cointegration and Tests for the Present Value Models, *Journal of Political Economy*, 95, 1062-1088.

CAMPBELL, J. Y. AND SHILLER, R. J. (1991), Yield Spreads and Interest Rates Movements: A Bird Eye View, *The Review of Economic Studies*, 58, 495-514.

ESTEVE, V. (2006), A note of nonlinear dynamics in the Spanish Term Structure of interest rates, *International Review of Economics and Finance*, 15 (3), 316-323.

GARRET, J. (2006), The liquidity effect across the short end of the term structure, *Applied Financial Economics Letters* 3, 159-163

KWIATKOWSKI, D, PHILLIPS, P.C.B.; SCHMIDT, P (1992), Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54,

MARTÍN, A. M., PÉREZ DE VILLARREAL, J. M. (1990), La estructura temporal de los tipos de interés: el mercado español de depósitos interbancarios¸ *Moneda y Crédito*, 191, 173-193.

MASSOT, M. AND NAVE, J. (2003), La hipótesis de las expectativas en el largo plazo: evidencia en el Mercado español de deuda pública, *Investigaciones Económicas*, 27(3), 553-564.

NOVELL, J. (2002), Ciclo de la Economía española y contenido informativo de los tipos de interés, *Estudios de Economía Aplicada*, 20(3), 583-598.

PATTERSON, K. (2000), An Introduction to Applied Econometrics. MaCMillan Press Ltd.

PÉREZ, J. V., SÁEZ, M. AND MURILLO, C. (1997), Expectativas y volatilidad condicionada. Los tipos de interés en el mercado interbancario, *Revista de Economía Aplicada*, 5, 83-107.

PERRON, P AND NG, S (1996), Useful modifications to unit root tests with dependent errors and their local asymptotic properties. *Review of Economic Studies*, 63, 435-65.

PRATS, M. A. AND BEYAERT, A. (1998), Testing the expectation theory in a market of short-term financial assets, *Applied Financial Economics*, 8, 101-109.

ROBLES, M. D. AND FLORES DE FRUTOS, R. (2000), Time varying term premia and risk: the case of Spanish interbank money market, *Applied Financial Economics,* 10, 243-260.

TSUJI, C. (2005) Does the tern structure predict real economic activity in Japan? *Applied Financial Economics Letters*, 1, 249-257

WANG, P. (2003), Financial Econometrics. Methods and Models. Routledge, London.

COINTEGRACIÓN POR UMBRALES, AJUSTES NO LINEALES Y ASIMETRÍAS EN LA TRANSMISIÓN DE PRECIOS: UNA APLICACIÓN AL SECTOR OVINO ESPAÑOL

MONIA BEN KAABIA
E-mail: monia@unizar.es
ANTONIO MONTAÑÉS
IRENE OLLOQUI
Departamento de Análisis Económico
Universidad de Zaragoza

Resumen

El objetivo de este trabajo consiste en analizar el proceso de transmisión de precios a lo largo de la cadena comercial en el sector del ovino en España. Para ello se han considerado el precio percibido por el productor y el pagado por el consumidor. El enfoque metodológico adoptado se basa en la estimación de un Vector de Corrección del Error con umbrales donde tanto el proceso de ajuste a largo plazo como la relación de cointegración a largo plazo siguen un proceso por umbrales. Los resultados indican que, a largo plazo, las elasticidades de transmisión dependen de la magnitud del margen de comercialización en el período anterior. Asimismo, en el corto plazo las reacciones de los precios, tanto en velocidad como en magnitud, son asimétricas.

Palabras Clave: Asimetrías, cointegración por umbrales, precios, ovino.
Area temática: Modelação Económica e Financeira.

Abstract

This paper aims to investigate the non-linear adjustments of prices between farm and retail prices in the lamb meet sector in Spain. The methodology used is based on the multivariate approach to specify and estimate a Threshold Autoregressive Model. The results indicate that, in the long run, price transmission is asymmetric. In the short-run, price adjustments between the farm and the retail levels are asymmetric.

Thematic Area: Economic and Financial Modelling.

1. INTRODUCCIÓN

El análisis de la transmisión de precios a lo largo de la cadena comercial es un tema que ha despertado un gran interés entre los economistas agrarios. Los precios son, quizás, el principal mecanismo que permite relacionar los diferentes eslabones de la cadena comercial, por lo que el análisis de la transmisión de precios se ha utilizado como medio para evaluar el funcionamiento general de los mercados. Este interés ha aumentado en los últimos años como consecuencia de los cada vez más complejos procesos de concentración, tanto horizontal como vertical, que han tenido lugar en los mercados agroalimentarios. En este contexto, los estudios sobre transmisión de precios pueden considerarse como una respuesta de los economistas agrarios ante la creciente preocupación social y pública sobre las consecuencias de dichos procesos de concentración.

Uno de los aspectos a los que se ha prestado más atención en los últimos años es a la posible existencia de asimetrías en los mecanismos de transmisión de precios a lo largo de la cadena de comercialización. El concepto de "asimetría" se asocia tanto a la magnitud como a la velocidad de respuesta de los diferentes precios. Aunque la existencia de asimetrías en la transmisión de

precios no tiene por qué asociarse a mercados ineficientes, lo que sí parece claro es que las respuestas rápidas y simétricas de los precios antes shocks inesperados, tanto de oferta como de demanda, suelen considerarse como representativas de mercados eficientes.

Dado que la teoría tradicional de los mercados agrarios asume la existencia de ajustes simétricos en precios, el conocimiento que se ha ido obteniendo sobre la posible existencia de asimetrías en la transmisión de precios procede de la acumulación de análisis empíricos. A partir de dichos trabajos, son numerosas las posibles causas que los investigadores han identificado para explicar la existencia de relaciones de precios asimétricas [Meyer y von Cramon-Taubadel (2004)]. Entre las más comunes, podemos destacar: el poder de mercado del sector detallista [Bailey y Brorsen (1989); Griffith y Piggott (1994); entre otros] [1]; los costes de transacción a nivel detallista [Chavas y Mehta (2002)]; el grado de sustitutibilidad de inputs en el proceso productivo [Bettendorf y Verboven (2000)]; el nivel de inventarios tanto a nivel productor como detallista (Reagan y Witzman, 1982); y la intervención pública [Kinnucan and Forker (1987)].

En cualquier caso, la literatura empírica no ofrece resultados concluyentes en una determinada dirección. Las conclusiones obtenidas difieren de un trabajo a otro dependiendo del producto analizado, del país de estudio, de la frecuencia de los datos, del período muestral y, sobre todo, del método de análisis empleado. En este sentido, y teniendo en cuenta los desarrollos que habían tenido lugar en las últimas dos décadas en el ámbito del análisis de series temporales, Von Cramon-Taubadel (1998) demostró, en un artículo pionero en este campo, que el enfoque tradicionalmente utilizado hasta la fecha, basado en los trabajos de Wolffram (1971), Houck (1977) y Ward (1982), no era compatible con la posible existencia de precios no estacionarios y cointegrados, especificando un modelo en forma de corrección del error asimétrico que era capaz de recoger la posible existencia de asimetrías tanto en el corto como en el largo plazo. Sin embargo, la existencia de costes de transacción puede generar ajustes en precios no lineales, que no pueden ser recogidos por un modelo lineal de corrección del error. Esta limitación ha dado lugar al desarrollo de propuestas alternativas capaces de compatibilizar la existencia de ajustes asimétricos y la presencia de relaciones de cointegración entre las variables. Entre estos, los más utilizados son: el modelo autorregresivo con regímenes cambiantes Markovianos (Markov Switching Autoregressive Model) (MS-AR) [Hansen (1997)], el modelo autorregresivo por umbrales (Threshold Autoregressive Model) (TAR) [Balke y Fomby (1997; Obstfeld y Taylor (1997; Goodwin y Piggott (2001)], y el Vector de Corrección del Error por umbrales (Threshold Vector Error Correction Model) (TVECM) [Ben Kaabia y Gil (2007)]; Lo and Zivot (2001; Goodwin and Piggott (2001; y Hansen y Seo (2002)]. Sin embargo, las aplicaciones los modelos de corrección del error por umbrales (TVECM)[2] en la mayoría de los trabajos empíricos, los procesos asimétricos se consideraban únicamente en los procesos de ajustes a corto plazo hacia el equilibrio a largo plazo, suponiendo una relación de cointegración a largo plazo entre los precios lineal. Mantener la hipótesis de linealidad en la relación de cointegración a largo plazo, permitiendo únicamente ajustes asimétricos (no lineales) a corto plazo no siempre son económicamente y econométricamente justificables. De hecho, podrían existir diferentes factores que hacen que la relación de equilibrio a largo plazo

[1] A pesar de que en la mayor parte de los trabajos, la existencia de asimetrías en la transmisión de precios se ha relacionado con comportamientos no competitivos, McCorriston et al. (2001) demostraron que esto no tenía por qué ser cierto siempre y que, en todo caso, había que tener en cuenta la relación entre poder de mercado y la existencia de economías de escala. Si la función de costes del detallista se caracteriza por la existencia de rendimientos de escala decrecientes, el efecto del poder de mercado se refuerza, provocando ajustes asimétricos. Sin embargo, en el caso de que la función de costes se caracterice por la existencia de rendimientos de escala crecientes, el efecto del poder de mercado se compensa, pudiendo aparecer ajustes simétricos.

[2] Los modelos TAR y TVECM se han convertido en un procedimiento muy popular dado que son relativamente más simples de especificar, estimar e interpretar que muchos otros modelos de series temporales no lineales (Hansen, 1997). El modelo TAR se basa en una estrategia en dos etapas, similar a la propuesta de Engle y Granger (1987), en el caso del análisis de cointegración. Asimismo, asume la existencia de una relación causal entre la variable dependiente y la variable independiente. El modelo TVECM genera estimaciones más robustas que el modelo TAR (Hansen y Seo, 2002) a la vez que considera como endógenas a las dos series de precios incluidas en el modelo. Una excelente descripción de los modelos de cointegración por umbrales puede encontrarse en Gonzalo y Pitarakis (2006).

no sea lineal, sino que dependen de que se den o no determinados condiciones en el mercado. Por ejemplo si las cualidades del producto o las condiciones de oferta o demanda cambian, pueden hacer que los precios cambian su relación de cointegración o hasta dejar de serlo. Gonzalo y Pitarakis (2006) proponen un estadístico que permite contrastar la hipótesis nula de cointegración lineal frente a la alternativa de cointegración por umbrales *(Threshold cointegration)*. Rechazar la hipótesis nula significa que la relación de cointegración entre las variables no estacionarias depende de la magnitud de una variable umbral que provoca la transición de un régimen a otro. En este sentido, Gonzalo y Pitarakis (2006) señalan que el uso del termino de "Cointegración por umbrales" en la especificación del TVECM es engañoso dado que estos modelos solamente permitan que los términos de corrección del error están sujetos a los efectos umbrales, mientras que la relación de cointegración se supone constante y lineal. En este trabajo, la transmisión de precios en el sector ovino en España se va a realizar a partir de la especificación y estimación de un modelo TVECM, cuya descripción presentamos en los párrafos siguientes.

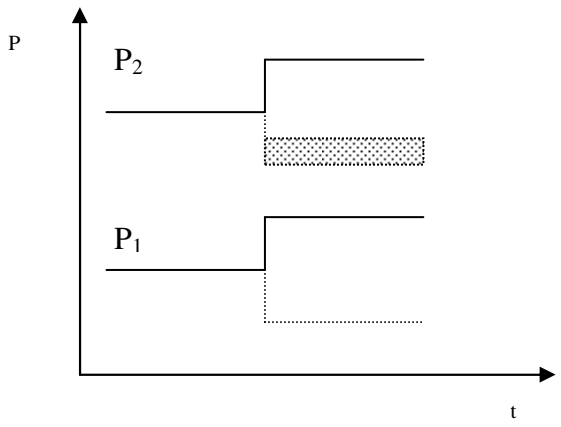
El objetivo de este trabajo consiste en analizar, utilizando el enfoque de cointegración por umbrales (*threshold cointegration)*, el mecanismo de transmisión de precios, tanto a corto como a largo plazo, en el sector del ovino en España. En concreto, partiendo de una situación de equilibrio en la que se produzca un shock inesperado en cualquiera de los dos extremos de la cadena comercial (productor – consumidor), enfocaremos nuestro trabajo hacia cuatro aspectos: 1) la magnitud de las respuestas en los dos niveles de la cadena; 2) la velocidad de ajuste hacia el equilibrio tras el shock en los dos extremos; 3) la naturaleza simétrica o asimétrica de la respuesta; y 4) la dirección de la respuesta, esto es, si los ajustes son diferentes dependiendo si el shock se transmite hacia arriba o hacia abajo de la cadena comercial.

2. CONCEPTO DE ASIMETRÍA Y POSIBLES CAUSAS

En términos generales podemos afirmar que se producen asimetrías en la transmisión de precios a lo largo de la cadena comercial cuando las respuestas de los precios ante shocks inesperados positivos o negativos son diferentes. En este sentido, las asimetrías pueden caracterizarse de acuerdo a la magnitud, velocidad o naturaleza de las respuestas.

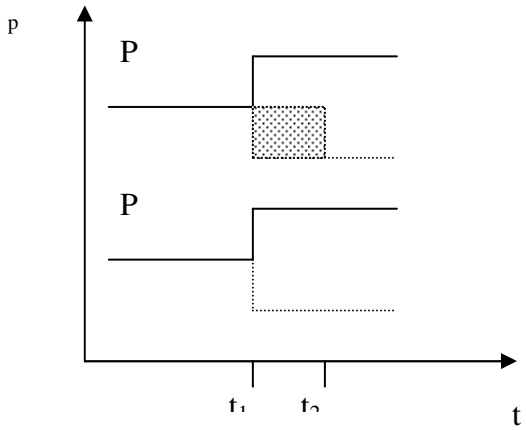
Tratemos de explicar este concepto con unos gráficos simples. Supongamos que tenemos dos precios P_1 (asociado, por ejemplo, al coste de un input) y P_2, asociado a la venta de un output. Diremos que existe asimetría en magnitud cuando la respuesta de P_2 ante un shock en P_1 es diferente dependiendo de si el shock inicial es positivo o negativo. Por ejemplo, como se puede apreciar en la Figura 1 cuando el shock en P_1 es positivo la respuesta de P_2 es de la misma magnitud. Sin embargo, cuando el shock es negativo, la respuesta es de menor intensidad, generándose, en este caso, asimetría positivas (aumenta el diferencial de precios).

Figura 1. Asimetrías en magnitud.



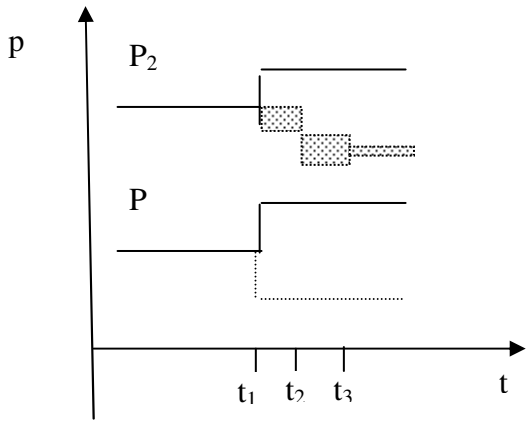
Por otro lado, diremos que existen asimetrías en velocidad cuando el número de períodos que se tarda en reaccionar es diferente según se trate de shocks positivos o negativos. La Figura 2 ilustra un ejemplo de este caso. Como se puede apreciar, la magnitud de la respuesta de P_2 ante un shock en P_1 es de la misma magnitud. Sin embargo, si el shock es negativo se tarda más tiempo en ajustarse que cuando el shock es positivo. En este caso, también se están produciendo asimetrías positivas.

Figura 2. Asimetrías en velocidad de ajuste.



Finalmente, se pueden generar múltiples combinaciones de relaciones asimétricas. La Figura 3 recoge un ejemplo posible. En el caso que se presenta se producen asimetrías tanto en magnitud como en velocidad.

Figura 3. Transmisión asimétrica por magnitud y velocidad.



En efecto, un incremento de P_1 genera una respuesta inmediata y de la misma magnitud de P_2. Sin embargo, cuando el shock es negativo, por un lado, el ajuste no es instantáneo y, por otro, ni siquiera es de la misma magnitud, sino que es menor. También en este caso, se generan asimetrías positivas. En caso que en las tres figuras anteriores las respuestas ante shocks positivos fueran de menor magnitud o velocidad que ante shocks negativos, se generarían asimetrías negativas.

Dado que la teoría tradicional de los mercados agrarios asume la existencia de ajustes simétricos en precios, el conocimiento que se ha ido obteniendo sobre la posible existencia de asimetrías en la transmisión de precios procede de la acumulación de análisis empíricos. A partir de dichos trabajos, son numerosas las posibles causas que los investigadores han identificado para explicar la existencia de relaciones de precios asimétricas (Meyer y von Cramon-Taubadel, 2004). Entre las más comunes, podemos destacar el poder de mercado y los costes de transacción, entre otros.

El poder de mercado del sector detallista (Bailey y Brorsen, 1989; Griffith y Piggott, 1994; Borestein et al.,1997; y Bettendorf y Verboven, 2000; entre otros). Sin emabrago, a pesar de que en la mayor parte de los trabajos, la existencia de asimetrías en la transmisión de precios se ha relacionado con comportamientos no competitivos, McCorriston et al. (2001) demostró que esto no tenía por qué ser cierto siempre y que, en todo caso, había que tener en cuenta la relación entre poder de mercado y la existencia de economías de escala. Si la función de costes del detallista se caracteriza por la existencia de rendimientos de escala decrecientes, el efecto del poder de mercado se reforzaba, provocando ajustes asimétricos. Sin embargo, en el caso de que la función de costes se caracterice por la existencia de rendimientos de escala crecientes, el efecto del poder de mercado se compensa, pudiendo aparecer ajustes simétricos.

Los costes de transacción a nivel detallista (Blinder et al., 1998; Buckle y Carlson, 2000; y Chavas y Mehta, 2004). A diferencia de otros productos, en el caso de la alimentación, la elección principal es la del establecimiento en el que hacer la compra. Si los precios no varían significativamente, el consumidor realiza la mayor parte de su compra en dicho establecimiento. Un concepto relacionado con este último son los denominados Costes de Menú, estos es los costes de promoción que realizan la mayor parte de los detallistas. A no ser que el incremento del precio del producto en cuestión sea muy significativo, a los detallistas no les interesa retirar su publicidad para introducir nuevos precios.

El grado de sustitutibilidad de inputs en el proceso productivo (Bettendorf y Verboven, 2000). Cuanto mayor sea el número de productos sustitutivos mayor probabilidad de que se generen relaciones asimétricas. El nivel de inventarios tanto a nivel productor como a nivel detallista (Reagan y Witzman, 1982). A medida que el nivel de almacenamiento es mayor, mayor es la probabilidad de que aparezcan relaciones de precios asimétricas. Finalmente, la intervención pública puede generar también la aparición de rigideces en la transmisión de precios (Kinnucan y Forker, 1987).

3. ENFOQUE METODOLÓGICO

Tradicionalmente, el análisis de la transmisión de precios, en la que explícitamente se considera la posibilidad de existencia de ajustes asimétricos entre los diferentes precios, se ha realizado empleando la metodología propuesta por Houck (1977) y Ward (1982), a partir de un trabajo anterior de Wolffram (1971). Dicho enfoque constituye el marco adecuado para analizar el mecanismo de transmisión de precios a lo largo de la cadena comercial siempre que las series de precios sean estacionarias. Sin embargo, la mayor parte de las series económicas son no estacionarias (su varianza no es invariante en el tiempo), si bien pueden existir entre ellas relaciones estables en el largo plazo (cointegración). Ender y Granger (1998) propusieron una metodología de cointegración que permitía ajustes no lineales y asimétricos. Basándose en esta idea, Balke et al. (1998); Lo y Zivot (2001) y Hansen y Seo (2002) propusieron, bajo la hipótesis de series cointegradas linealmente, especificar un Vector de Mecanismo de Corrección del Error por Umbrales (TVMCE). Este modelo constituye el marco adecuado para contrastar la posible presencia de relaciones asimétricas entre precios a lo largo de la cadena comercial.

Una de las limitaciones más importantes que presenta este enfoque es que asume que la relación de cointegración debe ser lineal, permitiendo únicamente que los ajustes en precios hacia el equilibrio a largo plazo sean no lineales. Sin embargo, la presencia de costes de transacción fijos puede generar reacciones a largo plazo no lineales. Esta limitación ha dado lugar al desarrollo de propuestas alternativas capaces de compatibilizar la existencia de ajustes asimétricos y la presencia de relaciones de cointegración no lineales entre las variables. Recientemente, Gonzalo y Pitarakis (2006) desarrollaron un contraste que permite contrastar la hipótesis nula de cointegración lineal frente a la alternativa de cointegración no lineal o por umbrales *(Threshold cointegration)*.

Para la aplicación de esta metodología, primero se debe contrastar el orden de integración de las variables. Si las variables resultan ser integradas de orden uno, a continuación, se aplica el procedimiento de Gonzalo y Pitarakis (2006) para contrastar la hipótesis nula de cointegración lineal frente a la alternativa de cointegración por umbrales.

3.1 Contrastes de cointegración no lineal

Una vez comprobado que las series de precios (precio al consumo (P^c) y el precio al productor (P^p)[3]) son I(1) mediante algunos de los procedimientos sugeridos por la literatura, tradicionalmente se procede a contrastar si existe cointegración entre las series de precios:

$$p_t^c = \beta_0 + \beta_1 p_t^p + u_t \quad \textbf{(1)}$$

Un problema importante en la estimación de la regresión de cointegración expuesta en (1) es el denominado "problema de normalización". Los resultados, tanto sobre la evidencia de cointegración como sobre la estimación del vector de cointegración, variarán dependiendo de cuál sea la variable elegida como "dependiente" en la regresión de cointegración. Este problema puede subsanarse parcialmente mediante la estrategia multivariante propuesta por Johansen. Para contrastar la existencia de una relación de cointegración por umbrales, Gonzalo y Pitarakis (2006) plantean una regresión alternativa a la expuesta en (1):

$$p_t^c = \beta_0 + \beta_1 p_t^p + (\lambda_0 + \lambda_1 p_t^p) I(q_{t-d} > \gamma) + \varepsilon_t \quad \textbf{(2)}$$

donde: q_{t-d} es la variable umbral estacionaria y retardada d periodos y que se utiliza como referencia para definir los dos regímenes, γ es el parámetro umbral que permite diferenciar entre los dos regímenes; (β_0, β_1) y (λ_0,λ_1) son los parámetros de la relación de cointegración en cada uno de los regímenes, y $I_t(q_{t-d}>)$ es indicador de la función tal que:

$$I_t = \begin{cases} 1 \text{ si } q_{t-d} > \gamma \\ 0 \text{ si } q_{t-d} \leq \gamma \end{cases} \quad \textbf{(3)}$$

A partir de (2) la hipótesis nula para contrastar si existe relación de cointegración lineal frente a la alternativa de cointegración por umbrales viene dada por:

$$\textbf{H}_0\textbf{:}\ \lambda_0 = \lambda_1 = 0 \textbf{, frente a la alternativa } \textbf{H}_a\textbf{:}\ \lambda_0 \neq 0 \text{ y } \lambda_1 \neq 0 \quad \textbf{(4)}$$

Para contrastar dicha hipótesis nula Gonzalo y Pitarakis (2006) proponen utilizar el siguiente estadístico

$$LM(\gamma) = \frac{U'MX_\gamma (X_\gamma' MX_\gamma)^{-1} X_\gamma' MU}{\tilde{\sigma}_0^2} \quad \textbf{(5)}$$

donde $M = I - X(X'X)^{-1}X'$ siendo X es la matriz de observaciones de los regresores en el modelo (1); U es el vector de perturbaciones del modelo (1); $X_\gamma = X * I_t(q_{t-d} > \gamma)$, y $\tilde{\sigma}_0^2$ es el estimador MCO de la varianza de la perturbación aleatoria del modelo (1).

El Estadístico $LM(\gamma)$ se calcula para todos los valores posibles de la variable umbral (q_{t-d}). Además para garantizar un número mínimo de observaciones en cada régimen, se fija un parámetro trimming π_0 de tal forma que $T_i / T \geq \pi_0$ donde T_i es el número de observaciones existentes en el régimen i-ésimo y T es el número total de observaciones. Andrews (1993) sugiere tomar valores de π_0 en el intervalo (0,05; 0,15). Finalmente el estadístico SupLM viene dado por la siguiente expresión:

$$SupLM = \sup_{\lambda \in [\lambda^L, \lambda^U]} LM(\lambda) \quad \textbf{(6)}$$

Dado que, bajo la hipótesis nula, en el modelo no lineal aparecen ciertos parámetros que no están restringidos y que no se encuentran en el modelo lineal (1), el estadístico (3) no sigue una

[3] A efectos de la explicación metodológica, nos vamos a centrar en la relación entre precio percibido por el productor y precio pagado por el consumidor. Evidentemente, la metodología expuesta puede extenderse a cualquiera de las relaciones expuestas en el apartado anterior.

distribución estándar (Davies, 1987; Hansen, 1999 y Hansen y Seo, 2002). Gonzalo y Pitarakis (2006) sugieren el uso de los valores críticos publicado en Andrews (1993). Los valores finalmente estimados de los umbrales son aquellos que optimizan la siguiente función:

$$\hat{\gamma} = \underset{\gamma \in [\gamma^L, \gamma^U]}{SupLM} \qquad \textbf{(7)}$$

El Rechazo de la hipótesis nula dada en (3) implica la existencia de una relación de cointegración por umbrales (no lineal) entre las dos series de precios. En principio, la magnitud de los parámetros ((β_1 y λ_1) de la regresión (2) da idea del grado de transmisión de precios, a largo plazo, entre los dos eslabones comerciales en cada uno de los dos regímenes. A este respecto hay que mencionar que, a pesar de que el estimador MCO de estos parámetros es superconsistente, la no estacionariedad de las variables impide hacer inferencias. Es por ello, que el análisis de cointegración debe completarse con la estimación del Vector de Mecanismo de Corrección del Error (VMCE), ya que ésta puede aportar una interpretación más rica y completa del comportamiento dinámica a corto plazo de las variables objeto de análisis. A pesar de estas ventajas, en los trabajos empíricos no se dedicó mucha atención a estos aspectos, especialmente, si nos movemos en un concepto de cointegración no lineal (cointegración por umbral).

3.2 Vector de mecanismo de corrección del error

En la mayoría de los trabajos dentro de la línea de cointegración no lineales, la representación del VMCE en el sentido de Granger, se hizo bajo el supuesto de que la relación de cointegration era por sí misma lineal y que la no-linealidad afectaba únicamente a los procesos de ajustes hacia el equilibrio a largo plazo (Lo y Zivot; Hansen y Seo, 2002). Otra suposición mantenida importante en esta línea de investigación era la existencia de un solo vector de cointegrating (see Balke y Fomby, 1997; Seo y Hansen, 2002; Seo, 2004).

Un sistema de variables cointegradas linealmente pero que presenta un proceso de ajustes no lineales, admite fácilmente la especificación de un Vector en forma de Modelo de Corrección no lineal (por umbrales) y que se puede escribirse como:

$$\Delta P_t = \begin{cases} \alpha^1 \omega_{t-1}(\beta) + \sum_{i=1}^{k-1} \Gamma_i^1 \Delta P_{t-i} + \varepsilon_t^1, & \text{si} \quad \omega_{t-1}(\beta) < \lambda^1 \\ \alpha^2 \omega_{t-1}(\beta) + \sum_{i=1}^{k-1} \Gamma_i^2 \Delta P_{t-i} + \varepsilon_t^2, & \text{si} \quad \lambda^1 \le \omega_{t-1}(\beta) \le \lambda^2 \\ \alpha^3 \omega_{t-1}(\beta) + \sum_{i=1}^{k-1} \Gamma_i^3 \Delta P_{t-i} + \varepsilon_t^3, & \text{si} \quad \omega_{t-1}(\beta) > \lambda^2 \end{cases} \qquad \textbf{(8)}$$

Esta representación se denomina Vector de Mecanismos de Corrección de Error con efecto umbral (Threshold Vector Error Correction model TVECM). Sin embargo, Gonzalo y Patarakis (2006) demostraron que un sistema de variables que se caracterizan por la existencia de una relación de cointegración por umbrales (no lineal), no puede admitir una representación de VECM en el sentido de Granger de la forma tradicional. Esto se debe a que no es posible obtener un modelo que garantiza que todos sus elementos sean estacionarios.

En este trabajo para poder estimar un VMCE que nos permita analizar la dinámica de ajuste de precios a largo plazo, se plantea una estrategia en dos etapas parecida a la de Engle y Granger (1987)[4]. Si existe suficiente evidencia a favor de un vector de cointegración por umbrales, el planteamiento es el siguiente. En una primera etapa, se trata de contrastar que el vector de cointegración en cada uno de los dos regímenes es estacionario con el fin de garantizar que los precios estén cointegrados en ambos regímenes. Una vez se demuestra que las variables están cointegradas en ambos regímenes. A continuación especificamos y estimamos un VMCE

[4] Si no se rechaza la hipótesis nula de cointegración lineal, a continuación, contrastamos si los procesos de ajustes a corto plazo son lineales frente a la alternativa de no lineales utilizando el procedimiento de Lo y Zivos (2001).

por umbrales (TVECM) siguiendo el procedimiento de Lo y Zivot (2001) y Hansen y Seo (2002):

$$\Delta P_t = \begin{cases} \alpha^1 \hat{\omega}_{t-1} + \sum_{i=1}^{k-1} \Gamma_i^1 \Delta P_{t-i} + \varepsilon_t^1, & \text{si} \quad q_{t\text{-}d} \le \hat{\lambda} \\ \alpha^2 \hat{\omega}_{t-1} + \sum_{i=1}^{k-1} \Gamma_i^2 \Delta P_{t-i} + \varepsilon_t^3, & \text{si} \quad q_{t\text{-}d} > \hat{\lambda} \end{cases} \quad \textbf{(9)}$$

donde :

$$\hat{\omega}_t = p_t^c - (\hat{\beta}_0 + \hat{\beta}_1 p_t^p) I_t(q_{t\text{-}d} \le \hat{\gamma}) + (\hat{\lambda}_0 + \hat{\lambda}_1 p_t^p)(1\text{-}I_t(q_{t\text{-}d} \le \hat{\gamma}) \quad \textbf{(10)}$$

Teniendo en cuenta que el valor umbral es conocido (estimado en la etapa anterior (ecuación 3)), el resto de parámetros del TVECM se pueden estimar por Mínimos Cuadrados. Finalmente, se puede llevar a cabo diferentes contrastes para determinar estadísticamente si, efectivamente, el modelo no lineal estimado se ajusta mejor al mecanismo de transmisión de precios que un modelo lineal. Para ello, podemos utilizar el estadístico de Wald tradicional con el objetivo de contrastar las siguientes hipótesis:

$$H_0 : \alpha^1 = \alpha^2 \Rightarrow \text{proceso de ajuste lineal}$$
$$H_0 : \alpha^1 \neq \alpha^2 \Rightarrow \text{proceso de ajuste no lineal} \quad \textbf{(11)}$$

$$H_0 : \Gamma_i^1 = \Gamma_i^2 \quad \text{i=1,..,k} \Rightarrow \text{dinámica a corto plazo lineal}$$
$$H_0 : \Gamma_i^1 \neq \Gamma_i^2 \Rightarrow \text{dinámica a corto plazo no lineal} \quad \textbf{(12)}$$

La estimación de este modelo resulta económicamente interesante. Su interés radica en que recoge la estructura dinámica del sistema de precios en los dos regímenes a corto y largo plazo. El número de retardos y la magnitud de sus coeficientes dan cuenta de la gradualidad con la que se produce el ajuste de los precios, dependiendo del régimen en el que nos encontramos. En todo caso, los parámetros realmente relevantes en los modelos de cointegración por umbrales son los coeficientes estimados de la matriz α asociados a la relación de equilibrio a largo plazo. El signo y la magnitud de estos coeficientes (en cada uno de los regímenes) nos permiten analizar la velocidad de ajuste de las variables dependientes hacia el equilibrio a largo plazo. En este sentido, las asimetrías pueden caracterizarse de acuerdo a la magnitud y significatividad de los parámetros de ajustes en cada uno de los regímenes considerado. Diremos que existe asimetría en magnitud cuando el proceso de ajuste de los precios es diferente en los dos regímenes. Por otro lado, diremos que existen asimetrías en velocidad cuando el número de períodos que se tarda en restablecer el equilibrio a largo plazo es diferente según se trate del régimen 1 o 2.

4. APLICACIÓN EMPÍRICA AL SECTOR DE LA CARNE DE OVINO

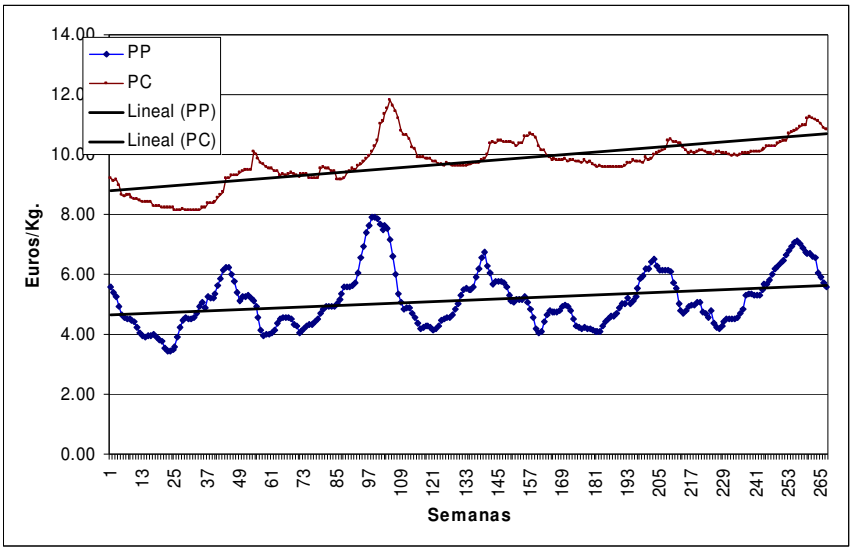
4.1 Datos y propiedades estocásticas de las series

En este apartado nos vamos a centrar en el análisis del proceso de transmisión de precios desde el origen (producción) hasta el consumo final. Con el fin de analizar el mecanismo de transmisión vertical de precios en el sector de la carne de cordero en España, se han considerado el precio percibido por el productor (PP) y el precio pagado por el consumidor (PC), ambos expresados en €/Kg. El período muestral se extiende desde Enero de 2004 hasta diciembre de 2008 considerando una frecuencia semanal. Ambas series se han obtenido del Observatorio de Precios del Ministerio de Medio Ambiente y Medio Rural y Marino (MARM). El precio al productor se refiere a animales vivos, mientras que el precio al consumo se refiere a cordero pascual 1ª. Hay que tener presente que es necesario ser riguroso al hacer comparaciones y análisis de las diferencias de precios entre distintos eslabones de la cadena comercial. No se debe confundir el margen bruto (diferencias entre dos precios) con el margen neto o beneficio. Que sería el resultado de restar al margen bruto los costes directos e indirectos asociados a los

sucesivos procesos de comercialización. El margen, por lo tanto, es estos eslabones, engloban distintos procesos necesarios todos ellos en el transcurso de la comercialización.

El Gráfico 1 recoge la evolución del precio al productor (PP) y al consumo (PC) a lo largo de las 265 semanas consideradas. Como se puede apreciar, ambos precios han experimentado un ligero crecimiento a lo largo del período considerado. Asimismo, se puede comprobar la existencia de un marcado componente estacional coincidente con el ciclo productivo del sector ovino

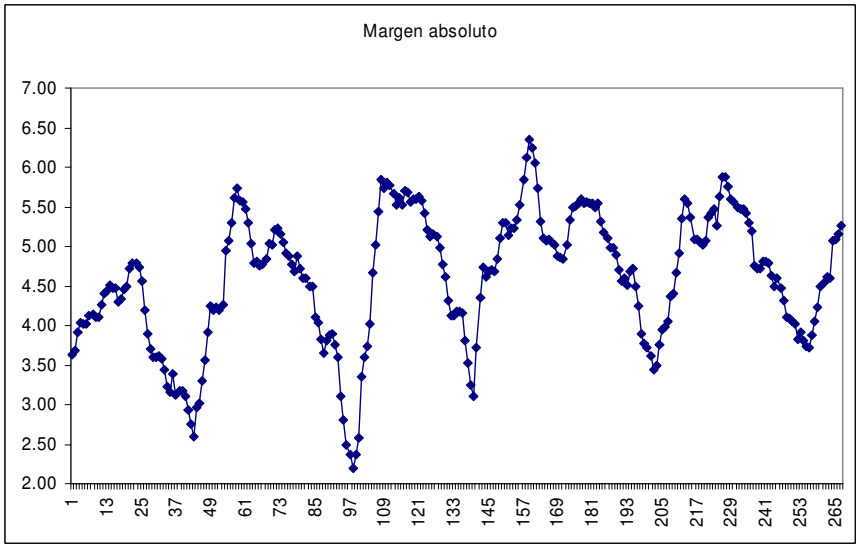
Gráfico 1. Evolución de los precios al productor y al consumo de carne de cordero en España (€/Kg).



Nota: PP = Precio Productor; PC = Precio Consumo
Fuente: Observatorio de Precios (MARM)

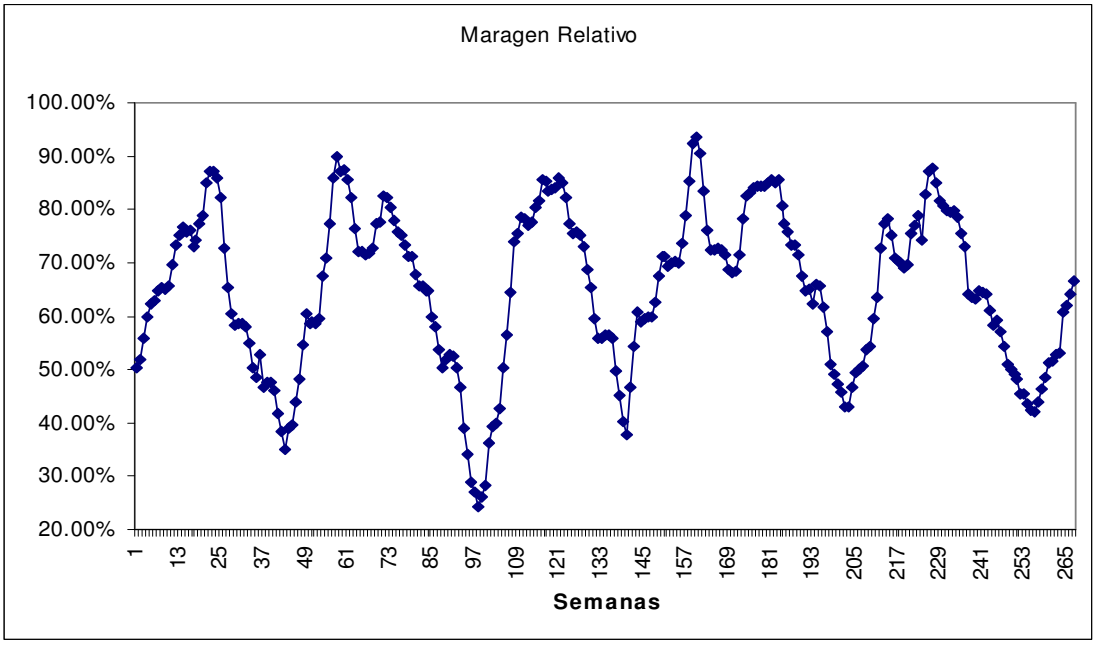
Por otro lado, la volatilidad del precio al productor es mucho más acusada que la del precio al consumo. En el primer caso, se aprecia una disminución de los precios durante la primavera, coincidiendo con el mayor volumen de oferta y un fuerte aumento hacia el otoño, previo a la campaña de Navidad, si bien la cantidad comercializada suele ser menor. En el caso del precio al consumo, la estacionalidad se genera a finales de año, coincidiendo con las fiestas Navideñas. Por tanto, la estacionalidad de final de año se produce un par de meses antes en el precio al productor.

Gráfico 2. Evolución del margen comercial absoluto en el sector del ovino en España (euros/Kg.)



Fuente: elaboración propia a partir de los datos del Observatorio de Precios (MARM)

Gráfico 3. Evolución del margen comercial relativo en el sector del ovino en España (% sobre precio al productor).



Fuente: elaboración propia a partir de los datos del Observatorio de Precios (MARM)

Como se puede apreciar en este Gráfico 1, en las épocas de precios al alza en otoño, los incrementos de precios al productor eran superiores a los del precio al consumo. Sin embargo, durante la mayor parte del año, ocurre justamente lo contrario y más concretamente durante la primavera. Como consecuencia de todo esto, observamos cómo los márgenes sufren fuertes oscilaciones motivados, principalmente por las oscilaciones de los precios al productor. Precios a la baja generan incrementos de márgenes mientras que un incremento del precio al productor genera reducciones de márgenes.

A pesar de que el aumento del diferencial de precios (margen absoluto) podría llevar al lector a pensar en una pérdida del poder de mercado del sector productor, un análisis más pormenorizado de la situación nos lleva a matizar mucho esta conclusión que, por otra parte, ha sido muy común en la mayor parte de los trabajos de divulgación y medios de comunicación. En el Gráfico 2 se recoge la evolución del diferencial de precios en relación con el precio percibido por el productor (margen relativo). Como se puede apreciar, a pesar de la existencia de ciclos, el margen relativo comercial se ha mantenido estable, alrededor de un 65%, a lo largo del período de estudio. Este resultado parece indicar que los operadores han trabajado con márgenes proporcionales estables a lo largo del tiempo. Esto explicaría el aumento del diferencial de precios absoluto, tal como se indica en el Gráfico 3, conforme ha venido aumentando el precio percibido.

Para concluir este apartado descriptivo, y teniendo en cuenta el objetivo de este trabajo, a partir de los análisis gráficos de los precios y los márgenes comerciales, podemos entresacar una conclusión importante. Conforme aumentan los precios percibidos, el margen disminuye y viceversa. Por el contrario, precios a la baja significan un aumento del diferencial de precios. En el largo plazo el diferencial de precios se mantiene estable, tal como hemos indicado anteriormente. Estos resultados intuitivos son los que trataremos, precisamente, de formalizar y contrastar en los próximos apartados.

Antes de proceder a la especificación y estimación del modelo adecuado para analizar el mecanismo de transmisión de precios, es necesario analizar con detalle las propiedades

estocásticas de las series mediante los contrastes de raíces unitarias. Para ello, se han utilizado los contrastes propuestos por Ng y Perron (2001) quienes proponen una versión modificada de los contrastes de Phillips y Perron (1988) teniendo en cuenta, asimismo, los resultados de Elliott et al. (1996). Para las dos series consideradas en este estudio, la aplicación de los mencionados contrastes indican que ambas series son I(1)[5].

4.2 Modelo de cointegración lineal

Dado que, como acabamos de mencionar, las series de precios no son estacionarias, el siguiente paso consiste en contrastar si dichas series están cointegradas. Para ello, en primer lugar se ha utilizado el procedimiento de Johansen (1988) para contrastar la existencia de cointegración lineal entre los dos precios. El modelo VAR se ha especificado con una constante restringida al espacio de cointegración. Con respecto al número óptimo de retardos del modelo VAR, en este trabajo se han utilizado dos criterios: el Criterio de Información de Akaike [Akaike (1971)] y el del Ratio de Verosimilitud propuesto por Tiao y Box (1981). Los resultados de la aplicación de ambos criterios indican que el sistema formado por las dos series de precios debería incluir cuatro retardos. En cualquier caso, se ha analizado la correcta especificación del modelo seleccionado utilizando el estadístico multivariante del Multiplicador de Lagrange de Doornik y Hendry (1997) para contrastar la presencia de autocorrelación de orden 1 y 7. Los resultados obtenidos de dichos estadísticos fueron de 3,03 y 28,78, respectivamente, ambos inferiores los valores críticos correspondientes para un nivel de significación del 5% (tal como indican los p-valores del Cuadro 1). Asimismo, se llevó a cabo el contraste de normalidad multivariante propuesto por Doornik y Hansen (1994). También en este caso, tal como se recoge en el Cuadro 1, el valor obtenido fue inferior al valor crítico al 5%, por lo que a partir del modelo estimado se podía proceder a contrastar si los dos precios estaban cointegrados.

En la parte superior del Cuadro 1, se recogen los resultados del estadístico de la Traza propuesto por Johansen (1988), así como sus correspondientes valores críticos. Para un nivel de significación del 5%, los resultados obtenidos no permiten rechazar la hipótesis nula de cointegración. La relación de cointegración normalizada por el precio al consumo aparece en la segunda fila del Cuadro 1. Como se puede apreciar, la elasticidad de transmisión de precios desde el mercado en origen al mercado detallista es de 1,29, algo alejada de la unidad. Asimismo, se observa que los parámetros de ajustes son significativos al 5% de nivel de significación. Estos resultado, nos indican que los dos precios (al consumo y al productor) no son débilmente exógenos respecto a los parámetros de largo plazo. El valor más elevado corresponde a α_{pp} indicando que el precio percibido por el productor responde más rápidamente ante una situación de desequilibrio. Así, ante una desviación respecto del equilibrio en t-1, el precio al productor ajusta el 12,4% en el mes siguiente y sólo el 4% en el caso del precio al consumo.

[5] Los resultados no se presentan debido a las limitaciones de espacio pero cualquier lector interesado puede solicitarlo a los autores.

Cuadro 1. Resultados de los contrastes de cointegración

Contraste de la Traza de Johansen		H_0: r=0	H_0: r=1
		53,23	7,17
	CV (5%)	20,16	9,14
Estimación de la relación de cointegración ($\hat{\beta}$)	$PC_t \underset{(-8,63)}{-1,293} PP_t \underset{(-0,772)}{-0,187} = \omega_t$		
Parámetros de ajustes: $\hat{\alpha} = \begin{pmatrix} \hat{\alpha}_{pc} \\ \hat{\alpha}_{pp} \end{pmatrix}$	$\begin{pmatrix} \underset{(-3,313)}{-0,027} \\ \underset{(5,608)}{0,124} \end{pmatrix}$		
Contrastes especificación multivariantes[a] Autocorrelación[b] No-normalidad[c]	A_LM(1) = 2,29 (0,68) A_LM(7) = 2,68 (0,61) N_LM = 16,37 (0,00)		

a. Los valores entre paréntesis se corresponden con los p-valores

b. A_LM(i) es el estadístico del Multiplicador de Lagrange para contrastar la hipótesis nula de ausencia de autocorrelación de orden i (Doornik y Hendry, 1997).

c. N_LM es el estadístico multivariante propuesto por Doornik y Hansen (1994) para contrastar la hipótesis de normalidad.

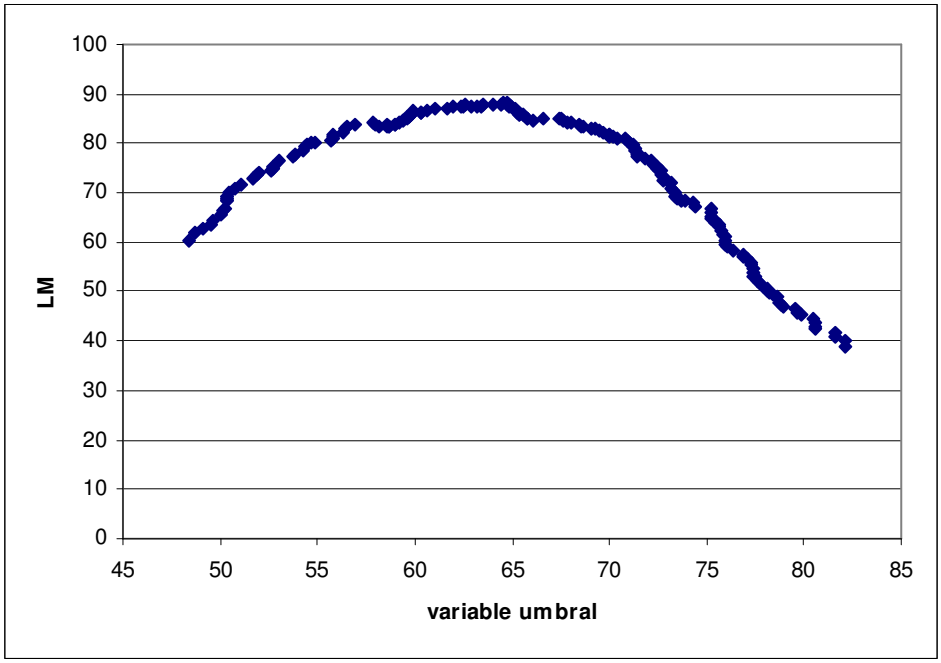
Fuente: Elaboración propia.

4.3 Cointegración por umbrales

Una vez comprobado que los precios están linealmente cointegrados, el siguiente paso consiste en contrastar si la transmisión de precios a largo plazo es lineal o sigue un proceso de transmisión por umbrales, utilizando el procedimiento propuesto por Gonzalo y Parakatis (2006), descrito en el apartado anterior. Dado que las dos series de precios evolucionan de forma semejante, la relación de cointegración puede ser estimada considerando cualquiera de los precios como variable dependiente (ver ecuación 1). Sin embargo, los resultados sobre la existencia de cointegración y las estimaciones de los parámetros del vector de cointegración, no son invariantes respecto a la normalización efectuada, eso es, a la variable elegida como dependiente. Además los resultados de la cointegración de Johansen indican que ninguno de los precios puede considerase como débilmente exógeno. En consecuencias, se han considerado las dos relaciones causales productor-consumidor y consumidor productor. En cuanto a la variable umbral se ha considerado el margen comercial relativo retardado un periodo[6]. Teniendo en cuenta estas consideraciones y fijando el parámetro trimming igual a 0.10, con el fin de asegurar que en cada régimen hay al menos 10% de las observaciones totales, se calcula el estadístico LM descrito en (3) para todos los valores de la variable umbral. Los resultados del contraste se observan en el Gráfico 2. Como se puede observar, el valor del estadístico supLM, en ambos regresiones se encuentra por encima de su valor crítico (12,27) para un nivel de significación del 5%. En consecuencia, se rechaza la hipótesis de linealidad en la relación de cointegración de los precios frente a la alternativa de una relación de cointegración por umbrales. Esto quiere decir que, teniendo en cuenta la expresión (2), el proceso de transmisión a largo plazo entre los dos precios va depender del margen comercial relativo en el periodo anterior. Para este valor supremo del estadístico LM, el parámetro umbral es igual a 64,5%. Desgraciadamente, todavía no existe un procedimiento validado en la literatura para realizar inferencia estadística sobre el estimador del parámetro umbral, dado que dicho parámetro posee una distribución limite no estándar que depende de los momentos de la muestra (Hansen, 1997). No obstante, Chan (1993) y Chan y Tsay (1998) demostraron que los parámetros umbrales son superconsistentes y, por tanto, sus estimadores pueden ser considerados como los verdaderos valores en las aplicaciones empíricas.

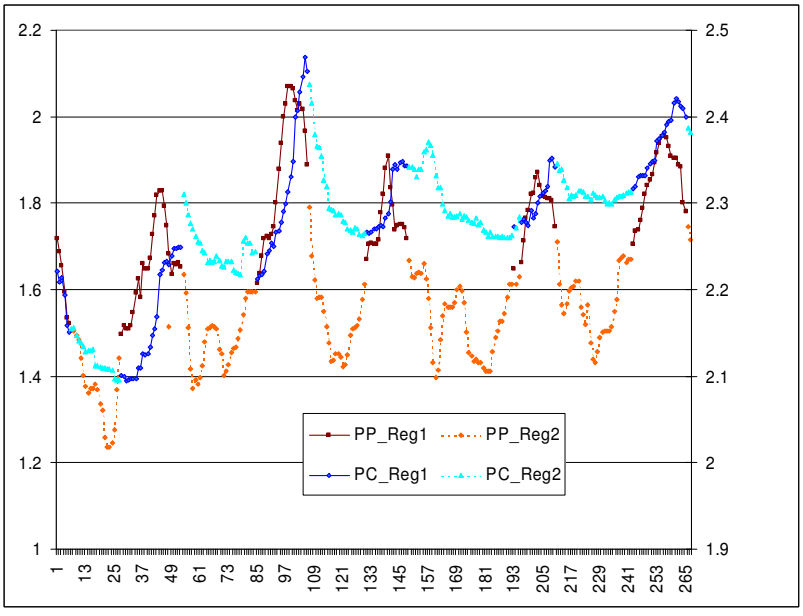
[6] La variable umbral se ha calculado de la siguiente forma: $q_t=100*(PC_t-PP_t)/PP_t$. Para garantizar la estacionariedad de dicha variables se han aplicado los contrastes de raíces unitarias. Los resultados, rechazan la hipótesis nula de una raíz unitaria.

Gráfico 4. Estimación del estadístico LM para todos los valores de la variable umbral de la regresión causal (PC-PP).



En el Gráfico 5 recogemos la evolución de los dos precios considerados así como la del margen comercial relativo para los dos regímenes establecidos. Como se puede apreciar, el primer régimen se asocia a períodos en los que el precio al productor ha llegado a su nivel más alto y empieza a descender (ultimas 28 semanas y las 4 primeras), coincidiendo con el período del año de mayor demanda. Por el contrario, el precio al consumidor se encuentra en una fase alcista. En definitiva, y como se indica, en el cuadro inferior izquierda del Gráfico 2, el primer régimen se identifica con márgenes comerciales relativo por debajo del 64,5%, mientras que, el segundo régimen se asocia a márgenes comerciales mayores del 64,5%. Por el contrario, y siguiendo un razonamiento similar, el segundo régimen se asocia a períodos en los que el precio al consumidor está bajando (generalmente desde la semana 4 a la 28) mientras que, el precio al productor se encuentra en su ciclo más bajo, coincidiendo con el periodo de menor demanda.

Gráfico 5. Clasificación de los precios en cada uno de los regímenes.



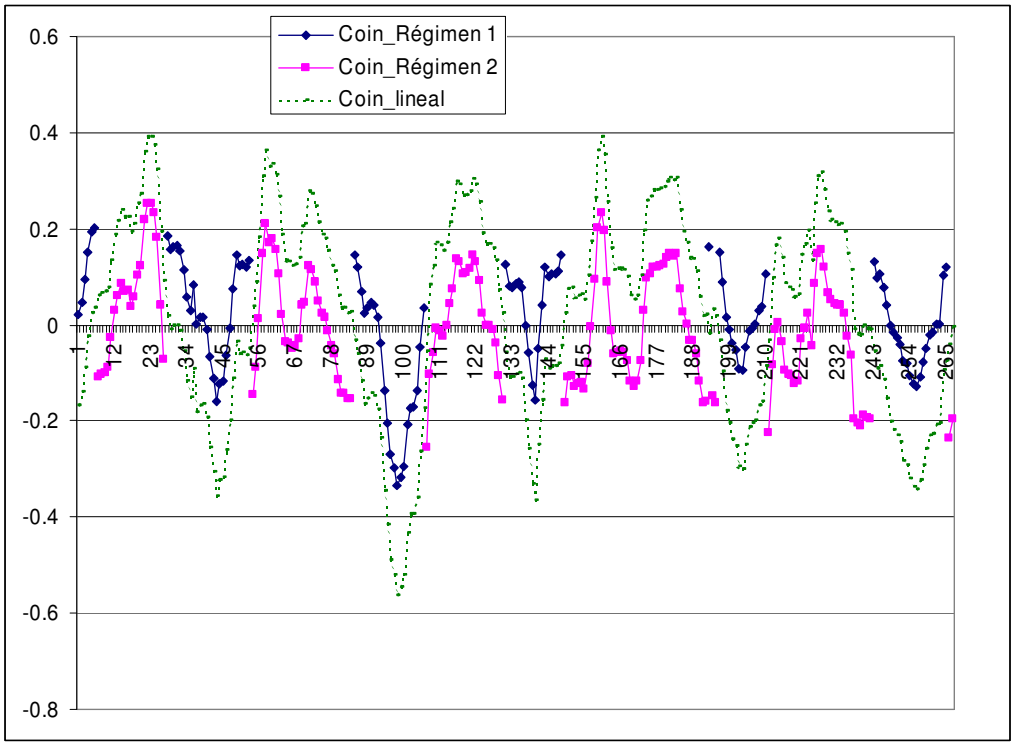
Cuadro 2. Cointegración por umbrales

	SupLM	$\hat{\gamma}$	Relación de cointegración
Relación PC-PP	88,33	64,5%	$\hat{\omega}_{t-1} = (Pc_{t-1} - 1{,}39Pp_{t-1})I_t(q_{t-1} \leq 64{,}5\%) + (Pc_{t-1} - 1{,}50Pp_{t-1})(1 - I_t)$
Relación: PP-PC	86,21	64,5%	$\hat{\omega}_{t-1} = (Pp_{t-1} - 0{,}77Pc_{t-1})I_t(q_{t-1} \leq 65{,}4\%) + (Pp_{t-1} - 0{,}66Pc_{t-1})(1 - I_t)$
Valor crítico (5%)	12,33		

Para cada uno de los regímenes se ha estimado la elasticidad de transmisión a largo plazo. El cuadro 2, recoge los resultados de la estimación. Como se puede observar, la elasticidad de transmisión para la relación causal PC-PP (PP-PC) es mayor (menor) en el régimen 2 que en el régimen 1. Atendiendo a la clasificación de los dos regímenes, podemos decir que la elasticidad de transmisión de precios a largo plazo es de mayor magnitud cuando los precios están bajando que cuando están subiendo. Dicho en otras palabras, un shock del 1% en el precio del productor (consumo) genera a largo plazo una variación del precio a nivel del consumidor (productor) del 1,28% (0,77%) cuando el margen comercial está por debajo de su nivel medio y del 1,52% (0,66%) cuando el margen está por encima del nivel medio.

En el Gráfico 5, aparecen las relaciones de cointegración tanto lineal como por umbrales (en cada uno de los regímenes considerados) para el caso de la relación causal PC-PP. Como se puede apreciar, las tres relaciones de cointegración muestran un comportamiento estacionario que giran entorno a una media nula. Los residuos estimados de las relaciones de cointegración por umbrales en cada uno de los regímenes se han utilizado para analizar si non estacionarios mediante el uso de los contrastes de raíces unitarias. Considerando los valores críticos para dos variables al 5% (-3,37), en los dos regímenes y para todas las regresiones causales, se rechaza la hipótesis nula de no cointegración. En consecuencia, se puede afirmar la cointegración de los precios en cada uno de los regímenes.

Gráfico 6. Relación de cointegración lineal y en cada uno de los regímenes.



4.4 Vector de mecanismo de corrección del error por umbrales

Tal y como se ha indicado en el apartado 3, debido a la imposibilidad de especificar un VMCE cuando las variables están cointegrados con umbrales, una forma de estimar el VMCE es suponiendo una relación de cointegración lineal. De hecho, si nos fijamos en los estimadores de las elasticidades de transmisión en ambos regímenes, se observa que el promedio [(1,28+1,50)/2] de las elasticidades de transmisión de los precios en los dos regímenes coincide exactamente con el valor estimado de la elasticidad de transmisión a partir de la relación de cointegración lineal (1,39). Asimismo, como se puede observar en el gráfico 3, la relación de cointegración lineal y la de cointegración por umbrales muestran evoluciones muy semejantes.

Basándonos en estos resultados, a continuación se estima un VMCE por umbrales de acuerdo a la ecuación (4). Para ello, partimos de la especificación de un TVECM con umbrales, suponiendo conocida "a priori" la relación de equilibrio a largo plazo existente entre los dos precios y tomando en consideración el valor umbral estimado en la etapa anterior ($q_{t-1} \gtrless 65,4\%$). Desde este punto de vista, y con el fin de determinar si los resultados son consistentes, hemos considerado dos especificaciones alternativas de dicha relación de cointegración. La primera es la obtenida a partir de los resultados de cointegración por umbrales, mientras que la segunda se basa en una relación de cointegración lineal. Denominaremos, TVECM_NL y TVECM_L a los modelos no lineales que incluyan la relación de equilibrio no lineal y lineal, respectivamente:

$$\Delta P_t = \begin{pmatrix} \Delta Pc_t \\ \Delta Pp_t \end{pmatrix} = \begin{cases} \begin{pmatrix} \alpha_1^1 \\ \alpha_2^1 \end{pmatrix} \hat{\omega}_{t-1} + \sum_{i=1}^{3} \Gamma_i^1 \Delta P_{t-i} + \varepsilon_t^1, & \text{si} \quad q_{t-1} \le 62,8\% \\ \begin{pmatrix} \alpha_1^2 \\ \alpha_2^2 \end{pmatrix} \hat{\omega}_{t-1} + \sum_{i=1}^{3} \Gamma_i^2 \Delta P_{t-i} + \varepsilon_t^2, & \text{si} \quad q_{t-1} > 62,8\% \end{cases} \qquad \textbf{(13)}$$

donde:

$$\hat{\omega}_{t-1} = Pc_{t-1} - 1,39 \mathrm{Pp}_{t-1} \quad \textbf{en el caso del TVECM_L} \qquad \textbf{(14)}$$

$$\hat{\omega}_{t-1} = Pc_{t-1} - 1,39 \mathrm{Pp}_{t-1} I_t(q_{t-1} \le 62,82\%) - 1,50 Pp_{t-1}(1 - I_t) \quad \textbf{en el caso del TVECM_NL} \quad \textbf{(15)}$$

Los resultados de la estimación de los dos modelos se recogen en el Cuadro 3. Únicamente se han incluidos los estimadores de los parámetros de ajustes a largo plazo. Así como los contrates de las hipótesis de ajustes lineales.

Cuadro 3. Estimación del VMCE por umbrales.

	Régimen 1	Régimen 2	Régimen 1	Régimen 2
Parámetros de ajustes: $\hat{\alpha} = \begin{pmatrix} \hat{\alpha}_{pc} \\ \hat{\alpha}_{pp} \end{pmatrix}$	$\begin{pmatrix} -0,041 \\ (-4,967) \\ 0,002 \\ (0,531) \end{pmatrix}$	$\begin{pmatrix} 0,001 \\ (-0,102) \\ 0,078 \\ (9,38) \end{pmatrix}$	$\begin{pmatrix} -0,043 \\ (-4,76) \\ 0,016 \\ (0,675) \end{pmatrix}$	$\begin{pmatrix} -0,002 \\ (-0,173) \\ 0,066 \\ (7,39) \end{pmatrix}$
Estadístico Wald (p valor) $H_0 : \Gamma_i^1 = \Gamma_i^2$ i=1,..,4 $H_a : \Gamma_i^1 \neq \Gamma_i^2$	98,08 (0,00)		87,53 (0,00)	
Estadístico Wald (p valor) $H_0 : \alpha^1 = \alpha^2$ $H_a : \alpha^1 \neq \alpha^2$	29,09 (0,00)		19,76 (0,00)	
% de observaciones	45%	56%	45%	56%

a Los valores entre paréntesis se corresponden con los t-ratios.

Fijándonos en los parámetros estimados de los dos modelos especificados, no se observan diferencias significativas en los resultados obtenidos. Asimismo, los resultados del estadístico de Wald rechazan las hipótesis nulas de igualdad de los coeficientes de corrección del error y de los parámetros de corto plazo entre los dos regímenes, respectivamente. Dada la similitud en los resultados de estimación de los dos modelos, a continuación, solamente interpretamos los del modelo TVMCE_L. Este modelo, será utilizado en la siguiente sección para obtener las Funciones-Impulso Respuesta no lineales.

En el Cuadro 3 se recogen los estimadores de los parámetros de ajustes hacia el equilibrio a largo plazo en cada uno de los regímenes, así como para cada de los modelos TVECM especificados anteriormente. A partir de los parámetros estimados ya podemos obtener algunas conclusiones preliminares sobre la posible existencia de relaciones asimétricas, cuestión que, en todo caso, será abordada de forma concluyente al calcular las denominadas funciones impulso, a las que dedicaremos la parte final de este apartado. Para ello, analicemos los coeficientes estimados de la matriz α asociados a la relación de equilibrio $\omega_{t-1}(\beta)$ (Cuadro 3). Estos coeficientes proporcionan dos tipos de información atendiendo a su significatividad, signo y magnitud. La significatividad es indicativa de que la variable en cuestión no es débilmente exógena respecto a los parámetros de largo plazo sino que reaccionan ante cualquier cambio inesperado para corregir el error o desajuste de la senda de equilibrio. Sin embargo, la no significatividad de un coeficiente de ajuste α significa que la variable no se ve afectada por el hecho de que las variables estén fuera del equilibrio a largo plazo, no reaccionando cuando se producen shocks inesperados en la relación de equilibrio a largo plazo.

Por otro lado, la magnitud de los coeficientes de ajustes (α) mide la velocidad de ajuste hacia el equilibrio a largo plazo. Teniendo en cuenta que los precios están expresados en logaritmos, estos parámetros pueden interpretarse como el porcentaje del error o desviación respecto del equilibrio a largo plazo en el período t-1 que el precio i corrige en un periodo. En consecuencia, un valor elevado indica que el ajuste ante una situación de desequilibrio se produce rápidamente.

Como se puede apreciar en el mencionado Cuadro, el coeficiente de ajuste (α) del precio al consumo resulta ser estadísticamente significativos en el primer régimen, pero por el contrario es débilmente exógeno en el segundo régimen. No obstante, ocurre totalmente lo contrario en el caso del precio al productor. Esto significa que en una situación de margen reducido (precios en fase alcista) el precio al productor es débilmente exógeno con respecto al equilibrio a largo plazo, por lo que no se ven afectados por las desviaciones que se producen respecto a la situación de equilibrio. Teniendo en cuanta que el régimen 1 corresponde a un ciclo de mayor demanda, este comportamiento, podría interpretarse como un mayor grado de liderazgo del productor y, en consecuencia, le produciría cierta ganancia transitoriamente.

Consideremos, a continuación, una fase de crecimiento del margen comercial que puede ser provocado bien por un aumento del precio al consumo, bien por una disminución del precio al productor (figuras derecha e izquierda de la parte inferior del Gráfico 3). Si el precio al consumo aumenta en t-1 por encima de lo que su relación de largo plazo con el precio al productor le permite, el precio del productor aumenta en t para restablecer la senda del equilibrio a largo plazo. Sin embargo, al disminuir el precio al productor los precios al consumo no reaccionan (α_{pc}^2 no es significativo), lo que genera un incremento del margen, en beneficio del sector detallista. En definitiva, lo que se deduce es que los precios al consumo son mucho más rígidos cuando los precios al productor disminuyen que cuando estos aumentan. Este resultado es consistente con una gran parte de la literatura existente sobre el tema, especialmente los trabajos de Ben-Kaabia y Gil (2007), von Cramon-Taubadel (1998) y Abdulai (2002). Finalmente, teniendo en cuenta la magnitud de los coeficientes de ajuste, podemos también afirmar que tanto el precio al consumo como el del productor reaccionan más rápidamente cuando el margen comercial aumenta (como consecuencia de un aumento del precio al consumo) que cuando éste se reduce (disminución del precio al consumo).

Finalmente, teniendo en cuenta la magnitud de los coeficientes de ajuste, podemos también afirmar que los precios al productor reaccionan más rápidamente ante una situación de desequilibrio que los precios al consumo.

A pesar de que los resultados obtenidos en el Cuadro 4 nos dan una idea aproximada de las asimetrías existentes en el mecanismo de transmisión de precios en el sector del tomate en España, el análisis de la dinámica del corto plazo nos va a permitir determinar con mayor precisión la magnitud y dirección de dichas relaciones asimétricas. A dicho análisis dedicaremos el próximo apartado.

4.5 Análisis de la dinámica de ajuste a corto plazo

En los modelos TVECM el análisis de la dinámica a corto plazo suele realizarse a partir del cálculo de las denominadas Funciones Impulso-Respuesta No lineales (FIRN) [Koop et al. (1996) y Potter (1995)]. La FIRN ante un shock $\varepsilon_t = \delta$ y un valor arbitrario de $P_{t-1}=\varphi_{t-1}$ (denominado como la historia o el pasado del sistema) viene dado por la siguiente expresión:

$$\begin{aligned} FIRN(n,\delta,\varphi_{t-1}) &= E\left[P_{t+n} \mid \varepsilon_t = \delta, \varepsilon_{t+1} = \ldots = \varepsilon_{t+n} = 0, \varphi_{t-1}\right] \\ &\quad - E\left[P_{t+n} \mid \varepsilon_t = 0, \varepsilon_{t+1} = \ldots = \varepsilon_{t+n} = 0, \varphi_{t-1}\right] \quad n = 0,1,\ldots N \end{aligned} \qquad \textbf{(16)}$$

Por tanto, las FIRN se calculan como la diferencia entre dos realizaciones de la serie P_t teniendo en cuenta el conjunto de información disponible hasta el período t-1. La primera es la realización de P_t después que haya ocurrido un shock $\varepsilon_t = \delta$ en el instante t, mientras que la segunda realización es el valor de P_t que se obtendría si no hubiera ocurrido ningún shock. Teniendo en cuenta esta definición, las FIRN varían en función de los valores que toman $\delta \in \varepsilon_t$ y $\varphi_{t-1} \in \Omega_{t-1}$ (Ω_{t-1} es el conjunto de información disponible en el instante t-1 utilizado para predecir los valores futuros de P_t). A partir de esta definición se pueden analizar las respuestas de los precios al consumidor (productor) ante un shock positivo (negativo) en los precios en origen (consumo) en cada uno de los regímenes identificados. Asimismo, las FIRN se pueden utilizar para evaluar el nivel de significatividad de los ajustes asimétricos a lo largo de los lapsos temporales. En este sentido, Potter (1995) definió una medida para evaluar el nivel de asimetría de las respuestas como la suma de las FIRN para un determinado shock y las FIRN para un shock de la misma magnitud pero con el signo opuesto. Dicha medida viene dada por la siguiente expresión:

$$ASY(n,\varepsilon_t,\varphi_{t-1}) = FIRN(n,+\delta_i,\varphi_{t-1}) + FRIN(n,-\delta_i,\varphi_{t-1}) \qquad \textbf{(17)}$$

En este trabajo hemos simulado las respuestas de cada precio antes shocks tanto positivos como negativos ($\pm 1\%$ $\pm 2\%$) en el otro precio. Por tanto, analizaremos en primer lugar las respuestas ante un cambio en el precio al productor (PP) (parte superior del Gráfico 7, para luego centrarnos en las derivadas de un cambio en el precio al consumo (PC) (parte inferior del Gráfico 8). Asimismo, para dar una idea de la posible existencia de asimetrías, en cada caso situamos a la derecha una medida de las asimetrías existentes que simplemente se obtiene sumando las respuestas de cada precio antes shocks de un 1 y un 2% tanto positivos como negativos. Por ejemplo, en el la parte inferior del Gráfico 7, ASI(PP) recogería la suma de las respuestas del precio al productor PP(1%), PP(-1%), PP(2%) y PP(-2%). Valores positivos (negativos) indican asimetrías positivas (negativas) mientras que valores alrededor de cero indicarían respuestas simétricas.

En la parte superior del Gráfico 7 se recogen las respuestas del precio al consumo y del precio al productor como consecuencia de un shock positivo y negativo (1% y 2%) en el precio al productor. En la parte inferior se recoge la medida propuesta por Potter (1995) para evaluar la asimetría de las respuestas obtenidas.

Un incremento inesperado en el precio al productor en el periodo t puede interpretarse como un exceso de demanda, mientras que una disminución como un exceso de oferta. El incremento (disminución) inicial en los precios puede ser a muy corto plazo e incluso de una magnitud superior al shock, dependiendo de cual sea el exceso de demanda (oferta) en el mercado. Como se puede observar en el gráfico las funciones impulso-respuesta del precio al productor ante un shock (tanto positivo como negativo) en si mismas reflejan un comportamiento esperado. El precio al productor responde en una proporción superior al propio shock inicial tanto positivo como negativo, alcanzando su valor máximo (mínimo) 3 semanas más tarde. A partir de este

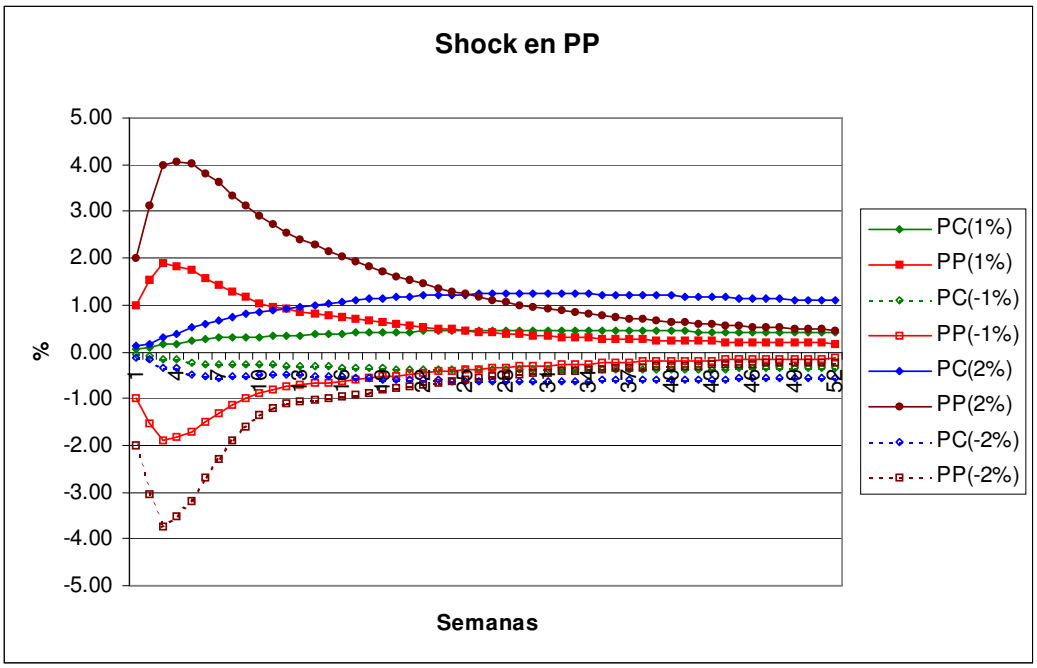
periodo el precio disminuye (aumenta) progresivamente hacia su nivel de equilibrio. Comparando la respuesta del precio al productor ante shocks positivos y negativos podemos extraer las siguientes conclusiones. En primer lugar, a lo largo de todo el lapso temporal las respuestas ante un exceso de demanda son de mayor magnitud que ante shocks de exceso de oferta, generando en consecuencia respuestas asimétricas. En segundo lugar, se observa que la asimetría positiva se va incrementando de forma progresiva alcanzando un valor máximo en la semana 10 (en torno al 1,7%). Esto indica que los productores, tras un shock positivo tardan mucho más tiempo en recuperar la situación de equilibrio que cuando se trata de un shock negativo. Estos resultados indican que un efecto positivo sobre los precios al productor se puede prolongar durante muchas semanas aprovechando el tirón de mercado, hasta que la decisión de aumentar la producción se manifiesta en el mercado a través de un incremento en la oferta. Este incremento en la oferta ejerce una presión a la baja de sobre los precios. Sin embargo, ante un shock negativo (disminución de los precios como consecuencia de exceso de demanda), los productores frenen el sacrificio de los corderos, con el consiguiente aumento de peso. Este hecho genera una disminución en la cantidad ofertada que, manteniendo una demanda constante, se traduce en un incremento de los precios.

Veamos ahora que ocurre con el precio al consumo ante shocks positivos y negativos en el precio al productor. Tanto ante shocks positivos como negativos, la respuesta del precio al consumo no es inmediata. En primer lugar, el precio al consumo requiere unas 2 semanas para shocks positivos y de 3 a 4 semanas en caso de shocks negativos. Esta respuesta retardada puede estar relacionada con el tiempo requerido por los detallistas en la obtención y transmisión de información (Coste de menú). En segundo lugar, ante shocks positivos en el precio al productor, la respuesta del precio al consumo es progresiva, llegando a situarse por encima del precio al productor a largo plazo. Sin embargo, las respuestas ante shocks negativos no sólo son de menor magnitud que los shocks positivos, sino además no proporcionales a la magnitud del shock. Este fenómeno genera a largo plazo asimetrías positivas.

Analicemos, para finalizar, las respuestas de ambos precios ante shocks positivos y negativos en el precio al consumo (Gráfico 8). En este caso, un incremento inesperado en el precio al consumo en el periodo t puede interpretarse como un exceso de demanda, mientras que una disminución como exceso de oferta.

La conclusión a la que llegamos es, en este caso diferente. En primer lugar, las respuestas son de menor magnitud, lo que corrobora que la mayor fuente de volatilidad en el sector se debe a las variaciones del precio al productor. En segundo lugar, shocks positivos (negativos) en el precio al consumo generan respuestas positivas (negativas), que van disminuyendo progresivamente en intensidad hasta alcanzar sus niveles de equilibrio a largo plazo. Por otra parte, los shocks positivos y negativos en el precio al consumo son inmediatamente respondidos por el precio al productor. Como se puede apreciar, un shock negativo del 1% en el precio al consumo se repercute casi íntegramente en el precio al productor 3 semanas más tarde. Sin embargo, shocks negativos tienen una menor repercusión en el precio al productor. Asimismo, shocks positivos se transmiten de forma más lenta que shocks negativos, lo que nos llevaría a pensar en la posible existencia de asimetrías negativas. Finalmente, si analizamos las respuestas de los dos precios ante shocks positivos y negativos del precio al consumo, concluimos que las respuestas ante shocks negativos son de mayor magnitud que las correspondientes a shocks positivos, generándose asimetrías negativas en caso del precio al productor. Este hecho pone en evidencia el poder que ejerce el mercado detallista en la fijación de los precios. Ante una baja de precios a nivel detallista, este último presiona al productor para que baje los precios y así recuperar su margen comercial.

Gráfico 7. Funciones Impulso Respuesta a shocks positivos y negativos del precio al productor (magnitud del shock: ± 1 y 2%)



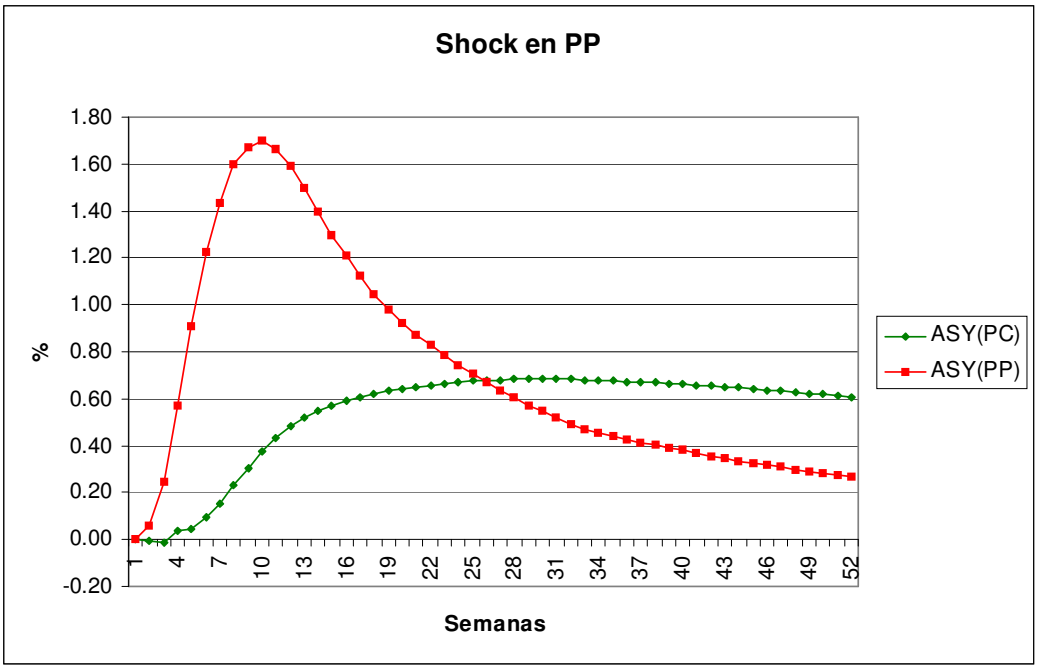
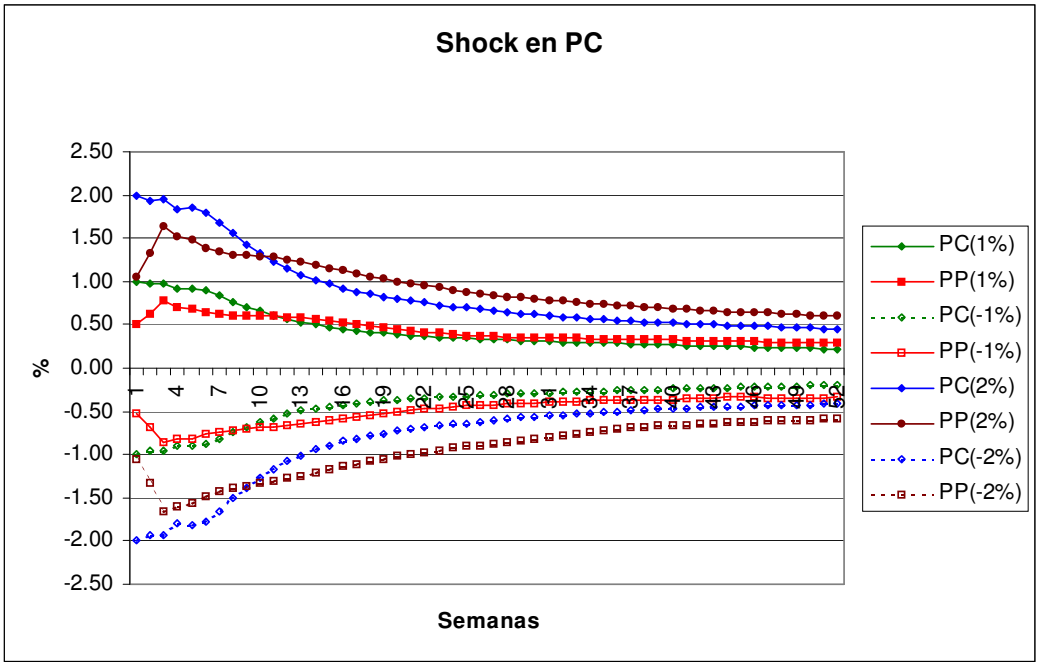
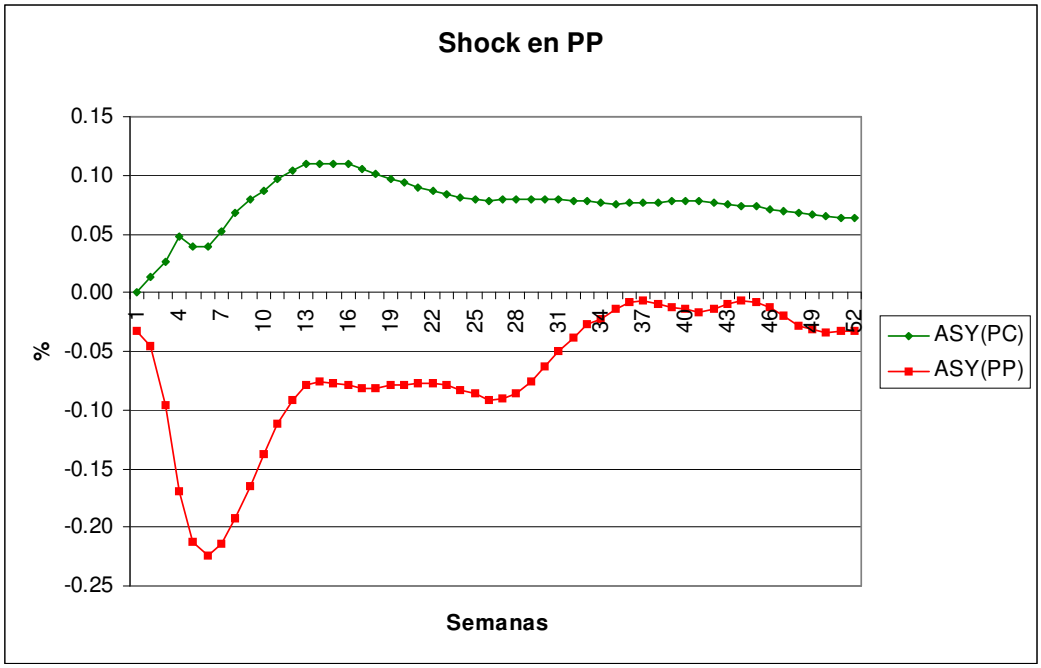


Gráfico 8. Funciones Impulso Respuesta a shocks positivos y negativos del precio al consumidor en el Régimen 3 (magnitud del shock: ± 1 y 2%)





BIBLIOGRAFÍA

ABDULAI, A. (2002). "Using threshold cointegration to estimate asymmetric price transmission in the Swiss pork market". *Applied Economics*, 34:679-687.

AKAIKE, H. (1971). "Fitting autorregresive models for prediction." *Annals of The Institute of Statistical Mathematics,* 21:243-247.

ANDREWS, D.W.K. (1993). "Tests for parameter instability and structural change with unknown change point". *Econometrica*, 61(4):821-856.

AZZAM, A. (1999). "Asymmetry and rigidity in farm-retail price transmission". *American Journal of Agricultural Economics,* 81:525-533.

BAILAY, D.V. Y BRORSEN, B.W. (1989). "Price asymmetric in spatial fed cattle markets." *Western Journal of Agricultural Economics*, 14:246-252.

BALKE, N.S. Y FOMBY, T.S. (1997). "Threshold cointegration". *International Economic Review*, 38:627-645.

BEAULIEU, J.J. Y MIRON, J.A. (1993). "Seasonal unit roots in aggregate U.S. data". *Journal of Econometrics*, 55:305-328.

BEN-KAABIA, M. Y GIL, J.M. (2007)."Asymmetric price transmission in the Spanish lamb sector". *European Review of Agricultural Economics,* 34(1):53-80.

CHAN, K.S. (1993). "Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model". *The Annals of Statistics,* 21:520-533.

CHAN, K.S. Y TSAY, R.S. (1998) "Limiting properties of the least squares estimator of a continuous threshold autoregressive model", *Biometrika*, 85(2):413-426,

CHAVAS, J.P. Y METHA, A. (2004). "Price dynamics in a vertical sector: the case of butter". *American Journal of Agricultural Economics,* 86:1078-1093.

DAVIES, R.B. (1987). "Hypothesis testing when a nuisance parameter is present only under the alternative". *Biometrika,* 74:33-43.

DOORNIK, J.A. Y HANSEN, H. (1994). *A practical test of multivariate normality*, unpublished paper, Nuffield College.

DOORNIK, J.A. Y HENDRY, D.F. (1997). *Modelling dynamic systems using PcFilm 9 for Windows.* Timberlake Consulting, London.

ELLIOTT, G., ROTHENBERG, T. J. Y STOCK, J. H. (1996). "Efficient tests for an autoregressive unit root". *Econometrica,* 64:813–836.

ENDERS, R.F. Y GRANGER, C.W.J. (1998). "Unit root tests and asymmetric adjustment with an example using the term structure of interest rates." *Journal of Business and Economic Statistics*, 16:304-312.

ENGLE, R.F. Y GRANGER, C.W.J. (1987). "Co-integration and error correction: representation, estimation and testing". *Econometrica* 55:251-276.

FRANSES, P.H. (1991). *Model selection and seasonality in time series*. Tinbergen Institute series nº 18. Erasmus University. Roterdam.

FRANSES, P.H. Y HOBIJN, B. (1997). "Critical values for unit root tests in seasonal time series". *Journal of Applied Statistics*, 24:25-46

FRANSES, P.H. Y VAN DIJK, D. (2000). *Non-linear time series models in empirical finance.* Cambridge University Press.

FROST, D. Y BOWDEN, R. (1999). "An asymmetry generator for error-correction mechanisms, with application to bank mortgage-rate dynamics". *Journal of Business and Economic Statistics*, 17(2):253-263.

GOODWIN, B.K. Y PIGGOTT, N.E. (2001). "Spatial market integration in the presence of threshold effects". *American Journal of Agricultural Economics*, 83(2):302-317

GONNZALO, J. Y PITARAKIS, J. (2006). "Threshold effects in multivariate error correction models". En: Mills, T.C. y Patterson, K. (eds) *Palgrave Handbook of Econometrics,* Vol I, Chapter 15:578-609.

GRANGER, C.W.J. Y LEE, T.H. (1989). "Investigation of production, sales and inventory relationships using multicointegration and non-symmetric error correction models". *Journal of Applied Econometrics*, 2:111-120.

GRIFFITH, G.R. Y PIGGOT, N.E. (1994). "Asymmetry in beef, lamb and pork farm-retail price transmission in Australia". *Agricultural Economics*, 10:307-316

HANSEN, B.E. (1997). "Inference in TAR models". *Studies in Nonlinear Dynamics and Econometrics,* 2:1-14.

HANSEN, B.E. (1999). "Testing for linearity". *Journal of Economic Surveys*, 13, 551-576.

HANSEN, B.E. Y SEO, B. (2002). "Testing for two-regime threshold cointegration in vector error correction models". *Journal of Econometrics*, 110:293-318.

HOUCK, J.P. (1977). "An approach to specifying and estimating nonreversible functions". *American Journal of Agricultural Economics*, 59:570-572.

ILMAKUNNAS, P. (1990). "Testing the order of differencing in quarterly data: an illustration of the testing sequence". *Oxford Bulletin of Economics and Statistics*, 52:79-88.

JOHANSEN, S. (1988). "Statistical analysis of the cointegration vectors". *Journal of Economic Dynamics and Control*, 12:231-254.

KINNUCAN, H.W. Y FORKER, O.D. (1987). "asymmetry in the farm-retail price transmission for major dairy products". *American Journal of Agricultural Economics*, 69:285-292.

KOOP, G., PESARAN, M.H. Y POTTER, S.M. (1996). "impulse response analysis in nonlinear multivariate models". *Journal of Econometrics*, 74:119-147.

LO, C. Y ZIVOT, E. (2001). "Threshold cointegration and nonlinear adjustments to the law of one price". *Macroeconomic Dynamics*, 5:533-576.

MAPA (varios años). *Boletín mensual de estadística*. Secretaria General Técnica, Madrid.

MAPA (2006). *Anuario de estadídtica agraria, 2004*. Secretaria General Técnica, Madrid.

MCCORRISTON, S., MORGAN, C.W. Y RAYNER, A.J. (2001). "Price transmission: the interaction between market power and returns to scale". *European Review of Agricultural Economics,* 28:143-159.

MEYER, J. Y VON CRAMON-TAUBADEL, S. (2004). "Asymmetric price transmission: a survey". *Journal of Agricultural Economics*, 55(3):581-611.

NG, S. Y PERRON, P. (2001). "Lag length selection and the construction of unit root tests with good size and power". *Econometrica*, 69:1519–1554.

OBSTFELD, M. Y TAYLOR, A. M. (1997). "Non-linear aspects of goods-market arbitrage and adjustment: Heckscher's commodity points revisited." *The Journal of the Japanese and International Economics,* 11:441-479.

PHILLIPS P.C.B. Y PERRON, P. (1988). "Testing for a unit root in time series regression". *Biometrika*, 75:335-346.

POTTER, S.M. (1995). "A nonlinear approach to U.S. GNP." *Journal of Applied Economics*, 10:109-125.

REAGAN, P.B. Y WEITZMAN, M.L. (1982). "Asymmetries in price and quantity adjustments by the competitive firm". *Journal of Economic Theory*, 27:410-420

SANJUÁN, A.I., GIL, J.M. Y ALBISU, L.M. (1996). *Una aproximación a la medición de la incidencia de los precios agrarios en los precios al consumo de alimentos (1987-1992*). Monografías INIA, 94. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Madrid.

SERRA, T. Y GOODWIN, B.K. (2002). "Specification selection issues in multivariate threshold and switching Models." Paper presented at the *AAEA-WAEA Annual Meeting*

TIAO, G.C. Y BOX, G.E. (1981). "Modeling multiple time series applications." *Journal of American Statistical Association*, 76:802-816.

TIFFIN, R. Y DAWSON, P.J. (2000). "Structural break, cointegration and the farm-retail price spread for lamb". *Applied Economics*, 32:1281-1286.

TSAY, R. (1998). "Testing and modelling multivariate threshold models". *Journal of the American Statistical Association*, 93(1):1188-1202.

VON CRAMON-TAUBADEL, S. (1998). "Estimating asymmetric price transmission with the error correction representation: An application to the German pork market". *European Review of Agricultural Economics*, 25:1-18.

WARD, R.W. (1982). "Asymmetry in retail, wholesale and shipping point pricing for fresh vegetable". *American Journal of Agricultural Economics*, 64:205-212.

WOLFFRAM, R. (1971). "Positivistic measures of aggregate supply elasticities: some new approaches – Some critical notes." *American Journal of Agricultural Economics*, 53:356-359.

THE VOTING PARADOX AND SOCIAL NETWORKS: AN EMPIRICAL ANALYSIS

JOÃO AMARO DE MATOS
PEDRO PITA BARROS
INÁCIA PACHECO PEREIRA
E-mail: ipimentel@fe.unl.pt
Universidade Nova de Lisboa

Abstract

The level of abstention in elections has been increasing over the years in the majority of democratic countries, whether compulsory voting exists or not. We provide empirical evidence for the role of social networks as a main factor of influence in the turnout of an election. More intense social networks and wider social networks are associated with a lower abstention rate.

Keywords: abstention, social networks.
Thematic Area: Economic and Financial Modelling.

1 INTRODUCTION

Each voter in an election has a negligible impact on the final outcome, and yet, many people take their time and make an effort to vote. This paradox has puzzled social scientists since Downs (1957). Several explanations have been given but it is essentially an empirical matter to identify the contribution of each one. This paper provides empirical evidence that social networking plays a significant role in the individual decision to vote or not to vote. By testing the empirical implications of a model that isolates the impact of the interaction between voters in the decision to vote (Amaro de Matos and Barros, 2003), this paper provides a number of factors that may help to understand the level of abstention. The empirical analysis is based on a set of recent data from the Portuguese political scene.

The study of voter behavior is as a topic of interest for Economic theory in at least two different dimensions. On one hand, it is a natural theoretical topic since by voting, people decide how to allocate scarce productive resources: those used by the government and those used in the voting process, namely, the time spent by electors. On the other hand, individuals in democracies determine who will be responsible to define the political, social and economic policies by voting. Elections define the governments and the legislative houses, and condition their decisions. Such decisions will affect all parties within the economic system. In turn, the economic performance may have an influence on the outcome of an election.

Before analyzing the composition of votes, it is wise to examine abstention. Democracies are legitimized by their people's vote. By voting, people choose those who will represent them in the decision-making bodies. When the abstention level is too high, the legitimacy of the regime may be questioned. As a result, the voting paradox has both economic and political implications.

Among the theories seeking to explain voting motivations, some have focused on the definition of the voting benefit, and, in particular, on identifying benefits that do not depend on the outcome of the election. Sieg and Schultz (1995) take a different route, by questioning the full rationality and the strong information requirements. Coate and Conlin (2002) present a distinct framework in which voting turnout results from a contest between two opposing groups. Their starting point was Harsanyi (1980), who suggested that voting turnout can be understood as people acting according to rule-utilitarianism. That is, each individual takes the action that, if adopted by all society members, maximizes social welfare, defined as the sum of individuals' utilities. These explanations, implicitly or explicitly, view the act of voting as a consumption

good.

The seminal work of Downs (1957) is closer to seeing voting as an investment. Downs relates income to turnout, arguing that the main cost of voting is usually the income lost by devoting time to it, and thus high-income people will tend to vote, as their marginal utility of income is lower. Moreover, his point of view reflects the fact that "if citizens act rationally on the basis of individually defined self-interest (…) they might very well abstain from voting in election".[1]

Several empirical experiments have attempted to explain the relationship between income and voting. Frey (1971) argues that the jobs of high-income voters endow them with superior information, which motivates higher participation. Palfrey and Rosenthal (1983, 1985) argue that, in a world of rational voters, the crucial factor is the expected benefit from voting. They provide some empirical evidence that public bureaucrats will vote more than the private-sector workers, and that these workers vote more than the unemployed citizens.

Other works have complemented the above results: Guttman, Hilger and Schachmurove (1994) came to the conclusion that more educated individuals are expected to vote more frequently than less educated individuals. Based on Myerson (1994), Castanheira (1999) uses a generalized model of Poisson games to explain turnout rates and why such rates might be decreasing with population size. Castanheira (2001) verifies that turnout is regularly higher in richer classes of the population.

Fielding (2000) uses a theoretical framework in which voters' choices are the result of utility maximization decisions, and estimates constituency-level regression equations to explain how the electoral shares of each main party depend on conditions in the constituency. Baum and Freire (2002) tested the hypothesis that aggregate levels of turnout vary in relation to an election's degree of importance for the functioning of the political system, using data for Portugal during the democratic period between 1975-2002, and concluded in favor of different levels of abstention for each election's degree of importance.

More recently, Coupé and Noury (2004) used a questionnaire organized by the National Research Council to isolate the effect of being well informed on the probability of participation in an election. They concluded that information is a major determinant of abstention, and that theoretical studies that assume that all voters have full information are likely to be unrealistic.

As an additional element to these explanations, Amaro de Matos and Barros (2003) argue the importance of the interaction between voters, and the capacity they have to influence each other, as a determinant factor for the abstention level. If a set of individuals has a positive attitude toward voting, they may influence the social network of people with whom they relate, each one of whom, in turn, will influence others, in a domino effect. The argument is based on mutual influences within the population, a kind of information sharing. By making explicit the channel by which social interactions influence voting behavior, empirical predictions are generated and taken to the data. We provide empirical evidence supporting this notion of social cascading. More recently, other works addressed the issue. Another line of approach is present in Blondel and Levy-Garboua (2008), where a generalization of the expected utility framework (cognitive consistency theory) is used. It has some similarities to the basic principle: receiving a message that one's vote is relevant leads to a revised probability on the value of voting. The origin of such messages is exogenous in Blondel and Levy-Garboua (2008), and the mechanism by which the effect takes place is completely different from Amaro de Matos and Barros (2003). Also, Feddersen and Sandroni (2006) present a model where high turnout rates at elections are obtained endogenously due to an extra benefit to voters from voting, associated with the process itself, which is a complementary, not competing, explanation.

This paper is organized in the following way. Section 2 presents the theoretical background and makes explicit the empirical implications of the theory. Section 3 describes the behavior of the Portuguese electoral system in recent decades and documents the evolution of the abstention rate in those years, namely in the different types of elections. Section 4 sets a reduced form model and presents the data used. Section 5 reports the econometric implementation. Section 6 discusses some robustness checks of the statistical results, and Section 7 concludes.

[1] See Goodin and Klingeman (1998: p. 225).

2 THE THEORETICAL FRAMEWORK

As pointed out in Amaro de Matos and Barros (2003), social interactions may have a determinant role on an election's outcome. The utility to each individual from voting was modelled as depending on the social influence to which each individual is subject. This social influence is decomposed into (1) direct social interaction between agents, (2) indirect social interaction between agents, and (3) personal values.

Suppose, for the moment, that agents do not interact directly, but that each individual faces an indirect social interaction, namely through some social pressure. By social pressure we mean social elements that are common to all the voters, and perceived by each of them in the same way, such as the existence of a majority, the presence of strong leaders, the influence of the media and factors such as ethical (or other social) values. Let h_c denote the intensity of this pressure, assumed to be felt equally by each agent. If there are no other reasons than this social pressure for each individual to vote or not to vote, then no matter how small the degree of pressure, it induces either massive voting or massive abstention. If the net effect of the pressure leads to a positive value of h_c, then everybody would vote. If the net effect leads to a negative h_c, then everybody abstains. This common element alone cannot explain different individual decisions regarding abstention.

Turning now to direct interaction between agents, it is assumed that, in the absence of social pressure, each individual chooses to vote or not to vote in order to maximize the degree of agreement between his social relations and himself. Let s_i express the voting attitude of individual *i*. If $s_i = +1$, individual *i* is willing to vote, if $s_i = -1$ individual *i* abstains. Consider a pair of voters *i* and *j*, who may either agree with respect to their voting attitude, or disagree. In the former case $s_i s_j = 1$, whereas in the latter $s_i s_j = -1$. The level of utility generated by agreement for a given pair of voters is given by $Js_i s_j$, with $J > 0$. It is assumed that, in the absence of social pressure, voter *i* chooses his/her attitude so that he/she maximizes the degree of agreement $J\sum_{j \in l_i} s_i\, s_j$, where $l_i = \{i_1, i_2, ..., i_n\}$ denotes the set of *n* voters with whom voter *i* interacts. As the total size of the voters' network *N* increases, both *J* and *n* may change.

To introduce the possibility that each person has his/her own idiosyncratic preferences that influence her attitude of voting, or not, a personal values variable $h_i s_i$ is added to the utility, where h_i reflects heterogeneity in the population, with density $p(h_i)$.

Taking all these factors together, individual *i* is assumed to choose the voting attitude (s_i) so as to maximize total utility,

$$H_i = h_c s_i + J\sum_{j \in l_i} s_i s_j + h_i s_{i.} \qquad (1)$$

In this paper, we provide empirical support for the relevance of social direct interaction in the decision to vote. The results reported here are valid for arbitrarily large networks of voters. In particular, we assume that *J=O(1/N)* and that *nJ=O(1)*, as the network size *N* grows without bound.

It can be shown that for a large network of voters, $m = \mathrm{E}(s_i)$, the expected voting attitude, is given by the solution to

$$m = \int p(h_i)\tanh[(nJm + h_c + h_i)/D]dh_i \qquad (2)$$

where *D* is a constant that may be interpreted as the degree of tolerance for different opinions.[2] We can isolate the effect of direct interaction by assuming that *J* is much larger than social pressure or personal values. Alternatively, we can make both $h \to 0$, to get the expected

[2] For further details of the model and on the role of *D* as indirect social interaction, see Amaro de Matos and Barros (2003) and Galam and Moscovici (1991).

voting attitude in a very large population to satisfy in equilibrium

$$m = \tanh(\frac{nJ}{D}m) \quad (3)$$

A high *D* means that interacting agents may more freely express different attitudes. A value of *D* close to zero leads to consensus about voting (*m*=1), for fixed values of *n* and *J*.

From this equation it follows that a high number of social contacts, *n,* and a high *J*, reflecting the intensity of influence due to exchange among individuals, makes it easier to reach consensus for a fixed *D* as well.

In order to generate a high turnout equilibrium it turns out that at least a certain number of individuals must believe in the importance of voting rather than abstaining. This minimum fraction of population with a positive attitude can be understood as a requirement for a minimum number of social interactions needed to trigger a general swing toward voting.

The model generates several empirical implications, namely a) the equilibrium turnout is smaller whenever the ability to accommodate divergence within society, increases. It becomes less important to follow the social norm, which results in a lower voting level; b) If the ability to allow for divergence is greater among young people, a higher level of abstention among younger voters is expected; c) An increase in the number of social contacts is expected to generate a higher turnout in elections; d) The abstention-dominant equilibrium (the $m<0$ solution) disappears if the intensity of social coercion, h_c, is sufficiently large; e) The development of new information and communication technologies, may influence the turnout by increasing the number of social contacts, that is, the equilibrium turnouts become closer to 100% participation or 0% participation.

A final issue to explore empirically is whether or not a small sub-set of the population deciding to vote is sufficient to induce high turnouts at elections.[3]

3 PORTUGUESE ELECTIONS

We now briefly describe briefly now the context in which Portuguese elections took place in the last decades and document the evolution of the abstention rates in those years. Since 1975, 39 elections have taken place in Portugal: 7 presidential elections, 11 parliamentary elections, 7 local elections, 7 elections for the administrative bodies of Azores and Madeira, 4 European elections, and 2 referenda.

All the elections taking place in Portugal follow the d'Hont system of proportional representation, except the presidential elections, which follow a run-off majority system.[4]

The Portuguese abstention rate is related to the country's political evolution. At the end of the 1970´s, participation in elections was very high, reflecting the revolutionary period that Portugal was going through. In the first election following the 1974 revolution (which restored democracy in Portugal) the abstention level was 10.3%. This level has increased since then, following a general trend in developed countries (Lane and Ersson, 1999; Gray and Caul, 2000; Dalton and Wattenberg, 2000; Bartolini, 2001). All of the Western democracies not having compulsory vote, apart from Spain, have shown an increase in abstention levels from 1970 to 1990. The abstention rate in countries where voting is compulsory was compared with the abstention rate in other countries (Freire and Magalhães, 2002, pp.45) and the abstention rate of the second group exceeds by 14.8 pp the first group. Despite this fact, it is worth mentioning that there is wide variance in terms of the abstention levels between industrialized democracies. For instance, the average abstention level in local elections in the USA from 1946 until 1998 is 52.3%, while in Italy it was only of 7.5% (Freire and Magalhães, 2002, pp.44). Also within each country there is a significant dispersion. In the Portuguese local elections of 2001, the abstention level ranged from 32% to 52% at the level of the administrative regions.

[3] This theoretical possibility is derived in Amaro-de-Matos and Barros (2003).

[4] If no candidate gets the majority of votes in the first ballot, a second election takes place with only the two candidates who received the most votes.

Regarding the number of registered voters, in 1973 only 21% of the total population were registered voters, in 1975 growing to 66%, in 1978 to 71%, and in the 1990s to 80% (Espírito-Santo, 2005). This is common to all elections.

Our data set comprises the results of the 2001 local elections, although data from European Parliament elections are used as well (in a way described below).[5]

Portugal is divided into administrative units called *distritos*. These units are subdivided into municipalities (*municipios)*, one per each township within each *distrito*. The municipalities are subdivided into parishes (*freguesias*) for voting and administrative purposes. The main elected officials are at the municipality level, which commands both the political and financial decision-making at local level. There are no officials elected at the level of the administrative unit (*distrito*), and those elected at the parish level are subject to financial and administrative rule by the municipality.

Local elections are held within municipalities and choose the bodies that govern the cities (or the areas dominated by the cities). We must distinguish between three governing bodies in these elections: the municipal executive (*câmara municipal*); the municipal assembly (*assembleia municipal*) and the assembly of the parish (*assembleia de freguesia*). Elections for these bodies take place into the 308 municipalities in which the country is divided. The average abstention rate in all local elections from 1974 until 1999 was less than 40%, (Espírito-Santo, 2005).[6] Silva and Costa (2006) and Veiga and Veiga (2004) also use Portuguese elections data. Using data from the 1997 local elections, Silva and Costa (2006) look at the hypothesis of rational ignorance, but do not find conclusive evidence. Veiga and Veiga (2004) document how economic conditions affect voting outcomes. None of these researchers address the specific issue of determinants of the abstention rate.

There are also the European Parliament elections. They began only in 1987, as Portugal became a member of the European Union in 1986. Looking at the European scenario, we see that almost half of the citizens do not vote in the European elections, and Portugal is not an exception, reaching an abstention level in these elections of 62.2% in the 1990s.[7] To justify this level of abstention, it is commonly observed that the result of these elections has no direct impact on the daily life of the European citizens, as do the presidential, national and local elections.

4 DATA AND METHODOLOGY

We use a reduced-form model for the abstention rate, A_k, where k labels the local municipalities and e_k is a (normally distributed) random error, accounting for neglected and non-specified factors. The basic relationship (2) is linearized to be expressed as,

$$A_k = \alpha_0 + \alpha_1 n_k + \alpha_2 J_k + \alpha_3 D_k + \alpha_4 p_k + e_k \qquad (4)$$

To empirically assess the implications of the model, we use as dependent variable the rate of abstention in the local elections of 2001 in Portugal (*ABST*). We do not observe direct or indirect social interactions, or personal values of voters. Therefore, we use proxies for these variables in order to test the model. We take independent variables clustered into three groups: the *political density group*, corresponding to proxies for the indirect social interaction; the *fraction of loyal electors' group*, to proxy for the personal values; and the *social networks group*, to proxy for the direct social interaction.

4.1 The abstention rate

The abstention rate in an election can be defined in different ways according to different methods of computation. We will restrict ourselves to the concept of legal abstention, defined as the difference between the number of registered electors, at a certain moment of time, and the

[5] Information on Presidential and Parliamentary elections in Portugal is presented in the Appendix.
[6] See Figure A1 in the Appendix.
[7] See Figure A2 in the Appendix.

number of registered voters who actually participate in elections.

The use of the legal abstention method could be a problem since municipalities with high mortality rate would present a higher abstention rate than those with a low mortality rate (as electoral records have a lag in the registration of valid voters). Knowing that the electoral registers were all revised in 1998 and that the registry law changed in 1999, specifically to avoid these mistakes, this type of issue should not be a major problem. The argument is reinforced when noting that the data used in this work are from the European Parliament elections of 1999 and the local elections of 2001.

4.2 Political density

We understand political density variables to be those related to social pressure. This group of variables comprises the following: percentage of voting obtained by each of the major parties in previous (1997) local elections (*MP*), the purchasing power per capita index (*PPI*) by municipality in Portugal in the year 2000, and the annual number of periodical publications by municipality in 2002 (*PUB*). [8] We now describe each variable in detail.

The variable *MP* is composed of the percentage of votes in the four major political parties in Portugal, the *Social Democratic Party* (PSD), the *Socialist Party* (PS), the *Popular Party* (PP), and the *Communist Party* (which is the major party in a left-wing coalition, termed CDU). To allow for a more flexible characterization, we also consider each party separately. Since 1985, these parties have accounted for almost 90% of all the votes in every legislative election.[9] In order to measure the impact of the concentration of votes in a limited number of political parties we use a variant of the *Herfindahl-Hirschman Index (HHI)*, defined as the sum of the squares of the shares of each party. As such, it can range from 0 to 1. As the *HHI* tends to one, there would be a greater domination by one party, a clear source of social pressure.

As a proxy for income we use a purchasing power index per capita (*PPI*). As argued by Downs (1957) and Frey (1971), higher income motivates greater participation. As the level of income per capita increases, people show higher awaremess of the political and economic issues that surround them, the Government's performance, and local issues, and will tend to vote, rather than not. Thus, the wealth level of a community would be a source of social pressure. However, we can also argue that, as the purchasing power increases, economic well-being is guaranteed and, as a result, there is less motivation to vote, since no substantial improvement is expected from the election of a specific party. A voter must sacrifice some working/leisure time, implying an opportunity cost of voting that increases with income. Whether abstention increases with wealth or not is essentially an empirical issue. As there are no data on the income per capita in each municipality, in order to identify the sign of this effect we opt to use a purchasing power index as a proxy.[10] A possible alternative is to use the unemployment rate per municipality, but the purchasing power index gives a more representative idea of the financial power throughout the country.

The influence of the media is common to all the voters in an election, and as such is a source of social pressure or a way to form common ethical and social values. People will tend to vote more when the pressure exerted by media in electoral times, especially radio, television and newspapers is greater. The variable *PUB*, defined as annual number of prints of periodical publications by municipality in 2002, seeks to proxy the intensity of media activity across municipalities. The more information people obtain, the more the interaction within the social network, as people get to know what the other voters think. For this reason, this variable should also belong to the social networks group. Note that local periodical publications and radio tend to focus on local issues as well, while television is more national in scope (only in the Azores and Madeira Islands does Portugal have regional television broadcasts). To avoid comparing a publication with restricted circulation to one with wide dissemination in another municipality,

[8] The use of data of the year 2002 is explained at the end of Section 5.2.

[9] For further information see Veiga and Veiga (2004).

[10] To calculate this index, Instituto Nacional de Estatistica (INE), uses 18 variables, such as consumption of electricity per capita and the number of basic digital telephone accesses per capita. For more details see INE.

the total yearly number of prints of periodical publications (per inhabitant) is taken as our variable of interest to capture the role of regional press. No data for the influence of the different type of media are available, and we must rely on *PUB* to capture this effect.

4.3 Loyal voters

In order to proxy for personal values, we also allow for a role to be played by loyal electors. We consider as loyal electors those who have an idiosyncratic preference toward voting. As a proxy for the fraction of such loyal electors we use the abstention rate in the 1999 European Parliament elections, *EURO*. European Parliament eleÿÿions are usually regarded those that hÿÿe the least direct impact on voters' everyday lives. Thus, it is expected that the higher the level of abstention in the European elections in a given location, the lower the number of loyal electors, and hence, the higher the level of abstention in the local elections.

The hypothesis that aggregate levels of turnout vary in relation to an election's degree of importance for the functioning of the political system has already been tested by Baum and Freire (2002) (using data for Portugal) for the period 1975-2001. These authors found empirical support for the belief that electors give differential importance to different types of elections, providing support for our assumption that voters in the European Parliament elections are loyal voters.

4.4 Social networks

Finally, the role of social networks is evaluated using several proxy variables: the population living in each municipality in 2001 segmented by age groups; the geographic density of each municipality in 2002; the percentage of people within each municipality with a university degree/college education completed; and the number of attendances to primary health care centers.

Regarding the age segmentation, the reasoning is the following. People experience different motivations toward their participation in the political decisions according to the stage of their lives, their cultural, social, economic status and employment situation. The partition of the total population according to age subgroups is related to the fact that people tend to interact with peers, namely people of the same age, sharing the concerns of their generation. Our intention is to capture a measure of this type of social interaction intensity for each municipality. We proxy this social interaction effect by obtaining the data for population living in each municipality in three age brackets: 15 to 24; 25 to 64; and more than 64 years old.[11]

Citizens younger than 25 in the year 2001 (*YG*) have lived their entire life in the current democratic regime. Also, this age group may reflect the enthusiasm of voting for the first time. As such, they may attach less weight to whether or not their friends vote. The fact that they have always lived in a stable democratic regime, may lead them to underestimate the importance of voting, as if their rights are guaranteed by social construction. Which of these effects dominates their voting attitude is a purely empirical matter.

For the group between 25 and 64 years of age in 2001 (*ADULT*) it would be optimal to have further discriminated data, distinguishing individuals between 25 and 45 (who have experienced only the democratic regime) and older than 45 (who had a political life before the 1974 revolution). People over 45 are therefore expected to be politically more active in any election, with a relatively low abstention rate. However, these data are not publicly available and we shall deal with the broader group.

The population living in each municipality with more than 65 years of age in 2001 (*SENIOR*) is the group of Portuguese voters more affected by health and therefore mobility problems. A high level of abstention is expected for this group. It is worth mentioning that this age group is characterized by the fact that each of its members was politically active before 1974.[12]

Similar hypotheses, regarding voting behavior according to age group, were tested by Baum and Freire (2002). They concluded that older citizens tend to vote much more regularly.

[11] The legal age for voting purposes is 18, but data available consider only the age group 15-24 at the local level of disaggregation.

[12] The use of data of the year 2002 is explained at the end of Section 5.2.

The geographic density of each municipality in 2002 (*GDEN*), defined as population per squared km, seeks to measure the extent to which the voters' characteristics of a given municipality are urban or rural. The higher the geographic density, the denser the social networks are and the lower the level of abstention in a certain municipality should be. A counter effect would be that in small towns people more often know, and discuss politics, with neighbours than in large towns. We let the data speak on this issue.[13]

The percentage of people with university degree/college education in each municipality (*UNIV*) attempts to capture the degree to which a person tends to express and discuss their opinion about politics and social issues with peers. It would then be expected that social networks are clustered by degree of education. Therefore, we also use this variable as a potential explanatory variable of abstention rate. In fact, Guttman, Hilger and Schachmurove (1994) provided evidence that the more educated individuals are, the more they are expected to vote. However, Baum and Freire (2002) tested the influence of education in the abstention level and the conclusion was that education is an inconsistent predictor in Portuguese elections; this may be due to the very low levels of education among older citizens, who tend to vote much more than the younger.

The number of attendances to primary health care centers, (*HC*), captures in a different way the role of social networks. It is a proxy for the number of people with whom each person may discuss any sort of question (including political issues). In a primary health care center, people may have to wait for a long time before they are seen by a doctor. During elections periods, exchange of ideas and mutual influence may occur. As the number of attendances increases, the social network tends to develop more.

4.5 Sources of information

The variable used as dependent variable is collected from the institution that oversees the voting process, National Election Comission (*Comissão Nacional de Eleições* - CNE). The variables used as independent variables are collected either from the National Institute of Statistics (*Instituto Nacional de Estatística* - INE) or from the CNE. From the INE, we obtain data for the purchasing power index per capita (*PPI*); the population living in each municipality with ages between 15 and 24 years (*YG*); the population living in each municipality with ages 25 to 64 years old (*ADULT*); the population living in each municipality with more than 65 years of age in 2001 (*SENIOR*); the percentage of people with university degree/college education in each municipality (*UNIV*); the number of attendances to primary health care center (*HC*); the annual number of printed issues of periodical publications by municipality in 2002 (*PUB*); and the geographic density for each municipality in 2002 (*GDEN*). From the CNE, we obtain data for the percentage of votes won by each of the major parties in 1997 (*MP*); the rate of abstention in the European elections in 1999 (*EURO*); and the rate of abstention in the local elections of 2001 (*ABST*).

The geographic density of a region is a very stable parameter between 2001 and 2002. As we could not find the data for the geographic density in 2001, we opt for using the geographic density of each municipality in 2002.

The data for the number of attendances to primary health care center (*HC*) and the annual number of periodical publications by municipality (*PUB*) are also dated from 2002. For *HC*, the total population that attended primary health care centers is the same; there was no increase in the number of centers available per municipality and no significant change in the health-problems panorama in the country. Hence, *HC* should be stable. For *PUB*, there is no evidence of any increase or decrease in the number of periodical publications by municipality in this period of time. In addition, it is not reasonable to think that election results imply any sort of impact on these variables. It therefore seems that using 2002 data for the year 2001 produces little change.

The variable *PUB* introduces an additional issue: this information is only available for only some of the municipalities. Whenever *PUB* is included we have a reduction in the sample size available for estimation. In addition, we also consider the values of *PUB* and *HC* normalized by total population in each municipality.

[13] The use of data of the year 2002 is explained at the end of Section 5.2 and the way this variable is defined can be found in INE.

There was a risk of a bias since an explanatory variable is not necessarily independent of shocks in the dependent variable. In this specific case a rise in the abstention rate could affect each of the parties involved in the election differently, taking into account that the data for the abstention rate (*ABST*) and the data for the percentage of voting obtained by each of the major parties in previous elections (*MP*) were initially from the same election, that is, the local elections of 2001. To avoid any reverse causality problem we prefer to use the data for the percentage of voting obtained by each of the major parties in the previous elections (*MP*).

Information on each variable is presented in Table A2 in the Appendix.

5 WHAT DETERMINES ABSTENTION RATES?

Estimation was performed using simple OLS. The estimated values for the coefficients in equation (5) and respective t-statistics can be found in Table 1 below, with column I holding the baseline regression.

Table 1: Empirical determinants of abstention rates

	(I)	(II)	(III)	(IV)	(V)	(VI)
PSD	-0.311	-0.106	-0.399			-0.340
	(-2.36)	(-1.81)	(-3.66)			(-2.61)
PS	-0.366	-0.135	-0.462			-0.403
	(-2,59)	(-2.29)	(-3.99)			(-2.89)
CDU	-0.191	-0.043	-0.282			-0.194
	(-1.37)	(-0.73)	(-2.40)			(-1.41)
PP	-0.356	-0.132	-0.454			-0.412
	(-2.58)	(-2.12)	(-4.07)			(-3.01)
PSD+PP+PS				-0.152		
				(-4.41)		
IND					3.238	
					(2.16)	
PPI	0.089	0.152	0.097	0.084	0.104	0.095
	(3.66)	(7.20)	(4.04)	(3.45)	(4.06)	(3.68)
PUB	-0.060			-0.105	-0.154	-0.028
	(0.24)			(-2.50)	(-3.54)	(-0.55)
EURO	0.586	0.458	0.578	0.596	0.578	0.596
	(5.76)	(7.50)	(5.68)	(5.86)	(5.34)	(5.63)
YG	-1.358	-1.129	-1.291	-1.432	-1.908	
	(-3.24)	(-2.38)	(-3.11)	(-3.45)	(-4.47)	
ADULT	0.459	0.562	0.445	0.509	0.659	
	(3.78)	(4.11)	(3.67)	(4.32)	(5.50)	
SENIOR	0.319	0.371	0.249	0.265	0.409	
	(1.47)	(1.82)	(1.19)	(1.27)	(1.84)	
POP						0.070
						(1.74)
UNIV	0.044	-0.563	-0.045	0.097	0.018	0.219
	(0.25)	(-3.35)	(-0.28)	(0.55)	(0.10)	(1.28)
HC	-34.595	-43.856	-30.800	-40.036	-46.372	-28.251
	(-2.96)	(-3.38)	(-2.73)	(-3.54)	(-3.87)	(-2.48)
GDENS	0.936	1.084	0.989	0.952	0.926	0.754
	(1.60)	(1.56)	(1.69)	(1.64)	(1.50)	(1.24)
const.	25.537	6.848	34.637	7.498	-7.629	27.913
	(1.77)	(1.04)	(2.82)	(1.06)	(-1.12)	(1.94)
obs.	117	295	117	117	117	117
Adj. R^2	0.619	0.495	0.618	0.618	0.567	0.582
RESET test (P-value)	0.79 (0.50)	7.58 (0.00)	1.19 (0.32)	0.90 (0.44)	2.63 (0.05)	1.26 (0.29)
White's test (P-value)	101.88 (0.54)	131.62 (0.00)	87.30 (0.56)	78.17 (0.13)	89.45 (0.02)	81.24 (0.35)

Regarding the number of observations used, it is important to note that only 117 of the 300 observations were included in the regression above. This is due to the fact that the variable *PUB* is available only for 117 municipalities. As a check on the results we have also estimated using the total number of observations and the 117 observations used above, but now excluding the variable *PUB*. There is no qualititative changes in the results, as we can see by comparing columns II and III with column I in Table 1.

The statistical significance of the individual coefficients shows that the regression results for the political density group did not confirm the expected effects.

Concerning the percentage of votes obtained by each of the main parties (MP), we came to the surprisingly conclusion that, with a significance level of 5%, the (previous) votes for the Communist Party (CDU) are irrelevant in the determination of the abstention rate. The undoubtedly significant results are the Socialist Party (PS), the Social-Democrats (PSD) and the Popular Party (PP), with coefficients indicating a negative relationship between the percentage of votes in these two political parties and the level of abstention. In other words, whenever any of these political parties is relatively more influential, the abstention rate is lower. This is somewhat surprising, especially when compared with the situation of the Communist Party. It is generally argued by political commentators that the Communist Party has highly loyal voters and therefore, there should be a negative association between abstention votes and the share of votes by the Communist Party.

Looking more closely at PP and CDU - two small parties of a similar dimension - the clear difference concerning their association with the abstention level is surprising. One possible interpretation is that where PP (right-wing party) has a higher share of votes, local society feels that more is at stake and participation in elections increases.

As mentioned above, the expected effect of the purchasing power index per capita (PPI) in the rate of abstention (*ABST*) is ambiguous. On the one hand, as the level of income per capita increases, people are more aware of the political issues and, as such, they will tend to vote rather than not to (Downs, 1957, Frey, 1971 and Castanheira, 1999). On the other hand, people's economic security is guaranteed for high levels of income and, as a result, people should feel no need to vote. (Downs, 1957 and Castanheira, 1999) Our empirical results above support the latter view, namely that as the PPI increases, the abstention rate will tend to increase as well. In richer local communities, abstention is higher on average.

With regard to the political density group we conclude in favor of a negative impact of the annual number of periodical publications by municipality (*PUB*) in the abstention rate. That is, the more active the local media is, the greater the negative influence in the level of abstention it seems to have, showing, in our interpretation, that social pressure conveyed in this way does not affect abstention rates. This result is not robust, as under slightly different specifications (as reported in columns IV – VI of Table 1, described below), the coefficient is negative and statistically significant, suggesting that a more active local press results in lower abstention rates.

Since the coefficients values for PSD, PS and PP seem to be similar at the significance level of 7%, we ran an equality test for these coefficients. The results do not reject the null hypothesis of equality of coefficients (see column IV in Table 1).

Next, we address the effect associated with significant concentration of votes in a limited number of political parties. In order to check for this effect, we use a *Herfindahl Index* computed over voting shares (*IND*) replacing parties' shares of the total vote. The estimates are in column V of Table 1.

The positive and statistically significant coefficient on *IND* shows that as the concentration of votes in one or more political parties increases, the abstention rate does too. One possible explanation for this positive relationship is as follows. Realizing the massive influence of one party, an elector may feel that the vote has even less weight than in normal circumstances in the local municipality, reinforcing the reasons not to vote.

The impact of loyal electors is in accordance with expectations: the level of abstention in the European Parliament elections is positively correlated with the rate of abstention in the local elections. Local municipalities facing a 1% increase in the abstention rate in the European

Parliament elections will see a rise the abstention level in the local elections by 0.59%. The fact that the European elections do not have any direct impact on the lives of the voters and that Portugal has a small influence on the European Parliament lead to a situation where people have a reduced incentive to vote in these elections. Those who do turn out are voters who feel their social and political obligations very strongly. As such, they will most likely vote in the local elections as well, which have a greater impact on their everyday lives, and influence others to vote as well. The relevance of social interaction matters more in local elections.

With respect to the social network group of variables, evidence was found that the younger the population, the lower the level of abstention, and that the share of the population with more than 65 years of age does not seem to influence the level of abstention across municipalities. To make our analysis more directly related to Castanheira (1999), we replace age groups by the municipal population itself. The regression results are reported in Table 1, column VI. We conclude for a positive impact in the abstention level as the total population per municipality increases, which matches the results obtained by Castanheira (1999).

We also find the expected effect on the abstention rate of the number of attendances to primary care health center (*HC*): as the number of attendances increase, there is a negative impact on the rate of abstention. An increase in the number of social contacts is expected to generate a higher turnout in elections.

Neither the percentage of people with university degree/college education in each municipality variable (*UNIV*) nor the geographic density variable (*GDEN*) do not have a significant impact on the abstention level.

6 ROBUSTNESS OF RESULTS

A relevant issue is heteroskedasticity. The 117 municipalities used in the main estimations are of similar size. However, we have used two independent variables to capture the dimension of each municipality, namely the geographic density and the total population living in each municipality. There is no evidence of heteroskedasticity, therefore the usual assumption that the random error term was uncorrelated and identically distributed is admissible. The White test and the RESET test for model misspecification do not reject the standard OLS assumption, except for the variants of the basic model reported in columns II and V in Table 1, which are rejected in favor of other models.

The nature of the data also suggests that a proportions data model could be used as the basis for the estimation strategy. The grouped data approach implies estimation with a dependent variable given by:[14] $\ln(ABST_k/(1-ABST_k))$. The results for the estimates of this new dependent variable on the same set of covariates (as described in previous sections) are in Table 2 below. The variables impacting in the abstention level, for a significance level of 5%, are exactly the same in this model and in the OLS estimates produced earlier. In both models the signs of the effects of the independent variables in the abstention are the same: *PSD*, *PS*, *PP*, *YG*, *PUB* and *HC* have a negative impact in the level of abstention; on the other hand, *PPI*, *EURO* and *CDU* contribute to an increase in the abstention in elections. Again, the variables representing the interactions between the variables, *GDEN*, *EURO*, *HC* and *UNIV*, have no impact in the abstention rate.

Our final set of estimates, in column I of Table 3, normalizes absolute values by population size, yielding the same qualitative results.

As a further robustness check on our results we allow for the existence of interaction effects between independent variables. Namely, we analyze the relationship between the density of social contacts and the existence of a positive attitude toward the participation in elections. We evaluate the effect of age groups on the marginal effect in the abstention level of *GDEN*, *EURO*, and *HC* variables in the abstention level. From the estimation, we conclude from Table 3, column II, that the variables representing the interactions between the above-mentioned variables have no impact on the abstention rate (for a significance level of 5 %).

[14] See Greene (1993).

Table 2 – Dependent variable ln(A/(1-A))

	(I)	(II)	(III)
PSD	-0.013	-0.005	-0.017
	(-2.32)	(-1.79)	(-3.63)
PS	-0.158	-0.006	-0.020
	(-2.56)	(-2.29)	(-3.97)
CDU	-0.008	-0.002	-0.012
	(-1.38)	(-0.81)	(-2.44)
PP	-0.015	-0.006	-0.019
	(-2.55)	(-2.07)	(-4.06)
PPI	0.004	0.007	0.004
	(3.66)	(7.19)	(4.05)
PUB	-0.003		
	(-1.22)		
EURO	0.025	0.020	0.025
	(5.62)	(7.38)	(5.53)
YG	-0.062	-0.051	-0.059
	(-3.40)	(-2.41)	(-3.26)
ADULT	0.021	0.025	0.019
	(3.85)	(4.04)	(-3.74)
SENIOR	0.015	0.018	0.012
	(1.56)	(1.94)	(1.27)
UD	0.002	-0.026	-0.002
	(0.22)	(-3.42)	(-0.33)
HC	-1.484	-1.914	-1.323
	(-2.91)	(-3.26)	(-2.67)
GDENS	0.038	0.044	0.041
	(1.49)	(1.40)	(1.59)
const.	-1.021	-1.887	-0.611
	(-1.61)	(-6.35)	(-1.13)
obs.	117	295	117
Adj. R^2	0.609	0.459	0.607
RESET test (P-value)	1.02 (0.39)	7.58 (0.00)	1.34 (0.27)
White's test (P-value)	103.99 (0.48)	117.80 (0.03)	91.14 (0.45)

Table 3: Robustness checks

	(I)	(II)		
PSD	-0.267	-0.381	EURO x YOUNG	0.013
	(-3.16)	(-2.38)		(0.12)
PS	-0.324	-0.384	EURO x ADULT	0.006
	(-3.57)	(-2.25)		(0.15)
CDU	-0.145	-0.269	EURO x SENIOR	0.020
	(-1.60)	(-1.60)		(0.25)
PP	-0.310	-0.379	HC x YOUNG	-5.628
	(-3.45)	(-2.28)		(-1.25)
PPI	0.056	0.030	HC x ADULT	1.272
	(1.65)	(0.94)		(0.75)
PUB	0.001	0.188	HC x SENIOR	2.150
	(0.17)	(1.63)		(0.58)
EURO	0.601	0.319	UD x YOUNG	0.112
	(5.75)	(1.42)		(1.10)
YG	-109.97	-4.125	UD x ADULT	-0.041
	(-2.64)	(-0.58)		(-1.38)
ADULT	-30.57	0.866	UD x SENIOR	0.010
	(-1.89)	(0.36)		(0.27)
SENIOR	-3.010	-1.793	GDENS x YOUNG	1.042
	(-0.12)	(-0.38)		(0.82)
UNIV	-1.852	0.74	GDENS x ADULT	-0.173
	(-3.24)	(1.75)		(-0.62)
HC	1.117	-52.382	GDENS x SENIOR	-0.362
	(1.97)	(-2.29)		(-1.05)
GDEN	52.082	0.694	Constant	50.063
	(3.80)	(0.40)		(2.36)
obs.	117	117		
Adj. R^2	0.602	0.642		
RESET test	0.57	0.41		
(P-value)	(0.64)	(0.75)		
White's test	92.90			
(P-value)	(0.39)			

7 CONCLUDING REMARKS

We document the importance of social interaction in elections turnout providing empirical support for the role of social networks in determining the level of abstention in elections. The weight of social pressure in the decision to vote (or to abstain) is clearly revealed by the analysis. Besides social pressure, other factors that facilitate contacts between the individual voters, allowing them to have more extensive networks, also seem to reduce abstention levels. One needs to be careful in the interpretation of these indicators, as they result from indirect inference. One limitation of the empirical application was the absence of direct measures on the average social network and on valuations of common opinions with members of one's social network. Also, we find a significant role for idiosyncratic (or personal) values in the decision to vote or not to vote.

Besides reporting the direct measure of the proxies for these three main forces underlying the decision to vote or not to vote, we can add a discussion of some of the empirical implications of the model. First, recall that the equilibrium turnout is supposed to be smaller whenever the ability to accommodate divergence within society increases. The more democratic a society is, the more room there should be for the expression of different opinions, the social pressure effect becoming less relevant. Our results clearly reveal that younger people tend to abstain less than older citizens, implying that more mature voters are more able to accommodate different opinions within their network of social contacts. Younger voters may be less open to different opinions, probably due to the enthusiasm typical of their age, and lack of the required experience to listen and to understand others.

Second, we obtain results similar to those of Castanheira (1999), with higher abstention levels as the total population per municipality increases. Abstention is higher in more populated areas, meaning that the networking effect is possibly less diluted in smaller population centres. Alternatively, we may think that more populous centers are more open to different ideas, decreasing the impact of social pressure, thus leading to a higher abstention according to the model. This suggests that there may be a correlation between the dilutions of the social network effects and the increase in the ability to accommodate divergence within society – a subject for further empirical analysis.

Regarding the social pressure effects, in contrast to the conclusion reached by Downs (1957) and Frey (1971), we found, surprisingly, that as the purchasing power increases, the abstention level increases. As explained above, this may be the result of an increase in individuals' opportunity cost of voting as their incomes rise.

There are many different directions for future theoretical and empirical research, namely to explain why both the percentage of people with university degree/college education in each municipality variable and the geographic density variable have no significant impact on the abstention level.

REFERENCES

AMARO DE MATOS, J. AND BARROS, P.P. (2003). "Social norms and the paradox of elections' turnout". *Public Choice* 121:239-255.

BARTOLINI, S. (2001), "A Integração Europeia Provocará uma Reestructuração dos Sistemas de Clivagens Nacionais?", *Sociologia - Problemas e Práticas*, 37: 91-114.

BAUM, M. AND FREIRE, A. (2002) "Second-Order Elections in Democratic Portugal, 1975-2001". Paper presented at the Conference on Portugal and Southern European Politics, University of California at Berkeley, Center for European Studies, October 31 – November 2.

CASTANHEIRA, M. (2001). "On the (non) paradox of (not) voting". Public Policy n.º 3126, Centre for Economic Policy Research.

CASTANHEIRA, M. (1999). "Should I stay or should I go?: Poisson games and the paradox of voting". Mimeo.IGIER.

COATE, S. AND CONLIN, M. (2002). "Voter turnout: Theory and evidence from Texas liquor referenda". NBER Working Paper 8720.

COUPÉ, T. AND NOURY, A. G. (2004), "Choosing not to Choose – On the link between information and abstention". *Economics Letters*, 84(2): 261-265.

DALTON, R.J. AND M. WATTENBERG (2000), editors, *Parties without partisans: Political Change in Advanced Industrial Democracies (Comparative Politics)*, Oxford University Press.

DOWNS, A. (1957). "*An economic theory of democracy"*. New York: Harper & Row.

ESPÍRITO SANTO, P. (2005), "As eleições autarquicas e regionais em Portugal", ISCTE, mimeo.

FEDDERSEN, T. AND SANDRONI, P. (2006). "A theory of participation in elections with ethical voters". *American Economic Review*, 96(4): 1271 – 1282.

FIELDING, D. (2000) "Social and economic determinants of English voter choice in the 1997 general election" *Public Choice*, 102:271-295.

FREIRE, A. AND MAGALHÃES, P. (2002). "*A abstenção eleitoral em Portugal"*, ICS, mimeo.

FREY, B. (1971). "Why do high income people participate more in politics?" *Public Choice* 11:10-15.

GALAM, S. AND MOSCOVICI, S. (1991). "Towards a theory of collective phenomena: consensus and attitude changes in group", *European Journal of Social Psychology*, 21: 481 – 495.

GOODIN, R.E. AND KLINGEMAN, H.-D. (1998). *A new handbook of political science*, Oxford University Press.

GRAY, M AND M. CAUL (2000). "Declining Voter Turnout in Advanced Industrial Democracies, 1950 to 1997", *Comparative Political Studies*, 33(9): 1091-1122.

GREENE, W. (1993) *Econometric Analysis*, Macmillan.

GUTTMAN, J., HILGER, N. AND SCHACHMUROVE, Y. (1994). "Voting as investment vs. voting as consumption: New evidence". *Kyklos* 47:197-207.

HARSANYI, J., (1980). "Rule Utilitarianism Rights, Obligations and the Theory of Rational Behavior." *Theory and Decision*, 11(2): 115 – 133.

INE, *Online statistics*, Instituto Nacional de Estatística, (www.ine.pt)

LANE, J.E. AND S. ERSSON (1999). *The New Institutional Politics: Outcomes and Consequences*. Routledge.

MYERSON, R. (1994). "Population uncertainty and Poisson games". Center for Mathematical Studies in Economics and Management Science Discussion Paper 1102R, Northwestern.

PALFREY, T.R. AND ROSENTHAL, H. (1985). "Voter participation and strategic uncertainty". *American Political Science Review 62*: 25-42.

PALFREY, T.R. AND ROSENTHAL, H. (1983). "A strategic calculus of voting". *Public Choice* 41: 7-53.

SIEG, G. AND SCHULTZ, C. (1995). Evolutionary dynamics in the voting game. *Public Choice 85:* 157-172.

SILVA E., COSTA J. (2006). "Are voters rationally ignorant? An empirical study of Portuguese local elections". *Portuguese Economic Journal*, volume 5- Number 1, pp. 31-44.

VEIGA, LINDA AND VEIGA, FRANCISCO (2004). "The Determinants of Vote Intentions in Portugal", *Public Choice*, 118: 341-364.

Appendix

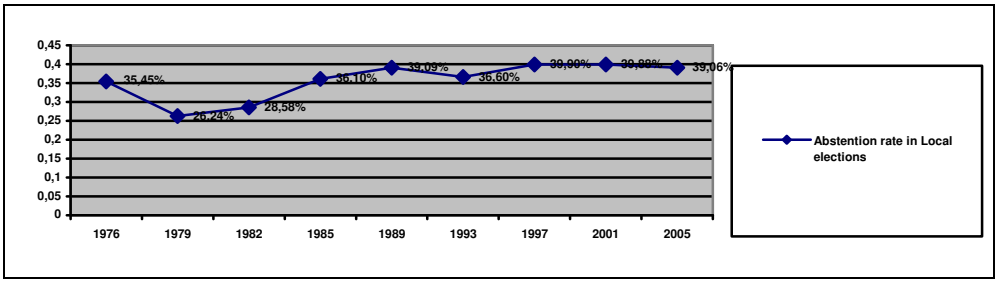
Table A1. Expected Effect in the rate of abstention in Local elections

Variable	Expected effects in rate of abstention in local elections
Percentage of voting obtained by each of the major parties (*MP*)	Positive
Purchasing power parity per capita index (*PPI*)	Uncertain
Annual number of periodical publications by municipality in 2002 (*PUB*) (million prints)	Negative
Rate of abstention in the European elections (*EURO*)	Positive
Population with ages between 15 and 24 years old (*YG*)	Positive
Population with ages between 25 and 64 years old (*ADULT*) ('000)	Negative
Population more than 65 years old in 2001 (*SENIOR*) ('000)	Positive
Geographic density of each municipality in 2002 (*GDEN*) (population per squared km)	Negative
The percentage of people with university degree/college education in each municipality (*UNIV*).	Negative
Number of attendances to primary health care center (*HC*) (millions of cases)	Negative

Table A2. Descriptive Statistics

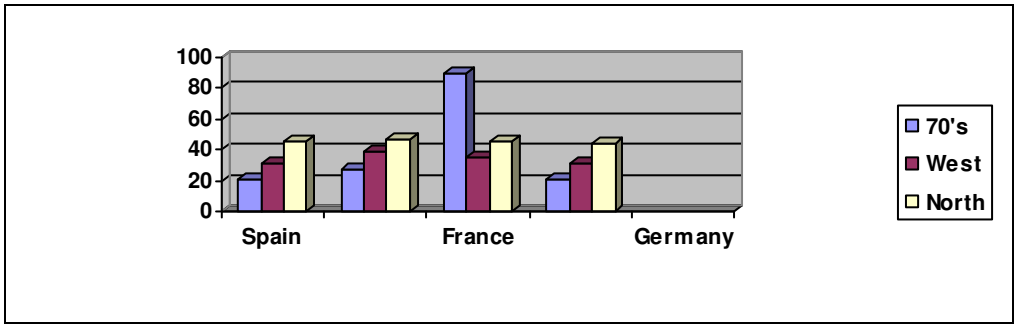
Variable	Mean	Std. Dev.	Min	Max
Abstention rate	38.051	7.363	21.700	54.640
PSD	37.277	14.825	0.000	67.440
PS	40.589	13.159	0.000	6..97
PP	8.097	11.527	0.000	61.170
CDU	10.071	13.433	0.000	52.630
EURO	60.358	4.346	50.400	73.540
YG	9.422	10.522	0.906	71.634
ADULT	35.459	42.331	2.812	294.171
SENIOR	10.016	13.909	1.380	133.304
UNIV	8.002	15.809	0.117	139.182
HC	0.177	0.220	0.000	173.679
PUB	5.872	34.915	0.000	355.525
PPI	80.532	37.051	33.380	305.190
GDENS	0.611	1.198	0.015	7.587
IND	0.419	0.306	0.271	3.621
POP	62.899	80.658	5.428	63.829
%YOUNG	0.155	0.023	0.107	0.243
%ADULT	0.564	0.040	0.447	0.657
%SENIOR	0.185	0.055	0.104	0.397
HC/POP	2.982	0.841	0.001	4.590
PUB/POP	26.682	70.281	0.008	556.994
UD/POP	0.095	0.040	0.004	0.234
POP	54.898	65.610	5.098	499.109

Figure A1. Evolution of the abstention rate in Local elections



Source: CNE

Figure A2. Abstention rate in European Parliament elections



Source: CNE

Presidential elections

In contrast to other countries, such as Spain, where elections had a determinant role in bringing down authoritarian regimes, in Portugal democracy began with the 1974 revolution. The first free and democratic elections took place on April 25th, 1975, with an outstanding turnout of 92%.

The President is elected for a term of 5 years, after which can (s)he be re-elected for a second term, and then cannot be re-elected. Since the 1974 Revolution the Portuguese have elected 4 Presidents (the first 3 being re-elected).

Since the 1982 revision of the Constitution, Portugal has had a semi-presidential system. This means that (a) the President is elected by universal suffrage; (b) the President possesses considerable powers; (c) the President nominates a Prime Minister who heads a government with executive power and can stay in office only if the parliament does not declare its opposition. (Duverger, 1980, p. 166)

The abstention rate in presidential elections tends to be lower in countries with semi-presidential systems. (Freire and Magalhães, 2002) On the other hand, the existence or not of compulsory voting in these elections is another issue that contributes to a lower or higher rate of abstention.

The abstention rate in the Portuguese presidential elections was 24.6% in the 1970s and 26.5% in the 1990s. Bearing in mind that Portugal has a semi-presidential system with no compulsory voting, one can conclude for a low rate of abstention in these Portuguese elections.

Parliamentary elections

The Portuguese Parliament is composed of 230 members. The members of Parliament are elected for a term of four years, after which they can be re-elected. The Parliamentary elections are considered to be the most important ones, since they determine the basis of support of the Prime Minister and his or her Government, who will be responsible for the executive power.

As mentioned above, the Portuguese abstention rate is intimately related to Portugal's historical and political evolution. Once this political transition took place, the abstention level started to increase: in the 1983 elections it was 20%, increasing to 20.9% in 1995, 30.7% in 1999 and 37.7% in 2002, as illustrated in Figure A1.

ACERCA DAS CONSEQUÊNCIAS ECONÓMICAS DAS ELEIÇÕES EVENTUALMENTE ANTECIPADAS

ANTÓNIO CALEIRO
E-mail: caleiro@uevora.pt
Departamento de Economia & CEFAGE
Universidade de Évora (Portugal)

Resumo

Este trabalho analisa algumas das consequências económicas resultantes da existência (ou da mera possibilidade de existência) de eleições antecipadas. Com base num modelo de oferta agregada mostra-se, em primeiro lugar, que a data das eleições, quer determinada de forma endógena quer de forma exógena, é relevante na colocação em causa do conhecido princípio da neutralidade das políticas económicas em contexto de expectativas racionais. Mostra-se também que, mesmo que se admita somente a possibilidade de existência de eleições antecipadas, este facto origina uma diminuição das vantagens eleitorais dos partidos de direita face aos partidos de esquerda resultantes das suas políticas económicas enquanto partidos formando o governo.

Palavras-chave: Eleições, Inflação, Oferta agregada, Políticas económicas partidárias.
Área temática: Modelação Económica e Financeira.

Abstract

The paper offers an analysis of some of the consequences due to the possibility of early elections being called. Based upon an aggregate supply model, it is shown, in the first place, that elections, whether exogenously or endogenously determined, are relevant to challenge the well known neutrality principle of economic policies under rational expectations. Furthermore, it is also shown that, in the particular case of being possible to call early elections, the electoral advantages (in terms of the associated economic policies) of right-wing parties in relation to left-wing parties undergo a diminishment.

Key words: Aggregate supply, Elections, Inflation, Partisan economic policies.
Thematic area: Economic and Financial Modelling.

1. INTRODUÇÃO

Apesar do enorme desenvolvimento da literatura sobre os ciclos eleitorais, a determinação do momento óptimo para a ocorrência das eleições, ou mesmo o estudo das consequências de ser possível a convocação de eleições antecipadas, têm sido assuntos amplamente ignorados.[1] Algumas excepções a esta ignorância generalizada por parte da literatura são, do ponto de vista empírico, Cargill & Hutchison (1991) e Ito (1989,1990), Lupia & Strøm (1995), Smith (1996), e, do ponto de vista teórico, Balke (1990,1991), Chappell & Keech (1986), Chappell & Peel (1979), Ito & Park (1988), Keech & Simon (1985), Lupia & Strøm (1995), Smith (1996), Kayser (2005), Keppo et al. (2008), Bierbrauer & Mechtenberg (2008). Na maior parte destes estudos, em que a vertente económica foi tida em conta, foram considerados modelos da, chamada, 'primeira geração' de modelos de ciclos eleitorais, i.e. modelos a tempo contínuo, admitindo expectativas adaptativas (particularmente o tipo de modelos analisado no artigo pioneiro de Nordhaus, 1975). O modelo que será considerado neste trabalho é representativo da

[1] Ainda assim pode referir-se o facto de existirem obras cujo campo de análise é, exclusivamente, o tempo das eleições (por exemplo, Smith 2004).

'segunda geração' de modelos de ciclos eleitorais, ou seja, um modelo a tempo discreto, admitindo a existência de expectativas racionais.

O problema da determinação de uma frequência óptima das eleições pode ser abordado de duas formas, de acordo a quem, Governo ou sociedade, aquele carácter óptimo diz respeito. *Grosso modo*, os primeiros estudos sobre o assunto consideraram o ponto de vista da sociedade, uma vez que o problema foi resolvido através da determinação da duração do mandato que levaria a resultados que se aproximassem, o mais possível, daqueles obtidos por um Governo do tipo 'benevolente social'.[2] Os estudos mais recentes, em geral, não seguem esta abordagem, na medida em que a duração óptima do mandato é determinada endogenamente, i.e. fazendo parte do programa óptimo de um Governo com motivações eleitorais (Lächler, 1982).[3]

Ainda que a data das eleições não seja alvo de escolha (óptima) por parte do Governo, podendo, inclusive, ser antecipada em resultado da necessidade de sanar uma crise política, i.e. num momento do tempo, eventualmente, inconveniente do ponto de vista do Governo, a simples possibilidade de existência de eleições antecipadas tem, por si só, consequências (de natureza económica) importantes.

Como adiante se tornará mais claro, a possibilidade de ser convocada uma eleição antecipada introduz incerteza adicional, *em cada período do mandato*, com a qual o Governo tem que lidar. Em termos mais específicos, a incerteza dos resultados eleitorais gera uma 'surpresa' temporária nas variáveis de expectativas dos agentes económicos (por exemplo, na taxa de inflação esperada), a qual não pode ser atribuída a uma escolha deliberada do Governo. Neste sentido, enquanto em contexto de datas de eleições exogenamente determinadas, o Governo pode ignorar a incerteza quanto aos resultados eleitorais, em contexto de eleições com datas possivelmente antecipadas (por exemplo, endogenamente determinadas) é introduzida uma (outra) fonte de incerteza, em cada momento do tempo, que o Governo não pode ignorar.

Este trabalho pretende, assim, formalizar os argumentos acima apresentados, através da análise de algumas das consequências económicas resultantes da existência (ou da mera possibilidade de existência) de eleições antecipadas. Para tal apresenta a seguinte estrutura: na secção 2, procede-se à formalização da forma como a existência de eleições pode ser um factor conducente a alterações bruscas nas variáveis económicas relevantes. A secção 3 analisa, então, algumas das consequências económicas associadas à possibilidade de existência de eleições antecipadas. A secção 4 termina este trabalho através da apresentação das suas principais conclusões e das possibilidades de novas vias de análise.

2. FORMALIZAÇÃO GERAL

Como é sabido a partir da literatura, a presença de agentes económicos racionais (e prospectivos) induz soluções óptimas caracterizadas por 'perturbações' ou 'choques' em torno de momentos cruciais, tais como a data das eleições. Neste sentido, assumindo uma abordagem partidária, a incerteza intrínseca sobre os resultados das eleições, que, portanto, podem ser considerados 'novidades' [*news*], leva a um 'salto' súbito nas variáveis económicas relevantes, tais como as taxas de inflação ou de câmbio no dia das eleições.[4]

Aquele argumento pode ser formalizado (e generalizado), considerando um modelo estrutural da seguinte forma:

$$\begin{bmatrix} y_{t+1} \\ z^{e}_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} y_t \\ z_t \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{12} \end{bmatrix} x_t , \tag{1}$$

[2] Seguindo esta perspectiva, Keech & Simon (1985) mostram que a duração óptima do mandato, num modelo do tipo considerado em Nordhaus (1975), depende, fundamentalmente, da parametrização do modelo.

[3] A literatura costuma fazer referência às chamadas eleições 'de estalo' [*snap*] como sendo aquelas em que, por vontade do Governo, são convocadas para obter benefícios eleitorais como, por exemplo, aqueles resultantes de uma conjuntura económica favorável e/ou de uma crise política na oposição.

[4] Para o caso das taxas de câmbio veja-se, por exemplo, Gärtner (1986). O mesmo tipo de trajectória para as taxas de câmbio pode ser obtido, num contexto político, se o Governo usar a taxa de câmbio, enquanto instrumento, de forma a explorar uma estrutura em curva J numa pequena economia aberta (van der Ploeg, 1989; Caleiro, 2008).

onde

- y_{t+1} é uma variável de estado, pré-determinada em t;[5]
- z_t é uma variável de expectativas, livre de assumir qualquer valor no momento t, com valor esperado, determinado em t, para o próximo período, z_{t+1}^e ; e
- x_t é uma variável instrumental, sob controle do Governo.

Para o caso das expectativas racionais, i.e. quando $z_{t+1}^e = \mathrm{E}[z_{t+1} \mid I_t]$, Blanchard & Kahn (1980) mostram que a única solução – para o modelo de 'caminho-sela' [*saddle-path*] estável – é dada por:

$$\begin{aligned} y_t &= y_0 \quad \text{para } t = 0 \\ &= \lambda_1 y_{t-1} + b_{11} x_{t-1} + \left((\lambda_1 - a_{11})\lambda_1 - a_{12}\lambda_2\right)\sum_{i=0}^{\infty} \lambda_2^{-i-1} \mathrm{E}[x_{t+i-1} \mid I_{t-1}], \ \text{para } t > 0; \end{aligned} \tag{2}$$

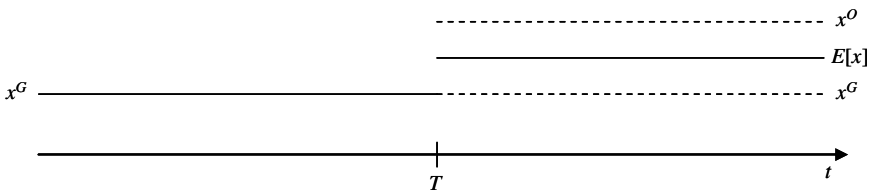
$$z_t = a_{12}^{-1}\left((\lambda_1 - a_{11})y_t + \left((\lambda_1 - a_{11})\lambda_1 - a_{12}\lambda_2\right)\sum_{i=0}^{\infty} \lambda_2^{-i-1} \mathrm{E}[x_{t+i} \mid I_t]\right), \ \text{para } t \geq 0, \tag{3}$$

onde λ_1 e λ_2 são, respectivamente, os valores próprios estável e instável da matriz **A** em (1), ou seja $|\lambda_1| < 1$ e $|\lambda_2| > 1$.

Assim, tal como se pode verificar a partir de (2) e (3), a dependência das variáveis de estado e de expectativas em relação aos valores esperados para as variáveis instrumentais torna claro como a existência de eleições, das quais podem resultar alterações na política económica, implica mudanças nas trajectórias das soluções para aquelas variáveis.

Vamos supor que o Governo (em função) se encontra comprometido com uma política económica fixa x^G e que uma eleição terá lugar em $t = T$, a qual, no caso de ser vencida pela oposição, irá corresponder a uma mudança na política económica para x^O. Isto significa que, até que a eleição ocorra, ou seja, para $t \leq T$, a variável de expectativas tomará um valor baseado em x^G, enquanto para depois das eleições tomará valores que dependem de x^G e de x^O, tal como se ilustra na figura 1.

Figura 1: A incerteza induzida pelas eleições



Porque z_t, para $t \leq T$, dependerá de valores conhecidos e esperados da política económica, então, em termos algébricos:

$$z_t = a_{12}^{-1}\left((\lambda_1 - a_{11})y_t + \left((\lambda_1 - a_{11})\lambda_1 - a_{12}\lambda_2\right)\left(\sum_{i=0}^{T-t} \lambda_2^{-i-1} x_{t+i}^G + \sum_{i=T-t+1}^{\infty} \lambda_2^{-i-1} \mathrm{E}[x_{t+i}]\right)\right).$$

Contudo, logo que a eleição tem lugar e, assim, a mesma política é aplicada x^G – no caso da reeleição do actual Governo – ou executada uma política diferente x^O – no caso de uma derrota eleitoral do actual Governo – então $\mathrm{E}[x_{t+i}]$ pode ser substituído por x^G ou por x^O, o que irá corresponder a um 'salto' na variável de expectativas e, por via de (2), também um 'salto' na

[5] Uma variável pré-determinada toma valores em função somente de valores conhecidos em t, ou seja de variáveis no conjunto de informação I_t. Assim, $y_{t+1} = y^e_{t+1}$, quaisquer que sejam as realizações das variáveis em I_{t+1}.

variável de estado. Na verdade, como o 'caminho-sela' depende dos valores esperados de variáveis de controlo, a ocorrência de eleições faz mudar a posição desta trajectória de equilíbrio. Assim, a fim de colocar o sistema sobre o 'novo' 'caminho-sela', as variáveis de expectativa terão que sofrer um 'salto'.

O que foi dito antes tem uma clara analogia com a análise de mudanças de política económica anunciadas ou não anunciadas (todavia, não no sentido de uma surpresa ou engano deliberado). Se o partido político vencedor das eleições literalmente mudar a política económica ao mesmo momento que sejam conhecidos os resultados eleitorais, este facto pode associar-se ao caso de uma alteração inesperada ou imprevista da política económica. Pelo contrário, se existir um desfasamento relevante antes de o partido vitorioso assumir o poder e começar a implementar novas políticas económicas, tal poderá estar relacionado com o caso de uma alteração previsível na política económica. Neste último caso, as 'novidades' são conhecidas no dia de eleição, mas dizem respeito a uma alteração na política económica que se verificará em algum momento no futuro.

De modo geral, o facto de os resultados eleitorais poderem ser considerados 'novidades' é de importância decisiva para a visão dos ciclos eleitorais partidários em contexto de expectativas racionais (Alesina, 1987; Alesina et al., 1997; Heckelman, 2002). Vamos então prosseguir com o estudo de algumas das consequências económicas associadas à possibilidade de convocação de eleições antecipadas, a partir de uma perspectiva racional partidária.

3. ALGUMAS DAS CONSEQUÊNCIAS ECONÓMICAS DE SER POSSÍVEL A CONVOCAÇÃO DE ELEIÇÕES ANTECIPADAS

De forma a melhor avaliar as consequências de uma convocação de eleições antecipadas, comecemos por considerar o caso em que a data das eleições se encontra exogenamente pré-determinada, enquanto situação de referência.

Assuma-se, então, que a popularidade do Governo em cada período do mandato, o qual decorre entre $t = 1$ e $t = T$, depende dos níveis de inflação, π_t, e de produto, Y_t, medido a partir do nível natural, $\overline{Y}$, i.e. $y_t = Y_t - \overline{Y}$, tal como se segue:

$$V = \sum_{t=1}^{T} \beta^t \left(-\tfrac{1}{2}\pi_t^2 + \theta y_t \right), \quad [6] \qquad (4)$$

onde, de acordo com uma curva de oferta agregada (Lucas, 1973):

$$y_t = \alpha\left(\pi_t - \pi_t^e\right), \qquad (5)$$

em que π_t^e representa a taxa de inflação esperada, para o período t, no momento t - 1, dada por

$$\pi_t^e = \mathrm{E}\left[\pi_t \mid I_{t-1}\right], \qquad (6)$$

e I_{t-1} é o volume de informação disponível no final do período t - 1.

Como se tornará adiante mais evidente, as consequências da convocação de eleições antecipadas são mais interessantes se existirem diferenças ideológicas, de acordo com o tipo de partido no poder. Considerem-se, então, dois tipos possíveis de governos, $i = L, R$, os quais diferem na ponderação atribuída ao nível de produto nas suas funções objectivo, as quais são dadas por:

[6] Desde Nordhaus (1975), esta é, efectivamente, a função objectivo mais considerada na área da macroeconomia política. Ainda assim, reconheça-se que a consideração de uma função objectivo duplamente quadrática pode conduzir a resultados qualitativamente diferentes, sobretudo quando o nível de produto exibe persistência (Caleiro, 2009), o que não é o caso em questão.

$$V^i = \sum_{t=1}^{T} \beta^t \left(-\tfrac{1}{2}\pi_t^2 + \theta^i y_t \right), \qquad (7)$$

onde $\theta^L > \theta^R$. Por outras palavras, os governos de direita ($i = R$) favorecem resultados menos inflacionários do que os governos de esquerda ($i = L$).

Tomando a inflação esperada, π_t^e, como um dado na optimização de (4) sujeita a (5) e (6), o Governo determina a taxa de inflação temporalmente consistente π_t^i tal como se segue:

$$\pi_t^i = \alpha\theta^i, \quad i = L, R. \qquad (8)$$

Para $2 \leq t \leq T$, a racionalidade das expectativas implica que $\pi_t^e = \pi_t$, o que resulta em $y_t = 0$, ou seja no funcionamento da economia ao seu nível natural.

Considerando eleitores racionais e prospectivos, o conhecimento das funções objectivo dos dois tipos de Governo permite o cálculo das suas taxas de inflação óptimas (8). Embora estas políticas económicas sejam conhecidas, uma vez que é assumido que a distribuição das preferências dos eleitores é desconhecida (Alesina et al., 1997: 55), os resultados eleitorais são desconhecidos. Por isso, para o período imediatamente após as eleições, $t = 1$, a taxa de inflação esperada deverá ser uma média de $\alpha\theta^L$ e $\alpha\theta^R$, ponderada pela probabilidade de cada tipo de Governo, i.e. de que cada partido esteja no poder.[7] Dito isto, para $t = 1$, a taxa de inflação esperada não coincidirá, em geral, com a taxa de inflação real, dado que

$$\pi_1^e = p\pi_1^L + (1-p)\pi_1^R = p\alpha\theta_1^L + (1-p)\alpha\theta_1^R,$$

onde p representa a probabilidade de uma vitória eleitoral do partido de esquerda e, naturalmente, $(1 - p)$ representa a probabilidade de uma vitória eleitoral do partido de direita.

Atendendo a que $\theta^L > \theta^R$, um mandato a começar no período $t = 1$ será caracterizado por:

$$\pi_1^e < \pi_1^L \Rightarrow y_1 = (1-p)\alpha^2\left(\theta^L - \theta^R\right) > 0 \Rightarrow Y_1 > \overline{Y},$$

no caso de um Governo de esquerda, e por

$$\pi_1^e > \pi_1^R \Rightarrow y_1 = p\alpha^2\left(\theta^R - \theta^L\right) < 0 \Rightarrow Y_1 < \overline{Y},$$

no caso de um Governo de direita, enquanto que para, $2 \leq t \leq T$, ter-se-á

$$\pi_t^e = \pi_t^L \Rightarrow y_t = 0 \Leftrightarrow Y_t = \overline{Y},$$

no caso de um Governo de esquerda, e

$$\pi_t^e = \pi_t^R \Rightarrow y_t = 0 \Leftrightarrow Y_t = \overline{Y},$$

no caso de um Governo de direita.

[7] Tal como na maior parte da literatura, vamos assumir probabilidades de eleição exógenas. Ellis (1991) é, tanto quanto sabemos, o único estudo a considerar probabilidades endógenas (veja-se também Ellis & Thoma, 1991).

Por outras palavras, em termos do chamado *efeito partidário*, o nível de produto é inferior/superior ao seu nível natural no primeiro período de um mandato de um Governo de direita/esquerda. Todos os outros períodos até a próxima eleição, como as expectativas se ajustam perfeitamente, o produto assumirá o seu nível natural, independentemente do tipo de Governo. A inflação, por sua vez, será sempre maior durante um Governo de esquerda. Assim, neste caso, o decréscimo/acréscimo na duração do mandato resultará, em caso de rotação no poder, numa maior/menor volatilidade no nível de produto e nas taxas de inflação.[8]

Dito isto, é possível determinar a utilidade registada pela sociedade durante o mandato de cada tipo de Governo. Assumindo que a utilidade social instantânea é dada por:

$$U_t = -\tfrac{1}{2}\pi_t^2 + \theta y_t, \tag{9}$$

e que a sociedade avalia os seus níveis de utilidade (9) de acordo com uma taxa de desconto (social) ρ, o nível de bem-estar social, durante um Governo de esquerda, é dado por:

$$\begin{aligned} U^L &= \rho\left(-\tfrac{1}{2}\left(\alpha\theta^L\right)^2 + \theta\left(\alpha(1-p)\left(\alpha\theta^L - \alpha\theta^R\right)\right)\right) + \sum_{t=2}^{T}\rho^t\left(-\tfrac{1}{2}\left(\alpha\theta^L\right)^2 + \theta(0)\right) \\ &= \rho\left(-\tfrac{1}{2}\left(\alpha\theta^L\right)^2 + \theta\left(\alpha^2(1-p)\left(\theta^L - \theta^R\right)\right)\right) - \tfrac{1}{2}\left(\alpha\theta^L\right)^2\frac{\rho^2 - \rho^{T+1}}{1-\rho}, \end{aligned} \tag{10}$$

enquanto que, no caso de um Governo de direita, é dado por:

$$\begin{aligned} U^R &= \rho\left(-\tfrac{1}{2}\left(\alpha\theta^R\right)^2 + \theta\left(\alpha^2 p\left(\theta^R - \theta^L\right)\right)\right) + \sum_{t=2}^{T}\rho^t\left(-\tfrac{1}{2}\left(\alpha\theta^R\right)^2 + \theta(0)\right) \\ &= \rho\left(-\tfrac{1}{2}\left(\alpha\theta^R\right)^2 + \theta\left(\alpha^2 p\left(\theta^R - \theta^L\right)\right)\right) - \tfrac{1}{2}\left(\alpha\theta^R\right)^2\frac{\rho^2 - \rho^{T+1}}{1-\rho}. \end{aligned} \tag{11}$$

Claramente, para o bem-estar da sociedade, o tipo de partido no poder é, em geral, um facto relevante já que

$$U^L - U^R = \tfrac{1}{2}\left(\theta^L - \theta^R\right)\alpha^2 p\frac{2\theta(1-\rho) + \left(\rho^T - 1\right)\left(\theta^L + \theta^R\right)}{1-\rho} \tag{12}$$

pode assumir valores positivos ou negativos. Obviamente, o bem-estar social é o mesmo, não importando o tipo de Governo, quando (12) assume um valor nulo, o que ocorre para:

$$\theta^* = \frac{1-\rho^T}{1-\rho}\frac{\theta^L + \theta^R}{2}, \tag{13}$$

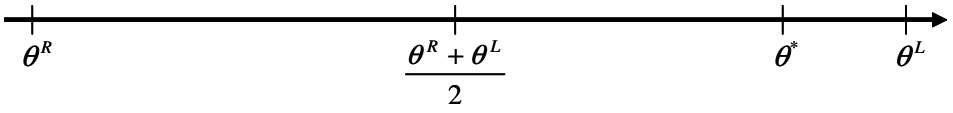
o que significa que

[8] Tal significa que podem existir valores intermédios de taxas de inflação $\pi^R < \pi^* < \pi^L$ tal que ambos os tipos de Governos registariam melhores resultados se escolhessem π^* em vez de π^L ou π^R porque a sub-otimalidade introduzida pelas flutuações na taxa de inflação e no produto seria reduzida. Como acabámos de referir, uma forma de reduzir essas flutuações, em média, (mas que não significa necessariamente uma redução no nível de perda) é através do aumento da duração do mandato. Alesina (1987) mostra que π^* é decrescente/crescente com a probabilidade de uma vitória eleitoral de direita/esquerda, pois o aumento do poder de negociação de cada partido irá fazer aproximar π^* de π^L ou π^R. Para que este mecanismo seja eficaz, ou seja considerado credível, tem que se assumir um horizonte temporal suficientemente longo para os dois tipos de partidos e uma suficientemente baixa taxa de desconto do futuro. Quase o mesmo argumento é considerado na literatura respeitante à utilização estratégica dos *défices* orçamentais (Milesi-Ferretti & Spolaore, 1994).

$$\theta^* > \frac{\theta^L + \theta^R}{2}.$$

Assim, de forma a ser indiferente perante os dois tipos de Governo, a sociedade deverá ser caracterizada por um parâmetro θ^* mais próximo de θ^L do que de θ^R. Este facto reflecte a vantagem eleitoral que os partidos de direita têm neste género de modelos, na medida em que o eleitor mediano preferirá, mais provavelmente, o Governo de direita do que o Governo de esquerda (veja-se a figura 2).

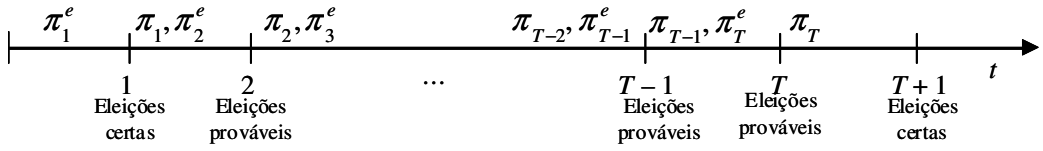
Figura 2: O espectro político para θ (caso 1)



Até agora assumiu-se uma duração exógena, ou seja, fixa do mandato, i.e. excluindo a possibilidade de convocação de eleições antecipadas. De forma a admitir esta possibilidade considere-se, em primeiro lugar, uma duração (fixa) máxima do mandato $\left(t = \overline{T}\right)$ e que τ representa a probabilidade de uma eleição ser convocada no momento *t*, a qual será vencida pelo partido de esquerda com probabilidade *p* e pelo partido de direita com probabilidade (1 – *p*).

A sequência de acontecimentos que irá ser considerada é a seguinte. No momento $t = 0$, é determinada a taxa de inflação esperada para o período seguinte, π_1^e. No momento $t = 1$ verifica-se a ocorrência de uma eleição, após a qual se determina a taxa de inflação por parte do partido que venceu essa eleição. O mandato tem uma duração máxima definida *a priori*, o que significa que, se não for convocada qualquer eleição, entre $t = 2$ e $t = \overline{T}$, então existirá uma eleição, certamente, em $t = \overline{T} + 1$. Em cada momento, a taxa de inflação esperada é determinada antes da ocorrência de uma eventual eleição antecipada e, consequentemente, antes da realização efectiva da taxa de inflação. A figura 3 mostra esta sequência de acontecimentos a partir de um novo mandato.

Figura 3: A sequência dos acontecimentos



Assim, pode facilmente verificar-se que

$$\pi_1^e = p\pi_1^L + (1-p)\pi_1^R$$

e que, para qualquer período, $2 \le t \le \overline{T}$,

$$\pi_t^e = (1-\tau)\pi_t^L + \tau\left(p\pi_t^L + (1-p)\pi_t^R\right), \qquad (14)$$

se o partido de esquerda estiver no poder no momento *t* – 1, enquanto que

$$\pi_t^e = (1-\tau)\pi_t^R + \tau\left(p\pi_t^L + (1-p)\pi_t^R\right), \qquad (15)$$

se o partido de direita estiver no poder no momento $t - 1$.

As expressões (14) e (15) para a taxa de inflação esperada mostram que:

- A simples possibilidade de ser convocada uma eleição antecipada implica, por si só, que os mandatos de Governos de esquerda se caracterizem por um valor superior de taxa de inflação esperada e, via (8), também por uma taxa de inflação superior àquelas que correspondem ao mandato de um Governo de direita. Somente em $t = \overline{T}$, ou se os agentes económicos tiverem a certeza que irá ser convocada uma eleição antecipada, é que a taxa de inflação esperada é a mesma, independentemente do partido que esteve no poder no momento anterior.[9]

- Enquanto existirem diferenças partidárias, ou seja $\pi_t^L \neq \pi_t^R$, *qualquer* mandato será caracterizado por 'surpresas' inflacionárias *em cada momento*, $\pi_t^e \neq \pi_t^i$, $\forall\ t,\ i = L,R$.[10] Esta é uma diferença substancial para o caso em que a data das eleições é exógena. Além disso, o bem conhecido princípio da neutralidade da política económica com expectativas racionais não é, por este facto, verificável. Em particular, durante um mandato de um governo de esquerda/direita o nível de produto estará acima/abaixo do seu nível natural já que $\pi_t^e < \pi_t^L / \pi_t^e > \pi_t^R$. De facto, é possível verificar que:

$$\pi^R < \pi^e_{2,\ldots,T} \quad (\text{se } R \text{ estiver no poder}) \quad < \pi_1^{\text{e}} < \pi^e_{2,\ldots,T} \quad (\text{se } L \text{ estiver no poder}) \quad < \pi^L.$$

Assim, assumindo uma função objectivo do tipo:

$$U = \sum_{t=1}^{\overline{T}} \rho^t \left(-\tfrac{1}{2}\pi_t^2 + \theta y_t \right),$$

é menos provável que os Governos de direita obtenham melhores resultados do que os Governos de esquerda do que no caso das eleições terem uma data fixa, exogenamente determinada.

Para melhor analisar este facto, suponhamos que, apesar de ser possível a convocação de uma eleição antecipada, ainda assim não se verifica qualquer eleição deste tipo durante todo o mandato. Se assim for, o bem-estar social, no caso de um Governo de esquerda, é dado por:

$$U^L = \rho\left(-\tfrac{1}{2}\left(\alpha\theta^L\right)^2 + \theta\left(\alpha(1-p)\left(\alpha\theta^L - \alpha\theta^R\right)\right)\right) + \sum_{2}^{T} \rho^t \left(-\tfrac{1}{2}\left(\alpha\theta^L\right)^2 + \theta\left(\alpha\left(\alpha\theta^L - \left((1-\tau)\alpha\theta^L + \tau\left(p\,\theta^L + (1-p)\theta^R\right)\right)\right)\right)\right),$$

enquanto que no caso de um Governo de direita é dado por:

$$U^R = \rho\left(-\tfrac{1}{2}\left(\alpha\theta^R\right)^2 + \theta\left(\alpha p\left(\alpha\theta^R - \alpha\theta^L\right)\right)\right) + \sum_{2}^{T} \rho^t \left(-\tfrac{1}{2}\left(\alpha\theta^R\right)^2 + \theta\left(\alpha\left(\alpha\theta^R - \left((1-\tau)\alpha\theta^R + \tau\left(p\,\theta^L + (1-p)\theta^R\right)\right)\right)\right)\right).$$

[9] Note-se que a diferença algébrica entre as expressões (15) e (16) é $(\pi_t^L - \pi_t^R)(1 - \tau) \geq 0$ e que esta não depende das probabilidades, p e, naturalmente, também não de $(1 - p)$.

[10] Note-se que para π_t^e, dada por (15) ou (16), ser igual a π_t^L ou π_t^R ter-se-ia que ter $\pi_t^L = \pi_t^R$.

Neste caso, a diferença entre os dois níveis de utilidade social é dada por:

$$U^L - U^R = \frac{1}{2}\left(\theta^L - \theta^R\right)\alpha^2\rho\frac{\left(\theta^L + \theta^R\right)\left(\rho^T - 1\right) + 2\theta\left(1 - \rho + \tau\rho - \tau\rho^T\right)}{1-\rho},$$

o que mostra que, para que a sociedade seja indiferente entre os dois tipos de Governo, deve ser caracterizada por um parâmetro

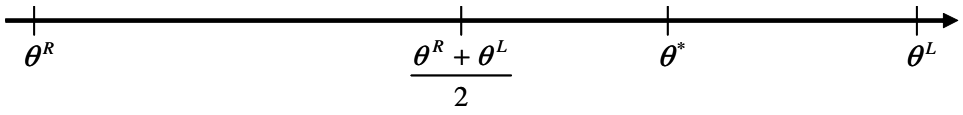
$$\theta^* = \frac{1-\rho^T}{1-\rho+\tau\left(\rho-\rho^T\right)}\frac{\theta^L + \theta^R}{2}, \qquad (16)$$

o que, mais uma vez, significa que

$$\theta^* > \frac{\theta^L + \theta^R}{2}.$$

Comparando (13) com (16) pode verificar-se que, neste caso, existe menor probabilidade de o eleitor mediano preferir o Governo de direita (comparem-se as figuras 2 e 4).

Figura 4: O espectro político para θ (caso 2)



Para finalizar, gostaríamos de salientar que a probabilidade de convocar uma eleição antecipada foi considerada constante. Em termos teóricos, tem sido reconhecido que os incentivos para a convocação de uma eleição antecipada aumentam à medida que se aproxima a data máxima para a ocorrência de uma eleição, $t = \overline{T}$, o que significa que a probabilidade τ deverá ser monotonicamente crescente com $t \to \overline{T}$ (Balke, 1990).[11] Esse facto parece ser também confirmado empiricamente (Caleiro, 2000). Assim, partindo do princípio de que τ é monotonicamente crescente com *t*, tal mostra imediatamente que, durante um Governo de esquerda/direita o nível esperado da inflação π_t^e irá aumentar/decrescer de π_t^L / π_t^R para $\pi_t^* \equiv p\pi_t^L + (1-p)\pi_t^R$.[12]

4. CONCLUSÃO

Este trabalho apresentou algumas das consequências associadas à possibilidade de existência de eleições antecipadas. Em primeiro lugar, mostrou-se que as eleições, quer a sua data seja exógena ou endogenamente determinada, são relevantes na colocação em causa do bem conhecido princípio da neutralidade da política económica perante expectativas racionais. Por outro lado, no caso particular de ser possível admitir eleições antecipadas, as vantagens eleitorais dos partidos de direita em relação aos partidos de esquerda sofrem uma diminuição.

[11] Como se mostra em Balke (1990), o Governo tem tanto mais incentivos para convocar uma eleição antecipada quanto há mais tempo esteja no poder (veja-se também Kayser, 2005). Em estreita ligação com este resultado, quanto maior for a duração máxima do mandato, menor o incentivo para convocar uma eleição no momento *t*.

[12] De facto, este aspecto pode ser alvo de uma análise mais formal e/ou completa, o que poderá constituir uma nova via de análise posterior a este trabalho.

Em termos de novas vias de análise, existem, pelo menos duas, que se julgam merecedoras de consideração. Em primeiro lugar, poder-se-á admitir uma alteração na função objectivo dos Governos no sentido de as considerar duplamente quadráticas (i.e. quadráticas também em relação ao nível de produto). Em segundo lugar, poder-se-á considerar (também) um modelo em que o produto exibe persistência, ou seja em que o produto do momento *t* dependa também do valor registado no momento anterior, tal como os mais recentes modelos de ciclos eleitorais têm vindo a admitir (Gärtner, 2000; Caleiro, 2009).

REFERÊNCIAS BIBLIOGRÁFICAS

ALESINA, ALBERTO (1987), "Macroeconomic Policy in a Two-Party System as a Repeated Game", *The Quarterly Journal of Economics*, Vol. CII, Agosto, pp. 651-678.

ALESINA, ALBERTO, NOURIEL ROUBINI, & GERALD D. COHEN (1997), "Political Cycles and the Macroeconomy", The MIT Press, Cambridge (MA).

BALKE, NATHAN S. (1990), "The Rational Timing of Parliamentary Elections", *Public Choice*, Vol. 65, pp. 201-216.

BALKE, NATHAN S. (1991), "Partisanship Theory, Macroeconomic Outcomes, and Endogenous Elections", *Southern Economic Journal*, Vol. 57, nº 4, Abril, pp. 920-935.

BLANCHARD, OLIVIER JEAN, & CHARLES M. KAHN (1980), "The Solution of Linear Difference Models under Rational Expectations", *Econometrica*, Vol. 48, nº. 5, Julho, pp. 1305-1311.

BIERBRAUER, FELIX, & LYDIA MECHTENBERG (2008), "Winners and Losers of Early Elections: On the Welfare Implications of Political Blockades and Early Elections", *SFB 649 Discussion Paper* 2008-071, Dezembro (disponível em http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2008-071.pdf).

CALEIRO, ANTÓNIO (2000), "Para uma Cronologia das Eleições - O que está para além de uma análise de sobrevivência?", *in* "Homenagem ao Professor Augusto da Silva", Departamento de Sociologia, Universidade de Évora, Dezembro, pp. 31-43.

CALEIRO, ANTÓNIO (2008), "The Political Economics Side of the J-Curve", *Documento de Trabalho nº 2008/22*, Departamento de Economia, Universidade de Évora (a aparecer em *The Icfai University Journal of Applied Economics*). (disponível em http://www.decon.uevora.pt/get_file.php3?id=1179)

CALEIRO, ANTÓNIO (2009), "How upside down are political business cycles when there is output persistence", *Research in Economics*, Vol. 63, nº 1, pp. 22-26.

CARGILL, THOMAS F., & MICHAEL M. HUTCHISON (1991), "Political Business Cycles with Endogenous Election Timing: Evidence from Japan", *The Review of Economics and Statistics*, Vol. LXXXIII, nº 4, Novembro, pp. 733-739.

CHAPPELl, D., & D. A. PEEL (1979), "On the Political Theory of the Business Cycle", *Economics Letters*, Vol. 2, pp. 327-332.

CHAPPELL Jr., HENRY W., & WILLIAM R. KEECH (1986), "Party Differences in Macroeconomic Policies and Outcomes", *The American Economic Review (Papers and Proceedings)*, Vol. 76, nº. 2, Maio, pp. 71-74.

ELLIS, CHRISTOPHER J. (1991), "Endogenous Voting in a Partisan Model with Rational Voters", *Journal of Macroeconomics*, Vol. 13, nº 2, Primavera, pp. 267-278.

ELLIS, CHRISTOPHER J., & MARK A. THOMA (1991), "Partisan Effects in Economies with Variable Electoral Terms", *Journal of Money, Credit, and Banking*, Vol. 23, nº 4, Novembro, pp. 728-741.

GÄRTNER, MANFRED (1986), "Some Political Economy of Flexible Exchange Rates", *European Journal of Political Economy*, Vol. 2, nº2, pp. 153-168.

GÄRTNER, MANFRED (2000), "Political Macroeconomics: A Survey of Recent Developments", *Journal of Economic Surveys*, Vol. 14, nº 5, Dezembro, pp. 527-561.

HECKELMAN, JAC C. (2002), "Variable Rational Partisan Business Cycles: theory and some evidence", *Canadian Journal of Economics*, Vol. 35, nº 3, Agosto, pp. 568-585.

ITO, TAKATOSHI, & JIN HYUK PARK (1988), "Political Business Cycles in the Parliamentary System", *Economics Letters*, Vol. 27, pp. 233-238.

ITO, TAKATOSHI (1989), "Endogenous Election Timings and Political Business Cycles in Japan", *National Bureau of Economic Research*, Working Paper No. 3128, Setembro.

ITO, T. (1990), "The Timing of Elections and the Political Business Cycles in Japan", *Journal of Asian Economics*, Vol. 1, nº 1, pp. 135-146.

KAYSER, MARK A. (2005), "Who surfs, who manipulates? The determinants of opportunistic election timing and electorally motivated economic intervention", *American Political Science Review*, Vol. 99, pp. 17-28.

KEECH, WILLIAM R., & CARL P. SIMON (1985), "Electoral and Welfare Consequences of Political Manipulation of the Economy", *Journal of Economic Behavior and Organization*, Vol. 6, Junho, pp. 177-202.

KEPPO, JUSSI, LONES SMITH, & DMITRY DAVYDOV (2008), "Optimal Electoral Timing: Exercise Wisely and You May Live Longer", *Review of Economic Studies*, Vol. 75, nº 2, pp. 597-628.

LÄCHLER, U. (1982), "On political business cycles with endogenous election dates", *Journal of Public Economics*, Vol. 17, pp. 111-117.

LUCAS Jr., ROBERT E. (1973), "Some International Evidence on Output-Inflation Tradeoffs", *The American Economic Review*, Vol. 63, nº 3, Junho, 326-334.

LUPIA, ARTHUR, & KAARE STRØM (1995), "Coalition termination and the strategic timing of parliamentary elections", *American Political Science Review*, Vol. 89, n. 3, pp. 648-665.

MILESI-FERRETTI, GIAN MARIA, & ENRICO SPOLAORE (1994), "How Cynical can an Incumbent be? Strategic Policy in a Model of Government Spending", *Journal of Public Economics*, Vol. 55, pp. 121-140.

NORDHAUS, WILLIAM D. (1975), "The Political Business Cycle", *The Review of Economic Studies*, Vol. 42(2), nº 130, Abril, pp. 169-190.

PLOEg, F. Van Der (1989), "The Political Economy of Overvaluation", *The Economic Journal*, Vol. 99, Setembro, pp. 850-855.

SMITH, ALASTAIR (1996), "Endogenous election timing in majoritarian parliamentary systems", *Economics and Politics*, Vol. 8, pp. 85-110.

SMITH, ALASTAIR (2004), "Election Timing", Cambridge University Press, Cambridge.

EFFICIENCY OF RESTRTUCTURED, MERGED AND ACQUIRED BANKS: THE CASE OF TURKEY

İMRE ERSOY
E-mail: iersoy@marmara.edu.tr
Marmara University
European Union Institute

Abstract

The restructuring of the Turkish Banking System started in 2001 following devastating financial crises. The mergers & acquisitions of the illiquid banks during restructuring are proceeded by voluntary M&As particularly after 2005 as the start of negotiations with the EU boosted foreign direct investment inflows to Turkey. The aim of this paper is to search the effects of restructuring and M&As on the efficiency of banks in Turkey. To estimate stochastic frontiers and to predict firm-level efficiencies relative to the estimated frontier, a Translog Stochastic Production Function is employed to the the balanced panel data of banks for the period of 2002-2007. The efficiency estimates suggest a steady increase of efficiency in Turkish banks, the marking of quite high efficiency scores by state-owned banks, cream- skimming effect in some voluntary M&As, the mattering of origin of foreign banks and the robust and statistically significant relation between the size and efficiency of banks.

Key-Words: Bank Efficiency; Restructuring and M&A; Stochastic Frontier Efficiency
Thematic Area: Economic and Financial Modelling.

1. INTRODUCTION

The efficiency concept is developed by Farell (1957) with his decomposition of economic efficiency of a firm into technical and allocative efficiencies. Frontier efficiency measurement techniques are developed in late 1970s and since then have been applied in many areas and sectors. Both parametric and nonparametric frontier efficiency measurement investigations for financial services sector in general and banking in particular have increased considerably during the last couple of decades.[1] These studies followed the developments in the financial markets in terms of liberalization, privatization, crises, restructuring and mergers and acquisions (M&A's).

The cycles of the Turkish Banking Sector (TBS) is not any different from most of the other emerging economies in terms of liberalization and privatization, followed by banking crises and eventually proceeded first by restructuring and then by M&As. The TBS started to be liberalised and privatised by mid-1980s. Almost 10 years later, the unsustainably high public sector borrowing requirement (PSBR) and current account (C/A) deficit due to unresolved structural problems caused a decade long macroeconomic instability in the country. During 1990s the overweight of the state-owned banks, the exchange risk exposures; the liquidity and interest rate risk due to asset-liability mismatches; the insufficiency of capital, the small and scattered banking sector infrastructure; the weak risk perception and management, the insufficient regulatory and supervisory infrastructure coupled with moral hazard and adverse selection problems led to continuous weakening of the TBS. The first banking crises in 1994 could be

[1] Berger et al (1997) reviewed 130 studies that apply parametric and non-parametric frontier efficiency measurement techniques to financial institutions in 21 different countries.

quickly overcome by the liquidation of three small privatea banks and the introduction of a full deposit insurance scheme. [2]

A devastating financial crisis, however, hit the open foreign exchange positions and highly leveraged securities portfolios of banks in Turkey in 2001.[3] Banking Regulatory and Supervisory Agency (BRSA, 2009) reports that four small non-depository investment and development banks are liquidated and 22 commercial banks are overtaken by Saving Deposits Insurance Fund (SDIF) .[4] The restructuring and the M&A's of the Turkish banks started just after the calming down of the crises in 2001.The period of restructuring and M&As in Turkey is a period of structural, legal and institutional reforms not only for the banking industry but also for the Turkish economy.[5]

The aim of this paper is to scrutinize the effects of restructuring and M&A's on the efficiency of banks in Turkey by applying stochastic frontier efficiency measurement technique for the period of 2002-2007.[6] The rest of the paper unfolds as follows: Section 2 portrays the developments in the TBS in terms of restructuring and M&A's. Section 3 describes the Stochastic Frontier Efficiency Measurement methodology applied and the dataset employed for the estimation of efficiency of Turkish banks. Section 4 discusses the findings of the empirical research. Section 5 concludes.

2. THE RESTRUCTURING AND THE MERGERS & ACQUISITIONS IN THE TURKISH BANKING SYSTEM

The restructuring of the TBS started with the enactment of the new Banking Law, decree 4389 ratified in June 1999, which authorized the newly founded BRSA to start implementing the "Banking Sector Restructuring Program". The program consisted of restructuring of the state-owned banks, the resolution of the banks overtaken by SDIF, the rehabilitation of the private banks, the enhancement of the regulatory and supervisory framework and the improvement of efficiency of the TBS.[7] (BRSA, 2006).

To restructure the state-owned banks, capital is injected, their duty losses are refined, Ziraat Bank merged with state-owned Emlakbank and the management of Halkbank and Ziraatbank are transferred to one newly appointed Executive Board and in two years time a radical reduction of 32, 5% in the number of branches and 50% in the number of employees is realized. (BRSA, 2009) The problem concerning the 22 illiquid commercial banks that are nationalized by SDIF is resolved by M&A's or by liquidations.[8]

The rehabilitation of the domestic commercial banks on the other hand, included –this time seriously-limiting the interest and exchange rate risks of banks, the gradual recapitalization of

[2] During 1990s the intermediation function of TBS deteriorated considerably with loans to total assets ratio declining to 32, 8% in 2000 from 47% in 1990. (BRSA 2009)

[3] In November 2000 Turkish banks lived a liquidity crisis but in February 2001 the speculative attack to foreign currency forced a shift to "Floating Exchange Rate" regime and the deepening of the crisis caused a systemic banking crises in Turkey.

[4] As of today 2 banks are still in the process of liquidation.

[5] The enactement of Central Bank independence, the new Banking Law, the increase in foreign direct investments to the banking sector following the start of accession negotiations with the EU and the price stability focused monetary policy implications are only a few but important developments of the period. These IMF supported reforms and the international conjucture were the main reasons behind the decrease in inflation rate from an average 75, 9% and the increase in growth rate from an average 2, 9 % between 1991-2001 to the levels of 13, 9% and 6,8 % respectively between 2002-2007. (BRSA, 2009)

[6] The effect of the recent developments between 2002 and 2007 in terms of restructuring and M&As in the TBS is not, at least to our knowledge, examined in depth by utilizing Stochastic Frontier Production function.

[7] The total cost of financial and operational restructuring exceeded 1/3 of Gross Domestic Production (GDP) of the country.

[8] The assets of these 22 banks accounted for almost 15% of the asset size of the TBS and 3 of these commercial banks are liquidated, 12 are merged and 5 are rehabilitated and sold.

private banks with their own resources, the restructuring of bad debts with the "Istanbul Approach", the establishment of internal control and risk management systems and the encouragement of M&A's of banks by offering tax incentives.

The encouragement of M&A's facilitated the mergers of banks belonging to the same holding company in search for economies of scale.[9] Table 1 depicts the voluntary mergers of banks that realized during the 2001-2005 period.[10]

Table 1: The Mergers in Turkish Banking Sector

Bank M&A	Bank M&A	The New Bank	Year
Morgan Guaranty Trust Co.	The Chase Manhattan B.(branch)	MGCB	2001
Osmanlı Bankası A.Ş	Körfezbank	Osmanlı Bank	2001
Osmanlı Bankası A.Ş	Garanti Bank	Garanti Bank	2001
Sınai Yatırım Bankası A.Ş.	T.S.K.B	T.S.K.B.	2002
Finansbank A.Ş	Fiba Bank A.Ş	Finansbank	2003
Credit Agricole Indosuez T.A.Ş	CLCAIndosuez	CA Indosuez	2004
Akbank T.A.Ş	Ak Uluslararası Bank A.Ş	Akbank	2005

Source: BRSA, 2006

Table 2 illustrates the M&A's particularly in the form of transfer of shares of Turkish commercial banks to foreign ownership. Foreign direct investment inflows toTurkey increased considerably after the start of accession negotiations with the EU in 2005.

Table 2: Transfer of Shares in Turkish Banking Sector

Name of the Bank	FDI from	Date	Country of Origin	Transfer of Shares
Koçbank	Unicredito	2002	Italy	Transfer of 49,5% of shares
T.Ekonomi Bank	BNP Paribas	2004	France	Transfer of 42,1% of shares
T.Dış Ticaret Bank	Fortis Bank NV	2005	Lux-Belgium	Transfer of 89,3% of shares
Yapı ve Kredi Bank	Koç-Unicredito	2005	Turkey-Italy	Transfer of 57,4% of shares
Garanti Bank	General Electric	2005	USA-France	Transfer of 25,5% of shares
Finansbank	N.Bank of Greece	2006	Greece	Transfer of 46% of shares
C Bank	Tarshish Hippolim H.	2006	Israel	Transfer of 57,6% of shares
Arap-Türk Bank	Libian Foreign Bank	2006	Libian-Kuwait	Transfer of 58,6% of shares
Denizbank	Dexia Participation B	2006	Belgium-France	Transfer of 75% of shares
Tat Yatırım Bank	Merrill Lynch E.A.H	2006	USA	Transfer of 99.95% of shares
Akbank	Citibank Overseas	2006	USA	Transfer of 20% of shares
Şekerbank	Bank Turanalem JSC	2006	Kazakhstan	Transfer of 33.98% of shares
MNG Bank	Arap Bank-BankMed	2006	Jordan-Lebanon	Transfer of 91% of shares
Tekfenbank	EFG Eurobank	2007	Greece	Transfer of 70% of shares

Source: (BRSA, 2006)

The restructuring and M&A's decreased the share of state -owned banks in total assets to 26%, and increased the foreign ownership in TBS to 43% from 5%. (BRSA, 2008)

Table 3 summarizes the major developments in the TBS before and after the restructuring.[11]

[9] Between 2001 and 2005 Osmanlı Bank, Körfezbank and Garanti Bank that belonged to Doğus Group, Finansbank and Fiba Bank belonging to Fiba Holding and Akbank and Ak International Bank under the ownership of Sabancı group merged.

[10] One of two two foreign bank mergers was between branches and the other between subsidiaries. T.S.K.B is a rather big development bank compared to Sınai Yatırım Bank which is also a development bank.

[11] It is worth noting that the number of banks decreased to 49 during restructuring from 79 in 2000. The loans to assets ratio of banks which gives an indication of the proper functioning of the intermediation role of banks increased to 50% in 2008 from 23% in 2001.

Table: 3 Fundamental Indicators of Turkish Banking Sector

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Total Assets (in bio TL)	36,8	72,1	104,1	173,4	212,7	249,7	306,4	406,9	499,7	581,6	732,5
Loans (in bio TL)	13,1	19,9	31,8	38,0	49,0	66,2	99,3	156,4	219,0	285,6	367,4
Deposits (in bio TL)	21,8	43,4	58,9	110,4	138,0	155,3	191,1	251,5	307,6	356,9	454,6
Capital (in bio TL)	33,3	4,2	7,2	18,3	25,7	35,5	46,0	54,7	59,5	75,8	86,4
Number of banks	75	81	79	63	54	50	48	51	50	50	49
Number of branches	7740	8104	8298	7386	6160	6029	6440	6240	6911	7700	8768
Capital Adequacy (%)	13	8,24	9,29	20,78	25,34	30,93	28,24	23,73	21,91	18,94	17,98
NPL/ Loans (%)	7,2	10,5	11,1	25,2	17,6	11,5	6,0	4,8	3,8	3,5	3,7
Loans / Deposits (%)	60,3	45,9	54,1	34,4	35,5	42,6	52,0	62,2	71,2	80,0	80,8
Profits (in bio TL)	0,8	-0,3	-3,1	-10,5	2,9	5,6	6,5	6,0	11,4	14,9	13,4
ROA (%)	2,1	-0,4	-3,0	-6,1	1,4	2,5	2,3	1,7	2,5	2,8	2,0
ROE (%)	23,1	-7,2	-43,7	-57,5	11,2	18,1	15,8	12,1	21,0	24,8	18,3
Concentration of 10 (%)	67,6	67,5	69,2	79,5	80,8	82,3	84,0	82,9	83,5	82,5	82,8
Foreign Ownership (%)	5,6	7,1	3,4	3,0	3,3	3,0	3,5	6,3	13,1	14,0	17,0
Funds in F. C. (bio$)	2,5	3,2	3,8	2,8	11	17	22	36	49	61	62

Source: BRSA, 2009

3. THE MEASUREMENT OF BANK EFFICIENCY

To measure the efficiency of production of a Decision Making Unit (DMU) frontiers are estimated. The frontier analysis consists of the estimation of the best –practice frontier of the optimal input mix for maximized outputs and to predict the firm- level efficiencies through the comparison of the efficiency of the DMU with that frontier.[12]

3.1. The methodology

To evaluate the impact of restructuring and M&A's in Turkish Banking Sector, the frontier of the most efficient banks are estimated and the efficiency levels of banks are measured by their distance from the efficiency frontier.[13] To this end the Stochastic Frontier Efficiency function is employed as it takes the possible measurement errors and other noise into account while measuring inefficiencies. Following Battesa and Coelli (1995) and Demir et al. (2005) the translog Stochastic Frontier Production frontier with the Technical Inefficiency Effects (TIE) of banks is employed. The generalized truncated normal distribution assumption is chosen according to the likelihood ratio tests. The mean technical efficiency of the banks and the bank specific determinants of technical inefficiencies of size and loans to assets ratio are estimated using a single equation of the Frontier model proposed by Batese and Coelli (1995).

[12] For a comprehensive understanding of efficiency measurement and the techniques implemented, the reader is advised to refer to Coelli and Battese (1998)

[13] The measurement of efficiency by the use of frontier analysis necessitates a number of decisions. Non-parametric frontier efficiency methods such as Data Envelopment Analysis (DEA) or parametric frontier efficiency techniques such as Stochastic Frontier (SF) Efficiency functions may be employed. Technical Efficiency (TE) may be output oriented or input oriented. Production / Profitability or Cost functions may be employed.

The translog stochastic production frontier (SPF) that is used may be expressed as follows:

$$\ln(Q_{i,t}) = \beta_0 + \sum_j \beta_j \ln(x_{j,i,t}) + \sum_j \sum_k \beta_{j,k} \ln(x_{j,k,i,t}) + v_{i,t} - u_{i,t} \quad (1)$$

where $Q_{i,t}$ is the output of bank i in period t defined as total loans and total securities , $x_{j,k,i,t}$ is a vector of input quantities representing deposits and other borrowings (D), labour (L) and capital (C). v_{it} are random variables which are assumed to be iid $N(0, \sigma v^2)$ and independent of the u_{it} . u_{it} are non-negative random variables which are assumed to account for technical inefficiency in production and are assumed to be independently distributed as truncations at zero of the $N(\mu_{it}, \sigma u^2)$ distribution which may be expressed as:

$$\mu_i = \delta_0 + \sum_m \delta_m Z_{m,i} \quad (2)$$

These Z variables of the Technical InefficienciesEffect (TIE) Model include total assets of the bank as an indication for the size of the bank and the ratio of loans to total assets.[14] Maximum Likelihood Estimation Method is used to estimate the parameters of the model of the 2 equations simultaneously.[15]

3.2. The dataset

The source of database for the audited balance sheets is Turkish Bankers' Association. Balanced panel data for 30 banks is used.[16] The time period employed is between 2002 and 2007. Intermediary approach is employed where deposits and other borrowings as well as capital and labor are considered as inputs and loans and securities are considered as output.

Table 4: Descriptive Statistics of Input and Output and z variables in the SPF Model

	LO	LL	LC	LD	TA	L/A
Mean	14.43398	6.955861	12.98547	14.32860	11550284	34.59964
Median	14.44215	7.201143	12.86687	14.54200	2573015	33.82932
Maximum	18.11562	10.05750	16.17673	18.03880	80942311	73.24822
Minimum	9.267477	2.639057	9.228992	8.035282	24661.00	0.286104
Std. Dev.	2.190757	2.065974	1.821205	2.210627	18547297	19.18383

Source: Own calculations

3.3. The empirical findings

The Maximum Likelihood Estimates of the Technical Efficiency Effects (TE) Frontier of the SPF are depicted in table 5.

[14] The measures of Production Efficiency relative to the Production Frontier is:
EFF = $(Y_i^* | \mathbf{U}_i, \mathbf{X}_i)$ **/ E** $(Y_i^* | \mathbf{U}_i = \mathbf{0}, \mathbf{X}_i)$**,** where Y_i^* is the production of the i-th firm which will be equal to $\exp(Y_i)$ when the dependant variable is in logs, with 0>EFF>1.

[15] Frontier 4,1 is used for the maximum likelihood estimations.

[16] Out of the 33 commercial banks, 30 are used as the rest had 0 values as input and/or output

Table 5: Mle of the TE Effects Frontier of the SPF Model

	Coefficient	Standard Error	t-statistics
β_0	0.48776790E+01	0.79381676E+00	0.61445907E+01
β_L	-0.59167739E-08	0.52209351E-09	-0.11332786E+02
β_C	-0.35881418E+00	0.11660314E+00	-0.30772256E+01
β_D	-0.28268810E-09	0.13587815E-09	-0.20804530E+01
β_{LxL}	0.37724025E+00	0.41837909E-01	0.90167091E+01
β_{CxC}	0.22931592E-08	0.58131163E-09	0.39448018E+01
β_{DxD}	0.39650963E+00	0.37587695E-01	0.10548921E+02
β_{LxC}	0.28831688E-08	0.50401215E-09	0.57204351E+01
β_{CxD}	0.35806478E-01	0.94190467E-02	0.38014970E+01
β_{LxD}	0.17000046E-08	0.51482967E-09	0.33020719E+01
δ_0	0.78772426E+01	0.22460854E+01	0.35070984E+01
$\delta_{T.A}$	-0.64401915E-01	0.22247272E-01	-0.28948229E+01
$\delta_{L/A}$	0.84532261E-08	0.67744241E-08	0.12478147E+01
σ^2	0.11920752E+01	0.41927143E+00	0.28432063E+01
γ	0.97726201E+00	0.11531690E-01	0.84745776E+02

Source: Own calculations

The estimated parameters are all statistically significant. To test if Cobb-Douglas production functions would be an adequate representation of the data, Ho: $\beta_{ij} = 0$ for all i≤j=1, 2, 3 is tested and rejected.[17] To test the null hypothesis that there are no Technical Inefficiency Effects in the Model, ***Ho***: $\gamma = 0$ is tested and rejected. [18] The estimated $\delta_{T.A}$ parameter of the TIE model is significantly different from 0. The coefficient of $\delta_{T.A}$ is negative which indicate that the larger the values of Total Assets or in other words size, the smaller the values of inefficiency effects will tend to be. The Maximum Likelihood estimates of $\delta_{T.A}$ and $\delta_{L/A}$ is larger relative to their standard errors. Ho: $\delta_{T.A}$ and $\delta_{L/A} = 0$ are tested to see if their true value is 0 and Ho is rejected based on LL estimates which proves that the parameters of technical inefficiency should not be deleted from the model.[19] White's general heteroscedasticity test is performed to find evidence of heteroscedastic error structure and the results proved negative.

[17] LL_{C-D} = -0.15160453E+03 and $LL_{TRANSLOG}$ = -0.82097376E+02

λ = -2((-0.15160453E+03)-(-0.82097376E+02)) = 129

129 > upper 5% point for the χ^2_6 distribution which is 12.592. (χ^2_6 is used as 6 regressors are restricted.)

[18] LR test of the one sided error is estimated as 0.76059024E+02, whereas the **critical value** $\chi^2_1(2\alpha)$ at 5% significance = 9,48773 for 4 degrees of freedom.

[19] The new LL functions are all smaller than the initial LL function. Otherwise, the concerned variables would be deleted from the model.

γ = 0.97726201 (0.01) indicates that the vast majority of residual variation is due to inefficiency effect u_i, and that the random error v_i is approximately zero. It shows that the stochastic frontier may not be significantly different from the deterministic frontier. [20]

The mean efficiency of banks in Turkey is estimated to be 66, 2% for 2002-2007 and 78, 9 % for 2007. The technical efficiency scores of Turkish Banks suggest that efficiency increased steadily from 2002 to 2007. The high technical inefficiencies in 2002 can be explained with the crises of 2001 and the increasing efficiency scores with the restructuring in Turkish Banking Sector. Table 6 reveals the efficiency scores of commercial banks in Turkey.

Table 6: The Efficiency Scores of Banks in Turkey.

Banks	2002	2003	2004	2005	2006	2007
State-owned Bank 1	0.7800	0.8452	0.8831	0.89470	0.8952	0.9414
State-owned Bank 2	0.8745	0.8810	0.8142	0.94143	0.9125	0.9097
State-owned Bank 3	0.6300	0.9065	0.9414	0.91795	0.9254	0.9402
Private Commercial Bank 1	0.8948	0.9244	0.9252	0.94283	0.9025	0.8646
Private Commercial Bank 2	0.7804	0.7425	0.7214	0.90605	0.8417	0.8760
Private Commercial Bank 3	0.7349	0.4696	0.8473	0.73111	0.8114	0.8071
Private Commercial Bank 4	0.7408	0.7598	0.6133	0.79530	0.9392	0.8720
Private Commercial Bank 5	0.6678	0.5863	0.5818	0.82341	0.7390	0.7644
Private Commercial Bank 6	0.1102	0.2163	0.1986	0.19347	0.6775	0.5911
Private Commercial Bank 7	0.3512	0.5327	0.7195	0.86812	0.8704	0.9441
Private Commercial Bank 8	0.9070	0.9094	0.9519	0.92497	0.9260	0.8958
Private Commercial Bank 9	0.6484	0.9018	0.8984	0.86441	0.8380	0.8815
Private Commercial Bank 10	0.6770	0.6559	0.6776	0.90494	0.9173	0.8978
Transition Bank of SDIF	0.6878	0.9460	0.8724	0.74366	0.9208	0.9048
Foreign Bank Branch 1	0.1521	0.1926	0.4628	0.68219	0.6970	0.6841
Foreign Bank Branch 2	0.1912	0.3596	0.6845	0.74119	0.6204	0.8407
Foreign Bank Branch 3	0.8003	0.1973	0.9092	0.48187	0.5663	0.5993
Foreign Bank Branch 4	0.2757	0.1553	0.4428	0.56762	0.8849	0.9173
Foreign Bank Branch 5	0.5432	0.4689	0.1693	0.57760	0.6709	0.8188
Foreign Bank Subsidiary 1	0.3334	0.4153	0.4646	0.46871	0.9060	0.4175
Foreign Bank Subsidiary 2	0.2923	0.5450	0.7176	0.65057	0.7623	0.8256
Foreign Bank Subsidiary 3	0.6063	0.5802	0.7971	0.74758	0.8862	0.8318
Foreign Bank Subsidiary 4	0.8965	0.9162	0.2772	0.53444	0.6628	0.3303
Foreign Bank Subsidiary 5	0.8643	0.8145	0.8960	0.89132	0.7024	0.8963
Foreign Bank Subsidiary 6	0.5932	0.8239	0.8872	0.77751	0.8661	0.3996
Foreign Bank Subsidiary 7	0.5381	0.5680	0.8006	0.58241	0.9018	0.8428
Foreign Bank Subsidiary 8	0.1520	0.1272	0.4596	0.85756	0.7547	0.4284
Foreign BankSubsidiary 9	0.4655	0.7748	0.8810	0.87684	0.9294	0.8862
Foreign Bank Subsidiary 10	0.3120	0.3356	0.3395	0.38914	0.7880	0.8609
Foreign Bank Subsidiary 11	0.2972	0.6236	0.4367	0.49631	0.7588	0.6443

Source: Own calculations

[20] The OLS Estimates of the Model

	Coefficient	Standard error	t-ratio
Beta 0	0.12997851E+01	0.97571987E+00	0.13321294E+01
Beta 1	-0.44862250E-08	0.69190838E-09	-0.64838425E+01
Beta 2	0.13512880E+00	0.14806737E+00	0.91261704E+00
Beta 3	0.34412462E-09	0.18708088E-09	0.18394430E+01
Beta 4	0.47691486E+00	0.55567684E-01	0.85825939E+01
Beta 5	0.59728254E-09	0.32647418E-09	0.18294940E+01
Beta 6	0.40356841E+00	0.50611573E-01	0.79738365E+01
Beta 7	0.23580360E-08	0.48324954E-09	0.48795412E+01
Beta 8	0.45358886E-02	0.12911539E-01	0.35130504E+00
Beta 9	-0.25487233E-09	0.45739625E-09	-0.55722437E+00

4. DISCUSSION OF THE EMPIRICAL RESULTS

The empirical finding about the increasing technical efficiency of TBS is consistent with Aysan and Ceyhan (2008) who claim that the restructuring after the 2001 crises robustly account for the improvement in the efficiency of banks in Turkey. [21] The estimated mean efficiency of 66,2% and 78,9% is consistent with Yunter and Caner (2004) who finds a mean efficiency rate of 59% for 2000, 58% for 2001-2002 .The mean geometric efficiency of Turkish commercial banks is estimated as 57% featuring high variability for the 1992-1997 period. (El-Gamal and Inanoğlu, 2002) The mean technical efficiency under constant returns to scale for 48 Turkish banks in 1998 is estimated as 67 %. (Jackson and Fethi, 2000) [22]

The State-owned banks are marking quite high efficiency scores, over the mean efficiency. This high efficiency of state-owned banks may be explained by the immediate radical reforms during restructuring and by their size.This result is consistent with the findings of some other recent empirical findings. Aysan and Ceyhan (2007) used DEA and Malmquisit TFP Change index and employed intermediation approach in their study of Turkish banks for the period between 1990 and 2006 and they claim that state-owned banks in Turkey became the most efficienct banks after 2001, replacing the foreign banks. Matousek et al. (2008) analyzed the efficiency of TBS in 2000-2005 by applying Stochastic Cost Frontier Efficiency method to a panel data of banks and their results show a decrease in inefficiency and transformation of the state banks to more efficient banks with the restructuring. The results of some other studies even for the period before the restructuring in Turkey show that state owned banks are more efficient than their private counterparts and that foreign bank are inefficient. (El-Gamal and Inanoğlu, 2002)

All privately owned domestic commercial banks have high average efficiency scores especially after 2004, excluding one which is the only small private commercial bank. The result of our empirical analysis of a robust and positive impact of size on bank efficiency in Turkey is in line with some other studies. Jackson and Fethi (2000) applied DEA and Tobit regression analysis for 48 banks in 1998 and their findings suggest a robust and positive relation between size and efficiency for Turkish banks. Lopez et al. (2002) investigated the efficiency of 450 cooperative banks in Italy covering the period 1995-1999 and their DEA findings suggest that larger banks show higher efficiency.

Out of the seven foreign bank subsidiaries which used to be privately owned Turkish banks, three proves the cream screaming affect in bank mergers & acquisitions as their efficiency was already high before the change of ownership. This finding is concistent with Poghosyan and Borovick (2006) who investigated the cream-skimming effect using a sample of 275 banks from 19 transition economies by implementing first a panel probit model and then a stochastic efficiency frontier function. Their findings put forward the fact that cream skimming effect is significant in transition aconomies.

[21] The non-parametric and a only a few parametric frontier functions employed to search for the efficiency of banks focused mainly on the effects of liberalization of TBS. Mercan and Yolalan (2000), Mahmud and Zaim (1998) to site only a few, empirically show that liberalization increased the level of efficiency in TBS. Jackson, et al. (1998) researched the 1992-1996 period and found productivity growth in the post-liberalization era of Turkey, with the exception of the crises hit year 1994. Fethi et al. (2001) used DEA and DEA based Malmquist indices to search for the efficiency and productivity change and their results show that the financial reforms have a positive effect on both technical and allocative efficiencies.

[22] The technical efficiency estimates for other countries reveal parallel results. Lopez, J. and Salvo, R. (2002) investigated the efficiency of 450 cooperative banks in Italy covering the period 1995-1999 and their DEA findings show that the cost efficiency is below 70% with a high variability between 40 and 100%. Kamberoglou et al. (2004) studied cost efficiency of Greek banks for the period of 1993-1999 by applying distribution free approach to the annual balanced panel data of 20 banks and found an average of technical efficiency of 56% with relative efficiency estimates ranging between 14% and 80%.

Out of the remaining four (now) foreign bank subsidiaries two bank had rather important changes in efficiency after the M&A. The efficiency of one improved and the other deteriorated radically.This result is consistent with the findings of Rezitis (2007) who investigated the Total Factor Productivity of Greek banks by utilizing Malmquist productivity index to find the impact of acquisitions on efficiency. The result of his study suggests a negative relationship as the Technical Efficiency of merged banks is found to decrease and that of non-merged increase. Amel et al. (2003) investigated the effects of consolidation in the financial sector for a period of twenty years and could neither find evidence of increase in managerial efficiency nor of consolidations yielding economies of scope.

CONCLUSIONS

The study empirically investigated the efficiency of restructured, merged and acquired banks in Turkish Banking Systen for the period of 2002-2007. The major findings of the study may be summarized as follows: The efficiency in Turkish Banking System increased during the 2002-2007 period. The mean efficiency is 66,2% for 2002-2007 and 78,92% for 2007. The efficiency increase is mainly due to restructuring rather than M&A's. The size has a robust, positive and statistically significant relation with efficiency in Turkish Banking System. State-owned Turkish banks rank first in efficiency.The private commercial banks have high efficiency scores except one bank which is the only small private commercial bank. The cream-skimming effect is more obvious than efficiency increase in the merged or acquired banks. Banks in Turkey has still a large potential for improvement in terms of efficiency.

REFERENCES:

AMEL, D., BARNES, C., PANETTA, F. AND SALLEO, C. (2003), "Consolidation and Efficiency in the financial sector: a review of the international evidence, *Journal of Banking and Finance*, Volume: 28,Issue. 10, October 2004

AYSAN, A. AND CEYHAN, P., (2008), "What determines the banking sector performance in globalized financial markets? The case of Turkey", *PhysicaA* 387 (2008) 1593–1602

AYSAN, A. AND CEYHAN, P., (2007), "Market Disciplining Role of Crisis on the Restructuring of the Turkish Banking Sector", *Munich Personal RePEc Archive Pape*r *(MPRA)* No.5492 http://mpra.ub.uni-muenchen.de/5492/

BANKING REGULATORY AND SUPERVISORY AGENCY, (2001), "Bankacılık Sektörü Yeniden Yapılandırma Programı",["Banking Sector Resrtructuring Program"] www.BRSA.org.tr

BANKING REGULATORY AND SUPERVISORY AGENCY, (2003), "Bankacılık Sektörü Yeniden Yapılandırma Programı Gelişme Raporu-(VII)", October 2003 ["Banking Sector Resrtructuring Program [Progress Report-(VII)] www.BRSA.org.tr

BANKING REGULATORY AND SUPERVISORY AGENCY, (2006) Bankacılıkta Yapısal Gelişmeler [Structural Changes in Banking] Issue No: 1, December 2006, ISSN 1307-5691

BANKING REGULATORY AND SUPERVISORY AGENCY, (2007), Bankacılıkta Yapısal Gelişmeler [Structural Changes in Banking], Issue No: 2 December 2007, ISBN: 1307-5691

BANKING REGULATORY AND SUPERVISORY AGENCY, (2008), Finansal Piyasalar Raporu, [Financial Markets Report] No: 11, September 2008, ISBN: 1307-0347

BANKING REGULATORY AND SUPERVISORY AGENCY, (2009), Krizden İstikrara Türkiye Tecrübesi,[From Crisis to Stability:The Experience of Turkey] www.BRSA.org.tr

BERGER, ALLEN .BROCKETT, P., COOPER, W AND PASTOR,J.(1997),"New approaches for Analyzing and Evaluating the Performance of Financial Institutions" *European Journal of Operational Research* 19, No2-3

BONIN, J., HASAN, I. AND WACHTEL, P. (2004) "Bank Performance, Efficiency and Ownership in Transition Countries" *Journal of Banking and Finance* 29 (2005)

COELLI, T. (1996), "A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation", *CEPA working paper 96/07*, University of New England

COELLI, T. AND BATTESE, G.(1998), *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers, Massachusetts,USA

CRISTOPOULOS, D. AND TSIONAS, E., (2001), "Banking Economic Efficiency in the Deregulation Period: Results from Heterosccedastic Stochastic Frontier Models, *The Manchester School* Vol.69, No. 6, December 2001

DEMIR. N, MAHMUD, S. AND BABUSCU, S. (2005), "The Technical Inefficiency Effects of Turkish Banks After Financial Liberalization", *The developing Ecoomies*, XLIII-3 (September 2005):396-441

EL-GAMAL, M. AND INANOGLU, H. (2002), **"**Efficiencies and Unobservable Heterogeneity in Turkish Banking: 1990-2000", seminar presentation in September 2002

FARRELL,M. J. (1957), "The measurement of productive efficiency", *Journal of Royal Statistical Society* Series A (General) 120, 2, p.253-281.

FETHI, M., JACKSON, P.AND WEYMAN-JONES, T. (2001), "An Empirical Study of Stochastic DEA and Financial Performance: the Case of the Turkish Commercial Banking Industry", INFORMS International Hawaii Conference, Maui, Hawaii, USA, June 17-20, 2001.

ISIK, H. AND HASSAN K., (2002a), "Cost and Profit Efficiency of the Turkish Banking Industry: An Empirical Investigation", *The Financial Review* 37 (2002), p.257-280

ISIK H, HASSAN K. (2002b) "Technical, scale and allocative efficiencies of Turkish Banking Industry", *Journal of Banking and Finance* 26 (2002), p.719-766

IŞIK, I, (2007), "Bank ownership and productivity developments: evidence from Turkey" *Studies in Economics and Finance,* Bradford: 2007. Volume 24, Issue: 2

JACKSON, P. AND FETHI, M.(2000), "Evaluating the efficiency of Turkish commercial banks: An Application of DEA and Tobit Analysis", International DEA Symposium, Univ3ersity of Queensland, Brisbane, Australia, 2-4 July 2000.

JACKSON, P. FETHI, M. AND İNAL,G.(1998), "Efficiency and Productivity Growth in Turkish Commercial Banking Sector: A non-parametric approach", The European Symposium on:Data EnvelopmentAnalysis-Recent Development and Applications held at the Hochschule Harz,Wernigerode, Germany, 16-18 October, 1998.

KAMBEROGLOU, N., LIAPIS, E., SIMIGANNIS, G.AND TZAMOURANI, P. (2004) "Cost Efficiency in Greek Banking", *Bank of Greece Working Papers* No: 9, 2004

KAYA, Y. AND DOĞAN, E. (2005), "Dezenflasyon Sürecinde Türk Bankacılık Sektöründe Etkinliğin Gelişimi",

["The Developments in the Efficiency of Turkish Banking Sector during Desinflation Period"]Banking Regulatory and Supervisory Agency, 2005

KRAFT, E.AND TIRTIROGLU, D., (1998), "Bank Efficiency in Croatia: A Stochastic-Frontier Analysis", *Journal of Comparative Economics,* Vol: 26, Issue: 2, pp. 282-300, http://ideas.repec.org

LOPEZ, J. AND SALVO, R. (2002) "Italian Mutual Banks: Performance, Efficiency and Mergers and Acquisitions", Societe Universitaire Europeene de Recherches Financieres, Vienna 2002.

MAHMUD, S.F. AND O. ZAIM (1998), "Cost structure of Turkish private banking industry: A GL restricted cost function approach", *METU Studies in Development* 25/3 (1998), 435-446.

MATOUSEK, R., DASCI, S. AND SERGI,B.(2008) "The Efficiency of the Turkish Banking system in 2000-2005", *International Journal of Economic Policy in Emerging Economies*, Vol:1, No.4, 2008

MERCAN, M. AND YOLALAN.R.(2000), "Turk bankacilik sisteminde olcek ve mulkiyet yapilari ile finansal performansin iliskisi", ["The Relation between Ownership Structure and Economies of Scale and Financial Performance in Turkish Banking System"],*YKB Tartisma Tebligleri[YKB Working Papers]* Serisi 1, 2000.

POGHOSYAN, T. AND BOROVICKA, J. (2006), "Banking Efficiency and Foreign Ownership in Transition: Is there an Evidence of a Cream-Skimming Effect?" CERGE-EI, Prague: The Center for Economic Research and Graduate Education, Charles University and the Economics Instituteof the Academy of Sciences of the Czech Republic.

REZITIS, A. (2007), "Efficiency and productivity effects of bank mergers: Evidence from the Greek banking industry", *Economic Modelling* XX (2007)

YUNTER, P. AND CANER, S. (2004), Efficiency of Turkish Commercial Banks in years 1999-2002, Bilkent University Worrking Paper

CORRECCIÓN POR SESGO EN LA ESTIMACIÓN MCO DE LA MAGNITUD DE OUTLIERS EN VOLATILIDAD EN MODELOS GARCH

JULIO A. AFONSO-RODRÍGUEZ
E-mail: jafonsor@ull.es
Facultad de Ciencias Económicas y Empresariales
Universidad de La Laguna

Resumen

En este trabajo se propone una generalización de las propuestas de Franses y Ghijsels (1999) y Charles y Darné (2005) para la estimación de la magnitud y localización de outliers de tipo aditivo e innovacional, respectivamente, en la ecuación de volatilidad tipo GARCH(1,1), introduciendo ouliers en volatilidad de cambio de nivel y cambio temporal. A partir de esta generalización del procedimiento de identificación y estimación de dichas observaciones anómalas empleando una ecuación estimable por MCO o MCG, se generaliza el análisis de Battaglia (2006) en el caso de modelos ARIMA, sobre la corrección por sesgo de estas estimaciones. En el caso de un modelo GARCH(1,1) se identifican los sesgos en la estimación de la magnitud de cada uno de los cuatro tipos de outliers considerados en los siguientes casos: (*a*) bajo el supuesto de parámetros GARCH conocidos cuando el outlier en volatilidad presente es realmente del tipo considerado, (*b*) con parámetros conocidos, cuando el outlier en volatilidad difiere del tipo considerado, y (*c*) evaluación del sesgo en el caso de parámetros GARCH estimados.

Palabras Clave: Modelos GARCH, outliers en volatilidad, corrección por sesgo.
Área Temática: Modelación Económica y Financiera.

Abstract

In this paper we propose a generalization of the proposals by Franses and Ghijsels (1999) and Charles and Darné (2005) to estimate the magnitude and location of additive and innovational outliers in a GARCH(1,1)-type volatility equation, by introducing level shifts and temporary changes. Using this generalized version of the procedure for identification and estimation of abnormal observations through an estimable equation by OLS or GLS, we apply the analysis by Battaglia (2006) in ARIMA models to derive the expressions for the bias induced by several sources and the way to correct the resulting estimates for this bias. In the case of a GARCH(1,1)-type model, we get the expression for the bias in the estimation of the magnitude for the four types of outliers considered in the following cases: (*a*) under the assumption of known GARCH parameters and the true outlier contamination, (*b*) with known GARCH parameters and wrong outlier mechanism, and (*c*) bias due to the estimation of GARCH parameters.

Key words: GARCH models, volatility outliers, bias correction.
Thematic Area: Economic and Financial Modelling.

1. INTRODUCCIÓN

En este trabajo analizamos parcialmente las propiedades de la estimación de la magnitud de un outlier que perturba la estructura de dinámica de la volatilidad de un proceso GARCH(1,1). Recientemente, ha surgido un gran interés por determinar las propiedades de los estimadores de los parámetros de un modelo GARCH basados en máxima verosimilitud (MV), especialmente utilizando como distribución condicional distribuciones de colas pesadas para las innovaciones, con el objetivo de explicar el aparentemente gran número de outliers u observaciones anómalas presentes en las series estudiadas. De entre la literatura existente, cabe destacar la aportación de Berkes y Horváth (2003) que establecen la consistencia de los estimadores QMV bajo

normalidad condicional en el caso de varianza finita de las innovaciones y si, además, existe al menos el momento de cuarto orden la normalidad asintótica. Hall y Yao (2003) estudian estas propiedades en el contexto de distribuciones de colas pesadas y recientemente, Mikosch y Straumann (2006) establecen resultados similares en el contexto de distribuciones asintóticas estables cuando las innovaciones tienen varianza infinita. Si la verosimilitud condicional se asume no gaussiana, que es una práctica común en muchas aplicaciones recientes, poco se sabe sobre la consistencia de las estimaciones QMV y no se han establecido resultados generales que permitan su verificación en la práctica. Así, en presencia de series temporales caracterizadas posiblemente por un gran número de outliers, como alternativa a la estimación MV, distintos autores han propuesto la estimación robusta de modelos GARCH que, en general, presentan ciertas dificultades computacionales (ver, por ejemplo, Muler, N., V.J. Yohai (2002), Muler, N., V.J. Yohai (2008), Jiang, J., Q. Zhao, Y.V. Hui (2001), Peng, L., Q. Yao (2003) y Carnero, M.A., D. Peña, E. Ruiz (2007) para algunas de las propuestas más recientes). Una alternativa a esta práctica puede consistir en tratar de corregir las series observadas por la posible influencia de outliers en volatilidad previamente a la estimación definitiva de los parámetros GARCH mediante QMV gaussiana. Los trabajos de Hotta y Tsay (1998) y Doornik y Ooms (2005) desarrollan procedimientos secuenciales basados en RV y ML, respectivamente, para la detección de outliers en volatilidad de tipo aditivo, mientras que Franses y Ghijsels (1999) adaptan la metodología de Chang, Tiao y Chen (1988) y Chen y Liu (1993) para la detección y corrección de outliers en modelos ARMA al caso GARCH(1,1) basada en una regresión mínimo cuadrática. Charles y Darné (2005) extienden el procedimiento al caso de outliers de tipo innovacional y en base al estudio de simulación en el caso AR(1)-GARCH(1,1) de Verhoeven y McAleer (2000) proponen utilizar una regla empírica simple para identificar un outlier en volatilidad, identificando el tipo según la estructura de la regresión considerada. En este trabajo estudiamos distintos aspectos de esta metodología. En primer lugar, planteamos que la extensión al caso de outliers de tipo cambio de nivel (o cambio permante) y de cambio transitorio (o temporal) no plantea ninguna dificultad adicional. En segundo lugar, motivado por el reciente trabajo de Battaglia (2006), estudiamos el posible sesgo del estimador MCO de la magnitud del outlier en volatilidad para cada uno de los cuatro modelos de contaminación considerados, base del procedimiento secuencial de detección y corrección de outliers en volatilidad propuesto por Franses y Ghijsels (1999). Encontramos que por la particular estructura de un proceso GARCH contaminado por un outlier en volatilidad, estas estimaciones son sesgadas incluso en el caso de suponer que los parámetros GARCH sean conocidos. En esta situación obtenemos las expresiones del sesgo en muestras finitas y del sesgo asintótico, tanto en el caso de suponer que el outlier presente es del tipo particular considerado como en el caso de suponer que el outlier en volatilidad presente es de tipo diferente al tipo considerado al implementar el procedimiento. Los resultados indican la posible inconsistencia de este mecanismo de detección de outliers en volatilidad y se señala posibles remedios como la corrección por el sesgo de naturaleza multiplicativo en la estimación de la magnitud del outlier. En tercer lugar tratamos de obtener estos mismos resultados en el caso general de parámetros GARCH estimados por QMV bajo normalidad condicional, puesto que para implementar el procedimiento del estimación de la magnitud del outlier se requiere disponer de valores para dichos parámetros para construir la estructura del regresor previa a la estimación MCO del parámetro de magnitud del proceso de contaminación. Encontramos que los resultados son, en general, comparables al caso de parámetros conocidos. Es decir, bajo ciertas restricciones muy débiles, la estimación QMV bajo normalidad condicional de un modelo GARCH para observaciones contaminadas por un outlier en volatilidad no aporta un sesgo adicional al procedimiento posterior de detección, si bien dichas estimaciones son no robustas, como ha probado, entre otros Sakata y White (1998) y de Melo Mendes (2000) que obtiene la curva de sensibilidad y función de influencia del estimador QMV. En este trabajo se aporta también un resultado en este sentido, el sesgo aproximado del estimador QMV de los parámetros de un modelo GARCH(1,1) bajo contaminación por un outlier en volatilidad de alguno de los cuatro tipos considerados.

2. MODELOS GARCH Y OUTLIERS EN VOLATILIDAD

En este epígrafe vamos a introducir la formulación general de un modelo GARCH y sus principales propiedades que serán utilizadas en el resto del trabajo.

2.1 Modelos GARCH

El modelo GARCH(p,q) fuerte de Bollerslev (1986), viene dado por

$$\varepsilon_t = \zeta_t \sqrt{h_t(\boldsymbol{\theta})} \tag{2.1}$$

donde ζ_t es el proceso de innovación, que se asume *i.i.d.* con media nula y varianza unitaria, con distribución independiente de ε_{t-i} $i \geq 1$ y, por tanto, independiente de la varianza condicional $h_t(\boldsymbol{\theta})$, dada por

$$h_t(\boldsymbol{\theta}) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}(\boldsymbol{\theta}) \tag{2.2}$$

donde $\boldsymbol{\theta} = (_{(0}, _{(1},\ldots,_{(p}, \beta_1,\ldots,\beta_q)$ es el vector de parámetros de la ecuación de volatilidad. Estableciendo las restricciones habituales de positividad de los coeficientes α_i, β_j (i=0,1,…,p, j=1,..,q) y bajo el supuesto de estacionariedad estricta, Berkes *et.al.* (2003) obtienen una representación recursiva única de la varianza condicional en términos del cuadrado de las observaciones de la forma

$$h_t(\boldsymbol{\theta}) = c_0(\boldsymbol{\theta}) + \sum_{i=1}^{\infty} c_i(\boldsymbol{\theta}) \varepsilon_{t-i}^2,$$

donde la estructura de estos coeficientes $c_i(\boldsymbol{\theta})$ en función de $\boldsymbol{\theta}$ pueden encontrarse en la referencia citada. Por otro lado, si se emplea la siguiente descomposición para el cuadrado de la serie ε_t,

$$\varepsilon_t^2 = h_t(\boldsymbol{\theta}) + (\zeta_t^2 - 1) h_t(\boldsymbol{\theta}) = h_t(\boldsymbol{\theta}) + v_t(\boldsymbol{\theta}) \tag{2.3}$$

donde $v_t(\boldsymbol{\theta})$ puede interpretarse como en el error en la aproximación del cuadrado de los valores observados del proceso por la varianza condicional GARCH, se obtiene entonces la representación ARMA(m,q) equivalente como

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^{m} \delta_i(\boldsymbol{\theta}) \varepsilon_{t-i}^2 + v_t(\boldsymbol{\theta}) - \sum_{j=1}^{q} \beta_j v_{t-j}(\boldsymbol{\theta}) \tag{2.4}$$

con $\delta_i(\boldsymbol{\theta}) = \alpha_i + \beta_i \, 1 \leq i \leq n$, $\delta_i(\boldsymbol{\theta}) = \alpha_i, n = q < i \leq p, = \beta_i, n = p < i \leq q$ con $n = min(p,q)$, y $m = máx(p,q)$. Empleando la notación del operador de retardos se tiene que $(1 - \sum_{i=1,m} \delta_i(\boldsymbol{\theta}) L^i) \varepsilon_t^2 = \alpha_0 + (1 - \sum_{j=1,q} \beta_j L^j) v_t(\boldsymbol{\theta})$, de forma que bajo la condición de estacionariedad débil, $0 < \sum_{i=1,p} \alpha_i + \sum_{j=1,q} \beta_j < 1$ o bien $0 < \sum_{i=1,m} \delta_i(\boldsymbol{\theta}) < 1$, entonces se tiene la siguiente representación AR(∞) para el proceso de error $v_t(\boldsymbol{\theta})$,

$$v_t(\boldsymbol{\theta}) = -\frac{\alpha_0}{1 - \sum_{j=1,q} \beta_j} + \frac{1 - \sum_{i=1,m} \delta_i(\boldsymbol{\theta}) L^i}{1 - \sum_{j=1,q} \beta_j L^j} \varepsilon_t^2 = \alpha_0(\boldsymbol{\theta}) + \sum_{i=0}^{\infty} \pi_i(\boldsymbol{\theta}) \varepsilon_{t-i}^2 \tag{2.5.1}$$

donde $\pi_0(\boldsymbol{\theta}) = 1$, o bien, de forma más compacta,

$$v_t(\boldsymbol{\theta}) = \alpha_0(\boldsymbol{\theta}) + \pi(L) \varepsilon_t^2 \tag{2.5.2}$$

con $\pi(L) = 1 - \pi_1(\boldsymbol{\theta}) L - \pi_2(\boldsymbol{\theta}) L^2 - \ldots$ Alternativamente a (2.4) se puede obtener también una representación MA(∞) de la forma

$$\varepsilon_t^2 = \frac{\alpha_0}{1-\sum_{i=1}^{m}\delta_i(\boldsymbol{\theta})} + \frac{1-\sum_{j=1}^{q}\beta_j L^j}{1-\sum_{i=1}^{m}\delta_i(\boldsymbol{\theta})L^i} v_t(\boldsymbol{\theta}) = \alpha_0^*(\boldsymbol{\theta}) + \psi(L)v_t(\boldsymbol{\theta})$$

Por definición, a partir de (2.3) se tiene que el término de error $v_t(\boldsymbol{\theta})$ tiene media nula, varianza $E[v_t^2(\boldsymbol{\theta})] = (E[\zeta_t^4]-1)\cdot E[h_t^2(\boldsymbol{\theta})] < \infty$ si $E[\zeta_t^4] < \infty$ y está serialmente incorrelacionado. Dado el conjunto de información pasado disponible en t, $I_{t-1} = \{\varepsilon_{t-i}, i \geq 1\}$, la distribución condicional se caracteriza por media condicional nula, $E_{t-1}[v_t(\boldsymbol{\theta})] = 0$ y es condicionalmente heterocedástico,

$E_{t-1}[v_t^2(\theta)] = (E[\zeta_t^4]-1)\cdot h_t^2(\theta)$

con covarianzas condicionales nulas.

2.2 Modelos GARCH contaminado por un Outlier en Volatilidad

A partir del trabajo de Hotta y Tsay (1998), la formulación de un modelo GARCH con un outlier en volatilidad viene dado por

$$\varepsilon_t^2(\omega) = \zeta_t^2 h_t(\boldsymbol{\theta}, \omega) + \tilde{\eta}_t(\omega, k) \tag{2.6}$$

donde $\varepsilon_t^2(\omega)$ son las observaciones al cuadrado del proceso GARCH, $h_t(\boldsymbol{\theta}, \omega)$ es la ecuación de varianza condicional para el proceso contaminado, dada por

$$h_t(\boldsymbol{\theta}, \omega) = \alpha_0 + \sum_{i=1}^{p}\alpha_i \varepsilon_{t-i}^2(\omega) + \sum_{j=1}^{q}\beta_j h_{t-j}(\boldsymbol{\theta}, \omega) \tag{2.7}$$

y $\tilde{\eta}_t(\omega, k)$ representa el proceso determinista de contaminación por un outlier en volatilidad, donde ω es la magnitud del outlier y k la localización.

Hotta y Tsay (1998) tratan la estimación quasi-máximo verosimil (QMV) de la magnitud de un outlier en volatilidad de tipo aditivo bajo el supuesto de normalidad condicional y derivan así un procedimiento secuencial de contrastación de la existencia de este tipo de outlier basado en el estadístico de multiplicadores de Lagrange. Posteriormente Franses y Ghijsels (1999) proponen emplear el método de detección de outlier de Chen y Liu (1993) para detectar y corregir por la preencia de este mismo tipo de outliers aditivos en modelos GARCH, que resulta ser mucho más fácil de implementar y que está basado en la estimación por MCO de una regresión lineal simple. Recientemente, Charles y Darné (2005) han propuesto extender esta metodología al caso de outliers en volatilidad de tipo innovacional. El aspecto clave de esta metodología es la analogía entre los modelos GARCH y los modelos ARMA a través de la representación ARMA para el cuadrado de la serie observada dada por (2.4). En lo que sigue planteamos la extensión de este procedimiento al caso de outliers en volatilidad de cualquiera de los tipos habitualmente considerados en los modelos ARMA de series temporales. Para ello, el proceso de contaminación en (2.6) viene dado por

$$\tilde{\eta}_t(\omega, k) = \omega\xi(L)I_t^{(k)} \tag{2.8}$$

donde ω es la magnitud del outlier y $\xi(L) = 1 + \xi_1 L + \xi_2 L + ...$ determina la dinámica del efecto del outlier e $I_t^{(k)}$ es la función impulso dada por $I_t^{(k)} = 1$, $t = k$, $I_t^{(k)} = 0$ $t \neq k$. Los cuatro modelos habituales de outliers a considerar son los siguientes:

$$AVO: \ \xi(L) = 1$$
$$IVO: \ \xi(L) = \pi^{-1}(L)$$
$$LSVO: \xi(L) = 1/(1-L)$$
$$TCVO: \xi(L) = 1/(1-\delta L) \quad (0 < \delta < 1) \tag{2.9}$$

donde AVO identifica un outlier aditivo en volatilidad (*additive volatility outlier*), IVO un outlier de tipo innovacional (*innovational volatility outlier*), con $\pi(L)$ dado por (2.5), LSVO un outlier con efecto de cambio de nivel permanente en volatilidad (*level shift volatility outlier*) y TCVO un outlier con efecto de cambio transitorio en volatilidad (*temporary change volatility*

outlier).

En el caso AVO, se tiene simplemente que

$$\tilde{\eta}_t(\omega,k) = \omega \cdot I_t^{(k)} \tag{2.10}$$

mientras que en el caso IVO resulta que

$$\tilde{\eta}_t(\omega,k) = \omega \cdot \pi^{-1}(L) I_t^{(k)} = \omega \cdot \psi(L) I_t^{(k)} = \omega \cdot \sum_{i=0}^{\infty} \psi_i(\boldsymbol{\theta}) I_{t-i}^{(k)}$$

con $\psi_0(\boldsymbol{\theta}) = 1$, de forma que $\tilde{\eta}_t(\omega,k) = \omega(I_t^{(k)} + \psi_{t-k}(\boldsymbol{\theta}) S_t^{(k+1)})$, donde $S_t^{(k+1)}$ es la función escalón $S_t^{(k+1)} = 1$, $t>k$, $S_t^{(k+1)} = 0$, $t \leq k$. Los coeficientes $\psi_j(\boldsymbol{\theta})$ resultan de la representación MA(∞) del cuadrado del proceso GARCH(1,1) para $\varepsilon_t(\omega)$ y se obtienen a partir de la ecuación $1 - \sum_{j=1}^{q} \beta_j L^j = (1 - \sum_{i=1}^{m} \delta_i(\boldsymbol{\theta}) L^i)(1 + \sum_{j=1}^{\infty} \psi_j(\boldsymbol{\theta}) L^j)$.

En el caso de un modelo GARCH(1,1) se tiene que los coeficientes $\psi_i(\boldsymbol{\theta})$ vienen dados por $\psi_0(\boldsymbol{\theta}) = 1$ y $\psi_i(\boldsymbol{\theta}) = \alpha_1(\alpha_1 + \beta_1)^{i-1}$ i = 1, 2, ..., de forma que

$$\tilde{\eta}_t(\omega,k) = \omega(I_t^{(k)} + \alpha_1(\alpha_1 + \beta_1)^{t-k-1} S_t^{(k+1)}) \tag{2.11}$$

En el caso de un cambio estructural permanente en volatilidad (LSVO), resulta que $\tilde{\eta}_t(\omega,k) = \omega \cdot \sum_{i=0}^{\infty} I_{t-i}^{(k)} = 0 \; t = 1,...,k-1$, y $\tilde{\eta}_t(\omega,k) = \omega, t = k, k+1, ..., T$, es decir

$$\tilde{\eta}_t(\omega,k) = \omega \cdot S_t^{(k)} \tag{2.12}$$

Finalmente, en el caso de un cambio temporal en volatilidad (TCVO), se tiene que $\tilde{\eta}_t(\omega,k) = \omega \cdot \sum_{i=0}^{\infty} \delta^i I_{t-i}^{(k)} = 0, t = 1,...,k-1$, y $\tilde{\eta}_t(\omega,k) = \omega\delta^{t-k}, t = k, k+1, ..., T$, de forma que

$$\tilde{\eta}_t(\omega,k) = \omega\delta^{t-k} S_t^{(k)}. \tag{2.13}$$

Teniendo en cuenta (2.6), en el caso GARCH(1,1) se tiene que la ecuación de varianza condicional

$$h_t(\boldsymbol{\theta},\omega) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2(\omega) + \beta_1 h_{t-1}(\boldsymbol{\theta},\omega) \tag{2.14}$$

puede expresarse en términos del proceso de contaminación como

$$h_t(\boldsymbol{\theta},\omega) = \alpha_0 \left(1 + \sum_{i=2}^{t} \prod_{j=1}^{i-1} A_{t-j}(\boldsymbol{\theta}) \right) + h_0(\boldsymbol{\theta},\omega) \prod_{j=1}^{t} A_{t-j}(\boldsymbol{\theta})$$
$$+ \alpha_1 \left(\tilde{\eta}_{t-1}(\omega,k) + \sum_{i=1}^{t-1} \tilde{\eta}_{t-1-i}(\omega,k) \prod_{j=1}^{i} A_{t-j}(\boldsymbol{\theta}) \right)$$

mediante sustituciones recursivas, y donde $A_{t-j}(\boldsymbol{\theta}) = \alpha_1 \zeta_{t-j}^2(\boldsymbol{\theta}) + \beta_1$ es, por construcción en el caso de un proceso GARCH fuerte, una secuencia de variables *i.i.d.* con valor esperado $E[A_{t-j}(\boldsymbol{\theta})] = \alpha_1 + \beta_1$. Asumiendo que $h_0(\boldsymbol{\theta},\omega) = h_0(\boldsymbol{\theta})$ no depende del proceso de contaminación por outlier, entonces se tiene que se puede escribir la ecuación GARCH(1,1) como

$$h_t(\boldsymbol{\theta},\omega) = h_t(\boldsymbol{\theta}) + F_t(\boldsymbol{\theta},\omega) \tag{2.15}$$

donde $h_t(\boldsymbol{\theta})$ es la ecuación de varianza condicional especificada para el proceso libre de contaminación y la función $F_t(\boldsymbol{\theta},\omega)$ recoge la influencia y dinámica del proceso de outlier en volatilidad,

$$F_t(\boldsymbol{\theta},\omega) = \alpha_1 \left(\tilde{\eta}_{t-1}(\omega,k) + \sum_{i=1}^{t-1} \tilde{\eta}_{t-1-i}(\omega,k) \prod_{j=1}^{i} A_{t-j}(\boldsymbol{\theta}) \right) \tag{2.16}$$

con valor esperado,

$$E[F_t(\boldsymbol{\theta},\omega)] = \alpha_1 \left(\tilde{\eta}_{t-1}(\omega,k) + \sum_{i=1}^{t-1} (\alpha_1+\beta_1)^i \tilde{\eta}_{t-1-i}(\omega,k) \right) \tag{2.17}$$

Para cualquiera de los modelo de outlier en volatilidad de los considerados con localización en el instante k, esta función de influencia toma valor no nulo para $t = k+1, \ldots, T$, perturbando entonces la dinámica de la volatilidad de distinta forma según la naturaleza del outlier. Así, en el caso AVO, se tiene que

$$F_t(\boldsymbol{\theta},\omega) = \alpha_1 \omega \prod_{j=1}^{t-k-1} A_j(\boldsymbol{\theta}) \quad t = k+1,\ldots,T \quad \left(\prod_{j=1}^{0} A_j(\boldsymbol{\theta}) = 1 \right)$$

de forma que la influencia del outlier, medida en términos esperados, viene dada por $E[F_t(\boldsymbol{\theta},\omega)] = \omega\alpha_1(\alpha_1+\beta_1)^{t-k-1} S_t^{(k+1)}$.

3. ESTIMACIÓN MCO DE LA MAGNITUD DE OUTLIERS EN VOLATILIDAD EN UN MODELO GARCH(1,1): CASO DE PARÁMETROS CONOCIDOS

En esta sección planteamos la formulación de la regresión auxiliar para la estimación por MCO (mínimos cuadrados ordinarios) del parámetro de magnitud de un outlier en volatilidad en el caso de contaminación de un modelo GARCH(1,1) y derivamos posteriorme las expresiones del sesgo para cada uno de los cuatro tipos de outliers considerados en el caso de parámetros GARCH conocidos.

3.1 Sesgo en la estimación MCO de la magnitud de un Outlier en Volatilidad

A partir de la representación dada en (2.6) para el proceso contaminado, se tiene que el término de error en la aproximación del cuadrado del proceso observado por la ecuación de varianza condicional tipo GARCH,

$$\tilde{v}_t(\boldsymbol{\theta},\omega) = \varepsilon_t^2(\omega) - h_t(\boldsymbol{\theta},\omega) \tag{3.1}$$

viene dado por

$$\tilde{v}_t(\boldsymbol{\theta},\omega) = (\zeta_t^2 - 1)h_t(\boldsymbol{\theta},\omega) + \tilde{\eta}_t(\omega,k) = v_t(\boldsymbol{\theta},\omega) + \tilde{\eta}_t(\omega,k) \tag{3.2}$$

donde $v_t(\boldsymbol{\theta},\omega)$ tiene las mismas propiedades que $v_t(\boldsymbol{\theta})$ en (2.3) en el caso de no contaminación. En el caso GARCH(1,1), empleando (3.1) la representación ARMA(1,1) equivalente viene dada por

$$\varepsilon_t^2(\omega) = \alpha_0 + (\alpha_1+\beta_1)\varepsilon_{t-1}^2(\omega) + \tilde{v}_t(\boldsymbol{\theta},\omega) - \beta_1 \tilde{v}_{t-1}(\boldsymbol{\theta},\omega)$$

que puede resolverse para el término de error de la misma forma que en (2.5),

$$\tilde{v}_t(\boldsymbol{\theta},\omega) = \alpha_0(\boldsymbol{\theta}) + \pi(L)\varepsilon_t^2(\omega)$$

y utilizando ahora (2.6) y (2.15) puede escribirse como

$$\tilde{v}_t(\boldsymbol{\theta},\omega) = \pi(L)\tilde{\eta}_t(\omega,k) + \tilde{v}_t(\boldsymbol{\theta}) \tag{3.3}$$

donde

$$\tilde{v}_t(\boldsymbol{\theta}) = v_t(\boldsymbol{\theta}) + \pi(L)\zeta_t^2 F_t(\boldsymbol{\theta},\omega) \tag{3.4}$$

empleando la descomposición general de la varianza condicional GARCH bajo contaminación dada en (2.15). La ecuación (3.3) puede escribirse alternativamente en forma de regresión teniendo en cuenta la representación (2.8) para el proceso general de contaminación como

$$\tilde{v}_t(\boldsymbol{\theta},\omega) = \omega \cdot x_t(\boldsymbol{\theta},k) + \tilde{v}_t(\boldsymbol{\theta}) \tag{3.5}$$

donde

$$x_t(\boldsymbol{\theta},k) = \pi(L)\xi(L)I_t^{(k)} \tag{3.6}$$

Teniendo en cuenta las distintas formulaciones de modelos de outliers en (2.9), se tiene que $x_t(\boldsymbol{\theta},k)$ viene dado en cada caso por

AVO: $$x_t(\boldsymbol{\theta},k) = I_t^{(k)} + \pi_{t-k}(\boldsymbol{\theta}) \cdot S_t^{(k+1)}, \quad \pi_{t-k}(\boldsymbol{\theta}) = -\alpha_1\beta_1^{t-k-1} \tag{3.7-1}$$

IVO: $$x_t(\boldsymbol{\theta},k)=I_t^{(k)} \tag{3.7-2}$$

LSVO: $$x_t(\boldsymbol{\theta},k)=I_t^{(k)}+\gamma_{t-k}(\boldsymbol{\theta})\cdot S_t^{(k+1)},\quad \gamma_{t-k}(\boldsymbol{\theta})=1-\sum_{j=1}^{t-k}\pi_j(\boldsymbol{\theta})$$

$$x_t(\boldsymbol{\theta},k)=I_t^{(k)}+\left\{1+\alpha_1\left(\frac{1-\beta_1^{t-k}}{1-\beta_1}\right)\right\}\cdot S_t^{(k+1)} \tag{3.7-3}$$

TCVO: $x_t(\boldsymbol{\theta},k)=I_t^{(k)}+\gamma_{t-k}(\boldsymbol{\theta},\delta)\cdot S_t^{(k+1)},\quad \gamma_{t-k}(\boldsymbol{\theta},\delta)=\delta^{t-k}-\sum_{j=1}^{t-k}\delta^{t-k-j}\pi_j(\boldsymbol{\theta})$

$$x_t(\boldsymbol{\theta},k)=I_t^{(k)}+\left\{\delta^{t-k}+\alpha_1\left(\frac{\delta^{t-k}-\beta_1^{t-k}}{\delta-\beta_1}\right)\right\}\cdot S_t^{(k+1)} \tag{3.7-4}$$

Así, la ecuación (3.5) puede emplearse como una regresión auxiliar base para implementar un procedimiento de estimación de la magnitud del outlier y de contrastación de su significación. Con parámetros GARCH conocidos, el estimador MCO de la magnitud del outlier en volatilidad vendría dado por

$$\hat{\omega}_T(\boldsymbol{\theta},k)=\frac{\sum_{t=1}^{T}x_t(\boldsymbol{\theta},k)\tilde{v}_t(\boldsymbol{\theta},\omega)}{\sum_{t=1}^{T}x_t^2(\boldsymbol{\theta},k)} \tag{3.8}$$

Teniendo en cuenta que $E[v_t(\boldsymbol{\theta},\omega)]=E[(\zeta_t^2-1)h_t(\boldsymbol{\theta},\omega)]=0$, y que a partir de (2.8) y (3.2), $E[\tilde{v}_t(\boldsymbol{\theta},\omega)]=\omega\xi(L)I_t^{(k)}$, la expresión general del sesgo del estimador de ω es

$$E[\hat{\omega}_T(\boldsymbol{\theta},k)-\omega]=-\omega\frac{\sum_{t=1}^{T}x_t(\boldsymbol{\theta},k)[x_t(\boldsymbol{\theta},k)-\xi(L)I_t^{(k)}]}{\sum_{t=1}^{T}x_t^2(\boldsymbol{\theta},k)}$$

Así, salvo en el caso del modelo IVO, donde el estimador toma la forma simple

$$\hat{\omega}_T(\boldsymbol{\theta},k)=\tilde{v}_k(\boldsymbol{\theta},\omega) \tag{3.9-1}$$

en el resto de casos, el estimador tiene la siguiente estructura

$$\hat{\omega}_T(\boldsymbol{\theta},k)=\frac{\tilde{v}_k(\boldsymbol{\theta},\omega)-\sum_{t=k+1}^{T}\phi_{t-k}(\boldsymbol{\theta})\cdot\tilde{v}_t(\boldsymbol{\theta},\omega)}{1+\sum_{t=k+1}^{T}\phi_{t-k}^2(\boldsymbol{\theta})}=\frac{\tilde{v}_k(\boldsymbol{\theta},\omega)-\sum_{j=1}^{T-k}\phi_j(\boldsymbol{\theta})\cdot\tilde{v}_{k+j}(\boldsymbol{\theta},\omega)}{1+\sum_{j=1}^{T-k}\phi_j^2(\boldsymbol{\theta})} \tag{3.9-2}$$

donde $-\phi_j(\boldsymbol{\theta})$ tiene la forma particular dada para el regresor $x_t(\boldsymbol{\theta},k)$ en (3.7-1), (3.7-3) y (3.7-4). Por tanto, salvo en el caso IVO, donde el estimador de ω es insesgado,

$$E[\hat{\omega}_T(\boldsymbol{\theta},k)]=E[\tilde{v}_k(\boldsymbol{\theta},\omega)]=\omega \tag{3.10-1}$$

en el resto de casos, AVO, LSVO y TCVO, se tiene que

$$E[\hat{\omega}_T(\boldsymbol{\theta},k)]=\frac{1}{1+\sum_{j=1}^{T-k}\phi_j^2(\boldsymbol{\theta})}\left(\tilde{\eta}_k(\boldsymbol{\theta},\omega)-\sum_{j=1}^{T-k}\phi_j(\boldsymbol{\theta})\cdot\tilde{\eta}_{k+j}(\boldsymbol{\theta},\omega)\right) \tag{3.10-2}$$

Empleando las expresiones en (2.10),(2.12) y (2.13) para $\tilde{\eta}_t(\boldsymbol{\theta},\omega)$ y de (3.7-1), (3.7-3) y (3.7-4) para los coeficientes $\phi_j(\boldsymbol{\theta})$, a continuación presentamos las expresiones del sesgo en muestras finitas y asintótico para la estimación de la magnitud de un outlier en volatilidad de cada uno de los cuatro tipos analizados.

3.1.1 CASO AVO

$$E[\hat{\omega}_T(\boldsymbol{\theta},k)-\omega]=-\omega\alpha_1^2\left(1+\alpha_1^2\frac{1-\beta_1^{2(T-k)}}{1-\beta_1^2}\right)^{-1}\left(\frac{1-\beta_1^{2(T-k)}}{1-\beta_1^2}\right)<0 \tag{3.11-1}$$

La magnitud del sesgo depende básicamente de la magnitud del parámetro AR, α_1, y de la localización del outlier, *k*, a través de $\beta_1^{2(T-k)}$. Se observa también que el sesgo no desaparece asintóticamente y viene dado por:

$$\lim_{T\to\infty} E[\hat{\omega}_T(\boldsymbol{\theta},k)-\omega] = -\omega\frac{\alpha_1^2/(1-\beta_1^2)}{1+\alpha_1^2/(1-\beta_1^2)} \tag{3.11-2)}$$

3.1.2 CASO LSVO

$$E[\hat{\omega}_T(\boldsymbol{\theta},k)-\omega] = -\omega\frac{\sum_{j=1}^{T-k}(\gamma_j^2(\boldsymbol{\theta})-\gamma_j(\boldsymbol{\theta}))}{1+\sum_{j=1}^{T-k}\gamma_j^2(\boldsymbol{\theta})} = -\omega\alpha_1\frac{\sum_{j=1}^{T-k}\gamma_j(\boldsymbol{\theta})\left(\frac{1-\beta_1^j}{1-\beta_1}\right)}{1+\sum_{j=1}^{T-k}\gamma_j^2(\boldsymbol{\theta})} \tag{3.12-1)}$$

donde

$$\begin{aligned}\sum_{j=1}^{T-k}\gamma_j(\boldsymbol{\theta})\left(\frac{1-\beta_1^j}{1-\beta_1}\right) &= \frac{\alpha_1(T-k)}{1-\beta_1}\left(1+\frac{\alpha_1}{1-\beta_1}\right)+\frac{\alpha_1^2\beta_1^2}{(1-\beta_1)^2(1-\beta_1^2)}(1-\beta_1^{2(T-k)})\\ &-\frac{\alpha_1\beta_1}{(1-\beta_1)^2}\left(1+\frac{2\alpha_1}{1-\beta_1}\right)(1-\beta_1^{T-k})\end{aligned} \tag{3.12-2)}$$

y

$$\sum_{j=1}^{T-k}\gamma_j^2(\boldsymbol{\theta}) = (T-k)\left(1+\frac{\alpha_1}{1-\beta_1}\right)^2+\frac{\alpha_1^2\beta_1^2}{(1-\beta_1)^2(1-\beta_1^2)}(1-\beta_1^{2(T-k)})-\frac{2\alpha_1\beta_1}{(1-\beta_1)^2}\left(1+\frac{\alpha_1}{1-\beta_1}\right)(1-\beta_1^{T-k}) \tag{3.12-3)}$$

Aplicando L'Hopital, se tiene que el sesgo asintótico viene dado por

$$\lim_{T\to\infty} E[\hat{\omega}_T(\boldsymbol{\theta},k)-\omega] = -\omega\alpha_1\frac{\frac{\alpha_1}{1-\beta_1}\left(1+\frac{\alpha_1}{1-\beta_1}\right)+\frac{\alpha_1^2\beta_1^2}{(1-\beta_1)^2(1-\beta_1^2)}-\frac{\alpha_1\beta_1}{(1-\beta_1)^2}\left(1+\frac{2\alpha_1}{1-\beta_1}\right)}{1+\left(1+\frac{\alpha_1}{1-\beta_1}\right)^2+\frac{\alpha_1^2\beta_1^2}{(1-\beta_1)^2(1-\beta_1^2)}-\frac{2\alpha_1\beta_1}{(1-\beta_1)^2}\left(1+\frac{\alpha_1}{1-\beta_1}\right)} \tag{3.12-4)}$$

3.1.3 CASO TCVO

$$E[\hat{\omega}_T(\boldsymbol{\theta},k)-\omega] = -\omega\frac{\sum_{j=1}^{T-k}(\gamma_j^2(\boldsymbol{\theta},\delta)-\delta^j\gamma_j(\boldsymbol{\theta},\delta))}{1+\sum_{j=1}^{T-k}\gamma_j^2(\boldsymbol{\theta},\delta)} \tag{3.13-1)}$$

donde

$$\begin{aligned}\sum_{j=1}^{T-k}\gamma_j^2(\boldsymbol{\theta},\delta) &= \frac{1-\delta^{2(T-k)}}{1-\delta^2}\left(\delta+\frac{\alpha_1}{1-(\beta_1/\delta)}\right)^2\\ &-\frac{\alpha_1\beta_1}{1-(\beta_1/\delta)}\left(\frac{\alpha_1\beta_1}{1-(\beta_1/\delta)}\frac{1-\beta_1^{2(T-k)}}{\delta^2(1-\beta_1^2)}+2\frac{1-(\delta\beta_1)^{T-k}}{1-\delta\beta_1}\right)\end{aligned} \tag{3.13-2)}$$

y

$$\sum_{j=1}^{T-k}\delta^j\gamma_j(\boldsymbol{\theta},\delta)=\delta\frac{1-\delta^{2(T-k)}}{1-\delta^2}\left(\delta+\frac{\alpha_1}{1-(\beta_1/\delta)}\right)-\frac{\alpha_1\beta_1}{1-(\beta_1/\delta)}\frac{1-(\delta\beta_1)^{T-k}}{1-\delta\beta_1} \qquad (3.13\text{-}3)$$

de forma que

$$\sum_{j=1}^{T-k}\gamma_j^2(\boldsymbol{\theta},\delta)-\sum_{j=1}^{T-k}\delta^j\gamma_j(\boldsymbol{\theta},\delta)=\frac{\alpha_1}{1-(\beta_1/\delta)}\left[\left(\delta+\frac{\alpha_1}{1-(\beta_1/\delta)}\right)\frac{1-\delta^{2(T-k)}}{1-\delta^2}-\beta_1\left(\frac{1-(\delta\beta_1)^{T-k}}{1-\delta\beta_1}+\frac{\alpha_1\beta_1}{1-(\beta_1/\delta)}\frac{1-\beta_1^{2(T-k)}}{\delta^2(1-\beta_1^2)}\right)\right] \qquad (3.13\text{-}4)$$

El sesgo asintótico viene dado, en este caso, por:

$$\lim_{T\to\infty}E[\hat{\omega}_T(\boldsymbol{\theta},k)-\omega]=-\frac{\omega\alpha_1}{1-(\beta_1/\delta)}\times\frac{\left(\frac{\delta+\alpha_1-\beta_1}{1-(\beta_1/\delta)}\right)\frac{1}{1-\delta^2}-\beta_1\left(\frac{1}{1-\delta\beta_1}+\frac{\alpha_1\beta_1}{1-(\beta_1/\delta)}\frac{1}{\delta^2(1-\beta_1^2)}\right)}{1+\frac{1}{1-\delta^2}\left(\frac{\delta+\alpha_1-\beta_1}{1-(\beta_1/\delta)}\right)^2-\frac{\alpha_1\beta_1}{1-(\beta_1/\delta)}\left(\frac{\alpha_1\beta_1}{1-(\beta_1/\delta)}\frac{1}{\delta^2(1-\beta_1^2)}+\frac{2}{1-\delta\beta_1}\right)} \qquad (3.13\text{-}5)$$

De la misma forma, estudiamos el posible sesgo en la estimación de ω derivado de asumir un tipo de outlier en volatilidad distinto del que realmente pudiese existir en la serie observada. En particular, presentamos los resultados de la estimación de ω basado en el supuesto de un AVO, existiendo realmente un outlier en volatilidad de tipo innovacional, y el caso contrario, búsqueda de un outlier innovacional cuando el outlier existente es de tipo aditivo.

3.1.4 CASO AVO vs. IVO

Suponiendo que **se** busca un outlier en volatilidad de tipo aditivo (AVO), de forma que $x_t(\boldsymbol{\theta},k)$ viene dado por la ecuación (3.7-1), y que realmente existe un outlier en volatilidad de tipo innovacional (IVO), de forma que el efecto $\tilde{\eta}_t(\omega,k)$ viene dado por la ecuación (2.11), se tiene entonces que, bajo el supuesto de parámetros GARCH conocidos, el valor esperado del estimador MCO de la magnitud viene dado por

$$E[\hat{\omega}_T(\boldsymbol{\theta},\omega)]=\omega\frac{1+\alpha_1^2\left(\frac{1-(\beta_1(\alpha_1+\beta_1))^{T-k}}{1-\beta_1(\alpha_1+\beta_1)}\right)}{1+\alpha_1^2\left(\frac{1-\beta_1^{2(T-k)}}{1-\beta_1^2}\right)} \qquad (3.14\text{-}1)$$

con sesgo,

$$E[\hat{\omega}_T(\boldsymbol{\theta},\omega)-\omega]=-\omega\alpha_1^2\frac{\left(\frac{1-\beta_1^{2(T-k)}}{1-\beta_1^2}-\frac{1-\beta_1^{T-k}(\alpha_1+\beta_1)^{T-k}}{1-\beta_1(\alpha_1+\beta_1)}\right)}{1+\alpha_1^2\left(\frac{1-\beta_1^{2(T-k)}}{1-\beta_1^2}\right)} \qquad (3.14\text{-}2)$$

El sesgo asintótico viene dado en este caso por

$$\lim_{T\to\infty} E[\hat{\omega}_T(\boldsymbol{\theta},\omega)-\omega] = \omega\frac{\alpha_1^2}{1-\beta_1^2}\frac{\dfrac{\alpha_1\beta_1}{1-\beta_1^2}}{\left(1+\dfrac{\alpha_1^2}{1-\beta_1^2}\right)\left(1-\dfrac{\alpha_1\beta_1}{1-\beta_1^2}\right)} \quad (3.14\text{-}3)$$

que será siempre positivo, de forma que en este caso el resultado es una sobreestimación de la magnitud del outlier en volatilidad.

3.1.5 Caso IVO vs AVO

Por último, consideramos el caso contrario al anterior, es decir, cuando se parte del supuesto de existencia de un outlier en volatilidad de tipo innovacional, de forma que $x_t(\boldsymbol{\theta},k)$ viene dado por la ecuación (3.7-2) y realmente existe un outlier de tipo aditivo, cuando $\tilde{\eta}_t(\omega,k)$ se ajusta al efecto descrito por la ecuación (2.10). Entonces, se tiene que $E[\hat{\omega}_T(\boldsymbol{\theta},\omega)]=\omega$, de forma que en este caso la estimación puntual no se ve afectada por la incorrecta especificación del tipo de outlier en volatilidad. Obviamente, el error de especificación si que tiene consecuencias sobre el procedimiento estándar de contrastación de la significación de la magnitud del outlier a través de la varianza incorrecta del estimador MCO.

El resultado más relevante común a todos los casos analizados es que el sesgo es proporcional a la magnitud del outlier en volatilidad y al parámetro AR α_1 de la especificación del modelo GARCH(1,1). Así, siempre que la magnitud del outlier en volatilidad sea positivo, el empleo del estimador (3.8) resultará en una infraestimación de ω lo que afectará al procedimiento de corrección de las observaciones en base a dicha estimación. Por esa razón, en la sección 6, en el estudio numérico de simulación se utilizan valores relativamente pequeños para ω, para comprobar que incluso en estos casos la inferencia puede verse afectada.

3.2 Corrección por sesgo en la estimación MCO de ω

Puesto que en todos estos casos se tiene que, en general, el valor esperado del estimador MCO de la magnitud de un outlier en volatilidad viene dado por

$$E[\hat{\omega}_T(\boldsymbol{\theta},k)] = \omega\cdot b_T(\boldsymbol{\theta},k) \quad (3.15)$$

con $b_T(\boldsymbol{\theta},k)=1+(E[\hat{\omega}_T(\boldsymbol{\theta},k)-\omega])/\omega$, entonces teniendo en cuenta la expresiones resultantes para el sesgo obtenidas anteriormente, el estimador insesgado de la magnitud de un outlier en volatilidad en el caso de parámetros GARCH conocidos viene dado por

$$\tilde{\omega}_T(\boldsymbol{\theta},k) = b_T^{-1}(\boldsymbol{\theta},k)\cdot\hat{\omega}_T(\boldsymbol{\theta},k) \quad (3.16)$$

Así, por ejemplo, en el caso AVO, se tendría que el estimador insesgado de ω es

$$\tilde{\omega}_T(\boldsymbol{\theta},k) = \left(1+\alpha_1^2\frac{1-\beta_1^{2(T-k)}}{1-\beta_1^2}\right)\cdot\hat{\omega}_T(\boldsymbol{\theta},k) \quad (3.17)$$

En el resto de caso, el procedimiento es idéntico aunque con expresiones más complicadas como resulta de las ecuaciones (3.12) y (3.13) anteriores. La relevancia de esta corrección resulta inmediata observando que el sesgo en todos los casos no desaparece asintóticamente y que es tanto mayor cuánto mayor es el valor del parámetro AR de la ecuación GARCH(1,1), α_1.

4. ESTIMACIÓN MCO DE LA MAGNITUD DE OUTLIERS EN VOLATILIDAD EN UN MODELO GARCH(1,1): CASO DE PARÁMETROS DESCONOCIDOS

En esta sección presentamos resultados aproximados relativos al sesgo en la estimación QMV de los parámetros de un modelo GARCH(1,1) cuando se ignora la existencia de un outlier

en volatilidad y como este error se traslada a la estimación de la magnitud de dicho outlier en el procedimiento descrito en la sección anterior.

4.1 Sesgo aproximado del estimador QMV de θ bajo normalidad condicional en el caso de contaminación por un Outlier en Volatilidad

Si se considera la estimación de los parámetros GARCH(1,1), $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1)$ a partir de las observaciones del proceso GARCH contaminado por un outlier en volatilidad dado por (2.6) bajo el supuesto de normalidad condicional en la distribución de las innovaciones, la función de log-verosimilitud condicional para la observación *t*-ésima viene dada por

$$\ell_t(\boldsymbol{\theta}) = -\frac{1}{2}\left(\ln h_t(\boldsymbol{\theta},\omega) + \zeta_t^2(\boldsymbol{\theta})\right) \quad t = 1,...,T \tag{4.1}$$

donde el término de innovación viene dado por $\zeta_t(\boldsymbol{\theta}) = \varepsilon_t(\omega)/\sqrt{h_t(\boldsymbol{\theta},\omega)}$, y, por tanto, bajo omisión de la perturbación por un outlier en volatilidad se tiene que recoge el efecto de la omisión a través de la siguiente relación con el término de innovación *i.i.d.* ζ_t, $\zeta_t^2(\boldsymbol{\theta}) = \zeta_t^2 + \tilde{\eta}_t(\omega,k)/h_t(\boldsymbol{\theta},\omega)$. El estimador de $\boldsymbol{\theta}$ se denomina quasi (o pseudos) MV (QMV) puesto que la verdadera distribución de las innovaciones no tiene que coincidir con la asumida. Bajo correcta especificación del modelo, diversos autores han probado que aún bajo una distribución incorrecta, el estimador QMV es consistente y asintóticamente normal, siempre que sea simétrica y bajo una serie de condiciones de regularidad relativas a la distribución de las innovaciones (ver, por ejemplo, Newey y Steigerwald (1997), Berkes *et.al.* (2003), Hall y Yao (2003), y Mikosch y Straumann (2006)). El estimador QMV viene dado por la solución de la ecuación estimable

$$\mathbf{0} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{s}_t(\hat{\boldsymbol{\theta}}_T) = -\frac{1}{2T}\sum_{t=1}^{T} \frac{\boldsymbol{u}_t(\hat{\boldsymbol{\theta}}_T,\omega)}{h_t(\hat{\boldsymbol{\theta}}_T,\omega)}(1-\zeta_t^2(\hat{\boldsymbol{\theta}}_T)), \quad \boldsymbol{u}_t(\hat{\boldsymbol{\theta}}_T,\omega) = \left.\frac{\partial h_t(\boldsymbol{\theta},\omega)}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_T}$$

donde $\boldsymbol{s}_t(\boldsymbol{\theta}) = \partial \ell_t(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$ es el score de la log-verosimilitud condicional normal. Un desarrollo en serie de Taylor de primer orden en torno a $\boldsymbol{\theta}$, permite obtener la siguiente forma linealizada aproximada para el estimador QMV,

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}) = -\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \boldsymbol{Q}_T^{-1}(\boldsymbol{\theta})\boldsymbol{s}_t(\boldsymbol{\theta}) \tag{4.2}$$

donde $\boldsymbol{Q}_T(\boldsymbol{\theta})$ es una matriz definida positiva no singular. Considerando que esta matriz tiene la estructura del producto exterior del score de la log-verosimilitud (como en el caso de emplear el algoritmo BHHH, Berndt *et.al.* (1974)), $\boldsymbol{Q}_T(\boldsymbol{\theta}) = -\sum_{t=1,T} \boldsymbol{s}_t(\boldsymbol{\theta})\boldsymbol{s}_t'(\boldsymbol{\theta})/T$, y teniendo en cuenta que, de forma general, el proceso de contaminación en (2.6) puede escribirse como $\tilde{\eta}_t(\omega,k) = \omega\eta_t(\boldsymbol{\gamma},k)$, con $\eta_t(\boldsymbol{\gamma},k) = I_t^{(k)} + q_{t-k}(\boldsymbol{\gamma})S_t^{(k+1)}$, donde $q_0(\boldsymbol{\gamma}) = 1$ y $q_{t-k}(\boldsymbol{\gamma}) = 0$ si el outlier es de tipo aditivo (AVO) y $q_{t-k}(\boldsymbol{\gamma}) \neq 0$ para $t–k$ = 1, …, $T–k$, y vienen dados por las ecuaciones (2.12)-(2.14) en el resto de casos, entonces se puede obtener la siguiente representación aproximada para el sesgo del estimador QMV de $\boldsymbol{\theta}$ bajo contaminación por un outlier en volatilidad[1],

$$B[\hat{\boldsymbol{\theta}}_T] = E[\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}] = -\frac{\omega}{2T}\sum_{t=k}^{T} q_{t-k}(\boldsymbol{\gamma})\boldsymbol{Q}^{-1}(\boldsymbol{\theta},\omega)E\left[\frac{\boldsymbol{\Sigma}_t(\boldsymbol{\theta},\omega)}{h_t^2(\boldsymbol{\theta},\omega)}\right], \tag{4.3}$$

con $\boldsymbol{\Sigma}_t(\boldsymbol{\theta},\omega) = \boldsymbol{u}_t(\boldsymbol{\theta},\omega)\boldsymbol{u}_t'(\boldsymbol{\theta},\omega)$ y $\boldsymbol{Q}(\boldsymbol{\theta},\omega)$ el límite en probabilidad de la matriz $\boldsymbol{Q}_T(\boldsymbol{\theta})$, dada por

[1] Los detalles se pueden consultar en: Afonso Rodríguez, J.A. (2008). Sesgo aproximado del estimador QMV en GARCH(1,1) con outliers en volatilidad. *Manuscrito* y Afonso Rodríguez, J.A. (2008). Estimación MCO de la magnitud de outliers en volatilidad en un modelo GARCH(1,1): caso de parámetros desconocidos. *Manuscrito.*

$$\boldsymbol{Q}(\boldsymbol{\theta},\omega)=-\frac{1}{4}\left(E(\zeta_0^4-1)E\left(\frac{\boldsymbol{\Sigma}_1(\boldsymbol{\theta})}{h_1^2(\boldsymbol{\theta})}\right)+\omega^2E\left(\frac{\boldsymbol{\Sigma}_1(\boldsymbol{\theta})}{h_1^4(\boldsymbol{\theta})}\right)\right)$$

Puesto que $\boldsymbol{\Sigma}_t(\boldsymbol{\theta},\omega)=O_p(\omega^2)$, $h_t(\boldsymbol{\theta},\omega)=O_p(\omega)$ y $\boldsymbol{Q}(\boldsymbol{\theta},\omega)=O(\omega^2)$, el sesgo aproximado es $E[\hat{\boldsymbol{\theta}}_T-\boldsymbol{\theta}]=O((\omega T)^{-1})$, de forma que aumenta tanto con la magnitud del outlier como con el tamaño muestral. Si $\omega = o(T)$, entonces aún bajo contaminación el estimador QMV resultará ser consistente. En cualquier caso, incluso un outlier de magnitud moderada puede causar un sesgo importante para un tamaño muestral moderado o incluso relativamente grande.

4.2 Sesgo en la estimación de la magnitud del Outlier en Volatilidad

En el caso de que los parámetros GARCH sean desconocidos, para implementar el procedimiento de detección de outliers se requiere la estimación de $\boldsymbol{\theta}$. Así, en lugar de (3.5), se tiene ahora que la regresión para estimar ω viene dada por

$$v_t(\hat{\boldsymbol{\theta}}_T,\omega)=\omega x_t(\hat{\boldsymbol{\theta}}_T,k)+v_t(\hat{\boldsymbol{\theta}}_T) \tag{4.4}$$

donde $v_t(\hat{\boldsymbol{\theta}}_T,\omega)=\varepsilon_t^2(\omega)-h_t(\hat{\boldsymbol{\theta}}_T,\omega)$. Entonces, para evaluar el sesgo de ω se tiene que a partir de un desarrollo en serie de Taylor en torno a $\boldsymbol{\theta}$ para la ecuación GARCH(1,1) estimada, de la forma $h_t(\hat{\boldsymbol{\theta}}_T,\omega)=h_t(\boldsymbol{\theta},\omega)+(\hat{\boldsymbol{\theta}}_T-\boldsymbol{\theta})'\boldsymbol{u}_t(\boldsymbol{\theta},\omega)$, se pueden obtener expresiones apropiadas para $v_t(\hat{\boldsymbol{\theta}}_T,\omega)$ y $x_t(\hat{\boldsymbol{\theta}}_T,k)$ como sigue

$$v_t(\hat{\boldsymbol{\theta}}_T,\omega)=\tilde{v}_t(\boldsymbol{\theta},\omega)-(\hat{\boldsymbol{\theta}}_T-\boldsymbol{\theta})'\boldsymbol{u}_t(\boldsymbol{\theta},\omega) \tag{4.5}$$

$$x_t(\hat{\boldsymbol{\theta}}_T,k)=x_t(\boldsymbol{\theta},k)+(\hat{\boldsymbol{\theta}}_T-\boldsymbol{\theta})'\frac{\partial x_t(\boldsymbol{\theta},k)}{\partial\boldsymbol{\theta}} \tag{4.6}$$

empleando un desarrollo similar. Así, se tiene que, de forma aproximada

$$\begin{aligned}E[v_t(\hat{\boldsymbol{\theta}}_T,\omega)x_t(\hat{\boldsymbol{\theta}}_T,k)]&=x_t(\boldsymbol{\theta},k)(\omega\eta_t(\boldsymbol{\gamma},k)-B[\hat{\boldsymbol{\theta}}_T']E[\boldsymbol{u}_t(\boldsymbol{\theta},\omega)])\\&+\frac{\partial x_t(\boldsymbol{\theta},k)}{\partial\boldsymbol{\theta}'}\left(\omega\eta_t(\boldsymbol{\gamma},k)B[\hat{\boldsymbol{\theta}}_T]-(Var(\hat{\boldsymbol{\theta}}_T)+B[\hat{\boldsymbol{\theta}}_T]\cdot B[\hat{\boldsymbol{\theta}}_T'])E[\boldsymbol{u}_t(\boldsymbol{\theta},\omega)]\right)\end{aligned} \tag{4.7}$$

y

$$\begin{aligned}E[x_t^2(\hat{\boldsymbol{\theta}}_T,k)]&=x_t^2(\boldsymbol{\theta},k)+Tr\left((Var(\hat{\boldsymbol{\theta}}_T)+B[\hat{\boldsymbol{\theta}}_T]\cdot B[\hat{\boldsymbol{\theta}}_T'])\frac{\partial x_t(\boldsymbol{\theta},k)}{\partial\boldsymbol{\theta}}\frac{\partial x_t(\boldsymbol{\theta},k)}{\partial\boldsymbol{\theta}'}\right)\\&+2x_t(\boldsymbol{\theta},k)\frac{\partial x_t(\boldsymbol{\theta},k)}{\partial\boldsymbol{\theta}'}B[\hat{\boldsymbol{\theta}}_T]\end{aligned} \tag{4.8}$$

pudiendo evaluar el sesgo aproximado en la estimación de ω, empleando el estimador

$$\hat{\omega}_T(\boldsymbol{\theta},k)=\frac{\sum_{t=1}^T v_t(\hat{\boldsymbol{\theta}}_T,\omega)x_t(\hat{\boldsymbol{\theta}}_T,k)}{\sum_{t=1}^T x_t^2(\hat{\boldsymbol{\theta}}_T,k)}$$

a partir de (4.7) y (4.8). Puesto que del análisis de estas expresiones resulta que $E[v_t(\hat{\boldsymbol{\theta}}_T,\omega)x_t(\hat{\boldsymbol{\theta}}_T,k)]=O(\omega)$ y $E[x_t^2(\hat{\boldsymbol{\theta}}_T,k)]=O(1/(\omega T))$, se puede concluir que en el caso donde la magnitud del outlier sea moderada, $\omega = o(T)$, el sesgo en la estimación de ω es aproximadamente de la misma magnitud que en el caso de parámetros GARCH conocidos. Particularizando estas expresiones al caso IVO, se tiene que

$$\sum_{t=1}^{T}E[v_t(\hat{\boldsymbol{\theta}}_T,\omega)x_t(\hat{\boldsymbol{\theta}}_T,k)]=\omega-B[\hat{\boldsymbol{\theta}}_T']E[\boldsymbol{u}_k(\boldsymbol{\theta},\omega)]=\omega-O(1/T)$$

mientras que $\sum_{t=1,T}E[x_t^2(\hat{\boldsymbol{\theta}}_T,k)]=1$, de forma que como ocurre en el caso de parámetros GARCH conocidos, el estimador de ω es (casi) insesgado. En los resultados de simulación del siguiente epígrafe podrá comprobarse numéricamente este resultado.

6. RESULTADOS DE SIMULACIÓN

En este epígrafe estudiamos los resultados numéricos de un experimento de simulación, cuantificando la magnitud estimada de un outlier en volatilidad en un proceso GARCH(1,1) contaminado, bajo distintos supuestos sobre la distribución de las innovaciones y valores de los parámetros GARCH. Con el objeto de medir también la posible influencia del sesgo adicional debido a la estimación de los parámetros GARCH, consideramos las dos situaciones posibles tratadas en los epígrafes 4 y 5. En el caso de parámetros GARCH desconocidos, empleamos la estimación QMV bajo el supuesto de normalidad condicional puesto que en este caso, bajo no contaminación, la estimación resultante es consistente.

En todos los casos analizados, se ha tratado de medir directamente la magnitud del sesgo considerando que la localización del sesgo es conocida aunque se genera en una posición aleatoria dentro de la muestra. El proceso de contaminación viene dado por

$$\tilde{\eta}_t(\omega, k) = \omega\xi(L)b_t(\pi)$$

donde $b_t(\pi)$ es un proceso i.i.d. de Bernoulli, con $\pi = P(b_t(\pi)=1)$. Tomando $\pi = 1/T$, se consigue un único outlier en volatilidad en la serie de T observaciones. En cada caso se ha considerado también que la magnitud de ω es fija y toma valores relativamente pequeños, ω = 1.5, 2.5 y 5.0. Los tamaños muestrales considerados son T = 100, 250, 500, 1000 y 2000 y la distribución de las innovaciones consideradas son: $N(0,1)$, $ST(d)$ (T de Student estandarizada) con d = 6, 9 y 12 y $SGED(\upsilon)$ (distribución generalizada del error estandarizada) con υ = 0.5, 1.0, 1.5 y 3.0. Con estas dos últimas distribuciones se pretende evaluar si existe algún tipo de efecto adicional relevante sobre el sesgo del estimador MCO de ω al considerar distribuciones de colas pesadas para las innovaciones del proceso GARCH(1,1), pero sin incumplir el requerimiento mínimo para la consistencia de existencia de la varianza de ζ_t. También se presentan algunos resultados para distintas combinaciones de valores de los parámetros GARCH, manteniendo aproximadamente el mismo grado de persistencia, medido por $\alpha_1+\beta_1$.

Los siguientes cuadros 1 a 4 recogen una selección de los principales resultados obtenidos en el caso de estimación MCO, mientras que los cuadros 5 a 7 presentan los resultados de la estimación MCO corregida por el sesgo multiplicativo.

Cuadro 1. ESTIMADOR MCO. Caso de parámetros GARCH conocidos (AVO)
$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1) = (0.05, 0.15, 0.65)$

1.A	Distribución de las innovaciones						
ω = 1.5		*ST*(*d*)			*SGED*(υ)		
	N(0,1)	*d*= 6	*d*= 9	*d*= 12	υ = 0.5	υ = 1.0	υ = 3.0
T = 100	1.4428	1.4392	1.4525	1.4142	1.4016	1.4041	1.4123
	0.3492	0.4675	0.4699	0.4007	1.1241	0.5426	0.1825
	0.3544	0.4876	0.4314	0.4049	1.2941	0.5422	0.1726
T = 250	1.4595	1.4268	1.4413	1.4277	1.3869	1.4329	1.4245
	0.4009	0.5843	0.5976	0.4532	0.9000	0.5412	0.3385
	0.3639	0.5434	0.4616	0.4246	1.9259	0.5341	0.1703
T = 500	1.4501	1.4345	1.4451	1.4192	1.4267	1.4206	1.4447
	0.4029	0.7407	0.4643	0.5393	3.1679	0.5537	0.2096
	0.3669	0.5738	0.4695	0.4359	2.4752	0.6149	0.1699
T = 1000	1.4393	1.4258	1.4435	1.4159	1.3994	1.4349	1.4353
	0.3631	0.5313	0.4562	0.9656	2.0606	0.8785	0.3038
	0.3684	0.5976	0.4759	0.4382	3.2679	0.6314	0.1692
T = 2000	1.4480	1.4323	1.4436	1.4444	1.4328	1.4645	1.4094
	0.3996	0.5830	0.4387	0.4627	3.6475	0.7177	0.6826
	0.3699	0.6195	0.4799	0.4442	4.1198	0.6429	0.1693

Cuadro 1. ESTIMADOR MCO. Caso de parámetros GARCH conocidos (AVO)
$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1) = (0.05, 0.15, 0.65)$
(*continuación*)

1.B	Distribución de las innovaciones						
$\omega = 2.5$		*ST*(*d*)			*SGED*(υ)		
	N(0,1)	*d*= 6	*d*= 9	*d*= 12	υ = 0.5	υ = 1.0	υ = 3.0
T = 100	2.4024	2.3375	2.3339	2.3396	2.3085	2.3329	2.3592
	0.3599	0.4869	0.3944	0.4142	1.3827	0.5080	0.1963
	0.3545	0.4900	0.4322	0.4086	1.2882	0.5451	0.1726
T = 250	2.4069	2.3703	2.3882	2.3768	2.3676	2.3619	2.3863
	0.3785	0.5405	0.5048	0.4214	1.6578	0.6657	0.1869
	0.3635	0.5445	0.4577	0.4252	1.9307	0.5923	0.1704
T = 500	2.4141	2.3929	2.3854	2.3939	2.2578	2.3706	2.3986
	0.4191	0.6240	0.6679	0.4969	2.4125	0.6019	0.1824
	0.3665	0.5733	0.4724	0.4354	2.5788	0.6149	0.1698
T = 1000	2.3994	2.3411	2.4014	2.3999	2.3404	2.3774	2.3934
	0.3627	3.1429	0.4968	0.5855	1.2409	0.7133	0.6433
	0.3679	0.5943	0.4789	0.4383	3.1235	0.6333	0.1694
T = 2000	2.3864	2.3049	2.3819	2.3687	2.2809	2.3500	2.4023
	0.4608	2.3802	1.1187	1.7578	1.2737	0.8569	0.1829
	0.3701	0.6177	0.4847	0.4433	3.8729	0.6433	0.1691

1.C	Distribución de las innovaciones						
$\omega = 5.0$		*ST*(*d*)			*SGED*(υ)		
	N(0,1)	*d*= 6	*d*= 9	*d*= 12	υ = 0.5	υ = 1.0	υ = 3.0
T = 100	4.7087	4.6662	4.6899	4.7028	4.6559	4.6688	4.6874
	0.4265	0.8022	0.5339	0.4771	1.1976	0.5613	0.2435
	0.3529	0.4903	0.4297	0.4073	1.3030	0.5429	0.1728
T = 250	4.7559	4.7398	4.7496	4.7267	4.6652	4.7188	4.7616
	0.4079	0.5512	0.4971	0.9278	1.6779	1.3124	0.2119
	0.3628	0.5447	0.4559	0.4278	1.8964	0.5967	0.1705
T = 500	4.7832	4.7795	4.7830	4.7817	4.7006	4.7575	4.7541
	0.4444	0.5748	0.5041	0.4755	1.5785	0.6726	0.7773
	0.3669	0.5803	0.4728	0.4353	2.5709	0.6159	0.1698
T = 1000	4.7791	4.8008	4.7505	4.7865	4.8053	4.8033	4.7908
	0.6868	0.9754	1.7612	0.4809	2.5337	0.7018	0.2034
	0.3687	0.6065	0.4797	0.4421	3.3526	0.6291	0.1693
T = 2000	4.7772	4.7531	4.7847	4.7835	4.7528	4.7643	4.8007
	1.1431	1.7425	0.6473	0.6390	2.5943	1.3347	0.2689
	0.3699	0.6137	0.4851	0.4435	3.8379	0.6438	0.1692

Cuadro 2. ESTIMADOR MCO. Caso de parámetros GARCH desconocidos (QMV-G) (AVO)
$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1) = (0.05, 0.15, 0.65)$

2.A	Distribución de las innovaciones						
$\omega = 1.5$		*ST*(*d*)			*SGED*(υ)		
	N(0,1)	*d*= 6	*d*= 9	*d*= 12	$\upsilon = 0.5$	$\upsilon = 1.0$	$\upsilon = 3.0$
T = 100	1.4006	1.3900	1.3936	1.4130	1.3561	1.4058	1.4090
	0.3734	0.6215	0.4508	0.4170	1.0569	0.5844	0.1983
	0.3674	0.4689	0.4139	0.3973	0.9715	0.5129	0.1817
T = 250	1.4334	1.4057	1.4052	1.4226	1.3957	1.4242	1.4258
	0.4461	0.4919	0.4159	0.4354	1.4999	0.5665	0.2743
.	0.3616	0.5019	0.4341	0.4169	1.2778	0.5576	0.1826
T = 500	1.4383	1.3973	1.4073	1.4150	1.3797	1.4123	1.4435
	0.5964	0.5435	0.4695	0.5269	1.7974	0.5688	0.2079
	0.3782	0.5445	0.4678	0.4314	1.6066	0.5803	0.1732
T = 1000	1.4205	1.4258	1.3884	1.4179	1.3105	1.4358	1.4359
	1.0101	0.5237	1.4341	0.9059	4.4090	0.8348	0.2837
	0.4036	0.5722	0.4944	0.4436	1.8127	0.6109	0.1758
T = 2000	1.4304	1.3879	1.4058	1.4427	1.3812	1.4602	1.4195
	0.5609	1.5355	1.3244	0.4642	2.1992	0.7200	0.2205
	0.3882	0.6154	0.5364	0.4394	2.1928	0.6200	0.2169

2.B	Distribución de las innovaciones						
$\omega = 2.5$		*ST*(*d*)			*SGED*(υ)		
	N(0,1)	*d*= 6	*d*= 9	*d*= 12	$\upsilon = 0.5$	$\upsilon = 1.0$	$\upsilon = 3.0$
T = 100	2.3342	2.3289	2.3207	2.3204	2.2933	2.3015	2.3358
	0.4340	0.5443	0.4845	0.5080	1.3803	0.6379	0.2591
	0.3734	0.4797	0.4378	0.4219	1.0025	0.5376	0.1986
T = 250	2.3817	2.3429	2.3815	2.3636	2.3321	2.3396	2.3765
	0.4335	0.5229	0.5409	0.4529	1.8707	0.6974	0.2079
	0.3725	0.5146	0.4535	0.4228	1.3169	0.5651	0.1790
T = 500	2.3799	2.3654	2.3799	2.3852	2.1864	2.3589	2.3953
	0.4686	0.7437	0.6231	0.5039	2.1487	0.6303	0.1887
	0.3827	0.5389	0.4765	0.4293	1.6113	0.5835	0.1739
T = 1000	2.3759	2.3528	2.3965	2.3976	2.2900	2.3710	2.3975
	0.5114	0.5738	0.5133	0.5461	1.5939	0.7210	0.4656
	0.3783	0.5762	0.4686	0.4495	1.8601	0.6088	0.2455
T = 2000	2.4037	2.3921	2.3848	2.3715	2.2402	2.3465	2.4020
	0.3847	0.7663	0.9358	1.5775	1.7088	0.8023	0.1851
	0.3705	0.6209	0.4898	0.4615	2.1713	0.6419	0.1702

Cuadro 2. ESTIMADOR MCO. Caso de parámetros GARCH desconocidos (QMV-G) (AVO)
$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1) = (0.05, 0.15, 0.65)$ (*continuación*)

2.C	Distribución de las innovaciones						
$\omega = 5.0$		*ST*(*d*)			*SGED*(υ)		
	N(0,1)	*d*= 6	*d*= 9	*d*= 12	$\upsilon = 0.5$	$\upsilon = 1.0$	$\upsilon = 3.0$
T = 100	4.6047	4.5565	4.5548	4.5803	4.6307	4.5493	4.5023
	0.7265	1.0869	0.9149	0.8413	1.7354	0.9993	0.6341
	0.4291	0.5640	0.4977	0.4831	1.0813	0.5884	0.2635
T = 250	4.7133	4.6707	4.6977	4.6737	4.5634	4.6682	4.6859
	0.5073	0.7531	0.6499	0.8595	2.0269	1.2491	0.3511
	0.3963	0.5440	0.4788	0.4685	1.3214	0.6206	0.2020
T = 500	4.7682	4.7279	4.7541	4.7644	4.5468	4.7159	4.7288
	0.4826	0.6514	0.5601	0.5259	1.7847	0.7587	0.6696
	0.3815	0.5528	0.4752	0.4461	1.7184	0.6083	0.2015
T = 1000	4.7744	4.7805	4.7464	4.7762	4.6437	4.7963	4.7812
	0.6564	0.9554	1.3258	0.5033	2.9004	0.7391	0.2235
	0.3825	0.5897	0.5165	0.4453	1.9663	0.6490	0.1762
T = 2000	4.7819	4.7544	4.7831	4.7824	4.6503	4.7643	4.7985
	0.8763	1.4849	0.6252	0.6231	2.8447	1.1954	0.2699
	0.3934	0.6385	0.4963	0.4503	2.2014	0.6454	0.1746

Notas a los cuadros 1 y 2: Los resultados están basados en 2000 replicaciones. En la primera fila para cada tamaño muestral se muestra el promedio de las estimaciones MCO de la magnitud del outlier, $\hat{\omega}_T(\theta, k)$. En la segunda y tercera filas se presentan, respectivamente, la desviación estándar de las replicaciones y el promedio de la desviación estándar teórica del estimador MCO de ω basado en la distribución asumida para las innovaciones.

Cuadro 3.ESTIMADOR MCO (Caso TCVO) ($\delta = 0.50$). $\theta = (\alpha_0, \alpha_1, \beta_1) = (0.05, 0.15, 0.65)$

	Distribución de las innovaciones: *N*(0,1)					
	Caso parámetros GARCH conocidos			Caso parámetros GARCH estimados		
	$\omega = 1.5$	$\omega = 2.5$	$\omega = 5.0$	$\omega = 1.5$	$\omega = 2.5$	$\omega = 5.0$
T = 100	1.2739	2.1401	3.7209	1.1387	1.8127	3.3746
	0.3920	0.9738	0.7711	0.3861	0.5339	0.9992
	0.2758	0.2766	0.2767	0.2848	0.3109	0.3994
T = 250	1.2511	2.1538	4.0115	1.1754	1.9929	3.8314
	0.4528	0.5149	0.7411	1.4221	0.4945	0.8533
	0.2821	0.2801	0.2820	0.3086	0.3009	0.3349
T = 500	1.2862	2.1236	4.1408	1.2455	2.0608	4.0464
	0.4073	0.5036	0.7598	0.4012	0.5206	0.8129
	0.2849	0.2851	0.2848	0.2865	0.2902	0.3099
T = 1000	1.2872	2.1322	4.2077	1.2785	2.0894	4.1646
	0.4196	0.6949	0.7973	0.4094	0.4894	0.8143
	0.2852	0.2857	0.2855	0.2859	0.2882	0.3024
T = 2000	1.2833	2.1398	4.2355	1.2738	2.0765	4.2125
	0.4170	0.5162	0.7741	0.4219	1.7882	0.7837
	0.2867	0.2868	0.2866	0.2869	0.3153	0.2937

Nota: Los resultados están basados en 2000 replicaciones. En la primera fila para cada tamaño muestral se muestra el promedio de las estimaciones MCO de la magnitud del outlier, $\hat{\omega}_T(\theta, k)$. En la segunda y tercera filas se presentan, respectivamente, la desviación estándar de las replicaciones y el promedio de la desviación estándar teórica del estimador MCO de ω basado en la distribución asumida para las innovaciones.

Cuadro 4. ESTIMADOR MCO: Efecto de los valores de θ (Caso AVO)

4.A	Distribución de las innovaciones: $N(0,1)$			Caso parámetros GARCH estimados		
	$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1) = (0.05, 0.25, 0.45)$			$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1) = (0.05, 0.45, 0.25)$		
	$\omega = 1.5$	$\omega = 2.5$	$\omega = 5.0$	$\omega = 1.5$	$\omega = 2.5$	$\omega = 5.0$
$T = 100$	1.3469	2.2257	4.2691	1.2730	2.0513	3.8151
	0.3587	0.5112	1.1539	0.5569	0.8184	1.7974
	0.2577	0.2910	0.3912	0.3065	0.3474	0.5357
$T = 250$	1.3685	2.2856	4.4980	1.2351	2.0589	3.9807
	0.3743	0.4099	0.7990	0.6028	0.7655	1.4799
	0.2604	0.2647	0.3203	0.3045	0.3235	0.4257
$T = 500$	1.3861	2.2933	4.5539	1.2178	2.0399	4.0427
	0.4838	0.7222	0.6166	0.6898	1.0864	1.5217
	0.2627	0.2871	0.2847	0.3304	0.3440	0.3913
$T = 1000$	1.3729	2.2919	4.5737	1.1796	2.0543	4.0573
	0.4226	0.6530	1.1021	1.2324	0.7706	1.3057
	0.2742	0.2749	0.3117	0.3523	0.3366	0.3675
$T = 2000$	1.3960	2.3173	4.6389	1.2179	2.0207	4.1161
	0.3148	0.3525	0.6115	0.8157	1.3065	1.6889
	0.2586	0.2692	0.2724	0.3514	0.3609	0.3959

4.B	Distribución de las innovaciones: $ST(d = 6)$			Caso parámetros GARCH estimados		
	$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1) = (0.05, 0.25, 0.45)$			$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1) = (0.05, 0.45, 0.25)$		
	$\omega = 1.5$	$\omega = 2.5$	$\omega = 5.0$	$\omega = 1.5$	$\omega = 2.5$	$\omega = 5.0$
$T = 100$	1.3472	2.2051	4.2864	1.2692	2.0249	3.9668
	0.4422	0.6240	1.3700	0.6771	1.0543	2.2021
	0.3256	0.3557	0.4665	0.3696	0.4313	0.6179
$T = 250$	1.3420	2.2387	4.4389	1.2538	2.0621	3.9921
	0.3677	0.6338	1.1574	0.8737	0.9673	1.9587
	0.3603	0.3735	0.4298	0.4309	0.4438	0.5085
$T = 500$	1.3578	2.2998	4.5639	1.2211	2.0315	4.0883
	0.4051	0.9782	0.9868	0.9325	1.2856	3.0781
	0.3775	0.3927	0.4183	0.4834	0.4840	0.5272
$T = 1000$	1.3634	2.3249	4.5701	1.2109	2.0675	4.1092
	0.9359	1.7681	0.7955	0.8952	1.2921	2.4596
	0.4369	0.4089	0.4202	0.5176	0.5242	0.5419
$T = 2000$	1.4048	2.3346	4.5757	1.2051	2.0188	4.0297
	0.7619	1.8591	0.7821	0.8244	0.9411	1.5774
	0.4363	0.4469	0.4289	0.6052	0.5623	0.5794

Nota: Los resultados están basados en 2000 replicaciones. En la primera fila para cada tamaño muestral se muestra el promedio de las estimaciones MCO de la magnitud del outlier, $\hat{\omega}_\tau(\theta, k)$. En la segunda y tercera filas se presentan, respectivamente, la desviación estándar de las replicaciones y el promedio de la desviación estándar teórica del estimador MCO de ω basado en la distribución asumida para las innovaciones.

Cuadro 5. ESTIMADOR MCO CORREGIDO POR SESGO (Caso AVO)
Parámetros GARCH conocidos
$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1) = (0.05, 0.15, 0.65)$

		Distribución de las innovaciones					
			ST(*d*)		*SGED*(υ)		
		N(0,1)	*d*= 6	*d*= 9	υ = 0.5	υ = 1.0	υ = 3.0
T = 100	ω = 1.5	1.4948	1.4824	1.4579	1.3983	1.4578	1.4664
	2.5	2.4984	2.4908	2.4383	2.3968	2.4221	2.4495
	5.0	4.9962	4.9791	4.8765	4.8336	4.8479	4.8671
T = 250	ω = 1.5	1.4937	1.5039	1.4828	1.4699	1.4884	1.4797
	2.5	2.4989	2.5078	2.4716	2.4593	2.4532	2.4786
	5.0	4.9835	4.9975	4.9249	4.8461	4.9014	4.9455
T = 500	ω = 1.5	1.4838	1.5055	1.4777	1.4723	1.4758	1.5008
	2.5	2.4852	2.4834	2.4645	2.3454	2.4626	2.4917
	5.0	4.9897	4.9919	4.9590	4.8831	4.9423	4.9683
T = 1000	ω = 1.5	1.5027	1.4924	1.4739	1.5474	1.4907	1.4912
	2.5	2.4875	2.4934	2.5103	2.4314	2.4698	2.4864
	5.0	4.9803	5.0102	4.9859	4.9922	4.9901	4.9771
T = 2000	ω = 1.5	1.5098	1.4903	1.4854	1.4566	1.5215	1.4643
	2.5	2.4853	2.4870	2.4939	2.3696	2.4413	2.4958
	5.0	4.9947	4.9814	4.9722	4.9379	4.9498	4.9875

Cuadro 6. ESTIMADOR MCO CORREGIDO POR SESGO (Caso AVO)
Parámetros GARCH desconocidos (QMV-G)
$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1) = (0.05, 0.15, 0.65)$

		Distribución de las innovaciones					
			ST(*d*)		*SGED*(υ)		
		N(0,1)	*d*= 6	*d*= 9	υ = 0.5	υ = 1.0	υ = 3.0
T = 100	ω = 1.5	1.4703	1.4720	1.4685	1.4357	1.4904	1.4694
	2.5	2.4595	2.4625	2.4577	2.4518	2.4339	2.4459
	5.0	4.8991	4.9344	4.8694	4.9729	4.8303	4.8262
T = 250	ω = 1.5	1.4874	1.4934	1.4888	1.4802	1.4898	1.4826
	2.5	2.4789	2.4755	2.4699	2.4999	2.4562	2.4746
	5.0	4.9382	4.946	4.9189	4.8248	4.9066	4.9176
T = 500	ω = 1.5	1.4974	1.4943	1.4782	1.4739	1.4741	1.5008
	2.5	2.4807	2.4627	2.4689	2.2846	2.4591	2.4908
	5.0	4.9635	4.9513	4.9523	4.8182	4.9277	4.9286
T = 1000	ω = 1.5	1.4959	1.4720	1.4768	1.4999	1.4961	1.4915
	2.5	2.4798	2.4975	2.5072	2.4342	2.4699	2.4905
	5.0	4.9832	4.9577	4.9818	4.9413	4.9925	4.9722
T = 2000	ω = 1.5	1.4939	1.4891	1.4910	1.4616	1.5185	1.4752
	2.5	2.4656	2.4808	2.4932	2.3696	2.4405	2.4953
	5.0	4.9822	4.9853	4.9742	4.8945	4.9559	4.9865

En general se observa que el efecto del sesgo en la estimación puntual de ω se traslada también al cálculo de la estimación de la desviación estándar del estimador, provocando una infraestimación de dicha magnitud, excepto en los casos de distribuciones con colas muy anchas para las innovaciones GARCH, donde opera al contrario. Esto afectará, por tanto, también a la inferencia estándar para la detección de outlier en este procedimiento. En los cuadros 5 a 7 se presentan los resultados de la estimación MCO de ω corregida por el sesgo en el caso aditivo bajo distintos supuestos distribucionales. Para todas las combinaciones de valor nominal de ω, tamaños muestrales, distribuciones y combinaciones de valores de los parámetros GARCH, se consigue una corrección total del sesgo. En estos cuadros no se presenta la estimación de la desviación estándar de los estimadores en la simulación puesto que es (casi) proporcional, en la

magnitud del factor de corrección multiplicativo, a la desviación estándar del estimador sin corregir.

En general se aprecia, en el caso de parámetros desconocidos, como el sesgo positivo que induce la estimación MV de los parámetros del modelo GARCH(1,1) contaminado por un outlier en volatilidad es también compensado por la corrección del sesgo calculada con esas mismas estimaciones, de forma que la corrección opera incluso mejor en esta situación que en el caso de parámetros conocidos.

Cuadro 7. ESTIMADOR MCO CORREGIDO POR SESGO (Caso AVO)
$\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1) = (0.05, 0.25, 0.45)$

	Caso parámetros GARCH conocidos Distribución de las innovaciones				Caso parámetros GARCH estimados Distribución de las innovaciones			
			ST(*d*)				*ST*(*d*)	
		N(0,1)	*d*= 6	*d*= 9		*N*(0,1)	*d*= 6	*d*= 9
T = 100	ω = 1.5	1.4522	1.4537	1.4505	ω = 1.5	1.4678	1.4757	1.4724
	2.5	2.4208	2.3937	2.4255	2.5	2.4536	2.4339	2.4618
	5.0	4.8308	4.8199	4.8196	5.0	4.8295	4.8199	4.8292
T = 250	ω = 1.5	1.4729	1.4555	1.4806	ω = 1.5	1.4793	1.4611	1.4898
	2.5	2.4773	2.4387	2.4578	2.5	2.4799	2.4382	2.4744
	5.0	4.9303	4.8938	4.9169	5.0	4.9208	4.8908	4.9050
T = 500	ω = 1.5	1.4625	1.4731	1.5008	ω = 1.5	1.4681	1.4702	1.4973
	2.5	2.4717	2.4898	2.4767	2.5	2.4681	2.5087	2.4801
	5.0	4.9498	4.9696	4.9381	5.0	4.9387	4.9592	4.9119
T = 1000	ω = 1.5	1.4777	1.4688	1.4820	ω = 1.5	1.4789	1.4779	1.4819
	2.5	2.4453	2.4929	2.4998	2.5	2.4558	2.5269	2.5002
	5.0	4.9056	4.9824	4.9958	5.0	4.9245	4.9504	4.9905
T = 2000	ω = 1.5	1.4590	1.5122	1.5041	ω = 1.5	1.4699	1.5186	1.5094
	2.5	2.5012	2.5192	2.4843	2.5	2.5043	2.5215	2.4869
	5.0	4.9966	4.9537	4.9914	5.0	4.9957	4.9457	4.9861

8. CONCLUSIONES Y EXTENSIONES FUTURAS

En este trabajo hemos estudiados distintos aspectos del tratamiento del modelo GARCH(1,1) bajo contaminación por un outlier en volatilidad. Los principales resultados se refieren a la evaluación del sesgo en la estimación de la magnitud de un outlier en volatilidad siguiendo la propuesta de Franses y Ghijsels (1999) y Charles y Darné (2005) y el posible perjuicio para la consistencia de este procedimiento de detección y corrección de outliers y la obtención de una expresión aproximada para el sesgo del estimador QMV bajo normalidad condicional de los parámetros de un modelo GARCH(1,1).

La situación analizada, tanto teórica como numéricamente en el ejercicio de simulación, se refieren al caso más restrictivo posible bajo el supuesto de conocimiento de la localización del outlier en volatilidad, $k = k_0$. El propósito es extender este estudio al caso más general de estimación de la magnitud de un outlier en volatilidad en el caso de localización desconocida, que se obtiene como resultado de un procedimiento secuencial de estimación de ω que recorre toda la muestra, $k = 1, 2, \ldots, T$, y posterior estimación de la localización mediante la regla

$$\hat{k}_T = k(\boldsymbol{\theta}, \omega) : \hat{\tau}_T(\boldsymbol{\theta}, \hat{k}_T) = \max_{1 \le k \le T} |\hat{\tau}_T(\boldsymbol{\theta}, k)|, \text{ donde } \hat{\tau}_T(\boldsymbol{\theta}, k) = \frac{\hat{\omega}_T(\boldsymbol{\theta}, k)}{\hat{s}_T(\boldsymbol{\theta}, k)} \quad k = 1, 2, \ldots, T$$

siendo

$$\hat{s}_T^2(\boldsymbol{\theta}, k) = \frac{\hat{\sigma}_T^2(\omega)}{\sum_{t=1}^{T} x_t^2(\boldsymbol{\theta}, k)}, \hat{\sigma}_T^2(\omega) = \frac{1}{T}\sum_{t=1}^{T} \hat{\upsilon}_t^2(\omega)$$

la habitual estimación MCO de la varianza del estimador de ω. La evidencia de simulación preliminar indica que en este caso el efecto puede ser más severo y se manifiesta además con un

sesgo en la estimación de la magnitud, pero de signo positivo, es decir, una sobreestimación de la magnitud del outlier en volatilidad.

BIBLIOGRAFÍA

AFONSO RODRÍGUEZ, J.A. (2007), "Estimación QMV de modelos GARCH versus estimación no paramétrica de la volatilidad local en series temporales financieras con outliers". XXI Reunión ASEPELT-España, Valladolid.

BATTAGLIA, F. (2006), "Bias Correction for Outlier Estimation in Time Series". Journal of Statistical Planning and Inference, 136, pp.3904-3930.

BERKES, I., L. HORVÁTH, P. KOKOSZKA (2003), "GARCH Processes: Structure and Estimation". Bernoulli, 9(2), pp.201-227.

BERNDT,E.K., B.H. HALL, R.E. HALL, J.A. HAUSMAN (1974), "Estimation and Inference in Nonlinear Structural Models". Annals of Economic and Social Measurement, 3/4, pp.653-665.

CARNERO, M.A., D. PEÑA, E. RUIZ (2006), "Effects of Outliers on the Identificatiion and Estimation of GARCH Models". Journal of Time Series Analysis, 28(4), pp.471-497.

CARNERO, M.A., D. PEÑA, E. RUIZ (2007), "Estimating GARCH Volatility in the Presence of Outliers". Manuscrito.

CHANG, I., G.C. TIAO, C. CHEN (1988), "Estimation of Time Series Parameters in the Presence of Outliers". Technometrics, 30(2), pp.193-204.

CHARLES, A., O. DARNÉ (2005), "Outliers and GARCH Models in Financial Data". Economics Letters, 86, pp.347-352.

CHEN, C., L.M.LIU (1993), "Joint Estimation of Model Parameters and Outlier Effects in Time Series". Journal of the American Statistical Association, 88(421), pp.284-297.

DE MELO MENDES, B.V. (2000), "Assessing the Bias of Maximum Likelihood Estimates of Contaminated GARCH Models". Journal of Statistical Computation and Simulation, 67, pp.359-376.

DOORNIK, J.A., M. Ooms (2005), "Outlier Detection in GARCH Models". Tinbergen Institute Discussion Paper No.05-092/4. Tinbergen Institute.

FRANSES, P.H., H. Ghijsels (1999), "Additive Outliers, GARCH and Forecasting Volatility". International Journal of Forecasting, 15, pp.1-9.

HALL, P., Q. YAO (2003), "Inference in ARCH and GARCH Models with Heavy-Tailed Errors". Econometrica, 71(1), pp.285-317.

HOTTA, L.K., R.S. TSAY (1998), "Outliers in GARCH Processes". Manuscript, Graduate School of Business. University of Chicago.

JIANG, J., Q. ZHAO, Y.V. HUI (2001), "Robust Modelling of ARCH Models". Journal of Forecasting, 20, pp.111-133.

MIKOSCH, T., D. STRAUMANN (2006), "Stable Limits of Martingale Transforms with Application to the Estimation of GARCH Parameters". The Annals of Statistics, 34(1), pp.493-522.

MULER, N., V.J. YOHAI (2002), "Robust Estimates for ARCH Models". Journal of Time Series Analysis, 23(3), pp.341-375.

MULER, N., V.J. YOHAI (2008), "Robust Estimates for GARCH Models". Journal of Statistics, Planning and Inference, 138(19), pp.2918-2940.

NEWEY, W.K., D.G. STEIGERWALD (1997), "Asymptotic Bias for Quasi-Maximum-Likelihood Estimators in Conditional Heteroskedasticity Models". Econometrica, 65(3), pp.587-599.

PENG, L., Q. YAO (2003), "Least Absolute Deviations Estimation for ARCH and GARCH Models". Biometrika, 90(4), pp.967-975.

SAKATA, S., H. WHITE (1998), "High Breakdown Point Conditional Dispersion Estimation with Application to S&P 500 Daily Returns Volatility". Econometrica, 66(3), pp.529-567.

VERHOEVEN, P., M. MCALEER (2000), "Modelling outliers and extreme observations for ARMA-GARCH processes". Working paper, University of Western Australia.

CÁLCULO AUTÓNOMO DEL RIESGO DE CRÉDITO EN BASILEA II: UN MODELO DE CÁLCULO DE PROBABILIDAD DE MORA

JON FRÍAS MENDI
E-mail: jfriasm@ibercaja.es
Departamento de Mercado de Capitales
Caja de Ahorros y Monte de Piedad de Zaragoza, Aragón y Rioja

Resumen

This article releases a credit default statistical model for the obtaining of the credit default probabilistic, individually as entity or sectorially, and independiently by the own credit entities. The main objetive of this publication is therefore try to help to understand this kind of models, step by step with a statistical software (SPSS), to any financial risk professional interested in and to help credit entities to the independient generation of them.
The realeased model pretends to achieve a simply assessment of credit solicitors' default probability, fist step for the assessment of the capital expected and unexpected loss from the own entities, obligated step for the entity capital requirement assessment. The main goal of this entire process is therefore the capital requirement management improve, obtaining this improvement from the own credit entities' branches.

Palabras clave: Capital Económico, Riesgo de Crédito, Regresión Logit.
Area temática: Modelação Económica e Financeira.

1. INTRODUCCIÓN

El modelo estadístico que analiza este artículo está destinado a calcular la Probabilidad de Mora (PD) de los solicitantes de crédito en una entidad financiera, primer paso para el cálculo por parte de esa entidad tanto de su pérdida esperada como inesperada. Este será el objetivo último, necesario para calcular las necesidades de capital, aunque para llegar al cálculo interno de la pérdida esperada, las entidades deberán ser capaces de calcular internamente mediante modelos la PD, la Severidad y la Exposición

En la actualidad se trabaja para la obtención del modelo avanzado (Probabilidad de Mora, Severidad, Exposición y Pérdida Esperada calculadas internamente por la entidad) dentro de los plazos previstos por el Comité de Supervisión de Basilea, superando así la etapa del modelo básico, consistente en que la entidad calcule su PD y reciba del Supervisor (Banco de España, en el caso español) la Severidad y la Exposición.

El modelo de cálculo de PD a analizar parte de una base de datos de solicitantes de crédito (con una profundidad mínima de entre cuatro y cinco años) que será tratada mediante software SPSS para poder realizar posteriormente el estudio estadístico sobre ella. Se realizará un análisis de las variables y posteriormente se ejecutará una regresión de tipo Logit. Este tipo de regresión asume que la probabilidad se distribuye logísticamente, al contrario que la Probit que asume una distribución Normal Acumulada (Altman et al., 1995). La regresión Logit presenta frente a la Probit una mayor sencillez de aplicación, debido en gran medida a que busca determinar la PD siguiendo la lógica de los modelos lineales, requiriendo además una menor potencia de cálculo frente a la regresión Probit. Sin embargo debe tenerse presente que existe evidencia empírica de que el comportamiento real escapa muchas veces del comportamiento real entre las variables (Greene, 1992), lo que no ha sido inconveniente para que el Logit sea el modelo más extendido para la determinación de la PD. Existen modelos lineales muy extendidos como el modelo

Z-Score Model (Altman, 1968, 2000) construido sobre la base de información contable y dirigido al análisis de contrapartidas empresariales. Este modelo busca una relación lineal entre las variables explicativas y la dependiente empleando un modelo discriminante (Magdala, 1993) para determinar las empresas con mayor y menor probabilidad de incumplir sus obligaciones de pago. La PD que busca el modelo de estudio es la exigida por el Comité de Supervisión Bancaria de Basilea, que deberá ser una Probabilidad de Mora calibrada a un año y ajustada al ciclo económico, por lo que una vez obtenidos los resultados de la regresión, habrá que calibrarlos y ajustarlos al ciclo.

Este artículo pretende ilustrar los pasos a seguir para la construcción de estos modelos, deteniéndose especialmente en los aspectos teóricos y económicos en los que se basan los mismos. También se hace especial hincapié en cómo ejecutar dicho proceso a nivel práctico, para los que se utilizará el software estadístico SPSS sobre una base de datos que replicaría la base de clientes de una entidad española de tamaño medio. El objetivo es la obtención de la PD calibrada y ajustada al ciclo, en un modelo de Scoring (Riesgo de crédito de particulares). Esto constituirá el primer escalón en el cálculo propio de la pérdida esperada e inesperada por parte de las entidades de crédito, y suministrará a las mismas un modelo de riesgo de crédito que les permitirá aplicar un criterio homogéneo en la decisión de concesión de préstamos.

2. GESTIÓN EFICIENTE DEL CAPITAL

Este tipo de modelos implica fundamentalmente dos ventajas para las entidades de crédito: La mejora de la eficiencia desde las propias oficinas de la entidad y el cálculo del Capital Económico de manera autónoma.

La primera ventaja debe enfocarse mediante la creación de una aplicación estadística, en el que la red comercial de la entidad recoja diferentes datos de cada solicitante y con ello pueda obtener una probabilidad de mora del solicitante de financiación y un dictamen para la operación. Este tipo de modelos está siendo utilizado sobre todo por entidades financieras para medir su exposición al riesgo de crédito. Sin embargo en los últimos tiempos han mostrado también interés por ellos empresas no financieras. Así, grandes empresas como eléctricas o gasistas estarían interesándose en estos modelos para determinar a qué clientes (empresas o particulares) les interesa vender sus servicios y a quiénes no.

Respecto al cálculo del Capital Económico, la posibilidad de que la propia entidad pueda obtenerlo de forma independiente, implica una mejora en la gestión de la entidad, que evitará tener un exceso de capital inmovilizado, lo que incidirá directamente en la rentabilidad final obtenida del capital. El objetivo básico de los reguladores, es la estabilidad del sistema financiero. La regulación por la que nacen estos modelos busca por tanto evitar la insolvencia de las entidades financieras, para proteger tanto a inversores como a depositantes.

En relación con estos objetivos, se definen varios tipos de capital. Capital Económico (Capital en Riesgo) como aquel capital necesario para cubrir las pérdidas inesperadas de una entidad, y Capital Regulatorio, que será el capital exigido por el regulador

El Nuevo Acuerdo de Capital de Basilea busca el acercamiento entre Capital Económico y Capital Regulatorio para conseguir una mejora en la eficiencia del sistema basando todo el análisis de capital en la distribución de pérdidas de las entidades. La Pérdida Esperada resultará de la esperanza matemática de la distribución de pérdidas, y se cubrirá mediante provisiones. Más concretamente, será aquella cantidad que en media cuantifica las pérdidas por morosidad, y que pretende obtener el gasto que supone a una entidad el hecho de operar con contrapartes de una determinada calidad crediticia.

Mientras tanto, por Pérdida Inesperada se entenderá la pérdida total menos la pérdida esperada y deberá ser cubierta por el Capital Económico ya mencionado. El cálculo de la Pérdida Inesperada de forma rigurosa requerirá el conocimiento previo de la función de densidad de probabilidad (FDP) por pérdidas de crédito (Knop et al., 2004). La FDP encierra la frecuencia con que puede ocurrir cualquier posible pérdida por crédito que pudiera darse en la cartera o en el nivel de agregación elegido.

La pérdida esperada consiste en una función que relaciona Probabilidad de Mora, Severidad y Exposición:

EL = f (PD, LGD, EAD) = PD x LGD x EAD

Siendo: EL: Pérdida Esperada
PD: Probabilidad de Mora (calibrada y ajustada al ciclo)
LGD: Severidad
EAD: Exposición

La exposición muestra a cuánto ascenderían las pérdidas sobre la posición mantenida con una contrapartida si esta se declarase en quiebra y no se diera ninguna recuperación.

La severidad mientras tanto, es el porcentaje sobre la exposición que finalmente no se recupera.

$$Severidad = \frac{Valor_\text{Re}cuperación - Gastos_\text{Re}cuperación}{Valor_Exposición}$$

Basilea contempla diversos métodos para la medición del riesgo de crédito en función del grado de sofisticación de la entidad. A mayor innovación por parte de la entidad, menor intervención del regulador en el cálculo del Capital Regulatorio.

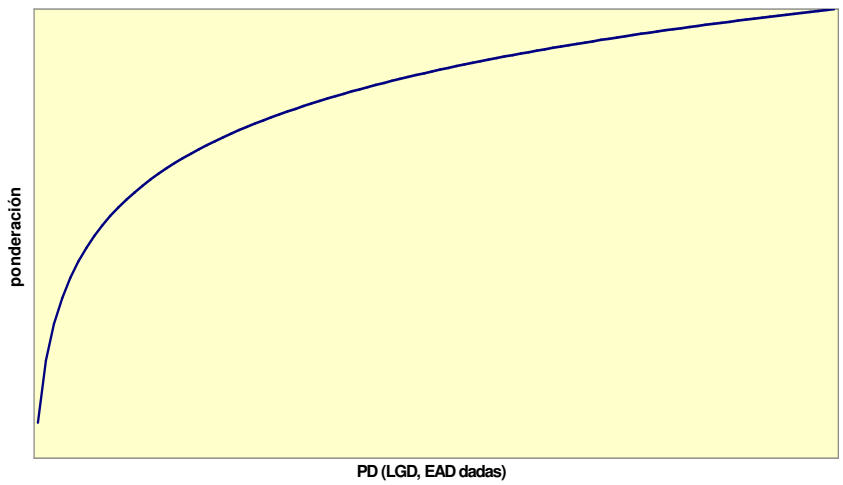
Cuadro 1: Grado de sofisticación en cada modelo de riesgo de crédito

	Prob. Mora PD	Severidad LGD	Exposición EAD	Pérd.Esperada EL	Pérd. Inesperada UL
ENFOQUE ESTANDAR	Supervisor	Supervisor	Supervisor	Supervisor	Supervisor
IRB BÁSICO	Entidad	Supervisor	Supervisor	Supervisor	Supervisor
IRB AVANZADO	Entidad	Entidad	Entidad	Entidad	Supervisor
IRAB COMPLETO	Entidd	Entidad	Entidad	Entidad	Entidad

Como se puede observar en el Cuadro 1, a medida que la entidad opta por métodos más sofisticados, es capaz de acercarse más a su Capital Económico, ya que aporta inputs propios y no externos (facilitados por el Supervisor).

En el Modelo Básico, los datos del solicitante se procesarían mediante un algoritmo que hayamos definido obteniendo de este modo su PD. Si el dictamen es Aprobar se buscará la ponderación que debe aplicarse a esa operación para una PD ya obtenida (con la Exposición y la Severidad dadas por el Supervisor) (Ver Gráfico 1)

Gráfico 1: Relación entre PD y ponderación (Modelo Básico)



Así para cada solicitante o cartera de solicitantes, una vez calculada su PD, se obtendría su ponderación, que dictará a la entidad la cantidad de capital a inmovilizar.

3. TRATAMIENTO PREVIO DE LA MUESTRA DE ESTUDIO

Todo el proceso de creación del modelo parte de una base de datos histórica de los solicitantes de crédito. En esta base deberán venir recogidas una serie de variables explicativas de los solicitantes históricos de préstamos, que servirán posteriormente para el cálculo de la probabilidad de mora de nuevos clientes.

Partiendo de un fichero que contendrá la base de datos sobre la que queremos trabajar, se ejecutarán Sintax (órdenes) dentro de SPSS para tratar la base y poder trabajar con ciertas garantías sobre ella. Los Sintax fundamentales se analizan a continuación.

SELECCIÓN

Su objetivo es una primera organización del fichero de texto recibido. Permite seleccionar los datos que resulten relevantes para el análisis de la base. Estas órdenes (SPSS: "Data List File"; "Write outfile") nos ayudarán a organizar el fichero recibido separando por ejemplo la información de consumo, hipotecas y tarjetas, para obtener la deseada para el análisis.

LECTURA

Intenta facilitar la posterior lectura de los datos mediante el etiquetado de las variables. (para evitar que la información llegue a ser incomprensible, SPSS: *"Variable Label"*).

Asimismo es interesante describir el significado de cada categoría o agrupación de categoría para evitar futuros errores de interpretación. ("*Value Label*": Describe el significado de cada categoría en las variables categóricas). En este punto se recodificarán las variables que vengan mal definidas y se eliminarán los registros que pueden originar problemas al ejecutar los comandos en SPSS (Comando "Select if" dentro del Sintax SPSS).

BALANCE DE CAJA

El objetivo será calcular el balance de caja de los solicitantes, esto es, su capacidad de ahorro mensual. Se calcula como la diferencia entre los ingresos y gastos mensuales teniendo siempre en cuenta las diferencias en el poder adquisitivo por zona geográfica y el número de componentes de la unidad familiar.

El cálculo de los Ingresos no suele generar problemas ya que se supone que vienen bien definidos los ingresos de cada solicitante.

El cálculo de los Gastos es ligeramente más complicado ya que habrá que tener en cuenta que un mismo gasto por ejemplo en dos provincias distintas, no suponen lo mismo para el que gasta. Una posible estructura de gastos podría ser la siguiente

Gastos = G. Fijos + G. Alquiler + G. Primera necesidad + G. Ocio

Una vez definida la estructura se deberán definir las variables. Según esta estructura se definiría primero la variable "Gastos de alquiler". La variable "Factor Monetario" se definirá como un factor de descuento que dependerá de la provincia o región en la que viva el solicitante, variable necesaria para calcular tanto los gastos de primera necesidad como los de ocio, y en consecuencia el balance de caja.

Se definirá a continuación la variable "Factor de Reducción", que intenta plasmar el efecto de las economías de escala sobre los gastos, siendo utilizado en el cálculo de los gastos de primera necesidad. (Esto es debido a que un familia con tres componentes no gasta el triple, sino menos, que una familia de un único componente). El "Factor Monetario" y el "Factor de Reducción" deberán usarse en el cálculo de los gastos de primera necesidad y los gastos de ocio, para poder calcular correctamente los gastos y el balance de caja.

COHERENCIA

Necesario para localizar las posibles incoherencias que pueda haber en los datos de la base de datos, dos ejemplos claros de incoherencia podrían ser registros de casados o parejas de hecho con componentes de unidad familiar igual a "1", o solicitantes no clientes con la antigüedad de cliente informada

Las incoherencias más habituales que se buscarán se refieren habitualmente a:

-estado de la vivienda y a los gastos vinculados a cada estado

-importe, precio, plazo, cuota, ... variables que obligatoriamente deben estar informadas.

-fecha de solicitud, nacimiento y formalización, que deben seguir una lógica temporal.

-profesión, actividad económica y tipo de contrato laboral no deben ser contradictorios.

-variables como el número de años de residencia o de antigüedad en el empleo deben tener sentido al lado de la variable edad.

La dinámica a seguir para localizar estas incoherencias será la creación de variables de chequeo para cada variable original, lo que irá mostrando aquellos registros que tienen algún dato no coherente y cuáles son correctos. De este modo se le asignará un valor a los registros correctos y otro valor a los incorrectos. Posteriormente se creará una variable que otorgará un valor para cada registro dependiendo si cumple o no todas las coherencias que se hayan definido. Es aconsejable obtener las frecuencias de cada variable de chequeo para poder analizar en cada caso la cantidad del número de errores, y en definitiva, cuantos registros están bien completos y cuales tienen algún error.

Las tablas siguientes resultan de la obtención de esas frecuencias tras la aplicación de este proceso sobre nuestra base de datos:

Cuadro 2: Chequeo de frecuencias de incoherencias en la base de datos

chequeo del error contrato laboral-profesion

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	correcto	13091	100,0	100,0	100,0

chequeo de la vivienda vs gastos de alquiler

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	correcto	12147	92,8	92,8	92,8
	alquiler sin gastos de alquiler	110	,8	,8	93,6
	propia con gastos de alquiler	510	3,9	3,9	97,5
	padres u otros con gastos de alquiler	324	2,5	2,5	100,0
	Total	13091	100,0	100,0	

Se deberán obtener asimismo las frecuencias de las variables originales (no chequeadas, esto es, frecuencia de variables categóricas y continuas tramadas) para conseguir los valores mal informados en la base, y que vendrán definidos como “Perdidos”. Si la variable es categórica, estos valores perdidos se anotarán en la categoría que tenga la probabilidad de mora más parecida (explicado en el punto siguiente “Sintax Depuración”); si la variable es continua habrá que dividir sus valores en tramos (cien tramos suelen ser suficientes) para asignar a uno de esos tramos los valores perdidos. Esto se realizará mediante la ejecución en SPSS de un Crosstab (relación de dos variables, en este caso la variable elegida con la mora), mediante el cual se relacionará cada tramo con la morosidad, obteniendo así la mora de cada tramo, y conformando cada tramo una variable.

DEPURACIÓN

Si las incoherencias representan un porcentaje muy pequeño de registros se podrán eliminar. Se borrarán los registros con alguna incoherencia siempre que se cumplan dos condiciones: que el porcentaje de valores perdidos sea igual o inferior al 2% (porcentaje opcional a elegir) y que no varíe la tasa de mora de la cartera al eliminar esos registros. Asimismo sí alguna de las variables fundamentales (importe, cuota, plazo) no viene informada, se deberán eliminar esos registros. En aquellas variables que estén estimadas en el balance de caja no hará falta eliminar las operaciones con determinados tipos de incoherencia, ya que en el cálculo de dicho balance se asignan valores a esos registros (las incoherencias a las que se refiere esta última anotación son "Situación vivienda alquiler sin gastos de alquiler" y "Situación vivienda hipotecada sin gastos de otras cuotas")

Si representan un porcentaje elevado se intentarán arreglar para no mermar de forma importante el número de registros de la base de datos. Es recomendable evitar la eliminación de un gran número de registros ya que podría terminar por invalidarse la base de datos. Otro método habitual de solucionar registros con incoherencias será la asignación de cada grupo de incoherencias a una categoría en concreto. Esta asignación se realizará por similitud en la probabilidad de mora con una determinada trama o categoría de la variable.

Una vez solucionadas las incoherencias habrá que volver a ejecutar las variables de chequeo y volver a lanzar las frecuencias sobre ellas para comprobar que todo ha quedado en orden.

4. ANÁLISIS BIVARIANTE Y UNIVARIANTE

4.1 Análisis bivariante

El objetivo de este análisis será obtener las agrupaciones o tramas de las variables (dependiendo de si son categóricas o continuas). Se obtendrá asimismo la muestra madura de la que posteriormente nos valdremos para obtener la muestra histórica y la actual del análisis univariante.

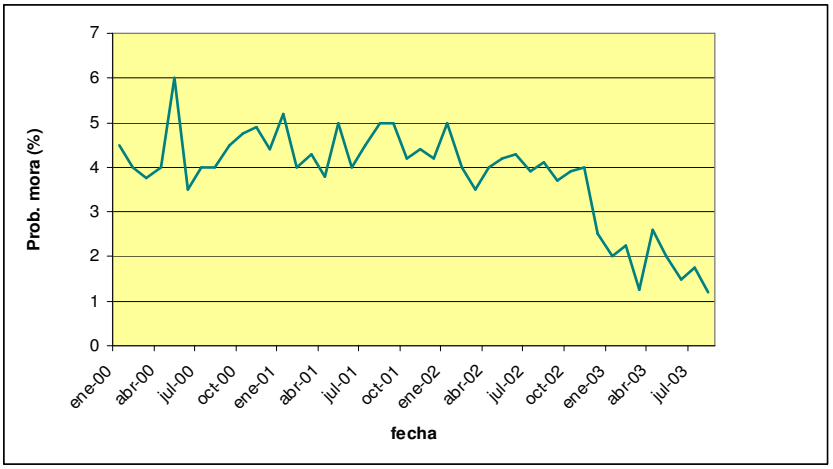
La selección de variables explicativas se hará sobre la base de una teoría económico-financiera sólida, ya que encontrar relaciones estadísticas no significa necesariamente relaciones de causalidad, y por lo tanto estas podrían carecer de capacidad predictiva (Calderón, 2004).

Las variables categóricas serán aquellas que pueden tomar un número finito de valores definidos en categorías, mientras que las continuas serán las variables que pueden tomar un número infinito de valores, por lo que deberán ser tramadas antes de realizar cualquier análisis.

Este análisis se realizará sobre la muestra madura, definiéndose ésta como aquel conjunto de operaciones en las que ha pasado un tiempo que consideramos suficiente para haber podido caer en mora.

Para seleccionar la muestra madura creamos una variable que contenga la fecha de las operaciones formalizadas. Se realiza posteriormente un Crosstab en SPSS y se lleva la tabla de solución a Excel. Allí se obtiene el porcentaje de morosos y no morosos quedando el gráfico siguiente:

Gráfico 2: Relación entre probabilidad de mora y antigüedad de las operaciones



Se observa que para un análisis que se esté realizando en 2003, el porcentaje de morosidad de esta serie se truncaría a partir de noviembre de 2002. Se consigue por tanto definir la muestra madura como las operaciones formalizadas antes de noviembre de 2002, esto significa que se considerará que las operaciones a partir de esa fecha no han pasado todavía un periodo suficientemente largo para estar en plenitud de condiciones de caer en mora La muestra no madura vendrá definida por aquellas operaciones a partir de noviembre de 2002, y hay que tener presente que en la generación de modelos en el análisis bivariante, se trabajará sobre la muestra madura

El objetivo de este análisis es determinar qué relación tiene cada variable de estudio con la mora. Para ello procederemos de la siguiente manera para cada variable independiente:

Dentro del Sintax de SPSS:

```
CROSSTAB
        /Tables = est.civil BY morosid
          /format= avalue tables (←tablas cruzadas)
          /cells= count row column total
```

o bien, dentro del Editor de SPSS:

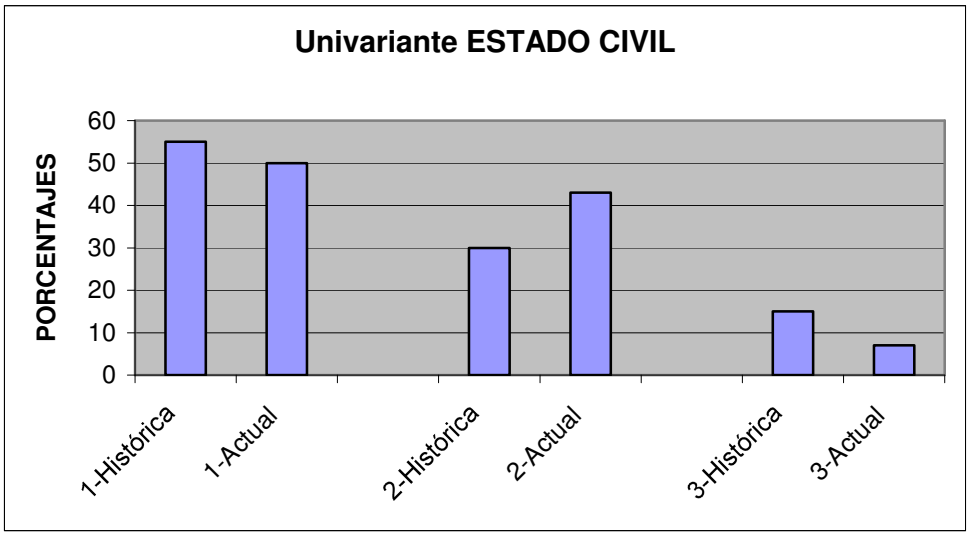
Analizar→Est. Descriptivo→tablas de contingencia

Este análisis permitirá definir las variables que mejor explican el comportamiento de la variable dependiente, esto es, la morosidad. Igualmente se evitará mediante este análisis introducir variables en la regresión que no sean determinantes del comportamiento de la mora.

4.2 Análisis univariante

El análisis univariante permite conocer la estructura de la cartera mediante la obtención de la distribución de frecuencias de todas las variables que componen la base de datos. Los usos que se le pueden dar a este análisis son variados, así puede mostrar variables o categorías con niveles de frecuencia muy bajos, a los que habrá que habrá que prestar especial atención, ya que si bien la probabilidad de mora de esos grupos es la real, puede que al ser un grupo muy pequeño, no sea representativa de la población. En el caso de variables categóricas con muchas categorías se deben agrupar para hacer el análisis (en este ejemplo con la variable estado civil sería correcto agruparla aunque no sería necesario).

Gráfico 3: Análisis univariante de la variable "Estado Civil"



En el Gráfico 3 se separa la variable Estado Civil en:
1-Casados, Parejas de Hecho
2-Solteros, Viudo
3-Separados, Divorciados

El análisis univariante: consiste en comparar dos muestras: la histórica y la actual. Se comparan las frecuencias relativas de las categorías de una variable en la muestra histórica con las frecuencias relativas de la misma en la muestra actual. En caso de variables continuas se procederá de igual modo teniendo en cuenta que se compararán tramas.

La obtención de la muestra histórica parte de la obtención de la fecha donde se trunca la morosidad. Se aplicará esa fecha sobre las operaciones solicitadas, esto es formalizadas y no formalizadas (denegadas, anuladas, ...) para diferenciar entre la muestra histórica y la actual, siendo la que quede anterior a esa fecha la muestra histórica, y la posterior la muestra actual.

Antes de realizar este análisis será necesario seguir los siguientes pasos:
1- Separar la muestra histórica y la muestra actual
2- Agrupar categorías si hace falta (si no hace falta no)
3- Sacar las frecuencias (absoluta y relativa) de esa variable en la muestra histórica sacar las frecuencias de esa variable en la muestra actual (sacar sus porcentajes)
4- Obtener los gráficos correspondientes

La agrupación de variables en categorías se realizará de forma diferente dependiendo si la variable a analizar es continua o categórica. Para poder agrupar variables continuas por similitud en mora se separará por percentiles en numerosas tramas. Este proceso de agrupación se realizará mediante un análisis que relacione cada percentil con la morosidad

Existen tres formas equivalentes de realizarlo en SPSS que se detallan a continuación:

SINTAX SPSS	(SINTAX SPSS	EDITOR SPSS
Frequencies	Freq	Analizar→Est.Descriptivo→frecuencia
Variables=edad	var=edad	
/ntiles=100	/ntiles=100	
/order=analysis	/order=analysis	

Posteriormente se realiza un análisis de morosidad para cada variable y todas las tramas obtenidas y se agrupan en relación a la similitud en la tendencia de la probabilidad de mora (comando "Piece Wise" en SPSS), separándose en grupos que tengan distinta tendencia.

A modo de ejemplo, si se separa la edad en pequeños grupos de un año cada uno:

Tramo	*nº tramo*	*probabilidad de mora*
[16-17)	*1*	*P1*
[17-18)	*2*	*P2*
[18-19)	*3*	*P3*

[72-73)	*57*	*P57*
[73-74)	*58*	*P58*
[74-75)	*59*	*P59*

Se obtiene la relación de cada tramo con la morosidad, observando por ejemplo para el primer tramo, cuántas personas de 16 a 17 años cayeron en mora sobre el total de personas en ese tramo de edad que obtuvieron el crédito (P). Si se diera el caso que hasta la edad 30, los grupos tienen parecida morosidad (tendencia plana); de 30 a 55 años hay una tendencia descendente; y de 55 a 75 existe una tendencia plana, con esta estructura crearíamos tres grupos diferentes:

Grupo 1 [16-30) *Edad1*
Grupo 2 [30-55) →*obteniendo tres variables* *Edad2*
Grupo 3 [55-75) *Edad3*

En el caso de las variables categóricas habrá que realizar previamente un análilis bivariante para poder agrupar correctamente las variables (agrupando las categorías con similar probabilidad de mora). Una vez agrupadas se obtendrán las frecuencias de cada agrupación con la muestra histórica y la actual y se compararán.

5. LANZAMIENTO DE LA REGRESIÓN

La regresión a utilizar en estos modelos será de tipo Logit como aproximación a la regresión Probit (Se evita utilizar esta última por requerir su uso una gran potencia de cálculo).

El objetivo por tanto es construir un modelo de regresión del siguiente tipo:

$$Z_i = \alpha + \beta \times X_i$$

$$P_i = \frac{1}{1+e^{-Z_i}} = \frac{1}{1+e^{-(\alpha+\beta \times X_I)}}$$

Donde Pi es la probabilidad de ser buen o mal pagador. El modelo proporcionará resultados entre 0 y 1, proporcionando directamente la probabilidad de incumplimiento.

Para ilustrar el lanzamiento de esta regresión, se supondrá un modelo extremadamente sencillo en el que tenemos dos variables: Estado Civil y Edad

Suponemos que la variable "Estado Civil" se ha agrupado anteriormente en tres grupos, que funcionan como variables independientes:

Estado1: Casado, Pareja de Hecho
Estado2: Soltero, Viudo
Estado3: Separado, Divorciado

La variable "Edad" también estará agrupada en tramos, que funcionan como variables independientes. Una posible división sería la siguiente:

Edad1: [16-30)
Edad2: [30-55)
Edad3: [55-75)

Así a modo de ejemplo, una persona de 50 años tomaría los siguientes valores para "Edad1", "Edad2" y "Edad3":

Edad1=30-16=14
Edad2=50-30=25
Edad3=0

una persona con edad 25 años:

Edad1=25-16=9
Edad2=0
Edad3=0

y una persona con 70 años:

Edad1=30-16=14

Edad2=55-30=25
Edad3=70-55=15

Para lanzar la regresión logística se necesita conocer la matriz Pi que estará formada por "1" (operaciones que resultaron morosas, esto es con caída en mora a más de 90 días) y por "0" (operaciones no morosas). De esta forma, una puntuación mayor significará una mayor probabilidad de mora. También se necesitarán formar matrices para las variables exógenas. Existirá una matriz por agrupación tanto en categóricas como en continuas (cada agrupación funciona como una variable independiente)

$$Z_i = \beta_0 + \beta_1 \times M_A + \beta_2 \times M_B + \beta_3 \times M_C + \beta_4 \times M_D + \beta_5 \times M_E + \beta_6 \times M_F$$

$$P_i = \frac{1}{1 + e^{-Z_i}}$$

Donde:
$\beta_0, \beta_1, ... \beta_6$: Betas del modelo que se obtienen al lanzar la regresión logística

MA: Matriz que tomará el valor "1" para aquellos registros que sean Casados o Parejas de Hecho, y "0" para aquellos que no lo sean
MB: Matriz que tomará "1" para los Solteros y Viudos, y "0" para el resto
MC: Matriz con "1" para Separados y Divorciados, y "0" para el resto
MD: Matriz de la variable "Edad1". Nos muestra en cada registro el valor de "Edad1"
ME: Matriz de la variable "Edad2". Nos muestra en cada registro el valor de "Edad2"
MF: Matriz de la variable "Edad3". Nos muestra en cada registro el valor de "Edad3"

Antes de lanzar la regresión habrá que realizar una selección aleatoria en la muestra separando entre una parte que servirá para lanzar la regresión, y otra que utilizaremos para validarla

80% → Regresión

20% → Validación

La muestra del 80% será utilizada para lanzar la regresión, obteniendo así los pesos (las betas) del modelo.

Una vez resuelto el modelo (obtenidas las betas) se habrá obtenido una puntuación no calibrada para cada operación. En variables categóricas la beta de la categoría que entrará en el modelo entra como peso directamente al cálculo de la puntuación. Sin embargo en variables continuas, el peso resultará de multiplicar cada beta para cada tramo que le corresponda.

El siguiente ejemplo simplifica el proceso anterior utilizando dos variables:

Estado Civil (categórica) y Edad (continua)

Una vez lanzada la regresión logística obtenemos los siguientes resultados:
B0= 0,3

Estado civil: Soltero, Viudo:	0	Edad: 16-30: 0
Casado, Pareja de Hecho:	-0,15	30-55: -0,0055 x (Edad-30)
Separado, Divorciado:	0,12	55-75: 0,1375

Ahora supongamos que un solicitante casado y con 42 años nos pide un préstamo y queremos calcular su puntuación mediante este modelo. Procederemos de la siguiente manera:

Ptos = 0.3 - 0.15 +(30-16) x 0 – 0.0055 x (42- 30) = 0,084

Esta puntuación sería una probabilidad de mora (un 52,09 %)

$$Puntuación = \frac{1}{1+e^{-Ptos}} = \frac{1}{1+e^{-0,084}} = 0,52098$$

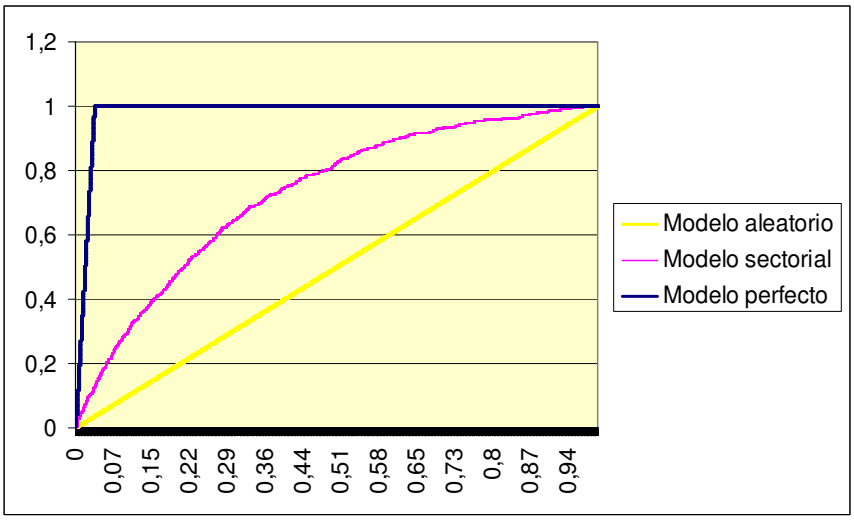
6. EFICIENCIA DEL MODELO

Una vez obtenido el modelo nos deberíamos preguntar cuál es el grado de eficiencia alcanzado por el mismo y si los estudios realizados con dicho modelo pueden ser fiables. Para responder a estas preguntas se recurrirá al Estadístico del Poder (Power Stats) aunque otras medidas como el Índice de Gini o la R^2 nos podrían dar también una idea clara de la validez de los resultados obtenidos.

El Estadístico del Poder permite conocer el grado de ajuste del modelo a través de la relación entre dos áreas: la que va desde la curva del modelo perfecto hasta la curva del modelo aleatorio (área 2); y la que va desde la curva de nuestro modelo (le llamaremos sectorial) hasta la curva del aleatorio (área 1). Se calculará como el cociente del área 1 entre el área 2.

El valor de este estadístico siempre estará entre 0 y 1. Valores próximos a 0 significan que nuestro modelo dista mucho de parecerse al modelo perfecto, mientras que valores cercanos a 1 significan que el modelo realizado es muy parecido al hipotético modelo perfecto.

Gráfico 4: Grado de ajuste de los diferentes modelos



El eje Y representa el porcentaje de morosos reales de nuestra cartera. Mientras tanto, el eje X muestra el porcentaje de observaciones acumuladas ordenadas de mayor a menor puntuación (probabilidad de mora no calibrada). Si nuestro modelo fuese muy fiable, las operaciones que han resultado morosas deberían coincidir con las operaciones que nuestro modelo les ha asignado una mayor probabilidad de mora (las que están al principio del eje X). Eso es precisamente lo que ocurre en el modelo perfecto: si la muestra de estudio tiene un porcentaje de morosos de un 2%, sólo deberían aparecer morosos en el primer 2% de los casos (en el 2% de la cartera con más probabilidad de mora según nuestro modelo).

Modelo Aleatorio: Ajusta el porcentaje de operaciones en mora al porcentaje de operaciones observadas, esto es, ante una observación de un 20% de las operaciones, el modelo habrá registrado al 20% de los morosos. En definitiva, es una relación lineal que no tiene en cuenta la probabilidad de mora para cada registro estimada por el modelo.

Modelo Perfecto: Ajusta perfectamente el 100% de las operaciones realmente morosas con las operaciones a las que nuestro modelo ha asignado una mayor probabilidad de mora.

Modelo Sectorial (el realizado): Cuanto más se aproxime este modelo al modelo perfecto será un síntoma claro de que es un modelo que estima de forma fiable la probabilidad de mora. De esta manera, las operaciones con una mayor probabilidad de mora asignada serían las que realmente resultan morosas a posteriori.

Antes de empezar a calcular este estadístico es interesante excluir las operaciones denegadas de la base de datos a utilizar para este análisis de eficiencia. La inclusión de denegadas no es recomendable ya que estas son operaciones con una puntuación muy alta pero que no han resultado morosas (precisamente porque no se han formalizado), por lo que el estadístico obtenido con operaciones formalizadas y no formalizadas no sería fiable.

El primer paso para calcular este estadístico será ordenar todos los registros de la muestra que hemos utilizado, de mayor a menor probabilidad de mora, indicando en una columna adyacente si esa operación ha sido morosa (1) o no morosa (0). A continuación se crea una columna con el número de observaciones acumuladas y otra con el porcentaje de observaciones acumuladas que hay en la muestra. Ahora se crea una columna (los registros siguen ordenados como en el primer paso de este apartado) que nos vaya marcando los registros morosos acumulados en la muestra. A su lado calculamos el porcentaje acumulado de registros morosos. Se estaría ya en condiciones de calcular el área entre la curva del modelo sectorial y la del modelo aleatorio. El cálculo será por aproximación: se calculará la diferencia en cada registro como el porcentaje de observaciones acumuladas menos el porcentaje de operaciones en mora acumuladas. Una vez realizadas las restas, se realizará el sumatorio de todas ellas consiguiendo así una aproximación al área que hay entre la curva de modelo sectorial y la de modelo aleatorio (área 1).

El siguiente paso será calcular el área entre la curva del modelo perfecto y la del modelo aleatorio (área 2). Para ello se crea una columna para el modelo perfecto en la que todos los registros morosos estarán concentrados al principio del estudio (relacionados con los registros con mayor probabilidad de mora). Así cuando se lleve observado un porcentaje de casos igual al porcentaje de mora de la cartera, estarán ya definidos el 100% de los morosos de nuestra cartera

El área entre la curva del modelo perfecto y la del modelo aleatorio (área 2) se aproximará como el sumatorio de las diferencias entre la columna calculada en el punto 5 y la columna de % observaciones acumuladas. El Estadístico del Poder resultará del cociente entre el área 1 y el área 2, siendo siempre un número entre 0 y 1, ya que el área 2 será siempre por lo menos tan grande como el área 1. Consideraremos que el modelo es fiable será aquel con Power por encima del 40%, aunque este criterio dependerá del nivel de exigencia que se tenga frente al modelo.

Ejemplo de obtención del Estadístico del Poder en una hipotética cartera con 30 operaciones y una probabilidad de mora del 10%.

Cuadro 3: Obtención del Estadístico del Poder

morosos	3										
total	30	Tasa de mora muestra	10,00%						Estadístico Poder	65,43%	
	MUESTRA				Sectorial	Perfecta		sumatorios	883,33%	1350,00%	
Morosidad	Probabilidad estimada	Observaciones acumul	en %		malos acum	%malos acum.			sumas para area 1	sumas para area 2	
0	0,32469	1	0,033		0	0	0,333		-3,33%	0,3000	
0	0,32123	2	0,067		0	0	0,667		-6,67%	0,6000	
0	0,31798	3	0,100		0	0	1		-10,00%	0,9000	
1	0,31652	4	0,133		1	0,333	1		20,00%	0,8667	
1	0,31647	5	0,167		2	0,667	1		50,00%	0,8333	
0	0,31299	6	0,200		2	0,667	1		46,67%	0,8000	
0	0,31032	7	0,233		2	0,667	1		43,33%	0,7667	
0	0,30977	8	0,267		2	0,667	1		40,00%	0,7333	
0	0,30072	9	0,300		2	0,667	1		36,67%	0,7000	
0	0,29296	10	0,333		2	0,667	1		33,33%	0,6667	
1	0,29059	11	0,367		3	1	1		63,33%	0,6333	
0	0,28997	12	0,400		3	1	1		60,00%	0,6000	
0	0,28986	13	0,433		3	1	1		56,67%	0,5667	
0	0,28862	14	0,467		3	1	1		53,33%	0,5333	
0	0,27746	15	0,500		3	1	1		50,00%	0,5000	
0	0,27502	16	0,533		3	1	1		46,67%	0,4667	
0	0,27309	17	0,567		3	1	1		43,33%	0,4333	
0	0,27102	18	0,600		3	1	1		40,00%	0,4000	
0	0,26805	19	0,633		3	1	1		36,67%	0,3667	
0	0,26331	20	0,667		3	1	1		33,33%	0,3333	
0	0,26259	21	0,700		3	1	1		30,00%	0,3000	
0	0,2609	22	0,733		3	1	1		26,67%	0,2667	
0	0,2609	23	0,767		3	1	1		23,33%	0,2333	
0	0,2609	24	0,800		3	1	1		20,00%	0,2000	
0	0,2597	25	0,833		3	1	1		16,67%	0,1667	
0	0,25945	26	0,867		3	1	1		13,33%	0,1333	
0	0,25909	27	0,900		3	1	1		10,00%	0,1000	
0	0,25899	28	0,933		3	1	1		6,67%	0,0667	
0	0,25824	29	0,967		3	1	1		3,33%	0,0333	
0	0,25586	30	1		3	1	1				

La morosidad de la cartera es 10%. En la perfecta cuando llevemos un 10% de observaciones llevaremos el 100% de morosos.

7. CALIBRACIÓN

Una vez lanzada la regresión y obtenidos los pesos para cada categoría y trama, la obtención de la puntuación para cada registro es prácticamente inmediata. Esta puntuación es una probabilidad de mora, pero no es la Probabilidad de Mora que se está pidiendo desde Basilea. Esta última se define como la probabilidad de morosidad calibrada a un año. Por tanto, el paso posterior a la obtención de la probabilidad de mora sin calibrar (puntuación) será calibrarla. Para ello lo primero que habrá que hacer será segmentar la base, para luego calcular para cada segmento la probabilidad de mora a un año a partir de la información disponible.

Un posible criterio de segmentación sería el siguiente (partiendo de un análisis sobre una cartera de créditos al consumo):

1-vinculación del solicitante a la entidad
2-puntuación
3-plazo

VINCULACIÓN:	muy vinculados					poco vinculados				
PUNTUACIÓN:	0-10	10-30	30-50	50-70	70-100	0-10	10-30	30-50	50-70	70-100
PLAZO:	c	l	c	l	c	l	c	l	c	l	c	l	c	l	c	l	c	l	c	l
EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0	EDF0
EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1	EDF1
EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2	EDF2

EDF 0= (Nº morosos en i) / (Nº operaciones vivas i→i+12 – Morosas anteriores en i)

Donde:
i: fecha de referencia
i+12: fecha de observación
Operación morosa: Si la fecha de entrada en mora cae entre la fecha de referencia y la fecha de observación
Operación viva: Si la fecha de vencimiento es mayor que la fecha de referencia
Operación morosa anterior: Si la fecha de entrada en mora es anterior a la fecha de referencia

EDF 1= (Nº morosos en i+1) / (Nº operaciones vivas i+1→i+13 – Morosas anteriores en i+1)

En este caso i+1 sería la fecha de referencia, y i+13 la fecha de formalización.

EDF 2=(Nº morosos en i+2) / (Nº operaciones vivas i+2→i+14 - Morosas anteriores en i+2)

...

```
_I___I_____________________________________________I___I_________
 i  i+1                                           i+12  i+13
```

Es importante observar que las operaciones morosas anteriores se restan en el denominador a las operaciones vivas, ya que lo que se calcula es la probabilidad de primera entrada en mora a

un año. Además no se considerarán en el cálculo aquellas operaciones con madurez inferior a un año.

Para cada una de las 20 salidas anteriores, se calcularía su probabilidad de mora empírica, EDF, media (que es la probabilidad de mora media histórica de solicitantes con las mismas características que el solicitante que nos ocupa en cada caso)

$$\frac{EDF0 + EDF1 + ... + EDFn}{n}$$

De este modo se obtienen las medias de las probabilidades de mora empíricas para cada segmento y esa será la PD que se está buscando para cada caso. De esta manera al obtener una puntuación para un solicitante concreto, dependiendo de en que segmento se encuentre se le asignará a una PD determinada.

Al final se está asignando una probabilidad de mora teniendo en cuenta el producto, la puntuación y el plazo (en este ejemplo).

Sin embargo, en el ejemplo de más arriba, haciendo el supuesto de dos solicitantes con préstamos de similares características, uno solicitando un préstamo al consumo a corto plazo y con una puntuación de 10, y el otro también con un préstamo al consumo a corto plazo y con una puntuación igual a 29, se llega a una misma probabilidad de mora para los dos. Para evitar esto, sería interesante ajustar mediante una función la nube de puntos que nos relaciona la puntuación con la probabilidad de mora, en préstamos al consumo a corto plazo. Esta función debería de ser de forma creciente y continua, por lo que acabaríamos asignando dos PD distintas para dos solicitantes distintos.

8. AJUSTE AL CICLO

Los resultados que resulten de la calibración reflejan el momento del ciclo económico en que se han recogido esas probabilidades de mora empíricas. Esto es, si las probabilidades de mora están tomadas en la parte baja del ciclo, esas probabilidades estarán sesgadas, y serán superiores a las que se habrían obtenido en caso de recogerlas durante un ciclo completo. Asimismo, si se utiliza un periodo de bonanza al calcular las probabilidades de mora empíricas, la morosidad será mínima y estaríamos subestimando la PD frente al caso de prolongar los resultados en ese periodo al ciclo completo. En este caso las dotaciones y reservas resultarían insuficientes.

Por todo esto se busca que las dotaciones sean acíclicas, esto es, que tengan en cuenta el impacto medio del ciclo. Podría también utilizarse el modelo para realizar simulaciones de nuestra cartera bajo diferentes escenarios futuros, lo que nos acercaría al cálculo de las dotaciones contracíclicas, caso que la entidad quisiera cubrirse ante la previsión de una recesión.

Es por esto que Basilea pide una PD aciclíca, que no tenga en cuenta el momento del ciclo en que nos encontramos. Para ello se ajustará un modelo macroeconómico con datos reales para obtener la relación de las distintas variables macroeconómicas con la PD. A modo de ejemplo podrían incluirse las siguientes variables: Tipos de interés, paro, crecimiento del PIB y la inflación.

Habrá que retirar del modelo las variables no significativas así como aquellas que estén fuertemente correlacionadas. El objetivo es obtener la relación existente de las caídas en mora con las variables elegidas (resumidas en "puntuación" y "maduración") y con las variables macroeconómicas.

$$Y_T = \beta_0 + \beta_1 \times Puntuación + \beta_2 \times Maduración + \beta_3 \times Variación_PIB + \beta_4 \times Variación_Tasa_Paro + \beta_5 \times Variación_Euribor1año + \beta_6 \times Variación_Inflación$$

De esta forma obtenemos las betas para cada variable, lo que nos permitirá según estimaciones macroeconómicas, calcular probabilidades de mora para la parte del ciclo que no

$$\Pr ob.Mora = \frac{1}{1+e^{-Y_T}}$$

conocemos. Para terminar este proceso de ajuste al ciclo podría realizarse una media entre la

Probabilidad de Mora real conocida de una parte del ciclo, y la probabilidad de mora estimada para la parte del ciclo que desconocemos.

9. FRONTERAS Y POLÍTICAS DE DECISIÓN

Una vez calculada la PD para una operación, deberá determinarse si dicha operación será aprobada o denegada. Para ello nos valdremos de los dictámenes. A continuación se expone una propuesta de dictamen de operaciones (aunque se podría realizar con otras variables a las elegidas aquí)

-Dictamen de Puntuación: Relacionado con la probabilidad de mora sin calibrar (puntuación)

-Dictamen de Capacidad: Relacionado con la capacidad de pago

-Dictamen Final: Relacionado con los dos dictámenes anteriores y con una serie de reglas que definiré posteriormente.

Estos tres dictámenes se ejecutarán en cascada, de arriba a abajo, no pudiendo nunca ser un dictamen más favorable que el anterior.

A través de los dos primeros dictámenes se llega al Dictamen Final que será el que indique si el dictamen del Scoring será Aprobar (AP), Duda (DU) o Denegar (DE).

El fondo de esta metodología radica en la importancia que se le da a la puntuación (probabilidad de mora) a la hora de dictaminar una determinada operación. Si el solicitante tiene una alta probabilidad de mora (Dictamen Puntuación = Denegar), le será indiferente tener un balance de caja elevado, ya que ese primer dictamen criba ya a los siguientes, y el Dictamen Final será Denegar.

DICTAMEN DE PUNTUACIÓN

En primer lugar deben definirse las fronteras de denegación, aprobación y duda, así como las políticas que seguiremos a partir de estas fronteras.

Las fronteras y políticas definidas servirán para determinar si la puntuación obtenida por ese cliente da lugar a un Dictamen de Puntuación igual a Aprobar, a Duda o a Denegar.

Así, por ejemplo, suponiendo la determinación de unas fronteras en 4 y en 8, y unas políticas por las que puntuaciones por encima de 8 darán un dictamen de Denegar; por debajo de cuatro será Aprobar; y entre 4 y 8 un dictamen de Duda. Según la puntuación que obtengamos tendremos un determinado Dictamen de Puntuación.

DICTAMEN DE CAPACIDAD DE PAGO

Habrá que establecer, como en el caso anterior, unas fronteras y unas políticas concretas. Este segundo dictamen dependerá de la capacidad de pago del solicitante, que vendrá determinada a partir del Ratio de Balance de Caja, definido como las veces que se podrá pagar la cuota del préstamo solicitado con el balance de caja del solicitante.

$$Ratio_Balance_Caja = \frac{Balance_Caja}{Cuota}$$

El Dictamen de Capacidad resultará de cruzar el Dictamen de Puntuación con la decisión por Capacidad de Pago:

Intención de Pago					
		AP	DU	DE	
Capacidad de Pago	AP	**AP**	**DU**	**DE**	**DICTAMEN DE CAPACIDAD DE PAGO**
	DU	**DU**	**DU**	**DE**	
	DE	**DE**	**DE**	**DE**	

Se observa que el Dictamen por Capacidad de Pago será siempre o igual o más restrictivo que el Dictamen de Puntuación. Este hecho se repetirá con el Dictamen Final respecto del Dictamen de Capacidad de Pago.

DICTAMEN FINAL

El Dictamen Final consiste en un "filtro" que penalizará el dictamen anterior en determinadas ocasiones. Estas reglas se fijarán por la observación de ciertos comportamientos de pagos que no vengan recogidos en el modelo. Por tanto podrán variar de un modelo de scoring a otro.

A continuación se citan a modo de ejemplo algunas reglas que se podrían seguir en nuestro modelo:

Si profesión es:
Parado y Dictamen de Capacidad = DU→ Dictamen Final = DE
Estudiante y Dictamen Final = AP → Dictamen Final = DU

Se edad + plazo es:
>75 años y Dictamen de Capacidad = AP → Dictamen Final = DU

De esta forma obtendremos el dictamen del Scoring que según la entidad, será o no vinculante aunque siempre se presentará como un modelo consultivo (en cualquier caso si el dictamen final es de duda la decisión quedará en manos del analista de riesgos).

Una vez completado el modelo hay que tener en cuenta la posibilidad de que para una misma operación existan varios titulares y/o avalistas. La puntuación y la PD que se le asignará a la operación deberá ser siempre la del mejor titular. Esa PD será la que se utilizará posteriormente para que la entidad calcule los requerimientos de capital para esa operación.

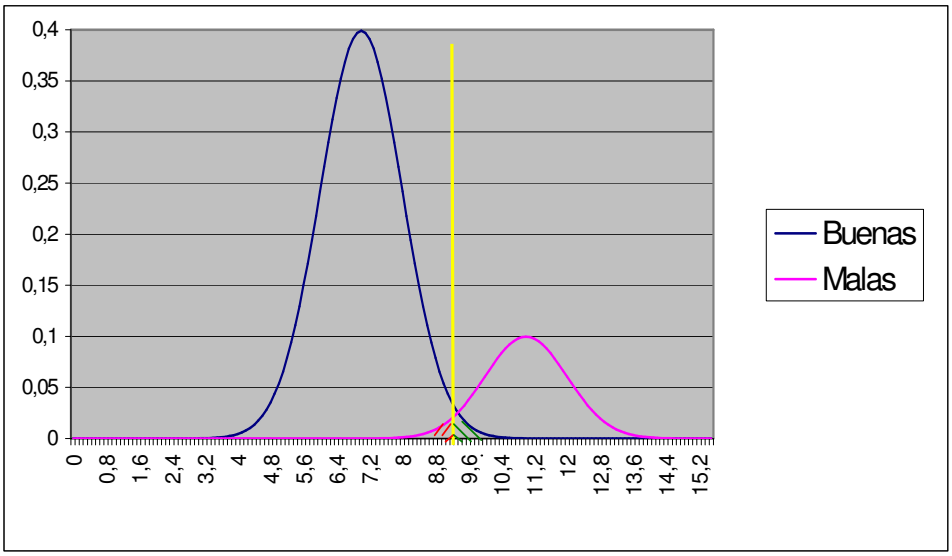
Es habitual también la existencia de reglas para casos especiales como que el titular sea menor de 25 años con avalistas y el dictamen sea denegar. En estos casos si el avalista es el padre, la madre o el hermano, y ese avalista tiene un dictamen Aprobar, influirá muy positivamente en el dictamen final de la operación. Este tipo de reglas serán habituales para evitar denegar operaciones a determinados segmentos de la población, que si bien tendrían perfil negativo de pagador, en realidad pocas veces resultan morosos.

Sin embargo, ¿cómo decide una entidad de crédito las fronteras sobre las que desea aplicar esos dictámenes? Aquí habría que realizar un análisis para cada caso ya que todo dependerá de los objetivos particulares de cada entidad. Así, una entidad que esté asentada en una región en la que posea una alta cuota de mercado, estará interesará en marcar unas fronteras de aprobación más bien restrictivas ya que querrá asegurarse que un gran número de sus operaciones aprobadas finalmente no resulten morosas (sin preocuparle en exceso denegar algunas operaciones que finalmente no hubieran caído en mora).

Por otro lado, una entidad con interés en entrar en una región en la que tenga un competidor ya asentado, puede querer llevar una política más agresiva y aprobar operaciones que la otra deniega, ya que estará dispuesta a asumir que aprobará ciertas operaciones que entrarán en mora a cambio de no denegar operaciones que finalmente no hubieran resultado morosas.

Como se puede ver en el Gráfico 5, esta elección se realizará sobre distribuciones de probabilidad de clientes que han terminado resultando morosos como de los que no lo han hecho. Así cada entidad estará dispuesta a asumir un porcentaje de error, y marcará el percentil que más se ajuste a sus objetivos, definiendo de esta forma una puntuación sobre la que definirá sus políticas de aceptación, duda o denegación.

Gráfico 5: Distribución de probabilidad de operaciones pagadas e impagadas



10. CONCLUSIONES

La generación de modelos de riesgo de crédito es una actividad que se ha extendido ampliamente durante los últimos años entre las entidades financieras debido en gran medida a la necesidad de un sector financiero estable por parte de toda la sociedad. Este tipo de modelos permite a las entidades homogeneizar sus decisiones de concesión de financiación, ayudándolas a alcanzar una mayor eficiencia en la concesión del crédito y en la gestión del capital.

Desde el punto de vista del conocimiento, este artículo demuestra que la generación de estos modelos resulta sencilla e intuitiva, lo que los hace accesibles a cualquier persona o empresa interesada en ellos. Aunque la aplicación de la metodología Logit presenta ciertas carencias frente a la Probit, ha sido y es por su sencillez y buenos resultados, la más utilizada en la generación de modelos de riesgo de crédito.

Tecnológicamente, la potencia de cálculo requerida y el software necesario es asimismo muy accesible, siendo la mayor complicación la obtención de una base de datos suficientemente amplia y con una profundidad histórica adecuada. Aquí es donde las entidades tendrán que dirigir sus mayores esfuerzos ya que la obtención de una base fiable acortará de manera importante el tiempo de generación del modelo, además de incrementar su eficacia.

BIBLIOGRAFÍA

ALTMAN E.I., (1968), "Financial Ratios, Discriminant Analisys and the Prediction of Corporate Bankruptcy", The Journal of Finance Vol. XXIII nº4

ALTMAN E.I. (2000), "Predicting Financial Distress of Companies: Revisiting the Z-Score and Zeta", Unpublished, NY University, Nueva York

ALTMAN E.I. Y HALDEMAN R. (1995), "Corporate Credit Scoring Models: Approaches and Standards for Succesful Implementation", Working Paper Series S-95-6, NY University, Nueva York

CALDERÓN, M. (2004), "La Gestión Bancaria en los Riesgos de Crédito. Una aplicación al Caso Peruano para Empresas y Corporaciones", Colección Tesis, Fondo de Desarrollo Editorial, Universidad de Lima, Lima

GREENE W.A., (1992), "Statistical Model for Credit Scoring", NY University Working Paper No. EC-92-29, Nueva York

KNOP R., ORDOVÁS R. Y VIDAL J., (2004), "Medición de Riesgos de Mercado y Crédito", Editorial Ariel, pp. 189-191, 202-206

MAGDALA, G.S. (1993), "Limited-Dependent and Qualitative Variables in Econometrics", Econometric Society of Monographs 3, Cambridge University Press, Cambridge

MODELAÇÃO LINEAR DA PROCURA TURÍSTICA EM PORTUGAL

TELMO NUNO MARTINS MACHADO
E-mail: telmo.nuno.machado@gmail.com
JOÃO PAULO RAMOS TEIXEIRA
E-mail: joaopt@ipb.pt
PAULA ODETE FERNANDES
E-mail: pof@ipb.pt
Instituto Politécnico de Bragança (IPB)

Resumo

Nos últimos anos, dada a relevância do sector do turismo para a economia de Portugal, a modelação e previsão de séries económicas relacionadas com o turismo têm tido um crescente interesse. Assim, o principal objectivo do presente trabalho assenta no desenvolvimento de um modelo univariado que explicitamente considere as características peculiares que as séries de turismo evidenciam, tais como, a sazonalidade e a tendência. Recorreu-se para o efeito à série mensal que mede a procura turística: "Dormidas Mensais registadas em Portugal", para o período compreendido entre Janeiro de 1990 e Dezembro de 2008.
Os modelos lineares desenvolvidos apresentaram qualidades estatística e de ajustamento satisfatórias, pelo que se procedeu à sua utilização para efeitos previsionais. Neste contexto, efectuou-se uma comparação dos resultados previstos com os reais, para os anos de 2007 e 2008, apresentando, para o período em análise, um EPAM de 4.2% e um coeficiente de correlação de 0,988.

Palavras-chave: Procura Turística; Séries Temporais; Modelação Linear; Previsão.
Área Temática: Modelação Económica e Financeira.

Abstract

The modulation and forecast of economic data series related with tourism showed an increase interesting, in last years, due to the relevance of the tourism sector for the Portuguese economy. Therefore, the main goal of the present work consists in the development of a model that explicitly considers the peculiar characteristics that the tourism data series evidence, such as seasonality and tendency. For these purpose, the monthly data series that measures the touristic search was used: "Monthly Guest Nights in the Hotels of Portugal", for the period between January 1990 and December 2008.
The developed linear models showed satisfactory statistical qualities of fitting. Thus the model was used for forecast purpose. Using the months of 2007 and 2008 to measure the prediction ability, the model reached a performance measured by the linear correlation coefficient between predicted and target data of 0.988, and an MAPE of 4.2%.

Key words: Tourism Demand; Time Series; Linear Modelling; Forecasting.
Thematic Area: Economic and Financial Modelling.

1. INTRODUÇÃO

O turismo sendo um factor estruturante da dinâmica da economia de qualquer país, se for implementado de uma forma sustentada pode vir a proporcionar benefícios a longo prazo, reflectindo-se na criação de postos de trabalho e na captação de divisas para os diferentes sectores da actividade económica a que está ligado.

À semelhança de outros países, no decurso das últimas décadas, também em Portugal o turismo se transformou num fenómeno singular apesar da situação económica e política lhe

serem, muitas vezes, desfavoráveis. Dados estatísticos de 2007 (INE, 2008; Turismo de Portugal, 2009), indicam que o turismo contribuiu com 10.4% do PIB (para Portugal), em 2007, e a proporção do Valor Acrescentado gerado pelo turismo no PIB da economia portuguesa foi de 4.6% em 2005, 4.9% em 2006 e 5.1% em 2007. Ainda, a procura turística, medida pelo consumo, ultrapassou os 15 mil milhões de euros, em 2006, mais 1,3 mil milhões que no ano anterior e no período entre 2000 e 2006, o consumo turístico teve um crescimento médio anual de 3.7%, correspondendo no final do período a mais 3 mil milhões de euros, comparativamente com o ano de 2000. Em 2007, o crescimento situou-se nos 11.4%, ano em que o consumo turístico ultrapassou 17 mil milhões de euros.

Neste sentido e dado o crescimento substancial deste sector de actividade em Portugal, será de todo proveitoso desenvolver modelos susceptíveis de serem utilizados para efectuar previsões fiáveis da procura turística, dado que esta assume um papel relevante no processo de planeamento e de tomada de decisões tanto no contexto do sector público como do privado. Actualmente, na área da previsão, encontra-se disponível uma grande multiplicidade de métodos que têm vindo a emergir para fazer face às mais variadas situações, apresentando características e metodologias diferentes, que vão desde as abordagens mais simples às mais complexas (Fernandes, 2005).

Assim, surgindo no contexto de um estudo mais abrangente sobre um sistema de previsão de séries temporais com redes neuronais artificiais (RNA), o propósito do presente trabalho prende-se com o desenvolvimento de um modelo de regressão linear univariado mensal, no sentido de se querer estudar a relação entre duas variáveis. Para tal, e para estimar os parâmetros do modelo, utilizou-se o método dos mínimos quadrados ordinários. A série temporal, de turismo, que serviu de referência ao estudo foi: "Dormidas Mensais registadas em Portugal". Os dados observados reportam-se ao período compreendido entre Janeiro de 1990 e Dezembro de 2008, correspondendo assim a um total 228 observações mensais, 19 para cada mês.

A presente estudo encontra-se estruturado da seguinte forma: a par desta introdução; apresenta-se uma secção com a modelação linear da procura turística em Portugal, onde se analisa a série temporal "Dormidas Mensais registadas em Portugal", descreve-se o modelo de regressão linear univariado e todo o processo que esteve na base da sua construção, apresenta-se ainda a previsão da procura turística, com base no modelo desenhado previamente, para os anos 2007 e 2008, analisando-se a *performance* do mesmo; por último, serão apresentadas as conclusões e desenvolvimentos futuros.

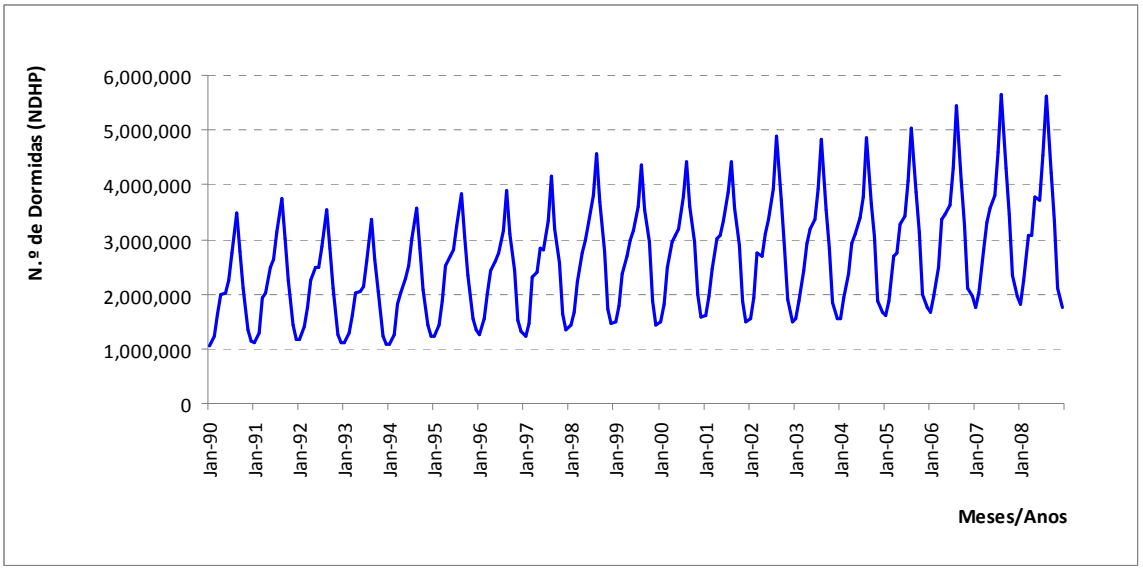
2. MODELAÇÃO LINEAR DA PROCURA TURÍSTICA EM PORTUGAL

2.1. Apresentação e análise da série

Na Figura 1, pode observar-se a série temporal "Dormidas registadas em Portugal" (NDHP), cujos dados estão disponíveis, no EUROSAT (2009). Os dados recolhidos consideram o período compreendido entre Janeiro de 1990 e Dezembro de 2008, correspondendo a 228 dados mensais ao longo de 19 anos (Tabela A.1, do Anexo I).

Optou-se por estudar esta série uma vez que tem sido considerada como significativa da actividade turística devido às suas características e contempla os visitantes que usufruíram das capacidades turísticas de Portugal, nesse período. Analisando o seu comportamento verifica-se a existência de oscilações periódicas em torno de um valor médio crescente, sugerindo a presença de sazonalidade (valores máximos nos meses de Verão e mínimos nos meses de Inverno), isto é, os valores das dormidas vêm em função da altura do ano.

Figura 1. Dormidas dos turistas nas Unidades de Alojamento, em Portugal, no intervalo [Jan-90:Dez-08].



A Figura 1 mostra que o turismo em Portugal, em particular a procura de estabelecimentos hoteleiros, segue um comportamento bastante típico, evidenciando para cada ano, um pico no mês de Agosto, seguido de uma queda abrupta até Janeiro e novamente uma subida na direcção dos meses de Verão. Esta característica, conhecida como sazonalidade, resulta da associação que existe entre a procura turística e as condições climáticas que a favorecem. Segundo Fúster (1989), a sazonalidade é uma das características do fenómeno turístico. Butler (2001) refere que a sazonalidade pode apresentar-se sob duas formas: a natural e a institucional. Relativamente à primeira, pode resultar de alterações nas condições climáticas; já a segunda, pode resultar de aspectos relacionados com as instituições sociais, religiosas, éticas e culturais, sendo menos previsível que a sazonalidade natural.

Também se pode constatar, que existe uma tendência crescente ao longo do intervalo de tempo abrangido neste estudo, sendo esta manifestamente mais evidente em meses com maior valor absoluto de dormidas. A variação é mais acentuada para os valores do índice, relativos ao mês de Agosto, se compararmos com a evolução do mês de Janeiro ao longo dos anos.

A tendência registada é resultado do crescimento económico e dos investimentos no sector turístico, que se têm verificado em Portugal nos últimos anos. Verifica-se, no entanto, que esta tendência não é tão linear quanto isso. O início da década de noventa apresenta uma tendência decrescente, que se prolonga até o ano de 1994. O mesmo aconteceu na transição do ano 1998 para 2000. Estas tendências decrescentes coincidem com períodos de recessão ou abrandamento económicos.

Apesar disso, consegue-se perceber que existe uma relação directa positiva entre a variável independente, o tempo, e a série temporal NDHP, que sugere a aplicação de métodos lineares, nomeadamente a regressão, para efeitos previsionais.

2.2. Modelo univariado

De forma a prever a evolução do valor mensal de dormidas, precedeu-se à determinação das rectas que definem a tendência para cada mês individualmente. Assim sendo, a actual variável independente n, será alvo de uma transformação, passando a ser definida no intervalo $[1:i]$, com n a representar o número de observações e i o número de anos.

Considere-se, S_n, o vector que contém os valores da série representada na Figura 1. A partir de S_n, define-se a matriz $Y_{i,j}$:

$$Y_{i,j} = \begin{bmatrix} S_1 & S_2 & \cdots & S_j \\ S_{j+1} & S_{j+2} & \cdots & S_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ S_{(i-1)j+1} & S_{(i-1)j+2} & \cdots & S_{ij} \end{bmatrix} \quad [1]$$

Onde, *j* representa o número de meses, e varia entre 1 - Janeiro e 12 - Dezembro

Para cada coluna da matriz de $Y_{i,j}$, são calculados os coeficientes de regressão linear simples utilizando para a sua estimação o método dos mínimos quadrados ordinários[1]. Pretende-se assim minimizar a soma do quadrado dos erros observados de forma a obter as equações do tipo (Johnston & Dinardo, 2000):

$$Y_j = \beta_0 + \beta_1 X_j + \varepsilon_j \quad [2]$$

Onde, β_0, representa a ordenada na origem e β_1, o declive da recta. Sendo, X_j, a variável independente *n* para cada mês *j*, a variar no intervalo $[1:i]$ e o termo ε, os erros ou resíduos do modelo.

À semelhança do que acontece com outros modelos, também o modelo de regressão tem por base alguns pressupostos, sabendo-se que os principais pressupostos[2] recaem sobre a componente do erro do modelo, ε_j, e a validação é realizada através dos resíduos (Gujarati, 2003).

O modelo de regressão até aqui definido, será aplicado ao intervalo [Jan-90:Dez-06], de forma a reservar os dois últimos anos para efectuar a previsão e validar o modelo. Desta forma as rectas de regressão Y_j, serão definidas para X_j a variar no intervalo [1:17]. Uma vez definidas as rectas de regressão (Tabela 1), procede-se à reconstrução do restante intervalo, calculando a projecção de Y_j, para $X_j = [18{:}19]$. Ainda de referir que os valores obtidos para os parâmetros dos modelos são significativamente diferentes de zero.

Tabela 1. Equações das rectas de regressão linear, mensais; valor da estatística *t-Student* e respectivo valor de prova, para cada coeficiente; coeficiente de correlação.

Y_j	β_0 *[t-Student;valor de prova]*	β_1 *[t-Student;valor de prova]*	$\beta_0+\beta_1 X_j$	r
Y_{Jan}	1 008 736 *[31,679; 0,0000]*	40 017 *[12,877; 0,0000]*	1 008 736+40 017 *X*	0.95763
Y_{Fev}	1 165 255 *[24,584; 0,0000]*	51 619 *[11,159; 0,0000]*	1 165 255+51 619 *X*	0.94472
Y_{Mar}	1 607 897 *[19,029; 0,0000]*	64 758 *[7,854; 0,0000]*	1 607 897+64 758 *X*	0.89687
Y_{Abr}	1 901 811 *[20,729; 0,0000]*	74 502 *[8,321; 0,0000]*	1 901 811+74 502 *X*	0.90660
Y_{Mai}	2 106 436 *[25,259; 0,0000]*	77 885 *[9,5701; 0,0000]*	2 106 436+77 885 *X*	0.26968
Y_{Jun}	2 217 542 *[25,311; 0,0000]*	84 188 *[9,847; 0,0000]*	2 217 542+84 188 *X*	0.93060
Y_{Jul}	2 641 517 *[28,347; 0,0000]*	93 323 *[10,262; 0,0000]*	2 641 517+93 323 *X*	0.93559

[1] Coeficientes calculados com base na função *regress* da *toolbox* do *matlab*. Esta função retorna os coeficientes de regressão na forma, $[\beta_1;\beta_0]$, para as entradas Y_i, X_i.

[2] [*i*] Assume-se que o erro tem média nula e variância constante; [*ii*] Assume-se que os elementos do erro são independentes entre si; [*iii*] O erro deve ter uma distribuição Normal para que se possa usar a inferência estatística e generalizar à população os resultados obtidos a partir da amostra (Gujarati, 2003).

Tabela 1 (cont.)

Y_j	β_0 *[t-Student;valor de prova]*	β_1 *[t-Student;valor de prova]*	$\beta_0+\beta_1 X_j$	r
Y_{Ago}	3 193 696 *[33,259; 0,0000]*	118 627 *[12,659; 0,0000]*	3 193 696+118 627 *X*	0.95625
Y_{Set}	2 630 092 *[30,520; 0,0000]*	78 564 *[9,342; 0,0000]*	2 630 092+78 564 *X*	0.92376
Y_{Out}	1939 979 *[26,554; 0,0000]*	76 480 *[10,727; 0,0000]*	1939 979+76 480 *X*	0.94057
Y_{Nov}	1 225 761 *[22,966; 0,0000]*	50 934 *[9,779; 0,0000]*	1 225 761+50 934 *X*	0.92973
Y_{Dez}	1 019 961 *[24,912; 0,0000]*	44 875 *[11,231; 0,0000]*	1 019 961+50 934 *X*	0.95763

Nas Figuras 2 a 5, apresentam-se as rectas sobre a dispersão que define a procura turística de cada mês, ao longo de 17 anos. Esta representação permite criar uma perspectiva diferente de abordar o problema, simplificando a sua análise.

Figura 2. Representação das rectas de regressão para os meses de Janeiro, Fevereiro e Março, no intervalo [1990:2006].

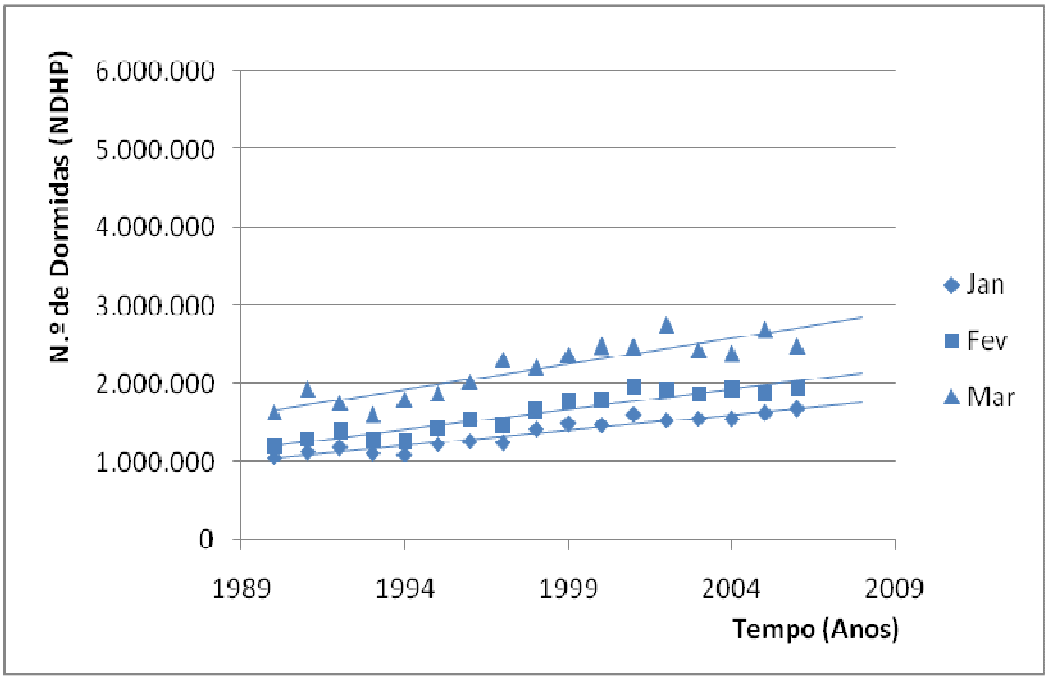


Figura 3. Representação das rectas de regressão para os meses de Abril, Maio e Junho, no intervalo [1990:2006].

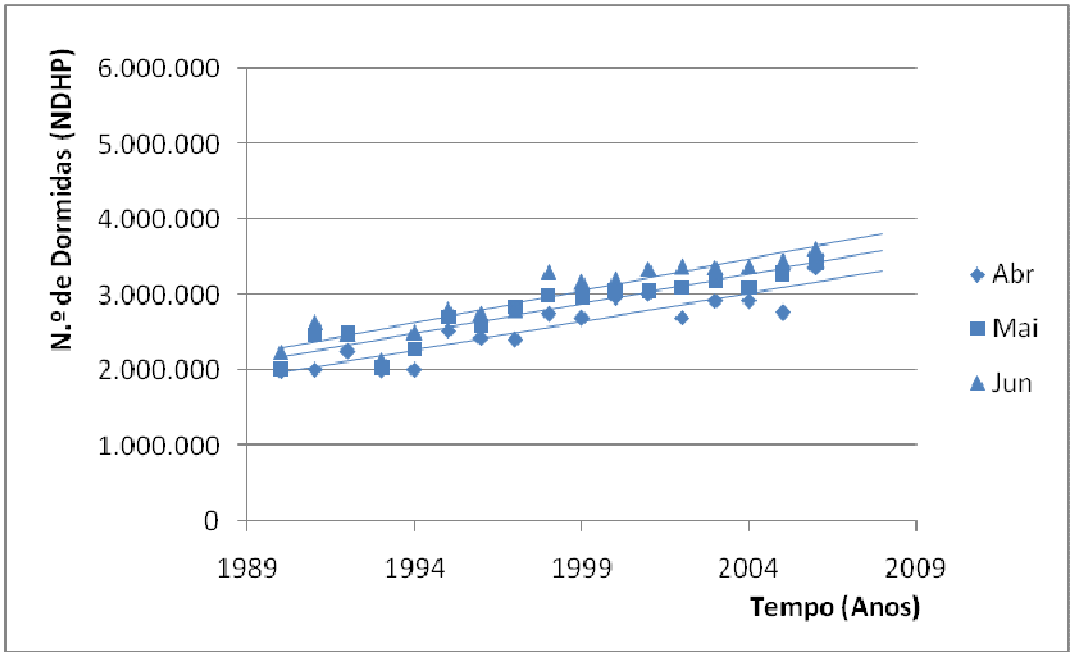


Figura 4. Representação das rectas de regressão para os meses de Julho, Agosto e Setembro, no intervalo [1990:2006].

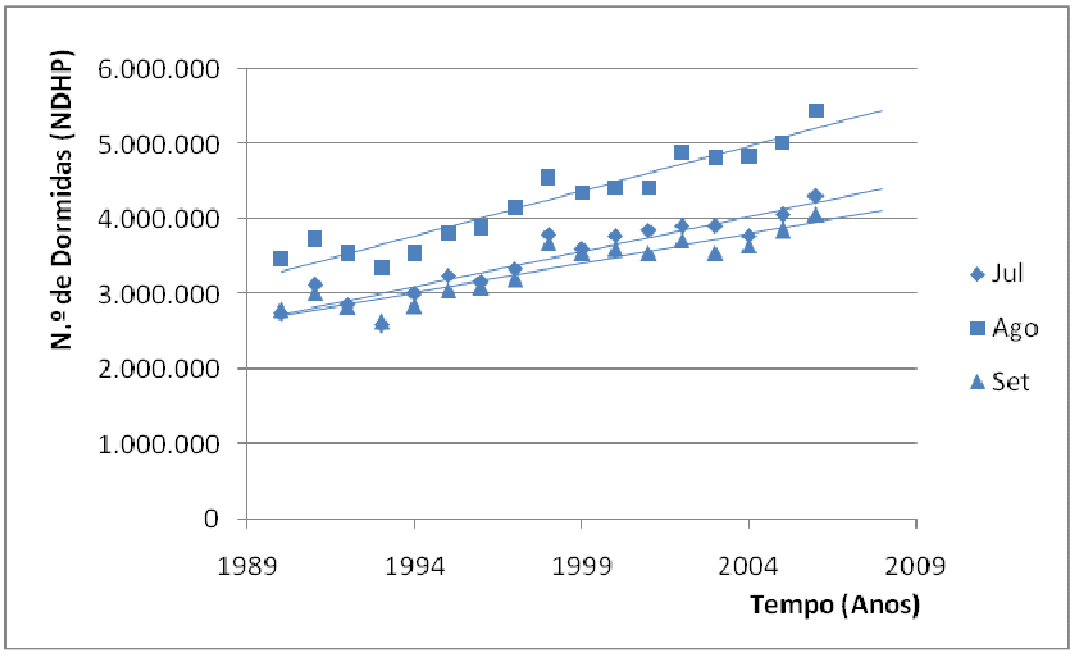
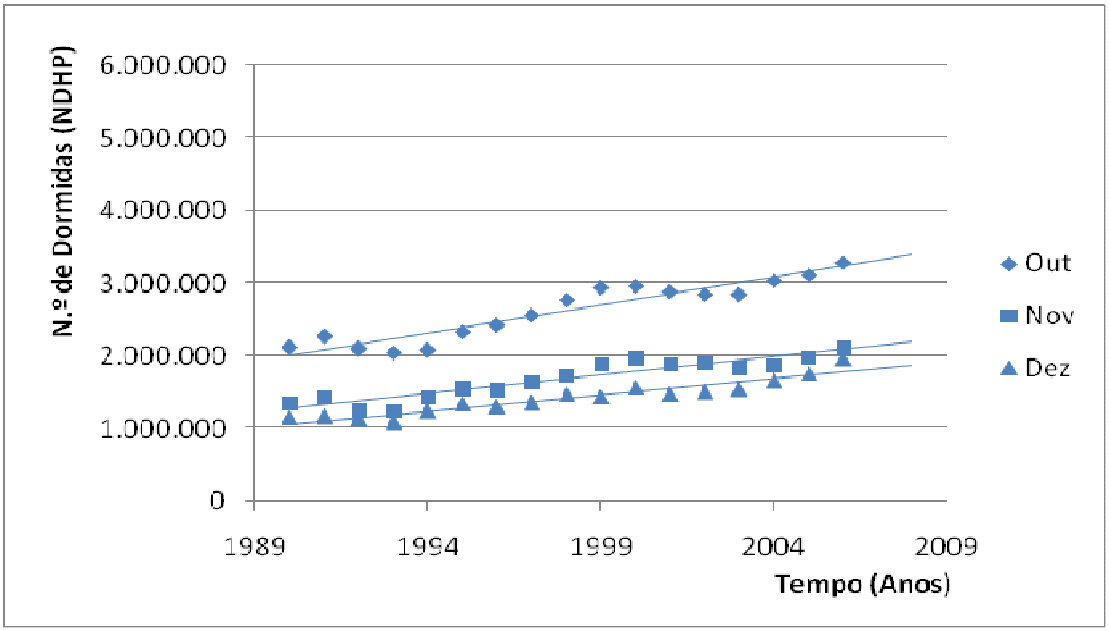
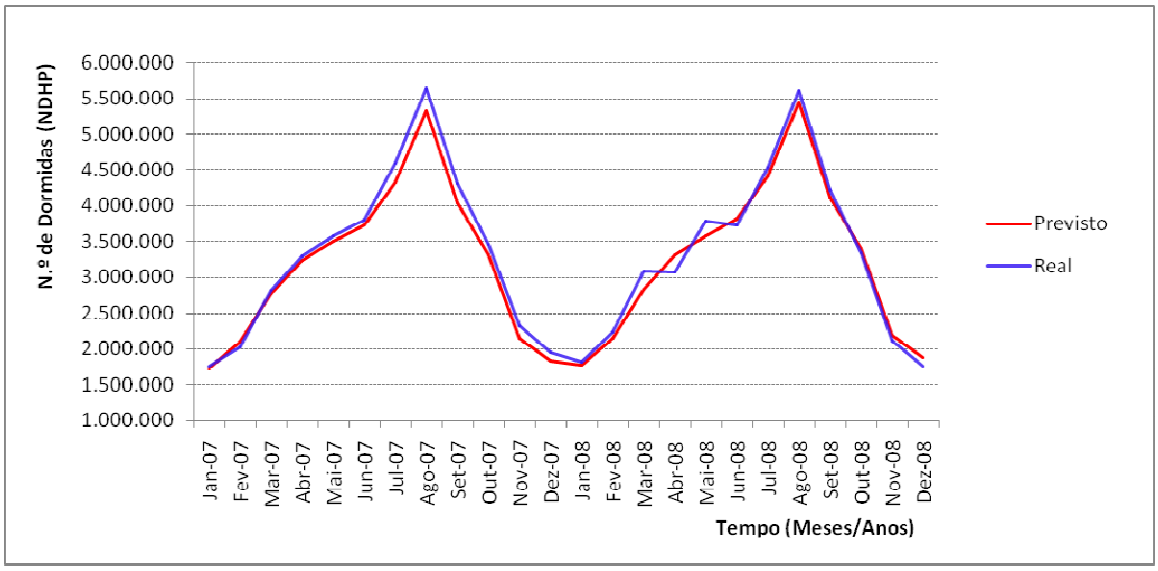


Figura 5. Representação das rectas de regressão para os meses de Outubro, Novembro e Dezembro, no intervalo [1990:2006].



Efectuada a projecção dos períodos referentes aos anos de 2007 e 2008, procedeu-se à inversão da transformação inicial definindo-se novamente a variável independente no intervalo [Jan-07:Dez-08], obtendo-se assim $Y_j^{'}$ com a previsão da série NDHP, conforme ilustrado na Figura 6.

Figura 6. Representação da série NDHP e a respectiva previsão para os anos de 2007 e 2008.



2.3. Análise e discussão dos resultados

De forma a validar a qualidade de ajustamento dos resultados obtidos foi calculado o erro percentual absoluto médio definido como:

$$EPAM = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i^{'} - Y_i}{Y_i}\right| \qquad [3]$$

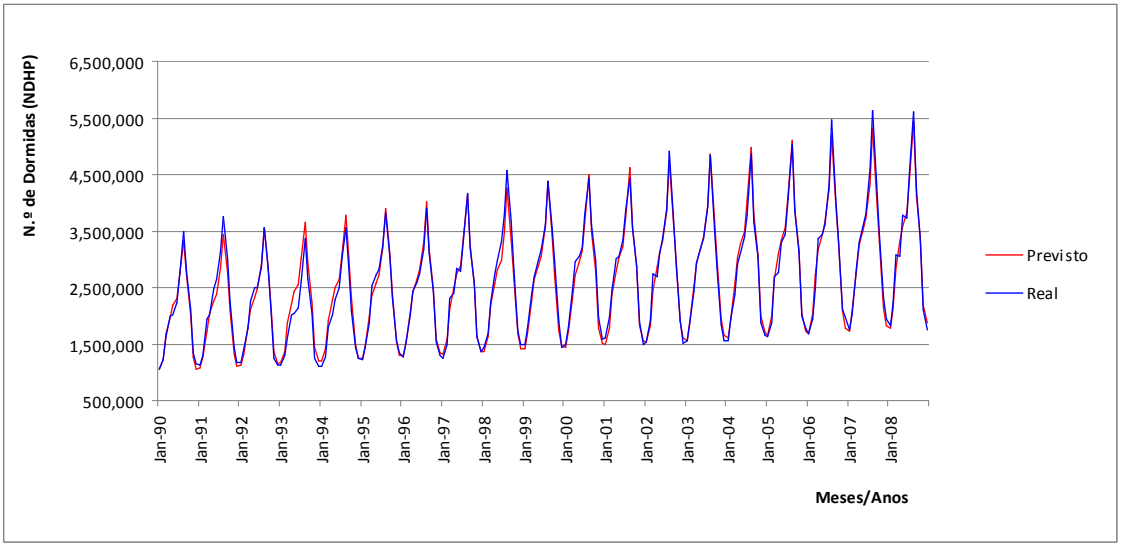
Onde, $Y_i^{'}$ representa o conjunto de valores previstos e Y_i os valores reais. No presente caso, para $n = 24$, o erro obtido foi de 4.16% (Tabela A.4, do Anexo).

O valor obtido, utilizando esta técnica de regressão, corresponde, a um resultado satisfatório, se tivermos em conta a simplicidade do método. Tendo por base a classificação do MAPE[3], ou EPAM, proposto por Lewis (1982), pode mesmo dizer-se que o modelo produziu previsões de precisão elevada, uma vez que para o EPAM obteve-se um valor inferior a 10%.

O principal problema deste método assenta no facto de ser incapaz de modelar determinados comportamentos de natureza estocástica. Todo o conjunto de amostras que não segue o padrão da maioria, será classificado com um erro acentuado, nomeadamente, as variações identificadas nos meses de Março e Abril de 2008, que registam um erro absoluto de 7.81% e 8.27%, respectivamente (Tabela A.4, do Anexo). É também superior à média o erro associado aos meses de Julho, Agosto e Setembro de 2007, de respectivamente 6.04%, 5.58% e 6.11%, resultado da inversão de tendência nos últimos dois anos de S_n, face à tendência geral dos meses anteriores (Tabela A.4, do Anexo).

Projectando Y'_i em todo o domínio de S_n, consegue ter-se uma ideia mais abrangente dos resultados deste método.

Figura 7. Representação gráfica dos valores previstos sobre a totalidade do conjunto S_n.



Como se pode observar, na Figura 7, o erro é mais acentuado no ano de 1994, com um valor absoluto de 12.2% (Tabela A.3, do Anexo), sendo o período que registou uma estimativa menos satisfatória.

Também se pode verificar da mesma observação, que existem meses que seguem uma tendência mais próxima do linear, como é o caso do mês de Janeiro, e outros que descrevem um comportamento não tão linear. Março é o mês que ao longo de todo o intervalo, apresenta um maior desvio em relação à tendência calculada (Tabela A.3, do Anexo).

3. CONCLUSÃO E DESENVOLVIMENTOS FUTUROS

A série temporal aqui estudada apresenta um comportamento linear quando analisado por mês. Assim, pode concluir-se com este estudo, que os métodos de modelação linear, mais

[3] Do inglês *Mean Absolute Percentage Error.*

concretamente a regressão, permitem estimar com uma precisão bastante satisfatória, a procura turística de estabelecimentos hoteleiros em Portugal. O facto de se ter analisado a evolução de cada mês de forma isolada, permite lidar com a especificidade característica de cada mês e assim aumentar a qualidade da estimativa.

No entanto, o modelo apresenta alguns problemas característicos dos modelos lineares, nomeadamente a impossibilidade de estimar determinados padrões que seguem um comportamento menos determinista, ou simplesmente não linear.

Existem algumas abordagens conhecidas para ultrapassar estas limitações, como é o caso dos modelos de regressão polinomial, médias móveis, entre outros.

Outras metodologias como as Redes Neuronais Artificiais (RNA) têm mostrado ser uma alternativa bastante viável na área da modelação de séries temporais. A sua capacidade para modelar padrões complexos é um trunfo no campo da ciência previsional.

Pode pensar-se, em trabalhos futuros, inclusive num modelo conjunto, que faça uso dos métodos de regressão linear, combinando-os com RNA, de forma a criar uma heurística válida à entrada da rede, que potencie a qualidade da previsão de séries temporais.

4. REFERÊNCIA BIBLIOGRÁFICAS

BUTLER, R.W.;(2001); "*Seasonality in Tourism: Issues and Implications*"; Ed. Baum, T.; Lundtorp, S., Oxford, Pergamonm.

EUROSTAT; (2009); *Statistical Office of the European Communities;* http://epp.eurostat.ec.europa.eu (acedido em 10 de Março de 2009).

FERNANDES, PAULA ODETE; (2005); "*Modelling, Prediction and Behaviour Analysis of Tourism Demand in the North of Portugal*"; Ph.D. Thesis in Applied Economy and Regional Analysis; Valladolid University - Spain.

FÚSTER, L.F.; (1989); "*Introducción a la Teoría y Técnica del Turismo*"; Madrid, Alianza.

GUJARATI, DAMODAR N.; (2003); "*Basic Econometrics*"; 4th Edition, McGraw-Hill.

INE; (1987-2008); *Anuários Estatísticos do Turismo de 1987 a 2008*. Lisboa.

JOHNSTON, J. e DINARDO, J.; (2000); "*Métodos Econométricos*", 4ª Edição, McGraw-Hill.

LEWIS, C.D.; (1982); "*Industrial and Business Forecasting Method"*; Butterworth Scientific. London.

Matlab; *Statistics Toolbox*.

TURISMO DE PORTUGAL; (2009); "*Conta Satélite do Turismo - Turismo: Actividade Económica 2000-2007*"; http://www.turismodeportugal.pt/ (acedido em 13 de Abril de 2009).

ANEXO

Tabela A.1. Dormidas registadas em Portugal (NDHP) no período [Jan-1990:Dez-2008].

	Jan	*Fev*	*Mar*	*Abr*	*Mai*	*Jun*	*Jul*	*Ago*	*Set*	*Out*	*Nov*	*Dez*
1990	1064694,00	1216039,00	1628090,00	1998486,00	2025169,00	2246786,00	2748904,00	3479306,00	2781673,00	2118146,00	1351765,00	1154476,00
1991	1125600,00	1296211,00	1936815,00	2009631,00	2485060,00	2637840,00	3131248,00	3749157,00	3027464,00	2266829,00	1428966,00	1166172,00
1992	1182262,00	1405960,00	1767046,00	2255123,00	2497848,00	2489164,00	2867486,00	3553068,00	2826538,00	2097758,00	1247996,00	1123793,00
1993	1119576,00	1288201,00	1605577,00	2006287,00	2038164,00	2142270,00	2592531,00	3362820,00	2634412,00	2045901,00	1239153,00	1090936,00
1994	1097050,00	1268858,00	1803275,00	2009921,00	2293984,00	2506588,00	3021774,00	3556404,00	2840808,00	2080653,00	1432524,00	1242112,00
1995	1227438,00	1428667,00	1878477,00	2522411,00	2705234,00	2812203,00	3246540,00	3821775,00	3057635,00	2339465,00	1554192,00	1342805,00
1996	1265604,00	1547819,00	2027335,00	2436365,00	2596425,00	2747914,00	3173069,00	3898713,00	3102936,00	2426589,00	1533598,00	1306920,00
1997	1239368,00	1478048,00	2314002,00	2402960,00	2844861,00	2799762,00	3347065,00	4163378,00	3199612,00	2568769,00	1636918,00	1355540,00
1998	1427813,00	1669745,00	2226679,00	2744398,00	2989225,00	3309335,00	3800426,00	4562245,00	3686793,00	2778732,00	1733056,00	1476052,00
1999	1491872,00	1781463,00	2380190,00	2686488,00	2984547,00	3167089,00	3608203,00	4369943,00	3554816,00	2951672,00	1886679,00	1442033,00
2000	1484802,00	1805077,00	2481530,00	2966987,00	3064380,00	3201352,00	3789338,00	4424181,00	3599561,00	2970317,00	1977758,00	1575413,00
2001	1608556,00	1964165,00	2457972,00	3011554,00	3064136,00	3349125,00	3866128,00	4426781,00	3559937,00	2887563,00	1886345,00	1480329,00
2002	1537139,00	1926842,00	2752279,00	2703439,00	3108798,00	3371407,00	3909345,00	4900264,00	3731515,00	2853204,00	1909950,00	1504786,00
2003	1550720,00	1869726,00	2429128,00	2925012,00	3195173,00	3365166,00	3920883,00	4833007,00	3551038,00	2843491,00	1843415,00	1548712,00
2004	1549992,00	1934294,00	2382340,00	2918348,00	3105531,00	3383239,00	3788629,00	4850672,00	3654583,00	3041458,00	1871106,00	1660389,00
2005	1622322,00	1878765,00	2686969,00	2764258,00	3287038,00	3432908,00	4066733,00	5034593,00	3850305,00	3122789,00	1985737,00	1765439,00
2006	1676355,00	1947106,00	2484546,00	3367859,00	3440228,00	3616770,00	4305916,00	5456461,00	4072262,00	3287801,00	2111710,00	1969363,00
2007	1763320,00	2024858,00	2816553,00	3299232,00	3575744,00	3806145,00	4599228,00	5644162,00	4307242,00	3454104,00	2330737,00	1957329,00
2008	1822603,00	2239183,00	3078757,00	3063921,00	3782832,00	3721954,00	4530198,00	5619010,00	4224292,00	3339442,00	2112816,00	1754558,00

Tabela A.2. Valores Previstos para a série NDHP, para intervalo [Jan-1990:Dez-2008].

	Jan	*Fev*	*Mar*	*Abr*	*Mai*	*Jun*	*Jul*	*Ago*	*Set*	*Out*	*Nov*	*Dez*
1990	1048753.67	1216873.61	1672655.41	1976312.94	2184321.25	2301730.02	2734839.63	3312323.41	2708656.41	2016459.69	1276694.76	1064836.12
1991	1088770.95	1268492.25	1737413.56	2050814.46	2262206.11	2385917.58	2828162.75	3430950.40	2787220.60	2092940.00	1327629.01	1109711.53
1992	1128788.24	1320110.88	1802171.71	2125315.99	2340090.96	2470105.13	2921485.87	3549577.38	2865784.78	2169420.31	1378563.25	1154586.94
1993	1168805.52	1371729.52	1866929.85	2199817.51	2417975.81	2554292.69	3014808.99	3668204.37	2944348.96	2245900.62	1429497.49	1199462.35
1994	1208822.80	1423348.16	1931688.00	2274319.03	2495860.66	2638480.25	3108132.11	3786831.35	3022913.15	2322380.93	1480431.74	1244337.76
1995	1248840.09	1474966.79	1996446.15	2348820.55	2573745.51	2722667.80	3201455.23	3905458.34	3101477.33	2398861.24	1531365.98	1289213.18
1996	1288857.37	1526585.43	2061204.29	2423322.07	2651630.36	2806855.36	3294778.35	4024085.32	3180041.51	2475341.55	1582300.22	1334088.59
1997	1328874.66	1578204.07	2125962.44	2497823.60	2729515.21	2891042.91	3388101.47	4142712.31	3258605.70	2551821.87	1633234.46	1378964.00
1998	1368891.94	1629822.71	2190720.59	2572325.12	2807400.06	2975230.47	3481424.59	4261339.29	3337169.88	2628302.18	1684168.71	1423839.41
1999	1408909.23	1681441.34	2255478.74	2646826.64	2885284.91	3059418.03	3574747.71	4379966.28	3415734.07	2704782.49	1735102.95	1468714.82
2000	1448926.51	1733059.98	2320236.88	2721328.16	2963169.76	3143605.58	3668070.83	4498593.26	3494298.25	2781262.80	1786037.19	1513590.24
2001	1488943.79	1784678.62	2384995.03	2795829.68	3041054.61	3227793.14	3761393.95	4617220.25	3572862.43	2857743.11	1836971.43	1558465.65
2002	1528961.08	1836297.25	2449753.18	2870331.21	3118939.46	3311980.70	3854717.07	4735847.24	3651426.62	2934223.42	1887905.68	1603341.06
2003	1568978.36	1887915.89	2514511.32	2944832.73	3196824.31	3396168.25	3948040.19	4854474.22	3729990.80	3010703.73	1938839.92	1648216.47
2004	1608995.65	1939534.53	2579269.47	3019334.25	3274709.16	3480355.81	4041363.31	4973101.21	3808554.99	3087184.04	1989774.16	1693091.88
2005	1649012.93	1991153.17	2644027.62	3093835.77	3352594.01	3564543.37	4134686.43	5091728.19	3887119.17	3163664.36	2040708.40	1737967.29
2006	1689030.22	2042771.80	2708785.76	3168337.29	3430478.86	3648730.92	4228009.55	5210355.18	3965683.35	3240144.67	2091642.65	1782842.71
2007	1729047.50	2094390.44	2773543.91	3242838.82	3508363.71	3732918.48	4321332.67	5328982.16	4044247.54	3316624.98	2142576.89	1827718.12
2008	1769064.78	2146009.08	2838302.06	3317340.34	3586248.56	3817106.03	4414655.79	5447609.15	4122811.72	3393105.29	2193511.13	1872593.53

Tabela A.3. Erro percentual absoluto de NDHP no intervalo [Jan-1990:Dez-2008].

	Jan	*Fev*	*Mar*	*Abr*	*Mai*	*Jun*	*Jul*	*Ago*	*Set*	*Out*	*Nov*	*Dez*
1990	1.5%	0.1%	2.7%	1.1%	7.9%	2.5%	0.5%	4.8%	2.6%	4.8%	5.6%	7.8%
1991	3.3%	2.1%	10.3%	2.1%	9.0%	9.6%	9.7%	8.5%	7.9%	7.7%	7.1%	4.8%
1992	4.5%	6.1%	2.0%	5.8%	6.3%	0.8%	1.9%	0.1%	1.4%	3.4%	10.5%	2.7%
1993	4.4%	6.5%	16.3%	9.7%	18.6%	19.2%	16.3%	9.1%	11.8%	9.8%	15.4%	10.0%
1994	10.2%	12.2%	7.1%	13.2%	8.8%	5.3%	2.9%	6.5%	6.4%	11.6%	3.3%	0.2%
1995	1.7%	3.2%	6.3%	6.9%	4.9%	3.2%	1.4%	2.2%	1.4%	2.5%	1.5%	4.0%
1996	1.8%	1.4%	1.7%	0.5%	2.1%	2.1%	3.8%	3.2%	2.5%	2.0%	3.2%	2.1%
1997	7.2%	6.8%	8.1%	4.0%	4.1%	3.3%	1.2%	0.5%	1.8%	0.7%	0.2%	1.7%
1998	4.1%	2.4%	1.6%	6.3%	6.1%	10.1%	8.4%	6.6%	9.5%	5.4%	2.8%	3.5%
1999	5.6%	5.6%	5.2%	1.5%	3.3%	3.4%	0.9%	0.2%	3.9%	8.4%	8.0%	1.9%
2000	2.4%	4.0%	6.5%	8.3%	3.3%	1.8%	3.2%	1.7%	2.9%	6.4%	9.7%	3.9%
2001	7.4%	9.1%	3.0%	7.2%	0.8%	3.6%	2.7%	4.3%	0.4%	1.0%	2.6%	5.3%
2002	0.5%	4.7%	11.0%	6.2%	0.3%	1.8%	1.4%	3.4%	2.2%	2.8%	1.2%	6.6%
2003	1.2%	1.0%	3.5%	0.7%	0.1%	0.9%	0.7%	0.4%	5.0%	5.9%	5.2%	6.4%
2004	3.8%	0.3%	8.3%	3.5%	5.5%	2.9%	6.7%	2.5%	4.2%	1.5%	6.3%	2.0%
2005	1.7%	6.0%	1.6%	11.9%	2.0%	3.8%	1.7%	1.1%	1.0%	1.3%	2.8%	1.6%
2006	0.8%	4.9%	9.0%	5.9%	0.3%	0.9%	1.8%	4.5%	2.6%	1.5%	1.0%	9.5%
2007	1.9%	3.4%	1.5%	1.7%	1.9%	1.9%	6.0%	5.6%	6.1%	4.0%	8.1%	6.6%
2008	2.9%	4.2%	7.8%	8.3%	5.2%	2.6%	2.6%	3.1%	2.4%	1.6%	3.8%	6.7%

Tabela A.4. Erro percentual absoluto médio de NDHP no intervalo [Jan-2007:Dez-2008].

Meses	*2007*	*2008*	*EPAM*
Jan	1.94%	2.94%	4,16%
Fev	3.43%	4.16%	
Mar	1.53%	7.81%	
Abr	1.71%	8.27%	
Mai	1.88%	5.20%	
Jun	1.92%	2.56%	
Jul	6.04%	2.55%	
Ago	5.58%	3.05%	
Set	6.11%	2.40%	
Out	3.98%	1.61%	
Nov	8.07%	3.82%	
Dez	6.62%	6.73%	

LINEARIDADE E DEPENDÊNCIA DE LONGO PRAZO: UMA APLICAÇÃO A SÉRIES DE ÍNDICES BOLSISTAS

JOSÉ RAMOS PIRES MANSO
E-mail: pmanso@ubi.pt
Departamento de Gestão e Economia
Universidade da Beira Interior (Portugal)

VÍTOR MANUEL DE SOUSA GABRIEL
E-mail: vigab@ipg.pt
Escola Superior de Tecnologia e Gestão
Instituto Politécnico da Guarda (Portugal)

Resumo

Neste estudo é analisado o comportamento de índices bolsistas. As séries usadas correspondem aos preços de fecho de vários índices internacionais, no período de 3 de Janeiro de 2000 a 13 de Fevereiro de 2009. Para o efeito são confirmados diversos factos estilizados acerca das séries, são aplicados vários testes de aderência à normalidade, de linearidade, e é investigada a dependência de longo prazo, recorrendo à estatística R/S e ao coeficiente de Hurst.

Palavras-chave: Bolsa, normalidade, linearidade, dependência de longo prazo.
Area temática: Modelação Económica e Financeira.

Abstract

This paper analyze the behavior of stock market indexes. The used series correspond to daily closures of international indices in the period from January 1, 2000 to February 13, 2009. We report and analyze some facts concerning the series. Various tests are applied to inquire normality and linearity. To investigate the long range memory occurrences, we use the rescaled range analysis and the Hurst exponent.

Key-words: Stock market, normality, linearity, long-run dependency.
Thematic Area: Economic and Financial Modelling.

1. INTRODUÇÃO E OBJECTIVOS

O presente documento pretende analisar o comportamento de alguns dos mercados bolsistas mundiais, nomeadamente o BOVESPA (Brasil), o CAC (França), o DAX (Alemanha), o DJIA (Estados Unidos), o FTSE (Inglaterra), o IBEX35 (Espanha), o NIKKEI (Japão) e o PSI20 (Portugal).

De acordo com alguns autores, como Peters (1994) e Mandelbrot (2006), os preços dos activos financeiros bem como a sua volatilidade sofrem grandes variações ao longo do tempo, coexistindo períodos de uma certa estabilidade dos preços dos activos com outros períodos de variação mais frequente e, em alguns casos, muito acentuada. As distribuições empíricas da rentabilidade desses activos são leptocúrticas, assimétricas, e apresentam *clusters* de volatilidade, pois a períodos de elevada volatilidade seguem-se outros com volatilidade semelhante, e vice-versa, concentrando-se a volatilidade em determinados períodos de tempo.

Mandelbrot, ao estudar as bolsas de valores de algumas das maiores bolsas mundiais, concluiu que, se um índice cair num determinado mês, a probabilidade de cair no mês seguinte é ligeiramente maior. Do mesmo modo, se o índice subir, a probabilidade de continuar a subir é ligeiramente maior (Mandelbrot, 2006).

Os objectivos deste trabalho compreendem a confirmação destes factos estilizados. Para esse efeito, analisamos estatisticamente as séries dos índices e aplicamos vários testes de aderência, de modo a concluirmos acerca da adequação do ajustamento da distribuição normal às distribuições empíricas. Os objectivos do estudo compreendem ainda a detecção de dependência e de linearidade das séries, e a detecção de dependências de longo prazo.

2. METODOLOGIA

Para podermos confirmar os factos estilizados acerca da distribuição empírica dos índices, representamos graficamente e analisamos os dados das respectivas séries, sob a forma de valores de fecho e de taxa de rendibilidade, calculamos algumas estatísticas, e concluímos acerca da estacionaridade dos mesmos, recorrendo aos correlogramas das funções de autocorrelação, ao teste Q de Box-Pierce e ao teste Dickey-Fuller Aumento (ADF). De seguida aplicamos os testes de Jarque-Bera (J-B), de Kolmogorov-Smirnov (K-S) e Anderson-Darling (A-D), e desenhamos o gráfico de quantis (Q-Q plot), de modo a testar a bondade do ajustamento da distribuição normal às distribuições empíricas das oito séries dos referidos índices.

A seguir, aplicamos o teste BDS, desenvolvido em 1987 por W. A. Brock, W. Deckert e J. Scheinkman, através do qual se testa a hipótese nula de uma série ser independente e identicamente distribuida contra a alternativa não especificada.

Por fim, será aplicada a análise R/S, desenvolvida por E. Hurst, e mais tarde aprofundada por B. Mandelbrot, para se calcular os coeficientes de Hurst, de modo a se perceber se as séries são caracterizadas por dependências de longo prazo, persistentes ou anti-persistentes, ou, se pelo contrário, são caracterizadas por movimentos do tipo passeio aleatório.

2.1 Propriedades estatísticas dos índices

Para se ficar com uma primeira ideia acerca da bondade do ajustamento da distribuição normal às distribuições empíricas, começamos por calcular a assimetria e a curtose. A seguir, aplicamos os testes de Jarque-Bera, de Kolmogorov-Smirnov e Anderson-Darling, e construímos os gráficos de quantis, de modo a se concluir acerca da bondade do ajustamento da distribuição normal.

O teste de Jarque-Bera é realizado a partir dos resultados de assimetria e curtose. Considerando-se a tabela do χ^2, com dois graus de liberdade e significância de 5%. O teste é dado pela seguinte equação:

$$W = N\left[\frac{A^2}{6} + \frac{(C-3)^2}{24}\right]$$

Onde: N é o tamanho da amostra, A é o coeficiente de assimetria e C é o coeficiente de custose. As hipóteses levantadas são:

$H_0 : W < 5{,}991$, distribuição normal.

$H_1 : W \geq 5{,}991$, distribuição não-normal.

O teste de K-S também pode ser aplicado para se averiguar se a característica estudada da amostra tem origem numa população com distribuição normal. Este teste é de simples execução, baseando-se na maior diferença absoluta entre a frequência acumulada observada e a estimada pela distribuição normal.

Por seu lado, o teste A-D é uma generalização do teste de K-S, contudo atribuindo mais peso às caudas. Este é um teste de distância ou de distribuição empírica da função, e baseia-se

no conceito de que quando uma distribuição de hipótese subjacente é dada, os dados podem ser transformados numa distribuição uniforme. Os dados da amostra transformada podem, por meio de um teste de distância, ser testados quanto à sua uniformidade.

O gráfico Q-Q é um método gráfico frequentemente utilizado na verificação da normalidade de séries temporais, consistindo na comparação gráfica dos quantis teóricos da distribuição normal com os quantis dos dados amostrais.

2.2 Correlogramas e teste ADF para apreciar a estacionaridade das séries

De acordo com Brooks (2002), um processo é estacionário quando a sua média, variância e auto-covariância se mantêm constantes ao longo do tempo.

Das várias alternativas disponíveis para analisar a estacionaridade, um teste simples baseia-se na chamada função de autocorrelação (ACF), que envolve o desenvolvimento de testes individuais sobre a significância dos coeficientes de autocorrelação.

Estatisticamente, a presença de uma ou mais raízes unitárias indica não-estacionaridade no comportamento de uma série temporal. Para testar a presença de uma raiz unitária usamos o teste Dickey-Fuller Aumentado.

O teste ADF usa uma expressão do tipo:

$$\Delta y_t = \rho_0 + \rho_1 y_{t-1} + \rho_2 t + \sum \delta_i \Delta y_{t-i} + \varepsilon_t$$

A formulação das hipóteses nula e alternativa é feita do seguinte modo:

H_0: $\rho_1 = 0$ (a série possui raízes unitárias ou tem comportamento não-estacionário);

H_1: $\rho_1 < 0$ (a série não possui raízes unitárias ou tem comportamento estacionário);

2.3 Teste BDS para apreciar a linearidade das séries

O teste BDS foi desenvolvido em 1987 por W. A. Brock, W. Deckert e J. Scheinkman (Brock, Dechert & Scheinkman, 1987). Este teste é uma importante ferramenta para detectar dependência em séries temporais, testando a hipótese nula de uma série ser independente e identicamente distribuída contra a alternativa não especificada. O teste BDS não pode testar directamente o caos, mas só a linearidade, proporcionada se este for aplicado aos resíduos de um processo auto-regressivo de 1ª ordem (Peters, 1994). Contudo, a não linearidade é uma das indicações do caos, podendo ser utilizado o teste BDS para detectar esta indicação.

O teste BDS utiliza o conceito de correlação espacial da teoria do caos. O cálculo deste teste compreende os seguintes procedimentos:

1. Dada uma série temporal com N observações, calcula-se a primeira diferença dos logaritmos dos dados da série temporal;

$$\{x_i\} = [x_1, x_2, x_3, ..., x_N]$$

2. Escolhe-se o valor de *m* (dimensão de mergulho), mergulha-se a série nos vectores de dimensão m, escolhendo cada um dos *m* sucessivos pontos nas séries. Este procedimento converte a série de escalares em séries de vectores.

$$x_1^m = (x_1, x_2, ..., x_m)$$

$$x_2^m = (x_2, x_3, ..., x_{m+1})$$

$$\vdots$$

$$x_{N-m}^{m} = (x_{N-m}, x_{N-m+1}, ..., x_N)$$

3. Calcula-se o integral de correlação (integral correlation), o qual mede a correlação espacial dos pontos, adicionando o número de pares dos pontos (i, j), em que $1 \le i \le N$ e $1 \le j \le N$, no espaço de dimensão m que é "fechado" de que os pontos estão dentro do raio de tolerância ε de cada um.

$$C_{\varepsilon,m} = \frac{1}{N_m(N_m - 1)} \sum_{i \neq j} I_{i,j;\varepsilon}$$

Em que $I_{i,j;\varepsilon}$ = 1 se $I_{i,j;\varepsilon} \left\| x_i^m - x_j^m \right\| \le \varepsilon$

= 0, em caso contrário.

4. Brock, Dechert & Scheinkman (1987) concluíram que se uma série é IID,

$$C_{\varepsilon,m} \approx \left[C_{\varepsilon,1}\right]^m$$

Se o rácio $\frac{N}{m}$ for superior a 200, se os valores de $\frac{\varepsilon}{\sigma}$ estiverem compreendidos entre 0,5 e 2 (Lin, 1997), e os valores de m estiverem compreendidos entre 2 e 6 (Brock e Sayers, 1988), a quantidade $\left[C_{\varepsilon,m} - (C_{\varepsilon,1})^m\right]$ segue uma distribuição normal, com média zero e variância $V_{\varepsilon,m}$, definida como:

$$V_{\varepsilon,m} = 4\left[K^m + 2\sum_{j=1}^{m-1} K^{m-j} C_\varepsilon{}^{2j} + (m-1)^2 C_\varepsilon{}^{2m} - m^2 K C_\varepsilon{}^{2m-2} \right]$$

Em que:

$$K = K_\varepsilon = \frac{6}{N_m(N_m - 1)(N_m - 2)} \sum_{i<j<N} h_{i,j,N;\varepsilon} \; ;$$

$$h_{i,j,N;\varepsilon} = \frac{\left[I_{i,j;\varepsilon} I_{j,N;\varepsilon} + I_{i,N;\varepsilon} I_{N,j;\varepsilon} + I_{j,i;\varepsilon} I_{i,N;\varepsilon} \right]}{3}$$

5. O teste BDS pode ser apresentado do seguinte modo:

$$BDS_{\varepsilon,m} = \frac{\sqrt{N}\left[C_{\varepsilon,m} - (C_{\varepsilon,1})^m\right]}{\sqrt{V_{\varepsilon,m}}}$$

Brock, Hsieh e LeBaron (1991) concluíram que a distribuição normal é assimptoticamente uma boa aproximação para a distribuição da estatística BDS, sempre que se disponha de mais de 500 observações, como é o caso das séries estudadas. O teste BDS é um teste bi-caudal, devendo-se rejeitar a hipótese nula se a estatística do teste for inferior ao valor crítico (por exemplo, se se considerar α = 0,05, o valor critico é de ± 1,96).

A formulação das hipóteses nula e alternativa é feita do seguinte modo:

H_0: a série histórica possui variáveis independentes e identicamente distribuídas (IID), indicando que o modelo possui características lineares;

H_1: a série histórica não possui variáveis independentes e identicamente distribuídas (IID), indicando que o modelo possui características de não linearidade.

É importante sublinhar que a estatística BDS é muito influenciada pela escolha dos valores de □e de *m*. Com □grande (pequeno) a correlação espacial entre os pontos da série tende a ser elevada (reduzida). Quanto maior for a dimensão de mergulho, mais pequeno será o número de histórias não sobrepostas e, em consequência, os pontos definidos pelos vectores implícitos ficarão mais "próximos" e o valor da estatística BDS tenderá a ser maior.

2.4 Coeficiente de Hurst para apreciar a dependência de longo prazo

O expoente de Hurst foi proposto em 1951 por H. E. Hurst[1], para utilizar em análise fractal, e tem sido aplicado em várias áreas de investigação. Recentemente, tornou-se bastante popular no âmbito das finanças, em grande parte devido aos trabalhos de Peters (1994). Outros trabalhos importantes nesta área foram os desenvolvidos por May (1999), Corazza e Malliaris (2002) e Grech e Mazur (2004).

O coeficiente de Hurst é obtido a partir da seguinte lei de potência:

$$R/S(N) = (a * N)^H$$

Em que:

a : é uma constante

N : é o número de observações

H : é o exponente do coeficiente de Hurst

$R/S(N)$: é a estadística R/S, que depende do tamanho da série.

O coeficiente de Hurst é utilizado para a detecção de memória de largo prazo em séries temporais. No caso de ser detectada, é uma clara evidência de comportamento não linear.

O procedimento para o cálculo do coeficiente de Hurst é o que se descreve a seguir:

a) Usando as variações logarítmicas diárias dos índices bolsistas obtém-se uma série temporal que se descreve como se segue: $\{X_t\} t = 1,...,T$;

b) Particiona-se a amostra total em sub-amostras *i* de tamanho similar, em que

$$n = \frac{N}{i}$$

c) Calcula-se a média e o desvio padrão da série temporal para cada sub-amostra:

$$m(N,t_0) = \sum_{t=t_0+1}^{t_0+N} \frac{x_t}{N} \qquad S(N,t_0) = \sqrt{\frac{1}{N}\sum_{t=t_0+1}^{t_0+N} \left[x_t - m(N,t_0)\right]^2}$$

d) Obtém-se uma nova série para captar a diferença acumulada em relação à média do índice:

$$y(N,t_0,\tau) = \sum_{t=t_0+1}^{t_0+\tau} \left[x_t - m(N,t_0)\right] \qquad 1 \le \tau \le N$$

[1] Edward Hurst foi um importante hidrólogo que projectou a barragem de Assuan (Egipto) e estudou séries temporais relacionadas com o caudal do rio nilo e os problemas de armazenamento de água. Utilizou uma base de dados de 800 anos de arquivos e concluiu que havia uma tendência para a um ano de grande caudal se seguir outro de caudal ainda mais elevado, e para a um de baixo caudal se seguir outro ainda mais baixo. A partir desta constatação, criou o método estatístico (R/S).

e) Determina-se a diferença máxima e mínima de cada partição. A diferença destes valores é a amplitude de cada sub-amostra:

$$R(N,t_0) = \underset{\tau}{máx}\ y(N,t_0,\tau) - \underset{\tau}{mín}\ y(N,t_0,\tau)$$

f) Obtém-se a estadística R/S dividindo a amplitude pelo desvio padrão encontrado em c):

$$[R/S](N) = \frac{\sum_{t_0} R(N,t_0)}{\sum_{t_0} S(N,t_0)}$$

g) Para cada partição, calcula-se a média de R/S. Dado que o crescimento de R/S segue a lei de potência descrita anteriormente, a potência da equação é o coeficiente de Hurst, o qual se obtém a partir da seguinte regressão:

$$\log\ R/S = a + H.\log N$$

Com o objectivo de se verificar a dependência de longo prazo nas séries dos vários índices estudados, calculamos os rendimentos logaritmicos diários através da expressão $R_t = \log\left(X_t \big/ X_{t-1}\right)$, sendo X_t e X_{t-1} os valores de fecho dos índices nos momentos t e t-1, respectivamente. Pois, de acordo com Peters (1994), os rendimentos logaritmicos são mais apropriados do que os rendimentos percentuais dos índices. A amplitude utilizada na análise R/S é o desvio acumulado a partir da média, e os rendimentos logaritmicos somam ao rendimento acumulado, enquanto com as variações percentuais isso não acontece.

3. APLICAÇÃO EMPÍRICA

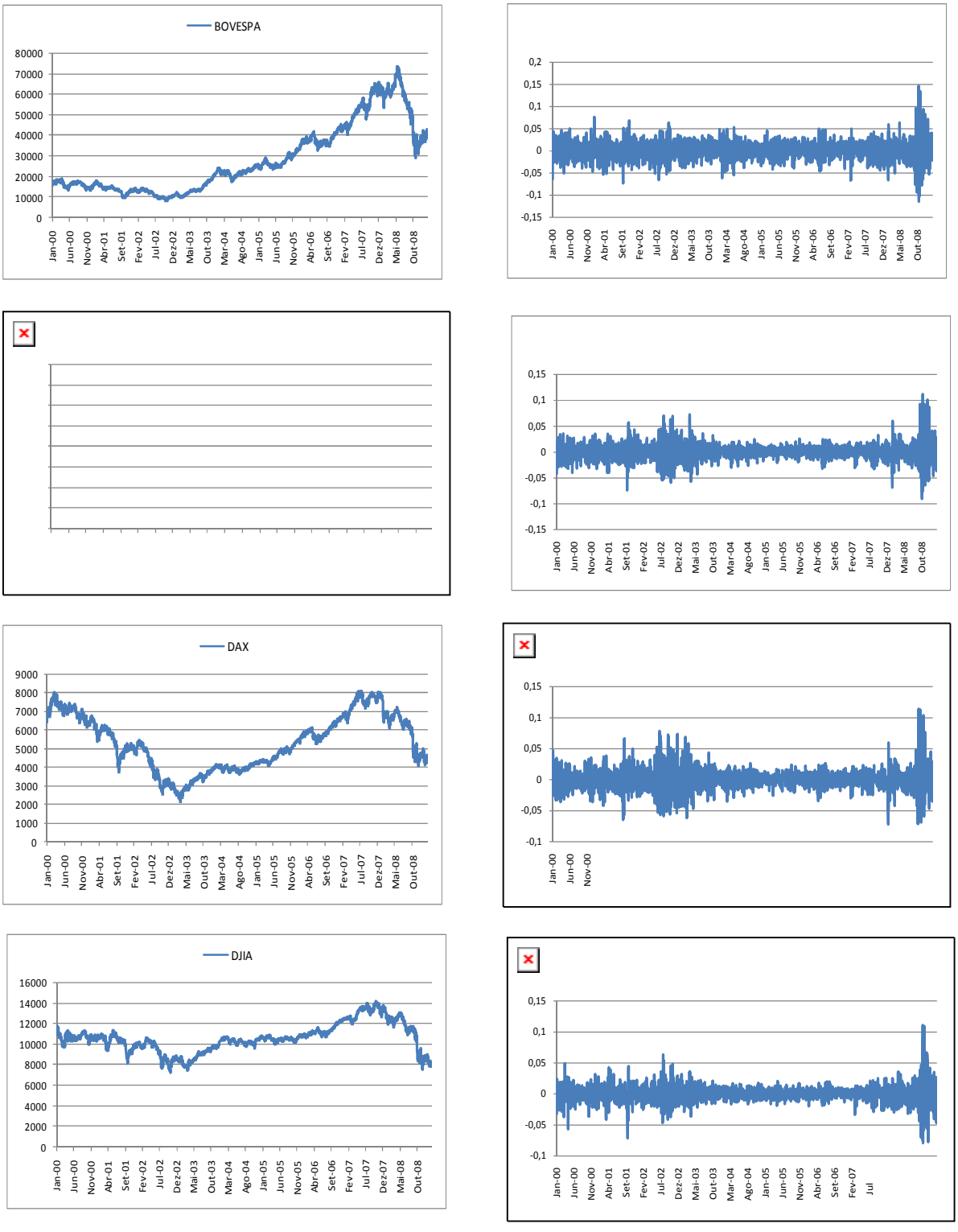
3.1 Base de Dados e suas Fontes

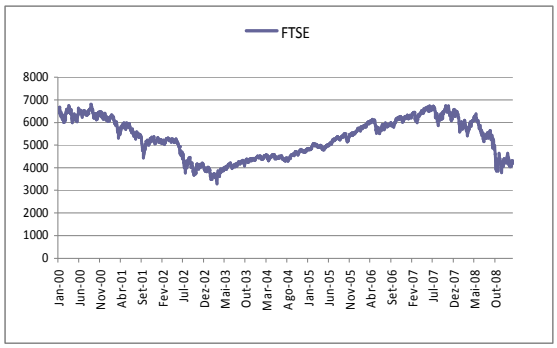
Os dados usados na aplicação empírica referem-se aos índices de alguns dos principais mercados bolsistas mundiais, nomeadamente o BOVESPA (Brasil), o CAC (França), o DAX (Alemanha), o DJIA (Estados Unidos), o FTSE (Inglaterra), o IBEX35 (Espanha), o NIKKEI (Japão) e o PSI20 (Portugal), no período compreendido entre 3 de Janeiro de 2000 e 13 de Fevereiro de 2009. As séries dos índices foram obtidas a partir do site: http://finance.yahoo.com/.

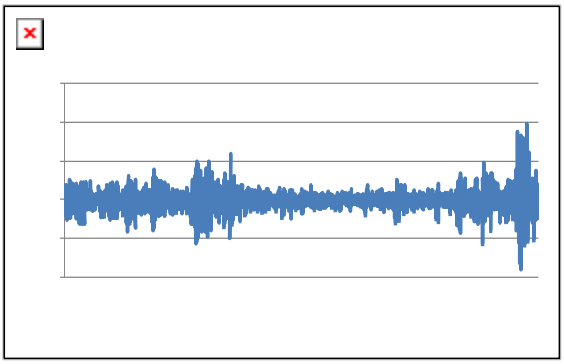
3.2 Análise de propriedades estatísticas dos índices

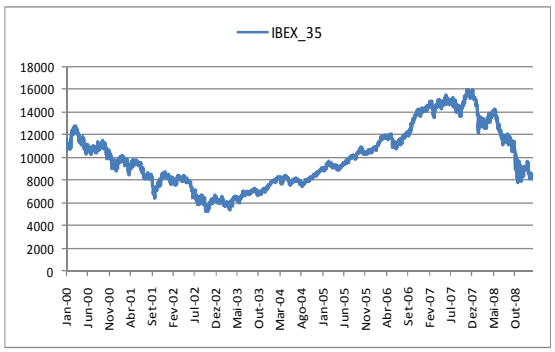
A evolução dos valores de fecho de cada índice (à esquerda) e das rendibilidades diárias (à direita) pode ser vista na figura 1.

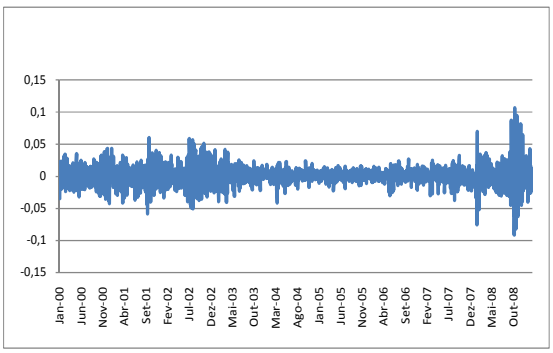
Figura 1: Evolução dos índices e respectivas taxas de rendibilidade (%)

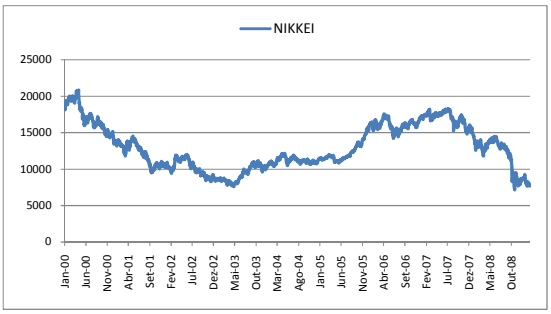


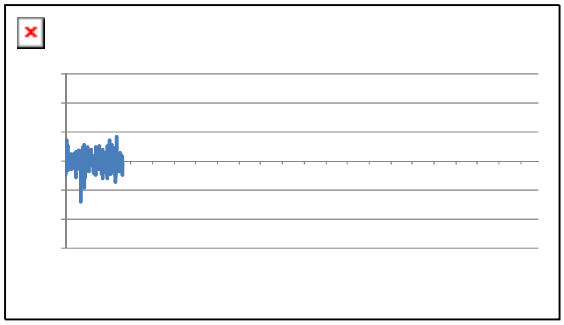


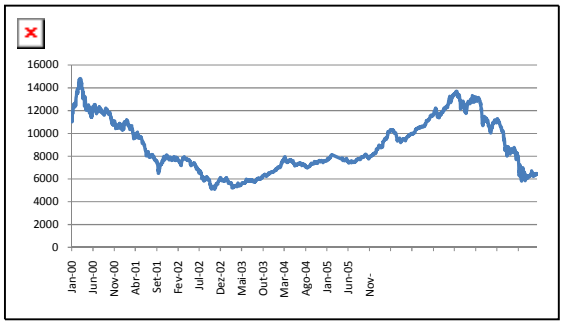


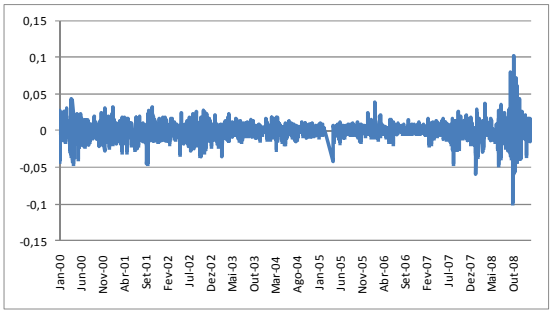












A análise gráfica dos índices permite concluir que os mesmos sofrem fortes variações ao longo do período em análise. Por outro lado, a análise das rendibilidades mostra a tendência para a volatilidade se concentrar em determinados períodos. A alguns períodos de uma certa acalmia sucedem-se outros caracterizados por uma certa turbulência e grande volatilidade, ou seja, as taxas de rendibilidade apresentam clusters de volatilidade. No período analisado, os meses de Setembro e Outubro de 2008 foram os de maior concentração de volatilidade, em resultado da crise dos mercados financeiros.

Quanto ao segundo facto estilizado, analisando os valores que constam do quadro 1, podemos concluir que todas as séries de rendibilidades evidenciam sinais de desvio face à hipótese de normalidade, já que os coeficientes de assimetria e de curtose são estatisticamente

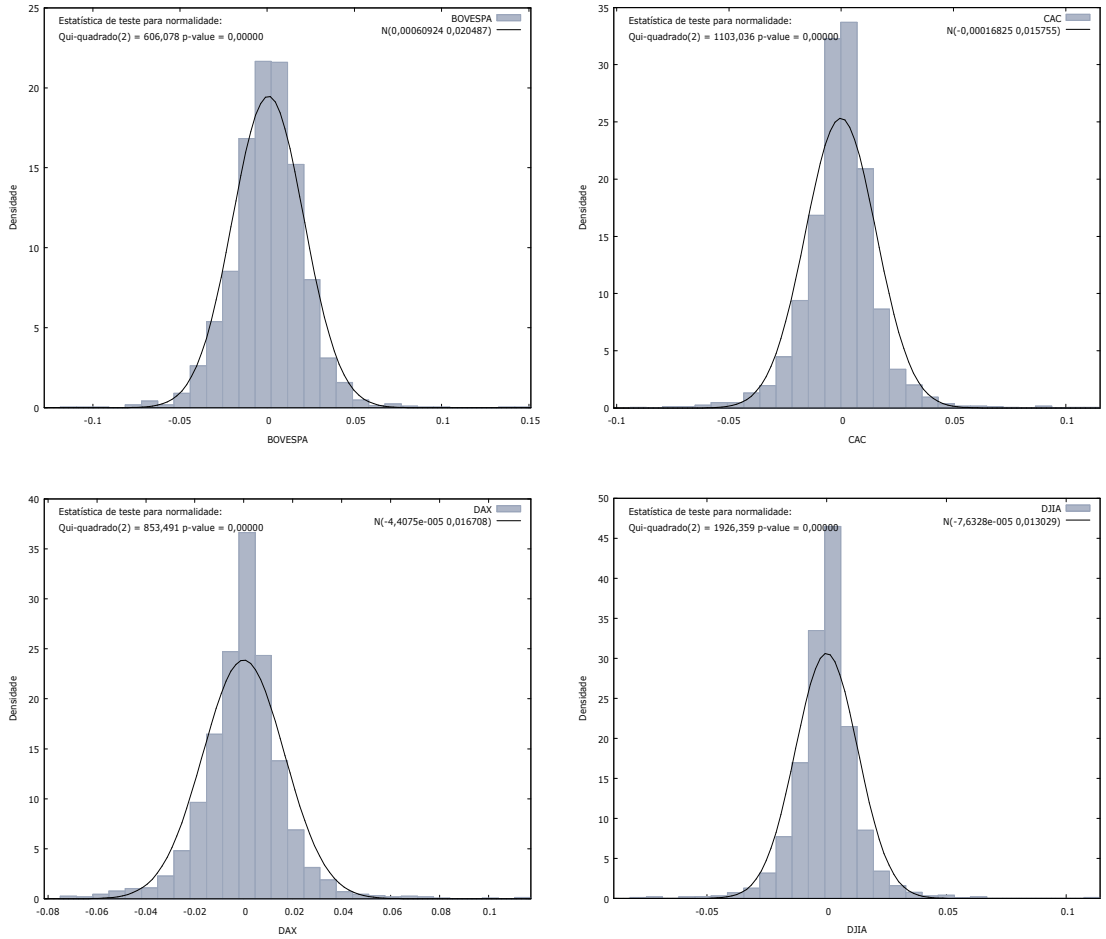
diferentes dos de uma distribuição normal[2]. As três séries são leptocúrticas e apresentam abas assimétricas. Com excepção do índice PSI20, todos os índices apresentam assimetria positiva.

Quadro 1: Medidas de estatística descritiva das taxas de rendibilidade

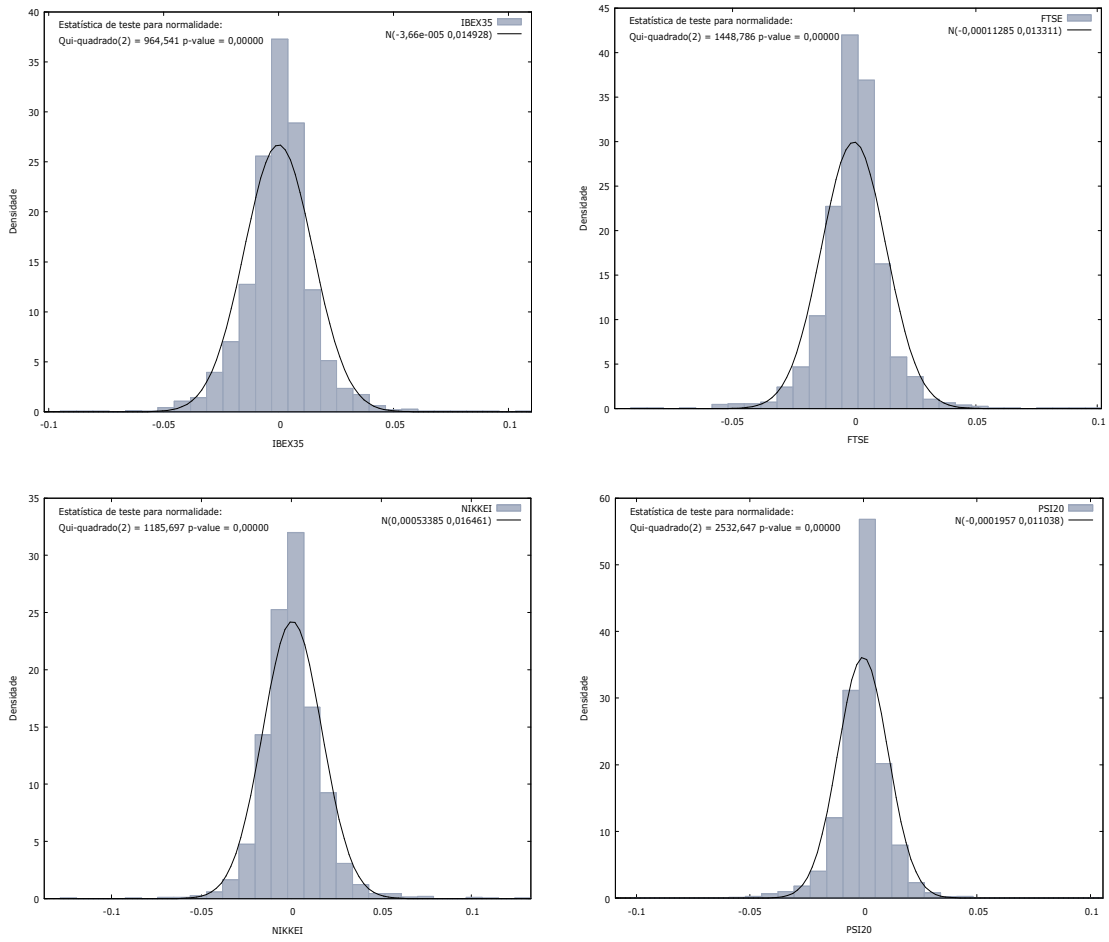
	BOVESPA	CAC	DAX	DJI	IBEX	FTSE	NIKKEI	PSI_20
Observações	2255	2327	2318	2292	2297	2303	2238	2252
Média	0,00061	-0,00017	-0,00004	-0,00008	-0,00004	-0,00011	0,00053	-0,00021
Mediana	0,00121	0,00008	0,00059	0,00032	0,00068	0,00013	0,00013	0,00018
Máximo	0,14656	0,11176	0,11402	0,11080	0,10647	0,09839	0,12875	0,10197
Mínimo	-0,11393	-0,09037	-0,07164	-0,07873	-0,09141	-0,08848	-0,12396	-0,09859
Desvio-padrão	0,02049	0,01576	0,01671	0,01303	0,01493	0,01331	0,01646	0,01132
Assimetria	0,07117	0,21047	0,26166	0,18151	0,14562	0,07215	0,54968	-0,13343
Curtose	6,71865	8,67299	7,80964	11,67315	8,13317	9,85711	9,95445	12,81226

A característica leptocúrtica das taxas de rendibilidade também pode ser constatada nos histogramas apresentados na figura 2.

Figura 2: Histogramas das séries com sobreposição da curva normal



[2] No caso de uma distribuição normal, deverá ter-se um valor de zero para o coeficiente de assimetria e um valor de três para o coeficiente de curtose.



Para se testar a adequação do ajustamento da distribuição normal às distribuições empíricas das oito séries, aplicamos os testes de aderência de J-B, K-S e A-D, cujos valores estatísticos podem ser vistos no quadro 2.

Quadro 2: Testes à normalidade da distribuição das taxas de rendibilidade

	BOVESPA	CAC	DAX	DJI	IBEX	FTSE	NIKKEI	PSI_20
J-B	1301,20	3137,58	2260,68	7196,43	2529,98	4513,96	4622,68	9040,99
K-S	1,88	3,40	3,16	4,02	3,38	3,77	3,09	4,178
A-D	8,48	28,20	26,78	37,75	26,16	34,62	19,96	39,002

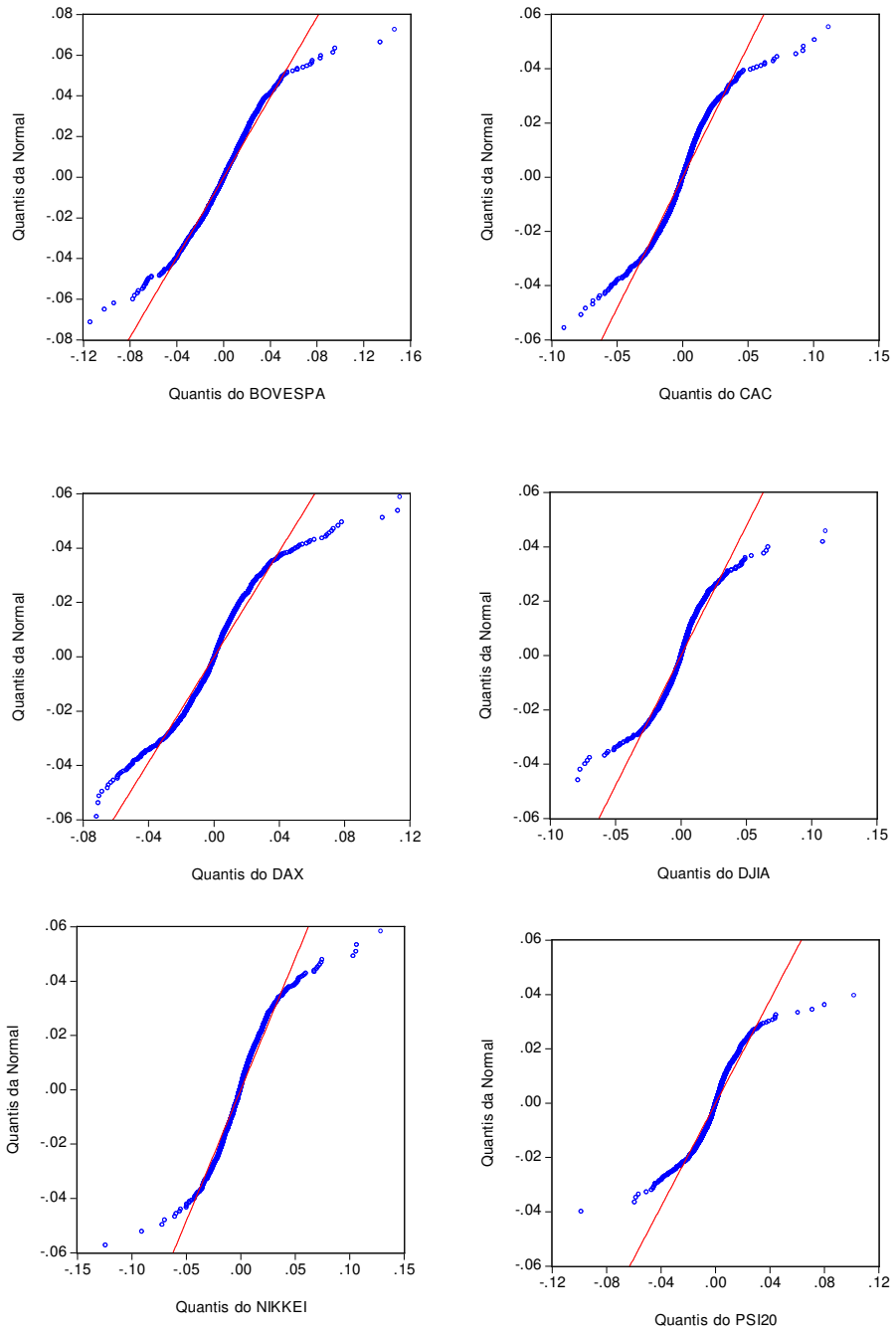
Nota: os valores dos testes JB e KS foram calculados com recurso ao software Eviews, versão 6.0, enquanto a estatística do teste AD foi calculada com recurso ao software SPSS, versão 16.0.

Tendo por base os valores dos testes e as respectivas probabilidades associadas (iguais a zero), concluímos que são todos estatisticamente significativos a 1%, rejeitando-se claramente a hipótese de normalidade das séries de rendimentos.

Para se poder confirmar a conclusão a que se chegou com os testes de aderência, desenhamos os gráficos de quantis[3] das séries estudadas, e que podem ser vistos na figura 3.

[3] Os gráficos de quantis foram criados através do software SPSS, versão 16.0.

Figura 3: Gráficos de quantis das séries de rendimentos percentuais



Os gráficos de quantis das rendibilidades percentuais dos índices mostram a existência de relações não-lineares entre os quantis teóricos e empíricos, sobretudo nas caudas das distribuições, indicando caudas mais pesadas na distribuição empírica.

Concluímos, então, que tanto os testes de normalidade como os gráficos de quantis rejeitam fortemente a hipótese de normalidade das séries estudadas. A não normalidade e a elevada curtose são um primeiro indício de que as séries são geradas por processos não lineares.

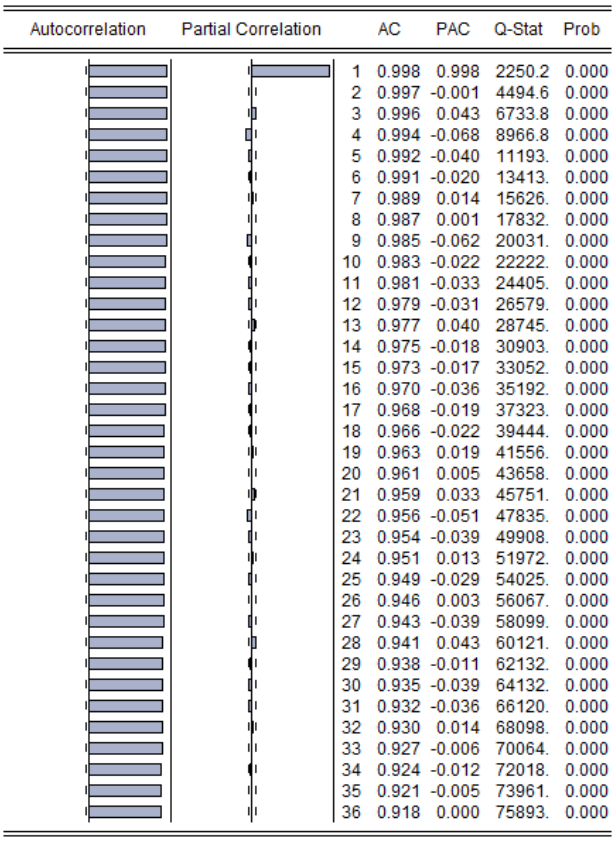
3.3 A Qualidade dos Dados

Com o objectivo de se perceber se um processo estocástico é estacionário, ou ainda para determinar a ordem de integração, pode-se analisar o teste visual ou gráfico das funções de

autocorrelação total (FAC) e parcial (PACF). Os resultados referentes à série do PSI20 são mostrados na figura 4.

Claramente, o processo revela uma convergência muito lenta da função de auto-correlação total, o que indicia a presença de uma raiz unitária, pelo que a série original do índice PSI20 é não estacionária. Observando-se a PACF, conclui-se que a série é caracterizada por um processo auto-regressivo de 1ª ordem, devido ao primeiro coeficiente retardado apresentar um valor muito alto. Para se concluir acerca da estacionaridade das séries, o mesmo tipo de análise poderia ser aplicada às restantes séries. No entanto, dispensamo-nos de apresentar esta análise.

Figura 4: Correlogramas da série PSI20



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.998	0.998	2250.2	0.000
		2	0.997	-0.001	4494.6	0.000
		3	0.996	0.043	6733.8	0.000
		4	0.994	-0.068	8966.8	0.000
		5	0.992	-0.040	11193.	0.000
		6	0.991	-0.020	13413.	0.000
		7	0.989	0.014	15626.	0.000
		8	0.987	0.001	17832.	0.000
		9	0.985	-0.062	20031.	0.000
		10	0.983	-0.022	22222.	0.000
		11	0.981	-0.033	24405.	0.000
		12	0.979	-0.031	26579.	0.000
		13	0.977	0.040	28745.	0.000
		14	0.975	-0.018	30903.	0.000
		15	0.973	-0.017	33052.	0.000
		16	0.970	-0.036	35192.	0.000
		17	0.968	-0.019	37323.	0.000
		18	0.966	-0.022	39444.	0.000
		19	0.963	0.019	41556.	0.000
		20	0.961	0.005	43658.	0.000
		21	0.959	0.033	45751.	0.000
		22	0.956	-0.051	47835.	0.000
		23	0.954	-0.039	49908.	0.000
		24	0.951	0.013	51972.	0.000
		25	0.949	-0.029	54025.	0.000
		26	0.946	0.003	56067.	0.000
		27	0.943	-0.039	58099.	0.000
		28	0.941	0.043	60121.	0.000
		29	0.938	-0.011	62132.	0.000
		30	0.935	-0.039	64132.	0.000
		31	0.932	-0.036	66120.	0.000
		32	0.930	0.014	68098.	0.000
		33	0.927	-0.006	70064.	0.000
		34	0.924	-0.012	72018.	0.000
		35	0.921	-0.005	73961.	0.000
		36	0.918	0.000	75893.	0.000

Na última coluna da figura 4 são apresentados os valores do teste Q de Box-Pierce. Todos os valores são inferiores a 5%, o permite rejeitar a hipótese nula de todos os coeficientes de autocorrelação serem nulos, pelo que com base neste teste se conclui que a série original PSI20 é estacionária. À mesma conclusão se chegaria se este teste fosse aplicado às restantes séries.

Com o objectivo de confirmar esta conclusão, aplicamos o tradicional teste ADF. A hipótese nula do teste estipula que a série tem raiz unitária, ou seja, que a série é integrada de ordem 1, face à hipótese alternativa de a série não possuir raiz unitária. O valor da estatística do teste ADF da série PSI20 (-1.151160) é, em valor absoluto, inferior aos módulos dos valores críticos de MacKinnon para os níveis de significância de 1%, 5% e 10%, respectivamente, -3.433056, -2.862622 e -2.567391, pelo que não se rejeita a hipótese nula de integração da série, concluindo-se que a série é integrada. Confirma-se, assim, a conclusão apresentada atrás. A mesma metodologia poderia ser aplicada às restantes séries, no entanto, dispensamo-nos de a apresentar.

Quanto à série das rendibilidades percentuais do índice PSI20, o valor do teste ADF (-44.60222) é, em valor absoluto, superior aos módulos dos valores críticos de MacKinnon para os níveis de 1%, 5% e 10%, respectivamente, -2.565990, -1.940964, -1.616605, pelo que se rejeita a hipótese nula de integração da série, concluindo-se que a série é estacionária. À mesma

conclusão se chega se o teste ADF for aplicado às restantes séries. No entanto, dispensamo-nos de apresentar os respectivos valores do teste.

3.4 Apreciação da linearidade das séries

Para se detectar a existência de possíveis estruturas não lineares nas séries estudadas dos oito índices, foi aplicado o teste BDS aos resíduos da auto-regressão estimada. Para cada uma das séries escolheu-se a especificação auto-regressiva que apresentou melhor ajustamento. Esta escolha teve em conta os critérios de informação de Akaike e Schwarz.

Os valores de m, no quadro 3, correspondem às dimensões analisadas. Os valores 0.5, 1, 1.5 e 2 correspondem ao valor das distâncias escolhidas, os quais são 0.5, 1, 1.5 e 2 vezes o valor do desvio-padrão da respectiva série.

Como foi descrito no ponto da metodologia, o teste BDS testa a hipótese nula de que uma série tem comportamento independente e identicamente distribuído. Assim, no caso de rejeição da hipótese nula, conclui-se que existe evidência estatística de que a série possui um comportamento não linear, devido à falta de independência nos rendimentos analisados.

Todos os valores do teste BDS, apresentados no quadro 3, são estatisticamente significativos a 1%, pelo que se conclui que as taxas de rendibilidade diárias dos oito índices rejeitam claramente a hipótese IID. Uma vez aplicado o filtro linear, esta decisão deve ficar a dever-se à existência de relações não lineares estatisticamente significativas. A existência deste tipo de relações lineares é, de acordo com Peters (1994), uma indicação de comportamento caótico nas séries de rendimentos dos índices.

Quadro 3: Estatísticas do teste BDS

		BOVESPA	CAC	DAX	DJIA	IBEX35	NIKKEI	PSI20
		AR (1)	AR (1)	AR (2)	AR (1)	AR (1)	AR (2)	AR (2)
m = 2	0,5	3,40388	10,03031	11,79534	8,02790	12,55241	3,96234	12,84726
	1	4,37910	10,36961	12,16129	10,35032	11,59401	4,31100	13,33795
	1,5	6,03499	10,57347	11,97578	12,52425	10,52796	6,35066	13,83376
	2	7,83048	10,96331	11,03457	13,25124	10,34617	8,71418	13,41405
m = 3	0,5	5,59879	15,37726	18,02960	12,75522	19,36038	6,83063	18,33323
	1	6,25125	15,38718	18,15166	14,65938	17,63300	7,50466	17,63484
	1,5	8,01700	15,15545	17,66443	16,50964	15,66233	9,73571	17,56739
	2	10,58498	15,56682	16,49175	17,66522	15,14256	12,09085	16,81722
m = 4	0,5	7,75899	21,39237	24,69665	16,55626	26,51838	10,45792	25,05796
	1	8,21404	19,98774	22,81222	18,04956	22,92101	10,30323	21,21702
	1,5	9,88019	18,57357	21,32435	19,15796	19,18419	11,91864	19,87587
	2	12,39732	18,28594	19,64848	19,90598	17,64062	13,78479	18,73874
m = 5	0,5	9,28782	27,91524	32,49706	21,43690	35,46108	14,65489	34,60909
	1	9,80883	24,05959	26,98778	21,33150	28,29469	12,69223	25,36979
	1,5	11,24022	21,22747	23,94314	20,89922	22,29769	13,36050	21,97466
	2	13,58395	20,25027	21,83793	20,70074	19,78063	14,61521	20,05727
m = 6	0,5	10,86913	36,74529	43,08104	28,02356	47,47566	19,73248	46,47502
	1	11,35074	28,68066	32,10332	25,26901	33,99315	15,46619	29,59582
	1,5	12,54089	23,78355	26,74332	22,96905	24,94941	15,11420	23,71121
	2	14,60019	22,09247	23,92536	21,92743	21,32791	15,82199	20,92484

Notas: m – dimensão de mergulho (*embedding dimension*); ε - distância entre pontos, medida em termos do número de desvios padrão dos dados originais; σ - desvio padrão.

3.5 Apreciação das dependências de longo prazo

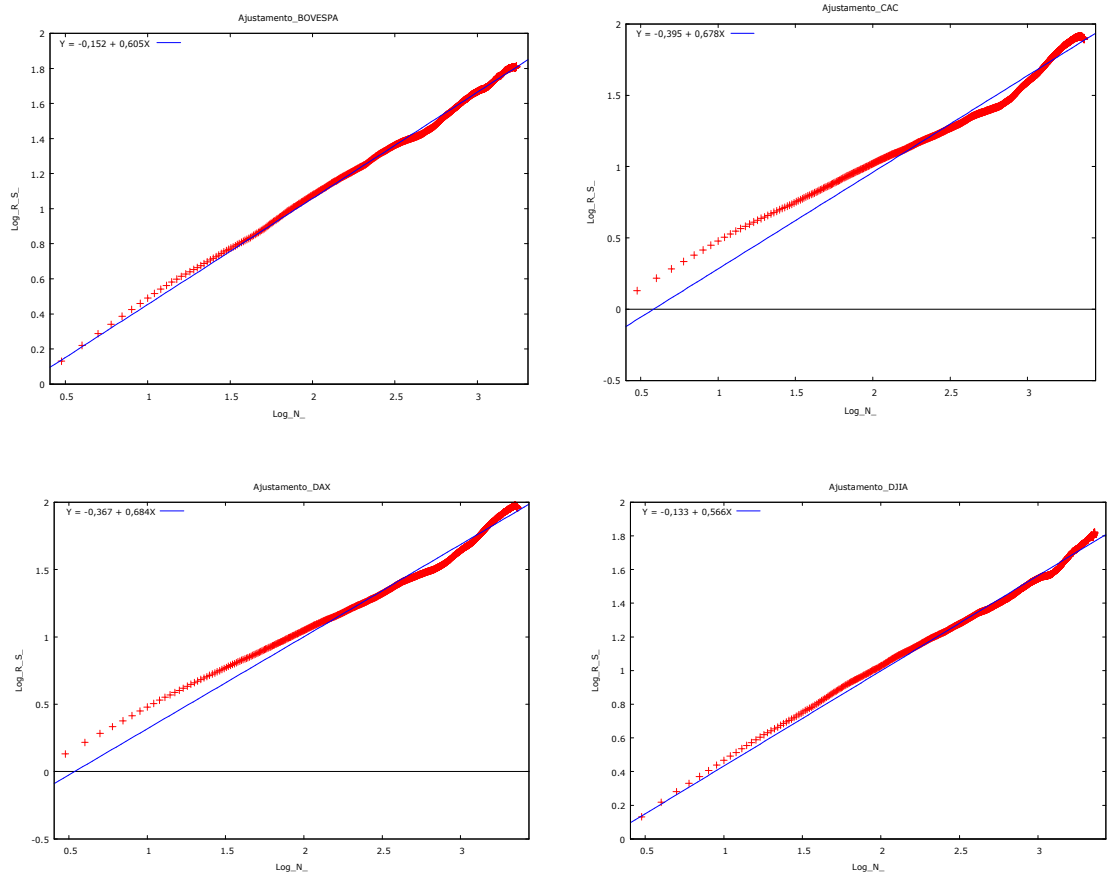
O quadro 4 mostra as estimativas dos coeficientes de Hurst obtidas através da metodologia descrita anteriormente. Por seu lado, a figura 4 apresenta o ajustamento pelo método dos mínimos quadrados aplicado a cada um dos índices, e que permitiu a obtenção dos coeficientes de Hurst.

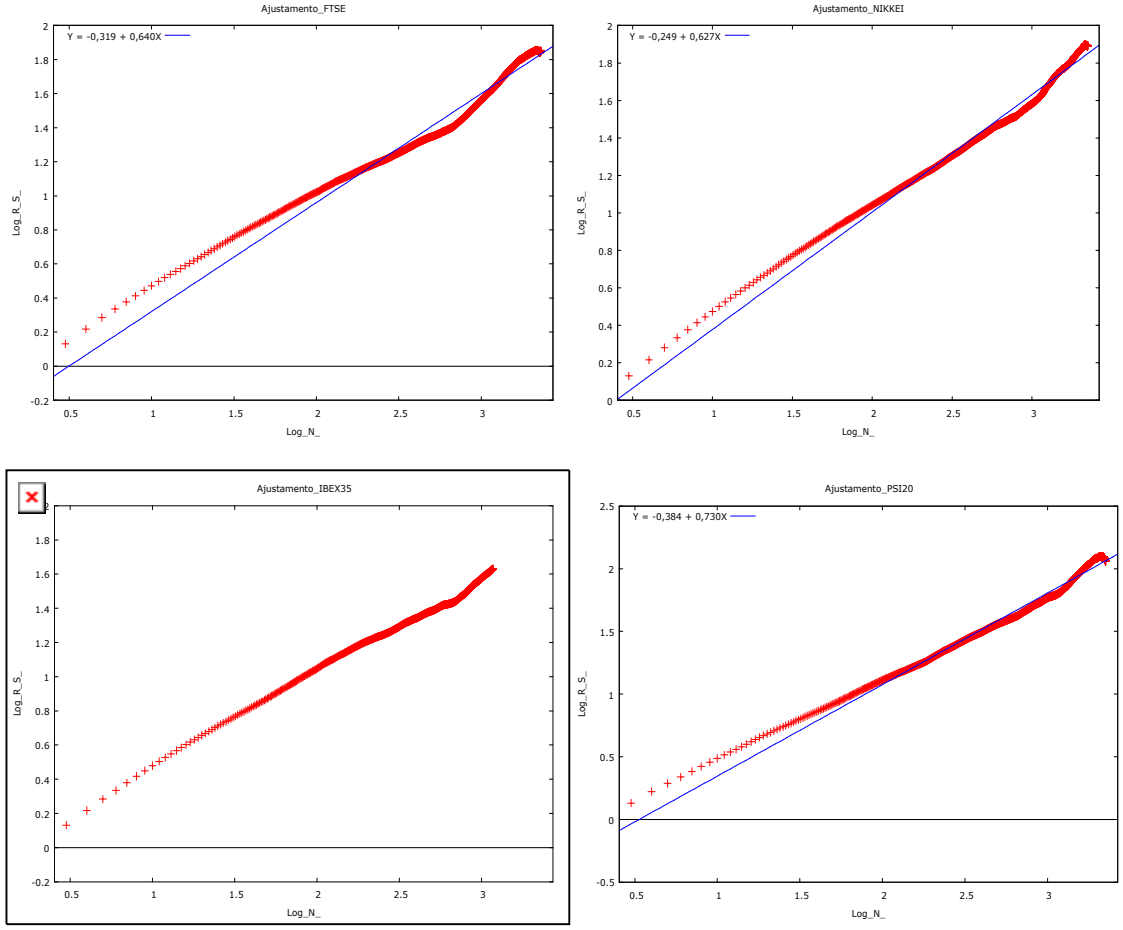
Quadro 4: Estimativa dos coeficientes de Hurst

	Hurst	Estatística-t	R2
BOVESPA	0,6052	685,1726	0,9963
CAC	0,6778	254,3832	0,9653
DAX	0,6842	301,2610	0,9751
DJIA	0,5657	450,2604	0,9888
FTSE	0,6397	262,2442	0,9677
IBEX35	0,6452	244,0212	0,9629
NIKKEI	0,6266	343,1067	0,9814
PSI20	0,7301	350,1392	0,9820

Notas: O cálculo da estatística R/S foi obtido através da macro Kaotixl, que funciona em ambiente Excel, e que permite calcular os valores Log(N) e log(R/S), sendo definidos tantos valores quanto o número de observações da amostra, isto é, para cada incremento de tempo são calculados novos valores para Log(N) e log(R/S). Por seu lado, o coeficiente de Hurst foi obtido a partir de uma regressão linear, apresentada no ponto da metodologia, recorrendo-se para o efeito ao software Eviews, versão 6.0.

Figura 4: Análise R/S para os rendimentos logaritmicos dos índices





Os valores do coeficiente de Hurst (superiores a 0,5 e inferiores a 1) permitem-nos concluir que as séries apresentam persistência (um período de crescimento é seguido por outro análogo). Estes valores são consistentes com os obtidos por Peters (1994) para o mercado norte-americano. Isto significa que há uma probabilidade mais elevada que a um crescimento se siga outro similar, e vice-versa. Por outro lado, pelo facto de as séries temporais serem persistentes, podemos concluir que são caracterizadas por dependências de longo prazo, e que apresentam um comportamento não linear. Teoricamente, o que acontece hoje tem impacto no futuro para sempre. Em termos de dinâmica caótica, uma série com estas características apresenta também elevada sensibilidade às condições iniciais.

4. CONCLUSÕES

A análise desenvolvida no presente trabalho, através da aplicação de vários testes de normalidade às taxas de rendimento dos índices, permite rejeitar fortemente a hipótese de normalidade, o que conjugado com a elevada curtose indicia que as séries são geradas por processos não lineares, confirmando-se deste modo os factos estilizados referidos pelos autores consultados.

Os resultados dos testes dos correlogramas das funções de autocorrelação parcial e total, bem como a aplicação do teste Dickey-Fuller Aumentado (ADF), permitem concluir que as séries originais são não-estacionárias e integradas de ordem 1, enquanto as séries das taxas de rendibilidade são estacionárias e não-integradas.

Por seu lado, a aplicação do teste BDS sugere que os rendimentos dos índices de acções não são estatisticamente independentes e identicamente distribuídos, dando uma primeira indicação de comportamento caótico das séries de rendimentos.

Por fim, os valores do coeficiente de Hurst permitem concluir que todas as séries apresentam persistência e, em consequência, são caracterizadas por dependência.

No conjunto, estas evidências são suficientes para se afirmar que as séries temporais apresentaram, durante o período analisado, um padrão fractal de comportamento

A conjugação da não normalidade, da não linearidade e da evidência de dependências nas séries financeiras, permite perceber que a previsão dos rendimentos dos índices não pode ser feita com base em modelos lineares nem com base em variáveis estatísticas que se distribuam de forma normal. Pelo que se manifesta a necessidade de procurar modelos, que para além de considerarem estas características, incorporem também as dependências encontradas nas referidas séries.

REFERÊNCIAS BIBLIOGRAFICAS

BROCK, W., DECHERT, W. e J. SCHEINKMAN. (1987), "A test for independence based on the correlation dimension", working paper, University of Winconsin at Madison, University of Houston, and University of Chicago.

BROCK, W. e SAYERS, C. (1988), "Is the business cycle characterized by deterministic chaos?", Journal of Monetary Economics, 22, 71-90.

BROCK, W., DECHERT, W., SCHEINKMAN, J., e LeBARON, B. (1991), "A Test for Independence Based on the Correlation Dimension", Social Systems Research Institute, University of Wisconsin, Março.

BROOKS, C. (2002), "Introductory Econometrics for Finance", Cambridge University Press.

LIN, K. (1997), "The ABC's of BDS", *Journal of Computational Intelligence in Finance*, July/August, 23-26.

MAY, C. (1999), "Nonlinear Pricing: theory and applications", New York: Wiley.

CORRAZA, M. E Malliaris, A., "Multi-fractality in Foreign Currency Markets", *Multinational Finance Journal*, 6, 65-98.

GRECH, D. e MAZUR, Z. (2004), "Can one make any crash prediction in finance using the local Hurst exponent idea?", *Physica A: Statistical Mechanics and its Applications*, 336, 133-145.

MANDELBROT, B. (1963), "The variation of certain speculative prices", *The Journal of Business*, 36, 394-419.

MANDELBROT, BENOIT (2006), "O Mau Comportamento dos Mercados: Uma Visão Fractal do Risco, da Ruína e do Rendimento", Gradiva.

PETERS, EDGAR (1991), "Chaos and order in the capital markets: a new view of cycles, prices, and market volatility", New York: Wiley.

PETERS, EDGAR (1994), "Fractal Market Analysis: Applying Chaos Theory to Investment and Economics", New York: Wiley.

TSAY, RUEY (2005), "Analysis of Financial Time Series", New York: Wiley.

O IMPACTO DA VARIÁVEL PÁSCOA NA PREVISÃO DA PROCURA TURÍSTICA

PAULA ODETE FERNANDES
E-mail: pof@ipb.pt
JOÃO PAULO TEIXEIRA
E-mail: joaopt@ipb.pt
Instituto Politécnico de Bragança
JOÃO JOSÉ FERREIRA
E-mail: jjmf@ubi.pt
SUSANA GARRIDO AZEVEDO
E-mail: sazevedo@ubi.pt
Universidade da Beira Interior

Resumo

Com este estudo pretende-se analisar o impacto do feriado móvel da Páscoa na previsão da procura turística, para as regiões Norte e Centro de Portugal. De salientar que a série temporal "Dormidas Mensais registadas nos estabelecimentos hoteleiros", considerada como significativa da actividade turística, devido às suas características, denota que os fenómenos influenciam de forma distinta a procura turística, nas regiões em estudo.
Assim, tendo por base modelos não lineares, sustentados pela metodologia das Redes Neuronais Artificiais (RNA), vai-se verificar se os resultados sofreram alterações significativas antes e após a utilização da variável dummy Páscoa. A inclusão desta nova variável no modelo prende-se com o facto de se ter detectado, em estudos anteriores, alguns valores atípicos na série temporal, dormidas mensais nos estabelecimentos hoteleiros nas regiões em estudo, pelo que se tenciona captar esse efeito.

Palavras-chave: Procura Turística; Redes Neuronais Artificiais; Modelação Não Linear; Previsão.
Área Temática: Modelação Económica e Financeira.

Abstract

This study aims to analyze the impact of the mobile holiday of Easter on the tourist demand' forecast, for the North and Centre regions of Portugal. To point out that the time series "Monthly Guest Nights in Hotels", considered as significant of the tourist activity due to its characteristics, it denotes that the phenomena influence in a different way the tourist demand, in the focused regions.
Thus, based on non-linear models, supported by the Artificial Neural Networks methodology, we intent to compare the results before and after the use of the dummy variable Easter. The inclusion of this new variable in the model is due to the fact that in previous studies on the same region, some atypical values in the time series "Monthly Guest Nights in Hotels" were detected. Being so, this paper intends also to catch this effect.

Key words: Tourism Demand; Artificial Neuronal Network; Non-Linear Modelling; Forecasting.
Thematic Area: Economic and Financial Modelling.

1. INTRODUÇÃO

O estudo da procura turística tem despertado o interesse da comunidade científica e empresarial pela sua importância para a economia mundial. Esta temática foi alvo da atenção da comunidade científica ainda na década de 70 do século XX no trabalho de Archer (1976). Duas décadas mais tarde Theobald (1998) apresentou o turismo como o fenómeno do século XX. A World Tourism Organisation (2000) também descreve o turismo como um dos fenómenos económicos e sociais mais marcantes do último século. Efectivamente, o turismo nas últimas décadas tem crescido consideravelmente no contexto internacional, pois este tem sido

reconhecido como o maior bem/serviço transaccionado no comércio internacional (Goeldner, *et al*., 1999).

Numa perspectiva macro económica este sector tem emergido como um dos mais dinâmicos e líderes das economias de muitos países. Este foi mesmo considerado na década de 90 como a principal fonte de receitas de economias como a Italiana e a Espanhola (Goeldner *et al*., 1999) e como o principal sector empregador nas economias do Canadá, Japão, Reino Unido (Hawkins *et al*., 1992). Ao nível local, o turismo também fornece oportunidades económicas para as comunidades locais beneficiarem de negócios complementares a esta actividade.

Para capitalização de uma forma bem sucedida dos benefícios e minimização dos potenciais impactos negativos do turismo, os decisores associados a este sector devem ter um conhecimento mais profundo das mudanças e desenvolvimentos que afectam as tendências no turismo. Frenchrling (1996), Faulkner e Valério (2000) e Gee e Fayos-Solá (1997) realçam que mais do que em qualquer outro sector de actividade, no turismo deve haver a preocupação com a redução do risco associado a tomadas de decisão. Logo, as actividades de monitorização e previsão da procura são particularmente importantes na gestão turística devido sobretudo ao carácter perecível do produto/serviço turístico e ao facto de a fase de produção e consumo do produto/serviço turístico se verificarem no mesmo momento.

De forma a ir ao encontro destas características únicas dos produto/serviço turísticos, uma resposta rápida deve ser dada pelos decisores mediante a observação dos níveis de oferta e procura para que consigam reagir em sintonia com a situação do mercado. De facto, uma previsão efectiva fornece informação credível e atempada para que os gestores turísticos possam fazer esse balanceamento entre a procura do mercado e a oferta turística. A previsão da procura turística apresenta-se assim estratégica ao permitir obter informação importante para que os decisores turísticos consigam antecipar potenciais crescimentos e riscos num futuro próximo para um país, região ou economia específica (Hu, 2002).

Com o presente estudo pretende-se assim explorar o campo de estudo da previsão da procura turística e evidenciar a utilidade da metodologia das Redes Neuronais Artificiais (RNA) no estudo específico do impacto do feriado móvel da Páscoa na previsão da procura turística, para as regiões Norte e Centro de Portugal.

A presente investigação está estruturada da seguinte forma: depois da presente introdução, são apresentadas, dentro da metodologia, as variáveis utilizadas neste estudo bem como as fontes onde estas foram recolhidas; seguidamente é feita a descrição do Modelo das Redes Neuronais Artificiais. Finalmente após a apresentação e discussão dos resultados obtidos algumas considerações finais são tecidas.

2. METODOLOGIA

2.1. Apresentação das Variáveis

A informação de suporte ao trabalho foi recolhida junto do Instituto Nacional de Estatísticas de Portugal (INE). Optou-se por considerar as séries Dormidas Mensais Registadas na Região Norte [DRN] e as Dormidas Mensais Registadas na Região Centro [DRC], por dois motivos: (*i*) por ser uma série que mede a procura turística efectiva; (*ii*) para comparar com resultados de estudos anteriores, produzidos pelos mesmos autores.

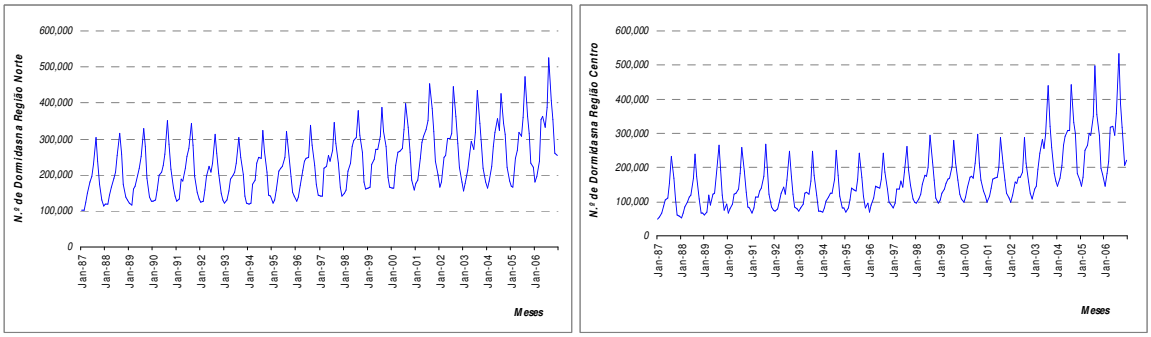
Assim, os dados recolhidos reportam-se ao período compreendido entre Janeiro de 1987 e Dezembro de 2006, correspondendo assim a um total de 240 observações mensais ao longo de 20 anos (ver Anexo A, Tabelas A.1 e A.2).

Pela análise à Figura 1, esta sugere a existência de oscilações irregulares indicando uma não estabilização da média e variância e a presença de sazonalidade (valores máximos nos meses de Verão e mínimos nos meses de Inverno). Ainda, para ambas as séries verifica-se:

- uma tendência crescente nos dois últimos anos, registando-se uma variação positiva, de 2005 para 2006, de 11,8% e de 6,4%, para a região Norte e para a região Centro, respectivamente;

- entre os meses de Março e Maio um pico intermédio que pode estar relacionado com a Páscoa que tradicionalmente atrai muitos turistas para as regiões.

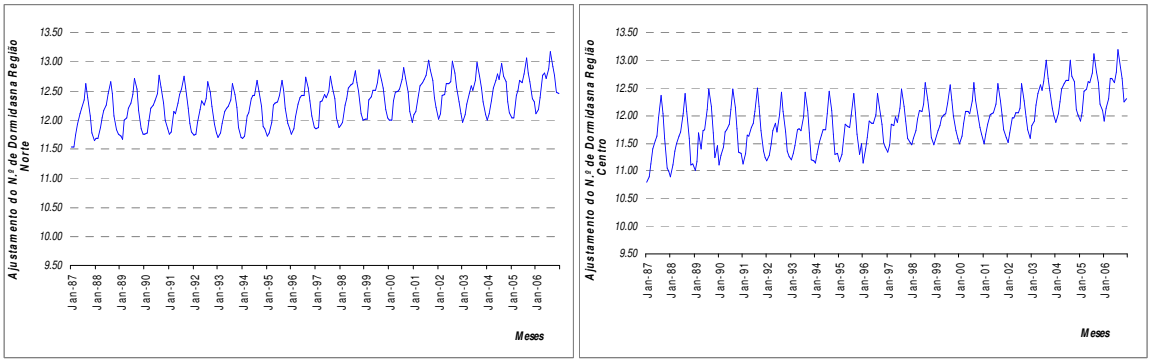
Figura 1. Cronograma das séries, para o período de 1987:01 a 2006:12.



O crescimento registado em 2005 e 2006 para valores nunca antes atingidos acarreta uma dificuldade muito grande para os modelos de previsão que se baseiam apenas no passado, como são os modelos aqui usados, uma vez que o modelo terá de prever um valor que nunca aconteceu.

Para tentar atenuar os efeitos nefastos para a previsão com RNA observados nas séries e no sentido de estabilizar a variância das séries, estas transformaram-se aplicando, a cada uma, o logaritmo natural: LRN e LRC, respectivamente para a região Norte e para a região Centro (Figura 2).

Figura 2. Cronograma das séries transformadas para o período de 1987:01 a 2006:12.



Pela análise da figura anterior pode observar-se que as séries continuam com uma certa estacionaridade que consiste num movimento cíclico, quase periódico, em torno de um valor médio crescente. Contudo, no domínio logarítmico as variações impostas pelo crescimento dos últimos anos aparecem agora menos pronunciadas.

Como se pode observar os resultados tornam-se mais satisfatórios para a região Norte [LRN] do que para a região Centro [LRC]. Assim e no sentido de atenuar o efeito da sazonalidade, ainda patente, foram também processados os dados de entrada pela aplicação das diferenças simples e sazonal, de acordo com a equação às diferenças [1], doravante designada de [LLRC], série 'limpa' (Fernandes & Teixeira, 2008).

$$y(n) = x(n) - x(n-1) + x(n-13) - x(n-12) \qquad [1]$$

Onde:

$x(n)$ é a série a limpar, já no domínio logarítmico;

$y(n)$ é a série limpa.

De referir que para a série LLRC, os valores previstos com a aplicação da metodologia que se propõe neste estudo, Redes Neuronais Artificiais (RNA), são posteriormente processados para inverter o processo de aplicação das diferenças simples e sazonais.

Assim, as novas séries que serviram de base a todo o estudo foram: LRN no domínio logarítmico e LLRC no domínio logarítmico com a aplicação de diferenças, uma simples e outra sazonal.

Ainda, tal como já foi referido, a análise gráfica (Figuras 1 e 2) das séries sugere alguns valores atípicos, ao longo das séries, que podem estar relacionados com época festiva da Páscoa que tradicionalmente gera em simultâneo um fluxo de turismo religioso e um mini período de férias associado às férias escolares. Tenciona-se captar esse efeito, com a inclusão no modelo de uma variável *dummy* [DP_i]. Como se conhece, este fluxo turístico e o período de mini férias ocorre habitualmente na semana anterior ao dia de Páscoa (semana Santa), pelo que se procurou que esta nova variável reflectisse essa realidade usando um peso proporcional aos dias dessa semana em cada mês (normalmente entre os meses de Março e Abril). Assim, de acordo com o Dia da Páscoa (DP) terá o valor apresentado na Tabela 1.

Tabela 1. Valor de DP.

Dia de Páscoa	*Valor de DP no mês da Páscoa*	*Valor de DP no mês anterior*
1	0.1	0.9
2	0.3	0.7
3	0.5	0.5
4	0.6	0.4
5	0.8	0.2
6	0.9	0.1
>6	1	0

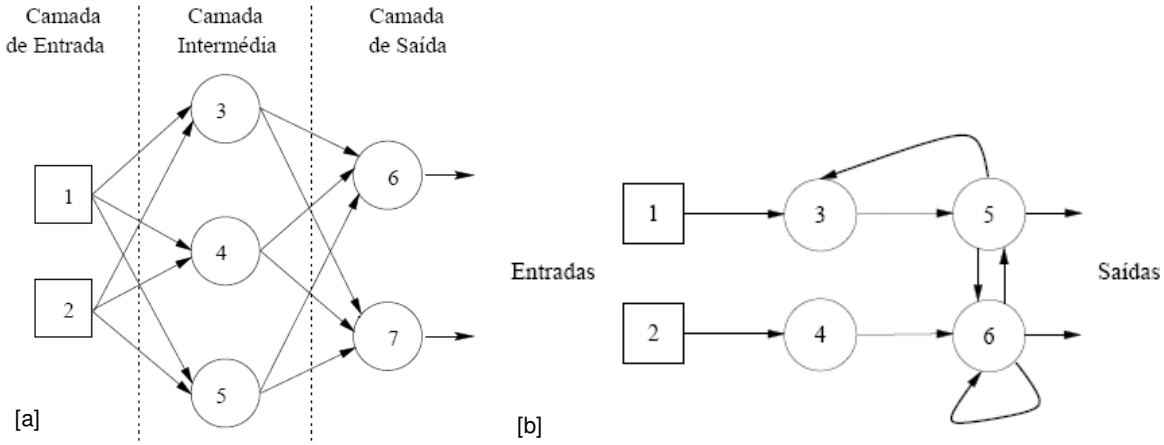
2.2. Redes Neuronais Artificiais

A modelação das séries em estudo vai ter como metodologia de base a utilização das Redes Neuronais Artificiais (RNA). As RNA são modelos que se destacam dentro de uma ampla área do conhecimento da inteligência artificial e têm como suporte modelos matemáticos de arquitectura semelhante à do cérebro humano.

Uma rede neuronal encontra-se composta por um conjunto de neurónios artificiais, nós ligados entre si - as conexões entre os nós denominam-se por *sinapses* -, que processam e transmitem informação através de funções de activação, sendo as mais utilizadas as funções lineares e as sigmoidais - logística e tangente hiperbólica - (Rodrigues, 2000; Cortez, 2002; Fernandes, 2005). Os neurónios de uma rede encontram-se estruturados em distintas camadas (mais conhecidas por camada de entrada, intermédia ou oculta e de saída), sendo as mais comuns, para a previsão de séries temporais, as multi-camadas ou também designadas por MLP[1] (Bishop, 1995; Thawornwong & Enke, 2004), por forma que um neurónio de uma camada encontra-se ligado com os neurónios da camada seguinte às que pode enviar informação, Figura 3. As redes, dependendo da forma como estão ligadas entre as diferentes camadas podem classificar-se em redes não recorrentes[2] (*feedforward*) e redes recorrentes[3] (*feedback*) (Faraway & Chatfield, 1998; Fernandes, 2005).

[1] *Multi-Layer Perceptron*.
[2] A informação flui de forma unidireccional de uma camada para outra, desde a camada de entrada para a camada oculta e posteriormente para a camada de saída (Fernandes, 2005).
[3] As conexões permitem que a informação regresse a lugares por onde já passou e admite-se também ligações intra-camada (laterais) (Fernandes, 2005).

Figura 3. Estrutura de uma Rede Neuronal Artificial: [a] *Feedforward*; [b] F*eedback* (Cortez, 2002).



Da especificação da rede neuronal também faz parte uma função de erro e um algoritmo para determinar o valor dos parâmetros que minimizam a função de erro. Desta forma, existem dois conceitos centrais: a parte física da rede, ou seja, a sua arquitectura, e o procedimento algorítmico que determina o seu funcionamento, ou seja, a forma como a rede se modifica consoante os dados fornecidos pelo ambiente (Haykin, 1999).

Assim, para as RNA aprenderem com a experiência têm que ser submetidas a um processo designado de treino, existindo diferentes algoritmos de treino, sendo um dos mais populares e utilizados na previsão de séries temporais o algoritmo *Backpropagation*[4] ou seus variantes, distribuindo-se os mesmos por duas classes (Gallant, 1993; Haykin, 1999; Fernandes, 2005):

(i) supervisionados, onde no processo de treino existe um 'professor' que fornece um conjunto de casos de treino, e um caso de treino consiste num vector de entrada X e no correspondente vector de saída Y, passando a aprendizagem pela minimização do erro da saída, o que é feito através do ajustamento dos pesos das conexões de acordo com uma determinada regra;

(ii) não supervisionados, onde existe um conjunto de entradas, pelo que o algoritmo de treino tenta agrupar os dados a partir de padrões por estes apresentados, seguindo assim uma regra de auto-organização.

Resumidamente um valor produzido por uma rede *feedforward*, com uma camada oculta, pode ser expressa da seguinte forma (Fernandes & Teixeira, 2007):

$$Y_t = b_{2,1} + \sum_{j=1}^{n} \alpha_j f\left(\sum_{i=1}^{m} \beta_{ij} y_{t-i} + b_{1,j} \right) \qquad [2]$$

Onde,

m, número de nós na camada de entrada;

n, número de nós na camada oculta;

f, função de activação sigmoidal;

$\{\alpha_j, j = 0,1,\ldots,n\}$, vector de pesos que liga os nós da camada oculta aos da camada de saída;

$\{\beta_{ij}, i = 0,1,\ldots,m; j = 1,2,\ldots,n\}$, pesos que relacionam os nós da camada de entrada aos da camada oculta;

[4] Este algoritmo procura o mínimo da função de erro no espaço de procura dos pesos das ligações entre os neurónios, baseando-se em métodos de gradiente descendente. A combinação de pesos que minimiza a função de erro é considerada a solução para o problema de aprendizagem. A descrição do algoritmo pode ser analisada em Rumelhart and McClelland (1986) e em Haykin (1999).

$b_{2,1}$ y $b_{1,j}$, indicam os pesos dos termos independentes (*bias*) associados a cada nó da camada de saída e da camada oculta, respectivamente.

A equação indica também a utilização duma função de activação linear na camada de saída.

2.3. Modelo de Redes Neuronais Artificiais

Para este estudo seleccionou-se um modelo de redes neuronais artificiais, já utilizado em estudos anteriores pelos mesmos autores (Fernandes *et al*., 2008), sendo o mesmo do tipo multi-camada, onde se utilizaram três camadas: de entrada, oculta e de saída, e com uma estrutura tipo *feedforward*. Na camada oculta foi utilizada a função de activação sigmoidal logística [*Logsig*] e na camada de saída utilizou-se a função de activação linear. No treino da rede utilizou-se o algoritmo *Resilient Backpropagation*, uma variante do algoritmo de treino *backpropagation*. A rede, utilizada neste estudo, para cada uma das séries, tem a seguinte arquitectura: 12 nós na camada de entrada, correspondentes aos últimos 12 valores da série, 4 nós na camada oculta e 1 na camada de saída, correspondente à previsão do valor do mês seguinte, ou seja, (1-12;4;1). A estimação/previsão produzida foi efectuada mensalmente. Contudo, para a previsão das observações relativas aos meses de 2005 e 2006 apenas se usaram valores conhecidos até Dezembro de 2004. Para as 12 entradas do modelo correspondentes aos 12 meses anteriores àquele para o qual se está a efectuar a previsão, posteriores a Dezembro de 2004, foram usados os valores previstos em iterações anteriores.

Desta forma, a previsão para 2005 e 2006 é realizada num processo iterativo, usando em cada iteração valores previstos em iterações anteriores. O processo de treino para actualização dos pesos fez-se utilizando o método *batch training*. Foi usado um conjunto de validação para interromper as iterações de aprendizagem quando a *performance* nesse conjunto não melhorava ao fim de 5 iterações sucessivas. Para cada rede realizaram-se 500 sessões de treino, seleccionando os resultados da melhor sessão de treino e escolhendo a RNA com melhores resultados no conjunto de validação. A realização de várias sessões de treinos justifica-se porque em cada sessão de treino os valores iniciais dos pesos são diferentes, chegando-se a soluções também diferentes, podendo ter *performances* significativamente diferentes. O critério que esteve na base para escolher o melhor modelo, para cada uma das séries em análise, foi a raiz do erro quadrático médio (REQM[5]) na comparação dos resultados obtidos pela rede com os valores observados, para o conjunto de validação.

De salientar que para ambas as séries, foi realizado um pré-processamento dos dados de entrada e dos dados de saída correspondente apenas a uma normalização entre -1 e 1.

As séries temporais, obtidas no domínio logarítmico, foram divididas em três conjuntos distintos:

1º - conjunto de teste, para ambas as séries, utilizaram-se as últimas 24 observações, correspondentes aos anos de 2005 e 2006;

2º - conjunto de validação, para a série LRN as 12 observações, correspondentes ao ano de 2004; para a série LLRC as observações: Janeiro de 1999, Fevereiro de 2004, Março de 2002, Abril de 1996, Maio de 2003, Junho de 2000, Julho de 1998, Agosto de 2004, Setembro de 1997, Outubro de 2001, Novembro de 1994 e Dezembro de 2003;

Por observação da série LLRC da Figura 2, verifica-se, apesar da passagem para o domínio logarítmico, que os valores dos últimos 4 ciclos, correspondentes aos anos de 2003 a 2006, assumem valores que se situam a um nível mais elevado. Sendo um dos preceitos para a utilização das RNA que os conjuntos de treino, validação e teste sejam uma boa representação do fenómeno a prever, estaríamos a desviar-nos muito deste preceito se os conjuntos de validação fossem as observações correspondentes ao ano de 2004, já que o

5

$$REQM = \sqrt{\frac{\sum_{t=1}^{n}\left(A_t - P_t\right)^2}{n}};\ com: A_t, valores\ reais\ no\ período\ t; P_t, valores\ previstos\ no\ período\ t; n, total\ de\ observações\ utilizadas.$$

teste têm obrigatoriamente que ser 2005 e 2006, restando para o treino as observações até Dezembro de 2003. Desta forma o conjunto de treino só disporia de 1 ano com os valores elevados e cerca de 15 anos com valores bastante inferiores. Por outro lado, o conjunto de validação só disporia de observações correspondentes aos anos com valores mais elevados. Com a distribuição do conjunto de validação usando os 12 meses do ano distribuídos por vários anos, estamos a usar mais observações com valores elevados no conjunto de treino equilibrando melhor estes dois conjuntos.

O conjunto de validação é usado como uma técnica alternativa de paragem do treino da rede para evitar uma sobre adequação ao conjunto de treino - '*overfitting*';

3º - conjunto de treino:

(i) Para a série LRN corresponde aos meses de Janeiro de 1988 a Dezembro de 2003, num total de 192 observações. Note-se que a primeira observação que este modelo pode usar é a de Janeiro de 1988, porque necessita das 12 observações anteriores na sua entrada;

(ii) Para LLRC são usadas 179 observações correspondentes aos meses entre Fevereiro de 1989 e Dezembro de 2004exceptuando os meses do conjunto de validação. Note-se que esta série dispõe de menos 13 observações que a anterior. Estas 13 observações são 'consumidas' no processo de limpar a série descrito na equação [1], uma vez que *n* começa em 14 para que $x(n-13)$ seja o primeiro mês (Janeiro de 1987).

3. APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Nesta secção pretende-se verificar se os resultados sofrem alterações antes e após a inclusão da variável *dummy* Páscoa [DP_i] no modelo, para ambas as regiões em estudo.

Para tal vão ser analisados os resultados para o conjunto de teste (anos 2005 e 2006), comparando os valores observados com os valores previstos, para as duas séries e utilizando a metodologia das redes neuronais artificiais, e não esquecendo o processo inverso das transformações efectuadas para produzir os valores previstos (tabela 2).

Como critério de análise da qualidade de ajustamento dos valores previstos para cada uma das situações teve-se por base a classificação do MAPE[6], ou EPAM, proposto por Lewis (1982), dado pela equação:

$$EPAM = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_t - P_t}{Y_t}\right|;\ Y_t, valor\ observado\ e\ P_t, valor\ previsto. \qquad [3]$$

Tabela 2. Avaliação da *Performance* com os modelos das RNA.

Destino Turístico	Conjunto de Teste	*r*		*EPAM*	
		Sem DP_i	Com DP_i	Sem DP_i	Com DP_i
$LRN_{RNA(1\text{-}12;4;1)}$	**2005**	0.967	0.979	6.89%	5.30%
	2006	0.962	0.953	8.70%	12.30%
$LLRC_{RNA(1\text{-}12;4;1)}$	**2005**	0.963	0.979	7.44%	5.27%
	2006	0.956	0.979	8.17%	5.71%

Fazendo uma leitura à Tabela 2, pode observar-se para as duas regiões que os modelos, antes e após a inclusão da variável *dummy*, apresentaram qualidades estatísticas e de ajustamento satisfatórias, evidenciando serem adequados para modelar e prever a procura turística das regiões em estudo. Ainda, observa-se que os modelos conseguem produzir resultados mais satisfatórios no curto prazo, neste caso para um ano, como seria de esperar, uma

[6] *Mean Absolute Percentage Error.*

vez que para prever um mês à frente entra-se em linha de conta com os meses anteriores e na previsão dos meses de 2006, o modelo tem na entrada apenas valores previstos em iterações anteriores para os 12 meses precedentes, que incorporam um desvio em relação à realidade e por conseguinte podem levar a um desvio significativo em relação às previsões para 2006.

Para a região Norte verifica-se, para o ano de 2005, uma melhoria de aproximadamente 2 pontos percentuais quando se utiliza a variável *dummy*, reflectindo-se também essa melhoria quando analisado o coeficiente de correlação. O mesmo já não se verifica para o ano de 2006, em que a inclusão da variável *dummy* penaliza os resultados em quase 4 pontos percentuais. Esta situação pode ter a ver com o facto de terem ocorrido outros factores que influenciam a série e que o modelo não considera nas suas entradas.

Relativamente à região Centro, para ambos os anos, incluir a variável *dummy* no modelo permite produzir melhorias significativas que rondam os 2 pontos percentuais, o que se traduz em resultados satisfatórios.

Contudo pela análise efectuada anteriormente e tendo por base os critérios estabelecidos por Lewis (1982), poder-se-á dizer que para as regiões em estudo, os modelos conseguiram produzir previsões de precisão elevada, sem a variável *dummy*, uma vez que o EPAM apresenta valores inferiores a 10%. Com a variável *dummy* introduzida no modelo, apenas para 2006 e para a região Norte o modelo consegui fazer previsões de boa precisão, uma vez que apresentou um EPAM entre 10 e 20%, tendo obtido previsões de elevada precisão para o ano de 2005, para ambas as regiões.

Na Tabela 3 apresenta-se o valor do erro percentual absoluto (EPA) para os meses da Páscoa nos anos de 2005 e 2006, respectivamente Março e Abril, para os modelos com e sem a variável DP, nas regiões Norte e Centro.

Tabela 3. EPA nos meses da Páscoa.

Conjunto de Teste	Região Norte		Região Centro	
	Sem DP_i	Com DP_i	Sem DP_i	Com DP_i
2005	10.7%	0.6%	9.9%	18.4%
2006	18.9%	23.5%	1.9%	3.2%

Pela análise da Tabela 3 constata-se que a inclusão da variável *dummy* também não trouxe melhorias de previsão para os meses particulares em que se esperaria que esta variável actuasse.

4. CONSIDERAÇÕES FINAIS

O turismo tem sido apresentada como uma actividade geradora de riqueza e cada dia que passa tem assumido um papel significativo na economia de um país, Portugal também não fica alheio a este facto. Deste modo, também as organizações, públicas ou privadas, particularmente ligadas ao sector turístico, implementadas nas regiões em estudo (Regiões Norte e Centro de Portugal) devem preocupar-se em construir mecanismos que lhes permitam obter uma visão antecipada da evolução da procura turística, no sentido de criar condições favoráveis aos visitantes que se deslocam a esses destinos turísticos.

Pretendia-se com esta investigação analisar o impacto que o feriado móvel da Páscoa tem na previsão da procura turística, tendo por base modelos não-lineares, sustentados pelas metodologias das RNA, para as respectivas regiões Norte e Centro de Portugal. Pretendeu-se assim verificar se os resultados sofreram alterações significativas antes e após a utilização da variável *dummy*.

Em conclusão, os resultados apresentados pelo modelo com a variável *dummy* Páscoa, não resultam numa melhoria significativa em relação ao modelo sem esta variável. A pequena variação observada (ver tabela 2) encontra-se dentro de uma pequena variação existente entre a escolha do melhor modelo (diferentes sessões de treino), pelo que a eventual melhoria observada não se pode considerar significativa. Além do mais os resultados para os meses da Páscoa do modelo com e sem a variável *dummy* mostram que a inclusão desta nova variável também não traz melhorias para os meses em que esta intervém. Assim, pode concluir-se a irrelevância desta variável para o modelo.

5. REFERÊNCIAS BIBLIOGRÁFICAS

BISHOP, C. M. (1995). "*Neural Networks for pattern recognition*". Oxford University Press. Oxford. London.

CORTEZ, P. (2002). "*Modelos Inspirados na Natureza para a Previsão de Séries Temporais"*. Tese de Doutoramento, Universidade do Minho, Portugal.

FARAWAY, J. e CHATFIELD, C. (1998). "Time series forecasting with neural networks: a comparative study using the airline data". *Applied Statistics*. N.º47, pp.231/250.

FAULKNER, B. and VALERIO, P. (2000). "*An integrative approach to tourism demand forecasting*". In C. Ryan and S. Page (Eds), "Tourism Management: Toward the new millennium". Oxford, UK: Perganon, Elsevier Science, Ltd.

FERNANDES, P. O. (2005). "*Modelling, Prediction and Behaviour Analysis of Tourism Demand in the North of Portugal*". Ph.D. Thesis in Applied Economy and Regional Analysis; Valladolid University - Spain.

FERNANDES, P. O. e TEIXEIRA, J. P. (2007). "A new approach to modelling and forecasting monthly overnights in the Northern Region of Portugal". *Proceedings of the 4th International Finance Conference.* Université de Cergy; Hammamet, Medina, Tunísia.

FERNANDES, P. O. e TEIXEIRA, JOÃO P. (2008). "Previsão da Série Temporal Turismo com Redes Neuronais Artificiais"; 5º Congresso Internacional Luso-Moçambicano de Engenharia -*CLME'2008-"A Engenharia no Combate à Pobreza, pelo Desenvolvimento e Competitividade*". Maputo-Moçambique.

FERNANDES, P. O.; TEIXEIRA, J. P.; FERREIRA, J. e AZEVEDO, S. (2008). "Modelling Tourism Demand: A Comparative Study between Artificial Neural Networks and the Box-Jenkins Methodology". *Romanian Journal of Economic Forecasting*, No 5(3), pp. 30/50.

FRECHTLING, D. (2001). "*Forecasting tourism demand: Methods and strategies*". Oxford, UK:Butterworth-Heinemann.

GALLANT, S. (1993). "*Neuronal Network Learning and Expert Systems*". MIT Press; Cambridge; USA.

GEE, C. and FAYOS-SOLÁ, E. (1997). "*International Tourism: A Global Perspective*". Madrid, Spain: World Tourism Organization.

GOELDNER, C., RITCHIE, B. and MCINTOSH, R. (1999). "*Tourism: Principles, practices, philosophies*". 8th. New York, NY: John Wiley & Sons.

HAWKINS, D. and RITCHIE, B. (1992). "*World travel and tourism review: indicators, trends and forecasts*". Wallingford, Oxon, UK:C.A.B. International.

HAYKIN, S. (1999). "*Neural Networks. A comprehensive foundation*". New Jersey, Prentice Hall.

HU, C. (2002). "*Advanced tourism demand forecasting: artificial Neural Network and Box-Jenkins Modeling*", Ph.D Thesis, Purdue University.

INE. "*Anuários Estatísticos do Turismo de 1987 e 2006"*. Lisboa.

LEWIS, C.D. (1982). "*Industrial and Business Forecasting Method"*. Butterworth Scientific. London.

RODRIGUES, S. (2000). "*Redes neuronais aplicadas à segmentação e classificação de leucócitos em imagens"*. Dissertação de Mestrado em Engenharia Electrónica e Telecomunicações. Universidade de Aveiro.

RUMELHART, D. E. e MCCLELLAND, J. L. (1986). "*Parallel Distributed Processing: Explorations in the Microstructure of Cognition*". Volume 1: Foundations. The Massachusetts Institute of Technology Press, Cambridge.

THAWORNWONG, S. e ENKE, D. (2004). "The adaptive selection of financial and economic variables for use with artificial neural networks". *Neurocomputing*. N.º6, pp. 205/232.

THEOBALD, W. (1998). "*Global tourism*". 2nd ed. Oxford. UK:Butterworth-Heinemann.

WORLD TOURISM ORGANIZATION. (1978-1981). "*World tourism statistics: Statistiques du tourism*". Madrid, Spain.

ANEXO

Tabela A.1. Valores da Série Original, para o período entre 1987:01 e 1996:12, Região Norte (cont.).

ANOS / MESES	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
Janeiro	102.447	118.011	122.217	126.671	126.826	124.194	121.469	118.606	122.480	126.910
Fevereiro	102.123	117.547	116.837	129.802	131.653	127.474	129.284	122.988	130.393	139.403
Março	125.401	142.687	160.658	158.701	188.999	157.536	154.734	175.261	156.645	172.393
Abril	150.042	167.118	169.326	197.757	182.290	196.087	189.142	185.525	209.263	213.973
Maio	180.430	189.823	199.158	207.876	219.187	223.918	198.402	232.075	218.666	239.142
Junho	197.113	207.729	218.595	227.159	251.295	207.907	207.216	248.237	222.720	245.264
Julho	229.293	254.523	252.634	257.633	273.927	231.801	231.453	246.274	247.589	248.398
Agosto	304.847	315.113	329.014	351.500	341.490	312.026	304.576	322.366	320.750	336.086
Setembro	238.542	258.287	278.074	284.867	283.378	259.023	249.583	266.094	269.433	280.769
Outubro	173.503	174.359	189.664	216.286	197.241	205.400	202.792	206.256	196.466	225.734
Novembro	130.187	137.933	138.683	162.062	152.554	149.289	141.976	144.803	152.340	175.438
Dezembro	114.229	128.774	127.730	139.683	132.802	130.963	120.748	139.706	140.643	143.163
TOTAL	***2.048.157***	***2.211.904***	***2.302.590***	***2.459.997***	***2.481.642***	***2.325.618***	***2.251.375***	***2.408.191***	***2.387.388***	***2.546.673***

Fonte: Elaboração própria com base nos dados facultados pelo INE.

Tabela A.1. Valores da Série Original, para o período entre 1997:01 e 2006:12, Região Norte.

ANOS / MESES	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Janeiro	140.430	148.218	163.696	162.389	176.690	165.653	155.527	162.900	168.100	180.700
Fevereiro	141.183	157.415	165.988	162.637	186.586	181.005	177.818	181.900	166.800	195.100
Março	219.465	209.929	228.149	226.010	245.261	249.214	214.106	224.600	247.000	237.200
Abril	224.382	232.767	242.744	262.865	291.395	253.274	258.519	279.800	268.500	352.600
Maio	253.833	280.326	269.854	264.497	306.743	302.028	293.531	317.300	316.900	361.200
Junho	238.334	296.612	270.126	273.881	325.568	301.465	271.454	355.300	307.700	331.500
Julho	266.993	303.866	306.031	324.962	351.955	314.560	318.706	324.400	358.500	388.400
Agosto	345.672	377.645	385.868	397.405	452.581	444.991	433.211	426.900	472.400	524.500
Setembro	288.409	309.700	321.248	331.155	383.793	361.181	343.534	342.100	362.200	406.500
Outubro	232.052	263.522	280.597	263.217	319.417	287.383	281.472	311.500	315.900	353.300
Novembro	166.835	180.796	193.062	186.445	238.925	221.910	219.463	221.200	233.400	258.800
Dezembro	141.349	161.273	166.990	157.210	202.351	179.766	178.439	182.800	221.300	254.700
TOTAL	***2.658.937***	***2.922.069***	***2.994.353***	***3.012.673***	***3.481.265***	***3.262.430***	***3.145.780***	***3.330.700***	***3.438.700***	***3.844.500***

Fonte: Elaboração própria com base nos dados facultados pelo INE.

Tabela A.2. Valores da Série Original, para o período entre 1987:01 e 1996:12, Região Centro (cont.).

ANOS / MESES	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
Janeiro	48.413	53.251	60.593	66.389	67.712	72.006	73.457	69.142	70.798	69.186
Fevereiro	53.932	66.257	70.923	78.898	81.963	78.873	82.466	80.463	81.326	89.418
Março	67.949	84.982	118.949	91.836	114.931	98.200	93.210	101.582	104.727	110.697
Abril	88.730	97.751	88.999	121.039	112.756	124.425	125.441	113.765	139.292	145.682
Maio	103.595	112.881	122.323	125.580	130.316	141.334	127.772	125.687	133.419	142.172
Junho	111.331	120.029	126.325	138.110	140.715	121.020	122.687	125.656	130.530	141.044
Julho	154.594	167.631	182.117	183.161	175.843	163.168	158.791	166.728	164.749	166.283
Agosto	233.117	240.183	263.974	259.879	267.754	247.192	247.527	250.555	242.433	241.940
Setembro	168.602	176.127	190.951	190.030	193.701	175.842	176.980	177.707	171.988	187.513
Outubro	106.730	107.174	118.864	127.891	123.425	121.295	118.980	116.944	116.247	137.972
Novembro	62.249	67.058	75.367	83.646	85.675	84.867	72.739	80.985	80.925	100.324
Dezembro	58.618	67.540	94.352	82.305	76.662	78.134	72.227	81.664	97.189	93.096
TOTAL	***1.257.860***	***1.360.864***	***1.513.737***	***1.548.764***	***1.571.453***	***1.506.356***	***1.472.277***	***1.490.878***	***1.533.623***	***1.625.327***

Fonte: Elaboração própria com base nos dados facultados pelo INE.

Tabela A.2. Valores da Série Original, para o período entre 1997:01 e 2006:12, Região Centro.

ANOS / MESES	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Janeiro	82.964	95.078	105.697	97.748	97.835	99.913	136.669	144.481	146.800	146.300
Fevereiro	95.439	106.779	123.941	112.210	117.057	118.807	146.512	169.494	172.000	184.500
Março	137.757	122.126	136.214	141.973	138.851	156.803	196.309	206.316	251.100	219.800
Abril	136.194	151.959	155.533	173.166	164.615	154.440	240.487	263.603	264.200	317.200
Maio	159.817	176.390	165.865	173.781	168.582	172.775	282.940	290.185	299.900	320.500
Junho	144.019	173.863	169.182	167.906	171.690	172.701	256.314	308.510	293.000	294.000
Julho	185.696	200.270	203.694	211.569	200.343	185.184	297.678	308.175	348.200	358.000
Agosto	262.815	294.081	280.780	296.264	287.122	288.336	439.293	442.413	496.700	534.200
Setembro	193.321	216.871	214.071	213.978	211.241	211.734	319.576	331.474	353.900	388.300
Outubro	147.357	162.655	161.856	162.932	163.283	158.020	257.783	300.534	294.200	316.100
Novembro	107.827	109.382	122.468	131.786	125.344	125.915	183.431	182.155	198.200	208.200
Dezembro	100.364	96.465	108.546	116.821	110.652	108.691	161.020	163.759	179.200	221.000
TOTAL	***1.753.570***	***1.905.919***	***1.947.847***	***2.000.134***	***1.956.615***	***1.953.319***	***2.918.012***	***3.111.099***	***3.297.400***	***3.508.100***

Fonte: Elaboração própria com base nos dados facultados pelo INE.

EFEITOS DO INVESTIMENTO DE CARTEIRA ESTRANGEIRO: EVIDÊNCIA PARA O MERCADO PORTUGUÊS

JOÃO DIONÍSIO MONTEIRO
E-mail: jdm@ubi.pt
JOSÉ RAMOS PIRES MANSO
E-mail: pmanso@ubi.pt
Departamento de Gestão e Economia
Universidade da Beira Interior

Resumo

Este trabalho tem por objectivo analisar os efeitos do investimento de carteira estrangeiro no mercado accionista português. Examinamos as causas da relação positiva contemporânea entre a rendibilidade do mercado e os fluxos de investimento estrangeiros. Subjacentes a esta relação positiva contemporânea poderão estar várias hipóteses que explicam esta relação, nomeadamente, a hipótese de alargamento da base de investidores, ou de diminuição do custo do capital local, a hipótese de pressão nos preços e a hipótese de revelação de informação pelos fluxos estrangeiros. Identificamos quais destas hipóteses melhor explicam estas relações contemporâneas.
Os resultados sugerem que os fluxos de carteira estrangeiros têm um impacto positivo e significativo nos preços locais. Os fluxos líquidos estrangeiros não antecipados, equivalentes a um por cento da capitalização accionista local, estão associados a um aumento contemporâneo nos preços locais de cerca de três por cento. No teste à hipótese de pressão nos preços pelos fluxos estrangeiros, os resultados sugerem que a relação contemporânea positiva entre fluxos e rendibilidades não se reverte nos meses seguintes. Pelo contrário, os resultados obtidos tendem a suportar a hipótese de alargamento da base de investidores, ou seja, a diminuição do custo do capital local: o aumento da partilha do risco e maior liquidez no mercado, motivado pelo investimento estrangeiro, tende a mover os preços para um nível de equilíbrio superior permanente.

Palavras-chave: Fluxos de investimento estrangeiro, diminuição do custo do capital, pressão nos preços, revelação de informação pelos fluxos.
Área temática: Modelação Económica e Financeira.

Abstract

The purpose of this paper is to analyze the effects of the foreign portfolio equity flows into the Portuguese stock market. Hence, this paper examines the causes of the positive contemporaneous relationships between the Portuguese stock market returns and the foreign portfolio equity flows. Underlying these contemporaneous relationships there might be various hypotheses that could explain this relation, namely, the investors' base broadening hypothesis, or the decrease of the cost of capital, the price pressure hypothesis and the flows information revelation hypothesis. We will try to uncover which one of these hypotheses is underlying and best explains these contemporaneous relationships.
In general, the results suggest that foreign portfolio equity flows have a positive significant impact on local prices. In the Portuguese stock market, non-anticipated net equity flows by foreign investors, equivalent at one percent of local market capitalization, are associated to a contemporaneous increase on local prices of about three percent. Testing the price pressure hypothesis by foreign flows, the results suggest that the positive contemporaneous relationship between flows and returns do not revert in the months ahead. On the contrary, the results obtained tend to support the investors´ base broadening hypothesis, ie, the decrease in the cost of local capital: a greater risk sharing and a higher liquidity on the market, motivated by the foreign investment, tend to move the prices to a higher permanent equilibrium level.

Key words: Foreign portfolio equity flows, Decrease on the cost of capital, Price pressure, Flows information revelation.
Thematic area: Economic and Financial Modelling.

1. INTRODUÇÃO

Na última década os investidores reconheceram os benefícios da diversificação internacional de carteiras. Portugal recebeu ao longo da última década montantes consideráveis de fluxos de investimento de carteira em acções. Os investidores estrangeiros reconheceram o mercado português como vantajoso para melhorar o *trade-off* risco - rendibilidade esperada das suas carteiras de investimento. A literatura sobre os fluxos de carteira internacionais e sobre a segmentação parcial dos mercados sugere que alterações significativas na base dos investidores activos nos mercados accionistas locais devem ter efeitos significativos nos preços locais. Em particular, devido à maior partilha do risco local e aumento de liquidez, as rendibilidades esperadas devem diminuir e os preços devem aumentar.

Neste artigo examinamos as causas da correlação positiva contemporânea entre a rendibilidade do mercado português e os fluxos líquidos estrangeiros. Para levar a cabo este objectivo baseamo-nos em várias hipóteses alternativas que podem explicar a relação contemporânea. Os resultados obtidos tendem a suportar a hipótese de diminuição do custo de capital no mercado português como o efeito explicativo da correlação contemporânea entre a rendibilidade local e fluxos estrangeiros.

O artigo está organizado do seguinte modo. Na secção 2 apresentamos a metodologia econométrica utilizada para quantificar o efeito associado a cada uma das hipóteses alternativas e enunciamos os correspondentes testes de hipóteses associados a cada uma. A metodologia usada é baseada na metodologia proposta por Warther (1995), também usada por Clark e Berko (1997). Consideramos a hipótese de diminuição do custo do capital do mercado português ou aumento da base de investidores, a hipótese de pressão nos preços locais resultantes de iliquidez temporária do mercado e a hipótese de revelação de informação pelos fluxos. Na secção 3 apresentamos os resultados das hipóteses e discutimos os resultados. Na secção 4 apresentamos as conclusões.

2 ESPECIFICAÇÃO DAS HIPÓTESES

2.1 Componentes esperados e não-esperados das aquisições líquidas

A teoria do mercado eficiente implica que a informação relevante disponível no início de um período deve ser prontamente reflectida no preço dos activos no início desse período. Assim, se é esperado que a procura estrangeira ou doméstica (investidores institucionais) tenha um impacto nos preços, levando-os para um nível de equilíbrio superior, mas os estrangeiros ou domésticos apenas investem gradualmente, os preços devem reflectir esta informação e aumentar em antecipação dos influxos efectivos.

Baseando-nos na ideia sugerida por Warther (1995), e também utilizada por Clark e Berko (1997), se os investidores domésticos estão incertos quanto à magnitude da nova procura estrangeira (vice-versa) por activos locais, a chegada de nova informação, ao levar os investidores a aumentar a sua estimativa dos influxos totais estrangeiros ou domésticos futuros, deverá impulsionar os preços para um nível superior[1].

Estas considerações, subjacentes à hipótese do mercado eficiente, sugerem que em vez de regredir as rendibilidades sobre as variações efectivas na base de investidores, dever-se-á regredir as rendibilidades sobre as revisões de expectativas relativas à evolução da base de investidores. Estas revisões de expectativas não são directamente observáveis. Porém, assumindo que os investidores prevêem a evolução futura da base de investidores analisando os fluxos observados, estes estimariam a posição de investimento dos investidores do seguinte modo:

[1] Merton (1987) reconhece este aspecto, e nota que a questão teria de ser analisada construindo uma versão dinâmica do seu modelo. Especificamente refere que (p.500)"…se um evento favorável implica uma revisão em alta nestas antecipações (isto é, expectativas sobre a evolução futura da dimensão da base de investidores), então o preço deve aumentar imediatamente, mesmo se existe um desfasamento antes dos novos investidores informados tomarem posições. De igual modo, um evento desfavorável implicando uma redução no crescimento antecipado na base de investidores deverá causar uma diminuição imediata no preço".

$$E_{t-1}(I_t^i) = F(a_1, a_2, ..., a_n, Al_{t-1}, ..., Al_{t-n}, \mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_m, \mathbf{y}_{t-1}, ..., \mathbf{y}_{t-m}) \quad [1]$$

onde I_t^i representa a posição de investimento dos investidores estrangeiros ou domésticos no mercado português no período t, Al_t representa as aquisições líquidas no período t, $F(.)$ é a função de previsão, a_j e b_j são coeficientes dos processos de desfasagens distribuídas $A(L)$ e $B(L)$ da série temporal para as aquisições líquidas:

$$Al_t = A(L)Al_{t-1} + B(L)\mathbf{y}_{t-1} + \varepsilon_t \quad [2]$$

e $\mathbf{y}_{t-1}$ seria um eventual vector de outras variáveis úteis para prever as aquisições líquidas. O componente não antecipado da série das aquisições líquidas, representando as inovações, serve como uma medida representativa das revisões das expectativas da evolução temporal dos fluxos líquidos. Tal como procedeu Warther (1995), definimos as aquisições líquidas não-esperadas, $U(Al_t)$, como sendo iguais às aquisições líquidas observadas menos as expectativas (valor esperado) dos investidores sobre as aquisições líquidas futuras, $E_{t-1}(Al_t)$, ou seja,

$$U(Al_t) \equiv Al_t - E_{t-1}(Al_t). \quad [3]$$

2.2 Hipótese de diminuição do custo do capital

Merton (1987) refere o resultado de que quando a base de investidores aumenta, ou seja, quando $\partial q_k > 0$, as rendibilidades requeridas de equilíbrio (esperadas) diminuem e os preços aumentam, ou seja:

$$\frac{\partial E(r_k)}{\partial q_k} = -\frac{E(r_k)\delta\sigma_k^2 x_k}{q_k^2} < 0 \quad [4]$$

$$\frac{\partial P_k}{\partial q_k} = \frac{P_k\delta\sigma_k^2 x_k}{q_k^2} > 0 \quad [5]$$

$$\frac{\partial V_k}{\partial q_k} = \frac{V_k\delta\sigma_k^2 x_k}{q_k^2} > 0 \quad [6]$$

Ceteris paribus, da primeira e terceira expressão, sai que um aumento na dimensão relativa da base de investidores da empresa, e por extensão, do mercado, reduz o custo do capital e aumenta o valor de mercado da empresa e, por extensão, do conjunto de empresas cotadas. Assim, segundo Merton (1987), os gestores das empresas têm um incentivo para expandir a sua base de investidores. A magnitude deste efeito será maior para as empresas menos conhecidas (com q_k reduzido) e para as empresas com variâncias de rendibilidades específicas elevadas.

Se considerarmos o mercado português como um activo financeiro da carteira mundial, a variável fundamental do modelo de Merton (1987), $q_k (\equiv N_k / N)$, mediria o rácio do número de investidores que estão "informados" sobre as acções portuguesas em relação ao número total de investidores, que pode ser definido pela expressão:

$$q_t \equiv \frac{n_{P,t} + n_{X,t}}{N_t} \quad [7]$$

onde n_P representa o número de investidores portugueses, que se assumem estarem "informados" sobre as acções portuguesas, n_X representa o número de investidores estrangeiros "informados" sobre as acções portuguesas, e N é o número total de investidores. Embora esta medida não seja observável, seguimos o procedimento sugerido por Clark e Berko (1997), e inferimos variações em q a partir de variações nas posições dos investidores estrangeiros e investidores domésticos em acções portuguesas em percentagem da capitalização do mercado português, definida como θ, e assumindo que o número de investidores domésticos se mantém constante quando os estrangeiros entram no mercado, e vice-versa.

O modelo de Merton (1987) implica que se os investidores domésticos e os investidores estrangeiros "informados" tiverem acesso à mesma informação, eles irão distribuir as suas carteiras de modo equivalente. Assim, θ será igual ao rácio dos investidores estrangeiros "informados" em relação ao total de investidores estrangeiros e investidores nacionais "informados":

$$\theta = \frac{n_X}{n_P + n_X} \quad \text{ou} \quad \theta = \frac{n_P}{n_P + n_X} \qquad [8]$$

Se, por exemplo, os investidores estrangeiros forem um terço dos investidores que estão "informados" sobre o mercado português, segundo o modelo de Merton, eles deverão representar um terço das posições de investimento em acções portuguesas. A medida q de Merton, por substituição e reorganizando, pode ser especificada como[2]:

$$q = \frac{n_P}{N}\left[\frac{1}{1-\theta}\right] \qquad [9]$$

Esta expressão implica que[3]:

$$\frac{\partial q}{q} = \frac{\partial \theta}{1-\theta}. \qquad [10]$$

Substituindo esta expressão na equação [5], e reorganizando, vem[4]:

$$\frac{\partial P_{k=}}{P_k} = \left[\frac{\delta\, x_k \sigma_k^2 N}{n_P}\right]\partial\theta \quad \text{ou} \quad \frac{\partial P_{k=}}{P_k} = \left[\frac{\delta\, x_k \sigma_k^2 N}{n_x^i}\right]\partial\theta \qquad [11]$$

Esta expressão é quase idêntica à expressão [5]; a diferença reflecte apenas a transformação realizada na medida q de Merton. Esta é a expressão utilizada por Clark e Berko (1997) para testar a hipótese de diminuição do custo de capital no mercado mexicano.

Notando que a posição de investimento estrangeira e doméstica, θ, varia com as aquisições líquidas de investidores estrangeiros e de residentes, respectivamente, de acções nacionais em percentagem da capitalização do mercado português, formulamos, à semelhança do realizado por Warther (1995) para a indústria norte-americana de fundos de investimento, a seguinte hipótese testável baseada na estática comparativa de Merton (1987), apresentada nas equações [4] a [6]. Partindo de uma regressão da forma:

$$R_PSI20_t = \beta_0 + \beta_1 Al_t + \varepsilon_t, \qquad [12]$$

onde R_PSI20_t é a rendibilidade no índice PSI-20 no período t. Caso a correlação contemporânea positiva entre fluxos líquidos estrangeiros e a rendibilidade do índice PSI-20 seja devida à diminuição do custo do capital no mercado português, é de esperar a rejeição da hipótese nula segundo a qual o coeficiente nas aquisições líquidas estrangeiras ou domésticas, β_1, é nulo a favor da hipótese alternativa unilateral: $\beta_1 > 0$[5].

---

[2] $q = \frac{n_P + n_X}{N} = \frac{n_P}{N}\left[\frac{n_P + n_X}{n_P}\right] = \frac{n_P}{N}\left[\frac{1}{1-(n_x^i/(n_P + n_x^i))}\right] = \frac{n_P}{N}\left[\frac{1}{1-\theta}\right]$.

[3] $\frac{\partial q}{\partial \theta} = \frac{n_P}{N}(1-\theta)^{-2} = \frac{n_P}{N}\left[\frac{1}{(1-\theta)^2}\right] \Leftrightarrow \frac{\partial q}{q} = \frac{n_P}{N}\left[\frac{1}{(1-\theta)^2}\right]\frac{\partial \theta}{q} = \frac{\partial \theta}{(1-\theta)}$.

[4] $\frac{\partial P_{k=}}{P_k} = \frac{\delta\, x_k \sigma_k^2}{q^2}\partial q = \frac{\delta\, x_k \sigma_k^2}{q}\frac{\partial \theta}{(1-\theta)} = \left[\frac{\delta\, x_k \sigma_k^2 N}{n_P}\right]\partial\theta$.

[5] Segundo Clark e Berko (1997) o parâmetro β_1 na equação [12] mede essencialmente o valor esperado das variáveis dentro de parênteses na equação [11]. Alguns dos parâmetros entre parênteses, em particular x_t, que no nosso caso representa o rácio da capitalização do mercado português em relação ao valor da carteira global, pode variar ao longo do período amostral. Clark e Berko (1997) adoptam a especificação de coeficiente fixo, para manter o paralelismo com as formas funcionais de Warther (1995), e também devido aos problemas dos erros de medição em relação à medida correcta do valor da carteira global para medir x_t. De igual modo, neste artigo, assumiremos que o coeficiente (os valores dos parâmetros) se mantém constante ao longo do período amostral.

No contexto da hipótese de mercado eficiente, Warther (1995) e Clark e Berko (1997) notam que mesmo se a hipótese de alargamento da base de investidores, ou seja, da diminuição do custo do capital, é verdadeira, o coeficiente β_1 na regressão [12] provavelmente subestima o impacto dos fluxos líquidos estrangeiros ou domésticos nos preços do mercado local.

Considerando a decomposição da série das aquisições líquidas no componente esperado e no componente não-esperado, esperamos que as rendibilidades covariem positiva e significativamente com as aquisições líquidas não-esperadas, mas não com as aquisições esperadas. Numa regressão da forma:

$$R_PSI20_t = \beta_0 + \beta_1 U(Al_t) + \beta_2 E_{t-1}(Al_t) + \varepsilon_t \quad [13]$$

esperaríamos que as rendibilidades evidenciassem uma covariância significativa com os fluxos não-esperados, e uma covariância menor com os fluxos esperados, i.é, que $\beta_1 > \beta_2 \geq 0$. Assim, caso a hipótese de diminuição do custo de capital seja verdadeira, será de esperar a rejeição da seguinte hipótese:

$$H_{0.1} : \beta_1 = 0 \quad [14]$$

e observar o coeficiente β_1 com sinal positivo. Dependendo da capacidade preditiva do modelo utilizado para prever os fluxos líquidos futuros, espera-se, ou não, aceitar a seguinte hipótese:

$$H_{0.2} : \beta_2 = 0. \quad [15]$$

Conforme referido por Clark e Berko (197) β_2 seria nulo se medíssemos com precisão as expectativas dos investidores. Porém, dado o reduzido período amostral de que dispomos, 1996:01 – 2004:06, estamos confinados à estimação do modelo de previsão neste período e à utilização das observações deste período no cálculo das aquisições líquidas esperadas. Ou seja, no período amostral completo, não utilizamos um sub-período para estimar o modelo de previsão e um sub-período não sobreposto ao primeiro para testar o modelo de previsão estimado. Neste caso podemos estar a atribuir aos investidores uma compreensão demasiado precisa do processo da série temporal para as aquisições líquidas, caso em que, o valor estimado das aquisições líquidas esperadas poderá incluir alguma componente de aquisições líquidas não-esperadas, o que nos levaria a rejeitar a hipótese de $H_0 : \beta_2 = 0$.

Assim, à semelhança de Warther (1995) e Clark e Berko (1997), estimamos uma especificação semelhante à equação [13] para testar a hipótese de diminuição do custo do capital e verificar se podemos atribuir a correlação positiva contemporânea entre os fluxos líquidos estrangeiros e a rendibilidade do índice PSI-20 à diminuição do custo do capital no mercado português. Utilizamos o processo autoregressivo univariado para representar a função de previsão dos investidores e com ele decompor o valor das aquisições líquidas dos grupos de participantes no mercado em dois componentes: o componente esperado e o não-esperado.

2.3 Hipótese de pressão (reversão) nos preços

Considerações baseadas na melhoria da partilha de risco e no aumento de liquidez sugerem que, em particular, o aumento da participação estrangeira deverá produzir, mantendo-se constantes as outras variáveis, uma redução permanente no prémio de risco e, assim, um aumento permanente de preço (Clark e Berko, 1997; Stulz, 1999). Uma teoria alternativa, que Warther (1995) e Clark e Berko (1997) também exploram e referem como hipótese de pressão nos preços, sugere que os aumentos nos preços associados ao aumento das aquisições líquidas são devidos a iliquidez temporária. Esta teoria conjectura que as aquisições líquidas induzem aumentos de preços que seriam posteriormente revertidos. Shleifer (1986) e Harris e Gurel (1986) documentam evidência de que aumentos nos preços resultantes do anúncio da inclusão de acções no índice *Standard & Poor´s* 500 são, pelo menos, parcialmente revertidos nos 30 a 60 dias seguintes de transacção. Esta hipótese de iliquidez temporária parece plausível no contexto do (pequeno) mercado português face a choques na procura por parte dos investidores estrangeiros e institucionais domésticos.

Para testar a hipótese de reversões nos preços, a abordagem de Warther (1995) é intuitiva e simples de concretizar. Clark e Berko (1997) também adoptam a abordagem de Warther (1995)

com a inclusão de valores desfasados das aquisições líquidas não-esperadas como regressores. A especificação empírica que iremos adoptar para testar a hipótese de reversão nos preços segue a abordagem utilizada por Clark e Berko (1997)[6]:

$$R_PSI20_t = \beta_0 + \sum_{i=1}^{4} \beta_i U_{t+1-i}(Al) + \varepsilon_t . \quad [16]$$

Se a hipótese de pressão nos preços é verdadeira, espera-se que as aquisições líquidas não-esperadas desfasadas tenham coeficientes negativos e significativos, rejeitando a hipótese nula:

$$H_0 : \beta_2 = \beta_3 = \beta_4 = 0 . \quad [17]$$

Warther (1995) testa esta hipótese e não confirma que as rendibilidades estejam relacionadas negativamente com os fluxos líquidos (mensais) anteriores. Esta incapacidade em detectar reversões nos preços é inconsistente com a presença de pressão nos preços embora, como nota Warther (1995), os testes não sejam poderosos. De igual modo, utilizando rendibilidades de frequência semanal, não confirma que os fluxos não-esperados correntes estejam relacionados negativamente com as rendibilidades (semanais) do mês seguinte. Assim, a evidência é consistente com a hipótese que os investidores individuais nos fundos de investimento têm informação ou transaccionam na mesma direcção que outros investidores que possuem informação. Clark e Berko (1997) também não confirmam que as variações de preços associadas com as aquisições líquidas não-esperadas se revertem nos meses seguintes.

2.4 Hipótese de revelação de informação e/ou efeitos de variáveis omitidas

Se se verificar que as aquisições líquidas estão correlacionadas com outros factores (sugeridos pela teoria do mercado eficiente) que exercem efeitos nos preços, como as variáveis de risco económico e financeiro (geralmente designadas de fundamentos do mercado), então a correlação entre as aquisições líquidas de alguns grupos de participantes e os preços pode reflectir essencialmente a influência destes terceiros factores, e não o efeito independente de alterações na procura de alguns grupos de participantes no mercado.

Para controlar o possível enviesamento decorrente da existência de variáveis omitidas, o que levaria a uma correlação espúria, seguimos a abordagem de Clark e Berko (1997) e Evans (2001), incluindo vários regressores adicionais. Como regressores adicionais consideramos uma variável representativa dos efeitos de contágio dos movimentos dos índices accionistas europeus no mercado português e várias variáveis de risco económico e financeiro relativo à economia portuguesa, e que podem ajudar a explicar a rendibilidade do mercado:

$$R_PSI20_t = \beta_0 + \beta_1 U_t(Fluxos) + \beta_2 E_{t-1}(Fluxos_t) + \boldsymbol{\beta}'\mathbf{z}_t + \varepsilon_t \quad [18]$$

onde $\mathbf{z}_t$ é o vector de regressores adicionais e $\boldsymbol{\beta}$ é o vector dos coeficientes. Sob a hipótese conjunta de que a hipótese de revelação de informação pelos fluxos e/ou de variáveis omitidas é verdadeira, e que incluímos correctamente as variáveis cuja informação nela contida é útil para prever a rendibilidade e está correlacionada significativamente com as aquisições líquidas, esperamos não rejeitar a hipótese nula conjunta que garante que o coeficiente das aquisições líquidas não-esperadas e esperadas numa regressão multivariada é igual a zero:

$$H_0 : \beta_1 = \beta_2 = 0 . \quad [19]$$

Se a diminuição do custo do capital for verdadeira, esperamos rejeitar esta hipótese nula, mesmo se as magnitudes dos coeficientes nas aquisições líquidas esperadas e não-esperadas nas regressões com e sem os regressores adicionais dos fundamentos do mercado diferem. Clark e Berko (1997) notam que existe a possibilidade de outras variáveis omitidas desempenharem o papel de influenciar as aquisições líquidas e as rendibilidades. O poder deste teste é claramente uma função da proporção da variação nas rendibilidades que somos capazes de explicar com a inclusão destas variáveis explicativas adicionais.

[6] Warther (1995) testa esta hipótese invertendo a regressão. Testa a existência de uma relação negativa entre fluxos não-esperados correntes e rendibilidades futuras, regredindo os fluxos não-esperados mensais nas rendibilidades das quatro semanas do mês seguinte.

3. RESULTADOS

3.1 Rendibilidade local e aquisições líquidas dos investidores estrangeiros e investidores domésticos

A regressão (1), na tabela 1, evidencia uma forte correlação contemporânea positiva entre a rendibilidade do índice PSI-20 e aquisições líquidas directas de acções nacionais pelos estrangeiros. Nesta especificação, as aquisições líquidas contemporâneas dos estrangeiros explicam cerca de 11,2% da variação mensal da rendibilidade do índice PSI-20 no período 1996:01 – 2004:06. A regressão (2) relaciona a rendibilidade do índice às aquisições líquidas contemporâneas e aquisições desfasadas de um, dois e três meses. Esta regressão mostra que o resultado obtido na regressão (1) é robusto quando incluímos os valores desfasados da variável $NEXTD_t$ na explicação da rendibilidade do índice PSI-20. De facto, a rendibilidade do PSI-20 está positiva e significativamente relacionada com as aquisições líquidas contemporâneas dos estrangeiros. Existe, todavia, uma relação positiva e significativa, ao nível de 10%, com as aquisições líquidas desfasadas de 3 meses e o coeficiente R_a^2 da regressão (2) é ligeiramente inferior ao obtido na regressão (1).

Tabela 1: Regressões da Rendibilidade e da Rendibilidade em Excesso, do Índice PSI-20 sobre as Aquisições Líquidas Directas de Acções Nacionais pelos Estrangeiros, 1996:01-2004:06.

Variável Dependente	R_PSI20_t			EXC_RPSI20_t	
(regressão)	(1)	(2)	(3)	(4)	(5)
β_0 (Constante)	-0,252	-0,860	-0,292	-1,068	-0,512
	(-0,377)	(-1,229)	(-0,434)	(-1,525)	(-0,762)
β_1 $(NEXTD_t)$	3,269***	2,828***		2,766**	
	(3,547)	(2,636)		(2,577)	
β_2 $(NEXTD_{t-1})$		1,271	2,059*	1,225	1,996*
		(1,258)	(1,813)	(1,208)	(1,758)
β_3 $(NEXTD_{t-2})$		-0,277	-0,718	-0,323	-0,755
		(-0,335)	(-0,831)	(-0,392)	(-0,858)
β_4 $(NEXTD_{t-3})$		1,511*	1,734*	1,451	1,669*
		(1,711)	(1,821)	(1,632)	(1,745)
R_a^2	0,112	0,103	0,032	0,095	0,030
Durbin-Watson	1,823	1,815	1,784	1,809	1,781

Esta tabela apresenta os resultados da regressão da rendibilidade, e da rendibilidade em excesso, do mercado português sobre as aquisições líquidas directas de acções nacionais pelos estrangeiros. A rendibilidade do mercado em excesso é calculada como a rendibilidade do índice PSI-20 menos a taxa mensal equivalente da taxa de juro de curto prazo de instrumentos de divida pública - 3 meses (anualizada), divulgada pelo Instituto de Gestão de Crédito Publico. As regressões são estimadas pelo método dos mínimos quadrados (OLS - *Ordinary Least Squares*) com variâncias-covariâncias consistentes à heterocedasticidade dos resíduos, método de White. As estatísticas-t encontram-se em parênteses. ***, **, *, denota significância ao nível de 1%, 5%, 10%, respectivamente.

Na regressão (2), a existência de uma relação contemporânea positiva e significativa entre a rendibilidade do PSI-20 e as aquisições líquidas dos estrangeiros, por um lado, e a inexistência de relações significativas entre a rendibilidade e os valores desfasados das aquisições líquidas dos estrangeiros, por outro, sugere, numa primeira instância, que os preços locais não reagem em demasia e inicialmente à entrada de capital estrangeiro e que posteriormente se revertem ou mantêm a tendência positiva. A relação positiva entre a rendibilidade e o valor desfasado de três

meses das aquisições líquidas dos estrangeiros, $\hat{\beta}_4 = 1{,}511$, poderá também significar que esta série (desfasada) representa a informação contida no componente esperado das aquisições contemporâneas e que a rendibilidade estará relacionada com o componente não-esperado das aquisições contemporâneas.

Com a regressão (3) vamos testar esta hipótese. Nesta regressão, as aquisições líquidas contemporâneas dos estrangeiros são omitidas. O coeficiente nas aquisições desfasadas de um mês, β_2, torna-se significativo ao n.s. de 10% e o coeficiente β_4 mantém quer a sua significância, quer a sua magnitude, quer o seu sinal. O R_a^2 desta regressão diminui significativamente. Estes resultados sugerem que são as aquisições líquidas desfasadas de um mês que tendem a conter a informação sobre o componente esperado das aquisições contemporâneas dos estrangeiros quando os valores contemporâneos das aquisições são omitidos, apesar da estimativa de β_4 aumentar ligeiramente. Nas regressões (4) e (5) usamos como variável dependente a rendibilidade do índice PSI-20 em excesso da taxa de juro isenta de risco. Esta última é representada pela taxa mensal equivalente da taxa de juro de curto prazo (passiva) a 3 meses (anualizada), divulgada pelo Instituto de Gestão de Crédito Público. Os resultados são semelhantes aos obtidos nas regressões (2) e (3). Porém, na regressão (4), comparativamente à regressão (2), a significância do coeficiente β_1 diminui e o coeficiente β_4 deixa de ser significativo. Porque, nas regressões, obtemos resultados muito semelhantes, usando quer a rendibilidade do PSI-20 quer a rendibilidade em excesso do PSI-20, nas secções seguintes apresentamos os resultados usando apenas a série R_PSI20_t para representar a rendibilidade do mercado.

Os dados na regressão (1), da tabela 1, mostram ainda que o coeficiente de regressão associado às aquisições contemporâneas regista uma magnitude inferior à evidenciada noutros estudos incluídos na literatura sobre a pressão nos preços e curva da oferta de acções com declive positivo. Estes estudos analisam a resposta dos preços de acções individuais a alterações não-esperadas na procura e oferta de acções. Harris e Gurel (1986) e Shleifer (1986) constatam que, no seguimento de anúncios relativos à inclusão de novas empresas no índice *Standard & Poor´s 500*, o preço das acções destas empresas aumentam, em geral, de cerca de 3% a 4%, provavelmente em resposta às compras realizadas por gestores de fundos de índices.

Warther (1995) obtém um coeficiente de resposta ainda maior nas suas regressões da rendibilidade do mercado norte-americano sobre as vendas líquidas contemporâneas e, em particular, sobre o componente não esperado das vendas líquidas contemporâneas de unidades de participação de fundos de investimento. Nas suas regressões, uma venda líquida mensal (*cash inflows*) não esperada equivalente a 0,1% da capitalização agregada do mercado está associada a um aumento de 5,2% no índice accionista. Clark e Berko (1997), na análise da relação entre o investimento estrangeiro no mercado mexicano e a variação do respectivo índice accionista, constatam que aquisições líquidas estrangeiras equivalentes a 1% da capitalização do mercado estão associadas a um aumento de cerca de 6% no índice de mercado.

No presente caso, a magnitude do coeficiente de resposta do índice PSI-20 às aquisições líquidas estrangeiras é cerca de metade do coeficiente observado por Clark e Berko (1997) e cerca de 16 vezes menos do que o observado por Warther (1995). Constata-se ainda pela regressão (1), na tabela 1, que aquisições líquidas directas estrangeiras, equivalentes a 1% da capitalização accionista portuguesa, estão associadas a um aumento de cerca de 3,3% no no valor do índice PSI-20.

Em suma, numa primeira análise, os resultados indicam que a entrada de investimento estrangeiro no mercado português tem um efeito positivo e significativo nos preços das acções.

3.2 Componentes esperados e não-esperados das aquisições líquidas

Para decompor as séries de fluxos líquidos em duas componentes, a antecipada e a não-antecipada, sobre-parametrizamos processos auto-regressivos (AR), começando por estimar, para cada série, processos AR de ordem 4. De seguida diminuímos a ordem de desfasamentos do processo AR para seleccionar o modelo que melhor capta o componente auto-regressivo da

série. O critério de escolha do processo AR que melhor capta o componente auto-regressivo baseia-se nos critérios informativos de *Akaike* e de *Schwartz* que minimizam o *lag* das duas estatísticas.

A tabela 2 apresenta as estimativas dos modelos de previsão (auto-regressivos) usados no cálculo do componente antecipado das aquisições líquidas de acções nacionais. A partir das estimativas dos coeficientes nestes modelos retiramos o componente antecipado, $\hat{Y}_t$, e o componente não-antecipado da série, $e = Y_t - \hat{Y}_t$. Aplicamos este procedimento às séries de aquisições líquidas dos investidores estrangeiros. Em cada modelo auto-regressivo estimado apresenta-se o R_a^2, o valor e significância da estatística F, a estatística de *Durbin-Watson*, e os valores das estatísticas usadas pelos critérios de informação de *Akaike* e *Schwartz*. O teste à autocorrelação de primeira ordem dos resíduos, baseado na estatística de *Durbin-Watson*, sugere que em nenhuma especificação os resíduos estão autocorrelacionados. Os modelos AR(4) e AR(1) na série *NEXTD* mostram-se significativos, aos n.s. de 5 e 1%, na explicação das aquisições líquidas, ambos com estatísticas F significativas ao n.s. de 5%. Estes resultados encaixam-se na autocorrelação positiva e significativa de primeira ordem das aquisições directas por parte dos estrangeiros.

Tabela 2: Equações de Previsão das Aquisições Líquidas de Acções Nacionais: Componente Esperada e Componente Não-Esperada.

Variável Dependente	*NEXTD*	
	AR(4)	AR(1)
Equação de previsão: $Y_{i,t} = \beta_0 + \beta_1 Y_{i,t-1} + ... + \beta_k Y_{i,t-k} + \varepsilon_{i,t}$		
β_0	0,1579**	0,1954***
	(2,015)	(3,226)
β_1	0,2644**	0,2438**
	(2,165)	(2,192)
β_2	-0,126	
	(-1,412)	
β_3	0,0278	
	(0,386)	
β_4	0,1889	
	(1,580)	
R_a^2	0,0734	0,0507
$Est.\,F$	2,923**	6,349***
Durbin-Watson	2,0353	1,9449
Akaike	2,0461	**2,0185**
Schwartz	2,1780	**2,0703**

Esta tabela apresenta as estimativas dos processos autoregressivos utilizados no cálculo do componente antecipado (esperado) e não-antecipado (ou não-esperado) das séries de aquisições líquidas de acções nacionais. As regressões são estimadas pelo método dos mínimos quadrados com variâncias-covariâncias consistentes à heterocedasticidade dos resíduos, método de White. Os valores em parênteses são as estatísticas-t. São apresentadas as estimativas dos processos autoregressivos de ordem 4, ordem máxima utilizada na selecção do modelo, e do processo autoregressivo que apresenta o menor valor segundo os critérios de informação de *Akaike* e de *Schwartz*. ***, **, *, denota significância ao nível de 1%, 5%, 10%, respectivamente.

A tabela 2 apresenta ainda os valores dos critérios informativos de *Akaike* e *Schwartz* utilizados na selecção do modelo que melhor capta o componente autoregressivo da série. Os resultados sugerem que, em ambos os critérios, os melhores resultados são obtidos com uma especificação simples AR(1), ou seja, aquele que melhor capta o componente autoregressivo das aquisições líquidas. É este o modelo utilizado para separar o componente previsível, ou

antecipado, do componente não antecipado das aquisições líquidas. Assim, usamos este modelo AR(1) para prever os valores das aquisições líquidas no período seguinte, procedimento que permite elaborar duas novas séries a partir da série original: a do componente antecipado, $\hat{Y}_t$, e a do componente não-antecipado, $e = Y_t - \hat{Y}_t$.

3.3 Hipótese de diminuição do custo do capital

A tabela 3 apresenta os resultados relativos à hipótese de diminuição do custo do capital exercido pelos investidores estrangeiros, começando por analisar se a rendibilidade do índice PSI-20 responde de modo diferente aos componentes esperado e não-esperado das aquisições líquidas dos estrangeiros. Na regressão (1), da tabela 3, relacionamos a rendibilidade do mercado às aquisições líquidas directas dos estrangeiros. Na regressão (2) relacionamos a rendibilidade ao componente não-esperado das aquisições líquidas dos estrangeiros. Quando passamos da regressão (1) para a regressão (2), o coeficiente das aquisições líquidas diminui ligeiramente, de 3,2 para 2,9, e o coeficiente de determinação R_a^2 diminui. Este resultado é consistente com a hipótese do componente esperado das aquisições líquidas deter algum poder explicativo sobre a rendibilidade. Sob a hipótese de diminuição do custo do capital, esperar-se-ia que a magnitude do coeficiente β_2, da regressão (2), aumentasse em relação ao coeficiente β_1, da regressão (1), porque este último estaria a representar o efeito simultâneo dos dois componentes. Num contexto de eficiência do mercado, a informação contida no componente esperado das aquisições líquidas seria incorporada nos preços imediatamente no início do período. De igual modo, este resultado poderá decorrer do processo auto-regressivo (utilizado para decompor a série nos componentes esperado e não-esperado) não captar adequadamente a expectativa de evolução das aquisições líquidas estrangeiras. Todavia os resultados fornecem evidência que tende a suportar a hipótese de diminuição do custo do capital, traduzida na rendibilidade na regressão (2) apresentar uma covariância positiva e significativa com o componente não-esperado das aquisições líquidas.

Na regressão (3) incluímos simultaneamente os dois componentes das aquisições líquidas. Constatamos que a hipótese de o coeficiente no componente esperado das aquisições ser nulo, tal como sugerido pela hipótese de diminuição do custo do capital, é rejeitada ao nível de 10%. Adicionalmente, a magnitude do coeficiente estimado no componente esperado é superior à observada no componente não-esperado, ou não-antecipado, apesar deste coeficiente ser significativo ao nível de 10% Esta evidência, numa primeira análise, não é inteiramente consistente com a verificação da hipótese de diminuição do custo do capital resultante da entrada de capital estrangeiro no mercado. De igual modo, na regressão (3), os dados não nos permitem rejeitar a hipótese de coeficientes idênticos para os dois componentes das aquisições líquidas estrangeiras. Na regressão (3), poderíamos esperar que β_3 fosse nulo se fossemos capazes de medir com precisão as expectativas dos investidores quanto à evolução futura das aquisições líquidas dos estrangeiros.

Porém, tal tarefa é difícil de concretizar, pois não existem orientações estabelecidas quanto ao método de captar as expectativas de evolução das aquisições. Acresce que, como no presente procedimento empírico do teste da hipótese não utilizamos um sub-período amostral para estimar um modelo de previsão (expectativas), e outro sub-período diferente para testar a hipótese utilizando o modelo de previsão estimado no sub-período anterior, podemos estar a atribuir aos investidores no mercado uma compreensão demasiado precisa do processo de evolução das aquisições líquidas por parte dos estrangeiros. Esta situação verifica-se provavelmente no nosso caso quando, na regressão (3), rejeitamos a hipótese nula $H_0 : \beta_3 = 0$ ao nível de 10%. Neste caso, é provável que estejamos atribuir aos investidores no mercado uma compreensão demasiado precisa do processo evolutivo relativo às aquisições dos estrangeiros.

Esta hipótese é parcialmente suportada pelos resultados obtidos na regressão (4) em que incluímos simultâneamente os componentes contemporâneos esperado e não-esperado e o componente não-esperado das aquisições líquidas do mês anterior. Os resultados mostram que apenas o coeficiente do componente não-esperado contemporâneo é significativo e ao nível de 1%. Nesta regressão o coeficiente no componente esperado perde a sua significância estatística. Constata-se ainda que o coeficiente positivo e significativo no componente esperado na

regressão (3) é influenciado, plausivelmente, pela relação entre o componente não-esperado no mês anterior e o componente esperado no período corrente.

Tabela 3: Regressões da Rendibilidade do PSI-20 sobre Aquisições Líquidas Antecipadas e Não Antecipadas dos Investidores Estrangeiros, 1996:01-2004:06.

Variável Dependente	R_PSI20_t			
(regressão)	(1)	(2)	(3)	(4)
β_0 (Constante)	-0,284	0,541	-1,335	-2,012
	(-0,429)	(0,857)	(-1,171)	(-0,598)
β_1 $(NEXTD_t)$	3,163***			
	(3,077)			
β_2 $(U(NEXTD_t))$		2,905***	2,905***	2,948***
		(2,873)	(2,759)	(2,744)
β_3 $(E(NEXTD_t))$			7,184*	9,894
			(1,888)	(0,762)
β_4 $(U(NEXTD_{t-1}))$				-0,656
				(-0,208)
Testes de Hipóteses			Significância	Significância
$H_0 : \beta_3 = 0$			0,055*	
$H_0 : \beta_3 = \beta_4 = 0$				0,150
$H_0 : \beta_2 = \beta_3$			0,286	
$H_0 : \beta_2 = \beta_3 = \beta_4$				0,424
R_a^2	0,096	0,074	0,098	0,090
Est. F	11,611***	8,986***	6,421***	4,623***
Durbin-Watson	1,819	1,799	1,815	1,822

Esta tabela apresenta os resultados da regressão da rendibilidade do índice PSI-20 sobre as aquisições líquidas directas de acções nacionais pelos investidores estrangeiros. As regressões são estimadas pelo método OLS com variância-covariância consistente à heterocedasticidade, método de White. Sob a hipótese $H_0 : \beta_3 = \beta_4 = 0$ o valor da significância é o valor da probabilidade associada à estatística do teste do rácio de verosimilhança de variáveis redundantes, cuja estatística segue a distribuição do Qui-Quadrado com número de graus de liberdade igual ao número de coeficientes a testar. O valor da significância para a hipótese $H_0 : \beta_2 = \beta_3 = \beta_4$ é o valor da probabilidade associada à estatística do teste de *Wald*, estatística que também segue a distribuição do Qui-Quadrado com o número de graus de liberdade igual ao número de coeficientes a testar. As estatísticas-t encontram-se em parênteses. ***, **, *, denota significância ao nível de 1%, 5%, 10%, respectivamente.

Esta associação traduz-se numa possível colinearidade entre os componentes no modelo de regressão. Esta colinearidade origina instabilidade na estimativa do coeficiente β_3 e o aumento do respectivo erro-padrão da estimativa quando passamos da regressão (3) para a regressão (4). A estimativa e a perda de significância do coeficiente β_3 na regressão (4), face ao evidenciado na regressão (3), sugerem que o respectivo componente, ao estar correlacionado com o componente não-esperado do período anterior poderá conter uma parcela de aquisições não-esperadas. Este facto traduzir-se-ia na impossibilidade de obtermos uma evidência imediata de suporte à hipótese de diminuição do custo do capital. Esta interpretação é suportada pelo facto de a estimativa do coeficiente do componente não-esperado contemporâneo se manter praticamente inalterado nas regressões (2), (3) e (4) e a estimativa do coeficiente ser significativa a 1%.

Concluímos que, embora os resultados das regressões não constituam evidência estatística inequívoca de suporte à hipótese de diminuição do custo do capital, os resultados tendem a conceder suporte à verificação desta hipótese. Em particular, a estimativa do coeficiente no componente não-esperado contemporâneo mantém-se idêntica e significativa nas várias

especificações de regressão, ao contrário do observado para o componente esperado. O reduzido período amostral e o método utilizado no teste da hipótese de diminuição do custo do capital, onde não é utilizado um sub-período para estimar o modelo de formação de expectativas e outro diferente para testar o modelo de previsão, podem atenuar a evidência de suporte à hipótese de diminuição do custo do capital.

3.4 Hipótese de pressão nos preços

A tabela 4 apresenta os resultados do teste à hipótese de pressão nos preços exercida pelas aquisições líquidas dos investidores estrangeiros. Os resultados não apresentam evidência de reversão nos preços exercida pelo componente não-esperado das aquisições líquidas dos estrangeiros. Na regressão (1) relacionamos a rendibilidade do mercado ao componente não-esperado contemporâneo das aquisições líquidas. Na regressão (2) explicamos a variação da rendibilidade do mercado em função dos componentes não-esperados, contemporâneo e dos três meses imediatamente anteriores, das aquisições líquidas. O coeficiente no componente não-esperado contemporâneo mantém inalterada, em relação à regressão (1), a significância e magnitude, sugerindo que o efeito contemporâneo deste componente nos preços de mercado é robusto aquando da inclusão dos seus valores desfasados. Esta evidência reforça a hipótese de este efeito estar associado à diminuição do custo do capital no mercado português.

Embora a verificação da hipótese de pressão nos preços, resultando possivelmente de uma iliquidez temporária, pareça plausível no contexto de um pequeno mercado, como o mercado português, a evidência nas aquisições líquidas dos investidores estrangeiros rejeita inequivocamente esta hipótese. Admitindo que a especificação empírica da hipótese capta adequadamente possíveis reversões nos preços, a evidência obtida não suporta a hipótese de os estrangeiros exercerem pressões nos preços, devida possivelmente a uma iliquidez temporária no mercado.

Tabela 4: Hipótese de Pressão nos Preços pelos Investidores Estrangeiros, 1996:01-2004:06.

Variável Dependente	R_PSI20		
(regressão)	(1)	(2)	(3)
β_0 (Constante)	0,541	0,543	0,550
	(0,857)	(0,849)	(0,875)
β_1 $(U(NEXTD_t))$	2,904***	2,902***	2,869***
	(2,872)	(2,681)	(2,709)
β_2 $(U(NEXTD_{t-1}))$		1,981**	1,757*
		(2,068)	(1,863)
β_3 $(U(NEXTD_{t-2}))$		0,173	
		(0,207)	
β_4 $(U(NEXTD_{t-3}))$		1,580*	
		(1,801)	
Teste de Hipótese		Significância	
$H_0 : \beta_2 = \beta_3 = \beta_4 = 0$		0,110	
R_a^2	0,074	0,101	0,096
Estat. F	8,986***	3,722***	6,237***
Durbin-Watson	1,800	1,815	1,800

Esta tabela apresenta os resultados da regressão da rendibilidade do índice PSI-20 sobre o componente não antecipado, contemporâneo e desfasado, das aquisições líquidas directas de acções nacionais pelos investidores estrangeiros. As regressões são estimadas pelo método OLS com variâncias-covariâncias consistentes à heterocedasticidade dos resíduos, método de *White*. Sob a hipótese $H_0 : \beta_2 = \beta_3 = \beta_4 = 0$, o valor da significância é o valor da probabilidade associada à estatística do teste do rácio de verosimilhança de variáveis redundantes, cuja estatística segue a distribuição do Qui-Quadrado com número de graus de liberdade igual ao número de coeficientes a testar. As estatísticas-t encontram-se em parênteses. ***, **, *, denota significância ao nível de 1%, 5%, 10%, respectivamente.

Na regressão (2) não obtemos evidência de que os componentes não-esperados desfasados das aquisições líquidas dos estrangeiros estejam relacionados negativa e significativamente à rendibilidade do mercado, como supõe a hipótese de pressão nos preços. Pelo contrário, na regressão (2), para além do coeficiente positivo e significativo no componente contemporâneo não-esperado, que tende a suportar a hipótese de diminuição do custo do capital, todos os restantes coeficientes são positivos. O coeficiente do componente não-esperado desfasado de um período apresenta um valor positivo e, financeira e estatisticamente, significativo ao nível de 5%. De igual modo, o coeficiente no componente não-esperado desfasado de três períodos é positivo e significativo ao nível de 10%. Na regressão (3) incluímos apenas o componente não-esperado do período anterior. A regressão mostra que este componente exerce um efeito positivo e significativo ao nível de 10% na rendibilidade do mercado, efeito que é contrário ao sugerido pela hipótese de pressão nos preços. Na regressão (2) os dados não nos permitem rejeitar a hipótese de os coeficientes serem nulos nos componentes não-esperados desfasados, apesar de, individualmente, os coeficientes nos componentes não-esperados desfasados de um e três períodos serem significativos aos níveis de 5% e 10%, respectivamente.

Em face da evidência obtida, concluímos que os investidores estrangeiros não tendem a exercer efeitos de pressão nos preços no mercado português mas, antes, tendem a reforçar a hipótese de os investidores estrangeiros exercerem efeitos de diminuição do custo do capital. Warther (1995), no seu estudo do efeito da venda líquida de unidades de participação de fundos de investimento mobiliários (*cash-inflows*) nos preços do mercado, também obtém evidência, ainda que reduzida, de que o componente não-esperado, desfasado de um período, dos fluxos líquidos está associado a uma rendibilidade positiva no período corrente. Clark e Berko (1997), na análise do efeito do investimento estrangeiro, principalmente norte-americano, no mercado accionista mexicano também obtêm evidência, ainda que reduzida, de que o componente não-esperado das aquisições líquidas, desfasado de um período, está associado a uma rendibilidade positiva no período corrente.

3.5 Hipótese de revelação de informação e/ou das variáveis omitidas

A investigação convencional do movimento de índices accionistas baseia-se no modelo de avaliação *standard* dos activos financeiros, representado pelo valor esperado actualizado dos *cash flows* futuros, que identifica choques nas expectativas dos fluxos de rendimento futuros, ou seja, que identifica variações não-esperadas nos *cash flows* futuros das empresas e nas taxas de desconto, como as forças que motivam a flutuação do preço. No essencial, qualquer força sistemática que afecta o pagamento esperado de dividendos aos accionistas pode influenciar os preços por razões que podem ser consideradas fundamentais ou relacionadas com o valor intrínseco (não observável). Porém, não existe uma medida directa, ou combinação de variáveis, representativa do valor intrínseco e, assim, nenhuma forma directa de explicar a variação do índice accionista.

Em geral, são identificadas variáveis observáveis (*ex ante*) consideradas adequadas para explicar o valor fundamental, não observável (*ex post*), do índice accionista. Os trabalhos econométricos de Fama (1981, 1990), Schwert (1990) e Chen, Roll e Ross (1986), Ferson e Harvey (1991), Lee (1992), Gjerde e Saettem (1999), Robotti (2002) utilizam as medidas de variação no índice de produção industrial (*leads*), a taxa de juro de curto prazo de obrigações (bilhetes do Tesouro a 30 dias), a estrutura de prazo de taxas de juro e o *default spread* (diferença entre o *yield* de obrigações de empresas com um *rating* de dívida de qualidade reduzida (maior risco) e de qualidade elevada (menor risco)) para explicar a variação fundamental do valor do índice. Estas são as variáveis tidas como determinantes do valor de equilíbrio dos índices accionistas, ou mais estritamente, as inovações fundamentais que influenciam a variação futura no pagamento de dividendos.

Nesta secção analisamos se a relação observada entre a rendibilidade do mercado e o componente não-esperado das aquisições líquidas dos estrangeiros no mercado é uma manifestação indirecta da correlação entre esta última variável e as variáveis fundamentais que se supõem exercer efeitos explicativos sobre os preços, mas omitidas nas análises de regressão anteriores. A hipótese alargada de mercado eficiente postula que alterações em variáveis que afectem o valor esperado dos *cash flows* futuros ou a taxa de desconto do processo de avaliação deverão ter implicações no movimento dos preços.

O interesse específico desta secção centra-se em analisar se a relação positiva e significativa entre a rendibilidade do mercado e o componente não-esperado das aquisições líquidas dos investidores estrangeiros observada anteriormente é, na verdade, motivada pela verificação da hipótese de diminuição do custo do capital exercida por estes fluxos ou se, pelo contrário, o efeito é motivado por uma correlação positiva entre o componente não-esperado dos fluxos e as variáveis fundamentais dos riscos económico e financeiro que influenciam a rendibilidade do mercado.

Especificando uma regressão que tenha como variável dependente a rendibilidade do mercado e englobando, simultâneamente, como regressores, os componentes esperado e não-esperado contemporâneos das aquisições líquidas e variáveis representativas dos riscos económico e financeiro, permite-nos controlar os possíveis efeitos da variação destas últimas na rendibilidade do mercado. Constatando que os regressores adicionais são significativos na explicação da rendibilidade do mercado e que estes absorvem os efeitos dos componentes não-esperados dos fluxos, tornando-os não significativos, esta evidência suportaria a hipótese de os fluxos não-esperados estarem a responder às alterações nas variáveis representativas dos fundamentos do mercado. Neste caso, os fluxos estariam a "revelar" informação ao mercado e a incorporar a informação nos preços. Tanto esta como a hipótese de diminuição do custo do capital inserem-se no âmbito da hipótese alargada de eficiência do mercado e ambas estão associadas a variações racionais nos preços, mas as variações seriam motivadas por causas diferentes.

3.5.1 Variáveis de risco económico e financeiro

Para testar a hipótese de revelação de informação pelos fluxos, incluímos como regressores adicionais variáveis representativas dos riscos económico e financeiro. As variáveis adicionais são as seguintes: 1) movimento nos preços de acções estrangeiras (europeias), representadas pelo índice *Dow Jones Euro Stoxx 50*; 2) o nível da taxa de juro (passiva) de curto prazo a 3 meses, portuguesa, taxa de juro anualizada, em percentagem; 3) a variação contemporânea da taxa de juro (passiva) de curto prazo a 3 meses, 4) a variação percentual na taxa de câmbio Euro – Dólar, 5) a volatilidade mensal dos preços das acções portuguesas, em percentagem; 6) a variação contemporânea no *yield-to-maturity* das obrigações do Tesouro português, a 10 anos, em pontos percentuais e, 7) a taxa de variação média mensal do índice de produção industrial português, em percentagem.

A inclusão da variação percentual no índice *Euro Stoxx 50* pretende controlar o efeito de contágio de variações nos preços de acções europeias no mercado português. Se os fluxos dos investidores estrangeiros no mercado português estão correlacionados com os movimentos nos preços de acções estrangeiras, então não esperamos rejeitar a hipótese de um coeficiente nulo no componente não-esperado das aquisições líquidas dos investidores estrangeiros. Neste caso, os fluxos líquidos estrangeiros poderiam estar a actuar como propagadores de estímulos estrangeiros no mercado português.

O nível da taxa de juro de curto prazo é incluído para captar o efeito preditivo deste na rendibilidade nominal esperada do mercado. Esta taxa de juro é uma taxa de referência para os instrumentos de dívida pública, com vencimento a 3 meses, elaborada pelo Instituto de Gestão do Crédito Público e divulgada no Boletim Estatístico do Banco de Portugal. É convencional na literatura utilizar o *yield* dos bilhetes do Tesouro a 30 dias. Aqui, não utilizamos este *yield* porque o Instituto de Gestão do Crédito Público suspendeu a emissão de bilhetes do Tesouro, a 30 dias, em 1999. Com a inclusão de uma medida de variação mensal contemporânea da taxa de juro de curto prazo pretende-se captar os possíveis efeitos nos preços das acções, resultantes de alterações, ao longo do mês, nas expectativas para a política monetária europeia.

O modelo de Merton (1987) sugere que os preços das acções devem covariar negativamente com a sua própria variância. Assim, utilizamos a volatilidade mensal do índice PSI-20 para captar o possível efeito desta no nível de preços. Como medida de percepção da variação do nível de risco financeiro do Estado Português utilizamos a variação mensal, em pontos percentuais, do *yield-to-maturity* das obrigações do Tesouro a 10 anos. Esta medida é uma estimativa alargada de alteração na avaliação pelos investidores da qualidade da dívida pública e da facilidade de acesso de Portugal a capitais estrangeiros para financiar a dívida pública. Esta medida poderá captar dois fenómenos distintos: a tendência de evolução nas taxas de juro internacionais e a variação no prémio de risco da dívida pública portuguesa. Espera-se que uma

variação positiva no *yield-to-maturity* esteja associada a uma diminuição nos preços das acções e a um aumento da rendibilidade esperada.

A taxa de variação média mensal do índice de produção industrial português é incluída para captar possíveis expectativas no crescimento económico do país. O efeito desta variável no índice accionista resultaria da inovação no índice de produção nos *cash flows* esperados das empresas. A medida de variação percentual na taxa de câmbio euro – dólar procura também captar o efeito exercido pela expectativa de alteração sobre a política de taxas de juro na zona euro no índice accionista nacional. Adicionalmente, esta medida pode influenciar os fluxos líquidos estrangeiros provenientes de mercados não pertencentes à zona euro.

3.5.2 Apresentação e discussão dos resultados

A tabela 5 apresenta os resultados da variação amostral explicada da rendibilidade do mercado atribuída aos regressores adicionais. As variáveis adicionais na regressão (1) explicam uma fracção significativa da variação amostral da rendibilidade, explicando conjuntamente cerca de 53% da variação total.

Nas regressões individuais, em geral, as variáveis apresentam os sinais esperados. A rendibilidade no índice *Dow Jones Euro Stoxx 50*, na regressão (2), explica cerca de 53% da variação da rendibilidade do PSI-20. Esta relação, positiva e fortemente significativa, em principio traduzirá o efeito de contágio/sentimento dos movimentos dos índices de acções europeus sobre o mercado português. O nível de taxa de juro de curto prazo portuguesa a 3 meses, na regressão (3), é significativo, ao nível de 10%, na explicação da rendibilidade do mercado, e o sinal positivo do coeficiente é consistente com a teoria que prevê que esta medida de taxa de juro tenha um efeito preditivo na rendibilidade nominal esperada do mercado.

Tabela 5: Rendibilidade do Índice PSI-20 e Variáveis Fundamentais do Mercado, 1996:01-2004:06.

Variável Dependente	R_PSI20_t				
(regressão)	(1)	(2)	(3)	(4)	(5)
β_0 (Constante)	-0,789	0,167	-1,576	0,534	3,511*
	(-0,432)	(0,380)	(-1,296)	(0,885)	(1,755)
β_1 ($R_EUXX\ 50_t$)	0,759***	0,751***			
	(9,386)	(10,498)			
β_2 ($TCPP_t$)	0,600*		0,578*		
	(1,892)		(1,819)		
β_3 ($\Delta TCPP_t$)	1,543		-4,310		
	(0,428)		(-1,050)		
β_4 ($\Delta OT10YMP_t$)	0,428			-1,708	
	(0,176)			(-0,627)	
β_5 (R_IPP_t)	-1,343				-0,025
	(-0,795)				(-0,014)
β_6 (Vol_PSI20_t)	-0,143				-0,632
	(-0,432)				(-1,335)
β_7 ($R_E/\$_t$)	0,287				
	(1,375)				
R_a^2	0,529	0,531	0,015	-0,007	0,028
$Est.F$	17,23***	115,34***	1,767	0,253	2,436*
Durbin-Watson	2,010	1,896	1,689	1,666	1,693

Esta tabela apresenta os resultados das regressões da rendibilidade do índice PSI-20 sobre um conjunto de variáveis representativas dos fundamentos do mercado e uma variável representativa do efeito de contágio de movimentos noutros mercados europeus. As regressões são estimadas pelo método

OLS com variâncias-covariâncias consistentes à heterocedasticidade dos resíduos, método de *White*. As estatísticas-t encontram-se em parênteses. ***, **, *, denota significância ao nível de 1%, 5%, 10%, respectivamente.

Na regressão (4) da tabela 5 o coeficiente da variação contemporânea no *yield-to-maturity* das obrigações do Tesouro português a 10 anos, é negativo mas não significativo. Este sinal é consistente com a teoria que estabelece uma relação negativa entre o aumento da percepção do risco financeiro do país e a rendibilidade no respectivo mercado. Em suma, exceptuando a rendibilidade do índice *Euro Stoxx 50* e o nível da taxa de juro (passiva) de curto prazo, as restantes variáveis, consideradas individual ou conjuntamente, não evidenciam poder explicativo significativo sobre a rendibilidade do índice PSI-20. A quase totalidade do poder explicativo contido na regressão (1) da tabela 5 é atribuída à variação da rendibilidade do índice *Euro Stoxx 50*.

A tabela 6 apresenta os resultados das regressões englobando simultâneamente os componentes esperados e não-esperados contemporâneos das aquisições líquidas dos investidores estrangeiros e o conjunto de variáveis representativas dos fundamentos de mercado. As regressões (1) e (2) da tabela 6, controlando a variação nas variáveis representativas dos fundamentos do mercado e nos efeitos de contágio (índice *Euro Stoxx 50*), mostram uma redução ligeira do impacto estimado do componente não-esperado das aquisições líquidas estrangeiras no índice PSI-20. O coeficiente estimado no componente não-esperado das aquisições estrangeiras diminui, mas a sua significância estatística aumenta, reforçando a evidência da hipótese da entrada de capital estrangeiro no mercado português exercer efeitos positivos sobre os preços por via da diminuição do custo do capital na bolsa portuguesa. De igual modo, na regressão (2), o coeficiente estimado do componente esperado diminui, permanecendo, todavia, significativamente diferente de zero ao nível de 10%. Globalmente, os modelos das rendibilidades nas regressões (1) e (2) explicam cerca de 56 e 57%, respectivamente, da variância da rendibilidade do índice PSI-20.

Tabela 6: Teste à Hipótese de Revelação de Informação e/ou de Variáveis Omitidas, 1996:01-2004:06.

Variável Dependente	R_PSI20_t			
	$NEXTD_t$			
(regressão)	(1)	(2)	(3)	(4)
β_0 (Constante)	0,814	0,273	-0,665	-2,297
	(0,465)	(0,156)	(-0,410)	(-1,118)
β_1 $(U(Fluxo_t))$	2,390***	2,583***	-8,513***	-8,221***
	(3,603)	(3,751)	(-3,485)	(-3,418)
β_2 $(E_{t-1}(Fluxo_t))$		5,193*		32,629*
		(1,691)		(1,793)
β_3 $(R_EUXX\,50_t)$	0,717***	0,709***	0,667***	0,641***
	(9,886)	(9,600)	(9,625)	(8,973)
β_4 $(TCP3MP_t)$	-0,094	-0,327	0,553*	0,469
	(-0,349)	(-1,082)	(1,849)	(1,552)
β_5 $(\Delta TCP3MP_t)$	-1,957	-1,021	-0,224	1,034
	(-0,576)	(-0,291)	(-0,063)	(0,288)
β_7 (Vol_PSI20_t)	-0,107	-0,102	-0,256	-0,269
	(-0,329)	(-0,304)	(-0,896)	(-0,940)
R_a^2	0,563	0,573	0,607	0,618
$Est.F$	22,50***	20,15***	26,77***	24,15***
Durbin-Watson	1,970	2,039	1,607	1,720

Esta tabela apresenta os resultados das regressões da rendibilidade do índice PSI-20 sobre um conjunto de variáveis representativas dos fundamentos do mercado e os componentes esperado e não-

esperado das aquisições líquidas dos investidores estrangeiros e domésticos. As regressões são estimadas pelo método OLS com variâncias-covariâncias consistentes à heterocedasticidade dos resíduos, método de *White*. As estatísticas-t encontram-se em parênteses. ***, **, *, denota significância ao nível de 1%, 5%, 10%, respectivamente.

Conclui-se assim que, (i) para além dos efeitos de propagação do sentimento/movimento dos índices accionistas europeus, os componentes esperados e não-esperados das aquisições líquidas dos investidores estrangeiros são significativos na explicação da rendibilidade do mercado português, (ii) que os resultados não permitem concluir que as relações positivas contemporâneas entre a rendibilidade do mercado e os efeitos dos componentes esperados e não-esperados das aquisições dos estrangeiros sejam uma manifestação indirecta da correlação existente entre estes componentes e as variáveis de risco económico e financeiro tidas como influentes no valor de equilíbrio do índice accionista doméstico. Estes resultados tendem a suportar a hipótese dos investidores estrangeiros diminuírem o custo do capital no mercado português.

4. CONCLUSÕES

Os resultados deste artio suportam a convição comumente aceite nos mercados financeiros que os fluxos estrangeiros para pequenos mercados, mesmo quando considerados desenvolvidos, têm um impacto positivo e significativo nos preços locais. No caso português, aquisições líquidas não esperadas de acções nacionais pelos investidores estrangeiros, equivalentes a um por cento da capitalização do mercado, estão associadas a um aumento contemporâneo dos preços locais de cerca de três por cento.

Neste artigo considerámos três hipóteses para explicar a correlação positiva contemporânea entre os fluxos estrangeiros e a rendibilidade local. Os resultados obtidos suportam a hipótese de alargamento da base de investidores, ou seja, a hipótese de diminuição do custo do capital: uma maior partilha de risco e aumento de liquidez no mercado resultante do investimento estrangeiro leva os preços para um nível de equilíbrio superior. No teste à hipótese de pressão nos preços pelos fluxos estrangeiros não se obtém evidencia de que a relação positiva contemporânea se reverte nos meses seguintes. Por sua vez, no teste à hipótese dos fluxos estrangeiros estarem a representar os efeitos de variáveis representativas dos fundamentos do mercado, mas omitidas nas regressões, constata-se que a relação entre os fluxos e os preços não é eliminada com a inclusão destas variáveis adicionais na regressão. Estes resultados levam-nos a concluir que os investidores estrangeiros têm um impacto considerável nos preços do mercado português, sendo muito provável que este efeito se deva à diminuição do custo do capital local. A entrada dos investidores estrangeiros no mercado português ao aumentar a base de investidores neste mercado, aumenta a partilha do risco do mercado local com os investidores estrangeiros, o que leva à diminuição do custo do capital e ao aumento dos preços.

BIBLIOGRAFIA

CHEN, NAI-FU; ROLL, RICHARD e STEPHEN A. ROSS, (1986), "Economic Forces and the Stock Market", *Journal of Business*, Vol.59, No.3, 383-403.

CLARK, JOHN e ELIZABETH BERKO, (1997), "Foreign investment fluctuations and emerging market stock returns: the case of Mexico", *Federal Reserve Bank of New York, Staff Reports*, No.24, May.

EVANS, LAWRENCE L. (2001), "*What Drives Equity Values: Fundamentals or Net Flows?An Empirical Analysis of the 1982-1999 U.S. Stock Market Boom*", PhD Dissertation, University of Massachusetts – Amherst, UMI Dissertations Services.

FAMA, EUGENE F., (1981), "Stock Returns, Real Activity, Inflation, and Money", *American Economic Review*, Vol.71, No.4, 545-565.

FAMA, EUGENE F., (1990), "Stock Returns, Expected Returns, and Real Activity", *Journal of Finance*, Vol.45, No.4, 1089-1108.

FERSON, WAYNE E. e CAMPBELL R. HARVEY, (1991), "The Variation of Economic Risk Premiums", *Journal of Political Economy*, Vol.99, No.2, 385-415.

GJERDE, OYSTEIN e FRODE SAETTEM, (1999),"Causal relations among stock returns and macroeconomic variables in a small, open economy", *Journal of International Financial Markets, Institutions and Money*, Vol. 9, 61-74.

HARRIS, LAWRENCE e EITAN GUREL, (1986), "Price and volume effects associated with changes in the S&P 500: new evidence for the existence of price pressure", *Journal of Finance*, Vol.41, No.4, 851-860.

LEE, B. (1992), "Causal Relations among Stock Returns, Interest Rates, Real Activity, and Inflation", *Journal of Finance*, Vol.47, No.4, 1591-1603.

MERTON, ROBERT C., (1987), "A Simple Model of Capital Market Equilibrium with Incomplete Information", *Journal of Finance*, Vol.42, No.3, 483-510.

ROBOTTI, CESARE, (2002), "Asset Returns and Economic Risk", *Federal Reserve Bank of Atlanta*, *Economic Review*, 2ºT, 13-25.

SCHWERT, G. WILLIAM, (1990), "Stock Returns and Real Activity: A Century of Evidence", *Journal of Finance*, Vol.45, No.4, 1237-1257.

SHLEIFER, ANDREI, (1986), "Do demand curves for stocks slope down?", *Journal of Finance*, Vol.41, No.3, 579-590.

STULZ, RENÉ M., (1999) "International Portfolio Flows and Security Markets", in M. Feldstein (ed.), *International Capital Flows* (p.257-293), NBER, University of Chicago Press, Chicago.

TESAR, LINDA L., (1999) "The Role of Equity in International Capital Flows", in M. Feldstein (ed.), *International Capital Flows* (p.235-257), NBER, University of Chicago Press, Chicago.

WARTHER, VINCENT A., (1995),"Aggregate mutual fund flows and security returns", *Journal of Financial Economics*, Vol. 39, 209-235.

ÁREA XIII

MÉTODOS CUANTITATIVOS PARA LA ECONOMÍA Y LA EMPRESA

MÉTODOS QUANTITATIVOS PARA A ECONOMIA E A EMPRESA

COMPARACIÓN DE DISTRIBUCIONES A PARTIR DE SU DESCOMPOSICIÓN DIFERENCIAL. APLICACIÓN AL ANÁLISIS DE LA DISTRIBUCIÓN PERSONAL DE INGRESOS DE CASTILLA-LA MANCHA.[1]

FRANCISCO JAVIER CALLEALTA BARROSO
E-mail: franciscoj.callealta@uah.es
Dpto. de estadística, Estructura Económica y O.E.I.
Facultad de CC.EE y EE. Universidad de Alcalá

JUAN FCO. ORTEGA DATO
E-mail: juanfco.ortega@uclm.es
Dpto. de Análisis Económico y Finanzas
Facultad de CC.EE y EE.. Universidad de Castilla-La Mancha

LUIS FELIPE RIVERA GALICIA
E-mail: luisf.rivera@uah.es
Dpto. de Estadística, Estructura Económica y O.E.I.
Facultad de CC.EE y EE. Universidad de Alcalá

Resumen

The divergence between distributions can be studied from different points of view with the help of appropriate measures of proximity defined for this purpose. Callealta (2006) introduced a new measure of proximity between distributions, as well as a methodology to study the divergence between them.
This paper analyzes the theoretical properties of this measure, deepening in the informational content of the differential distributions and the extent of proximity related to the concepts of inequality, overlapping and divergence of compared distributions.
Theoretical results are applied to analyze the divergence between the distribution of equivalent personal income distribution in the autonomous community of Castilla-La Mancha and the corresponding distribution of the rest of Spain, using data from the European Survey of Income and Living Conditions (EU-SILC).

Palabras clave: Divergencia, Diferencia Media de Gini, Desigualdad, Solapamiento.
Área Temática: Economía de la Energía, del Ambiente, de los Transportes y de los Recursos Naturales.

1. INTRODUCCIÓN.

El análisis dinámico del grado de proximidad y convergencia entre distribuciones manifestado por una adecuada medida de proximidad empleada a tal fin, puede facilitarse mediante el empleo de técnicas de escalamiento multidimensional. En Callealta (2006) se introduce una intuitiva medida de la proximidad entre distribuciones y se aplica para medir las proximidades o diferencias que presentan las diferentes distribuciones personales de ingresos equivalentes en la Unión Europea durante el periodo 1993-2000, con la finalidad de emplearla como indicador de una posible convergencia entre éstas, en ciertos aspectos que también investiga. De esta forma, el análisis se basaba fundamentalmente en la observación de las

[1] Esta investigación está siendo desarrollada en el contexto del Proyecto de I+D "UN PROTOTIPO PARA LA MONITORIZACIÓN DE LA DINAMICA DE DISTRIBUCIONES ECONOMICAS: ANÁLISIS DE LA CONVERGENCIA DE LAS DISTRIBUCIONES DE RENTA EN CASTILLA-LA MANCHA", referencia PAI08-0249-1400, del Plan Regional de Investigación Científica, Desarrollo Tecnológico e Innovación de Castilla–La Mancha (orden de 03/05/2007 de la Consejería de Educación y Ciencia de la JCCM).

formas y evolución de las trayectorias temporales de las distribuciones comparadas, que se visualizaban sobre un espacio factorial de dimensiones óptimamente reducidas.

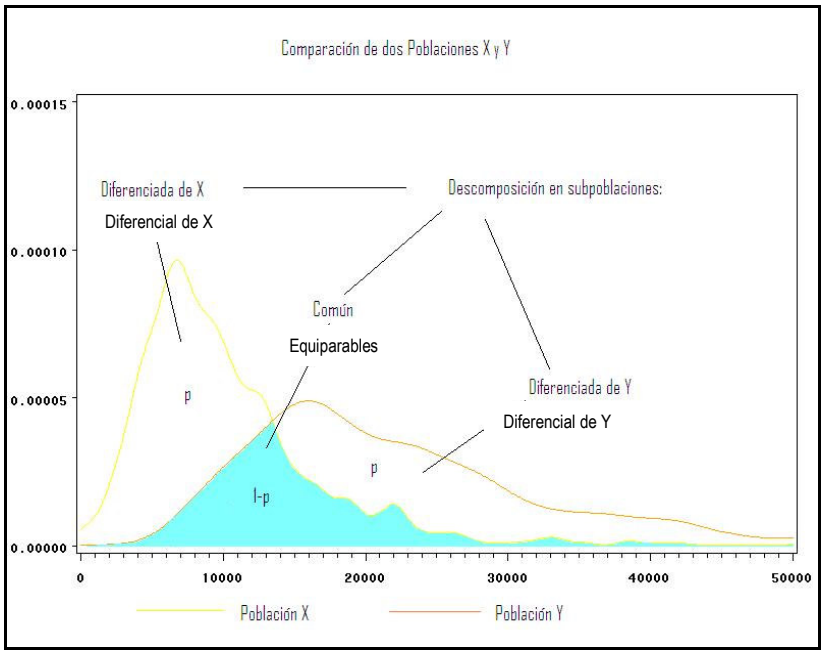
Ahora, consecuentemente con estos planteamientos, en este trabajo nos proponemos profundizar sobre la relación existente entre la citada medida de disimilaridad propuesta y la diferencia media de Gini calculada entre las poblaciones comparadas, bajo la hipótesis de que en ésta se mezclan informaciones relativas no sólo al grado de desigualdad existente en éstas, sino también a los grados de proximidad y de solapamiento entre ellas. Y finalmente, de forma adicional, se aplican los resultados teóricos obtenidos al análisis de la divergencia ente la distribución de ingresos personales equivalentes de la comunidad autónoma de Castilla-La Mancha y la correspondiente distribución observada en el resto del territorio español, de acuerdo con las Encuestas de Condiciones de Vida del INE.

2. DESCOMPOSICIÓN DIFERENCIAL DE LAS POBLACIONES COMPARADAS.

Siguiendo lo expuesto en Callealta (2006), supongamos que son las poblaciones P_X y P_Y las que se desean comparar a partir de sus comportamientos probabilísticos relativos a una cierta magnitud observada, representados por sendas variables aleatorias independientes y absolutamente continuas, X e Y, con funciones de densidad respectivamente asociadas $f_X(x)$ y $f_Y(y)$, bien definidas sobre $\Re$.

El gráfico nº 1 pretende mostrar de forma general los comportamientos comparados para ambas poblaciones, e ilustra de forma intuitiva la descomposición que se emplea en este trabajo y que se concreta a partir de los siguientes conceptos y definiciones.

Gráfico nº 1



Fuente: Elaboración propia.

2.1. Conceptos y Definiciones.

Según se observa en el gráfico nº 1, al comparar las poblaciones P_X y P_Y, de acuerdo con sus respectivas distribuciones de probabilidad, podemos considerar las siguientes subpoblaciones disjuntas de cada una de ellas, definidas como sigue:

a) *Subpoblaciones equiparables*:

Son aquellas dos subpoblaciones $P_{Cx} \subseteq P_X$ y $P_{Cy} \subseteq P_Y$, de igual tamaño relativo con respecto de sus respectivas poblaciones de procedencia P_X y P_Y, que presentan idénticas distribuciones de probabilidad y para las que nos es posible aumentar sus tamaños sin perder estas propiedades.

Este comportamiento "Común" de las dos subpoblaciones P_{Cx} y P_{Cy}, puede representarse mediante una variable aleatoria C, sugerida por la parte "común" sombreada en el gráfico nº 1, siendo su función de densidad asociada la siguiente:

$$f_C(t) = \frac{Min\left(f_X(t), f_Y(t)\right)}{1-p} = \begin{cases} \dfrac{f_X(t)}{1-p} & , si \quad f_X(t) \le f_Y(t) \\ \dfrac{f_Y(t)}{1-p} & , si \quad f_Y(t) \le f_X(t) \end{cases} \qquad [1]$$

En esta expresión, el valor *1-p* representa el tamaño relativo (proporción) de ambas poblaciones "equiparables", P_{Cx} y P_{Cy}, con respecto a los tamaños de sus respectivas poblaciones de pertenencia, P_X y P_Y:

$$1-p = \int_{\{y|f_X(y)>f_Y(y)\}} f_Y(y)dy + \int_{\{x|f_Y(x)>f_X(x)\}} f_X(x)dx \qquad [2]$$

b) *Subpoblación "diferencial, de X"*:

Es aquélla subpoblación $P_{X^*} \subseteq P_X$ cuyos elementos toman valores para la variable observada tales que, excluidos los elementos de la subpoblación P_{Cy}, no existe ningún otro elemento de la población P_Y que tome dichos valores. Consecuentemente, $P_{X^*} = P_X - P_{Cx}$.

El comportamiento de esta subpoblación "diferencial, de X", P_{X^*}, puede representarse mediante una variable aleatoria X^*, sugerida por la parte izquierda no sombreada en el gráfico nº 1, y cuya función de densidad vendría dada por:

$$f_{X^*}(x) = \frac{f_X(x) - f_Y(x)}{p} \cdot I\{f_X(x) > f_Y(x)\} =$$

$$= \begin{cases} \dfrac{f_X(x) - f_Y(x)}{p} & , si \quad f_X(x) > f_Y(x) \\ 0 & , si \quad f_X(x) \le f_Y(x) \end{cases} \qquad [3]$$

En esta expresión, el valor *p* representa el tamaño relativo (proporción) de la población diferencial P_{X^*} con respecto de su población de procedencia P_X; e $I\{\cdot\}$ es la función indicadora de la proposición entre llaves[2].

c) *Subpoblación "diferencial, de Y"*:

Es aquélla subpoblación $P_{Y^*} \subseteq P_Y$ cuyos elementos toman valores para la variable observada tales que, excluidos los elementos de la subpoblación P_{Cx}, no existe ningún otro elemento de la población P_X que tome dichos valores. Consecuentemente, $P_{Y^*} = P_Y - P_{Cy}$.

El comportamiento de esta subpoblación "diferencial, de Y", P_{Y^*}, puede representarse mediante una variable aleatoria Y^*, sugerida por la parte derecha no sombreada en el gráfico nº 1, y cuya función de densidad vendría dada por:

$$f_{Y^*}(y) = \frac{f_Y(y) - f_X(y)}{p} \cdot I\{f_Y(y) > f_X(y)\} =$$

$$= \begin{cases} \dfrac{f_Y(y) - f_X(y)}{p} & , si \quad f_Y(y) > f_X(y) \\ 0 & , si \quad f_Y(y) \le f_X(y) \end{cases} \qquad [4]$$

Obsérvese que el valor *p* representa también el tamaño relativo (proporción) de la población diferencial P_{Y^*} con respecto de su población de procedencia P_Y.

[2] La función indicadora para una proposición *A* vale $I\{A\} = \begin{cases} 1 & si \quad A \text{ es cierto} \\ 0 & si \quad A \text{ es falso} \end{cases}$

2.2. Propiedades.

A continuación se presentan algunas de las propiedades que cumplen las subpoblaciones y sus distribuciones asociadas resultantes de la descomposición diferencial propuesta:

1. Las subpoblaciones equiparables y diferenciales son complementarias dentro de sus poblaciones respectivas:

$$P_X = P_{X^*} \cup P_{Cx} \quad , \quad P_{X^*} \cap P_{Cx} = \phi$$
$$y \qquad [5]$$
$$P_Y = P_{Y^*} \cup P_{Cy} \quad , \quad P_{Y^*} \cap P_{Cy} = \phi$$

2. La variable X, observada sobre P_X, puede expresarse como una mixtura de las variables X* y C con pesos relativos de mixtura *p* y *1-p*, respectivamente; pudiendo expresarse su función de densidad y su función de distribución como:

$$f_X(x) = (1-p)\cdot f_C(x) + p\cdot f_{X^*}(x) \;;\; F_X(x) = (1-p)\cdot F_C(x) + p\cdot F_{X^*}(x) \quad \forall x \in \Re \qquad [6]$$

3. Análogamente, la variable Y, observada sobre P_Y, puede expresarse como una mixtura de las variables Y* y C con pesos relativos de mixtura *p* y *1-p*, respectivamente; pudiendo expresarse su función de densidad y su función de distribución como:

$$f_Y(y) = (1-p)\cdot f_C(y) + p\cdot f_{Y^*}(y) \;;\; F_Y(y) = (1-p)\cdot F_C(y) + p\cdot F_{Y^*}(y) \quad \forall y \in \Re \qquad [7]$$

4. Las variables X^*, Y^* y C son mutuamente independientes, al estar referidas a subpoblaciones diferentes.

5. La proporción de población calificada como "diferencial", tanto en P_X como en P_Y, es :

$$p = \frac{1}{2}\cdot\int_{-\infty}^{+\infty} |f_X(t) - f_Y(t)|\cdot dt \in [0, 1] \qquad [8]$$

siendo: $p = 0 \Leftrightarrow X \; i.d. \; Y$; y $p = 1 \Leftrightarrow f_x(t)\cdot f_y(t) = 0 \;\; \forall t \in \Re \; (c.s.)$

6. $E[X] = pE[X^*] + (1-p)E[C]$, $E[Y] = pE[Y^*] + (1-p)E[C]$ [9]
7. $(1-p)E[C] = E[X] - pE[X^*] = E[Y] - pE[Y^*]$ [10]
8. $E[X] - E[Y] = p\left(E[X^*] - E[Y^*]\right)$ [11]
9. $E[X] + E[Y] = p\left(E[X^*] + E[Y^*]\right) + 2(1-p)\cdot E[C]$ [12]

3. MEDIDA DE PROXIMIDAD ENTRE LAS POBLACIONES COMPARADAS.

3.1. Definición.

Con base en esta propuesta de descomposición de las distribuciones comparadas, en Callealta (2006) se propone, como medida de la proximidad entre las dos distribuciones X e Y, la diferencia media de Gini, $\Delta_{X^*Y^*}$, calculada entre las distribuciones asociadas X^* e Y^* de sus respectivas subpoblaciones diferenciales (no equiparables), ponderada por el producto de las proporciones que estas representan de las poblaciones originales X e Y[3].

$$d(X,Y) = p^2\cdot E\left[\left|Y^* - X^*\right|\right] = p^2\cdot\Delta_{X^*Y^*} =$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |y-x|\cdot\left(f_X(x) - f_Y(x)\right)\cdot I\{f_X(x) > f_Y(x)\}\left(f_Y(y) - f_X(y)\right)\cdot I\{f_Y(y) > f_X(y)\}\cdot dx\cdot dy$$

[3] Obsérvese que la introducción de la ponderación responde, en primer lugar, al interés de hacer que la medida sea lo más intuitiva posible (conduce a la evaluación directa de diferencias entre las unidades de las subpoblaciones diferenciales); así como, en segundo lugar, a introducir en la expresión el efecto de la magnitud relativa de las subpoblaciones que son realmente diferenciales o no equiparables (importancia relativa de las subpoblaciones diferenciales, establecida como el producto de las proporciones que éstas representan en cada una de las poblaciones originales comparadas).

3.2. Propiedades generales.

Ampliando las propiedades recogidas en Callealta (2006), a continuación se relacionan sus propiedades generales:

a) $d(X,Y)=0 \Leftrightarrow X \; i.d. Y$

b) La medida *d(X,Y)* crece con la diferencia entre X e Y; es decir, crece tanto con el aumento de la proporción (*p*) que representan las subpoblaciones diferenciales o no equiparables de X e Y, como con el aumento de la separación entre éstas.

c) La medida *d(X,Y)* es simétrica: $d(X,Y)=p^2 \cdot E\left[\left|X^*-Y^*\right|\right]=d(Y,X)$

d) $0 \le d(X,Y) \le p \cdot \Delta_{XY} \le \Delta_{XY}$, siendo Δ_{XY} la diferencia media de Gini entre las variables X e Y, ($\Delta_{XY}=E\left[\left|X-Y\right|\right]$).

e) Si $X>Y\,(\text{cs})$ ó $X<Y\,(\text{cs}) \Rightarrow d(X,Y)=\Delta_{XY}=\left|\mu_Y-\mu_X\right|$

f) Esta medida de disimilaridad es invariante ante traslaciones comunes para las variables comparadas: Si $\hat{X}=X+s,\ \hat{Y}=Y+s\ ,\ s\in\Re \Rightarrow d(\hat{X},\hat{Y})=d(X,Y)$

g) Además, es homogénea de grado 1, para cambios de escala comunes para las variables comparadas: Si $\hat{X}=\lambda \cdot X,\ \hat{Y}=\lambda \cdot Y\ ,\ \lambda\in\Re^+ \Rightarrow d(\hat{X},\hat{Y})=\lambda \cdot d(X,Y)$

h) Si dos distribuciones se trasladan alejándose respectivamente, su medida de disimilaridad aumenta: Si $\varepsilon\in\Re^+ \Rightarrow d\left(X,Y+\varepsilon\cdot(\mu_Y-\mu_X)\right)\ge d(X,Y)$

i) Esta medida de proximidad no cumple la propiedad triangular; por lo que no llega a ser estrictamente una medida de distancia[4].

Como complemento a estas propiedades, podemos también relacionar la medida de proximidad propuesta con la diferencia media de Gini en la población agregada formada por la mixtura de las poblaciones comparadas, así como con su índice de Gini; lo que, por su importancia desde el punto de vista del análisis económico, hacemos en los dos siguientes apartados.

3.3. Descomposición de la Diferencia media de Gini e interpretación de sus componentes.

Si X es una mixtura tal que $F_X(t)=p_U \cdot F_U(t)+p_V \cdot F_V(t)\quad , p_U\ge 0, p_V\ge 0, p_U+p_V=1$; si Y es otra mixtura tal que $F_Y(t)=p_W \cdot F_W(t)+p_S \cdot F_S(t)\quad , p_W\ge 0, p_S\ge 0, p_W+p_S=1$; y si las variables X e Y, U y V, W y S son respectivamente mutuamente independientes, entonces se cumple, de forma general, que:

$$\begin{aligned}\Delta_{XY} &= E\left[\left|X-Y\right|\right]= \\ &= p_U \cdot p_W \cdot E\left[\left|U-W\right|\right]+p_V \cdot p_W \cdot E\left[\left|V-W\right|\right]+p_U \cdot p_S \cdot E\left[\left|U-S\right|\right]+p_V \cdot p_S \cdot E\left[\left|V-S\right|\right]= \\ &= p_U \cdot p_W \cdot \Delta_{UW}+p_V \cdot p_W \cdot \Delta_{VW}+p_U \cdot p_S \cdot \Delta_{US}+p_V \cdot p_S \cdot \Delta_{VS}\end{aligned} \qquad [13]$$

$$\Delta_{XX}=p_U^2 \cdot \Delta_{UU}+2 \cdot p_U \cdot p_V \cdot \Delta_{UV}+p_V^2 \cdot \Delta_{VV} \qquad [14]$$

Teniendo en cuenta que las expresiones [6] y [7] expresan las variables X e Y como mixturas de las respectivas variables equiparables y diferenciales, supuesta la independencia de las variables X e Y, podemos emplear las expresiones [13] y [14] para demostrar las siguientes expresiones de la Diferencia Media de Gini relativas a las distribuciones comparadas:

$$\begin{aligned}\Delta_{XX} &= p^2 \cdot \Delta_{X^*X^*}+2 \cdot p \cdot (1-p) \cdot \Delta_{X^*C}+(1-p)^2 \cdot \Delta_{CC} \\ \Delta_{YY} &= p^2 \cdot \Delta_{Y^*Y^*}+2 \cdot p \cdot (1-p) \cdot \Delta_{Y^*C}+(1-p)^2 \cdot \Delta_{CC} \\ \Delta_{XY} &= p^2 \cdot \Delta_{X^*Y^*}+\frac{1}{2}\left\{\Delta_{XX}-p^2 \cdot \Delta_{X^*X^*}+\Delta_{YY}-p^2 \cdot \Delta_{Y^*Y^*}\right\}\end{aligned} \qquad [15]$$

[4] Si bien esta medida de proximidad no cumple la propiedad triangular, su raíz cuadrada sí parece cumplirla prácticamente en todas las aplicaciones prácticas realizadas hasta la fecha.

Consideremos ahora que las poblaciones P_X y P_Y tienen sendos pesos respectivos p_X y $p_Y = 1-p_X$ en la población conjunta $P_T = P_X \cup P_Y$, y que, entonces, ésta puede nuevamente considerarse como una mixtura de las anteriores. Así, aplicando la expresión [14] y sustituyendo en ella los resultados de [15], se obtiene la siguiente expresión que nos permite descomponer la Diferencia Media de Gini de una población compuesta por dos poblaciones a las que se les ha practicado la descomposición diferencial (en subpoblaciones "equiparables" y "diferenciales") definidas previamente:

[16]

$$\Delta_{TT} = (1-p)^2\Delta_{CC} + p_X^2\cdot p^2\Delta_{X^*X^*} + (1-p_X)^2\cdot p^2\Delta_{Y^*Y^*} + \\ + 2p_X\cdot p(1-p)\Delta_{X^*C} + 2(1-p_X)\cdot p(1-p)\Delta_{Y^*C} + 2\cdot p_X\cdot(1-p_X)\cdot p^2\cdot\Delta_{X^*Y^*} \quad [17]$$

Obsérvese que en dicha expresión aparecen las siguientes componentes a las que podemos dar, en principio, los siguientes significados (donde por μ_V y G_{VV} hacen referencia a la media e índice de Gini de la variable V):

- Proporción de población equiparable o "núcleo de solapamiento"
$$P = 1-p = 1-\frac{1}{2}\cdot\int_{-\infty}^{+\infty}\left|f_X(t)-f_Y(t)\right|\cdot dt$$
- Indicador de diversidad del "núcleo de solapamiento"
$$I_{CC} = (1-p)^2\cdot\Delta_{CC} = 2\cdot(1-p)^2\mu_C\cdot G_{CC}$$
- Indicador de integración de X^* con el "núcleo de solapamiento"
$$I_{X^*C} = p\cdot(1-p)\cdot\Delta_{X^*C} = \frac{1}{2}\left\{\Delta_{XX} - p^2\cdot\Delta_{X^*X^*} - (1-p)^2\cdot\Delta_{CC}\right\}$$
- Indicador de integración de Y^* con el "núcleo de solapamiento"
$$I_{Y^*C} = p\cdot(1-p)\cdot\Delta_{Y^*C} = \frac{1}{2}\left\{\Delta_{YY} - p^2\cdot\Delta_{Y^*Y^*} - (1-p)^2\cdot\Delta_{CC}\right\}$$
- Indicador de la diversidad de la subpoblación diferencial de X (de X^*)
$$I_{X^*X^*} = p^2\cdot\Delta_{X^*X^*} = 2p^2\cdot\mu_{X^*}\cdot G_{X^*X^*}$$
- Indicador de la diversidad de la subpoblación diferencial de Y (de Y^*)
$$I_{Y^*Y^*} = p^2\cdot\Delta_{Y^*Y^*} = 2p^2\cdot\mu_{Y^*}\cdot G_{Y^*Y^*}$$
- Indicador de disimilaridad diferencial, o divergencia, entre X e Y
$$\delta_{XY} = p^2\cdot\Delta_{X^*Y^*}$$

En función de estos indicadores, las expresiones de las Diferencias Medias de Gini relativas a las distribuciones consideradas serían:

$$\Delta_{XX} = I_{X^*X^*} + 2\cdot I_{X^*C} + I_{CC}\,; \quad \Delta_{YY} = I_{Y^*Y^*} + 2\cdot I_{Y^*C} + I_{CC}\,; \quad \Delta_{XY} = I_{CC} + I_{X^*C} + I_{Y^*C} + \delta_{XY}$$

$$\Delta_{TT} = p_X^2 I_{X^*X^*} + (1-p_X)^2 I_{Y^*Y^*} + I_{CC} + 2p_X I_{X^*C} + 2(1-p_X)I_{Y^*C} + 2p_X(1-p_X)\delta_{XY}$$

3.4. Descomposición del Índice de Gini e interpretación de sus componentes.

Teniendo en cuenta las relaciones matemáticas existentes entre las Diferencias Medias de Gini y los correspondientes Índices de Gini, pueden demostrarse las siguientes expresiones que relacionan los Índices de Gini de las distribuciones comparadas con los indicadores anteriormente definidos:

$$\mu_X\cdot G_X = p^2\cdot\mu_{X^*}\cdot G_{X^*} + p\cdot(1-p)\cdot(\mu_{X^*}+\mu_C)G_{X^*C} + (1-p)^2\cdot\mu_C\cdot G_C = \\ = \frac{1}{2}I_{X^*X^*} + I_{X^*C} + \frac{1}{2}I_{CC} \quad [18]$$

$$\mu_Y \cdot G_Y = p^2 \cdot \mu_{Y^*} \cdot G_{Y^*} + p \cdot (1-p) \cdot (\mu_{Y^*} + \mu_C) G_{Y^* C} + (1-p)^2 \cdot \mu_C \cdot G_C = \\ = \frac{1}{2} I_{Y^* Y^*} + I_{Y^* C} + \frac{1}{2} I_{CC} \qquad [19]$$

$$(\mu_X + \mu_Y) \cdot G_{XY} = p^2 \cdot \Delta_{X^* Y^*} + \mu_X \cdot G_X - p^2 \cdot \mu_{X^*} \cdot G_{X^*} + \mu_Y \cdot G_Y - p^2 \cdot \mu_{Y^*} \cdot G_{Y^*} = \\ = I_{CC} + I_{X^* C} + I_{Y^* C} + \delta_{XY} \qquad [20]$$

Si consideramos ahora la población conjunta $P_T = P_X \cup P_Y$, como una mixtura, de las anteriores, de la expresión [16] obtenemos que:

$$\mu_T \cdot G_T = p_X \cdot \mu_X \cdot G_X + p_Y \cdot \mu_Y \cdot G_Y + p_X \cdot p_Y \cdot p^2 \cdot \left\{ \Delta_{X^* Y^*} - \mu_{X^*} \cdot G_{X^*} - \mu_{Y^*} \cdot G_{Y^*} \right\}$$

de donde:

$$\boxed{\mu_T G_T = p_X \cdot \mu_X G_X \cdot \left(1 - p_Y p^2 \cdot \frac{\Delta_{X^* X^*}}{\Delta_{XX}} \right) + p_Y \cdot \mu_Y G_Y \left(1 - p_X p^2 \cdot \frac{\Delta_{Y^* Y^*}}{\Delta_{YY}} \right) + p_X \cdot p_Y \cdot \delta_{XY}} \qquad [21]$$

Podemos ver aquí cierta analogía con la descomposición que se presenta en Yitzhaki (1994)[5]. Pero, ahora, en la descomposición que se presenta en la expresión [21], el lugar de la aproximación *GB* al índice de Gini entre grupos a la que alude dicho autor, lo ocupa la medida de proximidad propuesta entre las poblaciones comparadas, δ_{XY}, ponderada por el producto de sus correspondientes pesos de mixtura.

Y el lugar de los indicadores de solapamiento de las variables sobre la población agregada, propuestos por Yitzhaki, lo ocupan ahora otros indicadores que miden también el grado de "solapamiento" de cada una de las poblaciones sobre la población agregada, pero en el sentido de su "equiparabilidad". Así, en este sentido:

- el indicador del Solapamiento de la población P_X con la población agregada P_T, sería:

$$\boxed{OVL_{X|T} = 1 - p_Y p^2 \cdot \frac{\Delta_{X^* X^*}}{\Delta_{XX}} = p_X + p_Y \left(1 - p^2 \cdot \frac{\Delta_{X^* X^*}}{\Delta_{XX}} \right) \Rightarrow p_X \le OVL_{X|T} \le 1} \qquad [22]$$

  cumpliéndose que:
  $OVL_{X|T}$=1 cuando X~Y~T están idénticamente distribuidas
  $OVL_{X|T}=p_X$ cuando $f_X(t) \cdot f_Y(t)=0$, $\forall t \in \Re$ (X e Y completamente diferenciales)

- el indicador de Solapamiento de la población P_Y con la población agregada P_T, sería:

$$\boxed{OVL_{Y|T} = 1 - p_X p^2 \cdot \frac{\Delta_{Y^* Y^*}}{\Delta_{YY}} = p_Y + p_X \left(1 - p^2 \cdot \frac{\Delta_{Y^* Y^*}}{\Delta_{YY}} \right) \Rightarrow p_Y \le OVL_{Y|T} \le 1} \qquad [23]$$

  cumpliéndose que:
  $OVL_{Y|T}$=1 cuando X~Y~T están idénticamente distribuidas
  $OVL_{Y|T}=p_Y$ cuando $f_X(t) \cdot f_Y(t)=0$, $\forall t \in \Re$ (X e Y completamente diferenciales).

4. ANÁLISIS DE LA PROXIMIDAD ENTRE LAS DISTRIBUCIONES PERSONALES DE INGRESOS DE CASTILLA-LA MANCHA Y DEL RESTO DE ESPAÑA.

El objetivo de este apartado es emplear los resultados teóricos anteriores como ayuda para el análisis de la posible proximidad existente entre las distribuciones personales de ingreso de los residentes en la Comunidad Autónoma de Castilla-La Mancha y la del resto de españoles, a modo de ilustración de sus posibilidades.

[5] En Yitzhaki (1994), la fórmula referenciada allí como [21] presenta la descomposición por grupos del Índice de Gini de la población agregada, como:

$$G_T = \frac{p_X \mu_X}{\mu_T} \cdot O_X \cdot G_X + \frac{p_Y \mu_Y}{\mu_T} \cdot O_Y \cdot G_Y + GB$$

siendo *GB* una aproximación al índice de Gini entre grupos, y O_X y O_Y los indicadores de solapamiento de X e Y sobre la población agregada, respectivamente.

4.1. Conceptos y Fuentes.

Se han empleado los datos de ingresos familiares de la Encuesta de Condiciones de Vida (ECV) que se realiza en España desde 2004, que proporciona datos sobre ingresos para los años anteriores a las encuestas. Así y de momento, disponemos de éstos solamente para los años 2003 a 2005, que serán los empleados para este estudio; lo que constituye una serie muy corta para poder hablar del proceso de convergencia o divergencia en el tiempo, pero que permite ilustrar la metodología propuesta para ello.

El concepto de ingreso usado como punto de partida para este análisis ha sido el de la Renta Disponible Total del Hogar en el año anterior al de la encuesta (variable HY002). Y para hacer comparables las distribuciones de ingresos considerados, en los distintos años de la encuesta, las distribuciones han sido consideradas en términos reales deflactando sus valores mediante los correspondientes IPC, base 2006.

Y como el bienestar de los hogares no sólo depende de sus ingresos, sino también de su tamaño y composición, hemos calculado finalmente, para cada año de la encuesta, la variable "Renta Disponible personal equivalente" (real, en euros constantes de 2006), dividiendo la Renta Disponible Total del Hogar (real) por el número de miembros equivalentes que resulta en cada familia al aplicar la escala de equivalencia de Oxford[6] al Número de Miembros del Hogar calculado a partir del Fichero de Datos Básicos de las Personas.

Estas "Rentas Disponibles personales equivalente" (reales), han sido asignadas a todos y cada uno de los miembros del mismo hogar, asumiendo la hipótesis de que todos los miembros de un mismo hogar disfrutan de un mismo nivel de bienestar económico. Desde este enfoque, la unidad de análisis es el individuo; por lo que, en cada año de la encuesta, la variable "Renta Disponible personal equivalente (real)" así construida ha sido ponderada por una variable "peso" construida como el producto del peso transversal del hogar (variable DB090) y el número de miembros del hogar, obtenido a partir del fichero básico de personas.

4.2. Descomposición diferencial de las distribuciones personales de ingresos de Castilla-La Mancha y del resto de España.

En primer lugar, se ha procedido a estimar las funciones de densidad de las distribuciones de "Renta Disponible personal equivalente (real)" de Castilla-La Mancha y del resto de España en los 3 años de observación, de forma no paramétrica, mediante núcleos ("kernels") gaussianos univariantes con anchos de banda óptimos, según el procedimiento de Silverman (1986). Para la estimación se ha empleado el procedimiento KDE de SAS/STAT[7], lo que nos ha permitido calcular las correspondientes estimaciones en cada uno de los 601 puntos equiespaciados en que hemos dividido el recorrido común considerado (de 0 a 60.000 euros constantes de 2006) que se ha prefijado como suficiente para todas las distribuciones de ingresos comparadas[8]. Sus representaciones gráficas obtenidas con el procedimiento GPLOT de SAS/GRAPH[7]se muestra en el gráfico nº 2.

Aunque el intervalo de observación es aún pequeño, en espera de ir incorporando información de años sucesivos en el ámbito del proyecto que financia este estudio, se aprecia en ellos una situación relativamente atrasada de la distribución de Castilla-La Mancha con respecto de la del resto de España en todos los años; a la vez que un sistemático desplazamiento entre años consecutivos, con relación al resto de España, de elementos de rentas más bajas en aquélla

[6] El número de personas equivalentes que resulta en la escala de equivalencia de Oxford para un hogar es igual a la raíz cuadrada del número de sus miembros.

[7] SAS/STAT® y SAS/GRAPH® son productos registrados de SAS Institute Inc., Cary, NC, USA.

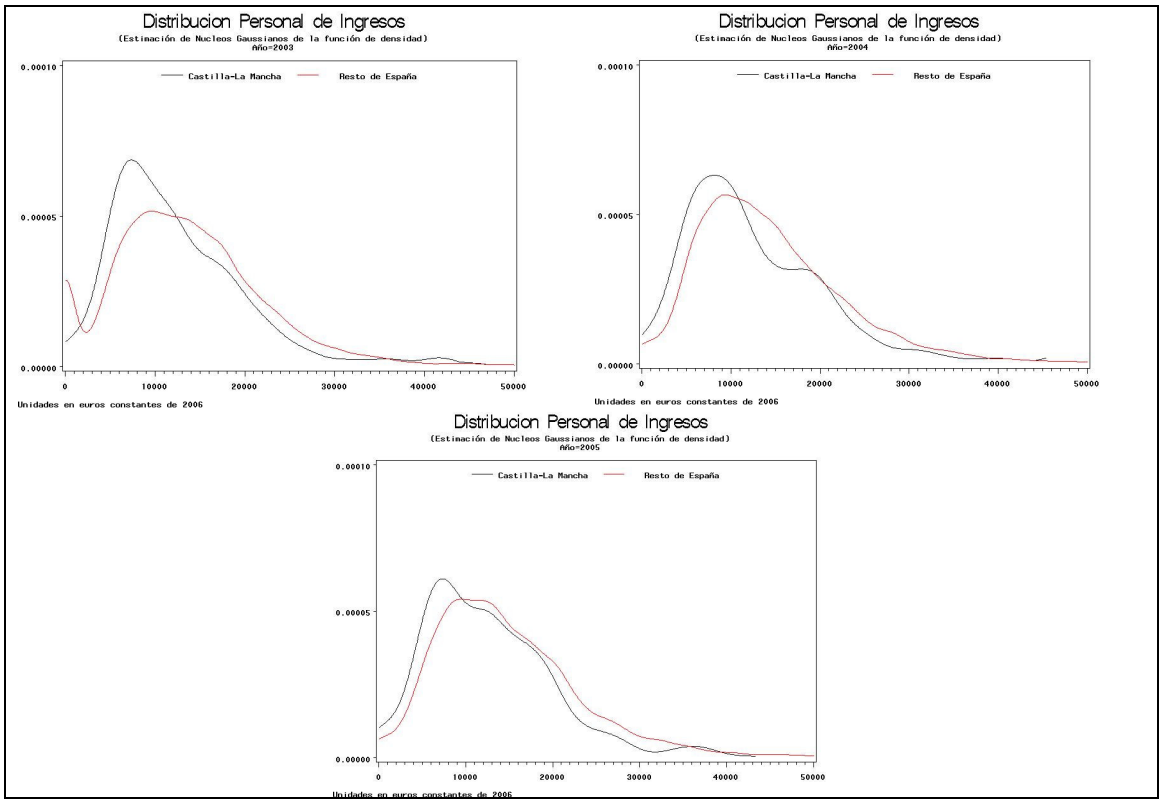
[8] Las estimaciones que los núcleos estocásticos dan para las funciones de densidad reproducen pequeñas desviaciones para las estimaciones de las medias poblacionales. En el convencimiento de que los tamaños muestrales de las ECV han sido calculados más para obtener estimaciones paramétricas que para otra función, hemos procedido, en cada caso, a corregir muy levemente la correspondiente función de densidad reestructurando el último intervalo de la cola derecha de la distribución, para el que se han determinado convenientemente su amplitud y marca de clase, de forma que su media reprodujera fielmente la correspondiente media estimada por el ECV.

comunidad hacia zonas más centrales de las distribuciones en los años siguientes. Lo que parece aventurar un posible proceso de convergencia entre ellas.

Para tratar de profundizar en esta primera impresión, analizaremos la información que suministra la descomposición diferencial propuesta para las distribuciones comparadas.

Así, la población agregada (España) puede verse como una mixtura de las de Castilla-La Mancha y del resto de España, con pesos de mixtura del 4,3% y del 95,7% respectivamente durante el período observado; y la descomposición diferencial de sus distribuciones personales de ingresos nos permite apreciar cómo, en líneas generales, los elementos diferenciales de las poblaciones (las componentes específicas o distribuciones CLM_e y R_E_e) se posicionan fundamentalmente de forma que las zonas de rentas bajas (hasta unos 10.000 € de 2006) la ocupa la distribución de Castilla-La Mancha, y a partir de ahí, presentando comportamientos distintos en dos intervalos adyacentes (entre 10.000 € - 20.000 € y entre 20.000 € - 60.000 €), los elementos del resto de España (ver Cuadro nº 1).

Gráfico nº 2



Fuente: Elaboración propia, a partir de los datos de la ECV.

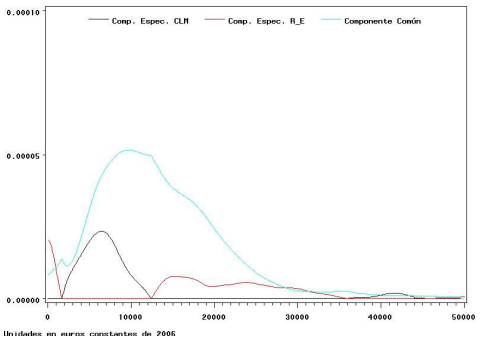
En este cuadro nº 1 se aprecia un comportamiento de los datos, cuando menos, extraño para el resto de España en 2003, y alrededor del cero. Este hecho motiva que su componente específica para este año presente un índice de Gini y una media sospechosamente alto y baja, respectivamente[9].

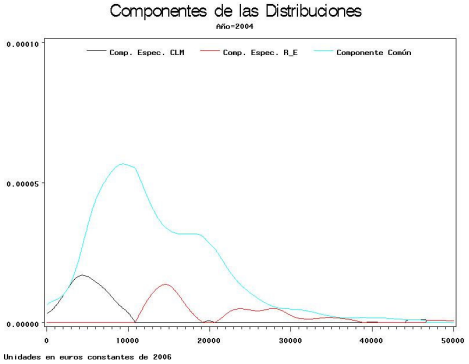
Salvando este aspecto, que obviamente también afecta al comportamiento relativo de la componente específica de Castilla-La Mancha, se aprecia que la proporción de población diferencial en estas dos poblaciones comparadas va decreciendo año a año (15%, 12,3%, 10,9%) apoyando la tesis de una presumible convergencia entre ambas; aunque la brecha entre las medias de las poblaciones especificas (diferenciales) comparadas, cada vez de menores tamaños, va creciendo paulatinamente (de 8.618 a 17.444 en 2003[10], de 5.635 a 23.454 en 2004, y de 5.758 a 26.766 en 2005).

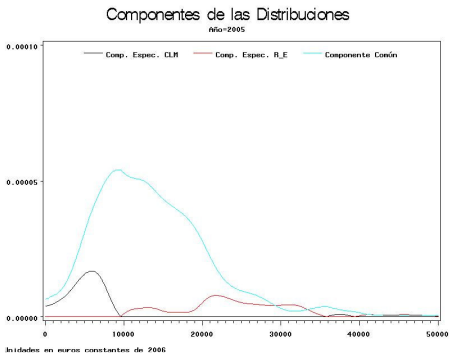
[9] La causa de esta anomalía debe ser necesariamente investigada; si bien no lo ha sido aún, por nuestra parte, a fecha de la redacción de este trabajo.

[10] Si bien, en nuestra opinión, estos valores deben ponerse en tela de juicio, ya que, presumiblemente, pueden estar influenciados por la anomalía antes referida.

Cuadro nº 1







	2003 T	2003 CLM	2003 R_E	2003 CLM_e	2003 R_E_e	2003 C
PROB	1	1	1	0,15	0,15	0,85
MEDIA	13827	12563	13883	8618	17444	13257
DT	9139	8057	9181	8644	14411	7744
DMGINI	9658	8495	9702	6228	15239	8335
P_05	600	2800	500	3000	0	2700
P_10	3700	4400	3700	3600	0	4800
P_25	7800	7000	7800	5000	800	7900
P_50	12700	10900	12800	6600	17600	12100
P_75	18300	16600	18400	8500	25000	17400
P_90	24500	22200	24600	10600	30700	22600
P_95	29400	27100	29500	39400	35100	26800
IGINI	0,349	0,338	0,349	0,361	0,437	0,314
P_MIXT		0,043	0,957			

	2004 T	2004 CLM	2004 R_E	2004 CLM_e	2004 R_E_e	2004 C
PROB	1	1	1	0,123	0,123	0,877
MEDIA	14739	12633	14834	5635	23454	13619
DT	9278	8184	9313	6734	13280	7883
DMGINI	9568	8832	9588	4893	12632	8560
P_05	3800	2300	3900	0	12400	3500
P_10	5400	3900	5500	600	13000	5200
P_25	8500	6800	8500	2800	14500	8000
P_50	13000	10900	13100	4700	17600	12000
P_75	19000	17300	19100	6800	27800	18200
P_90	25800	23000	26000	8600	35800	23600
P_95	31100	27900	31200	9700	50800	28500
IGINI	0,325	0,35	0,323	0,434	0,269	0,314
P_MIXT		0,043	0,957			

	2005 T	2005 CLM	2005 R_E	2005 CLM_e	2005 R_E_e	2005 C
PROB	1	1	1	0,109	0,109	0,891
MEDIA	14998	12816	15097	5758	26766	13676
DT	9305	8066	9345	7718	12957	7676
DMGINI	9579	8634	9608	5267	12178	8258
P_05	3800	2300	3900	0	12100	3600
P_10	5500	4200	5600	0	13800	5300
P_25	8700	7100	8800	2700	20400	8300
P_50	13400	11600	13500	5000	24200	12600
P_75	19300	17100	19400	6600	30100	17700
P_90	26000	22400	26100	7900	40400	22900
P_95	31400	27300	31500	8700	48400	27600
IGINI	0,319	0,337	0,318	0,457	0,227	0,302
P_MIXT		0,043	0,957			

Notas y leyendas: Todas las cantidades están expresadas en euros constantes de 2006.

T representa la Distribución de ingresos agregada (España)
CLM representa la Distribución de ingresos de Castilla-La Mancha
R_E representa la Distribución de ingresos agregada del Resto de España
CLM_e representa la Distribución diferencial de ingresos de Castilla-La Mancha
R_E_e representa la Distribución diferencial de ingresos del Resto de España
C representa la Distribución común (equiparable) de ingresos de Castilla-La Mancha

Fuente: Elaboración Propia, a partir de los datos de las ECV del INE

También se observa cómo esta reducción, en mayor medida, se va obteniendo por la translación de parte de la masa de elementos diferenciales de Castilla-La Mancha (de rentas más bajas: medias de 8.616 en 2003[10], 5.635 en 2.004 y 5758 en 2005 con menores Desviaciones Típicas y Diferencias Medias de Gini que las distribuciones de la región completa) hacia zonas más centrales, eliminándose prácticamente las diferencias en el intervalo 10.000-20.000 (lo que hace crecer en esta zona central la masa de elementos equiparables). Sin embargo, se observa también que, en menor proporción, las diferencias aumentan, a favor del resto de España, en las colas superiores de las distribuciones comparadas.

Así pues, es difícil responder sencillamente con una única respuesta a la pregunta de si éstas distribuciones están convergiendo o no. En algunos aspectos parece que sí, pero en otros no. ¿Cómo podríamos precisar en qué sentido lo hace y en qué sentido no? En el próximo apartado trataremos de emplear una vía de análisis en este sentido

4.3. Proximidad y dinámica de la distribución personal de ingresos de Castilla-La Mancha frente a la del resto de España.

Para tratar de responder sintéticamente a la pregunta de si están convergiendo las distribuciones, y tratar de profundizar sobre en qué sentido lo hace o deja de hacer, se ha adoptado la media de proximidad propuesta en el apartado 3.1 de este trabajo y se ha procedido a analizar los resultados de su aplicación, según su contenido informativo.

Así pues, en primer lugar se ha evaluado dicha medida de disimilaridad para cada pareja de funciones de densidad estimadas y que representan los comportamientos de las "rentas personales disponibles equivalentes" (reales) de las poblaciones comparadas en los 3 años de observación. Para ello hemos empleado el programa SAS/IML[11] con el que hemos construido la matriz que recoge la totalidad de los coeficientes de disimilaridad calculados entre todas ellas, y que se presenta en el Cuadro nº 2.

De forma general, se observa un rango de disimilaridades variado: alrededor de unas pocas decenas de euros reales de 2006, cuando comparamos la distribución de una población con la propia de los demás años; y en torno a unas pocas centenas de euros reales de 2006, cuando se comparan las distribuciones de las distintas poblaciones.

Cuadro nº 2

Disimilaridades entre Castilla-La Mancha y el resto de España, según las ECV.

	CLM-2003	R_E-2003	CLM-2004	R_E-2004	CLM-2005	R_E-2005
CLM-2003	0					
R_E-2003	339,72	0				
CLM-2004	48,01	269,22	0			
R_E.-2004	300,74	89,96	288,48	0		
CLM-2005	57,67	169,91	68,90	205,84	0	
R_E-2005	389,28	95,09	370,33	11,67	262,39	0

Fuente: Elaboración propia, a partir de los datos de la ECV.

Como puede apreciarse en los recuadros sombreados del cuadro nº 2, se confirma que la disimilaridad entre las regiones ha venido decreciendo año a año y que, consecuentemente, las distribuciones de rentas de las poblaciones comparadas están convergiendo en el sentido que mide la proximidad empleada[12].

Aunque el número de años de observación es aún reducido, la metodología que proponemos para analizar el proceso de convergencia entre las poblaciones comparadas está orientada, no sólo a medir la proximidad de las distribuciones en cada momento dado, sino también a describir el proceso dinámico por el que se ha llegado a esa situación a lo largo del tiempo: si las pautas de acercamiento o alejamiento se han mantenido sostenidamente a lo largo del periodo o no; si todas las poblaciones caminan hacia un patrón similar o no; si los caminos son relativamente parecidos o no; etc.

Así pues, con el objetivo de tener en cuenta la máxima información sobre la trayectoria de los comportamientos aproximativos, o de alejamiento, entre estas distribuciones de ingresos, es aconsejable considerar todas las disimilaridades calculadas entre las distribuciones de las poblaciones en los diferentes años de observación. En consecuencia, se ha empleado completamente la matriz triangular de coeficientes de disimilaridad calculados entre las 6

[11] SAS/IML® es producto registrado de SAS Institute Inc., Cary, NC, USA.

[12] Aparentemente, la convergencia se ralentiza desde el año 2004; pero recordemos que los datos de 2003 están en tela de juicio hasta aclarar la cifra sospechosamente alta de hogares con rentas nulas que presenta.

distribuciones de "Renta Personal Equivalente Disponible" (real) de las 2 poblaciones en cada uno de los 3 años observados de la encuesta y, para simplificar la interpretación de dichos coeficientes de disimilaridad, hemos empleado el modelo ALSCAL de la familia de técnicas de escalamiento multidimensional[13], según el procedimiento establecido en Young, F.W. Lewyckyj, R. y Takane, Y. (1986).

Este modelo tratará de encontrar, en cierto espacio p-dimensional, las coordenadas representativas de cada una de las distribuciones de las poblaciones en cada año de la encuesta (puntos) de forma que las distancias euclídeas entre cada dos de estos puntos, d_{ij}, o alguna posible transformación monótona de éstas, $T(d_{ij})$, reproduzcan lo más fielmente posible las disimilaridades realmente observadas, δ_{ij}, entre las distribuciones comparadas.

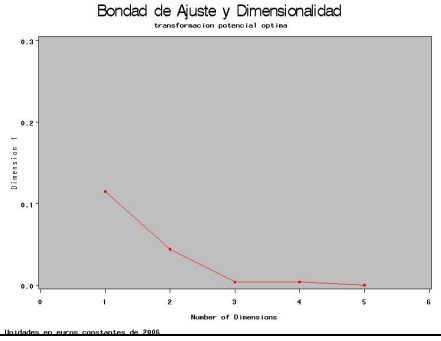
Para su estimación se ha empleado el procedimiento MDS de SAS/STAT[7], habiéndose probado varias transformaciones monótonas (identidad, afín, lineal, potencial y monótona-escalonada), así como varias dimensiones del espacio factorial de representación buscado (entre 1 y 5). El criterio de bondad de ajuste empleado ha sido la medida de Stress-1 de Kruskal[14] y, de acuerdo con él, se ha encontrado que en todos los espacios considerados, la mejor aproximación siempre la proporcionó el modelo de transformación potencial:

$$\delta_{ij} \cong T(d_{ij}) = s\cdot\left(d_{ij}\right)^{p} \qquad \text{o equivalentemente,} \quad \log\left(\delta_{ij}\right) \cong \log(s) + p\cdot\log\left(d_{ij}\right).$$

Este modelo, en función de la dimensión del espacio considerado, ha proporcionado los valores del criterio de bondad de ajuste que se recogen en el Cuadro nº 3.

Cuadro nº 3

Dimen-siones	Stress-1 de Kruskal
1	0.11464
2	0.04399
3	0.00367
4	0.00367
5	0.00000



Fuente: Elaboración propia, a partir de los datos de la ECV.

De acuerdo con estos resultados, el modelo ha sido finalmente ajustado para un espacio de 3 dimensiones, empleando la transformación logarítmica lineal,

$$\delta_{ij} \cong 33{,}77\cdot\left(d_{ij}\right)^{1{,}81} \qquad \text{o equivalentemente,} \quad \log\left(\delta_{ij}\right) \cong \log(33{,}77) + 1{,}81\cdot\log\left(d_{ij}\right).$$

Ello ha permitido obtener las coordenadas resultantes, en el espacio factorial óptimo, para cada distribución de ingresos de cada población en cada año de la encuesta, que se muestran en el cuadro nº 4.

[13] Los fundamentos del escalamiento multidimensional fueron propuestos por Torgenson, W.S. (1958). Para una introducción a estos métodos, puede consultarse Kruskal, J.B & Wish, M (1978).

[14] "Badness of Fit Criterion" en terminología del procedimiento MDS, cuando se seccionan las opciones Fit=1, Formula=1 y Coef=Identity, cuya formulación es:

$$S_1 = \left(\frac{\sum_{ij} \left(\delta_{ij} - T(d_{ij})\right)^2}{\sum_{ij} \delta_{ij}^2} \right)^{1/2}$$

Kruskal caracteriza la bondad de ajuste del modelo en base a este criterio como sigue: 0=perfecto, 0.025=excelente, 0.05=bueno, 0.1=aceptable, 0.2=pobre.

Cuadro nº 4

		DIM1	DIM2	DIM3
CLM	2003	1,9051	-0,5685	-0,2236
R_E	2003	-1,2324	1,1199	0,0174
CLM	2004	1,6858	0,1833	0,7025
R_E	2004	-1,4253	-0,5932	0,1226
CLM	2005	0,9982	0,3535	-0,5998
R_E	2005	-1,9314	-0,4949	-0,0189

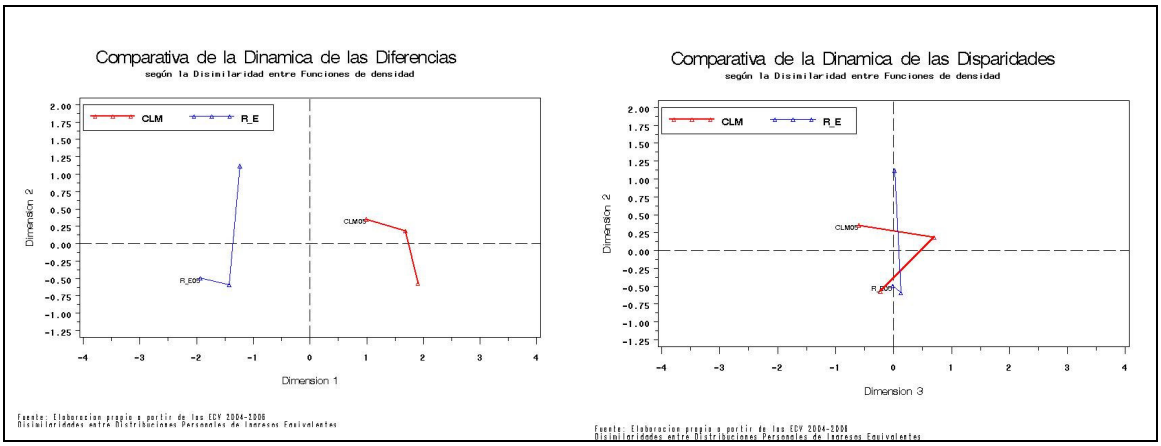
Distancias en cada dimensión		DIM1	DIM2	DIM3
CLM-R_E	2003	3,1376	1,6884	0,2410
CLM-R_E	2004	3,1110	0,7766	0,5799
CLM-R_E	2005	2,9296	0,8484	0,5809

Fuente: Elaboración propia, a partir de los datos de la ECV.

Así, uniendo consecutivamente las coordenadas que una determinada población presenta en los sucesivos años de la encuesta, podemos visualizar la trayectoria de su comportamiento y analizarla comparativamente con las de las demás.

En el gráfico nº 3 se representan dichas trayectorias en los planos de proyección formados por las dimensiones 1-2, y 3-2. Las trayectorias comienzan en las posiciones de las poblaciones en el año 2003 y evolucionan hasta las posiciones finales en el año 2005, marcadas con sus indicativos de nombre y año (CLM05 y R_E05).

Gráfico nº 3



Fuente: Elaboración propia, a partir de los datos de la ECV.

De una primera ojeada, se observa un movimiento de derecha a izquierda a lo largo de la 1ª dimensión, mayor en el caso de Castilla-La Mancha, de forma que las posiciones finales se encuentran más cercanas sobre esta dimensión que en el año 2003

En la segunda dimensión, las trayectorias evolucionan en sentidos contrarios proporcionando también situaciones relativas más cercanas en 2006 que en 2003, aunque de sentido opuesto.

En la tercera dimensión, la variación de posición de Castilla-La Mancha es errática, tendidendo a separarse de la posición del Resto de España; mientras que para ésta, la variación de su posición es prácticamente nula.

4.4. Interpretación de las dimensiones diferenciadoras.

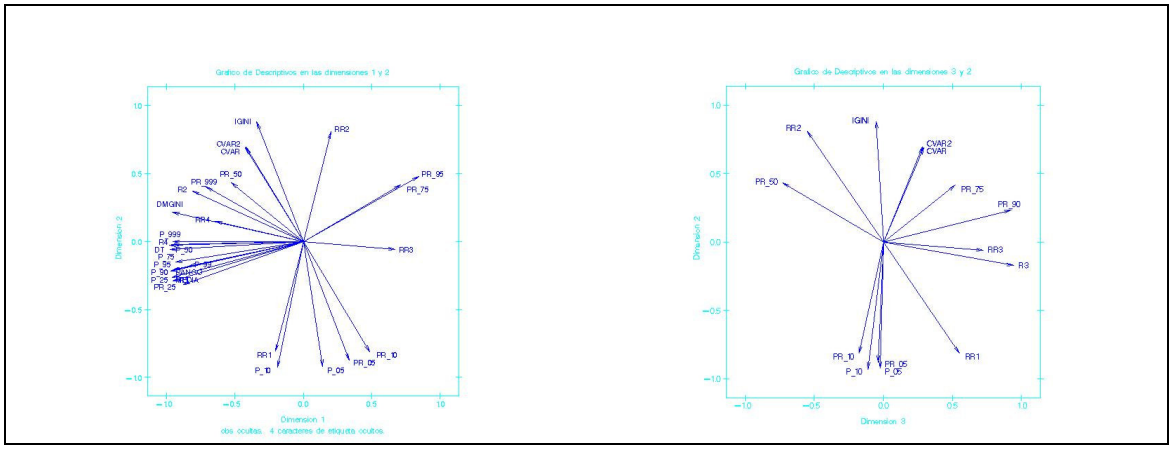
Para tratar de comprender mejor lo que muestran estos gráficos, podemos indagar sobre las posibles relaciones existentes entre las coordenadas de las poblaciones en cada momento y los indicadores estadísticos y económicos disponibles, que puedan ayudarnos a interpretar las dimensiones de este espacio donde se manifiestan los principales aspectos de sus diferencias.

Para ello, y a modo de ejemplo, se han calculado las correlaciones que presentan las 2 primeras dimensiones (las más diferenciadoras) del espacio factorial considerado con las siguientes variables descriptivas: media aritmética (media), cuantiles de órdenes 0.001, 0.01, …,0.99, 0.999 (P_001, P_01, P_05, P_10, P_25, P_50, P_75, P_90, P_95, P_99, P_999),

cocientes de estos cuantiles y la media aritmética (Pr_001, Pr_01, Pr_05, Pr_10, Pr_25, Pr_50, Pr_75, Pr_90, Pr_95, Pr_99, Pr_999), desviación típica (dt), rango o recorrido (rango), recorrido inter-cuantílico P_25-P_001 (r1), recorrido inter-cuantílico P_50-P_25 (r2), recorrido inter-cuantílico P_75-P_50 (r3), recorrido inter-cuantílico P_999-P_75 (r4), coeficiente de variación de Pearson (cvar), cocientes de los anteriores recorridos inter-cuantílicos y la mediana (rr1, rr2, rr3, rr4), distancia media de Gini (dmgini), índice de concentración de Gini (igini) y cuadrado del coeficiente de variación de Pearson (cvar2).

El gráfico nº 4 muestra las representaciones gráficas de dichas variables descriptivas, en función de sus correlaciones con las dimensiones (1-2 y 3-2) del espacio factorial; y nos permite profundizar en el significado estadístico de las mismas[15].

Gráfico nº 4



Fuente: Elaboración propia, a partir de los datos de la ECV.

Observando las proyecciones sobre las dimensiones 1 y 2, podemos ver que la primera dimensión se encuentra alta y negativamente correlacionada (con correlaciones cercanas a –1) con la práctica totalidad de medidas de localización y de dispersión absolutas. Por otro lado, con respecto de las medidas de localización y dispersión relativas, vemos que la correlación es negativa con los indicadores relativos a las posiciones extremas, y positiva con los indicadores relativos a la zona que va desde la mediana al percentil 99. Ello provoca que una población se situará tanto más negativamente en esta dimensión cuanto más se desplace y expanda hacia la derecha la totalidad de su distribución, acortando la diferencia relativa con respecto de la media existente entre la mitad izquierda de la distribución y la mayor parte de la mitad derecha, con la excepción de sus valores más extremos, que se distancian mucho más. Así, una coordenada más negativa sobre esta dimensión implica un incremento general de los ingresos y un reparto relativamente más igualitario en la población "regular", con cuyo término queremos hacer referencia a toda la población resultante una vez excluidos los valores extremos de la cola derecha, que concentran cada vez más recursos. En el sentido de que informa esta dimensión, la población agregada del Resto de España se manifiesta claramente más negativa que Castilla-La Mancha en todo el período evidenciando unos mayores ingresos y un mayor reparto de los mismos, si excluimos los recursos acumulados en los extremos de la cola derecha. En cualquier caso, las dos poblaciones evolucionan en la misma dirección hacia una mayor riqueza y mejor reparto entre la población "regular" y, según muestra el cuadro nº 4 para esta dimensión, su distancia se ha venido acortando lenta, pero progresivamente, a lo largo de los años observados (desde 3,14 a 2,93).

Si ahora nos fijamos en la segunda dimensión, los estadísticos más correlacionados positivamente con ella son las medidas de desigualdad y el recorrido relativo del segundo intervalo intercuartílico; y negativamente, los estadísticos informantes acerca de las zonas de menores ingresos (percentiles absolutos y relativos a la media de órdenes 5 y 10, y el recorrido relativo del primer intervalo intercuartílico). Así pues, una población se situará más negativamente sobre esta dimensión si disminuye su desigualdad, se estrecha el recorrido

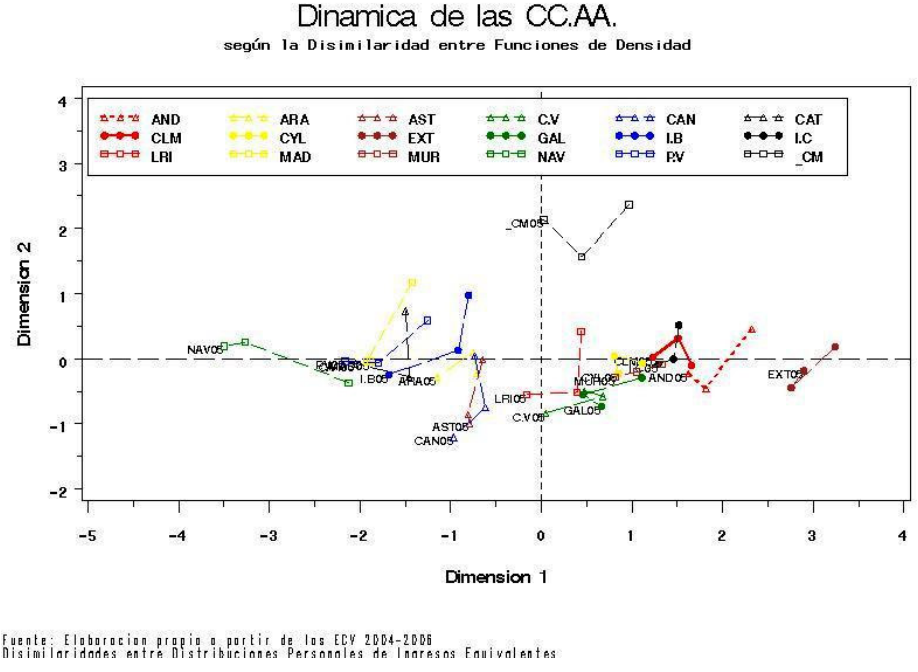
[15] Para no hacer innecesariamente densas estas representaciones gráficas, sólo se representan aquellas variables descriptivas para las que al menos una de las correlaciones con alguna dimensión representada es superior a 0.5.

relativo a la media entre el primer cuartil y la mediana, y se amplía el recorrido relativo a la media entre el mínimo y el primer cuartil, aumentando asimismo los percentiles de bajo orden; lo que significa que las clases bajas mejoran sus ingresos con relación a la media en mayor medida que lo hacen las clases medias-bajas, disminuyendo así la desigualdad. En este sentido, las poblaciones evolucionan en direcciones opuestas, observándose un primer movimiento sospechosamente mucho mayor del Resto de España, y que debemos considerar con precaución, pues ya hemos alertado, en el apartado 4.2, sobre la sospechosa acumulación de ceros para esta población en 2003, que desaparece rápidamente al año siguiente. En cualquier caso y a lo largo del período de observación, mientras la población agregada del resto de España evoluciona disminuyendo su desigualdad y mejorando la situación de sus elementos de rentas más bajas (hasta el primer decil), en Castilla-La Mancha, sus elementos homólogos de rentas más bajas se empobrecen ligeramente y el grado de desigualdad de la región no varía prácticamente. Pero sus posiciones iniciales, cualitativamente opuestas a las finales, eran más distantes; reduciéndose su distancia en esta dimensión a lo largo de los años observados (de 1,69 a 0,85).

Sobre la tercera dimensión, observando el gráfico de proyecciones de los indicadores estadísticos sobre el plano de dimensiones 3 y 2, vemos que las mayores correlaciones positivas se dan con aquéllos que informan sobre la zona que va desde la mediana al decil noveno; mientras que la mayor negativa se da, de forma singular, para el valor relativo de la mediana a la media. Así, una población se situará más positivamente sobre esta dimensión cuanto más se aleje la media de la mediana (menor cociente mediana/media), mayor amplitud absoluta y relativa tenga el tercer intervalo intercuartílico, y mayor valor relativo a la media presente el decil noveno; es decir, cuanto más mejoren las posiciones de las clases media-alta y alta (excluida la población extrema). Observando el gráfico nº 3, vemos que la población agregada del resto de España no varía prácticamente en esta dimensión, mientras que Castilla-La Mancha oscila con un movimiento de ida y vuelta alrededor de la posición de aquélla, informando de la extraña alteración que aparece y desaparece durante los años observados, en el intervalo 11.000-23.000 de sus funciones de densidad. En esta dimensión, las posiciones relativas se alejan como reflejo de un mayor crecimiento relativo en el Resto de España de las clases medias-alta y alta, excluyendo sus elementos extremos (de 0,24 a 0,58).

Es de observar que en el plano de las dos primeras dimensiones, las medidas relativas de dispersión basadas en los recorridos intercuartílicos actúan como verdaderos puntos cardinales (RR2-Norte, RR1-Sur, RR3-Este y RR4-Oeste), propiciando que la dimensión 1 informe fundamentalmente de las variaciones internas relativas entre las clases media-alta y alta; así como la dimensión 2 informa de las variaciones internas relativas entre las clases baja y media-baja.

Gráfico nº 5



Fuente: Elaboracion propia a partir de los ECV 2004-2006
Disimilaridades entre Distribuciones Personales de Ingresos Equivalentes

Fuente: Elaboración propia, a partir de los datos de la ECV.

Un análisis más detallado se está realizando en el contexto del proyecto al que se adscribe este trabajo, en el que se compara la dinámica de Castilla-La Mancha con las del resto de comunidades autónomas (incluidas Ceuta y Melilla), observándose pautas similares (ver gráfico nº 5): en la primera dimensión, movimiento de derecha a izquierda de prácticamente todas las comunidades; y en la segunda, movimientos generalizados de arriba abajo para casi todas las regiones, con muy pocas excepciones entre las que se encuentra Castilla-La Mancha. Todo ello no hace sino confirmar el análisis de la comparativa realizada entre Castilla-La Mancha y el resto de España agregado.

4.5. Análisis a partir de las componentes del índice de Gini de la población agregada.

Si atendemos exclusivamente a la renta media, observamos que las de las poblaciones comparadas divergen ya que, aunque ambas crecen, Castilla-La Mancha lo hace más lentamente. Si atendemos a los índices de Gini, se observa que Castilla-La Mancha parece estabilizada, mientras que en el resto de España parece decrecer ligeramente la desigualdad.

Considerados exclusivamente estos indicadores podríamos concluir que las distribuciones más bien divergen. Sin embargo, estos dos indicadores están muy influenciados por las observaciones extremas y pueden estar enmascarando lo que está ocurriendo con la gran mayoría de la población.

A analizar éste punto nos ayuda la descomposición [21] de la distancia media de Gini de la población agregada, en función de la descomposición diferencial propuesta en este trabajo.

De acuerdo con la expresiones [21], [22] y [23] del apartado 3 de este trabajo, donde se expresa la descomposición del índice de Gini de la población agregada en función de las poblaciones diferenciales y sus características, hemos procedido a extraer las señales que dicha descomposición nos proporciona y que hemos recogido en el cuadro nº 5.

Ahora vemos que las poblacionesÿÿiferenciajes han ido paulatinamente disminuyendo sus tamaños, año a año; y que sus distancias medias de Gini internas han ido paralelamente disminuyendo. Consecuentemente, las poblaciones diferenciales son cada año más pequeñas y menos dispersas, a la vez que más distantes entre sí.

Cuadro nº 5

	Pesos de Mixtura	Medias	Indices de Gini	Distancias	Medias de Gini	Proporciones poblaciones diferenciales	Indicadores de Solapamiento	Medida de Proximidad
Año		MU_T	IG_T		CLMe-R_Ee			
2003		13826,6	0,3493		15186,9			
2004		14738,9	0,3246		18914,6			
2005		14998,3	0,3193		22272,1			
Año	P_CLM	MU_CLM	IG_CLM	CLM	CLM_e	P	OV_CLM/_T	
2003	0,0430	12563,3	0,3381	8494,8	6228,1	0,1496	0,9843	
2004	0,0431	12633,1	0,3496	8832,0	4893,1	0,1235	0,9919	
2005	0,0431	12816,3	0,3368	8633,9	5266,7	0,1085	0,9931	
Año	P_R_E	MU_R_E	IG_R_E	R_E	R_E_e	P	OV_R_E/_T	δ_CLM _R_E
2003	0,9570	13883,4	0,3494	9702,0	15239,1	0,1496	0,9985	339,7
2004	0,9569	14833,7	0,3232	9588,3	12631,9	0,1235	0,9991	288,5
2005	0,9569	15096,5	0,3182	9607,8	12178,3	0,1085	0,9994	262,4

Notas y leyendas: Todas las cantidades están expresadas en euros constantes de 2006.

_T hace referencia a la Distribución de ingresos agregada (España)

_CLM hace referencia a la Distribución de ingresos de Castilla-La Mancha

_R_E hace referencia a la Distribución de ingresos agregada del Resto de España

_CLM_e hace referencia a la Distribución diferencial de ingresos de Castilla-La Mancha

_R_E_e hace referencia a la Distribución diferencial de ingresos del Resto de España

Fuente: Elaboración Propia, a partir de los datos de las ECV del INE

En cualquier caso, esta mayor diferencia de las poblaciones diferenciales queda mitigada por la cada vez menor importancia relativa de sus tamaños, de forma que la medida de proximidad propuesta en este trabajo, como medida de convergencia, indica que ambas poblaciones evolucionan aproximándose, año a año.

Ello lo corroboran también los indicadores de solapamiento, que se acercan progresivamente a 1 a lo largo de los años observados.

Así pues, en general, las distribuciones se aproximan en el sentido de que las poblaciones comparables crecen paulatinamente de tamaños (a la vez que decrecen las diferenciales), en el sentido de que los solapamiento de ambas poblaciones con la población agregada se aproximan paulatinamente a ser totales, y en el sentido de que la medida de disimilaridad propuesta, que evalúa conjuntamente los efectos conjuntos de los tamaños de las poblaciones diferenciales y su separación, está decreciendo también, año a año.

Consecuentemente, la divergencia observada en media y en desigualdad se debe prácticamente en exclusiva a las subpoblaciones diferenciales de cada población específicas (de tamaños decrecientes desde el 14,96% en 2003 hasta el 10,85% en 2005). Por una parte, la población diferencial del resto de España, con rentas medias que crecen de 17.444 a 26.766 € reales de 2006, frente a rentas medias del resto de esta población (población equiparable) que tan sólo crece de 13.257 a 13.676 €, hacen crecer la media de la población de 13.883 a 15.097 €. Y por otra parte, la población diferencial de Castilla-La Mancha, con rentas medias que decrecen de 8.618 a 5.758 € reales de 2006, frente a rentas medias del resto de esta población (población equiparable) que tan sólo crece de 13.257 a 13.676 €, hacen decrecer la media de la población de 12.563 a 12.816 €.

4.6. Conclusiones.

Emplear exclusivamente la media como indicador de convergencia entre poblaciones es, cuando menos, imprudente. La composición y dinámica interna de las distribuciones puede ser tan variada que aconseja realizar análisis multidimensionales de las comparaciones.

A ello nos puede ayudar los instrumentos y la metodología introducidas en este trabajo: la descomposición diferencial de las poblaciones comparadas, la definición de una medida de proximidad que informa de aspectos relevantes de las diferencias observadas, el establecimiento de una metodología para analizar multidimensionalmente las diferencias observadas, la descomposición de las diferencias en indicadores o señales con contenidos informativos útiles para la comparación, etc.

Desde el punto de vista de la aplicación de estos instrumentos y metodología, los resultados parecen razonables, habiéndose detectado ciertas cuestiones que deben ser abordadas en mayor profundidad. Por un lado, parece aconsejable revisar la abundancia de ceros en los datos originales, especialmente en la población agregada del Resto de España. Por otro lado, las estimaciones no paramétricas estándar de las funciones de densidad mediante núcleos estadísticos parecen ser muy sensibles a los datos de muestras relativamente reducidas, como es el caso de las regiones en las ECV; lo que sugiere la necesidad de ampliar la ventana de estimación de estos métodos para suavizar los cambios anuales de las distribuciones, o de emplear ajustes a modelos paramétricos adecuados.

REFERENCIAS BIBLIOGRÁFICAS.

Callealta, F.J. (2006) "A new measure of dissimilarity between distributions: application to the analysis of income distributions convergence in the European Union" en *Distribution Models Theory.* World Scientific Publishing. Chapter 8, pp. 125-160.

Dagum, C. (1980) "Inequality Measures Between income Distributions with Applications". *Econometrica*, Vol. 48 (7), pp. 1971-1803.

Kruskal,J.B.; Wish,M. (1978) *Multidimensional Scaling. Sage University Paper series on Qualitative Applications in the Social Sciences*, *07-011*. Sage Publications.

Pena, B.; Callealta, F.J.; Casas, J.M.; Merediz, A.; Nuñez, J.J. (1996) *Distribución Personal de la Renta en España*. Pirámide.

Shorrocks,A.F. (1982) "On the Distance between Income Distributions". *Econometrica*, Vol. 50(5), pp. 1337-1339.

Silverman, B.W.(1986) *Density Estimation for Statistics and Data Analysis.* London:Chapman and Hall.

Yitzhaki, S. (1994) "Economic distance and overlapping of distributions". *Journal of Econometrics,* 61. pp. 147-150.

Young,F.W; Lewyckyj,R.; Takane,Y. (1986) "The ALSCAL Procedure". *SUGI Supplemental Library User´s Guide. Version 5 Edition.* SAS Institute Inc.

ESPECIFICACIÓN DE UNA DISTRIBUCIÓN TRAPEZOIDAL, A PARTIR DE LAS ESTIMACIONES SUBJETIVAS DEL MÉTODO PERT

RAFAEL HERRERÍAS PLEGUEZUELO
E-mail: rherreri@ugr.es
JOSÉ MANUEL HERRERÍAS VELASCO
E-mail: jmherrer@ugr.es
Departamento de Métodos Cuantitativos para la Economía y la Empresa
Facultad de Ciencias Económicas y Empresariales
Universidad de Granada

Resumen

En primer lugar, en este trabajo se determina completamente una distribución de tipo Trapezoidal, Tp (a, m_1 , m_2 , b), a partir de las tres típicas estimaciones subjetivas del método PERT, sobre los valores mínimo, máximo y más probable del campo de variación de la variable objeto de estudio. Posteriormente, se destacan algunas de las peculiaridades de esta distribución que pueden resultar ventajosas en comparación a la distribución trapezoidal presentada por Callejón, Pérez y Ramos en 1996, en la que consideran el punto medio del intervalo del recorrido de la variable y su asimetría, para dilucidar los valores de m_1 , y m_2 . Por último se realizan comparaciones con algunas distribuciones que han sido utilizadas, como modelos probabilísticos, en el método PERT, tales como la distribución triangular, la distribución beta PERT y la distribución trapezoidal CPR.

Palabras clave: Estimación subjetiva, Método PERT, ambiente de incertidumbre, distribución trapezoidal.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

First, this paper completely determines the distribution of type Trapezoidal, Tp (a, m_1, m_2, b), from three typical subjective estimates of PERT method, the minimum, maximum and most probable values of the variable under study. Subsequently, it highlights some of the peculiarities of this distribution may be advantageous compared to the distribution presented by Callejón, Pérez and Ramos in 1996, which considered the midpoint of the range of the variable and its asymmetry to determine the values of m_1 and m_2. Finally comparisons are made with some distributions that have been used as probabilistic models in the PERT method, such as the triangular distribution, the beta PERT distribution and the trapezoidal CPR distribution.

Thematic Area: Quantitative Methods for Economics and Business.

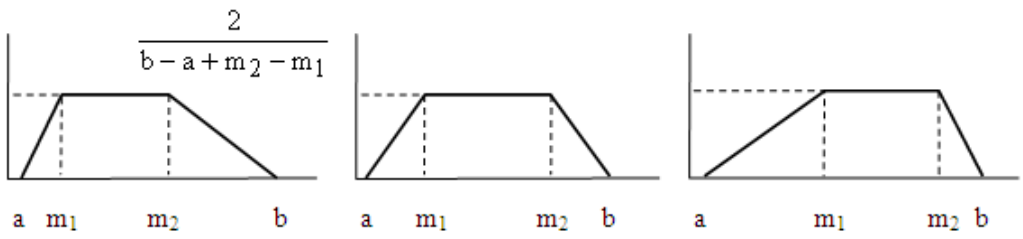
1. INTRODUCCIÓN

En la mayoría de los problemas en ambiente de incertidumbre: duración de un trabajo (T), valor de mercado de un bien inmobiliario o rústico (V), flujo neto de caja de un periodo futuro (Q_t), etc., no se disponen de datos suficientes, por no existir o por ser escasos, para realizar estimaciones fiables de los parámetros de la distribución de probabilidad que sigue la variable objeto de estudio: T, V, Q_t , etc. En cualquiera de estos dos supuestos sobre los datos, el investigador puede recurrir a las estimaciones subjetivas proporcionadas por un experto y relativas al menor valor (a), al mayor valor (b) y al valor modal (m) de la variable de interés. Con ayuda de estos tres valores y, a veces, con hipótesis adicionales se ajusta una distribución

de probabilidad para la variable. Entre las distribuciones utilizadas en este menester se encuentran las de tipo rectangular, triangular, trapezoidal, beta, normal y TSP (introducida por van Dorp y Kotz en 2002), con sus extensiones y generalizaciones respectivas.

La distribución trapezoidal, recibe este nombre por la forma de la gráfica de su función de densidad, véase figura 1, verifica el conocido aserto de Keynes: "más vale acertar aproximadamente que errar exactamente", en el sentido de que para estimar subjetivamente el valor modal de la variable, estimación necesaria e imprescindible para usar el método PERT, resulta menos comprometido para el experto determinarlo mediante un intervalo (m_1 , m_2) que por un solo valor puntual (m).

Figura 1: Distribución trapezoidal simétrica y con asimetría a la derecha e izquierda

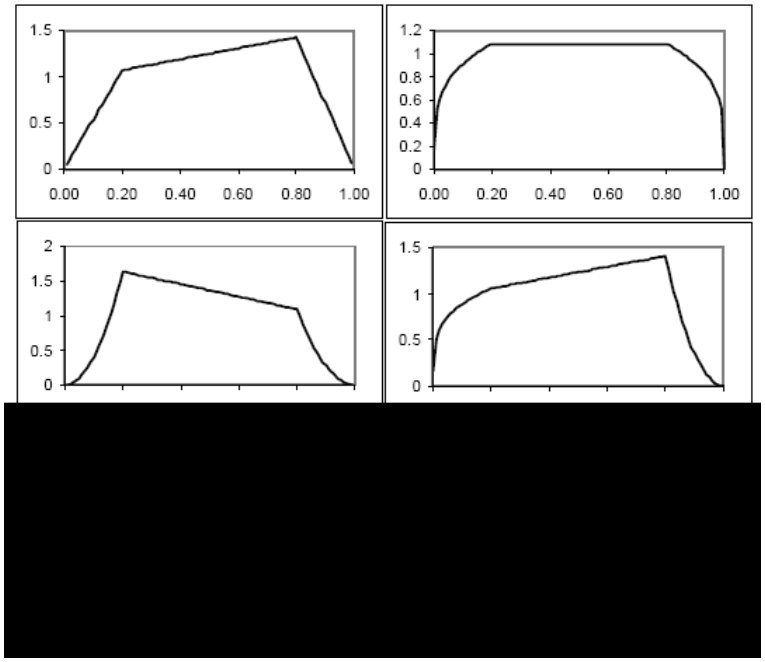


A la distribución trapezoidal se le han encontrado aplicaciones en los mismos campos en los que se han utilizado las distribuciones rectangular y triangular, esto es, en valoración de bienes inmobiliarios, véase Lozano (1996), que usa el método de valoración denominado de las dos betas, véase Ballestero y Caballer (1982), en selección de inversiones como modelo probabilístico de la variable flujo de caja neto, véase Herrerías y Calvete (1987), mediante el método del valor capital, véase Suárez (2007), también se usa como modelo en conjuntos difusos, junto a la distribución triangular, véase Kauffman y Gil Aluja (1986), más recientemente se ha introducido como una herramienta para el control, véase Akcayol, Elmas, Erden y Kart (2004).

Otros trabajos relativos a la distribución trapezoidal, en relación con el método PERT, son los siguientes: Herrerías, Palacios y Herrerías (2002), en el que se especifica la distribución beta utilizando la distribución trapezoidal, Callejón, Pérez y Ramos (1996), en el que se obtiene una distribución trapezoidal a partir de las tres estimaciones periciales del método PERT, y en relación con el método de valoración de las dos betas puede consultarse Herrerías, García, Cruz y Herrerías (2001).

Un renovado interés por esta distribución ha llevado a su generalización, por van Dorp y Kotz (2003), admitiendo tramos no rectilíneos, véase figura 2, en uno o varios de sus tres tramos: (a, m_1), (m_1 , m_2) y (m_2 , b) del recorrido (a, b) de la variable. En esta línea se sitúa el trabajo de van Dorp, Cruz, García y Herrerías (2007), en el que se presenta un procedimiento de elicitación para obtener los parámetros de la distribución trapezoidal con tramo central rectilíneo.

Figura 2: Distribución tipo trapezoidal con algún tramo no rectilíneo



El presente trabajo se centra en las distribuciones trapezoidales que caracterizadas por tener una gráfica del tipo de la figura 1. Se trata, en definitiva, de especificar los parámetros de la distribución trapezoidal Tp(a, m_1 , m_2 , b), conocidos o estimados subjetivamente los valores "a", "m" y "b". En otras palabras, se trata de encajar el valor de "m" en el intervalo (m_1 , m_2) de manera diferente a la realizada por Callejón, Pérez y Ramos en (1996) en la distribución bautizada como trapezoidal CPR.

2. LA DISTRIBUCIÓN TRAPEZOIDAL

La distribución trapezoidal responde a la siguiente función de densidad:

$$f(x)=\begin{cases} k\dfrac{x-a}{m_1-a} & \text{si } a < x \le m_1 \\ k & \text{si } m_1 \le x \le m_2 \\ k\dfrac{b-x}{b-m_2} & \text{si } m_2 \le x < b \end{cases} \qquad (1)$$

donde $k = 2\left(b-a+m_2-m_1\right)^{-1}$

Puede estandarizarse el recorrido (a, b) de la variable X, mediante el cambio de variable lineal:

$$Z=\frac{X-a}{b-a} \qquad (2)$$

al intervalo (0, 1) y esta nueva variable tiene una función de densidad:

$$f(z)=\begin{cases}\dfrac{k^{*}z}{m_1^{*}} & \text{si } 0<z\le m_1^{*}\\ k^{*} & \text{si } m_1^{*}\le z\le m_2^{*}\\ k^{*}\dfrac{1-z}{1-m_2^{*}} & \text{si } m_2^{*}\le z<1\end{cases} \qquad (3)$$

donde se ha notado por m_i^{*} (i=1, 2), los valores estandarizados de m_i (i=1, 2) obtenidos aplicando el cambio de variable visto en (2).

Obteniéndose la constante normalizadora, $k^{*}=2\left(1+m_2^{*}-m_1^{*}\right)^{-1}$, y los sucesivos momentos respecto al origen mediante la expresión:

$$\alpha_n=E[Z^n]=\frac{k^{*}}{(n+1)(n+2)}\left[\left(m_2^{*}\right)^{n+1}+\left(m_2^{*}\right)^{n}+....+1-\left(m_1^{*}\right)^{n+1}\right]=$$
$$=\frac{k^{*}}{(n+1)(n+2)}\left[\frac{1-\left(m_2^{*}\right)^{n+2}}{1-m_2^{*}}-\left(m_1^{*}\right)^{n+1}\right] \qquad (4)$$

sustituyendo en (4) n = 1 ó 2 se obtienen:

$$\alpha_1=E[Z]=\frac{1}{3}\left(\frac{\left(m_2^{*}\right)^2+m_2^{*}+1-\left(m_1^{*}\right)^2}{1+m_2^{*}-m_1^{*}}\right) \qquad (5)$$

$$\alpha_2=E[Z^2]=\frac{\left(m_2^{*}\right)^3+\left(m_2^{*}\right)^2+m_2^{*}+1-\left(m_1^{*}\right)^3}{6(1+m_2^{*}-m_1^{*})} \qquad (6)$$

por lo que la varianza de la variable estandarizada será:

$$\sigma^2=\frac{1}{18}\left(1+\left(m_2^{*}\right)^2+\left(m_1^{*}\right)^2-(1+m_2^{*})m_1^{*}-\frac{2(m_2^{*}-m_1^{*})(1-m_1^{*})m_2^{*}}{(1+m_2^{*}-m_1^{*})^2}\right) \qquad (7)$$

3. UNA NUEVA DETERMINACIÓN DE LA DISTRIBUCIÓN TRAPEZOIDAL

La distribución que se presenta en este trabajo surge como una modificación de la distribución trapezoidal CPR introducida por Callejón, Pérez y Ramos (1996). Estos autores, a partir de los tres típicos valores "a", "m" y "b" del método PERT, considerando la asimetría de la distribución, de acuerdo a que "m" sea menor o mayor que el punto central, $c=(a+b)/2$, del intervalo (a, b), determinan los extremos superior e inferior del intervalo modal de la distribución trapezoidal:

- $m_1=c$ y $m_2=m$, si $c<m$, o bien
- $m_1=m$ y $m_2=c$, si $c>m$ (8)

En este trabajo, con la misma información de partida: "a", "m" y "b", se determinan, en principio, los extremos superior e inferior del intervalo modal de la distribución trapezoidal de la siguiente forma:

- $m_1=m-(b-m)/2$ y $m_2=m+(b-m)/2$ si $c<m$

- $m_1 = m-(m-a)/2$ y $m_2 = m+(m-a)/2$ si $c > m$ (9)

En el caso de que $m \approx c$, la modificación propuesta en (9) conduce a distribuciones trapezoidales tales que $m_2 - m_1 > |m-c|$, por el contrario si m es muy diferente de c, el efecto es opuesto, es decir, $m_2 - m_1 < |m-c|$.

La distribución trapezoidal CPR vista en (8) y la modificación propuesta en (9) tendrán tramos horizontales iguales, cuando:

- $m = (3b+a)/4$, si $c < m$, o bien,
- $m = (3a+b)/4$, si $c > m$

Efectivamente, si $c < m$, para que $m_2 - m_1 = m-c$, tiene que ocurrir que $(m+(b-m)/2)-(m-(b-m)/2) = m-(a+b)/2$, lo que se cumple para $m = (3b+a)/4$.

Aplicando el mismo razonamiento, en el caso de que $c > m$, la igualdad de longitud de los tramos horizontales se cumple para $m = (3a+b)/4$.

Por lo anterior, siempre que $m < (3a+b)/4$ ó $m > (3b+a)/4$, la distribución trapezoidal presentada en (9) tiene un tramo horizontal más pequeño que el proporcionado por la distribución trapezoidal CPR.

En el caso de que $m \in [(3a+b)/4, (3b+a)/4]$ se tomarán por $m_1 = m-|m-c|/2$ y $m_2 = m+|m-c|/2$, con lo que la distribución propuesta tiene un tramo horizontal de igual longitud que el de la distribución trapezoidal CPR de Callejón, Pérez y Ramos.

Por todo lo que antecede, la propuesta final para determinar el intervalo (m_1 , m_2) de la distribución trapezoidal, Tp(a, m_1 , m_2 , b) , a partir de los valores "a", "m" y "b", es la siguiente:

- $\begin{matrix} m_1 = m-(m-a)/2 \\ m_2 = m+(m-a)/2 \end{matrix}$ si $m < (3a+b)/4$
- $\begin{matrix} m_1 = (3m-c)/2 \\ m_2 = (m+c)/2 \end{matrix}$ si $m \in [(3a+b)/4, c]$
- $\begin{matrix} m_1 = (m+c)/2 \\ m_2 = (3m-c)/2 \end{matrix}$ si $m \in [c, (3b+a)/4]$ (10)
- $\begin{matrix} m_1 = m-(b-m)/2 \\ m_2 = m+(b-m)/2 \end{matrix}$ si $m > (3b+a)/4$

Nótese que la distribución trapezoidal propuesta en (10) tiene como punto medio del intervalo modal, (m_1 , m_2), al propio valor de m, mientras que el punto medio del intervalo modal en la distribución trapezoidal CPR es $(m+c)/2$. Es decir, la distribución trapezoidal obtenida por (10) centra el valor modal suministrado por el experto, en contra de lo que ocurre con la trapezoidal CPR.

A partir de (2), se tiene que (9) puede expresarse de la siguiente forma:

- $m_1^* = (3m^*-1)/2$ y $m_2^* = (m^*+1)/2$ si $c^* = 0{,}5 < m^*$
- $m_1^* = m^*/2$ y $m_2^* = 3m^*/2$ si $c^* = 0{,}5 > m^*$ (11)

La distribución trapezoidal CPR y la propuesta en (9) tendrán tramos horizontales iguales, cuando $m^* = 3/4$ o bien $m^* = 1/4$

Aplicando la estandarización vista en (2) sobre el resultado (10) se obtiene la propuesta final para determinar una distribución trapezoidal estandarizada, Tp(0, m_1^* , m_2^* ,1) :

- $\begin{matrix} m_1^* = m^*/2 \\ m_2^* = 3m^*/2 \end{matrix}$ si $m^* < 1/4$

$$
\begin{array}{ll}
- & \begin{array}{l} m_1^* = (6m^* - 1)/4 \\ m_2^* = (2m^* + 1)/4 \end{array} \quad \text{si} \quad m^* \in [1/4, 1/2] \\
- & \begin{array}{l} m_1^* = (2m^* + 1)/4 \\ m_2^* = (6m^* - 1)/4 \end{array} \quad \text{si} \quad m^* \in [1/2, 3/4] \\
- & \begin{array}{l} m_1^* = (3m^* - 1)/2 \\ m_2^* = (m^* + 1)/2 \end{array} \quad \text{si} \quad m^* > 3/4
\end{array} \qquad (12)
$$

Al igual que la distribución trapezoidal propuesta en (10), la distribución trapezoidal estandarizada propuesta en (12) tiene como punto medio del intervalo modal, (m_1^*, m_2^*), el valor de m^*. Mientras que el punto medio del intervalo modal en la distribución trapezoidal CPR estandarizada es $(2m^* + 1)/4$.

4. VALORES DE LA MEDIA PARA LA DISTRIBUCIÓN TRAPEZOIDAL PRESENTADA Y COMPARACIÓN CON LAS MEDIAS DE LAS DISTRIBUCIONES CLÁSICAS DEL MÉTODO PERT

Teniendo en cuenta que las medias estandarizadas de las distribuciones más usadas en el método PERT son las siguientes:

- $\mu_U = 1/2$ de la distribución Uniforme
- $\mu_T = (1 + m^*)/3$ de la distribución Triangular
- $\mu_{BP} = (1 + 4m^*)/6$, de la Beta PERT
- La distribución Trapezoidal CPR tiene la siguiente distinción:

▪ $\mu_{TpCPR} = \dfrac{7 - 4m^{*2}}{6(3 - 2m^*)}$ si $0 < m^* \leq 1/2$

▪ $\mu_{TpCPR} = \dfrac{4m^{*2} + 4m^* + 3}{6(1 + 2m^*)}$ si $1/2 < m^* \leq 1$

En este apartado se comparan las medias estandarizadas de las distribuciones anteriormente mencionadas con la media estandarizada obtenida para la distribución trapezoidal introducida en el apartado anterior, a la que se denominará μ_{TpHH}.

Partiendo de (12), se tiene:

a) Cuando $0 < m^* \leq 1/4$

la distribución trapezoidal queda especificada como Tp(0, $m^*/2, 3m^*/2, 1$) luego a partir de (5) se tiene que $\mu_{TpHH} = \dfrac{4m^{*2} + 3m^* + 2}{6(m^* + 1)}$.

Es fácil comprobar que: $\mu_U > \mu_{TpCPR} > \mu_T > \mu_{TpHH} > \mu_{BP}$, además $\mu_T = \mu_{TpHH}$ en el caso de que $m^* = 0$.

b) Cuando $1/4 < m^* \leq 1/2$

se parte de la distribución trapezoidal Tp(0, $(6m^* - 1)/4, (2m^* + 1)/4, 1$) luego a partir de (5) se tiene que $\mu_{TpHH} = \dfrac{6m^* - 8m^{*2} + 5}{6(3 - 2m^*)}$.

Se demuestra que: $\mu_U > \mu_{TpCPR} > \mu_T > \mu_{TpHH} > \mu_{BP}$, además todas coinciden en el caso de que $m^* = 1/2$, como era de esperar.

c) Cuando $1/2 < m^* < 3/4$

partiendo de la distribución trapezoidal Tp(0, $(2m^*+1)/4$, $(6m^*-1)/4$,1) y teniendo en cuenta (5) se tiene que $\mu_{TpHH} = \dfrac{8m^{*2}+2m^*+3}{6(1+2m^*)}$.

Se comprueba que: $\mu_{BP} > \mu_{TpHH} > \mu_T > \mu_{TpCPR} > \mu_U$

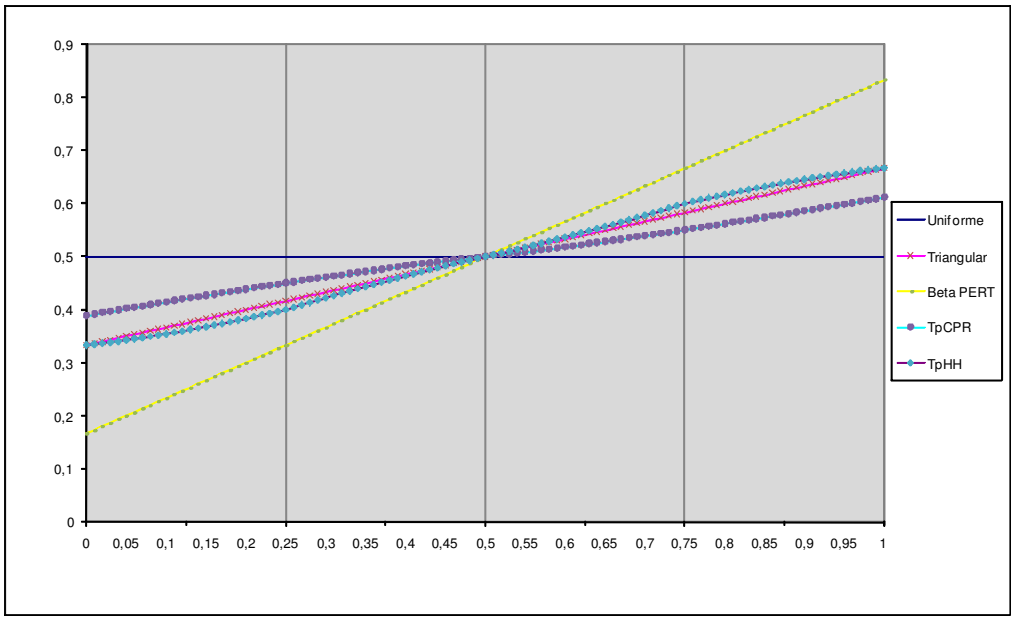
d) Cuando $3/4 \le m^* < 1$

la distribución trapezoidal es Tp(0, $(3m^*-1)/2$, $(m^*+1)/2$,1) luego a partir de (5) se tiene que $\mu_{TpHH} = \dfrac{5m^*-4m^{*2}+3}{6(2-m^*)}$.

Se comprueba que: $\mu_{BP} > \mu_{TpHH} > \mu_T > \mu_{TpCPR} > \mu_U$, además $\mu_T = \mu_{TpHH}$ en el caso de que $m^* = 1$.

Como puede comprobarse, véase gráfico 1, se ha obtenido que en el caso de que $0 < m^* < 1/2$ la media de la distribución Beta PERT es menor que cualquiera de las medias de las distribuciones utilizadas en este trabajo y en el caso de que $1/2 < m^* < 1$ se ha comprobado que la media de la distribución Beta PERT es mayor. En cualquier caso la media de la distribución introducida en este artículo es la más próxima a la media de la distribución Beta PERT de todas las estudiadas.

Gráfico 1: Representación de las medias de las distribuciones típicas del Método PERT



5. VALORES DE LA VARIANZA PARA LA DISTRIBUCIÓN TRAPEZOIDAL PRESENTADA Y COMPARACIÓN CON LAS VARIANZAS DE LAS DISTRIBUCIONES CLÁSICAS DEL MÉTODO PERT

Teniendo en cuenta que las varianzas estandarizadas de las distribuciones más usadas en el método PERT son las siguientes:

- $\sigma_U^2 = 1/12$ de la distribución Uniforme
- $\sigma_T^2 = \left(1 - m^*\left(1-m^*\right)\right)/18$ de la distribución Triangular
- $\sigma_{BP}^2 = 1/36$, de la Beta PERT

- La distribución Trapezoidal CPR tiene la siguiente distinción:

- $\sigma^2_{\text{TpCPR}} = \dfrac{37 - 90\,m^* + 112\,m^{*2} - 72\,m^{*3} + 16\,m^{*4}}{72\,(3 - 2\,m^*)^2}$ si $0 < m^* \leq 1/2$
- $\sigma^2_{\text{TpCPR}} = \dfrac{3 + 18\,m^* - 8\,m^{*2} + 8\,m^{*3} + 16\,m^{*4}}{72\,(1 + 2\,m^*)^2}$ si $1/2 < m^* \leq 1$

En este apartado se comparan las varianzas estandarizadas de las distribuciones anteriormente mencionadas con la varianza estandarizada obtenida para la distribución trapezoidal introducida en el apartado anterior, a la que se denominará σ^2_{TpHH}.

Partiendo de (12), se tiene:

a) Cuando $0 < m^* \leq 1/4$

la distribución trapezoidal queda especificada como Tp(0, $m^*/2, 3m^*/2, 1$) luego a partir de (7) se tiene que $\sigma^2_{\text{TpHH}} = \dfrac{4 + 6\,m^* - 5\,m^{*2} + 18\,m^{*3} + 7\,m^{*4}}{72\,(1 + m^*)^2}$

b) Cuando $1/4 < m^* \leq 1/2$

se parte de la distribución trapezoidal Tp(0, $(6m^* - 1)/4, (2m^* + 1)/4, 1$) luego a partir de (7) se tiene que $\sigma^2_{\text{TpHH}} = \dfrac{187 - 576\,m^* + 856\,m^{*2} - 576\,m^{*3} + 112\,m^{*4}}{288\,(3 - 2\,m^*)^2}$

c) Cuando $1/2 < m^* < 3/4$

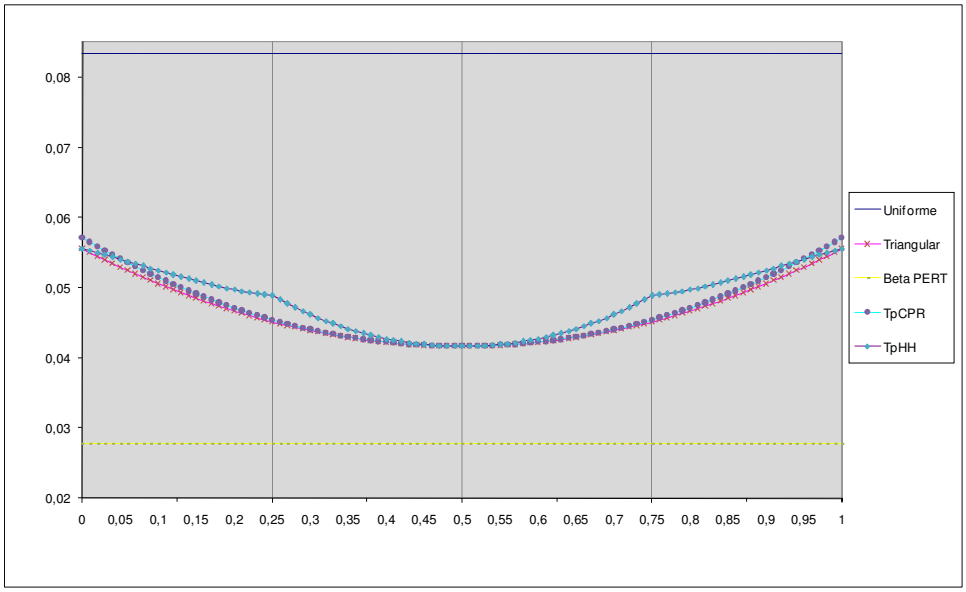
partiendo de la distribución trapezoidal Tp(0, $(2m^* + 1)/4, (6m^* - 1)/4, 1$) aplicada sobre (7) se tiene que $\sigma^2_{\text{TpHH}} = \dfrac{3 + 144\,m^* - 200\,m^{*2} + 128\,m^{*3} + 112\,m^{*4}}{288\,(1 + 2\,m^*)^2}$

d) Cuando $3/4 \leq m^* < 1$

la distribución trapezoidal es Tp(0, $(3m^* - 1)/2, (m^* + 1)/2, 1$) luego a partir de (7) se tiene que $\sigma^2_{\text{TpHH}} = \dfrac{30 - 78\,m^* + 91\,m^{*2} - 46\,m^{*3} + 7\,m^{*4}}{72\,(2 - m^*)^2}$

Como puede comprobarse, véase gráfico 2, se ha obtenido que en cualquier caso la varianza de la distribución Beta PERT es menor que cualquiera de las varianzas de las distribuciones utilizadas en este trabajo, y la varianza de la distribución Uniforme es la mayor de todas.

Gráfico 2: Representación de las varianzas de las distribuciones típicas del Método PERT



La varianza de la distribución trapezoidal introducida en este artículo es mayor que las varianzas tanto de la distribución Triangular como de la distribución Trapezoidal CPR, excepto en los intervalos $0 < m^* \leq 0{,}0532785$ y $0{,}946721 \leq m^* < 1$ donde queda por debajo de la varianza de la distribución Trapezoidal CPR. Además, cuando $m^* = 1/2$ coinciden las varianzas de estas tres distribuciones.

Desde un punto de vista conservador, el hecho de tener mayor varianza hace que esta nueva distribución sea preferida a las distribuciones habituales utilizadas en el Método PERT.

BIBLIOGRAFÍA

AKCAYOL, M.A., ELMAS, Ç., ERDEM, O.A. Y KURT, M. (2004) "An educational tool for fuzzy logic controller and classical controllers". Computer Applications in Engineering Education, Vol. 12, Issue 2, 126-135.

BALLESTERO, E. Y CABALLER, V. (1982) "Il método delle due Beta. Un procedimiento rapido nella stima dei beni fondari". Genio Rurale, vol. 45, nº 6, 33-36.

CALLEJÓN, J., PÉREZ, E. Y RAMOS, A. (1996) "La distribución trapezoidal como modelo probabilístico para la metodología PERT". Actas en CD-Rom de la X Reunión de ASEPELT-ESPAÑA celebrada en Albacete por la Universidad de Castilla la Mancha.

HERRERÍAS, R. Y CALVETE, H. (1987) "Una ley de probabilidad para el estudio de los flujos de caja de una inversión". Libro Homenaje al profesor Gonzalo Arnáiz Vellando. INE, Madrid, 279-296. Recogido en Herrerías, R. (2001, editor), 119-127.

HERRERÍAS, R., GARCÍA, J., CRUZ, S. Y HERRERÍAS, J.M. (2001). "Il modello probabilistico trapezoidale nel metodo delle due distribuzione della teoria generale de valutazioni". *Genio Rurale. Rivista di Scicienze Ambientali*, LXIV, 3-9.

HERRERÍAS, R., PALACIOS, F. Y HERRERÍAS, J.M. (2002). "Una determinación para la ponderación variable del valor más probable del Modelo Probabilístico Generalizado del método PERT". Actas III Reunión Científica ASEPELT - ESPAÑA. Análisis, Selección, Control de Proyectos y Valoración. Murcia

KAUFMANN, A. Y GIL ALUJA, J. (1986). "Introducción de la teoría de conjuntos borrosos a la gestión de empresas". Santiago de Compostela: Milladoiro.

LOZANO, J. J. (1996) "Tasación urbana: una metodología para informes de tasación masiva". Tesis Doctoral. Universidad Politécnica Madrid.

SUAREZ A., 2007. « Decisiones óptimas de inversión y financiación en la empresa". Pirámide, 21ª edición.

VAN DORP, J.R. Y KOTZ, S. (2002a). "The standard two sided power distribution and its properties: with applications in financial engineering". *The American Statistician*, 56, 90-99.

VAN DORP, J.R. Y KOTZ, S. (2002b). "A novel extension of the triangular distribution and its parameter estimation". *Journal of the Royal Statistical Society,* Ser. D, 51, 63-79.

VAN DORP, J.R. Y KOTZ, S. (2003). "Generalized trapezoidal distributions". *Metrika*, 58, 85-97.

VAN DORP, J.R., CRUZ, S., GARCÍA, J. Y HERRERÍAS, R. (2007). "An Elicitation Procedure for the Generalized Trapezoidal Distribution with a Uniform Central Stage", *Decision Analysis Journal*, Vol. 4, pp. 156 – 166.

Modelling Job Satisfaction Trajectories in the British Household Panel Survey using Latent Growth Curve Models

MARIA DE FÁTIMA SALGUEIRO
E-mail: fatima.salgueiro@iscte.pt
ISCTE - Instituto Superior de Ciências do Trabalho e da Empresa (Portugal)

Resumo

Nos últimos anos tem sido crescente o interesse sobre percepções de satisfação laboral, tendo sido utilizadas diferentes formas de a aferir. As abordagens estatísticas adoptadas consideram frequentemente a utilização de uma medida global de satisfação ou a construção de um *score*, e incluem modelos *ordered probit* e modelos de efeitos fixos.
O British Household Panel Survey (BHPS) é um inquérito anual, representativo dos agregados familiares na Grã-Bretanha, contendo diversas medidas de satisfação. O presente artigo propõe modelar trajectórias de satisfação laboral usando *latent growth curve models.* São utilizados dados do BHPS relativos aos anos 2000-2005, tendo os respondentes em análise estado empregados ao longo deste período.
São usados modelos com trajectórias não condicionadas e condicionadas, com o intuito de modelar a variação, intra e entre indivíduos, na satisfação laboral percepcionada. São possíveis determinantes de satisfação o estado de saúde, o número de crianças e o número de horas de trabalho.

Palavras Chave: BHPS, dados longitudinais, *latent growth curve models,* satisfação laboral.
Área Temática: Métodos Quantitativos para Economia e Empresa.

Abstract

In recent years there has been a growth in interest in perceptions of job satisfaction, and several measures have been used. Statistical approaches adopted in the literature to model job satisfaction often consider either using an overall measure or constructing a total score. Ordered *probit* models and fixed effects models have often been used.
The British Household Panel Survey (BHPS) is a national representative survey conducted on an annual basis, and contains several measures of job satisfaction. The current paper proposes modelling job satisfaction trajectories using latent growth curve models. Data from years 2000 to 2006 of the BHPS are used. All respondents have been employed throughout the period under analysis. Unconditional and conditional models are used to model both within-individual and between-individual level variation in perceived job satisfaction. Time-constant individual level predictors such as perceived health status, number of children and working hours are considered to explain growth.

Key words: BHPS, job satisfaction, latent growth curve models, longitudinal data.
Thematic Area: Quantitative methods for economics and business.

1. INTRODUCTION

In recent years there has been a growth in interest in subjective well-being by economists and social scientists. Particular attention has been devoted to perceptions of job satisfaction, and several measures have been proposed and used. Researchers have also been interested in identifying possible determinants of job satisfaction. The relationship between gender, qualification or social class and job satisfaction does not seem to be clear. Clark and Oswald (1996) show that job satisfaction decreases with an increasing number of working hours. Clark *et al.* (1996) propose a U-shaped relationship between age and job satisfaction. Vieira (2005) suggests that a good health status perception increases job satisfaction. García and García (2006) claim that marital status does not affect job satisfaction, whereas having children is considered an important determinant of job satisfaction.

Statistical approaches adopted in the literature to model job satisfaction often consider either using the overall measure or constructing a total score (by summing the scores of the dimensions of interest). Ordered *probit* models and fixed effects models have often been used (see, for example, Vieira *et al.,* 2005). Random effects models and structural equation models have also been proposed to model longitudinal survey data (see, for example, Berrington *et al.,* 2008).

The British Household Panel Survey (BHPS) is a national representative longitudinal survey conducted in Great Britain, since 1991, on an annual basis. Several measures of job satisfaction are available in the BHPS. Since 1998, in addition to a question on overall job satisfaction, respondents have been asked to rate their satisfaction levels with four job dimensions (total pay, security, work itself and hours worked), using a Likert-type response scale.

The current paper proposes modelling job satisfaction trajectories using latent growth curve models (see, for example, Bollen and Curran, 2006). Data from waves j (year 2000) to p (year 2006) of the BHPS are used. All respondents have been employed throughout the six-year period under analysis. Additionally, they were original sample members and they have fully answered all job satisfaction questions in all six waves considered – the current paper does not address the challenges of dealing with missing data.

Latent growth curve modelling is used to model both within-individual and between-individual level variation in the perceived job satisfaction. Time-constant individual level predictors are considered to explain growth. Indeed, possible determinants of job satisfaction considered in the paper are age, gender, marital status, perceived health status, qualifications, social class, number of children in the household, number of working hours and personal income. The statistical software Mplus 5 is used.

The structure of the paper is the following: Section 2 presents the main features of the sample under analysis. Section 3 gives an overview of latent growth curve models. Section 4 summarizes the main results of the statistical modelling undertaken. The paper ends with a discussion in Section 5.

2. THE SAMPLE UNDER ANALYSIS

The British Household Panel Survey (BHPS) is an on-going large nationally representative household panel survey of individuals in private domiciles in Great Britain. The main objective of the BHPS is providing information about social and economic change, both at the individual and at the household level. BHPS also collects data on behavioral variables, namely perceptions of life and of job satisfaction.

BHPS data have been collected in annual waves since 1991 (wave a or 1). Since 1998, in addition to a question on overall job satisfaction, respondents have been asked to rate their satisfaction levels with four job dimensions (total pay, security, work itself and hours worked) using a Likert-type scale from 1 (not satisfied) to 7 (completely satisfied).

The analyses conducted in this paper are based on a subsample of 1209 men and women, aged 16 years old or more, who were original sample members (*i.e.*, they have been interviewed in all waves a (1991) to o (2005), so that the longitudinal sampling design can be taken into account). Additionally, respondents gave full interview in waves j to o (years 2000 to 2005),

they have valid answers in all job satisfaction variables under analysis, and they have been employed throughout the six-year period considered: years 2000 to 2005.

Table 1 summarizes the distribution of the responses, in the year 2000, concerning the job satisfaction variables and Table 2 compares the mean perceived level of job satisfaction (overall and with each of the four dimensions) for the period 2000 to 2005. It is possible to conclude that the majority of respondents are satisfied or very satisfied with all dimensions. Mean satisfaction levels have remained stable throughout the six-year period under analysis. Job security and work itself are the dimensions with the highest mean satisfaction levels, whereas satisfaction with total pay has the lowest mean levels.

Table 1: Frequencies of response for the Job Satisfaction variables in the year 2000.

	JobSat 2000: total pay	JobSat 2000: security	JobSat 2000: work itself	JobSat 2000: hours worked	JobSat 2000: overall
	%	%	%	%	%
Not satis at all	2,3%	2,2%	1,0%	1,8%	1,1%
2	3,9%	3,6%	2,6%	2,9%	2,0%
3	15,1%	5,7%	6,5%	12,4%	6,8%
Not satis/dissat	5,9%	7,6%	8,0%	9,5%	7,0%
5	28,5%	22,9%	22,2%	22,5%	26,4%
6	37,5%	39,6%	46,8%	38,5%	48,1%
Completely satis	6,8%	18,4%	12,8%	12,4%	8,6%

Table 2: Means of the Job Satisfaction variables over the six-year period under analysis.

	2000	2001	2002	2003	2004	2005
JobSatisf: total pay	4.94	5.09	5.02	5.03	5.04	5.10
JobSatisf: security	5.38	5.39	5.42	5.45	5.50	5.50
JobSatisf: work itself	5.40	5.46	5.35	5.40	5.48	5.44
JobSatisf: hours worked	5.13	5.21	5.15	5.19	5.26	5.23
JobSatisf: overall	5.34	5.41	5.30	5.33	5.41	5.33

As far as gender is concerned, in year 2000 50.7% of the 1209 respondents are male and 49.3% are female. The distribution according to the age group is the following: 16 to 21 years old → 10.2%; 22 to 29 → 10.8%; 30 to 39 → 17.5%; 40 to 49 → 17.7%; the remaining 43.9% being aged 50 years old or more. The number of own children in the household ranges between none (in 73.1% of the cases) and 6 (0.1%); 95% of the respondents having up to 2 own children in the household. Regarding marital status, 66.1% of the respondents live with company and 33.6% live alone. In terms of perceived health status over the last 12 months, 24.8% consider it excellent; 54.3% very good; 16.7% fair and only 4.1% rate it as poor or very poor. The large majority of the respondents (79.9%) are employed full-time (30 hours or more). Figures 1 and 2 display the distribution of the highest educational qualification and social class of the respondents, respectively.

Figure 1: Distribution of the highest educational qualification, in year 2000.

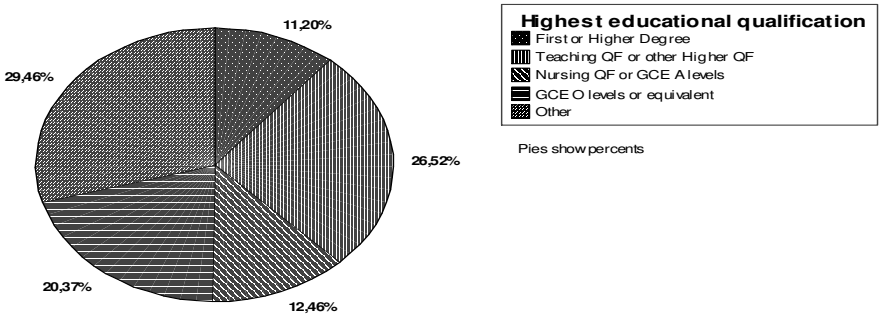
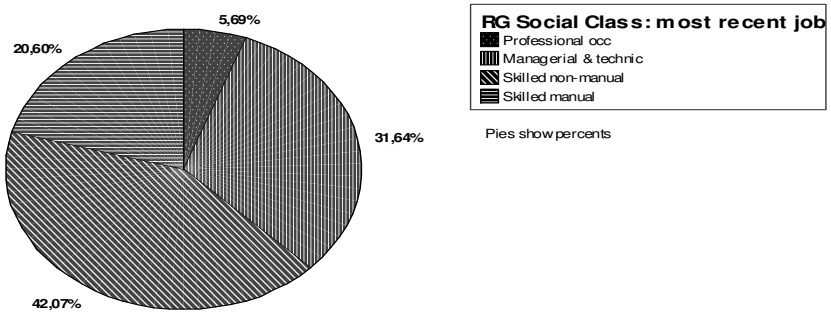


Figure 2: Distribution of the social class of the most recent job, in year 2000.



3. AN OVERVIEW OF LATENT GROWTH CURVE MODELS

Latent growth curve models (LGCM) are part of the general family of random effects models – for an introduction see Preacher *et al.* (2008). An important advantage of this methodology is that it is carried out using structural equation modeling (SEM) methodology, sharing many of its strengths and weaknesses (see Duncan *et al.*, 2006).

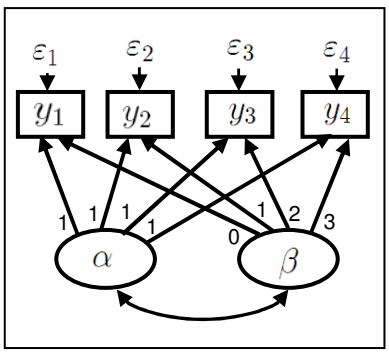
This allows the data analyst to take advantage of standard SEM features, such as global tests and assessment of model-data fit, multiple indicator latent variables and simultaneously modeling of multiple processes with regressions amongst the latent growth parameters. Yet, large sample sizes are still required.

LGCM permit the investigation of inter-individual differences in change over time, and allows for the investigation of the antecedents and consequences of change. The incorporation of both time-varying and time-invariant covariates is also possible. In LGCM time is incorporated in the model via constraints on the factor loadings in a latent variable model. Growth parameters are therefore specified as latent variables, and constraints are imposed on the factor loadings. The shape of the growth trajectory (linear, quadratic,…) depends on the number of latent variables specified in the model. The repeated observations are regarded as imperfect measures of an underlying latent trajectory. The shape of such trajectory can be estimated using the mean and the covariance structure of the observed repeated measures.

3.1. Models with Unconditional Linear Growth

Figure 3 displays the path diagram of an unconditional (with no predictors of growth in the model) LGCM. Four repeated measures (y_1, y_2, y_3, y_4) represent the observed indicators of the two latent variables: the intercept and the slope factors of the linear growth trajectory. To specify this linear growth the loadings of the intercept factor are constrained to 1, and the loadings of the slope factor equal the rank of the measurement occasion: 0, 1, 2, 3 in this example.

Figure 3: Unconditional LGCM, with linear growth



More generally, let

$$y_{it} = \alpha_i + \lambda_t \beta_i + \varepsilon_{it}$$

be the value for variable *y* for individual *i* at time *t*. A usual coding convention for linear growth is that

$$\lambda_t = t - 1$$

The parameters of special interest in this model are the means, variances and covariance of the random effects. The random intercept and the random slope are:

$$\alpha_i = \mu_\alpha + \zeta_{\alpha_i}$$

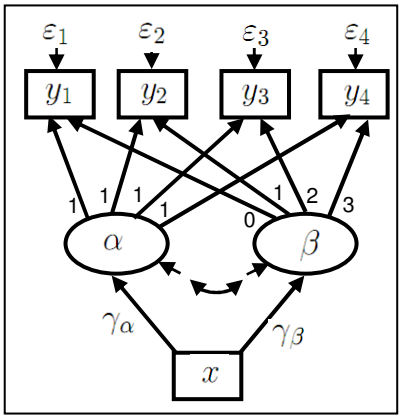
$$\beta_i = \mu_\beta + \zeta_{\beta_i}$$

The intercepts of the observed indicators are constrained to zero, and the means and variances of the intercept and growth factors are freely estimated. Thus, the mean of the intercept growth factor can be interpreted as the mean value of *y* at the first time point. The mean of the latent slope represents the mean rate of change in *y* for each unit increase in time, over the period considered. The variances of the latent intercept and slope represent individual heterogeneity around the group means: known as the inter-individual difference in the intra-individual change. The covariance between the latent intercept and the latent slope represents the relationship between initial status and rate of change, and is often negative.

3.2. Models with Conditional Linear Growth

Unconditional LGCM describe growth, whereas conditional LGCM try to explain growth, *i.e.,* to examine predictors of individual change over time. Figure 4 displays the path diagram of a conditional LGCM, where *x* is either a continuous or a dummy variable, which is a time-invariant explanatory variable for the random intercept and slope.

Figure 4: Conditional LGCM, with linear growth



More generally, the conditional LGCM is given by

$$y_{it} = \alpha_i + \lambda_t \beta_i + \varepsilon_{it}$$

with random intercept and random slope given by:

$$\alpha_i = \mu_\alpha + \gamma_\alpha x_i + \zeta_{\alpha_i}$$

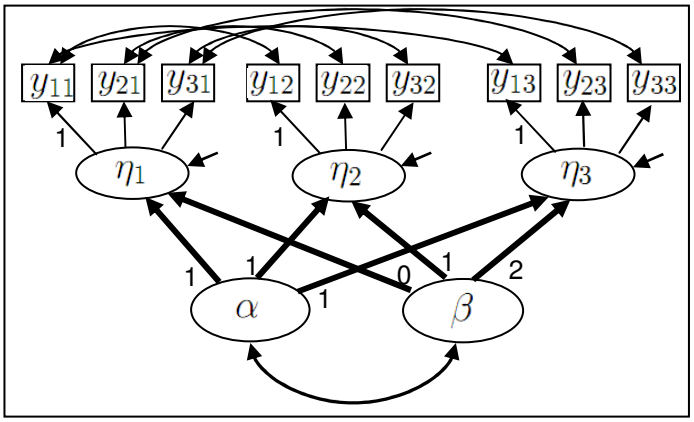
$$\beta_i = \mu_\beta + \gamma_\beta x_i + \zeta_{\beta_i}$$

The parameters of special interest now are, besides the means and variances of the random effects, the regression paths running from the predictor *x* to the latent growth parameters. The regression path on the random intercept denotes the expected change on the mean of the latent intercept for each unit increase in the fixed covariate *x*. The regression path on the random slope can be interpreted as the expected change in the mean growth rate. Standard tests can be used to determine the significance of the regression paths.

3.3. Second-order LGCM with Unconditional Linear Growth

Figure 5 displays the path diagram of an unconditional LGCM with linear growth for a latent variable η measured by three indicators at each time point. Such models are known in the literature as second-order LGCM or LGCM with multiple indicators. The focus is on examining the change in latent (unobserved) variables rather than modeling change in observed variables. Indeed, placing the repeated measures in a SEM framework allows us to take advantages of the major strengths of SEMs.

Figure 5: Second-order unconditional LGCM, with linear growth



The value for indicator *j* for individual *i* at time *t* is given by

$$y_{jit} = \nu_{jt} + \Lambda_{jt}\eta_{it} + \upsilon_{jit}$$

Where Λ_{jt} denotes the factor loading for indicator *j* at time *t*. The growth model for the latent variable η specifies that:

$$\eta_{it} = \alpha_i + \lambda_t\beta_i + \varepsilon_{it}$$

with random effects given by:

$$\alpha_i = \mu_\alpha + \zeta_{\alpha_i}$$

$$\beta_i = \mu_\beta + \zeta_{\beta_i}$$

As in Section 3.1, the parameters of special interest are mainly the means, variances and covariance of the random effects. Conditional second-order LGCM can also be specified, following the reasoning presented in Section 3.2. For a more detailed explanation see Boolen and Curran (2006).

4. RESULTS

The trajectory of each of the job satisfaction dimensions was first modeled individually, as an unconditional LGCM. Since within-individual and between-individual level variation in the perceptions of job satisfaction revealed significant, there was a need to fit conditional LGCM to explain growth – results are presented in Section 4.1. In a second step confirmatory factor

analysis (CFA) was used to assess the adequacy of a latent job satisfaction factor to account for the four dimensions measured at each wave – results are presented in Section 4.2. In a third step a second-order LGCM was considered to account for the growth trajectory of the latent job satisfaction factor. Unconditional and conditional models were fitted and results are presented in Section 4.3.

All statistical modeling was conducted using Mplus 5. Observed variables were declared as measured in an ordinal scale and the delta parameterization available in Mplus was used: equal thresholds of the outcome variable were imposed and the mean of the random intercept was fixed at zero. The robust weighted least squares estimator available in Mplus (WLSMV) was used.

4.1. Unconditional and Conditional LGCM for the Observed Job Satisfaction Dimensions

The linear trajectory (2000 to 2005) of the overall job satisfaction was first modeled as an unconditional LGCM (such models were presented in Section 3.1). The estimated mean of the slope was positive but non-significant. The variance of the intercept and of the slope were both significant, thus suggesting there are significant differences between individuals in the mean initial values of perceived overall job satisfaction, as well as in the growth rate of overall job satisfaction. The estimated covariance between the random intercept and the random slope was significantly negative, suggesting individuals with a higher mean of overall job satisfaction in 2000 increase their mean overall job satisfaction level at a lower rate during the six-year period under analysis.

Similar reasoning was followed to model the four job satisfaction dimensions. Table 3 summarizes the estimates of the means, variances and covariances between the latent intercept and the latent slope factors in the five unconditional LGCM that were fitted.

Table 3: Estimates of the means, variances and covariances of the intercept and slope factors for the five unconditional LGCM fitted. (*) denotes a non-significant estimate, at a 5% level.

	Overall JobSatisf	**JobSatisf: Total Pay**	**JobSatisf: Security**	**JobSatisf: Work itself**	**JobSatisf: Hours Worked**
Mean of the slope	0.002 (*)	0.017	0.021	0.008 (*)	0.013
Variance of the intercept	0.599	0.615	0.568	0.62	0.546
Variance of the slope	0.022	0.023	0.018	0.02	0.014
Covariance between intercept and slope	- 0.053	- 0.046	- 0.04	- 0.051	- 0.038

From Table 3 it is possible to conclude that there is significant variability within and between individuals that needs to be accounted for. Therefore, conditional LGCM have to be fitted.

The possible determinants of the job satisfaction growth trajectories considered in this paper are gender, age group, marital status, perceived health status, qualifications, social class, number of own children in the household, number of working hours and personal income. The logarithm of the income and the number of children in the household were treated as continuous variables. For the remaining 7 determinants dummy variables were computed and included in the conditional LGCM (recall such models were presented in Section 3.2).

Table 4 presents the significant estimates (at a 5% level) of the effects of the determinants of job satisfaction on random intercepts and random slopes, for the five conditional LGCM that were fitted. Marital status and social class had no significant effects, and for that reason these two variables have not been included in the table.

Table 4: Significant effects (at 5%) of the possible determinants of job satisfaction on the random intercepts and slopes, for the five conditional LGCM fitted.

	Overall JobSatisf		**JobSatisf: Total Pay**		**JobSatisf: Security**		**JobSatisf: Work itself**		**JobSatisf: Hours worked**	
	Intcpt	Slope	Intcpt	Slope	Intcpt	Slope	Intcpt	Slope	Intcpt	Slope
Gender:										
Female	0.234				0.174			0.026	0.143	
Working Hours:										
Part-time	0.238		0.263	-0.05	0.18		0.139		0.579	-0.06
Age group:										
22-29				0.067						
30-39				0.087						
40-49			-0.3	0.079						
>= 50										
Qualifications:										
Higher QF										
A levels		0.06				0.058				0.044
O levels		0.06				0.051		0.057		0.045
other						0.057		0.04		
Health Status:										
good	-0.19		-0.14		-0.33		-0.24		-0.15	
fair	-0.29		-0.28		-0.41		-0.37		-0.3	
poor	-0.33		-0.33		-0.36				-0.3	
Number of children		-0.02								-0.01
Logarithm of personal income						0.016				

The number of working hours for those working part-time has significant positive effects on the random intercepts of the five job satisfaction dimensions, thus suggesting individuals working part-time are, on average, more satisfied in year 2000 than those working full-time (the reference category). A negative slope associated with total pay suggests that those working part-time have a decrease on the average growth rate of satisfaction with total pay in the six-year period under analysis, when compared to those working full-time.

The health status also influences the initial average job satisfaction level in year 2000: the poorer the health status, the lower the average perception of job satisfaction. In year 2000 women are, on average, more satisfied than men, regarding overall job satisfaction, security and hours worked.

4.2. Confirmatory Factor Analysis of the Latent Job Satisfaction Factor

The four job satisfaction variables were then considered indicators of a single-factor model and a confirmatory factor analysis (CFA) model was fitted at each of the six waves under analysis. Manifest variables were treated as ordinal and the delta parameterization available in Mplus was used, with the WLSMV estimator. Table 5 presents the estimates of the factor loadings, in a standardized solution, and some of the goodness-of-fit measures that were obtained.

Table 5: CFA models for years 2000 to 2005 – standardized factor loadings and goodness-of-fit measures.

	2000	2001	2002	2003	2004	2005
	Standardized Factor loadings					
JobSatisf: Total pay	0.61	0.65	0.58	0.60	0.62	0.63
JobSatisf: Security	0.55	0.54	0.57	0.60	0.55	0.57
JobSatisf: Work itself	0.72	0.69	0.75	0.71	0.70	0.73
JobSatisf: Hours worked	0.72	0.69	0.72	0.69	0.71	0.69
	Goodness-of-fit measures					
CFI	0.985	0.99	0.99	0.99	0.984	0.986
TLI	0.97	0.981	0.976	0.98	0.97	0.973
RMSEA	0.094	0.072	0.087	0.072	0.092	0.092

The literature often recommends standardized factor loadings larger than 0.6. In our model the variable satisfaction with job security has, in some waves, factor loadings below the minimum recommended value, thus suggesting one should be very cautious when using the proposed job satisfaction single-factor. The goodness-of-fit measures suggest an acceptable model-data fit.

4.3. Unconditional and Conditional Second-order LGCM for the Latent Job Satisfaction Factor

The linear trajectory (2000 to 2005) of the latent job satisfaction factor proposed in Section 4.2 was modeled as a second-order LGCM with unconditional linear growth (as explained in Section 3.3, but with four manifest indicators at each wave).

Residuals of the repeated measures were allowed to correlate over time. The delta parameterization for categorical outcome variables available in Mplus was used, with a robust WLS estimator. Consequently, the mean of the random intercept term was fixed at zero, and so were the means of the six first order latent job satisfaction factors. Measurement invariance over time was imposed using equality constraints on the factor loadings. For the repeated categorical measures equal thresholds were imposed for the different time points.

The estimated mean of the slope was negative, but non-significant, thus suggesting a non-significant growth trajectory of the latent job satisfaction. The variances of the random intercept and of the random slope were both significant and, as expected, the two random effects were negatively correlated. The need to account for the within and between individuals variability suggested fitting a conditional second-order LGCM. The 9 possible determinants of job satisfaction previously used were considered once more.

Table 6 presents the estimates of the significant effects that were obtained (at a 5% level). The remaining five variables (age group, social class, marital status, number of children in the household and personal income) are not included in the table because they had no significant impact on either the intercept or the slope.

Table 6: Significant effects (at 5%) of the possible determinants of job satisfaction on the random intercept and slope of the fitted second-order conditional LGCM.

	Latent Job Satisfaction Factor	
	Intercept	Slope
Gender:		
Female	0.121	
Working Hours:		
Part-time	0.3	-0.032
Qualifications:		
Higher QF		
A levels		0.04
O levels		0.043
other		
Health Status:		
good	-0.24	
fair	-0.338	
poor	-0.252	

Results are somehow similar to those previously obtained in Section 4.1, when a conditional LGCM was fitted to explain the growth trajectories of the overall job satisfaction variable. In brief, in year 2000:

- Women are, on average, more satisfied than men;
- Those working part-time are, on average, more satisfied than those working full-time;
- Those with an excellent health status (the reference category) are, on average, more satisfied than those with lower perceived health conditions.

However, some differences are also noticeable:

- When modeling the overall job satisfaction variable, the number of children in the household had a negative impact on the average growth rate of satisfaction. Such impact is not significant when modeling the trajectory of the latent job satisfaction;
- Those working part-time have a significantly lower average growth rate of latent job satisfaction than those working full-time. Such effect was non-significant when the overall job satisfaction variable was modeled. However, such an effect did exist when the growth trajectories of the variables satisfaction with total pay and with hours worked were modeled (see Table 4);
- There is no significant age group effect either on the average satisfaction level in 2000 or on the satisfaction growth rate in the six-year period under analysis when job satisfaction is measured by an overall variable or by a latent factor. However, when job satisfaction with total pay is measured, in year 2000 those aged 40-49 years old are, on average, less satisfied than those aged 16-21 (the reference category). The positive slopes associated with different age groups indicate different average growth trajectories of satisfaction with total pay.

5. DISCUSSION

This paper has illustrated how to model perceptions of job satisfaction using unconditional and conditional LGCM. The growth trajectories of measured items have been modeled, and so have the trajectories of a latent variable proposed as a measure of job satisfaction. Such an approach allows accounting for measurement error in model specification. The potentialities of modeling longitudinal data within the SEM framework have been highlighted.

Future research includes considering the complex survey design and dealing with missing data. The possibility of including time-varying covariates in the model can also be considered.

REFERENCES

BERRINGTON, A., HU, Y., SMITH, P.W.F AND STURGIS, P. (2008), "A graphical chain model for reciprocal relationships between women's gender role attitudes and labour force participation". *J. R. Statist. Society*, A, 171(1), pp. 89-108.

BOLLEN, K.A. AND CURRAN, P.J. (2006), "Latent Curve Models – A Structural Equation Perspective", John Wiley & Sons, New Jersey.

CLARK, A. AND OSWALD, A. (1996), "Satisfaction and Comparison Income", *Journal of Public Economics*, 61, pp. 359-381.

CLARK, A., OSWALD, A. AND WARR, P. (1996), "Is Job Satisfaction U-shaped in Age?", *Journal of Occupational and Organizational Psycology*, 69, pp. 57-81.

DUNCAN, T.E., DUNCAN, S.C., STRYCKER, L.A. (2006), "An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues an Applications", 2nd ed., Lawrence Erlbaum Associates, Inc., New Jersey.

GARCÍA. E. AND GARCÍA, J. R. (2006), "Determinants Socioeconòmics de la satisfacció Laboral a Catalunya", *Nota d'Economia*, nº86, pp. 99-119.

PREACHER, K.J., WICHMAN, A.L., MACCALLUM, R.C AND BRIGGS, N.E. (2008), "Latent Growth Curve Modelling", Sage Publications, Inc.

VIEIRA, J. C. (2005), "Skill Mismatches and Job Satisfaction", *Economics Letters*, 89, pp. 39-47.

VIEIRA, J. C., MENEZES, A., GABRIEL, P. (2005), "Low Pay, Higher Pay and Job Quality: Empirical Evidence for Portugal", *Applied Economics Letters*, Vol. 12, pp. 505-511.

A COMPARISON OF AMBULATORY HEALTH CARE SYSTEMS

JEAN CIVET
e-mail: jean.civet@univ-fcomte.fr
Université de Franche-Comté

Abstract

We compare two ambulatory health care organisations from a theoretical point of view. The first one is a non-gatekeeping: patients have a free access to specialty care and both physicians are paid under fee-for-service (FFS). The second organisation is a gatekeeping system: a GP referral is compulsory to get to specialty care. We suppose that physicians remain reimbursed under FFS. We assume the insurer acts à la Laffont and Tirole (1993). The payment schemes of both physicians are computed to maximise the expected utility perceived of the treatment minus its social costs. We assume that the insurer observes neither the diagnosis effort nor the treatment carried out by GPs. We compute the payments schemes of both physicians which maximise the expected utility and satisfy the incentive and participative constraints. We conclude that in a gatekeeping system treatment costs are not necessarily lower.

Key-words: non gatekeeping, gatekeeping, fee-for-service
Thematic Area: Energy, Environment, Transports and Natural Resources Economics.

1. INTRODUCTION

Since the end of the 70s, developed countries have been facing increasing ambulatory and secondary health care expenditures. Reforms aiming at curbing ambulatory expenditures have been implemented only since the beginning of the 1990s.

These reforms depended on the prior health care organisation:

- in the UK GPs gatekeepers became fundholders,
- in Germany and in the Netherlands competition between insurers have been increased,
- in France since the reform which capped GPs' FFS fail to curb expenditures, GPs have become gatekeepers.

In this article we compare two health care organisations:

- the first one was found in France prior the health care reform of 2004 and in traditional insurance: GPs and specialists are paid under FFS and patients have a free access to specialty care,
- the second one is the current French ambulatory health care system and the Preferred Providers Organisations: GPs and specialists remain paid under FFS, but GPs are gatekeepers.

This paper is organised as follow. In the first section we give the hypotheses of the different health care systems. In the second part we compute the expected utilities and the optimal health care payment of both physicians. We suppose that insurers impose neither the diagnosis effort nor the treatment carried out by the GPs. In the last section we compare these two organisations. All our results are summed up in the conclusion.

2. The model

In this section we present the assumptions about patients', physicians' and insurers' behaviours. It enables us to take into account the main features of both health care systems.

2.1. Patients

We assume that the patient suffers from a mild or a severe illness which occur with the same

probability[1].We define two events:

- the patient suffers from a minor illness by m ,

- the patient suffers from a severe illness by s ,

and we obtain the following probabilities:

$$\Pr(s) = \Pr(m) = \frac{1}{2}$$

We suppose that:

- the minor illness is cured with m health care services which can be carried out by GPs and specialists,

- the severe illness is cured with s health care services only supplied by specialists, with $s > m$.

We define two events:

- the patient consults the GP by CG,

- the patient consults the specialist by CS .

In a health care organisation with free access to specialty care we suppose that the patient does not observes his illness. He chooses his physician randomly. Thus the consulting probabilities are:

$$\Pr(CG) = p, \Pr(CS) = 1 - p$$

with $p > \frac{1}{2}$, it is more likely that a suffering patient consults a GP than a specialist.

In the gatekeeping system, since the patient must consult his GP. The probabilities become:

$$\Pr(CG) = 1, \Pr(CS) = 0$$

2.2. Physicians

As both GPs and specialists produce ambulatory health care services we must take into account both types of physician, i.e.:

- their payment schemes,
- their knowledge and ability to treat the different kinds of illness.

2.2.1. Specialists

Specialists have a deeper knowledge of both illnesses:

- they cure both kinds of illness and always treat his patient
- they observe perfectly the illness of his patient and do not perform any diagnosis effort.

As specialists use high-tech equipment to treat illnesses we suppose they bear the treatment costs:

- if they treat a minor illness they support a cost c_1

- if they treat a patient suffering from a severe illness they support a cost c_2 , if the patient did not received a treatment before[2] or c_3 if the patient was treated unsuccessfully by his GP.

We suppose that $c_3 > c_2 > c_1$ [3].

If p_s is the fee for one service, their revenues are:

- $p_s(m+1) - c_1$ if they treat a minor illness,

[1]Unlike Bardey (2002), severity must be understood as different levels of a given illness (bacterial or viral sore throat for instance). This probability is common knowledge.

[2]It is the case when patient consults a specialist or when the GP refers him.

[3]When a patient suffering from a severe illness is referred too late to a specialist he must perform a deeper and costlier treatment, i.e. $c_3 > c_2$. Moreover it is more costly to treat a severe illness, i.e. $c_2 > c_1$.

- $p_s(s+1)-c_2$, if they treat a patient who is actually suffering from a severe disease and was referred by his GP or has a direct access to specialty care,

- $p_s(s+1)-c_3$, if they treat a patient who is actually suffering from a severe disease and was treated unsuccessfully before by his GP.

In both organizations specialists have a reservation wage ω_s with $\omega_s > \omega_g$.

2.2.2. GPs

Unlike specialists, GPs do not know exactly the illness of his patient. Moreover they are trained to treat the mild disease. It means that:

- GPs can not heal a patient suffering from the severe illness,
- they treat the mild illness at no cost,
- when a patient consults GPs they do not know if he suffers from a mild or a severe disease: to improve their knowledge they perform a diagnosis effort.

As in Jelovac (2001) we suppose that during the first consultation GPs carry out a diagnosis effort $e \in [0,1]$:

- this effort gives him a signal s^g on the illness of his patient,
- the signal is either $s^g = s$ (i.e. the patient suffers from a severe illness) or $s^g = m$ (i.e. the patient suffers from a minor illness).

We suppose that the higher the effort is, the sharper the signal is:

$$\Pr(s^g = s / s) = \Pr(s^g = m / m) = \frac{1+e}{2}$$

- GPs bear a disutility, $\Psi(e) = k\frac{e^2}{2}$ with $k > 0$.

Once the signal observed the GP chooses either to treat his patient or to refer him to a specialist. We suppose that patients suffering from a severe illness leave GPs who decided to treat them. It means that GPs suffer a loss of reputation r , with $r > p_g m$.

Finally we assume that GPs have a reservation wage ω_g , with $\omega_g < \omega_s$.

If p_g is the fee for one service, the revenues of GPs are:

- $p_g(m+1)$, if they treat a patient suffering from a minor illness,
- p_g , if they referred him to specialty care.

2.2. The Insurer

We suppose an insurer à la Laffont and Tirole (1993). His objective is to maximise social welfare and give the same weight to patient's utility and physicians' profits, taking into account the social cost of public funding τ .

As in procurement model the insurer takes into account the utility of the patient following the treatment. This utility depends on the timing of the game:

- if the patient is cured after one consultation the insurer assumes his utility is U_1 ,
- if the patient is cured after two consultations, i.e. the GP does not treat him and refers him to a specialist, the utility is U_2 ,
- if the patient is cured after two consultations, but the GP treats him before, the insurer assumes his utility is U_3 ,

with $U_1 > U_2 > U_3$.

The payments of physicians depend on the treatment carried out. The health insurer gives:

- $p_g(m+1)$ or p_g to GPs,
- $p_s(s+1)$, $p_s(s+1)$ or $p_s(m+1)$ to specialists.

In this paper, we focus on the imperfect information case, i.e. the insurer can impose neither the diagnosis effort nor the treatment carried out by GPs.

Our assumptions differ from the ones assumed in Marinoso and Jelovac (2003) and in Levaggi and Rochaix (2007):

- in the first paper, in the gatekeeping and non gatekeeping systems which are compared the payment scheme of GPs is based on three components. By contrast we assume that GPs are paid under FFS because we want to compare the actual health care systems
- in the second paper, the actual payment scheme is taken into account. But these payments are exogenous whereas in our model they are endogenous.

3. THE EXPECTED UTILITIES IN BOTH HEALTH CARE SYSTEMS

Before performing his diagnosis, the GP knows that both illnesses occur with the same probability and that patient chooses his physician randomly. Updating their belief with the Bayes rules, a GP diagnoses accurately a minor or a severe illness with the following probability:

$$p(s / s^g = s) = p(m / s^g = m) = \frac{1+e}{2}$$

and

$$p(s / s^g = m) = p(m / s^g = s) = \frac{1-e}{2}$$

We can compute for both health care organisations the expected utility of GPs U_g, the expected utility of specialists U_s and the expected utility of insurer U_r.

We limit ourselves to three treatment strategies. GPs can:

- always treat their patient,
- always refer them,
- treat them if they diagnose a mild disease[4].

3.1. In a Non Gatekeeping System

If GPs have a financial incentive to always treat their patient, their expected utility is:

$$U_g^{1'} = p\left(p_g(m+1) - k\frac{e^2}{2} - \frac{r}{2} \right)$$

The expected utilities of specialists and of the health insurer are:

$$U_s^{1'} = p_s\left[\frac{(s+1)}{2} + \frac{(1-p)(m+1)}{2} \right] - \left(\frac{(1-p)}{2}(c_1 + c_2) + \frac{p}{2}c_3 \right)$$

and

$$U_r^{1'} = U_1(1-\frac{p}{2}) + U_3\frac{p}{2} - \left(\left(\frac{1-p}{2} \right)(c_1 + c_2) + \frac{p}{2}c_3 \right)$$

$$-\lambda p_s\left[\frac{(s+1)}{2} + \frac{(1-p)(m+1)}{2} \right] - p\left(\lambda p_g(m+1) + k\frac{e^2}{2} + \frac{r}{2} \right)$$

If GPs have a financial incentive to always refer their patient, their expected utility is:

$$U_g^{2'} = p\left(p_g - k\frac{e^2}{2} \right)$$

The expected utilities of specialists and of the insurer are:

[4]This last treatment strategy is called the appropriate strategy.

$$U_s^{2'} = p_s\left[\frac{(s+1)}{2}+\frac{(m+1)}{2}\right]-\frac{1}{2}(c_1+c_2)$$

and

$$U_r^{2'} = U_1(1-p)+pU_2-\frac{1}{2}(c_1+c_2)$$

$$-\lambda p_s\left[\frac{(s+1)}{2}+\frac{(m+1)}{2}\right]-p\left(\lambda p_g-k\frac{e^2}{2}\right)$$

If GPs have a financial incentive to treat their patient only when they diagnose a mild illness, their expected utility is:

$$U_g^{3'} = \left(\frac{p}{2}\right)\left(p_g(m+2)\right)-p\frac{ke^2}{2}-r\frac{p(1-e)}{4}$$

The expected utility of secondary health care providers is:

$$U_s^{3'} = p_s\left[\frac{(s+1)}{2}+\frac{(m+1)}{2}\left(1-p+\frac{p(1-e)}{2}\right)\right]$$

$$-\left(\frac{(1-p)}{2}(c_1+c_2)+\frac{p}{4}\left((1-e)(c_1+c_3)+(1+e)c_2\right)\right)$$

The expected utility of the health care insurer is:

$$U_r^{3'} = U_1\left(1-p+\frac{p(1+e)}{4}\right)+U_2\left(\frac{p}{2}\right)+\frac{p(1-e)}{4}U_3$$

$$-p\frac{ke^2}{2}-r\frac{p(1-e)}{4}$$

$$-\lambda\left(\frac{p}{2}\right)p_g(m+2)$$

$$-\lambda p_s\left[\frac{(s+1)}{2}+\frac{(m+1)}{2}\left(1-p+\frac{p(1-e)}{2}\right)\right]$$

$$-\left(\frac{(1-p)}{2}(c_1+c_2)+\frac{p}{4}\left((1-e)(c_1+c_3)+(1+e)c_2\right)\right)$$

3.2. In a Gatekeeping System

Since both physicians are paid under FFS in both organisations only the probabilities of consulting GPs change. By replacing p with 1, we get the expected utilities in this gatekeeping system.

If GPs have a financial incentive to always treat their patient, the expected utilities become:

$$U_g^{1''} = p_g(m+1)-k\frac{e^2}{2}-\frac{r}{2}$$

$$U_s^{1''} = \frac{1}{2}\left(p_s(s+1)-c_3\right)$$

$$U_r^{1''} = \frac{1}{2}(U_1+U_3)-\frac{(1+\lambda)}{2}c_3$$

$$-\lambda p_s\frac{(s+1)}{2}-\left(\lambda p_g(m+1)+k\frac{e^2}{2}+\frac{r}{2}\right)$$

If GPs have a financial incentive to always refer their patient, the expected utilities are:

$$U_g^{2''} = p_g - k\frac{e^2}{2}$$

$$U_s^{2''} = p_s\left[\frac{(s+1)}{2} + \frac{(m+1)}{2}\right] - \frac{1}{2}(c_1 + c_2)$$

$$U_r^{2''} = U_2 - \frac{(1+\lambda)}{2}(c_1 + c_2) - \lambda p_s\left[\frac{(s+1)}{2} + \frac{(m+1)}{2}\right] - \lambda p_g - k\frac{e^2}{2}$$

If GPs have a financial incentive to treat their patient only when they diagnose a mild illness, the expected utilities are:

$$U_g^{3''} = \left(\frac{1}{2}\right)\left(p_g(m+2)\right) - k\frac{e^2}{2} - r\frac{(1-e)}{4}$$

$$U_s^{3''} = p_s\left[\frac{(s+1)}{2} + \frac{(m+1)}{2}\left(\frac{(1-e)}{2}\right)\right] - \left(\frac{1}{4}\left((1-e)(c_1+c_3) + (1+e)c_2\right)\right)$$

$$U_r^{3''} = U_1\left(\frac{(1+e)}{4}\right) + U_2\left(\frac{1}{2}\right) + \frac{(1-e)}{4}U_3 - k\frac{e^2}{2} - r\frac{(1-e)}{4} - \lambda\left(\frac{1}{2}\right)p_g(m+2) - \lambda p_s\left[\frac{(s+1)}{2} + \frac{(m+1)}{2}\left(\frac{(1-e)}{2}\right)\right] - (1+\lambda)\left(\left(\frac{(1-e)}{4}\right)c_1 + \left(\frac{(1+e)}{4}\right)c_2 + \frac{(1-e)}{4}c_3\right)$$

4. THE SECOND BEST CONTRACTS

In this section we compute second best contracts. We first determine the level of effort and the treatment carried out by GPs. Then we compute incentives constraints and finally second best payment schemes.

4.1 . The Implemented Treatment

In this section we compute in both health care systems the condition under which GPs:
- always treat their patient,
- always refer them,
- treat their patient if they diagnose a mild disease.

When GPs diagnose a severe illness, i.e. $s^g = s$, they can decide to treat their patient. If he suffers from a mild disease, which occurs with the probability $p(m/s^g = s)$, GPs get $p_g(m+1)$. If he suffers from the severe mild illness, which occurs with the probability $p(s/s^g = s)$, the treatment of GPs is useless. Since the patient shifts GPs, they bear a loss of reputation, and their utility is $p_g(m+1) - r$. GPs can also decide to refer their patient. In such

a case their payment is p_g , whatever the severity of the illness.

When GPs diagnose a mild illness, i.e. $s^g = m$, they first can decide to treat their patient. If he suffers from a mild disease, which occurs with the probability $p(m/s^g = m)$, the payment of GPs is $p_g(m+1)$. If he suffers from the severe mild illness, which occurs with the probability $p(s/s^g = m)$, the treatment of GPs is useless. Since the patient shifts GPs, they bear a loss of reputation, and their utility is $p_g(m+1) - r$. GPs can also decide to refer their patient. In such a case their payment is p_g , whatever the severity of the illness.

Since all updated probabilities do not depend on p , we have the same lemmas for both health care organisations.

Lemma A primary health care provider who diagnoses a mild illness treats his patient instead of referring him, if:

$$p(s/s^g = m)(p_g(m+1) - r) + p(m/s^g = m)p_g(m+1) \geq p_g$$

$$\Leftrightarrow$$

$$e \geq \frac{r - 2p_g m}{r} \equiv \hat{e}'$$

Lemma A primary health care provider who diagnoses a severe illness refers his patient instead of treating him, if:

$$p_g \geq p(s/s^g = s)(p_g(m+1) - r) + p(m/s^g = s)p_g(m+1)$$

$$\Leftrightarrow$$

$$e \geq \frac{2p_g m - r}{r} \equiv -\hat{e}'$$

Since e is a probability, we must take into account that $e \in [0,1]$. The following lemma sums up the different treatment strategies under this constraint.

Lemma The treatments carried out by GPs are the following ones:

i) if $r - 2p_g m > 0$, then there are two cases. If the level of effort is low, i.e. $e \in [0, \hat{e}')$, then GPs always refer their patient. If the level of effort is high, i.e. $e \in (\hat{e}', 1]$, then GPs treat their patient only when a mild illness is diagnosed,

ii) if $-r + 2p_g m > 0$, then there are two cases. If the level of effort is low, i.e. $e \in [0, -\hat{e}')$, then GPs always treat their patient. If the level of effort is high, i.e. $e \in (-\hat{e}', 1]$, then GPs treat their patient only when a mild illness is diagnosed.

In the following section we compute the level of diagnosis effort for each treatment strategy and for both organisations.

4.2. The Diagnosis Effort

Following the previous lemmas we know that the diagnosis efforts depend only on the treatment.

Lemma If a primary health care provider always treats or always refers his patient, the optimal level of effort is: $e^{1'} = e^{2'} = e^{1''} = e^{2''} = 0$,

and their expected utilities are:

$$U_g^{1'} = p\left(p_g(m+1) - \frac{r}{2}\right)$$

$$U_g^{2'} = pp_g$$

in the non gatekeeping system and,

$$U_g^{1''} = p_g(m+1) - \frac{r}{2}$$

$$U_g^{2''} = p_g$$

in the gatekeeping system.

This result is intuitive. Since GPs implement a blind strategy of treatment carrying out an effort is costly and useless. The best level of effort is the lowest one.

Lemma If GPs treat their patient only when they diagnose a mild disease, the optimal level of diagnosis effort is:

$$e^{3'} = e^{3''} = \min\{\max\{\frac{r}{4k}, \varepsilon'\}, 1\}$$

with

$$\varepsilon' = \max\{\hat{e}', -\hat{e}'\} = \frac{|2p_g m - r|}{r}$$

and their expected utilities are:

$$U_g^{3'} = \left(\frac{p}{2}\right)\left(p_g(m+2)\right) - pk\frac{(e^{3'})^2}{2} - r\frac{p(1-e^{3'})}{4}$$

in the non gatekeeping system and,

$$U_g^{3'} = \left(\frac{1}{2}\right)\left(p_g(m+2)\right) - k\frac{(e^{3'})^2}{2} - r\frac{1-e^{3'}}{4}$$

in the gatekeeping system.

Proof The optimal effort carries out by GPs is given by

$$\underset{e}{Max} \quad U_g^{3'}$$

$$\begin{cases} e \geq \varepsilon' \\ e \leq 1 \end{cases}$$

in a non gatekeeping system and,

$$\underset{e}{Max} \quad U_g^{3''}$$

$$\begin{cases} e \geq \varepsilon' \\ e \leq 1 \end{cases}$$

in a gatekeeping system.

Using these lemmas we compute in the following section the incentives constraints.

4.3. The Incentives Constraints

In this presentation we focus on the second best contracts under which GPs have a financial incentive to implement the adequate treatment strategy[5].

Lemma The primary health care provider decides to treat his patient when a mild illness is diagnosed instead of implementing a blind treatment strategy if:

$$e^{3'}(r - 2ke^{3'}) \geq |r - 2p_g m|$$

with $e^{3'} = \min\{\frac{r}{4k}, 1\}$.

[5]For some values of the parameters which are not given here it may be socially optimal to give financial incentives to GPs to carry out blind strategies.

The primary health care provider performs a level of diagnosis effort $e^{3'} = \min\{\frac{r}{4k},1\}$.

His expected utility is:

$$U_g^{3'} = \left(\frac{p}{2}\right)\left(p_g(m+2)\right) - pk\frac{(e^{3'})^2}{2} - r\frac{p(1-e^{3'})}{4}$$

in a non gatekeeping system and,

$$U_g^{3''} = \left(\frac{1}{2}\right)\left(p_g(m+2)\right) - k\frac{(\hat{e}^3)^2}{2} - r\frac{(1-e^{3''})}{4}$$

and in a gatekeeping system.

In the following section, we compute the optimal payment schemes in the gatekeeping and non gatekeeping systems.

4.4. The Payment Schemes

In this section we compute for both health care organisations the incentive payment scheme under which GPs carry out the most adequate treatment. Since there is no asymmetric information between specialists and the health care insurer specialists' FFS are the same as in the case with perfect information.

In both cases we suppose that $k \leq \frac{r}{4}$.

4.4.1. In the Non Gatekeeping System

In this case the health insurer must solve:

$$\underset{p_g, p_s}{Max} \quad U_r^{3'}$$

$$\begin{cases} U_g^{3'} \geq \omega_g \\ U_s^{3'} \geq \omega_s \\ e^{3'}(r - 2ke^{3'}) \geq | r - 2p_g m | \\ p_g \geq 0 \\ p_s \geq 0 \\ e^{3'} = \min\{\frac{r}{4k},1\} \end{cases}$$

Proposition The FFS of the second best contract are:

- if $\frac{m}{p}\omega_g \leq k \leq \frac{r}{4}$,

$$p_g^{3'} = \frac{k}{m}$$

for GPs and

$$p_s^{3'} = \frac{2\omega_s}{(s+1)+(m+1)(1-p)}$$

for specialists.

Under this contract the utility of the health care insurer is:

$$U_r^{3'} = U_1\left(1-\frac{p}{2}\right)+U_2\left(\frac{p}{2}\right)-\frac{pk}{2}$$
$$-\frac{\lambda pk(m+2)}{2m}$$
$$-\lambda\omega_s$$
$$-(1+\lambda)\left(\left(\frac{1-p}{2}\right)c_1+\left(\frac{1-p}{2}+\frac{p}{2}\right)c_2\right)$$

- if $k \le \min\{\frac{m}{p}\omega_g, \frac{r}{4}\}$,

$$p_g^{3'} = \frac{2}{p(m+2)}\left(\omega_g+\frac{pk}{2}\right)$$

for GPs and

$$p_s^{3'} = \frac{2\omega_s}{(s+1)+(m+1)(1-p)}$$

for specialists.

Under the utility of the health care insurer is:

$$U_r^{3'} = U_1\left(1-\frac{p}{2}\right)+U_2\left(\frac{p}{2}\right)$$
$$-(1+\lambda)\left(pk\frac{1}{2}+\left(\frac{1-p}{2}\right)c_1+\left(\frac{1}{2}\right)c_2\right)$$
$$-\lambda(\omega_g+\omega_s)$$

Proof In this case the incentive constraint is:

$$|r-2p_g m| \le e^{3'}(r-2ke^{3'})$$

with

$$e^{3'} = \min\{\frac{r}{4k}, 1\}$$

The participation constraint of GPs is:

$$p_g \ge \frac{2}{p(m+2)}\left(\omega_g+pk\frac{(e^{3'})^2}{2}+r\frac{p(1-e^{3'})}{4}\right)$$

The participation constraint of specialists implies:

$$p_s^3 = \frac{2\omega_s}{(s+1)+(m+1)\left(1-p+\frac{p(1-e_s^3)}{2}\right)}$$

Since we have assumed that $k \le \frac{r}{4}$ the optimal level of effort is:

$$e^{3'} = 1$$

In this case the FFS of specialists is:

$$p_s^{3'} = \frac{2\omega_s}{(s+1)+(m+1)(1-p)}$$

The incentive constraint of GPs becomes:

$$|r-2p_g m| \le r-2k$$
$$\Leftrightarrow$$
$$k \le p_g m \le r-k$$

These last inequalities are always satisfied since $k \le \frac{r}{4}$.

Taking into account GPs' participation constraint we conclude that there are two cases:

- if $\frac{k}{m} \geq \frac{2}{p(m+2)}\left(\omega_g + \frac{pk}{2}\right) \Leftrightarrow k \geq \frac{m}{p}\omega_g$

Then under

$$p_g^{3'} = \frac{k}{m}$$

GPs have a financial incentive to implement the most adequate treatment.
Under these conditions, their utility is:

$$U_g^{3'} = \left(\frac{p}{2}\right)\left(\frac{k}{m}(m+2)\right) - \frac{pk}{2}$$

The utility of the insurer is:

$$U_r^{3'} = U_1\left(1-\frac{p}{2}\right) + U_2\left(\frac{p}{2}\right) - \frac{pk}{2}$$

$$-\frac{\lambda pk(m+2)}{2m}$$

$$-\lambda\omega_s$$

$$-(1+\lambda)\left(\left(\frac{1-p}{2}\right)c_1 + \left(\frac{1-p}{2}+\frac{p}{2}\right)c_2\right)$$

- *if* $\frac{k}{m} \leq \frac{2}{p(m+2)}\left(\omega_g + \frac{pk}{2}\right) \Leftrightarrow k \leq \frac{m}{p}\omega_g$,

Then GPs have a financial incentive to implement the most adequate treatment only under

$$p_g^{3'} = \frac{2}{p(m+2)}\left(\omega_g + \frac{pk}{2}\right)$$

Under these conditions, their utility is:

$$U_g^{3'} = \omega_g$$

The utility of the insurer is:

$$U_r^{3'} = U_1\left(1-\frac{p}{2}\right) + U_2\left(\frac{p}{2}\right)$$

$$-(1+\lambda)\left(pk\frac{1}{2} + \left(\frac{1-p}{2}\right)c_1 + \left(\frac{1}{2}\right)c_2\right)$$

$$-\lambda(\omega_g + \omega_s)$$

?

In the following section we compute the optimal payment schemes in a gatekeeping system.

4.4.2. In the Non Gatekeeping System

In this case the health insurer must solve:

$$Max \quad U_r^{3''}$$

$$p_g, p_s$$

$$\begin{cases} U_g^{3''} \geq \omega_g \\ U_s^{3''} \geq \omega_s \\ e^{3''}(r-2ke^{3''}) \geq \mid r-2p_g m \mid \\ p_g \geq 0 \\ p_s \geq 0 \\ e^{3''} = \min\{\frac{r}{4k},1\} \end{cases}$$

Proposition The FFS of the second best contract are:
- if $m\omega_g \leq k \leq \frac{r}{4}$,

$$p_g^{3''} = \frac{k}{m}$$

for GPs and

$$p_s^{3''} = \frac{2\omega_s}{(s+1)}$$

for specialists.
Under the utility of the health care insurer is:

$$U_r^{3''} = \frac{1}{2}(U_1+U_2) - \frac{k}{2} - \frac{\lambda k(m+2)}{2m} - \lambda\omega_s - \frac{(1+\lambda)}{2}c_2$$

- if $k \leq \min\{m\omega_g, \frac{r}{4}\}$,

$$p_g^{3''} = \frac{2}{(m+2)}\left(\omega_g + \frac{k}{2}\right)$$

for GPs and

$$p_s^{3''} = \frac{2\omega_s}{(s+1)}$$

for specialists.
Under the utility of the health care insurer is:

$$U_r^{3''} = \frac{1}{2}(U_1+U_2) - \frac{(1+\lambda)}{2}(k+c_2) - \lambda(\omega_g+\omega_s)$$

?
Proof In this case the incentive constraint is:

$$\mid r-2p_g m \mid \leq e^{3''}(r-2ke^{3''})$$

with

$$e^{3''} = \min\{\frac{r}{4k}, 1\}$$

The participation constraint of GPs is:

$$p_g \geq \frac{2}{(m+2)}\left(\omega_g + k\frac{(e^{3'})^2}{2} + r\frac{(1-e^{3'})}{4}\right)$$

The participation constraint of specialists implies:

$$p_s^{3''} = \frac{2\omega_s}{(s+1)+(m+1)\left(\frac{(1-e_s^{3''})}{2}\right)}$$

Since we have assumed that $k \leq \frac{r}{4}$ the optimal level of effort is:

$$e^{3''} = 1$$

In this case the FFS of specialists is:

$$p_s^{3''} = \frac{2\omega_s}{(s+1)}$$

and the participation constraint of GPs is:

$$p_g \geq \frac{2}{(m+2)}\left(\omega_g + \frac{k}{2}\right)$$

The incentive constraint of GPs becomes:

$$| r - 2p_g m | \leq r - 2k$$

$$\Leftrightarrow$$

$$k \leq p_g m \leq r - k$$

These last inequalities are always satisfied since $k \leq \frac{r}{4}$.

Taking into account GPs' participation constraint we conclude that there are two cases:

- if $\frac{k}{m} \geq \frac{2}{(m+2)}\left(\omega_g + \frac{k}{2}\right) \Leftrightarrow k \geq m\omega_g$

Then under

$$p_g^{3''} = \frac{k}{m}$$

GPs have a financial incentive to implement the most adequate treatment.

Under these conditions, their utility is:

$$U_g^{3''} = \left(\frac{1}{2}\right)\left(\frac{k}{m}(m+2)\right) - \frac{k}{2}$$

The utility of the insurer is:

$$U_r^{3''} = U_1\left(\frac{1}{2}\right) + U_2\left(\frac{1}{2}\right) - \frac{k}{2}$$

$$-\frac{\lambda k(m+2)}{2m}$$

$$-\lambda\omega_s$$

$$-\frac{(1+\lambda)}{2}c_2$$

- if $\frac{k}{m} \leq \frac{2}{(m+2)}\left(\omega_g + \frac{k}{2}\right) \Leftrightarrow k \leq m\omega_g$

Then GPs have a financial incentive to implement the most adequate treatment only under:

$$p_g^{3''} = \frac{2}{(m+2)}\left(\omega_g + \frac{k}{2}\right)$$

The utility of the insurer is:

$$U_r^{3''} = U_1\left(\frac{1}{2}\right) + U_2\left(\frac{p}{2}\right)$$

$$-(1+\lambda)\left(k\frac{1}{2} + +\left(\frac{1}{2}\right)c_2\right)$$

$$-\lambda(\omega_g + \omega_s)$$

?

Since we have computed the optimal payment schemes for both health care organisations we can compare them.

5. COMPARISONS

In this section we study the impact of restricting access to specialty care on:

- GPs' FFS,
- specialists' FFS,
- the insurer's utility.

Proposition Shifting from non gatekeeping to gatekeeping system increases the FFS of specialists.

Proof The sign of the difference of FFS is:

$$p_s^{3''} - p_s^{3'} = \frac{2\omega_s(m+1)(1-p)}{(s+1)+(m+1)(1-p)} > 0$$

?

This result is rather intuitive: since in a gatekeeping system fewer patients consult their specialists, the insurer must increase their FFS to satisfy their participation constraint.

Proposition According to the disutility of effort shifting from non gatekeeping to gatekeeping systems have no effect or decrease the GPs' FFS.

Proof In the case of a small disutility of effort the sign of the difference of the FFS is:

$$p_g^{3''} - p_g^{3'} = \frac{2}{(m+2)}\omega_g\left(1 - \frac{1}{p}\right) < 0$$

In the other case we see that:

$$p_g^{3'} = p_g^{3''} = \frac{k}{m}$$

?

This result is intuitive, particularly with a low disutility of effort. In such a case, since it is not costly for a GP to implement the most adequate treatment and since more patients consult him, the insurer can decrease the FFS of the GP which satisfies his participation constraint.

With a high disutility of effort, the FFS is the same in the gatekeeping and non gatekeeping system, due to two conflicting effects:

- on the one hand, more patients consult him, which should a priori decrease the level of the FFS,
- on the other hand, since the disutility of effort is high, the insurer must increase the FFS of the GP to give a financial incentive to perform the most adequate treatment.

We can conclude that the gatekeeping system is not necessarily cheaper than the non gatekeeping, because shifting from non gatekeeping to gatekeeping systems:

- increases the FFS of specialists,
- decreases or does not change the FFS of GPs.

Proposition The insurer prefers implementing a non gatekeeping system instead of a gatekeeping system depending on the costs and the utilities of treatments.

Proof In the case of a small disutility of effort the sign of the difference of insurer's utilities is given by the following equation:

$$U_r^{3'} - U_r^{3''} = \left(\frac{1-p}{2}\right)\left[(U_1 - U_2) - (1+\lambda)(c_1 - k)\right]$$

We can conclude that the gatekeeping system is socially optimal if:

$$c_1 > \frac{(U_1 - U_2)}{(1+\lambda)} + k$$

In the case of a large disutility of effort the sign of the difference of the insurer's utility is given by the following equation:

$$U_r^{3'} - U_r^{3''} = \left(\frac{1-p}{2}\right)\left((U_1 - U_2) + k + \frac{\lambda k(m+2)}{m} - (1+\lambda)c_1\right)$$

We can conclude that the gatekeeping system is socially optimal if:

$$c_1 \geq \frac{1}{(1+\lambda)}(U_1 - U_2) + k + \frac{\lambda k(m+2)}{m}$$

?

If we compare gatekeeping and non gatekeeping system taking into account:

- their cost and,
- their benefit

we conclude that a gatekeeping system is socially optimal only if the cost of treatment of a minor illness by specialists is too high.

6. CONCLUSION

In this paper, we compared gatekeeping and non gatekeeping systems under imperfect information. We assume that in both organisations physicians are paid under FFS.

Our results are the following ones:

- shifting from a non gatekeeping to gatekeeping systems increases the FFS of specialists, since fewer patients consult them,
- shifting from a non gatekeeping to gatekeeping systems decreases the FFS of GPs only if the disutility of effort is low, since more patients consult them. But if the disutility of effort is high the FFS in a gatekeeping system is identical to the FFS in a non gatekeeping system, even if more patients consult them.

We can conclude that a gatekeeping system is not cheaper than a non gatekeeping system even if it avoids useless treatment. Moreover a non gatekeeping may still be socially optimal if we trade off between:

- the costs of treatment,
- the utility of a faster recovery due to the free access to specialty care.

Our future research follows one main direction. We are setting up another model to study the PPOs. In the model presented here we assume that patients must go to their GPs first. But in some PPOs insurers have implemented co-payments which allow patients to get to specialty care according to their symptoms. In this new model we first compute the optimal health care payment when the insurer perfectly observes patients' symptoms and GPs' decisions, i.e. the level of effort and the treatment carried out. Then we assume that the insurer observes neither GPs' decisions nor the symptoms of the patients. It means that the health care insurer is in a

situation with double moral hazard and adverse selection. In this case the health insurer first computes optimal co-payments under which patients consult primary health care providers if they have mild symptoms and consult secondary health care providers otherwise. Then he computes the optimal payments schemes under which GPs treat their patients when they diagnose the mild disease and to refer them otherwise. We finally compare this health care organisation to the genuine gatekeeping and non-gatekeeping systems.

7. BIBLIOGRAPHY

BARDEY D. (2002) - Demande induite et réglementation de médecins altruistes, *Revue économique*, 53, pages 581-588.

JELOVAC I. (2001) - Physicians' payment contracts, treatment decisions and diagnosis accuracy, *Health Economics*, 10, pages 9-25.

LEVAGGI R. AND ROCHAIX L. (2007) - Exit, Choice or Loyalty: Patient Driven Competition in Primary Care, *Annals of Public and Cooperative Economics*, 78, pages 501-535.

MARINOSO B.G. AND JELOVAC I. (2003) - GPs'payments contracts and their referral practice, *Journal of Health Economics*, 22, pages 617-635.

Queueing networks with infinite servers in each node in the management of a two echelons repair system

MANUEL ALBERTO M. FERREIRA
E-mail : manuel.ferreira@iscte.pt
MARINA ANDRADE
E-mail : marina.andrade@iscte.pt
ISCTE Business School

Resumo

Após uma apresentação de resultados sobre filas de espera com infinitos servidores, considerando essencialmente o período de ocupação, constrói-se um modelo baseado em redes de filas de espera com infinitos servidores em cada nó, para estudar um sistema de reparações com dois escalões de, por exemplo, uma frota de aviões, barcos ou camiões. Os clientes são as avarias. E o seu tempo de serviço é o que decorre a partir do instante em que elas ocorrem até que estejam completamente reparadas. As reparações das avarias ocorrem numa base ou numa estação remota. Todas as avarias detectadas na base são reparadas aí. Algumas das avarias detectadas na estação são lá reparadas e as outras na base. O modelo permite a determinação de medidas de performance do sistema. Nesta aplicação consideram-se modelos de Carrillo (1991) e Ferreira (1996) que são melhorados e completados. A teoria é ilustrada com um exemplo numérico.

Palavras-chave: $M|G|\infty$, rede de filas de espera, período de ocupação, logística.

Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

After a report of results about infinite servers queues, essentially on its busy period, a model is built, using networks of queues with infinite servers in each node, to study a two echelons repair system for example of a fleet of aircraft, of shipping or of trucks. The costumers are the failures. And its service time is the time that goes from the instant at which they occur till they are completely repaired. The failures repairs occur in a base or in a remote station. The whole failures detected in the base are repaired there. Some of the failures detected in the station are repaired there and the others in the base. The results referred above allow the determination of a two echelons repair system performance measures. In this application models of Carrillo (1991) and Ferreira (1996) are considered, improved and completed. The theory is illustrated with a numerical example.

Key Words: $M|G|\infty$, network of queues, busy period, logistics.

Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUCTION

In the $M|G|\infty$ queueing system the customers arrive according to a Poisson process at rate λ, each of them receives a service whose length is a positive random variable with distribution function $G(.)$ and mean value α. So

$$\alpha = \int_0^\infty [1 - G(t)]dt \qquad (1.1),$$

There are infinite servers. So when a customer arrives it always finds an available server. The service of a customer is independent of the services of the other customers' services and of the arrival process. An important parameter is the traffic intensity, called ρ, being

$$\rho = \lambda\alpha \qquad (1.2).$$

It is obvious that in an $M|G|\infty$ queue there are neither losses nor waiting. In fact there is no queue in the formal sense of the word. For these systems it is not so important to study the population process as for other ones with losses or waiting. Generally it is much more interesting to study other processes as, for instance, the busy period. The busy period of a queue system begins when a customer arrives there, finding it empty, and ends when a customer leaves the system letting it empty. During the busy period there is always at least one customer in the system. A network of queues is a collection of nodes arbitrarily connected by arcs, instantaneously traversed by costumers, where

- An arrival process is associated to each node.
- There is a commutation process that commands the different customers' paths.

Call J the network number of nodes. If $J < \infty$, the nodes are numbered $1,2,...,J$ and $U = \{1,2,...,J\}$. The arrival processes may be the result of exogenous arrivals, from the outside of the collection, and of endogenous arrivals, from the other collection nodes. A network is open if any customer can enter it or leave it. A network is closed if it has a fixed member of customers moving from node to node, with neither exterior arrivals nor departures. The networks opened for some customers and closed for others are called mixed.

In this paper open networks, with infinite servers in each node, and Poisson exogenous arrival rate λ are considered. So

$$\Lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_j \end{bmatrix} \qquad (1.3)$$

is the network exogenous arrival rate vector. The rate λ_j, is the exogenous arrival rate at node $j, j = 1,2,...,J$ and $\sum_{j=1}^{J} \lambda_j = \lambda$. The matrix

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1J} \\ p_{21} & p_{22} & \cdots & p_{2J} \\ \vdots & \vdots & & \vdots \\ p_{J1} & p_{J2} & \cdots & p_{JJ} \end{bmatrix} \qquad (1.4)$$

is the commutation process matrix, being p_{jl} the probability of a customer, after ending its service at node j, go to node l, $j,l = 1,2,...,J$. The probability $q_j = 1 - \sum_{l=1}^{J} p_{jl}$ is the probability that a customer leaves the network from node j, $j = 1,2,...,J$. It is supposed that P does not change with t and is independent of everything that is happening in the network.

It will be shown that a network of this kind is equivalent to a $M|G|\infty$ system with Poisson process arrivals at rate λ, where each customer service time is its sojourn time in network. But first some results about the $M|G|\infty$ systems busy period, that are also useful for these networks of queues evidently, are presented. Then a model is built, using these networks with infinite servers in each node, to study a two echelons repair system of a fleet of aircraft, of shipping or

of trucks. The customers are the failures. And its service time is the time that goes from the instant at which they occur till the one at which they are completely repaired. The results referred above, about the $M|G|\infty$ queue busy period, allow the determination of some system performance measures. The theory is illustrated with a very simple and short numerical example.

2. THE M| G| ∞ QUEUE BUSY PERIOD

Call B the $M|G|\infty$ queue system busy period length. The mean value of B is, see Takács (1962), whatever is $G(.)$,

$$E[B]=\frac{e^{\rho}-1}{\lambda} \qquad (2.1).$$

Being $R(t)$ the mean number of busy periods that begin in $[0,t]$ ($t=0$ is the beginning of a busy period) see Ferreira (1995),

$$e^{-\rho}(1+\lambda t)\le R(t)\le 1+\lambda t \qquad (2.2).$$

Be N_B the mean number of the customers served during a busy period in the $M|G|\infty$ queue systems. So, Ferreira (2001),

- If $G(.)$ is exponential

$$N_B^M = e^{\rho} \qquad (2.3).$$

- For any other service distribution function

$$N_B \cong \frac{e^{\rho(\gamma_s^2+1)}\left(\rho(\gamma_s^2+1)+1\right)+\rho(\gamma_s^2+1)-1}{2\rho(\gamma_s^2+1)} \qquad (2.4)$$

where γ_s, is the variation coefficient of $G(.)$.

3. THE SOJOURN TIME OF A CUSTOMER LAPLACE TRANSFORM IN A NETWORK OF QUEUES WITH INFINITE SERVERS IN EACH NODE AND POISSON EXOGENOUS ARRIVALS

Note that for these queueing networks

- The sojourn time of a costumer in each node is the service time, since there is not waiting,
- The sojourn times of a customer in the various nodes are independent.

The sojourn time of a customer in the network, known its path, is distributed as the convolution of the service times in the nodes for which it passes. So the sojourn time distribution will be the mixture of the convolutions, related to each possible path, being the weight of each one given by the respective path probability. These probabilities are known and do not depend on time since are known and do not depend on time the exogenous arrival rates to the various nodes, the commutation probabilities and the probabilities to leave the network.

It will be, generally, difficult to get an efficient formula to the sojourn time distribution function, based on the whole possible paths direct enumeration because

- The number of paths may be infinite or enormously great, even in the situation of few nodes networks.
- The convolutions can lead to intractable analytic expressions.

Using matrixes whose shape is suggested by (1.3) and (1.4) a simple formula is deduced to the sojourn time Laplace Transform, as a function of the service times in each node Laplace Transforms (Ferreira and Ramalhoto (1990) and Ferreira (1996)).

Be T the network sojourn time of a costumer, and S_j, its service time at node j, $j=1,2,...J$. Be $G(t)$ and $G_j(t)$, $j=1,2,...,J$ the T and S_j, $j=1,2,...,J$, distribution functions respectively, being $\overline{G}(s)$ and $\overline{G}_j(s)$, $j=1,2,...,J$ the T and S_j, $j=1,2,...,J$, Laplace Transforms, respectively.

Making

$$\Lambda(s)=\begin{bmatrix}\lambda_1\overline{G}_1(s)\\ \lambda_2\overline{G}_2(s)\\ \vdots\\ \lambda_J\overline{G}_J(s)\end{bmatrix} \tag{3.1}$$

and

$$P(s)=\begin{bmatrix}p_{11}\overline{G}_1(s) & p_{12}\overline{G}_2(s) & \cdots & p_{1J}\overline{G}_J(s)\\ p_{12}\overline{G}_1(s) & p_{22}\overline{G}_2(s) & \cdots & p_{2J}\overline{G}_J(s)\\ \vdots & \vdots & & \vdots\\ p_{J1}\overline{G}_1(s) & p_{J2}\overline{G}_2(2) & \cdots & p_{JJ}\overline{G}_J(s)\end{bmatrix} \tag{3.2}$$

it results

$$\overline{G}(s)=\sum_{n=0}^{\infty}\lambda^{-1}\Lambda^T(s)P^n(s)(1-P)\mathrm{A} \tag{3.3}$$

being A a column with $J1$`s.

Note that

- $\Lambda(0)=\Lambda$
- $P(0)=P$

In (3.3) the whole paths and the respective probabilities are considered, associating them simultaneously its nodes service times Laplace Transforms product. Each path begins in node j with probability $\lambda_j\lambda^{-1}$ and will finish in node k with probability $1-\sum_{j=1}^{J}P_{kj}$, $j=1,2,...,J$, $k=1,2,...,J$.

The final formula is obtained noting that (3.3) may be put in the form

$$\overline{G}(s)=\lambda^{-1}\Lambda^T(s)(I-P(s))^{-1}(1-P)\mathrm{A} \tag{3.4}$$

since $|I-P(s)|\neq 0$.

4. A TWO ECHELONS REPAIR SYSTEM

The results presented may be applied in logistics. Suppose a fleet of aircraft, of shipping or of trucks which failures repairs occur in a base or in a remote station. The whole failures detected in the base are repaired there. Some of the failures detected in the station are repaired in the base with probability p, being necessary to transport them to the base, and the others in the station. Here the service time is the time that goes from the instant at which the failure occur till the one at which it is completely repaired. When it is necessary to transport an item with a failure from the remote station to the base it is assumed that is immediately possible, being the service time, now, the time that the transport lasts. It is also supposed that the failures occur

according to a Poisson process at rate λ, being some detected in the remote station with probability q and the others in the base.

This situation may be modeled as a network of queues with three nodes

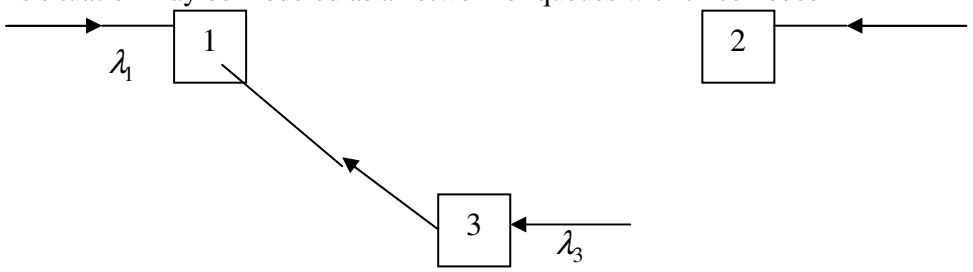


Figure 1

where

- 1 is the base,
- 2 is the remote station,
- 3 considers the necessary transports from the remote station to the base.

Representing the variables related with each node for the same letter, as in the former sections, but with an index corresponding to the node, obviously,

$$\Lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_J \end{bmatrix} = \begin{bmatrix} (1-q)\lambda \\ (1-p)q\lambda \\ pq\lambda \end{bmatrix} \qquad (4.1),$$

$$P = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad (4.2).$$

Being the system, globally, a $M|G|\infty$ queue and after (3.4)

$$\overline{G}(s) = \lambda^{-1}\left[(1-q)\lambda\overline{G}_1(s) \quad (1-p)q\lambda\overline{G}_2 \quad pq\lambda\overline{G}_3(s)\right]$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\overline{G}_1(S) & 0 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \left[(1-q)\overline{G}_1(s) \quad (1-p)q\overline{G}_2 \quad pq\overline{G}_3(s)\right]$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \overline{G}_1(S) & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \left[(1-q)\overline{G}_1(s) + (1-p)q\overline{G}_2(s) + pq\overline{G}_1(s)\overline{G}_3(s)\right]$$

So the service time distribution function is

$$G(t) = (1-q)G_1(t) + (1-p)qG_2(t) + pqG_{13}(t) \qquad (4.3)$$

being G_{13} the distribution function of the convolution of the service time distributions in nodes 1 and 3.

Three $M|G|\infty$ queues can also be considered

- One related with the repairs in the base, whose failures were detected there:

$$\begin{aligned} \lambda_b &= (1-q)\lambda \\ G_b(t) &= G_1(t) \\ \rho_b &= (1-q)\lambda\alpha_1 \end{aligned} \qquad (4.4)$$

- Other related with the repairs in the remote station:

$$\begin{aligned} \lambda_{st} &= (1-p)q\lambda \\ G_{st}(t) &= G_2(t) \\ \rho_{st} &= (1-p)q\lambda\alpha_2 \end{aligned} \qquad (4.5)$$

- And still other related with the repairs in the base after transport from the remote station

$$\begin{aligned} \lambda_{tr} &= pq\lambda \\ G_{tr}(t) &= G_{13}(t) \\ \rho_{tr} &= pq\lambda(\alpha_1+\alpha_3) \end{aligned} \qquad (4.6).$$

Concerning the application of (2.4), being σ_1^2, σ_2^2 and σ_3^2 the variances corresponding to $G_1(.)$, $G_2(.)$ and $G_3(.)$, respectively

$$\begin{aligned} \gamma_{sb} &= \gamma_{s1} \\ \gamma_{sst} &= \gamma_{s2} \\ \gamma_{str} &= \frac{\sqrt{\sigma_1^2+\sigma_3^2}}{\alpha_1+\alpha_3} \end{aligned} \qquad (4.7)$$

and the coefficient of variation corresponding to the distribution function given for (4.3)

$$\gamma_s = \sqrt{\frac{(1-q)\left(\sigma_1^2+\alpha_1^2\right)+(1-p)q\left(\sigma_2^2+\alpha_2^2\right)+pq\left(\sigma_1^2+\sigma_3^2+(\alpha_1+\alpha_3)^2\right)}{\left((1-q)\alpha_1+(1-p)q\alpha_2+pq(\alpha_1+\alpha_3)\right)^2}-1} \qquad (4.8).$$

5. EXAMPLE

Suppose that $G_1(.)$ and $G_2(.)$ are both exponential with mean 1 week and $G_3(.)$ is constant with value 1 week, $q=0.3$ and $\lambda=\frac{1}{4weeks}$ (1 per month). Now $p=0.9$ and considering 1 year (52 weeks) of operation is it possible to conclude if decreasing p (that is: increasing the station capacity of repairs) there would be any advantage?

Making $p=0.9;0.8;...0.1$

- For the global system

Table 1

p	$\frac{e^{\rho}-1}{\lambda}$	$e^{-\rho}(1+\lambda 52)$	$(1+\lambda 52)$	N_B
0.9	1.51	10	14	2.04
0.8	1.45	10	14	2.03
0.7	1.40	10	14	2.02
0.6	1.40	10	14	2.03
0.5	1.35	10	14	2.02
0.4	1.29	11	14	2.01
0.3	1.24	11	14	1.99
0.2	1.24	11	14	2.00
0.1	1.19	11	14	1.99

- For the remote station

Table 2

p	$\frac{e^{\rho_{st}}-1}{\lambda_{st}}$	$e^{-\rho_{st}}\left(1+\lambda_{st}52\right)$	$\left(1+\lambda_{st}52\right)$	N_B
0.9	1.00	1.38	1.39	1.01
0.8	1.01	1.75	1.78	1.02
0.7	1.01	2.12	2.17	1.02
0.6	1.02	2.48	2.56	1.03
0.5	1.02	2.84	2.95	1.04
0.4	1.02	3.19	3.34	1.05
0.3	1.03	3.54	3.73	1.05
0.2	1.03	3.88	4.12	1.06
0.1	1.03	4.22	4.51	1.07

- For the repairs in the base after transport

Table 3

p	$\frac{e^{\rho_{tr}}-1}{\lambda_{tr}}$	$e^{\rho_{tr}}\left(1+\lambda_{tr}52\right)$	$\left(1+\lambda_{tr}52\right)$	N_B
0.9	2.14	3.94	4.51	3.49
0.8	2.12	3.65	4.12	3.13
0.7	2.11	3.36	3.73	2.81
0.6	2.09	3.05	3.34	2.54
0.5	2.08	2.74	2.95	2.30
0.4	2.06	2.41	2.56	2.10
0.3	2.05	2.07	2.17	1.91
0.2	2.03	1.73	1.78	1.76
0.1	2.02	1.37	1.39	1.62

- For the base

$\frac{e^{\rho_b}-1}{\lambda_b}=1.09$, $e^{-\rho b}\left(1+\lambda_b 52\right)=8.48$, $1+\lambda_b 52=10.1$ and $N_B^M \cong 1.19$, whatever is p.

Of course, in the operation of a fleet, it interests big idle periods and little busy periods. And if these occur it is good that they are as rare as possible, with a short number of failures. So it is possible to conclude that the global system improves its performance as p decreases but very lightly. The remote station grows worse has it was expected and the repairs in the base after transport improve its performance. Note, for instance, that to guarantee less than two busy periods for the repairs in the base after transport it is essential that $p \leq 0.2$. But, for this, it is necessary to spend money in staff and material in the remote station. It is necessary to balance these expenses with the savings in transports.

6. CONCLUSIONS

To apply this mode it is necessary to verify if the failures occur according to a Poisson process. This hypothesis has to be tested. After this, is very simple to apply the other results.

Carrillo (1991) studied a model looking like the one presented here. But he did not consider either the possibility of transport from the station to the base or the busy period. So these conceptions allow, in a very simple way, to evaluate the performance of a fleet as the example presented has shown.

BIBLIOGRAPHY

ALBRIGHT, S.C. AND A. GUPTA, (1993), Steady-State Approximation of a Multiechelon Multi-Indentured Repairable-Item Inventory System with a Single Repair Facility, *Naval research logistics* 40 no. 4, 479-494.

CARRILLO, M. J. (1991), "Extensions of Palm's Theorem: A Review", *Management Science*, Vol. 37, n.º 6, p.739 – 744.

FERREIRA, M. A. M. (1995), "Comportamento transeunte e período de ocupação de sistemas de filas de espera sem espera", PhD Thesis discussed in ISCTE, Supervisor: Prof. Augusto A. Albuquerque.

FERREIRA, M. A. M. (1996), "Sistemas de filas de espera com infinitos servidores – Uma aplicação em logística", *Revista de Estatística*, Vol. 1, INE.

FERREIRA, M. A. M. (2001), "Mean number of the Customers Served During a Busy Period in the $M|G|\infty$ Queueing System", *Statistical Review*, Vol. 3, INE.

FERREIRA, M. A. M. E RAMALHOTO, M. E. (1990), "Algoritmo para o cálculo da Transformada Laplace-Stieltjes do tempo de permanência de um cliente em redes de filas de espera abertas, com distribuição de equilíbrio da forma produto, com tempos de permanência independentes em cada nó", *Actas da 1.ª Conferência em Estatística e Optimização*, CEA (INIC) e DEIOC, Universidade de Lisboa.

GROSS, D, D.R. MILLER AND R.M. SOLAND, (1983), A closed queueing network model for multi-echelon repairable item provisioning, *IIE Transactions* 15 no. 4, 344-352.

RAMALHOTO, M. E. AND FERREIRA, M. A. M., (1996), Some Further Properties of the Busy Period of an $M|G|\infty$ Queue, *Central European Journal for Operations Research and Economics*, Volume 4, Number 4, 251-278.

SLEPTCHENKO, A, M.C. VAN DER HEIJDEN AND A. VAN HARTEN, (2002), Effects of finite repair capacity in multi-echelon, multi-indenture service part supply systems, *International Journal of Production Economics* 79, 209-230.

TAKÁCS, L. (1962), "An Introduction to Queueing Theory", Oxford University Press, New York.

ZIJM, W.H.M., AND M. AVSAR, (2003), Capacitated two-indenture models for repairable item systems, *International Journal of Production Economics* 81-82, 573-588.

ANÁLISIS DE LA IMAGEN DE UN DESTINO TURÍSTICO MEDIANTE UN MODELO DE ESTRUCTURA LATENTE SIMULTÁNEO. APLICACIÓN A LAS CIUDADES MEDIAS DE ANDALUCÍA

MARCELINO SÁNCHEZ RIVERO
E-mail: sanriver@unex.es
Universidad de Extremadura

JUAN IGNACIO PULIDO FERNÁNDEZ
Universidad de Jaén

MIGUEL ÁNGEL FAJARDO CALDERA
Universidad de Extremadura

Abstract

En un entorno tan competitivo y exigente como el que caracteriza actualmente a la actividad turística, conseguir un buen posicionamiento a través de su imagen es un objetivo estratégico para los destinos turísticos que compiten en un mercado globalizado. Sin embargo, la imagen de un destino es un concepto latente, que no puede ser apreciado de una forma directa, y al que únicamente es posible aproximarse mediante un conjunto de indicadores.
El objetivo de esta comunicación es identificar la variable latente *imagen del destino* y, a partir de ella, segmentar la demanda turística en función de la imagen que los destinos turísticos crean en la misma. Para ello, se ha utilizado el análisis de estructura latente simultáneo que permite, además, contrastar el grado de homogeneidad o de heterogeneidad entre dos o más mercados turísticos. Tras presentar los aspectos metodológicos más significativos de esta técnica, se han utilizado los datos de un estudio sobre el comportamiento de la demanda turística en las ciudades medias del interior de Andalucía para determinar si los turistas con una marcada vocación de turismo cultural tienen una imagen similar (homogénea) de este destino turístico que el resto de turistas, empleando para ello un conjunto de cinco indicadores dicotómicos obtenidos de la valoración, en términos de satisfacción, otorgada por los visitantes a otros tantos aspectos característicos de las ciudades medias del interior de Andalucía.

Palabras Clave: imagen, destino turístico, análisis de clases latentes, análisis simultáneo, homogeneidad, turista cultural, gestión de destinos.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

1. INTRODUCCIÓN

La situación actual de los mercados turísticos se caracteriza, entre otras cosas, por una creciente competencia entre destinos, consecuencia de fenómenos como la globalización (a la que acompaña la desregulación de sectores económicos clave, como telecomunicaciones, electricidad, transporte o servicios financieros), la generalización en el uso de las TICs o los avances en las infraestructuras de transporte. En tal contexto, los gestores de los destinos turísticos están empezando a reorientar sus planteamientos estratégicos hacia aspectos como la imagen, a fin de generar ventajas competitivas más allá de los recursos e infraestructuras que el lugar posee.

La capacidad que tenga el destino para comunicar sus atractivos y, sobre todo, para adaptarlos a las necesidades, gustos y expectativas de los potenciales clientes, se convierte hoy en uno de los principales instrumentos estratégicos para garantizar una mejora de su

posicionamiento competitivo. Tanto más en destinos emergentes, con una imagen aún por construir, o muy débil en los mercados. Como reconocen Kotler el al. (1993), los lugares en general y los destinos turísticos en particular deben gestionarse desde una perspectiva estratégica para ser competitivos, y la imagen juega un papel fundamental en el proceso de posicionamiento (Calantone et al., 1989).

La investigación cuyos resultados se presentan en este trabajo se ha centrado en las ciudades medias del interior de Andalucía, caracterizadas por disponer de una oferta turística en la que los aspectos patrimoniales y culturales son los protagonistas. El principal objetivo es analizar si los turistas de marcada vocación cultural tienen una imagen de estos destinos turísticos diferente de la que tienen el resto de turistas. En la medida en que ello sea así, y dado que estas ciudades se intentan posicionar en los mercados de turismo cultural, pues son en los que tienen una mayor capacidad para competir debido a la cantidad y calidad de sus recursos, los agentes implicados en la gestión turística de estos territorios deberán desarrollar un conjunto de acciones que permitan potenciar los aspectos positivos valorados por estos turistas y mejorar los aspectos peor valorados, a fin de consolidar una imagen efectiva que permita mantener una ventaja competitiva sobre el resto de destinos.

Para ello, se analiza la valoración que ambos segmentos de turistas (los de marcada vocación cultural y el resto) realizan respecto a cinco atributos asociados al concepto de "imagen de un destino turístico". Se trata de la señalización, la relación calidad-precio, la limpieza, la tranquilidad y la atención y trato recibidos durante la visita. Se han utilizado técnicas de análisis de clases latentes, poco comunes en este tipo de investigaciones, lo que ya supone una aportación novedosa e interesante, completada con la posibilidad, no sólo de de delimitar y cuantificar los diferentes segmentos, sino, incluso, de conocer la probabilidad de que un individuo de cualquiera de los grupos identificados tenga un comportamiento concreto.

2. REVISIÓN DE LA LITERATURA

2.1 La imagen de un destino turístico. Concepto, dimensiones y medida

En una primera aproximación, el concepto de imagen aparece relacionado con el conjunto de impresiones, o la percepción global, que se tiene sobre productos, personas, organizaciones, ciudades o países (Altés, 1993; Foxall y Goldsmith, 1994; Hampton, Guy y Sinkula, 1987; Luque, Del Barrio, Ibáñez y Rodríguez, 2004).

Dowling (1986) conceptúa la imagen como "un conjunto de significados por los que se conoce un objeto, y mediante los cuales los individuos lo describen, recuerdan, y relacionan. Es el resultado de la interacción de creencias, ideas, sentimientos, e impresiones que una persona tiene sobre un objeto". Y Luque et al. (2004) señalan que "en general, una imagen se relaciona con un conjunto de adjetivos vinculados a una serie de conocimientos denominados creencias o estereotipos. Por esto, la imagen hacia un objeto es el resultado neto de las experiencias, creencias, sentimientos e información de la que disponen un grupo de individuos en relación con un objeto". Se trata, por tanto, de un concepto vinculado con aspectos cognitivos, pero que también tiene una clara influencia de aspectos afectivos (Dobni y Zinkhan,1990; Nguyen y LeBlanc, 2001).

La literatura científica sobre turismo no ha transitado por caminos muy diferentes. De hecho, una buena parte de las definiciones realizadas sobre imagen de un destino turístico se limitan a conceptuarla como la percepción global, o el conjunto de impresiones, sobre un lugar, o como el retrato mental de un destino (Alhemoud y Armstrong, 1996; Bigné, Sánchez y Sanz, 2005; Calantone, Di Benetton, Hakam y Bojanic, 1989; Fakeye y Crompton, 1991; Hunt, 1971, 1975; Milman y Pizam, 1995; Phelps, 1986; Valls, 1992), mientras que otros insisten en su descomposición en atributos (Dadgostar e Isostalo, 1992; Gartner, 1986, 1989; Seaton y Benett, 1996).

Fakeye y Crompton (1991) ponen de manifiesto el carácter de constructo mental de la imagen y el papel que desempeña el procesamiento de información en la construcción del mismo. En otros casos (Crompton, 1979; Dadgostar e Isostalo, 1992; Gartner, 1986; Gartner

1989; Kotler, Haider y Rein, 1993; Richardson y Crompton, 1988; Season y Benett, 1996), en vez de considerar la imagen turística como percepción global, se la considera como la suma de impresiones, o percepción de atributos, que se tiene sobre un destino. Algunas definiciones sugieren que la imagen turística es una representación mental, resaltándose el carácter subjetivo de la misma (Alhemoud y Arsmstrong, 1996; Crompton, 1979; Fakeye y Crompton, 1991; Moutinho, 1987; Seaton y Benett, 1996). Mientras que otras resaltan su carácter connotativo y emocional (Mazanec, 1994). Por contra, autores como Moutinho (1987) o Alhemoud y Armstrong (1996) mantienen una posición intermedia, reconociendo la existencia tanto de componentes afectivos como de elementos cognitivos.

Coincidiendo con Bigné, Sánchez y Currás (2007, p. 60), la imagen de un destino turístico puede considerarse como "la interpretación subjetiva de la realidad realizada por el turista, interviniendo en su formación tanto elementos cognitivos como afectivos". Es, en definitiva, "todo lo que evoca el destino en el viajero, cualquier idea, creencia, sentimiento o actitud que asocia con el lugar y que no se limita a la percepción de estímulos sensoriales, sino que incluye también códigos semánticos, y se ve afectada tanto por factores internos como externos al individuo" (Sánchez y Sanz, 2004, p. 468).

La mayoría de los trabajos revisados (Ashworth y Goodall, 1988; Bordas y Rubio, 1993; Cooper, Fletcher, Gilbert y Wanhill, 1993; Gunn, 1972; Seaton y Benett, 1996; Telisman-Kosuta, 1994; Waitt, 1996) coinciden en la existencia de dos tipos de imagen turística: la *imagen orgánica* y la *imagen inducida*. La primera se forma a partir de fuentes de información que no persiguen intencionadamente la promoción del destino (consejos de amigos, familiares o conocidos –comunicación boca-oreja-, información aparecida en medios de comunicación, o la influencia del sistema educativo). La imagen inducida, en cambio, es consecuencia de los esfuerzos deliberados de comunicación de las entidades turísticas (tanto públicas, como privadas), con el objeto de promocionar el destino e inducir al turista potencial a que lo visite.

Otros autores (Echtner y Ritchie, 1991; Fakeye y Crompton, 1991; Gunn, 1988) incorporan un tercer tipo de imagen turística, la *imagen compleja*, formada una vez que se visita el destino como consecuencia de la propia experiencia turística, que permite al viajero disponer de una imagen más completa del lugar, gracias a una información más detallada y real.

Autores como Alhemoud y Arsmstrong (1996); Bordas y Rubio (1993); Fakeye y Crompton (1991); Gartner (1989) y Gunn (1988) consideran que la formación de la imagen global es fruto de una evolución de la imagen orgánica a la inducida y, en su caso, a la compleja.

Esta distinción entre imagen orgánica e inducida es casi exclusiva de los destinos turísticos ya que, en el resto de productos, los mensajes promocionales son la principal fuente de información (Gunn, 1988). Y, en cualquier caso, no deja de ser un mero recurso teórico ya que, en la realidad, es imposible aislar la influencia de las distintas fuentes de información sobre la imagen global del destino. Además, el efecto total del conjunto de factores no equivale a la suma de los efectos individuales, debido a que se producen interacciones entre los mismos.

Finalmente, señalar que la literatura recoge dos tipos de técnicas para la medición de la imagen de un destino turístico: estructuradas y no estructuradas (San Martín et al., 2006). En la primera técnica, también denominada aproximación multi-atributo, el investigador fija de antemano un conjunto de atributos más o menos comunes para todos los destinos turísticos (Echtner y Ritchie, 1991, 1993). A continuación, mediante una escala Likert o de diferencial semántico, se mide la percepción del individuo para cada uno de los atributos, capturándose de esta forma el componente común de la imagen de un determinado destino. A este respecto, cabe puntualizar que la cantidad e identidad de los atributos, así como la amplitud de las escalas de medición, suelen variar de unos trabajos a otros (Bigné y Sánchez, 2001).

Por su parte, la técnica no estructurada se fundamenta en la utilización de preguntas abiertas que permiten al individuo realizar descripciones libres sobre el destino (Reilly, 1990; Tapachai y Waryszak, 2000). De este modo, esta técnica permite medir o capturar las impresiones holísticas del individuo sobre el lugar –componente holístico–, así como las características que a su juicio son únicas o distintivas del destino –componente único de la imagen– (Echtner y Ritchie, 1993).

A partir de la metodología propuesta por Echtner y Ritchie (1993), son varias las investigaciones recientes que han utilizado conjuntamente ambas técnicas para medir todos los

componentes de la imagen de un destino turístico (Choi, Chan y Wu, 1999; Baloglu y Mangaloglu, 2001; Bigné y Sánchez, 2001; Hui y Wan, 2003; O´Leary y Deegan, 2003; Hsu, Wolfe y Kang, 2004).

2.2 La importancia de la imagen en la competitividad de los destinos turísticos

Aunque la literatura científica sobre imagen de destinos turísticos es relativamente reciente (Pike, 2000), viene siendo abundante en los últimos años, lo que evidencia una creciente importancia de la imagen en el desarrollo de un destino turístico, entendido éste como el proceso que permite seleccionar, desarrollar y promocionar un territorio como producto turístico que cubra las necesidades de los distintos consumidores y usuarios (Bramwell y Rawding, 1996). Una imagen adecuada permite transmitir a los potenciales clientes los aspectos diferenciales del lugar, lo que genera ventajas competitivas (Goodall, 1990), a la vez que genera confianza en el destino, favoreciendo la atracción y lealtad de los turistas (Bigné, Sánchez y Sánchez, 2001; Bramwell y Rawding, 1996).

La imagen ejerce una fuerte influencia en el comportamiento de la demanda turística, tanto con anterioridad al viaje, como durante la estancia e, incluso, en el futuro (Beerli y Martín, 2004). Hay que tener en cuenta que los turistas tienen un conocimiento limitado de los destinos que no han visitado, por lo que una imagen fuerte, positiva, diferenciada y reconocible en los mercados de origen de estos potenciales turistas apoya significativamente el posicionamiento competitivo del destino (Hunt, 1975; Goodrich, 1978; Pearce, 1982; Woodside y Lysonsky, 1989; Ross, 1993).

Por otra parte, la experiencia vivida en el destino puede conducir a una modificación de la imagen inicial percibida por el turista, que sólo en el caso de ser positiva conseguirá potenciar su lealtad (Sánchez y Sanz, 2004). Por tanto, la *imagen compleja* del destino tiene una gran influencia en las intenciones de volver en un futuro y en la predisposición a recomendarlo (fidelidad) (Court y Lupton, 1997; Bigné et al., 2001), lo que tiene que ver con la capacidad del destino de proporcionar a los visitantes una experiencia que se corresponda con sus necesidades y encaje con la imagen que ellos tenían del mismo (expectativas) (Chon, Weaver y Kim, 1991; Court y Lupton, 1997; Bigné, Sánchez y Sánchez, 2001; Joppe, Martín y Waalen, 2001). De ahí la necesidad, no sólo de construir una imagen efectiva (Ahmed, 1996), sino de que ésta se pueda asociar a expectativas positivas posibles de satisfacer por parte del destino (Buhalis, 2000), por lo que es muy importante conocer de qué manera valora el turista aspectos básicos vinculados con la imagen del destino y con sus expectativas respecto al mismo.

De hecho, la identificación de los distintos componentes de la imagen puede contribuir de forma importante al desarrollo de estrategias de segmentación más efectivas y a la implementación de acciones promocionales dirigidas a segmentos clave (Ahmed, 1996). Por consiguiente, cada vez en mayor medida, los gestores turísticos tratan de desarrollar comunicaciones efectivas con su público objetivo, de manera que puedan controlar y decidir sobre la gestión de su imagen (Luque et al., 2004). Sin embargo, hasta ahora, la mayoría de las estrategias de marketing de los destinos turísticos se centran en la imagen global, en lugar de distinguir entre las diferentes dimensiones que la forman (Ahmed, 1996).

En esta línea, se viene analizando el papel de la cultura como una importante fuente de diferenciación y, consecuentemente, un factor determinante en la mejora de la imagen de los destinos, ya sean éstos propiamente culturales o no (Castaño, Moreno y Crego, 2006; MacKay y Fesenmaier, 2000; Richardson y Crompton, 1988; Royo y Serarols, 2005; De la Calle, 2002; Montero Gutiérrez y Díaz, 2001).

3. NECESIDADES DE INVESTIGACIÓN Y METODOLOGÍA

3.1. Caso de estudio y objetivo de la investigación

Andalucía recibe más de 25 millones de turistas al año y es una de las regiones más conocidas y consolidadas en los mercados turísticos internacionales gracias, fundamentalmente,

a su oferta de turismo de litoral y a los iconos urbanos de Córdoba, Granada y Sevilla. Pero, además, de estos destinos más consolidados en el panorama internacional, la región cuenta con un número importante de ciudades de tamaño medio que, por su significación histórica, su actual importancia funcional y administrativa y, sobre todo, la disponibilidad de un extraordinario volumen de recursos (fundamentalmente patrimoniales) susceptibles de puesta en valor turístico, constituyen hoy un elemento básico para la estructuración y vertebración territorial andaluza y, por su dinamismo social y económico, además, pueden convertirse en los próximos años en un elemento clave para la diversificación de la oferta turística andaluza.

Se trata, en su mayoría, de ciudades con una amplia oferta cultural (museos, música, teatro, etc.) y gastronómica, que se une a un riquísimo patrimonio monumental, lo que ha permitido en algunos casos (Ronda, Carmona, Úbeda-Baeza) organizar una oferta turística dirigida al segmento de turismo cultural, que empieza a ser muy reconocida en los mercados y muy bien valorada por su calidad y excepcionalidad. El resto de ciudades están aún en una etapa muy inicial del desarrollo turístico, aunque cuentan con un amplio potencial, lo que permite vislumbrar un futuro lleno de oportunidades en el segmento del turismo cultural. Pero, para ello, es necesario "repensar" el rol turístico que pueden jugar estas ciudades, definiendo nuevas estrategias de desarrollo turístico y nuevos modelos de gestión.

En este contexto, la definición de una imagen de marca fuerte, coherente, diferenciada y reconocible favorecerá, sin duda, el interés de los potenciales turistas por estos destinos, convirtiéndose, por tanto, en un factor clave para los gestores turísticos de estas ciudades, pues su habilidad para captar clientes dependerá de la capacidad del destino para satisfacer sus expectativas y necesidades, para lo que es imprescindible conocer la valoración que los actuales turistas dan a aspectos básicos de la imagen turística de estas ciudades, que influye, a su vez, en su decisión de visitar el destino y en su comportamiento durante la estancia.

Por tanto, la hipótesis que sustenta esta investigación es que existe un comportamiento diferencial entre los turistas con una marcada vocación cultural y el resto respecto a la valoración de aspectos clave de la imagen de estas ciudades como destino turístico. A partir de la formulación de esta hipótesis, el objetivo que se persigue es determinar si los turistas culturales y el resto de turistas que visitan las ciudades medias del interior de Andalucía tienen una imagen idéntica de estas ciudades, o, bien, por el contrario, esta marcada vocación cultural condiciona una imagen de estos destinos diferente a la que tienen el resto de turistas.

Para ello, utilizando técnicas de análisis de clase latente, se ha procedido a segmentar esta demanda, de manera que sea posible disponer de grupos con comportamiento homogéneo, a fin de facilitar la toma de decisiones de los gestores turísticos en sus estrategias de posicionamiento en los mercados. La novedad de este trabajo no está tanto en los segmentos a los que ha dado lugar el proceso desarrollado, cuanto en la metodología utilizada para realizar la segmentación, que, a diferencia de otras técnicas, como el análisis cluster, permite no sólo delimitar y cuantificar los diferentes segmentos, sino conocer la probabilidad de que cualquier individuo de un determinado grupo tenga comportamientos propios del turismo cultural.

El estudio se ha estructurado a partir de 1.822 encuestas válidas realizadas a turistas que han visitado alguna de las ciudades medias del interior de Andalucía analizadas: Alcalá la Real, Antequera, Arcos de la Frontera, Baeza, Carmona, Écija, Estepa, Loja, Linares, Lucena, Osuna, Priego de Córdoba, Ronda, Úbeda.

3.2 Aspectos metodológicos. El análisis de clases latentes simultáneo

El análisis de clases latentes simultáneo es una extensión del modelo de análisis de clases latentes, en el que se identifica una variable latente a partir de un conjunto de indicadores de naturaleza categórica. A partir de la hipótesis de independencia local, y considerando una variable latente X con un total de T clases o categorías, el modelo de análisis de clases latentes estima las probabilidades de clase latente y las probabilidades condicionadas. Las primeras determinan el tamaño relativo de cada clase latente, mientras que las segundas permiten caracterizar, en términos de probabilidad, cada una de las clases identificadas. Para una amplia descripción del modelo de análisis de clases latentes, puede consultarse McCutcheon (1987), Hagenaars (1993), Von Eye y Clogg (1994), Heinen (1996) o Hagenaars y McCutcheon (2002).

El análisis de clases latentes simultáneo, desarrollado por Clogg y Goodman (1985), permite comparar las estructuras latentes de varios grupos o poblaciones. Se trata, por tanto, de una técnica de análisis comparativo válida cuando se estudia un mismo fenómeno mediante el mismo conjunto de indicadores en dos o más poblaciones. Es, además, una técnica muy útil para el análisis de tendencias de fenómenos latentes cuando los grupos estudiados son muestras independientes de la misma población tomadas en distintos períodos de tiempo. Ejemplos de esta última utilidad del análisis de clases latentes simultáneo pueden encontrarse en McCutcheon (1986) o en Dayton (1998).

En el modelo de análisis de clases latentes simultáneo se pueden considerar diferentes tipos de estructuras latentes:

a) Estructuras latentes heterogéneas: son aquellas que surgen cuando el número de clases latentes de cada población es el mismo, pero no se imponen restricciones sobre ninguno de los parámetros. En este caso, el modelo de estructura latente recibe el nombre de **modelo no restringido heterogéneo**.

b) Estructuras latentes parcialmente homogéneas: son aquellas en las que se imponen determinadas restricciones de igualdad sobre las probabilidades condicionadas o de clases latentes de las diferentes poblaciones analizadas. El modelo que recoge este tipo de estructuras latentes se denomina **modelo restringido parcialmente homogéneo**.

c) Estructuras latentes totalmente homogéneas: estas estructuras latentes aparecen cuando se imponen restricciones de igualdad sobre la totalidad de parámetros del modelo. Se dice entonces que las estructuras latentes de las poblaciones comparadas son idénticas. El modelo que permite contrastar esta hipótesis se conoce con el nombre de **modelo restringido de homogeneidad completa**.

Pues bien, dados cuatro indicadores discretos de la variable latente X, que denominaremos A, B, C y D, respectivamente, con *i, j*, *k* y *l* categorías, respectivamente, y considerando una **variable grupal** G con *g* grupos o poblaciones, la expresión del modelo de análisis de clases latentes simultáneo es la siguiente:

$$\Pi_{ijkl\,gt}^{\overline{A}\overline{B}\overline{C}\overline{D}G\overline{X}} = \Pi_{g}^{G} \times \Pi_{igt}^{\overline{A}GX} \times \Pi_{jgt}^{\overline{B}GX} \times \Pi_{kgt}^{\overline{C}GX} \times \Pi_{l\,gt}^{\overline{D}GX} \times \Pi_{gt}^{G\overline{X}} \qquad [1]$$

donde $\Pi_{ijkl\,gt}^{\overline{A}\overline{B}\overline{C}\overline{D}G\overline{X}}$ es la probabilidad de que un individuo perteneciente al grupo *g* responda a los indicadores A, B, C y D en las categorías *i, j*, *k* y *l*, respectivamente, y pertenezca a la clase *t* de la variable latente X; Π_{g}^{G} es la probabilidad de pertenencia a cada población considerada; $\Pi_{igt}^{\overline{A}GX}$, $\Pi_{jgt}^{\overline{B}GX}$, $\Pi_{kgt}^{\overline{C}GX}$ y $\Pi_{l\,gt}^{\overline{D}GX}$ son las probabilidades condicionadas de que individuo del grupo *g* y de la clase latente *t* responda a los indicadores A, B, C y D en una categoría determinada de los mismos; y $\Pi_{gt}^{G\overline{X}}$ es la probabilidad de clase latente de que un individuo del grupo *g* pertenezca a la clase latente *t*.

Dada su naturaleza probabilística, los parámetros del anterior modelo deben verificar lo siguiente:

$$\sum_{g=1}^{G} \Pi_{g}^{G} = 1 \qquad [2]$$

$$\sum_{i=1}^{I} \Pi_{igt}^{\overline{A}GX} = \sum_{j=1}^{J} \Pi_{jgt}^{\overline{B}GX} = \sum_{k=1}^{K} \Pi_{kgt}^{\overline{C}GX} = \sum_{l=1}^{L} \Pi_{l\,gt}^{\overline{D}GX} = 1 \qquad [3]$$

$$\sum_{t=1}^{T} \Pi_{gt}^{G\overline{X}} = 1 \qquad [4]$$

El modelo [1] incluye, por tanto, G(T-1) probabilidades de clase latente y GT(I-1)+GT(J-1)+GT(K-1)+GT(L-1) probabilidades condicionadas. Por tanto, el número total de parámetro del modelo [1] es de G[(I+J+K+L-3)T-1]. La estimación de estos parámetros se realiza por el método de la máxima verosimilitud, siendo posible emplear tanto el algoritmo EM (McCutcheon, 1987, pp. 23-27) como el método de Newton-Raphson (Agresti, 1990; Hagenaars, 1990; Hagenaars, 1993).

Finalmente, para verificar la bondad de ajuste de este modelo, se suele utilizar el test estadístico de la razón de verosimilitud, a través del cual se comparan las frecuencias observadas (n_{ijkl}) con las correspondientes frecuencias estimadas ($\hat{m}_{ijkl}$) mediante la siguiente expresión, que sigue una distribución chi-cuadrado asintótica:

$$G^2 = 2\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L} n_{ijkl} \log\frac{n_{ijkl}}{\hat{m}_{ijkl}} \qquad [5]$$

El modelo no restringido heterogéneo [1] es un modelo de naturaleza **exploratoria**, puesto que no se han impuesto restricciones de igualdad (o de valor específico, según el caso) sobre los parámetros del mismo. Pero, una de las grandes potencialidades del análisis de clases latentes es la posibilidad de contrastar determinadas hipótesis relativas a la naturaleza de la variable latente mediante la imposición de restricciones sobre los parámetros del modelo. Esta contrastación de hipótesis pertenece al ámbito de lo que se denomina **análisis de clases latentes confirmatorio**.

En el modelo [1] se pueden imponer restricciones intragrupales como las siguientes:

- Todas las clases latentes de la población *g* poseen el mismo tamaño: $\Pi_{g1}^{G\overline{X}} = \Pi_{g2}^{G\overline{X}} = \cdots = \Pi_{gT}^{G\overline{X}} = \frac{1}{T}$.

- La probabilidad condicionada de situarse en la categoría 2 del indicador A es la misma en todas las clases latentes de la población *g*: $\Pi_{2g1}^{\overline{A}GX} = \Pi_{2g2}^{\overline{A}GX} = \cdots = \Pi_{2gT}^{\overline{A}GX}$.

- La probabilidad de respuesta de los individuos de la tercera clase latente de la población *g* es la misma para todas las categorías del indicador B: $\Pi_{1g3}^{\overline{B}GX} = \Pi_{2g3}^{\overline{B}GX} = \cdots = \Pi_{Jg3}^{\overline{B}GX}$.

Sin embargo, el análisis confirmatorio cobra especial sentido cuando se imponen restricciones intergrupales (también denominadas **restricciones de homogeneidad**), puesto que dichas restricciones permiten contrastar si las estructuras latentes son homogéneas en las distintas poblaciones consideradas.

En este sentido, se puede proponer un modelo de clases latentes simultáneo con homogeneidad parcial en las probabilidades condicionadas del indicador A mediante la imposición de la siguiente restricción de homogeneidad:

$$\Pi_{i1t}^{\overline{A}GX} = \Pi_{i2t}^{\overline{A}GX} = \cdots = \Pi_{iGt}^{\overline{A}GX} \text{ para } i=1, 2, \ldots, I;\ t=1, 2, \ldots, T \qquad [6]$$

La anterior restricción asume que la distribución del indicador A no depende de las diferentes poblaciones consideradas. Esta misma restricción de homogeneidad parcial puede imponerse sobre la distribución de los indicadores B, C y D.

Otra posibilidad que ofrece el análisis de clases latentes confirmatorio es contrastar que el tamaño de las clases latentes es el mismo en todas las poblaciones analizadas mediante la siguiente restricción sobre las probabilidades de clases latentes:

$$\Pi_{1t}^{G\overline{X}} = \Pi_{2t}^{G\overline{X}} = \cdots = \Pi_{Gt}^{G\overline{X}} \text{ para } t=1, 2, \ldots, T \qquad [7]$$

Los modelos que surgen tras la imposición de las restricciones [6] o [7] se conocen con el nombre de modelos restringidos parcialmente homogéneos.

También es posible considerar un modelo restringido de homogeneidad completa cuando se imponen simultáneamente todas las restricciones de homogeneidad en probabilidades condicionadas [6] a todos los indicadores y todas las restricciones de homogeneidad en probabilidades de clases latentes [7]. Este modelo presupone que las clases latentes tienen el mismo tamaño en todas las poblaciones consideradas y que la distribución de los indicadores es independiente de la población.

Por último, para contrastar si las restricciones impuestas se verifican o no, se compara la bondad de ajuste del modelo restringido (M_R) con la bondad de ajuste del modelo no restringido (M_0) mediante la siguiente expresión:

$$G^2(M_R/M_0) = G^2(M_R) - G^2(M_0) \qquad [8]$$

Esta diferencia de verosimilitudes sigue una distribución chi-cuadrado asintótica con un número de grados de libertad igual a la diferencia entre los grados de libertad de los modelos M_R y M_0. Si la significación de este test es inferior a α (normalmente, el 5%), se podrá

concluir que las restricciones impuestas contribuyen de forma significativa a la mejora de la bondad de ajuste del modelo no restringido.

4. RESULTADOS DEL ANÁLISIS DE LA HOMOGENEIDAD DE LA IMAGEN TURÍSTICA EN LAS CIUDADES MEDIAS DEL INTERIOR DE ANDALUCÍA

Como ya se expuso con anterioridad, se pretende determinar hasta qué punto la imagen de los destinos turísticos que conforman las ciudades medias del interior de Andalucía es homogénea en los dos segmentos de demanda considerados (los turistas de marcada vocación cultural y el resto). Para ello, se ha definido una variable grupal G, con dos categorías: turistas culturales (categoría 1) y resto de turistas (categoría 2).

Por otro lado, y suponiendo el carácter latente de la imagen de un destino turístico, se han considerado los siguientes indicadores de la misma:

S: valoración de la señalización del destino turístico.
R: valoración de la relación precio-calidad del destino.
L: valoración de la limpieza en las ciudades medias del interior de Andalucía.
T: valoración de la tranquilidad en el destino.
A: valoración de la atención y del trato recibido durante la visita al destino.

En todos los casos, los indicadores considerados son de naturaleza dicotómica, de forma que la categoría 1 representa una valoración muy buena del destino (8 o más puntos en una escala de 1 a 10), mientras que la categoría 2 representa una valoración media o mala (7 puntos o menos en una escala de 1 a 10).

A partir de una muestra de 1.822 encuestas realizadas a turistas que visitaron las ciudades medias del interior de Andalucía, las frecuencias observadas para las 32 modalidades de respuesta posibles en las dos poblaciones consideradas se muestran en la Tabla 1.

Tabla 1. Frecuencias observadas de los indicadores S, R, L, T y A en el grupo de turistas culturales y en el grupo del resto de turistas

Respuestas	Frecuencias		Respuestas	Frecuencias	
S R L T A	Culturales	Resto	**S R L T A**	Culturales	Resto
1 1 1 1 1	159	53	2 1 1 1 1	41	19
1 1 1 1 2	42	15	2 1 1 1 2	20	12
1 1 1 2 1	30	11	2 1 1 2 1	20	12
1 1 1 2 2	20	3	2 1 1 2 2	10	4
1 1 2 1 1	26	39	2 1 2 1 1	31	29
1 1 2 1 2	7	10	2 1 2 1 2	3	9
1 1 2 2 1	41	24	2 1 2 2 1	56	51
1 1 2 2 2	16	4	2 1 2 2 2	24	26
1 2 1 1 1	70	9	2 2 1 1 1	51	21
1 2 1 1 2	29	12	2 2 1 1 2	25	11
1 2 1 2 1	25	10	2 2 1 2 1	35	20
1 2 1 2 2	20	5	2 2 1 2 2	39	17
1 2 2 1 1	21	7	2 2 2 1 1	31	11
1 2 2 1 2	11	5	2 2 2 1 2	29	12
1 2 2 2 1	21	18	2 2 2 2 1	74	61
1 2 2 2 2	12	7	2 2 2 2 2	155	81
Muestra global de turistas culturales: 1.194 Muestra global del resto de turistas: 628					

Fuente: Elaboración propia.

El primer objetivo del análisis consiste en determinar si los 5 indicadores considerados son estadísticamente independientes, puesto que, si se da esta situación, no tiene sentido plantear un modelo de análisis de clases latentes simultáneo para profundizar en las relaciones entre dichos indicadores, pues dicho modelo parte de la existencia de independencia local, lo que significa la presencia de asociación estadística entre los indicadores, la cual está motivada por la relación existente entre cada uno de ellos y la variable latente "imagen del destino turístico". Pues bien, para determinar la independencia o dependencia entre los indicadores utilizados se ajustará un modelo de clases latentes con una única clase, lo que equivale a aceptar la independencia entre

S, R, L, T y A. Como puede observarse en la Tabla 2, el modelo de independencia arroja un ajuste muy malo a los datos observados (G^2=958,1183), lo que conduce claramente a su rechazo, puesto que la significación estadística asociada al modelo es claramente inferior a 0,05 (nivel de significación que se considerará por defecto).

Una vez comprobada la existencia de asociación estadística entre los 5 indicadores, el siguiente paso consiste en determinar el número óptimo de clases latentes de la variable X. Teniendo en cuenta el número de poblaciones consideradas (2), el número de indicadores (5) y el carácter dicotómico de estos últimos, el número de clases latentes (T) deberá estar comprendido entre 2 y 5. En la Tabla 2 se muestra el ajuste de los modelos de clases latentes simultáneo con 2, 3, 4 y 5 clases, respectivamente. Como puede observarse, los modelos con 2 y 3 clases latentes proporcionan un ajuste insatisfactorio a los datos observados. Por el contrario, los modelos con 4 y 5 clases latentes proporcionar un ajuste adecuado a las frecuencias de la Tabla 1, aunque, de los dos, el que arroja un ajuste ligeramente superior es el modelo con 5 clases latentes. Por consiguiente, las restricciones de homogeneidad que se presentan a continuación tendrán como base el modelo con 5 clases latentes (modelo no restringido heterogéneo), al que se denominará en adelante M_0.

Tabla 2. Bondad de ajuste de modelos de análisis de clases latentes simultáneo a las frecuencias observadas en las Ciudades Medias de Andalucía

Modelo	G^2	g.l.	Signif.
- Modelo de independencia (T=1)	958,1183	52	0,0000
- Modelo con T=2 clases latentes	181,0956	40	0,0000
- Modelo con T=3 clases latentes	72,9613	28	0,0000
- Modelo con T=4 clases latentes	24,5995	16	0,0772
- Modelo con T=5 clases latentes	8,3250	4	0,0804

Fuente: Elaboración propia a partir de cálculos realizados con LEM[1].

El análisis de clases latentes confirmatorio consiste en la imposición de restricciones de homogeneidad en las probabilidades condicionadas y de clase latente del modelo M_0.

Comenzando por las restricciones de homogeneidad parcial en las probabilidades condicionadas de los 5 indicadores (Tabla 3), se puede apreciar que existe homogeneidad en la valoración de la señalización (S) y de la atención y trato recibido (A) entre los turistas culturales y el resto del turista. Por el contrario, estas dos poblaciones son heterogéneas en la valoración que otorgan a la relación calidad-precio (R), a la limpieza (L) y a la tranquilidad (T).

Por su parte, la Tabla 4 presenta los contrastes de homogeneidad parcial en probabilidades de clases latentes y de homogeneidad completa. Como puede observarse, la hipótesis de homogeneidad en clases latentes puede ser aceptada a un nivel de significación de un 5%, lo cual significa que el tamaño relativo de las cinco clases identificadas es el mismo en la población de turistas culturales que en la población del resto de turistas. Por otro lado, y aunque el hecho de que no todas las probabilidades sean homogéneas ya desvela que no existe homogeneidad completa entre las dos poblaciones consideradas, se ha contrastado también si el modelo restringido de homogeneidad completa proporciona un ajuste satisfactorio a los datos observados. Como ponen de manifiesto los resultados de la Tabla 4, este modelo debe ser rechazado a un nivel de significación del 5%, lo que equivale a admitir que la valoración de la imagen de las ciudades medias del interior de Andalucía por parte de los turistas culturales y del resto de turistas no es idéntica. Será necesario, por tanto, indagar en las diferencias más significativas existentes entre ambos segmentos.

De esta forma, el modelo de análisis de clases latentes simultáneo final presenta las siguientes características:

a) Homogeneidad parcial en las probabilidades condicionadas de los indicadores S y A.
b) Heterogeneidad en las probabilidades condicionadas de los indicadores L, R y T.
c) Homogeneidad parcial en las probabilidades de clase latente.
d) Ausencia de homogeneidad completa.

[1] Latent class models were estimated using the free computer program LEM (Vermunt, 1997).

Tabla 3. Modelos con restricciones de homogeneidad parcial en las probabilidades condicionadas de los 5 indicadores utilizados

M_S : modelo de homogeneidad parcial en las probabilidades condicionadas del indicador S. $\Pi_{111}^{\bar{S}GX} = \Pi_{121}^{\bar{S}GX}$; $\Pi_{112}^{\bar{S}GX} = \Pi_{122}^{\bar{S}GX}$; $\Pi_{113}^{\bar{S}GX} = \Pi_{123}^{\bar{S}GX}$; $\Pi_{114}^{\bar{S}GX} = \Pi_{124}^{\bar{S}GX}$; $\Pi_{115}^{\bar{S}GX} = \Pi_{125}^{\bar{S}GX}$ - Diferencia en verosimilitudes: $G^2(M_S / M_0) =$ 5,1044 - Diferencia en grados de libertad: 5 - *p*=0,4033 $\Rightarrow$ No rechazo (aceptación) de la homogeneidad parcial.
M_R : modelo de homogeneidad parcial en las probabilidades condicionadas del indicador R. $\Pi_{111}^{\bar{R}GX} = \Pi_{121}^{\bar{R}GX}$; $\Pi_{112}^{\bar{R}GX} = \Pi_{122}^{\bar{R}GX}$; $\Pi_{113}^{\bar{R}GX} = \Pi_{123}^{\bar{R}GX}$; $\Pi_{114}^{\bar{R}GX} = \Pi_{124}^{\bar{R}GX}$; $\Pi_{115}^{\bar{R}GX} = \Pi_{125}^{\bar{R}GX}$ - Diferencia en verosimilitudes: $G^2(M_R / M_0) =$ 17,6384 - Diferencia en grados de libertad: 5 - *p*=0,0034 $\Rightarrow$ Rechazo de la homogeneidad parcial.
M_L : modelo de homogeneidad parcial en las probabilidades condicionadas del indicador L. $\Pi_{111}^{\bar{L}GX} = \Pi_{121}^{\bar{L}GX}$; $\Pi_{112}^{\bar{L}GX} = \Pi_{122}^{\bar{L}GX}$; $\Pi_{113}^{\bar{L}GX} = \Pi_{123}^{\bar{L}GX}$; $\Pi_{114}^{\bar{L}GX} = \Pi_{124}^{\bar{L}GX}$; $\Pi_{115}^{\bar{L}GX} = \Pi_{125}^{\bar{L}GX}$ - Diferencia en verosimilitudes: $G^2(M_L / M_0) =$ 16,829 - Diferencia en grados de libertad: 5 - *p*=0,0048 $\Rightarrow$ Rechazo de la homogeneidad parcial.
M_T : modelo de homogeneidad parcial en las probabilidades condicionadas del indicador T. $\Pi_{111}^{\bar{T}GX} = \Pi_{121}^{\bar{T}GX}$; $\Pi_{112}^{\bar{T}GX} = \Pi_{122}^{\bar{T}GX}$; $\Pi_{113}^{\bar{T}GX} = \Pi_{123}^{\bar{T}GX}$; $\Pi_{114}^{\bar{T}GX} = \Pi_{124}^{\bar{T}GX}$; $\Pi_{115}^{\bar{T}GX} = \Pi_{125}^{\bar{T}GX}$ - Diferencia en verosimilitudes: $G^2(M_T / M_0) =$ 16,5565 - Diferencia en grados de libertad: 5 - *p*=0,0054 $\Rightarrow$ Rechazo de la homogeneidad parcial.
M_A : modelo de homogeneidad parcial en las probabilidades condicionadas del indicador A. $\Pi_{111}^{\bar{A}GX} = \Pi_{121}^{\bar{A}GX}$; $\Pi_{112}^{\bar{A}GX} = \Pi_{122}^{\bar{A}GX}$; $\Pi_{113}^{\bar{A}GX} = \Pi_{123}^{\bar{A}GX}$; $\Pi_{114}^{\bar{A}GX} = \Pi_{124}^{\bar{A}GX}$; $\Pi_{115}^{\bar{A}GX} = \Pi_{125}^{\bar{A}GX}$ - Diferencia en verosimilitudes: $G^2(M_A / M_0) =$ 2,3895 - Diferencia en grados de libertad: 5 - *p*=0,7930 $\Rightarrow$ No rechazo (aceptación) de la homogeneidad parcial.

Fuente: Elaboración propia a partir de cálculos realizados con LEM.

Tabla 4. Modelo con restricciones de homogeneidad parcial en las probabilidades de clase latente y modelo de homogeneidad completa

M_{LC} : modelo de homogeneidad parcial en las probabilidades de clase latente.
$\Pi_{11}^{G\bar{X}} = \Pi_{21}^{G\bar{X}}$; $\Pi_{12}^{G\bar{X}} = \Pi_{22}^{G\bar{X}}$; $\Pi_{13}^{G\bar{X}} = \Pi_{23}^{G\bar{X}}$; $\Pi_{14}^{G\bar{X}} = \Pi_{24}^{G\bar{X}}$; $\Pi_{15}^{G\bar{X}} = \Pi_{25}^{G\bar{X}}$
- Diferencia en verosimilitudes: $G^2(M_{LC}/M_0) =$ 0,6404
- Diferencia en grados de libertad: 4
- p=0,9585 $\Rightarrow$ No rechazo (aceptación) de la homogeneidad parcial en clases latentes.
M_T : modelo de homogeneidad completa.
$\Pi_{111}^{\bar{S}GX} = \Pi_{121}^{\bar{S}GX}$; $\Pi_{112}^{\bar{S}GX} = \Pi_{122}^{\bar{S}GX}$; $\Pi_{113}^{\bar{S}GX} = \Pi_{123}^{\bar{S}GX}$; $\Pi_{114}^{\bar{S}GX} = \Pi_{124}^{\bar{S}GX}$; $\Pi_{115}^{\bar{S}GX} = \Pi_{125}^{\bar{S}GX}$
$\Pi_{111}^{\bar{R}GX} = \Pi_{121}^{\bar{R}GX}$; $\Pi_{112}^{\bar{R}GX} = \Pi_{122}^{\bar{R}GX}$; $\Pi_{113}^{\bar{R}GX} = \Pi_{123}^{\bar{R}GX}$; $\Pi_{114}^{\bar{R}GX} = \Pi_{124}^{\bar{R}GX}$; $\Pi_{115}^{\bar{R}GX} = \Pi_{125}^{\bar{R}GX}$
$\Pi_{111}^{\bar{L}GX} = \Pi_{121}^{\bar{L}GX}$; $\Pi_{112}^{\bar{L}GX} = \Pi_{122}^{\bar{L}GX}$; $\Pi_{113}^{\bar{L}GX} = \Pi_{123}^{\bar{L}GX}$; $\Pi_{114}^{\bar{L}GX} = \Pi_{124}^{\bar{L}GX}$; $\Pi_{115}^{\bar{L}GX} = \Pi_{125}^{\bar{L}GX}$
$\Pi_{111}^{\bar{T}GX} = \Pi_{121}^{\bar{T}GX}$; $\Pi_{112}^{\bar{T}GX} = \Pi_{122}^{\bar{T}GX}$; $\Pi_{113}^{\bar{T}GX} = \Pi_{123}^{\bar{T}GX}$; $\Pi_{114}^{\bar{T}GX} = \Pi_{124}^{\bar{T}GX}$; $\Pi_{115}^{\bar{T}GX} = \Pi_{125}^{\bar{T}GX}$
$\Pi_{111}^{\bar{A}GX} = \Pi_{121}^{\bar{A}GX}$; $\Pi_{112}^{\bar{A}GX} = \Pi_{122}^{\bar{A}GX}$; $\Pi_{113}^{\bar{A}GX} = \Pi_{123}^{\bar{A}GX}$; $\Pi_{114}^{\bar{A}GX} = \Pi_{124}^{\bar{A}GX}$; $\Pi_{115}^{\bar{A}GX} = \Pi_{125}^{\bar{A}GX}$
$\Pi_{11}^{G\bar{X}} = \Pi_{21}^{G\bar{X}}$; $\Pi_{12}^{G\bar{X}} = \Pi_{22}^{G\bar{X}}$; $\Pi_{13}^{G\bar{X}} = \Pi_{23}^{G\bar{X}}$; $\Pi_{14}^{G\bar{X}} = \Pi_{24}^{G\bar{X}}$; $\Pi_{15}^{G\bar{X}} = \Pi_{25}^{G\bar{X}}$
- Diferencia en verosimilitudes: $G^2(M_T/M_0) =$ 103,7347
- Diferencia en grados de libertad: 29
- p=0,0000 $\Rightarrow$ Rechazo de la homogeneidad completa.

Fuente: Elaboración propia a partir de cálculos realizados con LEM.

Las estimaciones máximo-verosímiles del modelo final se muestran en la Tabla 5. Un análisis detallado de estas estimaciones permite delimitar y caracterizar las cinco clases latentes de la siguiente forma:

Clase 1: **Demandantes de calidad** (27,70% del total). Este primer segmento tiene, en general, unas probabilidades altas de otorgar una valoración óptima a los indicadores de imagen turística considerados, con la única excepción de la relación calidad-precio de las ciudades medias del interior de Andalucía, para la que los turistas de este segmento tienen una probabilidad bastante reducida (superior sólo a las probabilidades condicionadas de la clase 4) de sentirse altamente satisfechos con este aspecto concreto. Por tanto, parece claro que la imagen del destino turístico pasa, para este grupo de turistas, por una mejora de la calidad y/o por una reducción de los precios.

Clase 2: **Reclamantes de sensibilización turística (resto)** (20,30%). Los turistas de este segmento muestran un alto grado de satisfacción con la relación calidad-precio y con la atención y el trato recibido. Sin embargo, es poco probable que otorguen una valoración excelente a la señalización turística. Por otro lado, aunque a la vista de sus probabilidades condicionadas, los turistas culturales tienen, en alta consideración la limpieza y la tranquilidad de las ciudades medias del interior de Andalucía, no sucede lo mismo con el resto de turistas. Para esta segunda población, sólo un 16% de turistas valora con 8 o más puntos la limpieza, mientras que esta proporción se reduce a menos de un 3% cuando se trata de valorar la tranquilidad del destino. Parece evidente, por tanto, que los turistas que realizan actividades no estrictamente culturales en estos destinos están reclamando una mejor señalización turística, una mayor limpieza y, sobre todo, más tranquilidad y menos ruido.

Clase 3: **Reclamantes de sensibilización turística (culturales)** (12,13%). Las características de este tercera clase son muy similares a las de la anterior, con la única diferencia de que son los turistas culturales de este segmento los que demandan más señalización turística, aunque lo más llamativo es su valoración de la limpieza y de la tranquilidad. En efecto, la probabilidad de que los turistas culturales valoren con 8 o más puntos estos dos aspectos de la imagen del destino es nula. Por consiguiente, y aunque este segmento es el que menor peso tiene en la composición de la demanda turística de las ciudades medias del interior de Andalucía,

parece claro que deben tomarse algunas acciones que mejoren la imagen percibida por los turistas culturales de este segmento en lo que a señalización, limpieza y tranquilidad se refiere.

Clase 4: **Decepcionados** (23,21%). Los turistas de esta clase presentan probabilidades muy bajas de valorar de forma excelente los cinco aspectos de la imagen turística que han sido tenidos en cuenta. Son, con notable diferencia, los que tienen peor imagen de las ciudades medias del interior de Andalucía como destino turístico. Son, por tanto, el colectivo sobre el que hay que incidir de forma prioritaria para mejorar la imagen de este destino turístico.

Clase 5: **Encantados** (16,67%). Esta última clase latente está integrada por viajeros con una imagen turística de estas ciudades radicalmente opuesta a la de los turistas de la clase anterior. De hecho, presentan unas probabilidades muy elevadas (por encima del 60% en todos los casos, alcanzando el 100% para algún indicador) de conceder 8 o más puntos a los aspectos considerados.

Tabla 5. Bondad de ajuste y estimaciones de las probabilidades condicionadas y de las probabilidades de clase latente del modelo de estructura latente final

Test estadístico de razón de verosimilitud: G^2=18,5059
Grados de libertad: 18
Significación: p=0,4228

a) Probabilidades condicionadas:

a.1) Indicador S

$\Pi_{111}^{\bar{S}GX} = \Pi_{121}^{\bar{S}GX}$ =0,4823
$\Pi_{112}^{\bar{S}GX} = \Pi_{122}^{\bar{S}GX}$ =0,3781
$\Pi_{113}^{\bar{S}GX} = \Pi_{123}^{\bar{S}GX}$ =0,3706
$\Pi_{114}^{\bar{S}GX} = \Pi_{124}^{\bar{S}GX}$ =0,0311
$\Pi_{115}^{\bar{S}GX} = \Pi_{125}^{\bar{S}GX}$ =1,0000

a.2) Indicador R:

$\Pi_{111}^{\bar{R}GX}$ =0,4080	$\Pi_{121}^{\bar{R}GX}$ =0,2326
$\Pi_{112}^{\bar{R}GX}$ =0,5310	$\Pi_{121}^{\bar{R}GX}$ =0,6009
$\Pi_{113}^{\bar{R}GX}$ =0,7772	$\Pi_{121}^{\bar{R}GX}$ =0,9732
$\Pi_{114}^{\bar{R}GX}$ =0,0000	$\Pi_{121}^{\bar{R}GX}$ =0,2186
$\Pi_{115}^{\bar{R}GX}$ =0,8128	$\Pi_{121}^{\bar{R}GX}$ =1,0000

a.3) Indicador L:

$\Pi_{111}^{\bar{L}GX}$ =0,8004	$\Pi_{121}^{\bar{L}GX}$ =0,6439
$\Pi_{112}^{\bar{L}GX}$ =0,5476	$\Pi_{122}^{\bar{L}GX}$ =0,1679
$\Pi_{113}^{\bar{L}GX}$ =0,0000	$\Pi_{123}^{\bar{L}GX}$ =0,4078
$\Pi_{114}^{\bar{L}GX}$ =0,1348	$\Pi_{124}^{\bar{L}GX}$ =0,0882
$\Pi_{115}^{\bar{L}GX}$ =0,9528	$\Pi_{125}^{\bar{L}GX}$ =0,6232

a.4) Indicador T:

$\Pi_{111}^{\bar{T}GX}$ =0,5598	$\Pi_{121}^{\bar{T}GX}$ =0,6155
$\Pi_{112}^{\bar{T}GX}$ =0,7714	$\Pi_{122}^{\bar{T}GX}$ =0,0260
$\Pi_{113}^{\bar{T}GX}$ =0,0000	$\Pi_{123}^{\bar{T}GX}$ =1,0000
$\Pi_{114}^{\bar{T}GX}$ =0,1285	$\Pi_{124}^{\bar{T}GX}$ =0,0407
$\Pi_{115}^{\bar{T}GX}$ =0,9032	$\Pi_{125}^{\bar{T}GX}$ =0,8963

Tabla 5 (cont.)

Test estadístico de razón de verosimilitud: G^2=18,5059
Grados de libertad: 18
Significación: p=0,4228

a) Probabilidades condicionadas:

a.5) Indicador A:

$\Pi_{111}^{\bar{A}GX} = \Pi_{121}^{\bar{A}GX}$ =0,5112
$\Pi_{112}^{\bar{A}GX} = \Pi_{122}^{\bar{A}GX}$ =0,9256
$\Pi_{113}^{\bar{A}GX} = \Pi_{123}^{\bar{A}GX}$ =0,7216
$\Pi_{114}^{\bar{A}GX} = \Pi_{124}^{\bar{A}GX}$ =0,2653
$\Pi_{115}^{\bar{A}GX} = \Pi_{125}^{\bar{A}GX}$ =0,8232

b) Probabilidades de clase latente:

$\Pi_{11}^{G\bar{X}} = \Pi_{21}^{G\bar{X}}$ =0,2770
$\Pi_{12}^{G\bar{X}} = \Pi_{22}^{G\bar{X}}$ =0,2030
$\Pi_{13}^{G\bar{X}} = \Pi_{23}^{G\bar{X}}$ =0,1213
$\Pi_{14}^{G\bar{X}} = \Pi_{24}^{G\bar{X}}$ =0,2321
$\Pi_{15}^{G\bar{X}} = \Pi_{25}^{G\bar{X}}$ =0,1667

c) Probabilidades de pertenencia al grupo:

Π_{1}^{G} =0,6553
Π_{2}^{G} =0,3447

Fuente: Elaboración propia a partir de cálculos realizados con LEM.

Por otro lado, la comparación de las probabilidades condicionadas de ambas poblaciones en aquellos indicadores en los que se haya contrastado la ausencia de homogeneidad permite establecer comportamientos diferenciales entre los turistas culturales y el resto de turistas, que son, a su vez, la base para proponer medidas de gestión turística diferenciadas.

Así, entre los **demandantes de calidad**, mientras los turistas culturales tienen una probabilidad superior al 40% de valorar muy positivamente la relación calidad-precio del destino turístico, para el resto de turistas esta probabilidad es sólo de un 23%. Dicho de otro modo, los turistas culturales no son los más preocupados por los precios de los productos y servicios turísticos ofertados por las ciudades medias del interior de Andalucía y por la calidad asociada a los mismos, puesto que esta preocupación es mucho más patente entre los turistas no culturales.

En el segmento de los **reclamantes de sensibilización turística (resto)**, las diferencias de probabilidades condicionadas más notables se dan en los indicadores de limpieza y de tranquilidad. Así, de la misma manera que son mayoritarios los turistas culturales que se muestran altamente satisfechos por estos dos aspectos de la imagen turística del destino, sólo el 16% de los turistas no culturales valora de forma significativamente elevada la limpieza, mientras que no llega al 3% el porcentaje de turistas no culturales que consideran muy destacable la tranquilidad y la ausencia de contaminación acústica en estas ciudades andaluzas.

Por su parte, entre los **reclamantes de sensibilización turística (cultural)** se detecta un importante sentimiento de insatisfacción con la limpieza y con la tranquilidad entre los turistas culturales, hasta el punto de que ninguno de ellos valora con más de 7 puntos estos dos aspectos de la imagen. En la población de turistas no culturales, por el contrario, más de un 40% destaca como muy satisfactoria la limpieza, si bien lo más significativo (por contraste con la probabilidad asociada a los turistas culturales) es que la totalidad de turistas no culturales considera sobresaliente la tranquilidad y el bajo nivel de ruidos de este destino turístico. Es evidente, pues, que los reclamantes de sensibilización turística son, exclusivamente, los turistas con una evidente inclinación por el turismo cultural.

Las principales diferencias en las probabilidades condicionadas en el segmento de **decepcionados** se producen en la relación calidad-precio, en la que la insatisfacción es considerablemente superior entre los turistas culturales (0,0% de turistas altamente satisfechos) que entre el resto de turistas (21,9%). Por el contrario, entre los turistas culturales la probabilidad de valorar positivamente la limpieza y la tranquilidad es superior que entre el resto de turistas.

Finalmente, a pesar de la heterogeneidad detectada, las diferencias en las probabilidades condicionadas en el segmento de **encantados** no son muy importantes. Así, por ejemplo, turistas culturales y no culturales tienen prácticamente la misma probabilidad de manifestar su gran satisfacción por la tranquilidad (un 90%, aproximadamente, en ambos colectivos). La mayor heterogeneidad se produce en la valoración de la limpieza del destino turístico, indicador para el que se computa una probabilidad de más de un 95% de valorar con 8 o más puntos este aspecto entre los turistas culturales, pero que se reduce en más de 30 puntos porcentuales (62,32%) entre el resto de turistas.

5. CONCLUSIONES

El presente trabajo perseguía como principal objetivo determinar si los turistas culturales que visitan las ciudades medias del interior de Andalucía tienen una imagen idéntica de estas ciudades diferente a la que tienen el resto de turistas. Los resultados del trabajo empírico permiten confirmar que, en efecto, la marcada vocación cultural de estos turistas se traduce en un comportamiento diferencial respecto al resto en lo que respecta a la valoración de los cinco atributos analizados sobre la imagen de estas ciudades.

Por otra parte, las estimaciones máximo-verosímiles del modelo final han contribuido a la delimitación y caracterización de cinco segmentos (clases latentes) de demanda turística con comportamientos diferentes, lo que permitirá a los gestores turísticos de estas ciudades diseñar una política de gestión dirigida a satisfacer las necesidades y expectativas de cada uno de estos grupos, asegurando un trato más individualizado y que responda a necesidades concretas de cada grupo.

Los resultados obtenidos confirman, además, que la política de gestión turística de cualquier destino, especialmente la política de marketing, debe estar guiada por la imagen objetivo que se desea alcanzar y debe estar apoyada sobre el conocimiento de la imagen actual que sobre el destino poseen los turistas, que, a su vez, deberá contrastarse con la imagen de los destinos competidores, a fin de asegurar un mejor posicionamiento competitivo.

En lo que respecta a las implicaciones de esta investigación, en la práctica, como ya se ha señalado, los resultados obtenidos han permitido disponer de un conocimiento más detallado de los aspectos de la imagen más valorados por los turistas que visitan las ciudades medias del interior de Andalucía, lo que supone un avance significativo para el diseño e implementación de políticas de gestión turística en estos territorios, especialmente en lo que respecta al marketing, a fin de proyectar la imagen adecuada a los mercados objetivo. Desde el punto de vista académico, este trabajo contribuye en una doble vertiente. Por un lado, se suma a las investigaciones que se vienen realizando en los últimos años sobre la valoración de la imagen de los destinos turísticos, en este caso aportando herramientas de análisis para destinos culturales emergentes. Por otro lado, hay que destacar el interés de la técnica utilizada durante el proceso de segmentación. A diferencia del análisis cluster, los modelos de segmentación latentes son sumamente atractivos para quienes tienen que tomar decisiones de gestión en el ámbito del turismo, ya que, además de facilitar información sobre los diferentes segmentos que conforman su clientela, permiten ordenarlos en función de la probabilidad de que los individuos que los componen mantengan un determinado comportamiento (en este caso, con respecto a la valoración de un conjunto de atributos de la imagen del destino), lo que facilita no sólo la identificación, accesibilidad y cuantificación de los segmentos, sino la priorización en el diseño e implementación de estrategias específicas para cada grupo.

REFERENCIAS

AGRESTI, A. (1990): *Categorical Data Analysis.* New York: John Wiley.

AHMED, Z. U. (1996): "The need for the identification of the constituents of a destination's tourist image: a promotional segmentation perspective". *The Tourist Review*, 2, pp. 44-57.

ALHEMOUD, A. M. y ARMSTRONG, E. G. (1996): "Image of tourism attractions in Kuwait". *Journal of Travel Research*, 34 (4), pp. 76-80.

ALTÉS, C. (1993): *Marketing y turismo. Introducción al marketing de empresas y destinos turísticos*. Madrid: Síntesis.

ASHWORTH, G. y GOODALL, B. (1988): "Tourist images: marketing considerations". En Goodall, B. y ASHWORTH, G. (eds.): *Marketing in the tourism industry. The promotion of destination regions*. London: Routledge, pp. 213-238.

BALOGLU, S. y MANGALOGLU, M. (2001): "Tourism destinations images of Turkey, Egypt, Greece, and Italy as perceived by US–based tour operators and travel agents". *Tourism Management*, 22 (1), pp. 1-9.

BEERLI, A. y MARTÍN, J. D. (2004): "Tourists´ characteristic and the perceived image of tourist destinations: a quantitative analysis. A case study of Lanzarote, Spain". *Tourism Management*, 25 (5), pp. 623-636.

BIGNÉ, E. y SÁNCHEZ, I. (2001): "Evaluación de la imagen de destinos turísticos. Una aplicación metodológica en la Comunidad Valenciana". *Revista Europea de Dirección y Economía de la Empresa*, 10 (2), pp. 189-200.

BIGNÉ, E., SÁNCHEZ, M. I. y SÁNCHEZ, J. (2001): "Tourism image, evaluation variables and after purchase bahaviour: inter-relationship. *Tourism Management*, 22 (6), pp. 607-616.

BIGNÉ, E.; SÁNCHEZ, I. y CURRÁS, R. (2007): "El papel de la imagen del destino en la valoración y comportamiento postcompra del turista de sol y playa". *Papers de Turisme*, 42, pp. 57-73.

BIGNÉ, E.; SÁNCHEZ, I. y SANZ, S. (2005): "La imagen de un destino como variable de segmentación: un modelo de relaciones post-compra". *Actas de XVII Encuentro de Profesores Universitarios de Marketing*. Madrid.

BORDAS, E. y RUBIO, M. L. (1993): "La imagen turística de España: un modelo de gestión a largo plazo". *Información Comercial Española*, 722, pp. 107-118.

BRAMWELL, B. y RAWDING, L. (1996): "Tourism marketing images of industrial cities". *Annals of Tourism Research*, 23 (1), pp. 201-221.

BUHALIS, D. (2000): "Marketing the competitive destination of the future". *Tourism Management*, 21 (1), pp. 97-116.

CALANTONE, R. J., DI BENETTON, C. A., HAKAM, A. y BOJANIC, D. C. (1989): "Multiple multinational tourism positioning using correspondence analysis". *Journal of Travel Research*, 28 (2), pp. 25-32

CASTAÑO, J. M.; MORENO, A. y CREGO, A. (2006): "Factores psicosociales y formación de imágenes en el turismo urbano: un estudio de caso sobre Madrid". *PASOS Revista de Turismo y Patrimonio Cultural*, 4 (3), pp. 287-299.

CHOI, W. M.; CHAN, A. y WU, J. (1999): "A qualitative and quantitative assessment of Hong Kong's image as tourist destination". *Tourism Management*, 20(3), pp. 361-365.

CHON, K. S.; WEAVER, P. A. y KIM, C. Y. (1991): "Marketing you community: image analysis in Norfolk". *Cornell Hotel and Restaurant Administration Quarterly*, 31 (4), pp. 31-37.

CLOGG, C.C. y GOODMAN, L.A. (1985): "Simultaneous latent structure analysis in several groups". En Tuma, N. B. (ed.): *Sociological Methodology*. San Francisco: Josey-Bass.

CLOGG, C.C. y GOODMAN, L.A. (1986): "On scaling models applied to data from several groups". *Psychometrika* nº 51, pp. 123-135.

COOPER, C.; FLETCHER, J.; GILBERT, D. y WANHILL, S. (1993): *Tourism. Principles and Practice*. Pitman Publishing, UK.

COURT, B. y LUPTON, R. A. (1997): "Costumer portfolio development: modelling destination adopters, inactives and rejecters". *Journal of Travel Research*, 36 (1), pp. 35-43.

CROMPTON, J. L. (1979): "An assessment of the image of Mexico as a vacation destination and the influence of geographical location upon that image". *Journal of Travel Research*, 17 (4), pp. 18-23.

DADGOSTAR, B. e ISOSTALO, R. M. (1992): Factors affecting time spent by near-home tourists in city destinations". *Journal of Travel Research*, 31 (2), pp. 34-39.

DAYTON, C.M. (1998): *Latent class scaling analysis.* Sage University Paper series on Quantitative Applications in the Social Sciences, 07-126. Newbury Park, CA: Sage.

DE LA CALLE, M. (2002): *La ciudad histórica como destino turístico*. Barcelona: Ariel Turismo.

DOBNI, D. y ZINKHAN, G. M. (1990): "In search of brand image: a foundation analysis". *Advances in Consumer Research*, 17, pp. 110-119.

DOWLING, G. (1986): "Perceived risk: the concept and its measurement". *Psychology and Marketing*, 3 (3), pp. 193-210.

ECHTNER, C. M. y RITCHIE, B. (1991): "The meaning and measurement of destination image". *The Journal of Tourism Studies*, 2 (2), pp. 2-12.

ECHTNER, C. M. y RITCHIE, B. (1993): "The measurement of destination image: an empirical assessment". *Journal of Travel Research*, 31 (4), pp. 3-13.

FAKEYE, P. C. y CROMPTON, J. L. (1991): "Image differences between prospective, first-time, and repeat visitors to the Lower Rio Grande Valley". *Journal of Travel Research*, 30 (2), pp. 10-16.

FOXALL, G. R. y GOLDSMITH, R. E. (1994): *Consumer psychology for marketing*. London: Routledge.

GARTNER, W. C. (1986): "Temporal influence on image change". *Annals of Tourism Research*, 13 (4), pp. 635-644.

GARTNER, W. C. (1989): "Tourism image: attribute measurement of state tourism products using multidimensional scale techniques". *Journal of Travel Research*, 28 (2), pp. 16-20.

GOODALL, B. (1990): "The dynamics of tourism place marketing". En Ashworth, G. y Goodall, B. (eds.): *Marketing tourism places*. London: Routledge, pp. 259-279.

GOODRICH, J. N. (1978): "A new approach to image analysis through multidimensional scaling". *Journal of Travel Research*, 16, pp. 3-7.

GUNN, C. (1972): *Vacationscape*. Austin: Bureau of Business Research, University of Texas.

GUNN, C. (1988): *Vacationscapes: Designing tourist regions*. New York: Van Nostrand Reinhold.

HAGENAARS, J.A. (1990): *Categorical longitudinal data: log-linear panel, trend, and cohort analysis*. Newbury Park, CA: Sage.

HAGENAARS, J.A. (1993): *Loglinear models with latent variables*. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-094. Newbury Park, CA: Sage.

HAGENAARS, J.A. y McCUTCHEON, A.L. (2002): *Applied latent class analysis.* Cambridge University Press.

HAMPTON, R. D.; GUY, B. S. y SINKULA, J. M. (1987): "Consumer images of financial institutions: a comparative study of Banks versus savings and loans". *Journal of Professional Services Marketing*, 2, pp. 83-100.

HEINEN, T. (1996): *Latent class and discrete latent trait models: similarities and differences.* Advanced Quantitative Techniques in the Social Sciences, volume 6. Sage Publications, Inc. Thousand Oaks, California.

HSU, C. H. C.; WOLFE, K. y KANG, S. K. (2004): "Image assessment for a destination with limited comparative advantages". *Tourism Management*, 25 (1), pp. 121-126.

HUI, T. K. y WAN, T. W. D. (2003): "Singapore's image as a tourist destination". *International Journal of Tourism Research*, 5, pp. 305-313.

HUNT, J. D. (1971): *Image: A Factor of Tourism*. Unpublished Ph.D dissertation. Fort Collins: Colorado State University.

HUNT, J. D. (1975): "Image as factor in tourism development". *Journal of Travel Research*, 13 (3), pp. 1-7.

JOPPE, M.; MARTÍN, D. W. y WAALEN, J. (2001): "Toronto's image as a destination: a comparative importance-satisfaction analysis by origin of visitor". *Journal of Travel Research*, 39 (3), pp. 252-260.

KOTLER, P.; HAIDER, D. H. y REIN, I. (1993): *Marketing Lugares*. New York: Free Press.

LUQUE, T.; DEL BARRIO, S.; IBÁNEZ, J. A. y Rodríguez, M. A. (2004): "La modelización de la imagen de Granada desde la perspectiva de los líderes de opinión externos". *Actas XVI Encuentro de Profesores Universitarios de Marketing*. Alicante.

MacKAY, K. J. y FESENMAIER, D. R. (2000): "An exploration of cross-cultural destination image assessment". *Journal of Travel Research*, 38, pp. 417-423.

MAZANEC, J. A. (1994): "Consumer behaviour in tourism". En Witt, S. y Moutinho, L. (eds.): *Tourism Marketing and Management Handbook*. Prentice-Hall, pp. 63-68.

McCUTCHEON, A.L. (1986): "Sexual morality, pro-life values, and attitudes toward abortion: a simultaneous latent structure analysis for 1978-1983". Paper presented at the Annual Meeting of the Eastern Sociological Society, New York, NY, April 1986.

McCUTCHEON, A.L. (1987): *Latent class analysis.* Sage University Paper series on Quantitative Applications in the Social Sciences, 07-064. Newbury Park, CA: Sage.

MILMAN, A. y PIZAM, A. (1995): "The role of awareness and familiarity with a destination: the Central Florida case". *Journal of Travel Research*, 33 (3), pp. 21-27.

MONTERO, I.; GUTIÉRREZ, D. y DÍAZ, R. (2001): "La cultura como componente de la oferta de los destinos turísticos maduros", *Estudios Turísticos*, 150, pp. 41-55.

MOUTINHO, L. (1987): "Consumer behaviour in tourism". *European Journal of Marketing*, 21 (10), pp. 5-44.

NGUYEN, N. y LEBLANC, G. (2001): "Corporate image and corporate reputation in customer's retention decisions in services". *Journal of Retailing and Consumer Services*, 8, pp. 227-236.

O'LEARY, S. y DEEGAN, J. (2003): "People, pace, place: qualitative and quantitative images of Ireland as a tourism destination in France". *Journal of Vacation Marketing*, 9 (3), pp. 213-226.

PEARCE, P. L. (1982): "Perceived changes in holiday destinations". *Annals of Tourism Research*, 19 (2), pp. 145-164.

PHELPS, A. (1986): "Holiday destination image-the problem of assessment: and example developed in Menorca". *Tourism Management*, 7 (3), pp. 168-180.

PIKE, S. (2002): "Destination image analysis-a review of 142 papers from 1973 to 2000". *Tourism Management*, 23 (5), pp. 541-549.

REILLY, M. D. (1990): "Free elicitation of descriptive adjectives for tourism image assessment", *Journal of Travel Research*, 28, (4), pp. 21-26.

RICHARDSON, S. L. y CROMPTON, J. L. (1988): "Cultural variations in perceptions on vacations attributes", *Tourism Management*, 9 (2), pp. 128-136.

ROSS, G. E. (1993): "Residents perception of the impact of tourism on an Australian city". *Journal of Travel Research*, 30, pp. 13-17.

ROYO, M. y SERAROLS, Ch. (2005): "El turismo rural-cultural: un modelo de gestión del marketing turístico a nivel local basado en la medida de la imagen del destino". *Cuadernos de Turismo*, 16, pp. 197-222.

SAN MARTÍN, H.; RODRÍGUEZ DEL BOSQUE, I. A. y VÁZQUEZ, R. (2006): "Análisis de la imagen en turismo mediante técnicas estructuradas y no estructuradas: implicaciones competitivas para los destinos turísticos". *Revista Asturiana de Economía*, 35, pp. 69-91.

SÁNCHEZ, I. y SANZ, S. (2004): "El papel de la construcción turística en la construcción de la imagen de un destino". En Blanquer. D. (dir.): *Turismo cultural y urbano*. 6º Congreso de Turismo, Universidad y Empresa. Valencia: Tirant lo Blanch, pp. 463-480.

SEATON, A. V. y BENETT, M. M. (1996): *Marketing tourism products. Concepts, issues, cases*. London: International Thomson Business Press.

TAPACHAI, N. y WARYSZAK, R. (2000): "An examination of the role of beneficial image in tourist destination selection". *Journal of Travel Research*, 39 (1), pp. 37-44.

TELISMAN-KOSUTA, N. (1994): "Tourist destination image". En Witt, S. y Moutinho, L. (eds.): *Tourism Marketing and Management Handbook*. Prentice-Hall, pp. 557-561.

VALLS, J. F. (1992): *La imagen de marca de los países*. Madrid: McGraw-Hill.

VERMUNT, J.K. (1997): *LEM: a general program for the analysis of categorical data.* Tilburg University.

VON EYE, A. and CLOGG, C.C. (1994): *Latent variables analysis.* Sage Publications, Inc. Thousand Oaks, California.

WAITT, G. (1996): "Marketing Korea as an international tourist destination". *Tourism Management*, 17 (2), pp. 113-121.

WOODSIDE, A. G. y Lysonsky, S. (1989): "A general model of travel destination choice". *Journal of Travel Research*, 27 (4), pp. 8-14.

IS EUROPEAN ECONOMIC GROWTH HETEROGENEOUS ACROSS COUNTRIES? A LATENT MARKOV ANALYSIS

JOSÉ G. DIAS
Email: jose.dias@iscte.pt
ISCTE
University Institute of Lisbon

Abstract

This paper analyses the economic growth in Europe. We set a panel Markov switching model that allows the joint estimation of heterogeneous latent trajectories. We apply the model to data from 16 European countries in the period 1950 to 2003. Three important results are obtained: we found no heterogeneity of patterns across European countries, that is, all countries follow the same dynamic trajectory; four regimes or states describe best the temporal heterogeneity, which allow a good description, monitoring and identification of business cycles in each country; we found that there is synchronization between European economies. This type of mapping and monitoring of economic processes is especially relevant for policy decision making.

The Markov switching model has become a standard tool for modeling and characterizing business cycles in most countries. Following Hamilton (1989), the Markov switching regime model has become extensively used to study nonlinearities in economic growth. Nonlinearities are important since real output responds to a shock depends on which state the economy is in; i.e. an expansion or a contraction (see, e.g., Tan and Habibullah, 2007). The transition from one state of the cycle to another is modeled as a regime switch (a latent variable), and the probability of changing regime is inferred from the data. The most common specification assumes a two-regime model, generally named expansion and recession stages. However, some authors have shown that more regimes can further improve the specification. For example, Artis *et al.* (2004) assumed three states representing recession, high growth and normal growth. They estimate Markov switching models individually for each country in the analysis: Germany, France, Italy, The Netherlands, Austria, Belgium, Spain, Portugal, and the United Kingdom. It is unclear whether there is a European business cycle, and for example, whether and to what degree the United Kingdom is converging on the cycles of its European partners (Hallett and Richter, 2006).

We propose to analyse economic growth using a dynamic latent class model that takes into account unobserved heterogeneity by means of time-constant and time-varying discrete latent variables (Dias *et al.*, 2008). Here, this methodology is used to help understand and identify changes in the historical growth performance of 16 European countries in the period from 1950 to 2003.

The article is organized as follows: Section 1 introduces the econometric model developed in this research; Section 2 describes the data and some stylized facts of the European GDP time series. Section 3 reports and discusses the results. The paper concludes with a summary of the main findings.

Key Words: economic growth, Markov Switching models, European economies, unobserved heterogeneity.
Thematic Area: Quantitative Methods for Economics and Business.

1. MARKOV SWITCHING MODELS

We model simultaneously the economic time series of n countries. Let y_{it} represent the response of observation (country) i at time point t, where $i \in \{1, \ldots, n\}$, $t \in \{1, \ldots, T\}$, and $y_{it} \in \Re$. In addition to the observed "response" variable y_{it}, the model contains two different latent variables: a time-constant discrete latent variable (latent class) and a time-varying discrete latent variable (regime). The former, which is denoted by $w \in \{1, \ldots, S\}$, is used to capture the unobserved heterogeneity across countries; that is, countries are clustered based on differences in their dynamics. The K -regime time-varying latent variable is denoted by $z_t \in \{1, \ldots, K\}$. Changes between the K states or regimes between adjacent time points are assumed to be in agreement with a first-order Markov or first-order autocorrelation structure.

Let $f(y_i;\varphi)$ be the (probability) density function associated with pattern of country i, where φ is the vector of parameters in the model. The density of the model is:

$$f(y_i;\varphi)=\sum_{w=1}^{S}\sum_{z_1=1}^{K}\cdots\sum_{z_T=1}^{K}f(w)f(z_1|w)\prod_{t=2}^{T}f(z_t|z_{t-1},w)\prod_{t=1}^{T}f(y_{it}|z_t). \qquad (1)$$

As in any mixture model, the observed data density $f(y_i;\varphi)$ is obtained by marginalizing over the latent variables. Because in our model these are discrete variables, this simply involves the computation of a weighted average of class-specific probability densities where the (prior) class membership probabilities or mixture proportions serve as weights (McLachlan and Peel, 2000). We assume that within latent class or cluster w, the sequence $z_1,\ldots,z_T$ is in agreement with a first-order Markov chain. Moreover, we assume that the observed return at a particular time point depends only on the regime at this time point; i.e, conditionally on the latent state z_t, the response y_{it} is independent of returns at other time points, which is often referred to as the local independence assumption. As far as the first-order Markov assumption for the latent regime switching conditional on cluster membership w is concerned, it is important to note that this assumption is not as restrictive as one may initially think. It does clearly not imply a first-order Markov structure for the responses y_{it}. The hidden Markov model and the Markov switching model (Baum *et al.*, 1970; Hamilton, 1989) are special cases of this model that is obtained by eliminating the time-constant latent variable w from the model, that is, by assuming that there is no unobserved heterogeneity. On the other hand, removing the time-varying latent variables the model reduces to a finite mixture model (McLachlan and Peel, 2000).

The statistical model depicted in Equation (1) is characterized by the following four functions:

1. $f(w)$ is the prior probability of belonging to a particular latent class or cluster w with multinomial parameter $\pi_w = P(W=w)$;
2. $f(z_1|w)$ is the initial-regime probability; that is, the probability of starting the sequence in regime k conditional on belonging to latent class w with multinomial parameter $\lambda_{kw} = P(Z_1 = k|W=w)$;
3. $f(z_t|z_{t-1},w)$ is a latent transition probability; that is, the probability of being in a particular regime at time point t conditional on the regime at time point $t-1$ and class membership; assuming a time-homogeneous transition process, we have $p_{jkw} = P(Z_t = k|Z_{t-1}=j, W=w)$ as the relevant multinomial parameter. In other words, within cluster w one has the transition probability matrix of the Markov switching mechanism

$$P_w=\begin{pmatrix} p_{11w} & \cdots & p_{1Kw} \\ \vdots & \ddots & \vdots \\ p_{K1w} & \cdots & p_{KKw} \end{pmatrix}$$

with $\sum_{k=1}^{K} p_{jkw}=1$ for every row of P_w. The model allows that each cluster has its specific transition or regime-switching dynamics, whereas in a standard Markov switching model it is assumed that all cases have the same transition probabilities;

4. $f(y_{it}|z_t)$, the probability density of having a particular observed y_{it}, conditional on the regime occupied at time point t, is assumed to have the form of a univariate normal (or Gaussian) density function. This distribution is characterized by the parameter vector $\theta_k=(\mu_k,\sigma_k^2)$ containing the mean ($\mu_1 k$) and variance (σ_k^2) for regime k. Note that these parameters are assumed to be equal across clusters, an assumption that may, however, be relaxed.

Since $f(y_i;\varphi)$, defined by Equation (1), is a mixture of densities across clusters w and regimes, it defines a flexible Gaussian mixture model that can accommodate deviations from normality in terms of skewness and kurtosis. The Markov switching model with S latent classes and K regimes has $SK^2 + 2K - 1$ free parameters to be estimated, including $S - 1$ class sizes, $S(K - 1)$ initial-regime probabilities, $SK(K - 1)$ transition probabilities, K conditional means, and K conditional variances.

Maximum likelihood (ML) estimation of the parameters of the model involves maximizing the log-likelihood function: $\ell(\varphi; y) = \sum_{i=1}^{n} \log f(y_i;\varphi)$, a problem that can be solved by means of the Expectation-Maximization (EM) algorithm (Dempster et al., 1977). In the E step, we compute the joint conditional distribution of the $T + 1$ latent variables given the data and the current provisional estimates of the model parameters. In the M step, standard complete data ML methods are used to update the unknown model parameters using an expanded data matrix with the estimated densities of the latent variables as weights. Since the EM algorithm requires us to compute and store the $S.K^T$ entries in the E step this makes this algorithm impractical or even impossible to apply with more than a few time points. However, a special variant of the EM algorithm has been proposed that is usually referred to as the forward-backward or Baum-Welch algorithm (Baum *et al.*, 1970). The Baum-Welch algorithm circumvents the computation of this joint posterior distribution making use of the conditional independencies implied by the model.

An important modeling issue is the selection of the value of S and K, the number of clusters needed to capture the unobserved heterogeneity and the number of regimes, respectively. The decision on model dimension is typically based on information statistics, the best known being BIC and AIC. In our application we select S and K that minimize the information criterion (IC) value defined as:

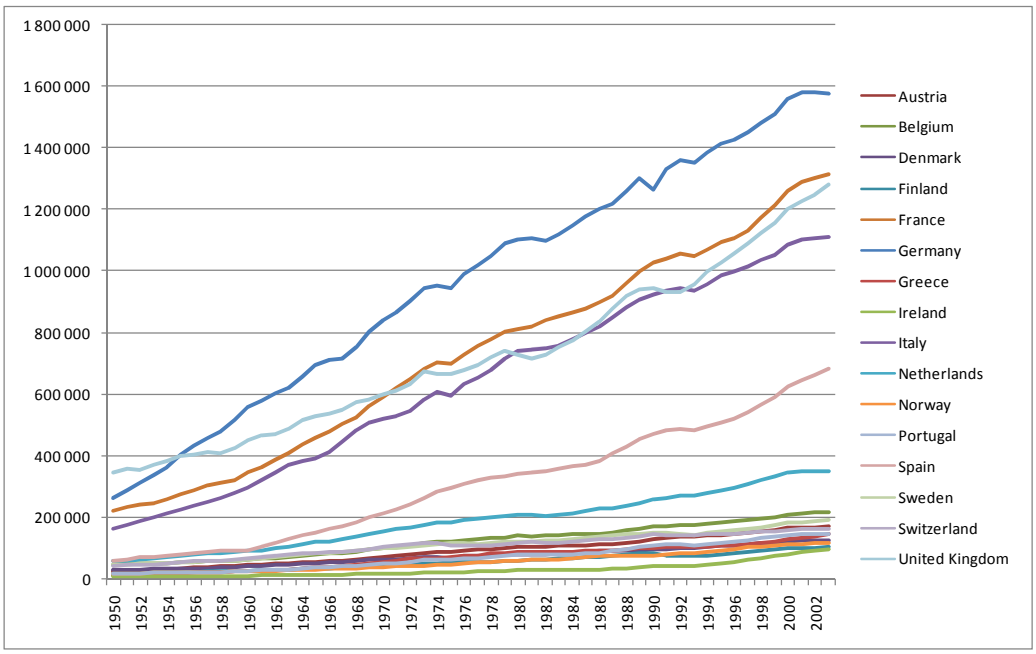
$$IC_{S.K} = -2\ell(\hat{\varphi}; y) + d \cdot N_{S.K},$$

where $N_{S.K}$ is the number of free parameters of the model concerned. The Akaike Information Criterion (AIC) sets $d = 2$ (Akaike, 1974). For the Bayesian Information Criterion (BIC1) of Schwarz (Schwarz, 1978), it takes $d = \log n$. An alternative definition of the penalization is $d = \log(nT)$, as it assumes the sample size being the product of the number of units by the length of the time series, resulting BIC2. A simulation study reported in Dias and Vermunt (2007) shows that overall BIC identifies the correct dimension of the model more often than BIC2. Similarly, CAIC1 and CAIC2 are the Consistent Akaike Information Criterion with $d = 1 + \log(n)$ and $d = 1 + \log(nT)$, respectively (Bozdogan, 1987). Bozdogan (1993) suggested AIC3 that takes $d = 3$. Dias (2007) showed that AIC3 tends to retrieve the correct model dimension more often than AIC and BIC.

2. DATA SET

The data set used in this article is taken from Maddison (2007). It comprises the period from 1950 to 2003 for 16 European economies. The time series are already available in 1990 International Geary-Khamis dollars, which allows a direct comparison. Figure 1 depicts the 16 time series and the countries involved in the analyses.

Figure 1. Economic growth in 16 countries between 1950 and 2003 (GDP in million 1990 International Geary-Khamis dollars)



Out of these 16 countries, the large size of four of them is clear: United Kingdom, Germany, France, and Italy. In this period Spain has shown a higher economic growth rate, and at less extent the Netherlands. These time series include periods of economic expansion and some periods of crises as the ones in the 70's, or more specific breaks as for example the case of German reunification.

Let GDP_{it} be the observed GDP for country i in period t. We model the GDP growth rates, and consequently

$$y_{it} = 100 \cdot \left(\frac{GDP_{it}}{GDP_{i,t-1}} - 1 \right), i = 1, \ldots, 16; t = 1951, \ldots, 2003.$$

Given that finite mixtures can mimic extremely complex patterns of observed data such as non-normality, asymmetry, excess kurtosis, we apply the Markov switching model to this data set.

3. RESULTS

This section summarizes the results of the application of the econometric model described in Section 2 to the panel data set introduced in the previous section. We identify shifts in mean growth rates and volatilities by estimating Markov switching models to GDP growth rates of 16 European countries.

Table 1. Model selection

Model	S	K	Log-likelihood	AIC	AIC3	BIC1	BIC2	CAIC1	CAIC2	Nr. of parameters
I	1	1	-2032.79	4069.57	4071.57	4071.12	4079.06	4073.12	4081.06	2
II	1	2	-1921.10	3856.21	3863.21	3861.61	3889.41	3868.61	3896.41	7
III	1	3	-1883.82	3795.64	3809.64	3806.45	**3862.04**	3820.45	**3876.04**	14
IV	1	4	-1863.76	3773.53	**3796.53**	3791.30	3882.61	**3814.30**	3905.61	23
V	1	5	-1847.52	3763.04	3797.04	**3789.31**	3924.29	3823.31	3958.29	34
VI	1	6	-1832.49	**3758.98**	3805.98	3795.29	3981.89	3842.29	4028.89	47
VII	2	1	-2032.79	4071.57	4074.57	4073.89	4085.80	4076.89	4088.80	3
VIII	2	2	-1918.57	3859.15	3870.15	3867.64	3911.32	3878.64	3922.32	11
IX	2	3	-1877.99	3801.97	3824.97	3819.74	3911.06	3842.74	3934.06	23
X	2	4	-1855.82	3789.64	3828.64	3819.77	3974.61	3858.77	4013.61	39
XI	2	5	-1832.93	3783.87	3842.87	3829.45	4063.70	3888.45	4122.70	59
XII	2	6	-1813.06	3792.12	3875.12	3856.25	4185.78	3939.25	4268.78	83

We estimated models characterized by different number of clusters ($S = 1,2$) and regimes ($K = 1, \dots, 6$), using for the estimation of each of them 100 different starting values for the parameters to avoid local maxima. Table 1 reports information criteria that allow the selection of the best model.

From Table 1 it turns out that there is disagreement between the information criteria on the proper model dimension. Models with one latent class ($S = 1$) always dominate models with more than one latent class. The fact that we have a single latent class is in itself a very important finding: GDP dynamics are homogeneous across European countries. It means that all countries share the same latent process.

The number of regimes needed ranges from 3 (BIC2 and CAIC2) to at least 6 (AIC). Given the results in Dias and Vermunt (2007) that show underfitting of BIC2 and CAIC2, and the tendency of AIC to overfit the data, we selected the 4-regime solution suggested by AIC3. Therefore, we have a solution with four temporal regimes that are enough for dynamic pattern description. Table 2 summarizes the estimates of model parameters. It provides the mean and variance estimates in each regime, the initial distribution and transition probabilities for the latent process and regime occupancy.

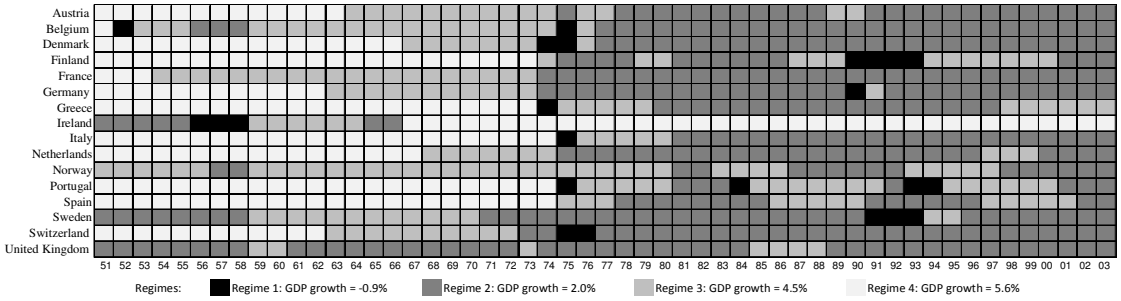
Table 2. Estimates of the parameters of the model

a) Observed process				
	Regime 1	**Regime 2**	**Regime 3**	**Regime 4**
Parameters				
GDP growth (mean)	-0.8957	1.9814	4.5232	5.5857
	(1.2570)	*(0.1345)*	*(0.1418)*	*(0.2594)*
GDP growth (variance)	6.1942	2.2739	1.529	8.4867
	(2.5803)	*(0.3038)*	*(0.2941)*	*(0.9512)*
b) Latent process				
Initial distribution				
P(Z$_0$)	0.0342	0.1724	0.0983	0.6952
	(0.0670)	*(0.1394)*	*(0.1419)*	*(0.1502)*
Transition probabilities				
Regime 1	0.4385	0.0648	0.4911	0.0056
	(0.1324)	*(0.0958)*	*(0.1406)*	*(0.0224)*
Regime 2	0.0235	0.911	0.0594	0.0061
	(0.0178)	*(0.0328)*	*(0.0332)*	*(0.0089)*
Regime 3	0.0266	0.1856	0.7865	0.0014
	(0.0362)	*(0.0619)*	*(0.0463)*	*(0.0054)*
Regime 4	0.0132	0.0009	0.0413	0.9446
	(0.0123)	*(0.0034)*	*(0.0188)*	*(0.0173)*
Regime occupancy				
P(Z)	0.0367	0.457	0.2465	0.2598
	(0.0349)	*(0.0688)*	*(0.0479)*	*(0.0709)*

Note: Standard errors in brackets.

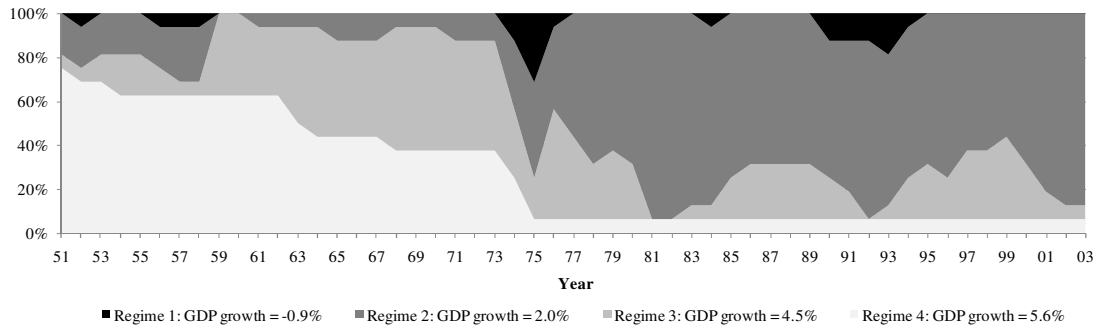
Four regimes were identified, whose means are nicely connected to traditional business cycle characteristics: economic contraction (GDP growth = -0.9%), low economic growth (GDP growth = 2.0%), medium economic growth (GDP growth = 4.5%), and high economic growth = 5.6%). From a variability viewpoint one observes that extreme regimes (1 and 4) show larger variance comparing with regimes 2 and 3. In 1951 more than 69.5% of the countries were in regime 4 (high growth), and only around 3.4% of the countries were in recession regime (estimated initial distribution). From transition probabilities, one observes that all regimes tend to be regime persistent with high retention probabilities. Quite interestingly, after a year of recession (regime 1) the most likely move is to regime 3 (medium economic growth). From regime 3 is more likely to move to regime 2 than to regime 4. Overall the most likely regime is low economic growth (0.457). On the other hand, the probability of a year of crisis is 0.0367.

Figure 2. Posterior modal regimes for each country



Figures 2 depicts the posterior latent modal regime; that is, the regime with maximum posterior probability. The model identifies the important dynamic stages in the European economic growth. There is a general synchronization of the European countries with country-especific characteristics. Most of the countries in the post-war period had high economic growth in the 1950s and 1960s, the exceptions being Belgium, Ireland, Sweden and United Kingdom. The 1970s is a decade of transition, starting a shift towards moderate economic growth rates. This decade combines both economic and political shifts in some countries. For example, Greece and Portugal start democratic experiences which created economic disruption and affected economic growth. Moreover, countries like Belgium, Denmark, Italy, and Switzerland felt a deeper impact of the oil shock crisis. Interestingly, the model can retrieve periods of crises such as in 1990, after the reunification in Germany. With exception of Ireland, most European economies kept a moderate or low economic growth in the recent years.

Figure 3. Proportion of regime occupancies from 1960 to 2003



Figures 3 shows the percentage of countries in each regime from 1951 to 2003. In this period European economic growth shifts from regimes of high economic growth in the post-war period to slow economic growth regimes in most of the countries after the oil shock crises in the 1970s. Until 2003, Ireland was able to keep a high economic growth rate. As expected, the slow economic growth regime has become dominant in recent years (conditional viewpoint).

4. CONCLUSIONS

Markov switching and finite mixture models have proven to be powerful tools for analyzing unobserved heterogeneity and asymmetry in time series data (see, for example, McLachlan and Peel, 2000; Dias *et al.*, 2008). This article proposes the application of an extension of the traditional Markov switching model that takes into account both time-constant unobserved heterogeneity between and hidden regimes within time series. The proposed approach is flexible in the sense that it can deal with the specific features of economic time series data, such as asymmetry, kurtosis, and unobserved heterogeneity.

This methodology was applied to the GDP growth time series of 16 countries from 1950 to 2003. The best fitting model was the one with one latent class, showing a homogeneous regime switching mechanism shared by the 16 countries. However, there is heterogeneity over time, as one needs four regimes to express each time series over time. Our results show that we can track and map the pattern of each country and identify for example synchronization between the economies which are relevant results for policy decision making.

BIBLIOGRAPHY

AKAIKE, H. (1974), A new look at statistical model identification, *IEEE Transactions on Automatic Control*, AC-19, 716-723.

ARTIS, M., H.-M. KROLZIG, J. TORO (2004), The European business cycle, *Oxford Economic Papers*, 56, 1-44.

BAUM, L.E., PETRIE, T., SOULES, G., WEISS, N. (1970), A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains, *Annals of Mathematical Statistics*, 41, 164-171.

BOZDOGAN, H. (1987), Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions, *Psychometrika*, 52(3), 345-370.

BOZDOGAN, H. (1993), Choosing the number of component clusters in the mixture-model using a new informational complexity criterion of the inverse-Fisher information matrix, *in* Opitz, O., B. Lausen, and R. Klar (Eds.), *Information and Classification, Concepts, Methods and Applications*, Berlin: Springer-Verlag, 40-54.

DEMPSTER, A.P., LAIRD, N.M., RUBIN, D.B. (1977), Maximum likelihood from incomplete data via the EM algorithm (with discussion), *Journal of the Royal Statistical Society*, B 39, 1-38.

DIAS, J.G. (2007), Complete data information criteria for the Naive-Bayes model: A Monte Carlo study, *Journal of the Korean Statistical Society*, 36(3), 435-445.

DIAS, J.G. AND J.K. VERMUNT (2007), Latent class modeling of website users' search patterns. Implications for online market segmentation, *Journal of Retailing and Consumer Services*, 14, 359-368.

DIAS, J.G., VERMUNT, J.K., and RAMOS, S. (2008). Heterogeneous hidden markov models, *in* Brito, P. (ed.), *COMPSTAT2008. Proceedings in Computational Statistics*, Heidelberg: Physica/ Springer-Verlag, 373-380.

HALLETT, A. H., and C. RICHTER (2006), Measuring the degree of convergence among European business cycles, *Computational Economics*, 27(2-3): 229-259.

HAMILTON, J. (1989), A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 57, 357-384.

MADDISON, A. (2007), *Contours of the World Economy, 1-2030 AD. Essays in Macroeconomic History*, Oxford: Oxford University Press.

MCLACHLAN, G.J., D. PEEL (2000), *Finite Mixture Models*, John Wiley & Sons, New York.

MISAS, M. and M.T. RAMÍREZ (2007), Depressions in the Colombian economic growth during the twentieth century: A Markov switching regime model, *Applied Economics Letters*, 14, 803-808.

SCHWARZ, G. (1978), Estimating the dimension of a model, *Annals of Statistics*, 6, 461-464.

TAN, S.-H. and M. S.HABIBULLAH (2007), Business cycles and monetary policy asymmetry: An investigation using Markov-switching models, *Physica A*, 380, 297-306.

ROBUSTNESS OF THE MODIFIED LEVENE'S TESTS WITH AND WITHOUT BOOTSTRAPPING

ISABEL PARRA-FRUTOS
E-mail: ipf@um.es

Departamento de Métodos Cuantitativos para la Economía y la Empresa
Facultad de Economía y Empresa
Universidad de Murcia

Abstract

This is a simulation study that analyzes the robustness properties of tests for equality of variances in terms of both Type I error rate and power. The well-known Levene's and Bartlett's tests, along with different modifications proposed for Levene's test with and without bootstrapping, are examined. Without bootstrapping, some of the modifications may produce better homogeneity tests. Bootstrapping Levene's test based on the median and the same test simultaneously modified by Keyes-Levy's adjustment and Satterthwaite's correction are always robust for the significance level under any kind of distribution and combination of sample sizes.

Keywords: Homoscedasticity tests, simulation, bootstrapping.
Thematic Area: Quantitative Methods for Economics and Business.

Resumen

Este trabajo consiste en un estudio de simulación que analiza la robustez de los tests de igualdad de varianzas en cuanto a la probabilidad del error tipo I y a su potencia. Los conocidos test de Levene y test de Bartlett junto con diferentes modificaciones propuestas para el primero con o sin bootstrapping son investigados. Sin bootstrapping, algunas de las modificaciones estudiadas pueden producir mejores tests the homogeneidad. El test de Levene basado en la mediana con bootstrapping y éste mismo conjuntamente con el ajuste de Keyes y Levy y la corrección de Satterthwaite son tests robustos en cuanto al nivel de significación independientemente de la distribución de probabilidad considerada y los tamaños muestrales utilizados.

Palabras Clave: Tests de homoscedasticidad, simulación, bootstrapping.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

1. INTRODUCTION

There is considerable statistical literature on testing homogeneity of variances. A great number of tests have been examined and simulated to determine their robustness at nominal significance levels. The tests that continue to receive the most attention are the *F*-test (two samples), Levene's (1960) test, and Bartlett's (1937) test. Simulation studies (e.g. Brown and Forsythe 1974, and Conover, Johnson and Johnson 1981) suggest there is no test that is uniformly best for all distributions and sample size configurations.

The *F*-test is extremely sensitive to the normality assumption. The true distribution may be symmetric and resemble the normal distribution and yet the true level of significance may be two or three times as high as it is supposed to be (Siegel and Tukey 1960, and Markowski and Markowski 1990). For unequal sample sizes and under normality, the *F*-test is not generally

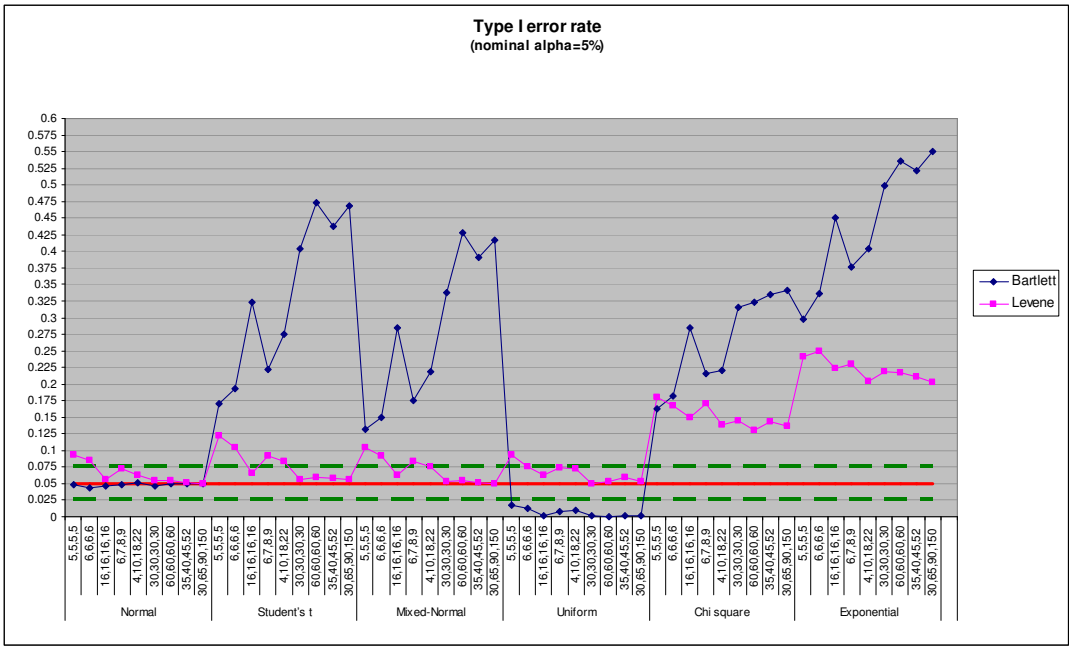
capable of detecting homoscedasticity violations that lead to poor performance of the *t*-test (Markowski and Markowski 1990).

Bartlett's test depends crucially on the assumption of normal distributions or on large samples whose kurtosis is equal to the normal. If the distributions have more kurtosis than the normal distribution, this would represent a liberal test; otherwise it is conservative (Boos and Brownie, 1989).

Levene's test is a better test in terms of power and robustness of the significance level under normality violations (Conover, Johnson and Johnson, 1981; Lim and Loh, 1996). It is often presented as an option for testing homoscedasticity in popular statistics packages.

To illustrate the behaviour, in terms of Type I error rates, of Bartlett's and Levene's test, a simulation study has been carried out. The results are presented in Figure 1. Two conclusions may be drawn from this figure: the Bartlett test is highly non-robust to non-normality and Levene's test is not recommendable in asymmetric distributions.

Figure 1 Type I error rate of the Bartlett and Levene tests, for 4 samples, under different settings: probability distributions and sample sizes.

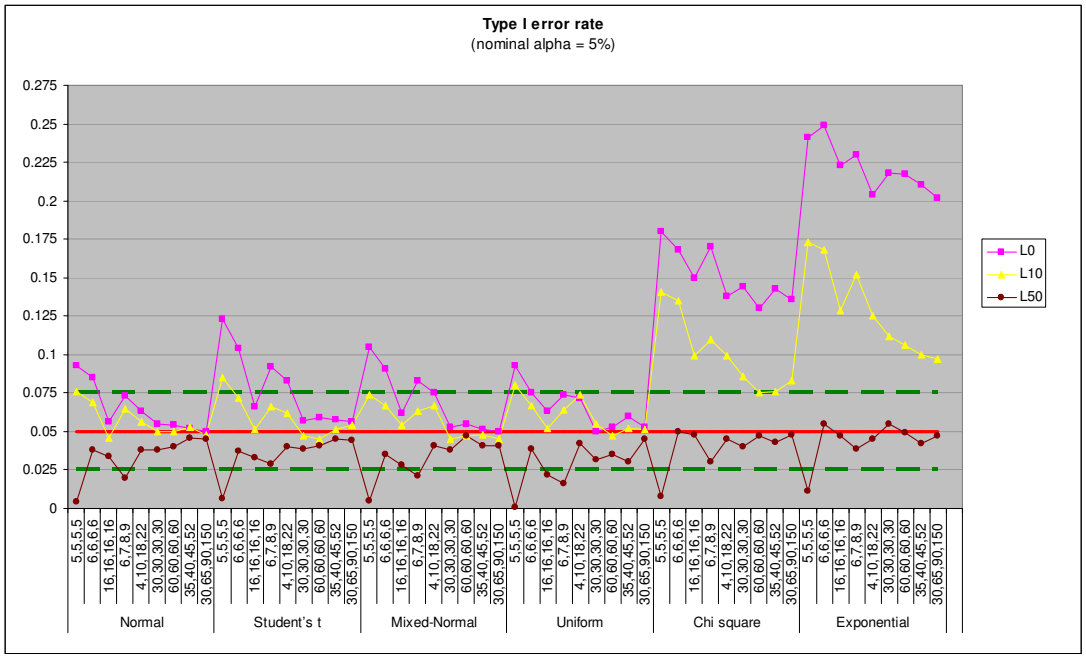


Let y_{ij}, $i=1,...,k$ and $j=1,...,n_i$, denote the *j*th observation from the *i*th group with size n_i. Levene's (1960) test is the best known homoscedasticity test, defined as the one-way analysis of variance (ANOVA) on the absolute residual $z_{ij}=\left|y_{ij}-\overline{y}_i\right|$, where $\overline{y}_i$ is the sample mean of the *i*th group.

New proposals emerged because of the problems of the classic homogeneity tests. Various modifications of Levene's test have been proposed, including investigations by Miller (1968), Brown and Forsythe (1974), O'Brien (1978, 1979), Carroll and Schneider (1985), Yitnosumarto and O'Neill (1986), Loh (1987), and Keyes and Levy (1997), to name but some.

Brown and Forsythe's (1974) modifications of Levene's test are the most popular. They consider the *median* and the *10 percent trimmed mean*, which are more robust estimates of location, as alternatives to the mean in computing absolute deviations. The former usually stands out in terms of power and robustness against non-normality (Brown and Forsythe 1974; Lim and Loh 1996). See, also, Figure2.

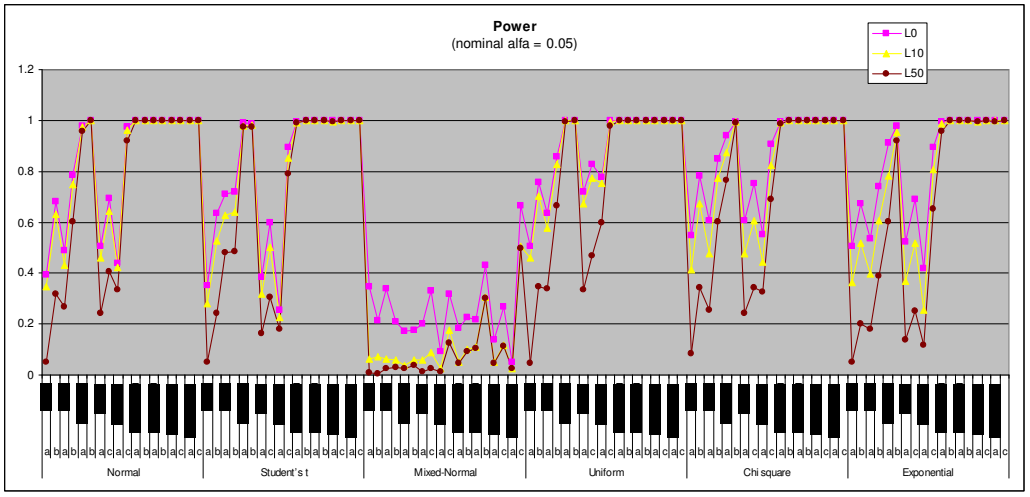
Figure 2 Type I error rate of the Brown and Forsythe's (1974) modifications of Levene's test.



NOTE: L0 (Levene's test); L10 (deviations from the 10 per cent fractional trimmed mean); and L50 (deviations from the median).

On examining power, we observe that, in small samples, the lower the Type I error, the lower the power. A tradeoff between Type I and Type II errors is showed (see Figure 3). In small samples (except in small, odd, equal sizes), Levene's test based on the median controls Type I error rate although a lower power is exhibited.

Figure 3 Estimated power level of the Brown and Forsythe's (1974) modifications of Levene's test.



NOTE: L0 (Levene's test); L10 (deviations from the 10 per cent fractional trimmed mean); and L50 (deviations from the median).

Loh (1987) examines the effectiveness of applying Satterthwaite's (1946) method of correcting degrees of freedom and data-based power transformations on Levene's test with group medians instead of group means. Loh finds that Satterthwaite's method may improve the robustness of Levene's test for small samples but data-based power transformations increase the significance level of the test. Yitnosumarto and O'Neill (1986) give another method for modifying the degrees of freedom.

Keyes and Levy (1997) prove that, with normal populations, only when sample sizes are equal is the null hypothesis of Levene's test homoscedasticity; otherwise the null hypothesis is equality of design-dependent multiples of the population variances:

$$H_0 : \left(1-1/n_1\right)\sigma_1^2 = \left(1-1/n_2\right)\sigma_2^2 = \cdots = \left(1-1/n_k\right)\sigma_k^2.$$

Consequently, they propose an adjusted test for correcting this problem and for the real null hypothesis to be always the desired one (homoscedasticity). An inflated significance level is, however, exhibited. They suggest using standard techniques like Satterthwaite's to adjust significance levels. They recommend regular use of their adjustment to Levene's test for unbalanced designs and speculate that the robustness of Levene's test is unaffected when used simultaneously with Satterthwaite's correction. We investigated robustness when both adjustments are applied.

Lim and Loh (1996) compare Levene's test and Bartlett's test and their bootstrap versions, the Box-Andersen (1955) test, and three jackknife tests. They recommend the bootstrap version of Levene's test with the sample median as an estimate of the location parameter. If the distribution does not have very high kurtosis, the bootstrap version of Bartlett's test is an alternative. They affirm that to apply the tests in practice it is highly desirable to have a balanced design with large sample sizes (at least 10 in each group) to achieve good power.

To assess robustness of tests we use Bradley's (1978) liberal criterion of robustness at nominal significant level, which establishes that a test is considered robust if its empirical Type I error rate falls within the interval $[0.025, 0.075]$ for a nominal level $\alpha = 0.05$. Not all authors agree with this criterion. Conover et al. (1981), for example, define a test to be robust if the maximum empirical Type I error rate is less than *0.10* for a *5* percent test.

Our aim is to compare the classic tests (Bartlett's test and Levene's test) with the modified versions of Levene's test, and we also consider bootstrapping. We extend the analysis to determine empirically how the modifications proposed for Levene's test affect robustness of the significant level (the ability to maintain the nominal level of the test under a variety of different settings) and power (the ability to reject H_0 when it is false). Various versions of Levene's test are analyzed: when residuals are calculated with respect to the mean, median, and fractional trimmed mean under different designs (balanced and imbalanced, small and large sample sizes, and for symmetric, asymmetric, and heavy-tailed distributions). The results are then compared with results obtained using Satterthwaite's correction, Keyes and Levy's adjustment and both at the same time.

2. DESCRIPTION OF TESTS

Originally Levene's test (L0 test) was defined as the one-way analysis of variance on $z_{ij} = \left| y_{ij} - \overline{y}_i \right|$, the absolute residuals $r_{ij} = y_{ij} - \overline{y}_i$, $i = 1,...,k$ and $j = 1,...,n_i$, where k is the number of groups and n_i the sample size of the i-th group. The test statistic is given by:

$$L = \frac{\left(N-k\right)\sum_{i=1}^{k}\left(\overline{z}_i - \overline{z}\right)^2 n_i}{\left(k-1\right)\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(z_{ij} - \overline{z}_i\right)^2} \rightarrow F_{k-1,N-k}\,, \tag{3}$$

where

$$N = \sum_{i=1}^{k} n_i\,,$$

$$\overline{z}_i = \sum_{j=1}^{n_i} z_{ij} \big/ n_i\,,$$

$$\overline{z} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} z_{ij} \big/ N\,.$$

Brown and Forsythe (1974) subsequently proposed the absolute deviations from the 10 percent trimmed mean $\tilde{y}_i$ (L10 test), thus $z_{ij} = \left| y_{ij} - \tilde{y}_i \right|$, and from the median of the *i*-th group Me_i (L50 test), so $z_{ij} = \left| y_{ij} - Me_i \right|$. Rather than the δ percent trimmed mean, the δ percent *fractional* trimmed mean is used, which is slightly different and may be defined as:

$$\tilde{Y} = \frac{(1-\gamma+g)Y_{(g+1)} + \sum_{i=g+2}^{n_i-g-1} Y_{(i)} + (1-\gamma+g)Y_{(n_i+1)}}{n_i - 2\gamma}$$

where $\gamma = n_i\delta/100$, g is γ rounded down to the nearest integer, and $Y_{(1)} < Y_{(2)} < \cdots < Y_{(n_i)}$ are the observations of the *i*-th sample written in ascending order.

Levene's test attempts to detect significant differences between the means of the absolute residuals, which are functions of the population variances and sample sizes (see Keyes and Levy 1997). When sample sizes are identical, the null hypothesis is homoscedasticity, but when sample sizes are different then the null hypothesis is equality of design-dependent multiples of the population variances. Keyes and Levy (1997) suggest that a remedy is to base the one-way ANOVA on the absolute value of the modified residual d_{ij} (L0(KL) test):

$$d_{ij} = \frac{r_{ij}}{\sqrt{1-1/n_i}}.$$

For this adjustment $E\left[\left|d_i\right|\right] = \left(2\sigma_i^2/\pi\right)^{1/2}$ and $Var\left[\left|d_i\right|\right] = (1-2/\pi)\sigma_i^2$.

The analysis is extended to study this modification of Levene's test, considering the 10 percent fractional trimmed mean (L10(KL) test) and the median (L50(KL) test) to calculate the residuals r_{ij}.

To apply the ANOVA test, several assumptions must be verified, including normal populations, homoscedasticity, and independent observations. The absolute residuals do not meet any of these assumptions, so Levene's test is an approximate test of homoscedasticity. Thus, the real significance level is different from the nominal significance level. To mitigate this effect, Keyes and Levy (1997) and Loh (1987) recommend Satterthwaite's correction for degrees of freedom. Instead of referring to the *F*-distribution with *k*-1 and *N-k* degrees of freedom, Satterthwaite suggests *k*-1 and *v*, where:

$$v = \frac{\left(\sum_i u_i\right)^2}{\sum_i u_i^2/(n_i-1)} \qquad u_i = \sum_j \left(z_{ij} - \overline{z}_i\right)^2.$$

We study the effect of Satterthwaite's correction on robustness of Levene's test when absolute deviations are calculated from the sample mean [L0(S) test], the 10 percent fractional trimmed mean [L10(S) test], and the median [L50(S) test]. We also examine the Keyes-Levy adjustment and the Satterthwaite correction when applied simultaneously [L0(KL-S), L10(KL-S), and L50(KL-S)].

Bartlett's (1937) test statistic is given by:

$$\frac{M}{1+C} \rightarrow \chi^2_{k-1}, \qquad (1)$$

where

$$M = (N-k)\ln S_a^2 - \sum_{i=1}^{k}(n_i - 1)\ln S_i^2$$

$$S_i^2 = \frac{\sum_{j=1}^{n_i}\left(y_{ij} - \overline{y}_i\right)^2}{n_i - 1}$$

$$S_a^2 = \frac{\sum_{i=1}^{k}(n_i - 1)S_i^2}{N-k}$$

$$C = \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\frac{1}{n_i - 1} - \frac{1}{N-k}\right).$$

3. DESIGN OF THE SIMULATION

The Type I error rate and the power of the tests are compared in a simulation using six distributions and nine configurations of group sizes (n_1, n_2, n_3, n_4). The distributions are: normal (symmetric), Student's t (symmetric, long-tailed and low kurtosis), mixed normal (symmetric and heavy-tailed), uniform (symmetric and very low kurtosis), chi-square (skewed, long-tailed and high kurtosis) and exponential (skewed, heavy-tailed and high kurtosis). The nominal 5 percent significance level is used throughout. The simulation results are based on 5,000 replications in nonbootstrap tests and on 500 bootstrap replications for 1,000 Monte Carlo replications in bootstrap tests. The S language is used.

The mixed (or contaminated) normal has a distribution function given by $F(x) = (1-\lambda)F_1(x) + \lambda F_2(x;\sigma)$, where $0<\lambda<1$, $F_1(x)$ is the standard normal distribution function and $F_2(x;\sigma)$ is the normal distribution function with standard deviation σ. This distribution is symmetric and quite similar to the normal distribution when λ is close to 0. The distribution differs from the normal in that we see outliers more often than would be expected for a normal distribution. In our simulation study $\lambda = 0.05$ and $\sigma = 3$.

Five configurations of small samples are studied for each distribution. Three of them are of equal sample size: (5,5,5,5), (6,6,6,6), and (16,16,16,16). Two of them are of unbalanced design: (6,7,8,9) and (4,10,18,22). Four configurations of large samples are also studied: (30,30,30,30), (60,60,60,60), (35,40,45,52), and (30,65,90,150). A null hypothesis of equal variances is studied along with three alternatives: $(\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2)$ = (1,6,11,16), (16,11,6,1), and (1,1,1,16).

A total of 54 settings are considered (9 sample size configurations extracted from each of 6 probability distributions). Twenty-six tests (13 non-bootstrap tests and 13 bootstrap tests) are studied in each of those settings. When estimating power, two different combinations of variances are assumed for the parent populations, thus 108 settings now emerge.

The bootstrap resampling procedure we use is based on the pooled bootstrap approach given by Boos and Brownie (1989). It is carried out as follows:

1. Compute the value of the test statistic *T* from the observed data.

2. Initialize $R = 0$.

3. Compute the residuals $r_{ij} = y_{ij} - \tilde{y}_i$, $i = 1,...,k$, $j = 1,...,n_i$, where $\tilde{y}_i$ is the 10 percent fractional trimmed mean of group i.

4. Draw *N* bootstrap samples r_{ij}^* with replacement from the set of pooled residuals $\{r_{ij} : i = 1,...,k,\ j = 1,...,n_i\}$.

5. If $n_i < 10$ for at least one sample size, smooth the bootstrap samples by setting $y_{ij}^* = (12/13)^{1/2}\left(r_{ij}^* + vU\right)$, where $v^2 = \sum\sum\left(y_{ij} - \overline{y}_i\right)^2 / N$ and U is an independent uniform random variable on $(-1/2, 1/2)$. Otherwise define $y_{ij}^* = r_{ij}^*$..

6. Compute one realization of the bootstrap statistic T^* from the bootstrap samples y_{ij}^*. If $T^* > T$, increment R to $R+1$.

7. Repeat steps 4, 5, and 6 a total of A times (A=500 is used in the simulation experiment).

The bootstrap *p*-value is given by R/A. The null hypothesis is rejected if R/A is below the nominal level α.

This bootstrap approach is also used by Lim and Loh (1996). Differences come from estimation of the population mean employed in computing residuals, $r_{ij} = y_{ij} - \tilde{\mu}_i$, from which to draw bootstrap samples. We apply the 10 percent fractional trimmed mean, Lim and Loh (1996) the 20 percent fractional trimmed mean, and Boos and Brownie (1989) the sample mean and the 20 percent trimmed mean.

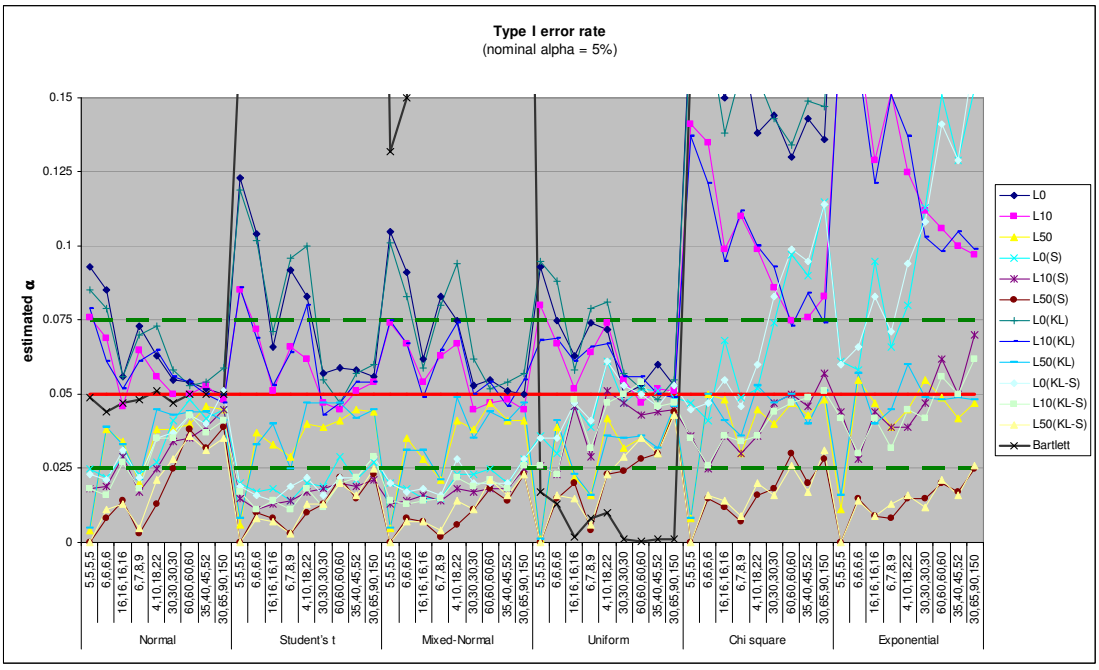
All tests are bootstrapped. The notation consists of adding the letter B as the first letter of the name of the test.

4. SIMULATION RESULTS

All the simulation results are presented in figures where the abscissa axe represents the settings studied and the ordinate axe represents the Type I error rate or the estimated power level. Bradley's (1978) liberal criterion of robustness to nominal significant level, which establishes that a test is considered robust if its empirical Type I error rate falls within the interval $[0.025, 0.075]$ for a nominal level $\alpha = 0.05$, is also represented.

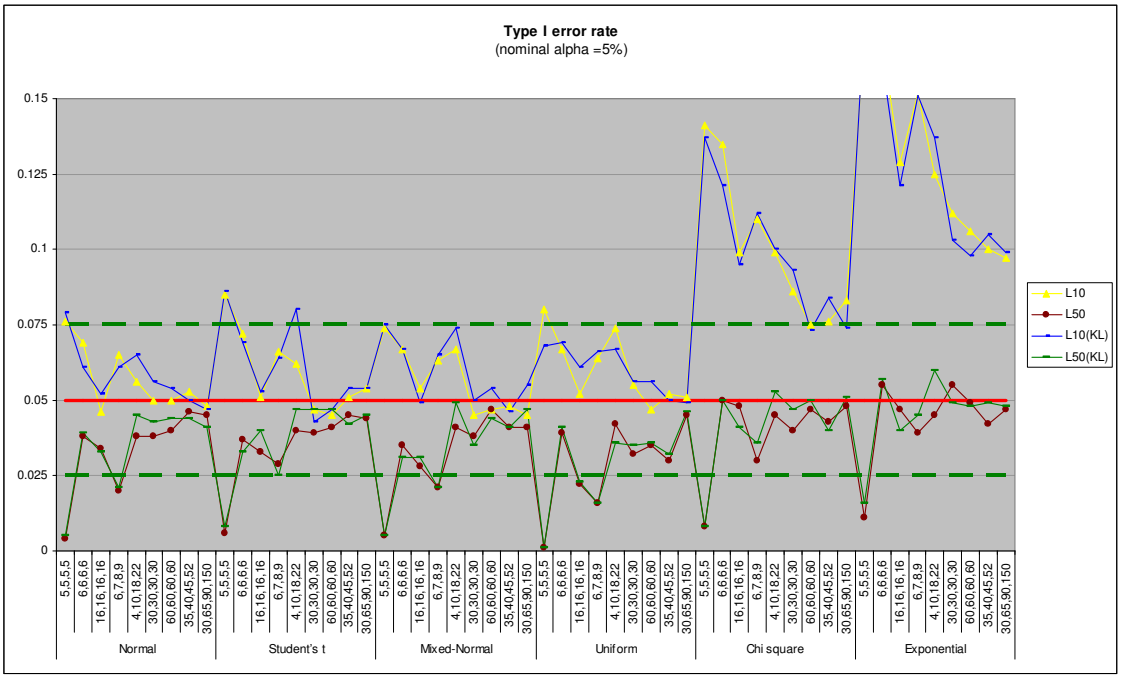
Figure 4 shows that non-bootstrap modified Levene's tests present problems of robustness. After eliminating those tests with severe problems of control of Type I error rate, the tests that stay are L10, L50, L10(KL) and L50(KL), as illustrated in Figure 5.

Figure 4 Type I error rate of the all non-bootstrap tests under 54 settings.



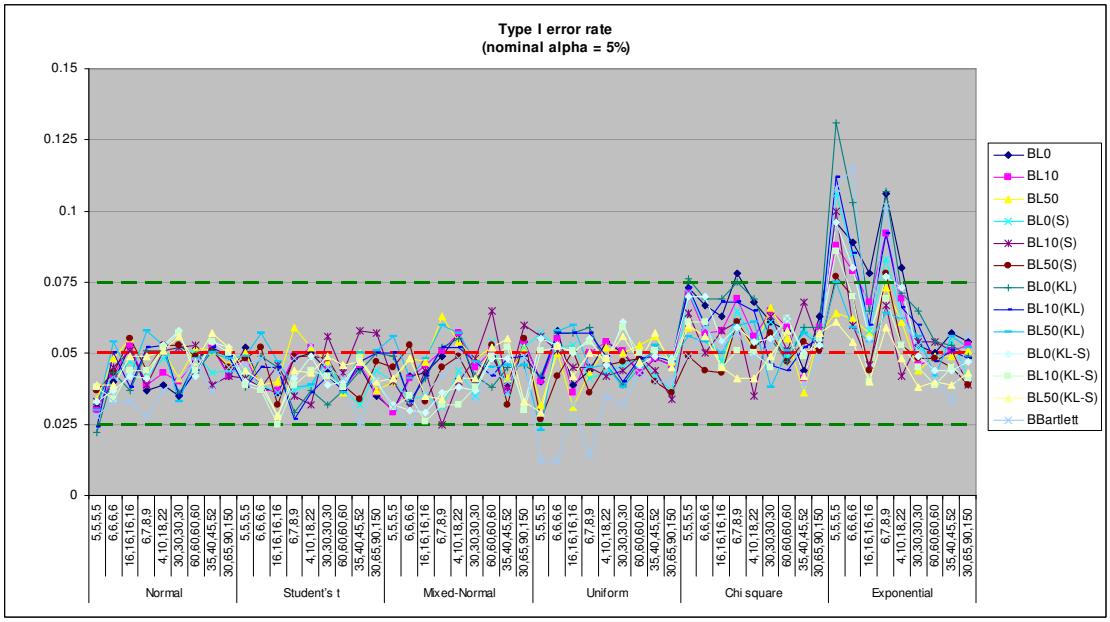
Problems are still present. L10 and L10(KL) cannot be used in asymmetric distributions and L50 and L50(KL) may be conservative especially when sample sizes are small, equal and odd. None of the non-bootstrap tests may be recommended for general use.

Figure 5 Type I error rate of the non-bootstrap tests that control Type I error rate better.



Simulations results for the bootstrap tests are in Figure 6. A general improvement in the robustness when bootstrapping can be observed. Problems appear in skewed, heavy-tailed and with high kurtosis distributions (exponential).

Figure 6 Type I error rate of the bootstrap tests under 54 settings.



The bootstrap tests that control Type I error are BL50, BL50(KL) and BL50(KL-S), as can be seen in Figure 7. The power of these tests is represented in Figure 8. In general, BL50 shows a higher power, although in small samples power of both tests is very low.

Figure 7 Type I error of the bootstrap tests that control Type I error rate.

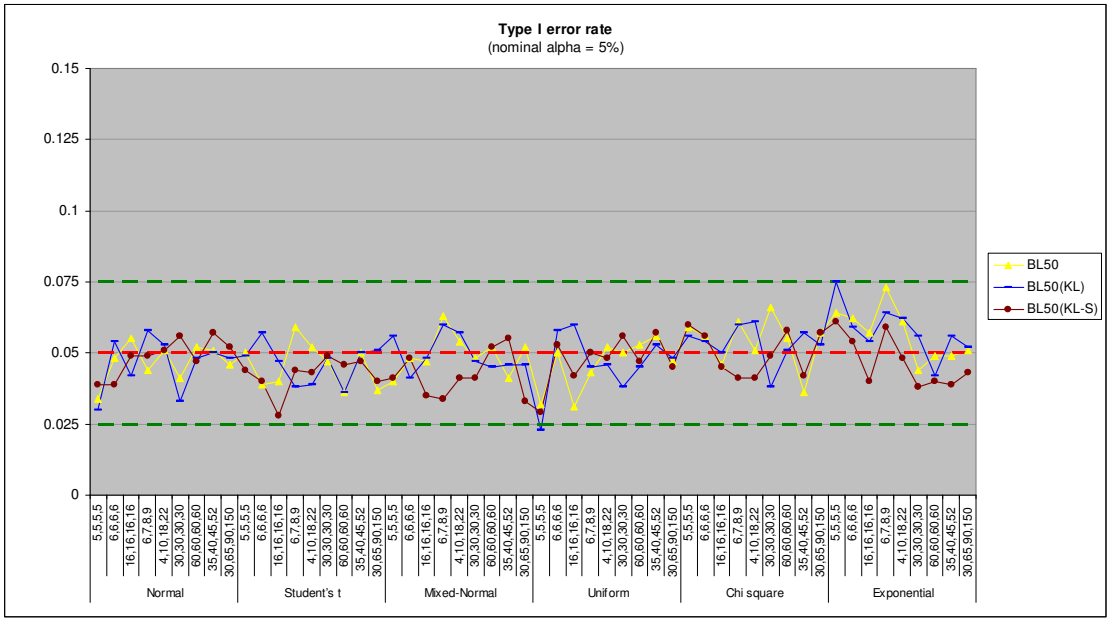
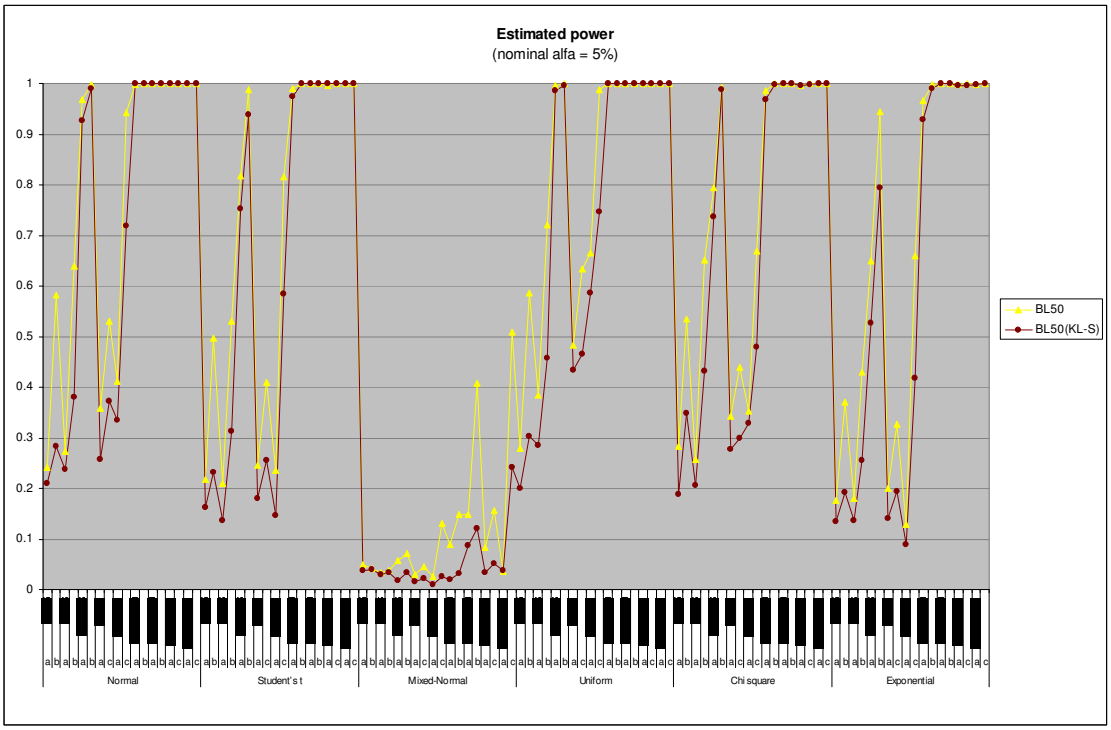


Figure 8 Power of the bootstrap tests that control Type I error rate in 108 settings.



5. CONCLUSIONS

From this study two tests may be recommended for general use because they control Type I error rate according to Bradley's liberal criterion under any kind of distribution (normal or non-normal) and any combination of sample sizes: the bootstrap Levene's test based on the median (BL50), and the bootstrap Levene's test based on the median modified at the same time by the Keyes-Levy adjustment and the Satterthwaite correction (BL50(KL-S)). Power may be considered acceptable in sample sizes that are not very small.

In large samples, and only when bootstrapping, all tests control Type I error rate. They are also very powerful across the different distributions analyzed, except when there are extreme values.

REFERENCES

BARTLETT, M.S. (1937), "Properties of sufficiency and statistical tests", Proceedings of the Royal Society. Series A, A160, pp 268-282

BOOS, D.D.; BROWNIE, C. (1989), "Bootstrap methods for testing homogeneity of variances", Technometrics 31, pp 69-82

BRADLEY, J.V. (1978), "Robustness?", British Journal of Mathematical and Statistical Psychology 31, pp 144-152

BROWN, M.B.; FORSYTHE, A.B. (1974), "Robust tests for equality of variances", Journal of the American Statistical Association, 69, pp 364-367

CARROLL, R.J.; SCHNEIDER, H. (1985), "A note on Levene's test for equality of variances", Statistics and Probability Letters 3, pp 191-194

CONOVER, W.J.; JOHNSON, M.E.; JOHNSON, M.M. (1981), "A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data", Technometrics 23, pp 351-361

KEYES, T.K.; LEVY, M.S. (1997), "Analysis of Levene's test under design imbalance", Journal of Educational and Behavioral Statistics 22, pp 227-236

LEVENE, H. (1960), "Robust tests for equality of variances". In *Contributions to Probability and Statistics. Essays in Honor of Harold Hotelling*. I. Olkin et al., eds. Stanford University Press, 278-292

LIM, T.S.; LOH, W.Y. (1996), "A comparison of tests of equality of variances", Computational Statistics & Data Analysis 22, pp 287-301

LOH, W.Y. (1987), "Some modifications of Levene's test of variance homogeneity", Journal of Statistical Computation and Simulation 28, pp 213-226

MARKOWSKI, C.A.; MARKOWSKI, E.P. (1990), "Conditions for the effectiveness of a preliminary test of variance", American Statistician 44, pp 322-326

MILLER, R.G., Jr. (1968), "Jackknifing variances", Annals of Mathematical Statistics 39, pp 567-582

O'BRIEN, R.G. (1978), "Robust techniques for testing heterogeneity of variance effects in factorial designs", Psychometrika, 43, pp 327-342

O'BRIEN, R.G. (1979), "A general ANOVA method for robust tests of additive models for variances", Journal of the American Statistical Association 74, pp 877-880

SATTERTHWAITE, F.E. (1946), "An approximate distribution of estimates of variance components", Biometrics Bulletin 2, 110-114

SIEGEL, S.; TUKEY, J.W. (1960), "A nonparametric sum of ranks procedure for relative spread in unpaired samples", Journal of the American Statistical Association 55, pp 429-444 (corrections appear in vol. 56, p. 1005),

YITNOSUMARTO, S.; O'NEILL, M.E. (1986), "On Levene's test of variance homogeneity", Australian Journal of Statistics 28, pp 230-241

CORRADO GINI (1884-1965) Y SU RAZON DE CONCENTRACIÓN (1908-1914)

JESÚS BASULTO SANTOS
E-mail: basulto@us.es
JOSÉ JAVIER BUSTO GUERRERO
E-mail: jjbusto@us.es
Facultad de Ciencias Económicas y Empresariales
Universidad de Sevilla

Resumen

A lo largo de los años 1908-1913, Gini propuso varias medidas de desigualdad hasta que en 1914 escribe "Sulla misura della concentrazione e della variabilità dei caratteri", donde recoge su Razón de Concentración. En 1908, Gini propuso el trabajo "Il diverso accrescimento delle classi sociali e la concentrazione della ricchezza", donde, estimulado por el concepto de desigualdad de Pareto (1848-1923) y de su parámetro alfa, Gini introdujo su definición de desigualdad y la medida de desigualdad delta. La memoria "Indici di Concentrazione e di Dependenza" fue escrita en 1910, donde Gini propuso otro criterio de desigualdad y varias medidas de concentración descriptivas. Gini publicó la memoria "Variabilità e mutabilità: contributo allo Studio delle distribuzioni e delle relazioni statistiche" en 1912, donde estudia la Media de las Diferencias (MD). Al final de esta memoria Gini llega al convencimiento de que debe buscar una medida de desigualdad que pueda ser aplicada a cualquier modelo estadístico. La publicación de Max Otto Lorenz (1876-1959) en 1905 de "Methods of mesuring the concentration of wealth" será lo que le lleve a definir a Gini su Razón de Concentración como el área entre la curva de Lorenz y el segmento de igualdad.

Palabras clave: reparto, desigualdad, concentración.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

During the period of 1908-1913, Gini proposed several inequality measures. Gini wrote the report "Sulla misura della concentrazione e della variabilità dei caratteri" in 1914, where he proposed his concentration ratio. In 1908, Gini published "Il diverso accrescimento delle classi sociali e la concentrazione della ricchezza", where he proposed his conception of inequality and his inequality measure δ. The interest of Gini in the income inequality was stimulated by the Type I model of income distribution and the inequality interpretation of parameter α of Pareto (1848-1923).The report "Indici di Concentrazione e di Dependenza" was written by Gini in 1910. In this paper, Gini proposed a new conception of inequality and several descriptive concentration measures. In 1912, Gini published "Variabilità e mutabilità: contributo allo Studio delle distribuzioni e delle relazioni statistiche", where he introduced the mean difference as an alternative way of perceiving the income inequality. After this report, Gini became convinced that he would try to find a distribution-free inequality measure. In 1905, M.O. Lorenz (1876-1959) had published "Methods of measuring the concentration of wealth", which was of very useful to Gini to define his concentration ratio as the area between the egalitarian line and the Lorenz curve.

Thematic Area: Quantitative Methods for Economics and Business.

INTRODUCCIÓN

A lo largo de los años 1908-1914, Gini introducirá sus conceptos de desigualdad y varias medidas de desigualdad hasta lograr su razón de concentración R .

En el mismo año 1908 que Gini publicó su tesis doctoral "Il sesso del punto di vista statistico", presentó en la II Reunión de la Sociedad para el Progreso de la Ciencia el trabajo sobre "Il diverso acrescimiento delle classi sociale e la concentrazione della ricchizza" en Florencia, que sería publicado en 1909 en el Giornale digli Economisti.

En este trabajo de 1908-1909, Gini introduce un nuevo criterio de desigualdad, como alternativa al criterio de "disminución de desigualdad" de Vilfredo Pareto, y propone su primera medida de desigualdad δ a partir de la primera ley de los repartos de las rentas de Pareto. Veremos que si el criterio de desigualdad es válido dentro de las curvas de Lorenz llamadas duales siempre que no se corten (las construidas con proporciones de individuos y rentas cuando ordenamos los individuos de mayor a menor valor de las rentas), la medida δ sólo es válida para la curva de Lorenz de la primera ley de Pareto.

En 1909 Gini fue nombrado profesor temporal de Estadística en la Universidad de Cagliari, accediendo a la cátedra en 1910. Fue en 1909 cuando Gini presentó, en la III Reunión de la Sociedad para el Progreso de las Ciencias, el trabajo de "Índice di Concentrazione e di Dipendenza" en Padua, que publicaría en 1910 en la Biblioteca dell`Economista como una memoria de investigación con el mismo título.

En esta memoria, Gini considera la población finita de n valores de una variable estadística no negativa, la renta, que ordena los n individuos de menor a mayor según sus rentas, para reducirlos a $n-1$ proporciones acumuladas, donde para un valor k , $1 \leq k \leq n-1$, define para los k individuos con mayores rentas su proporción sobre el total de individuos y la proporción del montante de renta que corresponden a estos k individuos. A partir de estas proporciones, Gini da una nueva definición de desigualdad[1], distinta de la que había propuesto en su trabajo de 1908. También, volverá a su medida δ de desigualdad introducida en 1908 y criticará el uso de índices de concentración construidos con distintos valores de k porque daban lugar a concentraciones entre ellos. De ahí que proponga el llamado índice de concentración medio.

En esta memoria de 1910, Gini relacionará su medida δ con la llamada α [2]de Pareto. Por medio de aproximaciones Gini afirmaba que $\delta = \frac{\alpha}{\alpha - 1}$, para $\alpha > 1$, del primer modelo de Pareto. A partir de esta relación, Gini interpreta el α en sentido contrario a como ño hace Pareto, es decir que cuando α disminuye entonces aumenta δ y esto significa, según Gini, un aumento de la desigualdad. Ahora bien, Gini señala que podría ocurrir que aumentarán a la vez δ y α , como en el caso de Amburgo, lo que sería una señal de que el modelo primero de Pareto ha fallado en este caso[3].

Otra consecuencia de la relación $\delta = \frac{\alpha}{\alpha - 1}$ es que α es poco sensible cuando sus valores están en el intervalo $(1,2)$ mientras que δ que tiene grandes variaciones. Sobre este asunto, Gini afirma que "Di quiera sorta presso molti l'idea, como vedenmono infondata, che la distribuzione della ricchezza fosse pressochè uguale in tutti i paesi e in tutti i tempe".

[1] La posible relación del concepto de desigualdad de Gini de 1908 con el artículo de Lorenz de 1905, sostenida por Zenga (1987) en su página 312, no es justificable, ya que Gini no volvió a citar dicho concepto de desigualdad en sus posteriores memorias.

[2] Veremos que este parámetro α era usado por Pareto en el sentido de que cuando "disminuía la desigualdad" entonces α disminuía.

[3] Por esto, es exagerada la insistencia en una visión única del comportamiento de α de Pareto que tanto repite Terencio Maccabelli (2002).

En 1912, Gini escribe la memoria "Variabilità e Mutabilità: contributo allo Studio delle distribución e delle relazioni statistiche" que publicará la Universidad de Cagliari.

Una posible justificación de porqué Gini escribió esta memoria sobre medidas de variabilidad estadística, es que su medida de desigualdad δ , elaborada a partir de proporciones acumuladas de población y renta, no mostraba directamente la dispersión entre las rentas individuales. Si Gini quería convencer de que la desigualdad en las rentas globales había aumentado, parece razonable que buscara medidas de dispersión existentes en la literatura estadística para probarlo.

Gini dedica gran parte de esta memoria al estudio de la medida de dispersión llamada media de las diferencias, que hoy llamamos media de las diferencias de Gini. Aporta varias fórmulas de esta dispersión, incluyendo una aplicable al caso de variables estadísticas continuas que le permitirá calcular su media de las diferencias en el modelo primero de Pareto. Gini obtendrá, para el modelo I de Pareto, que la media de las diferencias aumenta cuando α disminuye, con lo que zanjaba por fin la interpretación "errónea" de Pareto. Otro resultado fue la demostración de la relación $\delta = \frac{\alpha}{\alpha - 1}$. Por último, Gini prueba que su medida δ , que como sabemos daba lugar a la curva de Lorenz dual, era equivalente a la curva de Lorenz para el modelo I de Pareto. Este resultado le hizo ver que su medida δ era dependiente del modelo I de Pareto. Este resultado inesperado le llevó a "retroceder" al año 1908, ya que tuvo que emprender una nueva búsqueda de una medida de desigualdad que pudiera aplicarse a cualquier modelo estadístico.

En 1913, Gini se trasladó a la Universidad de Padua y después se traslada en 1927 a la de Roma.

Gini lograría su objetivo en 1914 con la memoria de "Sulla misura della concentrazione e della variabilità dei caratteri" que fue publicada en las Atti del Real Istituto Veneto di Science, Lettere ed Arti, año académico 1913-1914.

En el apartado dos de esta memoria, que Gini titula "Di una misura della concentrazione della distribuzione del carattere", recuerda nuestra interpretación. mencionada al final de la memoria de "Variabilità…", sobre las razones que indujeron a Gini a interesarse en la media de las diferencias.

Si hasta ese momento Gini había desarrollado su investigación estimulado por Pareto, la limitación de su medida δ le llevará a descubrir el artículo que M. O. Lorenz había publicado en 1905 sobre "Methods of measuring the concentration of Wealth". En su memoria, páginas 1230-1231, Gini menciona la parte del trabajo de Lorenz que le llevó a descubrir su medida de concentración R; dice así: "Questo metodo grafico presentava due incovenienti, debitamente riconosciceti de Lorenz e dal King (1912):

a) Di non dare una misura precisa della cocentrazione.

b) Di non permettere neppure, in certi casi, di giudicare dove o quando la concentrazione è più forte. Se, infatto, le curve di concentrazione si accavallano[4], non è sempre possibile dire se sia più accentuata l`una a l`otra.

Questo incovenienti, che si può ritenere trascurabile nei parragón tra la concentrazione dei fenomeni dello stesso genere (per es, di redditi di duo anni o di due paesi diversi), assume particolare gravità nei confronti tra la concentrazione di fenomeni differenti, per cui è diversa la forma della distribuzione".

Como respuesta a los inconvenientes, Gini propondrá el área de Lorenz (área entre la curva de Lorenz y el segmento de igualdad) como medida de concentración R .

A partir de aquí, recogemos en el parágrafo **1** una sucinta biografía de Corrado Gini; en el parágrafo **2** hacemos un resumen de las dos investigaciones de Pareto sobre leyes de los repartos de las rentas y de la desigualdad de las rentas. Es en el parágrafo **3** donde presentamos nuestros resultados sobre los trabajos que Gini hizo en el período 1908-1910; la memoria de 1912 la analizamos en el parágrafo **4** y, finalmente, en el parágrafo **5** estudiamos la memoria de 1914.

[4] Lorenz presentaba dos curvas continuas que se cortaban una vez.

1. PERFIL BIOGRÁFICO DE CORRADO GINI

Corrado Gini nació en Motta de Livenza (en la provincia de Treviso) en el Nordeste de Italia, el 23 de Mayo de 1884. Era hijo de unos ricos terratenientes. Murió en Roma el 13 de Marzo de 1965.

En 1905 se licenció en Derecho por la Universidad de Bolonia. Su tesis doctoral *Il sesso dal punto di vista statistico* (El sexo desde el punto de vista estadístico), que fue publicado en 1908, mostraba que sus inquietudes intelectuales no eran solo de naturaleza jurídica incluso "si su estudio de las leyes le proporcionó una cierta inclinación y capacidad para los argumentos sutiles tendiendo a someter los hechos a la lógica y, en consecuencia, a organizarlos y domeñarlos" (Castellano, V. (1965), p.3). En este trabajo discute las tasas de sexo en el momento del nacimiento. Comienza con una exposición de las teorías anteriores, a través de la información estadística existente formula nuevas hipótesis sugeridas por estos datos, y obtiene consecuencias que se derivan de esas hipótesis, construyendo una teoría que posteriormente contrasta con los datos estadísticos.

Durante su estancia en la Universidad recibió algunos cursos adicionales de matemáticas y biología. Con la base de esta formación desarrolló su carrera científica en dos direcciones principales: las ciencias sociales y la estadística.

Contempló la Estadística como un instrumento para la investigación científica y como un método para la utilización de técnicas en el largo proceso a través del cual el conocimiento avanza desde la percepción inicial hasta la observación y explicación. Como estadístico, sus intereses se situaron más allá de los aspectos formales de la construcción de modelos para llegar a las leyes que gobiernan los fenómenos biológicos y sociales en los campos de investigación que exploró durante 50 años.

Su carrera universitaria progresó con rapidez. En 1909 fue nombrado profesor temporal de Estadística en la Universidad de Cagliari y en 1910, a los 26 años, accedió a la cátedra de Estadística en la misma Universidad. En 1913 se trasladó a la Universidad de Padua y desde 1927 ocupó la cátedra de Estadística de la Universidad de Roma.

Durante y después de la primera guerra mundial se involucró cada vez más en los problemas económicos y sociales, derivados de la guerra y la reconstrucción del país, tales como pérdidas de la guerra, suministro de víveres, riqueza y renta nacional, inflación y depresión económica (Castellano, (1965)). Estos intereses estaban conectados con una brillante carrera: durante este tiempo fue consejero en calidad de experto del Gobierno Italiano en la Liga de las Naciones, y se dio a conocer internacionalmente. Entre 1917 y 1925 fue miembro de numerosos comités italianos e internacionales que trataban problemas de suministro de víveres, la medida de la riqueza y la renta en los estados miembros de la Liga de Naciones, trabajo, cuidados de la infancia, liquidación de las deudas de la guerra, etc. Pero Gini nunca permitió que esas ocupaciones le distrajeran de su trabajo teórico, de las cuales, por el contrario, pudo extraer nueva inspiración para sus intereses prácticos y fue guiado por ellos en direcciones fructíferas.

En el período 1926-1932 fue presidente del Instituto Central de Estadística (ISTAT) y en esos años organizó y coordinó los servicios de estadística en Italia, aportando un buen y eficiente nivel técnico a pesar de las diversas dificultades que tuvo que encarar. Su trabajo, relativo a la centralización de la información estadística, se encontró con los obstáculos erigidos por los burócratas que mantenían altas posiciones en varios ministerios y con la resistencia pasiva de la mayoría de los funcionarios públicos. De hecho la batalla fue tan dura que Mussolini, que en aquel tiempo era el Primer Ministro, tuvo que intervenir para sosegar las disputas. Esto deterioró las buenas relaciones que hasta entonces habían mantenido Mussolini y Gini. Además, las cada vez más frecuentes interferencias del régimen fascista eran una señal clara de que "los tiempos estaban cambiando y su espíritu impaciente e independiente no aguantaría más interferencias en su trabajo" (Castellano, V. (1965), p.7). Por todo ello, dimitió en 1932.

Por lo que se refiere al período en que Gini fue presidente del ISTAT, es preciso resaltar que aunque respetó al poder político, nunca se sometió a él. Por el contrario, se puede decir que

utilizó al régimen fascista para mejorar la organización y eficacia del servicio de estadística oficial (Leti, G. (1996)).

Mantuvo un interés considerable en la actividad editorial y en 1920 fundó *Metron,* una revista internacional de Estadística que fue editada hasta su muerte y, como la revista Biometrika bajo la dirección de Pearson, nunca aceptó artículos sin aplicaciones prácticas. Es precisamente esto lo que Gini pensaba que distinguía un trabajo de estadística teórica de un ejercicio en matemáticas, y para el estadístico las matemáticas son simplemente una herramienta en su trabajo y no un interés en si mismo. En 1934 fundó *Genus*, una revista del Comité Italiano para el Estudio de los Problemas de Población, institución que el fundó en 1929, y que dos años más tarde organizó en Roma el Primer Congreso sobre Población, seguido después de la segunda guerra mundial por una serie de congresos bajo el auspicio de las Naciones Unidas y de la Unión Internacional para el Estudio Científico de la Población. El programa de investigación del Comité se basaba en un concepto amplio de la Demografía, que incluía las relaciones de los fenómenos demográficos dentro de su entorno físico así como sus aspectos biológicos, económicos y sociales. El Comité sobrevivió a las dificultades de la posguerra gracias al interés que ponían en su trabajo, y la alta calidad del trabajo realizado hasta el comienzo de la guerra; las principales contribuciones a su crédito son la publicación de una serie de volúmenes de material fuente, *Fonti Archivistiche per lo studio dei problema della popolazione fino al 1943,* y las expediciones científicas para el estudio de grupos de población aislados, que Gini, como Presidente, organizó y dirigió. Aún hoy en día *Genus* sigue siendo le revista oficial del comité, que Gini fundó en 1934.

Otra revista que debió su origen a Gini fue *La Vita Economica Italiana*; esta revista fue fundada en 1926, registraba los desarrollos económicos diarios hasta el comienzo de la segunda guerra mundial y fue clausurada en 1943.

En 1936 fundó la facultad de Ciencias Estadísticas, Demográficas y Actuariales en la Universidad de Roma de la que fue decano hasta 1954. En 1955 fue nombrado Profesor Emérito.

Desde el punto de vista científico Gini tuvo varios intereses, además de la estadística, por ejemplo economía, sociología, demografía y biología, disciplinas que el consideraba estaban fuertemente interrelacionadas. En particular, concebía la población como el cuerpo de la sociedad hacia el que las contribuciones de otras disciplinas de las llamadas ciencias humanas deberían converger. Esta manera de pensar le permitió establecer los fundamentos teóricos de la sociología positiva que parte de la demografía y considera la sociedad como un organismo con un número de propiedades básicas que tiene en común con los organismos biológicos.

Desde el punto de vista estrictamente estadístico la principal contribución de Gini concierne a tres amplias áreas de estudio: valores medios, variabilidad y asociación entre variantes estadísticos. Se debe señalar que aunque el se hizo famoso por su razón de concentración (Giorgi, G.M. (1990)), también hizo otras importantes contribuciones científicas que desafortunadamente no siempre recibieron el reconocimiento que merecían.

De los trabajos de Gini podemos ver que nunca antepuso los aspectos formales a las ideas sustantivas. La mayoría de su trabajo metodológico surge de la necesidad de resolver problemas prácticos. En particular "el no se preocupó por el formalismo matemático complejo y los nombres altisonantes de algunos sabios extranjeros. En muchas ocasiones tuvo bastante éxito destruyendo una construcción muy elaborada mediante una simple nota...Aparte de raras excepciones consideraba a los matemáticos como gente extravagante que debía ser observada de cerca cuando tratan con un problema concreto" (Giorgi, G.M. (1996), p.14). En muchas ocasiones mostró que no estaba dispuesto a sacrificar las ideas sustantivas en aras de un formalismo matemático exagerado. En el campo de la economía por ejemplo su artículo de 1956 *Delusioni dell'econometria (Desengaños de la econometría)* creó una gran conmoción entre los econometras que tendían a exagerar los aspectos formales "porque son bien conocidas las dificultades de ir al grano en economía," y por esta razón "muchos veces un estudiante se ha sentido más cómodo al seguir el formalismo matemático o, como ocurre a menudo, realizar ejercicios matemáticos en vez de economía" (Georgescu-Roegen, N. (1966), p. 266).

2. LAS LEYES DE LOS REPARTOS DE LAS RENTAS Y EL CONCEPTO DE DESIGUALDAD DE PARETO

La dependencia de Gini de las investigaciones de Pareto es la razón por la que recogemos en este parágrafo los principales resultados que Pareto logró sobre las leyes que explican el reparto de las rentas, así como su concepto de desigualdad económica. Veamos, en primer lugar, ***las leyes de reparto de las rentas de Pareto.***

Se admite generalmente que la formulación de la "ley de Pareto" ha abierto nuevas perspectivas a la ciencia económica, las matemáticas y la estadística. Para comprender mejor esta ley es necesario situarnos en el contexto de la historia del pensamiento en el que se produce, así como las concepciones que animaban a Pareto en la elaboración de dicha ley.

Hasta 1894, el estudio del reparto de la riqueza y de las desigualdades económicas habían permanecido en un segundo plano en las investigaciones económicas y en las preocupaciones de los economistas. Era el problema de la producción el que se situaba en un primer plano y el del consumo como un problema accesorio, consecuencia del acto fundamental de la producción. Si bien se pueden encontrar entre los mercantilistas y algunos fisiócratas análisis del problema de la circulación bastante rigurosos, de los que se deduce la distribución del producto neto, se trata de leyes que determinan un reparto ideal en un estado de libre competencia perfecta. Aunque Adam Smith hace un desarrollo más realista del problema del reparto, no hay en su obra un desarrollo cuantitativo del mismo. Ricardo se pregunta en qué proporciones el producto total de la renta, de los salarios y de los beneficios debe ser repartido entre los propietarios, los capitalistas y los trabajadores (Stigler, 1946). Es preciso reconocer a la Escuela Histórica alemana y a los socialistas el situar en el primer plano de la investigación económica el problema de si es posible conciliar una distribución justa con una producción que crece de forma incesante, es decir si la evolución del capitalismo podría suscitar formas sociales donde la distribución de la riqueza fuera más justa y equitativa. Desde entonces el análisis económico no ha podido desentenderse de este problema, incluidos los economistas liberales.

Después de algunos titubeos e incertidumbres, los economistas liberales acabarán diciendo que los intereses del capital, del trabajo y de la propiedad acabarán por armonizarse de forma natural, en virtud del libre juego de la oferta y de la demanda.

Las concepciones que animaban a Pareto (Pareto, 1897, *Cours d'Economie Politique,* Tomo II, edición de Busino) hablando sobre el reparto de la riqueza afirma que "puede depender, de la naturaleza de las personas de las que se compone la sociedad, de la organización de ésta, y también, en parte, del *azar* (las *conjeturas* de Lasalle), es decir de ese conjunto de causas desconocidas tanto en un sentido, como en otro, a las cuales, en nuestra ignorancia de su verdadera naturaleza, damos el nombre de azar…Es la observación quien nos debe enseñar qué influencia tienen realmente esas causas en el reparto de la riqueza…"(obra citada, p. 304).

Para elaborar su ley, Pareto, toma como punto de partida la exploración de los datos de las declaraciones de los contribuyentes para el impuesto de la renta, y argumenta, que a pesar de la incertidumbre de esas declaraciones, es la base más segura de la que se dispone, después de discutir en los parágrafos precedentes (**950-956**) sobre la fiabilidad y adecuación de otros datos relativos a la renta. Esta labor de recogida de los datos se inicia en 1893, una vez que sucede a Leon Walras en la cátedra de Economía Política de la Universidad de Lausanne, y recopila datos de Francia, Inglaterra, Irlanda, Bélgica, Suiza, los Estados Unidos de América, Prusia... Tomando, en un principio, los datos de las declaraciones de la renta de dos países diferentes socialmente, como lo eran en aquella época Inglaterra e Irlanda (1893-94), razona, como se indica a continuación, con el objetivo de indagar si esas cifras se distribuyen al azar, o se agrupan siguiendo alguna ley.

Adoptando las notaciones de Pareto se indicará mediante x una cierta renta, medida en las correspondientes unidades monetarias del país de que se trate; se indicará mediante N_x el número de contribuyentes que tienen una renta igual o superior a x. En el plano cartesiano toma sobre el eje de abscisas los logaritmos de x y sobre el eje de ordenadas los logaritmos de N. Observa que los puntos así determinados $(\log x, \log N_x)$ tienen una tendencia muy

marcada a disponerse en línea recta, e incluso un hecho que Pareto considera más notable: que esas rectas son todas ellas casi paralelas. La ecuación de esta recta puede ser representada por

(2.1) $$\log N_x = \log A - \alpha \log x \ ,$$

lo que da origen a su primera ley de reparto

(2.1.1) $$N_x = \frac{A}{x^{\alpha}}$$

Para estimar los parámetros A y α utiliza una interpolación de Cauchy en (2.1), que dice ser suficiente en una primera aproximación, o incluso un método gráfico mediante papel milimetrado. Añade que la ecuación general de la curva quizás debería ser

(2.2) $$\log N_x = \log A - \alpha \log(a + x) - \beta x \ ,$$

pero es solamente para el caso del ducado de Oldenburg para el que encuentra un valor de β que no es despreciable y, en general, se tiene simplemente

(2.3) $$\log N_x = \log A - \alpha \log(a + x) \ ,$$

que define su segunda ley de reparto

(2.3.1) $$N_x = \frac{A}{(x + a)^{\alpha}}$$

Cuando se trata de la renta total, α es también, en general, despreciable, muy pequeño en la mayor parte de los casos, del orden de los errores de observación, lo que conduce a la ecuación (2.1). Cuando se trata de rentas mobiliarias el parámetro α no debe ser despreciado, incluso puede adquirir valores bastante considerables.

Una vez establecidas las fórmulas anteriores, Pareto se dedica a contrastar las fórmulas con otros datos y llega a las siguientes conclusiones:

a) Para los casos examinados de Prusia (1876, 1881, 1886, 1893-94, 1894-95), Sajonia (1880, 1886)…, además de los ya citados de Inglaterra y Holanda, la recta logarítmica se ajusta muy bien, en el sentido de que las desviaciones entre los logaritmos de N_x teóricos y los correspondientes logaritmos empíricos son muy pequeñas.

b) Las rectas de los logaritmos empíricos y las rectas ajustadas son casi paralelas. De hecho, observa Pareto, los valores calculados de α mediante interpolación (la pendiente de las rectas) varían poco, entre 1,13 (Augsburg ,1526) y 1,89 (Prusia, 1852)

c) De las consideraciones a) y b), concluye Pareto : "Es absolutamente imposible admitir que sean debidos solamente al azar. Hay ciertamente una *causa* que produce la tendencia de las rentas a disponerse siguiendo una cierta curva. La forma de esta curva parece depender de una forma bastante débil de las diferentes condiciones económicas de los países considerados, puesto que los efectos son mas o menos los mismos para países cuyas condiciones económicas son tan diferentes como las de Inglaterra, Irlanda, Alemania, ciudades italianas, e incluso Perú" (obra citada, §**960**, p.312).

d) No obstante, el carácter general de la afirmación anterior queda matizada y valorada en su justo término con la siguiente consideración: "…cuando se trata de leyes puramente empíricas, se debe de ser extraordinariamente prudente. En cualquier caso, las consecuencias que vamos a extraer de esta ley serán siempre válidas, al menos, para los pueblos para los cuales hemos visto que se verifican".

Corrado Gini partirá de la primera ley de los repartos de Pareto para llegar a definir su medida de desigualdad $I_{\delta} = \delta$. Gini relacionará su δ con el α de los modelos de Pareto, señalando que α *es muy poco sensible frente a* δ , es decir dudaba de que los modelos de Pareto fueran válidos para cualquier país y época.

Veamos a continuación los principales resultados que Pareto obtuvo en su investigación sobre ***el concepto de desigualdad.***

En la página 318 de la edición de Busino, parágrafo **964**, de su Curso de Economía Política, tomo II, publicado en Lausanne en 1897, Pareto propone su primera definición de "desigualdad" sobre las fortunas:

Definición I

"La disminución de desigualdad será, por tanto, definida por el hecho de que el número de pobres va *disminuyendo* en relación al número de ricos, o lo que es igual, en relación al número total de miembros de la sociedad. Es este el sentido que parece haber prevalecido, es por tanto el que nosotros adoptaremos".

Esta definición fue interpretada por Constantino Bresciani-Turroni (1882-1963), en su trabajo de 1939, como la proporción de la población que "no alcanza un mínimo decente de nivel de vida" (página 132), lo que correspondería a la llamada *línea de pobreza.*

Definición II

"En general, mientras que el número de personas que tienen una renta inferior a x *disminuye*[5] en relación al número de personas que tienen una renta superior a x , diremos que la desigualdad de la renta disminuye".

A pie de página, Pareto aclara su definición II de la forma siguiente. Llamando N_x al total de individuos con rentas mayores o iguales a la renta x y N_h el total del individuos con rentas mayores o iguales a la renta mínima, h , considerada, Pareto define la función siguiente,

$$u_x = \frac{N_x}{N_h} \qquad (2.4)$$

que es la proporción de individuos con rentas iguales o superiores a x. A partir de la función (2.4), Pareto interpreta sus definiciones diciendo que la ***desigualdad disminuye*** cuando el cociente u_x *aumenta* para todo $x \geq h$.

Bortkiewicz (1931) fue quien aportó una interpretación del concepto de desigualdad que Pareto recoge en su Definición II. Bortkiewicz observó que la expresión (2.4) es el complemento de la función de distribución de una variable aleatoria no negativa, también llamada *función de supervivencia*, (que escribimos como $1 - F(x)$, donde $F(x)$ es la función de distribución), viendo entonces que el criterio (2.4) de Pareto exigía que una de las funciones, que sería la de menor desigualdad, no fuera mayor de la otra para toda renta no negativa. Esta interpretación que hizo Bortkiewicz (1931) del criterio de "disminución de la desigualdad" de Pareto corresponde hoy al criterio de *Dominancia Estocástica de GradoI* [6](Paul D. Thistle (1989)). Una primera consecuencia del criterio de *Dominancia Estocástica de Grado I* es que si $F_2(x)$ exhibe dominancia estocástica de grado I sobre $F_1(x)$, entonces la rentas medias, μ_2 y μ_1 de los modelos $F_2(x)$ y $F_1(x)$ respectivamente, verifican que $\mu_2 \geq \mu_1$[7].

En la parágrafo **965** del curso, Pareto dice que: "una disminución de la desigualdad no puede producirse más que si el total de las rentas crece más rápidamente que la población". Es decir que la "disminución de la desigualdad" implica que la renta media debe aumentar. Una consecuencia probada por Bortkiewicz (1931). Vemos que el criterio de Pareto está comparando dos repartos de las rentas donde el reparto que tenga una renta media mayor será una condición necesaria para que según Pareto "sea la de menor desigualdad".

---

[5] En el original aparece "aumenta" como un error tipográfico. Este error fue corregido en la edición de 1919 del Manuale di Economia Politica (página 371).

[6] **Definición III**

Diremos que la función de distribución $F_2(x)$ exhibe dominancia estocástica de grado I sobre $F_1(x)$ siempre y cuando $F_2(x) \leq F_1(x)$para todo x no negativo, siendo la desigualdad estricta par algún valor de x para evitar que coincidan.

[7]Cuando la función $F(x)$ tiene esperanza matemática finita, su esperanza matemática es

$$E[X] = \int_0^\infty (1 - F(x))dx ,$$

que, al igual que probó Bortkiewicz, si se verifica, $F_2(x) \leq F_1(x)$ entonces $1 - F_2(x) \geq 1 - F_1(x)$, lo que conduce a que la renta media aumente al pasar de $F_1(x)$ a $F_2(x)$ ya que la definición II exige que la desigualdad sea estricta para algún valor de x.

¿Cómo entender que la desigualdad disminuye cuando lo que ocurre es que aumenta siempre la renta media?

En 1896-97 se establece una controversia entre Wilfredo Pareto y George Sorel. Esta controversia puede seguirse en el capítulo décimo de Marc Barbut (2007). Aquí nos interesa la discusión acerca del criterio de "disminución de la desigualdad".

En Sorel (1897, p. 586), éste nos dice que "M. Pareto écrit à ce sujet: Quelle est la vraie signification des termes: moindre inégalité des revenus, ou de ceux qu'on emploie à peu près dans le même sens; moindre inégalité des fortunes, moindre inégalité des conditions? S'il s'agissait d'une égalité complète des revenus, des fortunes ou des conditions, il n'y aurait pas d'équivoque posible. Mais on peut se rapprocher de cet état de deux manières essentiellement différentes: on s'en rapproche, aussi bien si les riches deviennent pauvres que si les pauvres deviennent riches. Ces sont deux phénomènes différents et, s'il s'agissait d'une science positive, où les faits sont tout et les mots rien, on n'hésiterait pas à designer par des termes different des choses aussi différentes. Mais l'économie politique n'est encore, souvent, qu'un genre de littérature. On y donne, aux dépens des faits, une grande importance aux mots. On se dispute donc pour savoir auquel des deux phénomènes indiqués doit être réservée la dénomination de moindre inégalité des fortunes." (Pareto, 1897, Tomo II, pp. 318-319)

Sobre la definición II de Pareto, Sorel (1987, pp. 586-587) dice "Tel lui paraît être ce qui ressort de plus clair des dissertations des écrivains les plus notables: mais il ne se fait peu d'illusion sur la valeur de cette définition; il m'écrit (en carta de Pareto a Sorel) "Le terme *moindre inégalité des conditions* ne me semble pas, plus qu'a vous indiquer la chose qu'on a en vue. J'y aurais volontiers substitué un autre terme; mais on m'a déjà assez reproché la malheureuse ophélimité et j'ai voulu éviter de nouvelles logomachies à mes critiques. Au moins, j'ai tenu à bien expliquer ce qu'était la chose désignée par ce terme". Y sigue diciendo Sorel que "Il aurait bien mieux valu employer une expresión neutre que celle que M. Pareto a empruntée aux littérateurs de l'économie politique. En effect, on arrive à cette conclusión que l'inégalité des conditions va en décroissent au fur et à mesure que se développe le capitalisme; on a beau dire qu'on a en vue seulement ce qui est exprimé dans la définition précédente, on ouvre la voie aux sophismes des apologistes du capitalisme". Y termina diciendo que "Je crois qu'il est préférable de donner une interprétation qui sois davantage le mot à mot des phénomènes observés".

Pareto no modificará la expresión de "disminución de la desigualdad" como puede verse en la edición de 1919 del Manuale di Economia Politica (pp. 371-372). Será Enrico Barone (1859-1924) el que modifique la expresión "disminución de la desigualdad". Tomando como ejemplo la primera ley de Pareto, Barone (1908-1909, p. 264) llama al parámetro α "convergencia", y como este parámetro disminuye cuando "disminuye la desigualdad", entonces el criterio de Pareto, basado en (2.4), es expresado, en la notación de Barone, como que "disminuye la convergencia".

Exponemos a continuación una posible explicación del criterio (2.4) de Pareto. Hoy sabemos que la dominancia estocástica de grado I implica la dominancia estocástica de grado II[8]. Ahora si $F_2(x)$ exhibe *segundo grado de dominancia estocástica* $F_1(x)$, de la proposición-4 de P.D. Thistle (1989) debe ocurrir que $G_2(p) \geq G_1(p)$ para $0 \leq p \leq 1$, donde $G_i(p), i = 1,2$ es la curva de Lorenz generalizada, esto es $G_i(p) = \mu_i L_i(p)$, donde $L_i(p)$ es la

[8] Si $F_2(x)$ exhibe *primer grado de dominancia estocástica* sobre $F_1(x)$ y si existen las esperanzas matemáticas μ_2 y μ_1, entonces se cumple las desigualdades siguientes:

$$S_2(x) = \int_0^x F_2(t)dt \leq \int_0^x F_1(t)dt = S_1(x)$$

para todo x no negativo.

La desigualdad anterior define que $F_2(x)$ exhibe *segundo grado de dominancia estocástica* sobre $F_1(x)$ cuando $S_2(x) \leq S_1(x)$ para todo *x* no negativo.

curva de Lorenz. La curva de Lorenz generalizada fue introducida por Shorrocks (1983) y Kakwani (1984). Las curvas de Lorenz generalizadas consideran a la vez cambios en las rentas medias (esperanza matemáticas) y en la desigualdad. Para apreciar mejor estos últimos resultados Shlomo Yitzhaki (1982) probó lo siguiente:

Los criterios de dominancia estocástica de grado I y II implican las dos consecuencias siguientes: (a) $\mu_2 \geq \mu_1$ y (b) $\mu_2 - \Gamma_2 \geq \mu_1 - \Gamma_1$, donde $\Gamma_i, i = 1,2,$ es la mitad de la media de las diferencias, con repetición, de Gini.

El resultado de Yitzhaki implica que el criterio de Pareto de "disminución de la desigualdad" llevará a elegir siempre una distribución de la renta con mayor renta media y que además tenga mayor valor de la diferencia entre la renta media y la mitad de las media de las diferencias de Gini.

La parte (b) del criterio de Yitzhaki puede también escribirse como $\mu_2(1 - R_2) \geq \mu_1(1 - R_1)$, donde $R_i, i = 1,2,$ es la razón de concentración de Gini. Vemos ahora que el criterio (b) pesa la renta media con $(1 - R)$ y así le impone una restricción al criterio (a). Cuando R disminuye entonces pesa más la renta media, lo que la hace aumentar, mientras que si R aumenta entonces pesa menos la renta media, lo que la hace disminuir.

Volviendo a Pareto. Si como hemos visto la expresión "disminución de la desigualdad" fue un problema; cuando Pareto aplicó su criterio de desigualdad a la primera y segunda ley de los repartos de la rentas, surgió un nuevo problema que creó mayores rechazos entre los críticos de Pareto.

Cuando Pareto aplicó su criterio de desigualdad a su primera ley de los repartos obtuvo lo siguientes: la "disminución de la desigualdad" equivale a que disminuya α o aumente la renta mínima h, en el sentido de que cuando un parámetro es constante el otro debe disminuir (caso de α) o aumentar (caso de h). Un resultado, un poco más complicado, fue obtenido para la segunda ley de los repartos de las rentas. En general Pareto probó que su criterio de "disminución de la desigualdad" conducía a la disminución del parámetro α en sus dos modelos de repartos. Pronto surgieron las primeras críticas.

Sorel (1897, p. 601), en la discusión que mantuvo con Pareto en 1896-97, concluyó que en el modelo I de Pareto, la desigualdad disminuía siempre que el parámetro α aumentaba. Sorel probó este resultado, contrario al de Pareto, usando distintos criterios de desigualdad, como el basado en el cociente entre la renta mínima y la renta media que equivale al cociente $\frac{\alpha - 1}{\alpha}$, viendo que cuando α disminuye entonces la parte de la renta minina en la renta media disminuye.

Rodolfo Benini publicó en 1897 un trabajo donde en la página 194 afirma los siguiente: "Come già dicemmo, quanto più grande è α, tanto più rapido è il descrescere dei possessori, o contribuenti in relazione al creceré del reddito". Una explicación de esta afirmación de Benini puede hacerse si tomamos la tangente de la función de supervivencia para el valor h de la renta mínima de la primera ley de los repartos de Pareto, resulta la expresión $-\frac{\alpha}{h}$, que disminuye siempre que α aumente o que la renta mínima disminuya, y así van disminuyendo los individuos con rentas cada vez más altas, lo que conduce a que las rentas se aproximen a la renta mínima, lo explica la disminución de la desigualdad. En 1906, Benini volveria a afirma que "Quanto maggiore è α, tanto più rapido è il diminuire dei titolari ad ogni gradito della scala del reddito, ciè tanto meno disuguale è la ripartizione delle ricchezze".

En 1908, Gini observará empíricamente un comportamiento inverso entre su δ y el α de Pareto, en el sentido de que cuando α disminuye entonces δ aumenta, lo que conduce a que *aumente la desigualdad,* frente a la expresión de Pareto de "disminución de la desigualdad". En 1910 Gini probará, usando una aproximación, que la relación entre su δ y el α de Pareto es de

la forma $\delta = \frac{\alpha}{\alpha - 1}$ lo que le reforzará en su creencia de que Pareto había caído en un "error" de interpretación. No será hasta 1912 cuando Gini aporte una demostración rigurosa de la relación entre su δ y el α de Pareto. Ahora bien, fue en 1912 cuando Gini descubre que su medida de desigualdad δ es dependiente del modelo I de Pareto, al igual que el α de Pareto. Así, Gini tuvo que comenzar de nuevo a buscar una medida de desigualdad que sirviera para cualquier modelo estadístico, lo que lograría en 1914 con su razón de concentración R .

3. EL ÍNDICE DE CONCENTRACIÓN δ DE C. GINI (1909-1910)

Nos proponemos en este parágrafo analizar los trabajos publicados por Gini en el período 1909-1910, que se refieren a la exposición de su concepto de concentración y la elaboración de un índice de concentración: *Il diverso accresimento delle classi sociali e la concentrazione della richezza*, Gini (1909); *Indici di concentrazione e di dependenza*, Gini (1910).

El primero de esos trabajos, Gini (1909), estudia el distinto crecimiento de las clases sociales y la concentración de la riqueza. Nosotros estudiaremos la segunda parte de este trabajo que se refiere a la exposición crítica de las distintas concepciones de la concentración de la riqueza, la definición de su propio concepto de concentración y la elaboración de un índice que mida la concentración, tomando como punto de partida la ley de V. Pareto que hemos expuesto en el parágrafo anterior, (Gini, 1909, pp. 69-83).

Por lo que se refiere a las distintas concepciones de la concentración de la riqueza, Gini (1909, p. 69) distingue dos:

1) "Algunos autores suponen que la distribución de la riqueza es tanto más desigual cuanto más pequeño es el número de los ricos en relación al número de los pobres. Estos autores se basan probablemente en la consideración sicológica de que la desigualdad de la riqueza es más sentida cuando pocos son los afortunados". Se observa que Gini está aludiendo, sin citar, a la concepción mantenida, entre otros, por Pareto en la Definición I y no en la Definición II (donde Pareto relaciona la población con la renta detentada por esa población).

2) "Otros autores entienden por el contrario que la riqueza será tanto más desigual cuanto más grande es el número de ricos en relación al número de pobres. La condición más próxima a una igualdad absoluta, dicen, sería aquella en la que hubiera una sola persona rica y todas las demás tendrían el mismo patrimonio (o la misma renta)".

Ambas definiciones le parecen imperfectas, así como todas aquellas que tengan en cuenta un único factor de la distribución de la riqueza, ya sea éste el número de ricos o de pobres o, únicamente, el montante de sus patrimonios o rentas. Por ello, Gini considera que una definición adecuada de la desigualdad o de la concentración debe tener en cuenta ambos factores: la relación entre el número de ricos y pobres y la relación entre el montante de la riqueza de los ricos y el montante de la riqueza de los pobres. Gini (1909, pp. 69-70) propone la siguiente definición para la concentración de la riqueza:

"La concentración de la riqueza en un país o en una época A es mayor que en un país o una época B , cuando la parte de la población que posee una parte $1/x$ de la riqueza nacional sea en A menor que en B o, viceversa, cuando la parte de la riqueza poseída por una parte de la población $1/y$ sea en B menor que en A "[9]. Es decir, si denominamos p_A la fracción de la población de A , q_A la fracción riqueza que posee esta fracción de la población y p_B la fracción de la población de B , q_B la fracción riqueza que posee esta fracción de la población,

[9] Esta definición no es utilizada posteriormente por Gini. Además en algunos casos no es aplicable, como en el ejemplo que se muestra. En efecto, sean dos poblaciones con distribuciones de renta, por ejemplo $A = \{4,6,10\}$, $B = \{4,6,20\}$ se observa que de $q_A\{4,6\} = q_B\{4,6\}$ implica que $p_A = p_B$ y $p_A\{4,6\} = p_B\{4,6\}$ implica que $q_A = q_B$.

entonces la concentración de la riqueza es mayor en A que en B siempre y cuando se verifique:

1) Si $q_A = q_B = 1/x$, entonces $p_A < p_B$

2) Si $p_A = p_B = 1/y$, entonces $q_A > q_B$

Después de esta definición, Gini trata de determinar un índice apropiado que de cuenta de la relación entre $1/x$ y $1/y$ razonando como sigue. Sea N_x el número de individuos que tiene una renta mayor o igual x , A_x el montante total de su renta y x_0 la renta mínima. Siguiendo a Pareto en su artículo *Aggiunta allo studio sulla curva delle entrate* (Cours d'économie politique, Tomo III, cap. 1) establece la siguiente relación

(3.1) $$\log N_x = \log K - \alpha \log x \ , \qquad x \geq x_0 > 0$$

lo que se puede escribir en la forma

(3.2) $$\frac{N_x}{N_0} = \left(\frac{x_0}{x}\right)^{\alpha}$$

donde el parámetro α es mayor que cero y *K* es una constante positiva que depende de α y de la renta mínima x_0. Los parámetros K y α son determinados en cada caso según los datos disponibles. Es decir, aplica la fórmula de Pareto al número de individuos con una renta superior a una renta dada.

La novedad que introduce Gini es la aplicación de la misma ley de Pareto, pero esta vez al montante de la renta poseída por esos individuos, a saber

(3.3) $$\log A_x = \log H - \beta \log x \ , \qquad x \geq x_0 > 0$$

lo que se puede escribir en la forma

(3.4) $$\frac{A_x}{A_0} = \left(\frac{x_0}{x}\right)^{\beta}$$

donde el parámetro β es mayor que cero y *H* es una constante positiva que depende de β y de la renta mínima x_0. Los parámetros H y β son determinados en cada caso según los datos disponibles.

Puesto que el objetivo de Gini es relacionar la tasa de individuos con renta superior a una renta dada y la tasa del montante de la renta que poseen, combina las fórmulas (3.2) y (3.4) para obtener

(3.5) $$\frac{N_x}{N_0} = \left(\frac{A_x}{A_0}\right)^{\delta} \ , \qquad \delta = \frac{\alpha}{\beta}$$

Es a este índice, que denominaremos $I_\delta = \delta$, al que Gini llama *índice de concentración*.

En cuanto a las propiedades de este índice se pueden enumerar las siguientes:

1) En la ecuación (3.5), al ser la proporción de renta mayor que la proporción de población correspondiente se deduce, salvo en los casos extremos, que $I_\delta \geq 1$.

2) Cuando $I_\delta = 1$ se produce la igualdad, es decir la proporción de población coincide con la correspondiente proporción de renta y, por ello, todos los individuos tienen la misma renta (igual a la renta media).

3) A medida que I_δ aumenta debe aumentar la concentración de la renta. Esto se debe a que la proporción de la población está más alejada de la proporción de renta y así debe aumentar δ para lograr la identidad de la fórmula (3.5).

4) Si en la ecuación (3.5) expresamos la proporción de renta en función de la proporción de población, se obtiene

(3.6) $$\frac{A_x}{A_0} = \left(\frac{N_x}{N_0}\right)^{\frac{1}{\delta}} \ ,$$

que es la función de Lorenz dual. Desconocemos si en ese momento Gini conocía la obra de Lorenz.

Para ***la estimación del índice*** δ ***a partir de los datos empíricos***, Gini debe ajustar la fórmula (3.1) para obtener una estimación de α y ajustar la fórmula (3.3) para obtener una estimación de β. A partir de (3.5) obtiene δ, es decir debe ajustar dos modelos. Estos cálculos son bastante laboriosos y costosos, además no aseguran la identidad de (3.5) en todo el recorrido de la renta x. Si estos ajustes no son buenos, las mencionadas propiedades del índice δ pueden dar resultados anómalos.

En la memoria de Gini de 1910, ***Indici di concentrazione e di dependenza***, se propone "encontrar índices de distribución y de relaciones entre los fenómenos cuantitativos, dotados de suficiente sensibilidad y aplicabilidad a las estadísticas usuales sin cómputos excesivamente laboriosos y sin hipótesis excesivamente alejadas de la realidad" (Gini, 1910, p.5). Con otro enfoque distinto al expuesto en 1909 aborda la tarea de construcción de estos índices y propone su conocido índice de concentración. Para ello, dice (obra citada, p.5), "es necesario establecer con claridad qué es lo que se entiende por concentración, y cómo se puede dar una medida de ella" y da la siguiente definición de concentración: "…habitualmente decimos que una sustancia está tanto más concentrada cuanto mayor es la cantidad de ella contenida en un espacio dado o, dicho de otro modo, cuanto menor es el espacio que de ella contiene una cantidad determinada…Entre los fenómenos que someteremos a examen están algunos para los que habitualmente hablamos de concentración. Se dice, por ejemplo que, en una población, la riqueza está muy concentrada cuando buena parte de esta está depositada en pocas manos".

En general, si imaginamos una población de n individuos ordenados según un carácter cuantitativo dado, X, de menor a mayor, y expresamos mediante x_i la intensidad del carácter X para la persona singular $i = 1,2,\dots,n$

$$x_1 \leq x_2 \leq \cdots \leq x_n ;$$
$$\frac{B_n}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

expresará el carácter medio. Elijamos ahora las últimas m personas de la ordenación; éstas presentarán el carácter en cuestión con la mayor intensidad; $x_{n-m+1}, x_{n-m+2}, \dots, x_n$ expresarán la intensidad creciente (o al menos no decreciente) del carácter del grupo así elegido,

$$\frac{B_m}{m} = \frac{x_{n-m+1} + x_{n-m+2} + \cdots + x_n}{m}$$

indicará el carácter medio. En razón de la ordenación de menor a mayor, se verifica

(3.7) $$\frac{B_n}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n} < \frac{x_{n-m+1} + x_{n-m+2} + \cdots + x_n}{m} = \frac{B_m}{m}$$

y por tanto

(3.8) $$\frac{x_{n-m+1} + x_{n-m+2} + \cdots + x_n}{x_1 + x_2 + \cdots + x_n} > \frac{m}{n}$$

Cuánto más fuertes sean las desigualdades (3.7) y (3.8) diremos que es mayor la concentración del carácter. Si ahora definimos las siguientes distancias,

$$d_m = \frac{B_m}{m} - \frac{B_n}{n}, m = 1,2,\dots,n,$$

se deduce de las consideraciones anteriores que:

(3.8.1) $$\frac{B_n}{n} - x_k = (n-k)d_{n-k} - (n-k+1)d_{n-k+1}$$

La expresión (3.8.1) determina la distancia entre la media aritmética y cualquier valor de la variable, como una combinación lineal de las distancias definidas. Puesto que n y la media aritmética son constantes, se observa a partir de la desigualdad (3.7) que:

a) Si las desigualdades (3.7) se convierten en igualdades, es decir si las distancias $d_k, k = 1,2,\dots,n-1$ son nulas, entonces de (3.8.1) se deduce que todos los valores de la variable coinciden con la media aritmética, y nos encontramos con una situación de igualdad.

b) Si las desigualdades (3.7) son máximas, es decir las distancias d_m toman los valores,

$$d_m^- = \frac{B_n}{m} - \frac{B_n}{n},$$

de (3.8.1) se sigue que $x_k = 0, k = 1,2,\ldots,n-1$ y $x_n = B_n$, es decir una situación de máxima concentración.

Se trata ahora de buscar, entre los dos miembros de una y otra desigualdad, una relación que valga cualquiera que sea m y que se pueda expresar mediante una sola constante. Tal constante se denominará *índice de concentración*. Será conveniente determinarla de tal modo que sea más alta cuanto mayor sea la concentración del carácter. La relación podrá asumir diversas formas según los distintos caracteres examinados. Para algunos caracteres (por ejemplo, renta global, renta profesional, salario, etc.) se encuentra la relación

$$\left(\frac{x_{n-m+1} + x_{n-m+2} + \cdots + x_n}{x_1 + x_2 + \cdots + x_n}\right)^{\delta} = \frac{m}{n} \tag{3.9}$$

Para otros la relación es de la forma

$$1 - \frac{x_{n-m+1} + x_{n-m+2} + \cdots + x_n}{x_1 + x_2 + \cdots + x_n} = \left(1 - \frac{m}{n}\right)^{\delta} \tag{3.10}$$

o sea

$$\frac{x_1 + x_2 + \cdots + x_{n-m}}{x_1 + x_2 + \cdots + x_n} = \left(\frac{n-m}{n}\right)^{\delta} \tag{3.11}$$

Para las "*rentas globales*", vale la fórmula

$$\left(\frac{x_{n-m+1} + x_{n-m+2} + \cdots + x_n}{x_1 + x_2 + \cdots + x_n}\right)^{\delta} = \frac{m}{n};$$

δ es *el índice de concentración. Éste es el exponente al que hay que elevar la fracción de las* m *rentas más altas para obtener la fracción de los individuos que las poseen".*

Por lo que se refiere a ***la estimación del parámetro*** δ, Gini propone dos métodos:

a) Calcular para cada valor de m el valor de

$$\delta_m = \frac{\log\left(\frac{m}{n}\right)}{\log\left(\frac{A_m}{A_n}\right)}, \qquad m = 1,2,\ldots,n-1$$

y toma como δ la media aritmética de los δ_m. A esta estimación de δ, Gini lo denomina ***índice de concentración medio***.

b) Ajustar el modelo $\log N_x = \delta \log A_x - \alpha \log K$, donde N_x es el número de individuos cuya renta es mayor o igual a x, A_x es el montante total de la renta de esos individuos. Esta ecuación se deduce de (3.9).

C. Gini después de varios trabajos empíricos sobre la distribución de las rentas en distintos Estados y en distintos momentos del tiempo, llega a la conclusión de que "*la distribución de las rentas globales a lo largo del tiempo y de Estado a Estado es enormemente diversa"* (los valores de δ en los casos estudiados por Gini varían entre 1,6 y 6). Y "que esta conclusión está en notable contraste con lo que para muchos constituye, los pensadores más notables, el más inesperado de los resultados de la investigación de lo que se llama *economía inductiva"* (Gini, 1910, p.40). De hecho, sigue comentando, "Pareto, tratando las series de las rentas con el método matemático, llega a la conclusión de que la distribución de las rentas es casi idéntica para todos los Estados y a lo largo del tiempo. Lo que trae como consecuencia de capital importancia, que la distribución de la renta de un Estado es independiente o casi de su constitución económica. Tal conclusión está basada en el hecho de que el índice de la distribución de las rentas globales de Pareto, α, difieren poco de lugar a lugar y de tiempo en tiempo, oscilando, en los casos estudiados por Pareto entre 1,50 y 1,13".

Por ello propone estudiar la relación que existe entre el índice de Pareto y el índice por él propuesto. Así, partiendo de la fórmula de Pareto en su expresión logarítmica (2.1), demuestra

(obra citada, pp. 41-42) mediante algunas aproximaciones basadas en cálculos empíricos e integración que

$$\delta = \frac{\alpha}{\alpha - 1} \qquad (3.12)$$

La expresión (3.12) nos da la relación teórica que existe entre el índice de concentración de Gini δ y el parámetro α de la distribución de Pareto de tipo I, en el supuesto de que los valores de la variable sigan esta distribución. Más adelante (obra citada, p. 42) vuelve a insistir en el carácter teórico de esta relación y en su carácter aproximado, consecuencia del método que utiliza en su demostración, a saber: aproximaciones empíricas y paso del caso discreto al continuo al usar la integración

Se debe observar que de (3.12) se deduce que cuando $\alpha \to \infty$ entonces $\delta \to 1$ y cuando $\alpha \to 1$ entonces $\delta \to \infty$. Con carácter general, el sentido de variación de estos dos parámetros es inverso el uno del otro.

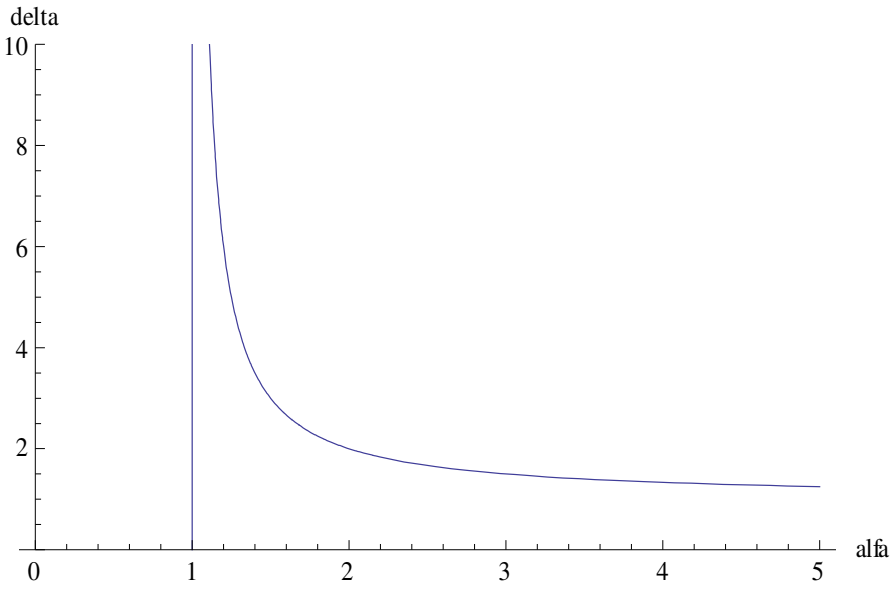
Esta relación, entre el índice de las rentas propuesto por Pareto y el índice de concentración, es la que permite a Gini examinar si los resultados de Pareto autorizan realmente a llegar a la conclusión de la uniformidad de la distribución de la riqueza. Para ello en la tabla que sigue Gini expone los valores de α (columna 3) calculados por Pareto y los valores de δ (columna 4) que se corresponden teóricamente de acuerdo con (3.12).

Tabla XXI de la obra citada, p. 43.

1	2	3	4
Estado	Año	Valores de α	Valores teóricos de δ
Prusia	1852	1,89	2,12
Prusia	1876	1,721	2,39
Prusia	1881	1,726	2,38
Prusia	1886	1,679	2,47
Prusia	1890	1,60	2,67
Prusia	1894	1,598	2,67
Sajonia	1880	1,58	2,72
Sajonia	1886	1,51	2,96
Hamburgo	1891	1,1308	8,64
Bremen	1891	1,1814	6,51
Schaumburg-Lippe	1893	1,2983	4,35
Sajonia-Weimar	1892	1,6305	2,59
Sajonia-Meiningen	------	1,7342	2,36

Es necesario observar que pequeñas diferencias en los valores de δ se corresponden con grandes diferencias en las distribuciones de la renta y que a pequeñas variaciones de α se corresponden mayores variaciones de δ: en la tabla, los valores de α, calculados por Pareto, varían entre un valor mínimo de 1,13 y un valor máximo 1,89, mientras que los valores de δ teóricos, obtenidos mediante (3.12) a partir de los correspondientes valores de α, oscilan entre 2,12 y 8,64.

Para una mejor comprensión de la relación entre los parámetros alfa y delta, mostramos la gráfica de la función (3.12) donde se puede observar que cuando α varía entre 1 y 2 el parámetro δ varía desde infinito a 2, es decir en esta variación el parámetro δ es mucho más sensible que α; mientras que cuando α varía entre 2 a infinito el parámetro δ varía desde 2 hasta 1, es decir en esta variación el parámetro α es mucho más sensible que δ.



4. LA MEDIA DE LAS DIFERENCIAS DE C. GINI (1912)

De la memoria de C. Gini (1912), *Variabilitá e mutabilitá: contributo allo Studio delle distribuzioni e delle relazioni statistiche,* nos vamos a interesar por la parte dedicada a los Índices de Variabilidad y, en particular, nos centraremos en la justificación que hace Gini para la introducción de la medida de variabilidad, hoy conocida como *Media de las Diferencias de Gini*.

La justificación para la búsqueda de un nuevo índice de variabilidad aparece en (Gini, 1912, § **11-13,** pp.225-228) y dice que "el estudio de la variación de un carácter se puede hacer para dos categorías de caracteres, que son notablemente distintas".

"Unos primeros caracteres cuya intensidad observada se nos presenta con diferentes modalidades cuantitativas y donde las causas de error en la medida de las observaciones son debidas a causas accidentales o sistemáticas. Por ejemplo, las sucesivas medidas de la altura de una montaña difieren más o menos entre sí por causas particulares atribuibles a los instrumentos de medida o a las tendencias personales del observador (causas sistemáticas de error) o de circunstancias naturales que originan intensidades extremadamente variables de ocurrencia imprevisible, como las bruscas variaciones del tiempo atmosférico, las oscilaciones del terreno, etc. (causas accidentales). La mayor parte de las observaciones de las ciencias físicas tienen por objeto, con bastante frecuencia, caracteres de esta primera categoría". Por otra parte, "existe una segunda categoría de caracteres que a su vez presentan, en la realidad, durante el curso de sus observaciones diferentes modalidades cuantitativas. La estatura, el peso, la renta, el patrimonio, son distintos para cada persona o para una misma persona, según que las observaciones hayan sido hechas en distintos lugares o en distintos momentos del tiempo. Son estos segundos caracteres de los que se ocupa la estadística, en sus aplicaciones a la biología, a la demografía, a la antropología, a la economía".

"El objeto de la medida de su variación es distinto para las dos categorías del carácter. En el primer caso, el problema que interesa es: ¿En cuánto difieren las medidas de las distintas observaciones del carácter del tamaño real del carácter? En el segundo caso, en vez de la pregunta anterior, el problema que se debe plantear es: ¿En cuánto difieren entre si los distintos valores efectivos del carácter?"

Por lo que se refiere a los índices de variación de las variables que se agrupan en la segunda categoría, hemos visto que para la variación de los índices que se presentan en la demografía, antropología, biología, economía la pregunta apropiada era: ¿En cuánto difieren entre si los distintos valores efectivos del carácter?. En consecuencia con el modo en que formula la

pregunta, Gini propone como medida la media de las diferencias absolutas entre los valores de la variable[10].

En Gini, 1912, (§ **14-16**, pp. 228-232) define ***la media de las diferencias***, y encuentra distintas expresiones de la misma, razonando como sigue. Se considera una variable cuantitativa no negativa (por ejemplo, la renta), variable que ha sido observada sobre los individuos $i = 1,2,\dots n$, tomando valores que han sido ordenados de menor a mayor

$$x_1 \leq x_2 \leq \cdots \leq x_n$$

se propone encontrar una fórmula que le permita expresar *la media aritmética de las diferencias absolutas entre los* n *valores de la variable,* es decir:

$$\Delta = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j| \tag{4.1}$$

Para ello, demuestra (obra citada, pp. 228-230) en primer lugar que:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j| = 2 \sum_{i=1}^{i^*} [(n+1-2i)(x_{n+1-i} - x_i)]$$

donde i^* es igual a $(n+1)/2$ si n es impar y $n/2$ si n es par. Se puede observar que los valores de la variable, x_{n+1-i}, x_i , son *simétricos respecto a la mediana* de los datos (cuando n es impar la mediana es el valor $x_{\frac{n+1}{2}}$, cuando n es par la mediana es el valor $(x_{\frac{n}{2}} + x_{\frac{n}{2}+1})/2$. En consecuencia la media de las diferencias se expresará:

$$\Delta = \frac{2}{n(n-1)} \sum_{i=1}^{i^*} [(n+1-2i)(x_{n+1-i} - x_i)] \tag{4.2}$$

Si tomamos todas las diferencias entre todos los valores de la variable (incluidas las diferencias nulas) se obtiene la diferencia media con repetición entre los valores de la variable:

$$\Delta_R = \frac{2}{n^2} \sum_{i=1}^{i^*} [(n+1-2i)(x_{n+1-i} - x_i)] \tag{4.3}$$

Es fácil comprobar que la relación entre la *diferencia media sin repetición,* Δ , y la *diferencia media con repetición*, Δ_R , viene dada por

[10] En cuanto a los antecedentes históricos de la media de las diferencias (David, 1998, pp. 373-374) es necesario remontarse a una serie de artículos sobre la estimación de la dispersión, publicados en la década de los 1860 en la revista alemana *Astronomische Nachrichten (Noticias Astronómicas)*, que se inician con la figura W. Jordan, profesor en Calsruhe. Propone mejorar las estimaciones de Gauss basándose no en las mismas estimaciones $x_1, x_2, \dots, x_n$, sino más bien sobre las $n(n-1)/2$ diferencias absolutas

$$|x_i - x_j|, \qquad i, j, = 1,2, \dots, n, \qquad i \neq j$$

Concretamente propone, en notación moderna, el estadístico (Jordan, 1869)

$$G_k = \frac{1}{n(n-1)/2} \sum_{i<j}^{n} |x_i - x_j|^k$$

Posteriormente centra su atención en $G_1 = g$. Von Andrae (1872) hace un estudio detallado de g , proponiendo la siguiente fórmula para su cálculo

$$g = \frac{2}{n(n-1)} \sum_{i=1}^{n/2} (n - 2i + 1)(x_{n+1-i} - x_i)$$

$$\Delta_R = \frac{n-1}{n}\Delta \tag{4.3}$$

"La diferencia media sin repetición es, en muchos casos, el índice apropiado para el estudio de la variación de los fenómenos de la demografía, antropología, biología y economía; la diferencia media con repetición sirve para poner de manifiesto la diferencia entre la diferencia media y la desviación simple media respecto a la mediana y respecto a la media aritmética…", (obra citada, p. 231).

A partir de la fórmula (4.2), Gini obtiene la siguiente expresión

$$\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|}{2(n-1)X_n} = \frac{2\sum_{i=1}^{n}(i-1)x_i}{(n-1)X_n} - 1\,, \tag{4.2.1}$$

donde X_n es la suma de todos los valores de la variable, y el segundo miembro es precisamente la fórmula que mostraremos más adelante (5.8) de la razón de concentración. Pero en este momento (1912), Gini aún no había introducido la razón de concentración (1914). Sin saberlo en este momento, en la fórmula (4.2.1) demuestra la relación existente entre la media de las diferencias y la razón de concentración.

Una forma más simple para llegar a otra fórmula más sencilla de la media de las diferencias, es considerar la suma de las desviaciones simples de todos los valores de la variable respecto a uno de los valores de la variable x_h , por ejemplo, (Gini, 1912, pp. 247-248):

$$\sum_{k=1}^{n}|x_k - x_h| = \sum_{k=h}^{n}(x_k - x_h) + \sum_{k=1}^{h-1}(x_h - x_k)$$

Operando se puede escribir

$$\sum_{h=1}^{n}\sum_{k=1}^{n}|x_k - x_h| = 2\sum_{h=1}^{n}(2h - 1 - n)x_h = 4\sum_{h=1}^{n}\left(h - \frac{n+1}{2}\right)x_h$$

En consecuencia, la media de las diferencias se puede expresar:

$$\Delta = \frac{4}{n-1}\sum_{h=1}^{n}\left(\frac{h}{n} - \frac{n+1}{2n}\right)x_h \tag{4.4}$$

Y la media de las diferencias con repetición se puede expresar:

$$\Delta_R = \frac{4}{n}\sum_{h=1}^{n}\left(\frac{h}{n} - \frac{n+1}{2n}\right)x_h \tag{4.5}$$

Si consideramos que la variable en consideración toma una infinidad de valores de forma continua, con función de distribución $F(x)$ y función de densidad $f(x)$, la expresión de ***la media de las diferencias en el caso continuo*** es

$$\Delta = \int_{x_0}^{\infty}\int_{x_0}^{\infty}|x - y|f(x)f(y)dxdy$$

donde x_0 es la renta mínima. Esta integral doble es difícil de calcular directamente, pero tomando límites continuos en la fórmula (4.5), se obtiene para la media de las diferencias en el caso continuo, la fórmula:

$$\Delta = 4\int_{x_0}^{\infty}\left(F(x) - \frac{1}{2}\right)xf(x)dx \tag{4.5.1}$$

Por lo que se refiere a la relación entre el parámetro alfa de Pareto y la media de las diferencias, es preciso recordar, que en su trabajo de 1909 Gini introducía el índice de desigualdad delta, siguiendo el modelo de Pareto. Las dos conclusiones fundamentales eran:

1. El parámetro δ variaba según los países y a lo largo del tiempo, mientras que Pareto consideraba que su parámetro α era constante.
2. Las variaciones del parámetro δ estaban en relación inversa a las variaciones del parámetro α de Pareto.

Pero las conclusiones de Gini presentaban el problema de que los datos no aportaban suficiente evidencia sobre la primera, y respecto a la segunda el índice δ no era del tipo de los índices de variabilidad que existían en la literatura estadística.

En la memoria de 1910, Gini aportaba una prueba matemática aproximada de la relación entre los índices α y δ, $\delta = \frac{\alpha}{\alpha - 1}$. También, en esta memoria, Gini mejora la forma de introducir su índice, aunque vuelve a seguir los pasos de Pareto.

Veamos cómo procede Gini en la memoria de 1912 para relacionar el parámetro α con la media de las diferencias (pp. 275-277). Considera que la sucesión de los datos x_i es continua, x_0 es la renta mínima, y que los datos se distribuyen según una ley con función de densidad con parámetro h

(4.6) $$f(x) = Vx^{-h}$$

donde V es una constante que depende de la renta mínima x_0 y del parámetro h. Esta función de densidad es el modelo de Pareto con $\alpha = h - 1$. A partir de estas consideraciones, utilizando la fórmula integral (4.5.1), llega a la relación buscada

(4.7) $$\Delta_R = \frac{2(h-1)x_0}{(h-2)(2h-3)}$$

Si en la expresión anterior hacemos $\alpha = h - 1$, se obtiene:

(4.8) $$\Delta_R = \frac{2\alpha x_0}{(\alpha - 1)(2\alpha - 1)}$$

Es preciso poner de manifiesto, en este momento, que a partir de la fórmula (4.8) Gini obtiene ***la relación entre el parámetro*** α ***del modelo de Pareto*** (cuyo significado trata de comprender) ***con la media las diferencias*** (una medida de dispersión): si α aumenta entonces Δ_R disminuye.

Por otra parte, como la media de la distribución de Pareto viene dada por $\mu = \frac{\alpha x_0}{\alpha - 1}$, se deduce de (4.8) que:

(4.9) $$\frac{\Delta_R}{2\mu} = \frac{1}{2\alpha - 1}$$

En esta memoria de 1912 (pp. 288-289) da una nueva demostración de la relación entre el índice α de Pareto y su índice δ, $\delta = \frac{\alpha}{\alpha - 1}$. En el proceso de esta demostración parte de su curva de Lorenz dual, para definir su índice, para llegar a la curva de Lorenz que define el modelo de Pareto (la prueba la hace en las dos direcciones). ***Tiene que admitir que su medida de desigualdad delta se basa en la suposición, hasta entonces no evidente para Gini, de que los datos siguen una distribución de Pareto***. Este resultado, para él inesperado, le sitúa en el comienzo de sus investigaciones sobre las medidas de desigualdad (1908) y se plantea la tarea de buscar una nueva medida que sea independiente de la distribución que sigan los datos.

5. LA RAZÓN DE CONCENTRACIÓN DE C. GINI (1914)

En la memoria de C. Gini (1914), *Sulla misura della concentrazione e della variabilità dei caratteri,* es donde define la nueva medida, *la razón de concentración,* con la pretensión de que sea independiente de la distribución que sigan los datos. Además, aporta tres resultados de gran interés: en el primero interpreta geométricamente su razón de concentración como el cociente entre el área de Lorenz-Gini y el máximo de esta área; en el segundo relaciona su razón de concentración con la media de las diferencias y en el tercero relaciona la media de las diferencias con el área de Lorenz- Gini.

Para la introducción de la ***razón de concentración*** razona como sigue (Gini, 1914, § **1-2**, pp. 413-417). Se considera una variable cuantitativa no negativa (por ejemplo, la renta), variable

que ha sido observada sobre los individuos $i = 1,2,\dots n$, tomando valores que han sido ordenados de menor a mayor

$$x_1 \leq x_2 \leq \cdots \leq x_n$$

Se define la variable

$$X_i = \sum_{k=1}^{i} x_k \quad i = 1,2,\dots,n \tag{5.1}$$

que es la cantidad acumulada de los valores de la variable por los i primeros individuos.

Por lo que, la variable

$$Q_i = \frac{X_i}{X_n} = \frac{\sum_{k=1}^{i} x_k}{\sum_{k=1}^{n} x_k} \tag{5.2}$$

es la fracción acumulada de los valores de la variable por los i primeros individuos, respecto a la cantidad total acumulada de los valores de la variable.

La variable

$$P_i = \frac{i}{n} \tag{5.3}$$

representa la fracción de observaciones hasta el individuo i , respecto al número total de observaciones.

Si $i < n$, en razón de la ordenación de menor a mayor de los valores de la variable, se tiene que la media de los i primeros valores es menor que la media de los $n - i$ últimos valores, es decir,

$$\frac{1}{i}\sum_{k=1}^{i} x_k < \frac{1}{n-i}\sum_{k=i+1}^{n} x_k$$

Por otra parte, la media de todos los valores

$$\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k = \frac{i}{n}\frac{1}{i}\sum_{k=1}^{i} x_k + \frac{(n-i)}{n}\frac{1}{(n-i)}\sum_{k=1+i}^{n} x_k$$

Aplicando la desigualdad anterior, se tiene

$$\bar{x} = \frac{i}{n}\frac{1}{i}\sum_{k=1}^{i} x_k + \frac{(n-i)}{n}\frac{1}{(n-i)}\sum_{k=1+i}^{n} x_k > \frac{i}{n}\frac{1}{i}\sum_{k=1}^{i} x_k + \frac{(n-i)}{n}\frac{1}{i}\sum_{k=1}^{i} x_k = \frac{1}{i}\sum_{k=1}^{i} x_k$$

Por tanto,

$$\frac{1}{n}\sum_{k=1}^{n} x_k > \frac{1}{i}\sum_{k=1}^{i} x_k$$

Lo que es equivalente a

$$\frac{\sum_{k=1}^{i} x_k}{\sum_{k=1}^{n} x_k} < \frac{i}{n}$$

Es decir $Q_i < P_i$, y la línea poligonal cuyos vértices son

$$(0,0), (P_i, Q_i), i = 1,2,\dots,n-1, \tag{5.4}$$

queda por debajo de la bisectriz del primer cuadrante y tiene una forma convexa. Esta línea poligonal recibe el nombre de *poligonal de Lorenz.*

Es obvio, como afirma Gini, "*que la concentración del carácter será tanto más fuerte cuanto más fuertes sean las desigualdades*

$$Q_i < P_i \quad i = 1,2,\dots,n-1 \text{ "} \tag{5.5}$$

Los casos extremos de concentración se dan para

a) Si $x_i = 0, i = 1,2,\dots,n-1$ $y\, x_n = X$, entonces $Q_i = 0, i = 1,2,\dots,n-1$ y $P_i - Q_i = P_i$, $i = 1,2,\dots,n-1$. En este caso la concentración del carácter es perfecta o de máxima desigualdad.

b) Si $Q_i = P_i$ $i = 1,2,\dots,n-1$, hay una equidistribución del carácter o de máxima igualdad.

La desigualdad (5.5) es tanto más fuerte cuanto mayor sea la diferencia $P_i - Q_i$; y, *en términos relativos, cuánto más alta es la fracción*

(5.6) $$R_i = \frac{P_i - Q_i}{P_i}$$

R_i representa el coeficiente por el que es necesario multiplicar la fracción P_i de los casos, en los que la intensidad del carácter no supera un valor límite, para obtener la diferencia entre esta fracción y la fracción del montante del carácter que se espera en tales casos.

Se podrá obtener una medida de la concentración del carácter al tomar la media ponderada de los $n-1$ valores de R_i, en los que cada R_i entra con un peso proporcional al valor de P_i . Es decir,

$$R = \frac{\sum_{i=1}^{n-1} R_i P_i}{\sum_{i=1}^{n-1} P_i} = \frac{\sum_{i=1}^{n-1} \frac{[(P]_i - Q_i)}{P_i} P_i}{\sum_{i=1}^{n-1} P_i} = \frac{\sum_{i=1}^{n-1} [(P]_i - Q_i)}{\sum_{i=1}^{n-1} P_i}$$

El coeficiente

(5.7) $$R = \frac{\sum_{i=1}^{n-1} [(P]_i - Q_i)}{\sum_{i=1}^{n-1} P_i}$$

se le denomina ***Razón de concentración de Gini*** y verifica las siguientes propiedades:

1) Para la máxima concentración, $Q_i = 0, i = 1,2,\dots,n-1$, se verifica que $R = 1$.
2) Para la mínima concentración, $P_i = Q_i, i = 1,2,\dots,n-1$, se verifica que $R = 0.$
3) En la medida que las desigualdades $Q_i < P_i$ son más grandes, entonces R aumenta.

Para hacer más operativa la expresión (5.7), introduce (Gini, (1914), § **3**, pp. 417-418) la siguiente fórmula de ***la razón de concentración para datos desagregados en frecuencias absolutas unitarias***

(5.8) $$R = \frac{2\sum_{i=1}^{n}(i-1)x_i}{(n-1)X_n} - 1$$

La deducción que hace de esta fórmula es sustituyendo las definiciones de los Q_i, P_i en (5.7). Se debe observar que esta fórmula coincide con el segundo miembro de la expresión (4.2.1), obtenida por Gini en 1912.

Si consideramos la variable rango, que denotamos por $y = i - 1$, $i = 1,2,\dots,n$, se puede comprobar, a partir de (5.8), y que

(5.9) $$R = \frac{2}{(n-1)\bar{x}} Cov(y, x)$$

La fórmula (5.9) muestra la relación que existe entre la razón de concentración, la covarianza entre la variable rango y la variable en cuestión, y la media aritmética de la variable.

La expresión (5.8) de la razón de concentración se puede expresar como

(5.10) $$R = 1 - \frac{\bar{x}^*}{\bar{x}}$$

donde

$$\bar{x}^* = \sum_{i=1}^{n} \left(\frac{(n-i)}{\frac{n(n-1)}{2}} \right) x_i$$

es una media ponderada de todos los valores de la variable, con coeficiente de ponderación afectado por el factor $(n-i)$, siendo éstos:

$n-1\ para\ x_1, n-2\ para\ x_2, \dots, 1\ para\ x_{n-1}, 0\ para\ x_n$

y

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

la media aritmética de todos los valores de la variable.

La expresión (5.8) sólo es válida si los datos están desagregados, por lo que es preciso adaptarla cuando los datos están agregados. Adoptamos las siguientes notaciones; n, el número total de individuos u observaciones; r, las clases o categorías en las que se dividen los n individuos; $k, k = 1,2,\dots,r$, el número de orden de cada una de las clases comenzando por la más baja; n_k, el número de individuos que entra en la clase k ; S_k, la suma de los valores de la variable de los n_k individuos que entran en la clase k ; L_k, el límite superior de la clase k , con $L_{k-1} < z_{kl} \leq L_k$; $z_{kl} = x_k$, es el valor de la variable para los individuos $l = 1,2,\dots,n_k$ de la clase k ; N_k , el número de individuos con renta inferior o igual a L_k . Se puede comprobar, utilizando (5,8), que ***la razón de concentración para datos agregados en frecuencias absolutas*** es

$$R = \frac{\sum_{k=1}^{r}(N_{k-1} + N_k - 1)z_{kl}n_k}{(n-1)X_n} - 1 \qquad (5.11)$$

En la mayor parte de los casos se dispone de datos estadísticos agrupados en distintas clases o intervalos, donde solo conocemos el valor de la variable para las distintas clases más o menos amplias. Pero en otros casos, se dispone también la suma de los valores de la variable dentro de cada una de las clases. En lo que sigue vamos a ver la fórmula de la ***razón de concentración para datos agregados por clases donde solo se conoce la suma de los valores de la variable para cada una de las clases***. Denotamos por

$$\delta_{kl} = z_{kl} - \frac{S_k}{n_k}, \qquad k = 1,2,\dots,r; l = 1,2,\dots,n_r \qquad (a)$$

la desviación del valor de la variable, del individuo l de la clase k , respecto a la media aritmética de los valores de la variable de los n_k que forman parte de la misma clase. Se puede comprobar que

$$\sum_{l=1}^{n_k} \delta_{kl} = 0 \qquad (b)$$

Denotamos por

$$\varepsilon_{kl} = (N_{k-1} + l - 1) - \frac{N_k + N_{k-1} - 1}{2}, \qquad k = 1,2,\dots,r; l = 1,2,\dots,n_r \qquad (c)$$

la desviación de orden, del individuo $y_{kl} = (N_{k-1} + l - 1)$ de la clase k , respecto al orden de la media aritmética de los n_k individuos de su misma clase. Se puede comprobar que:

$$\sum_{l=1}^{n_k} \varepsilon_{kl} = 0 \qquad (d)$$

Haciendo las sustituciones pertinentes para este tipo de datos en (5.11), se obtiene para la razón de concentración de este tipo de datos la fórmula:

$$R = \frac{\sum_{k=1}^{r}(N_k + N_{k-1} - 1)S_k + 2\sum_{k=1}^{r}\sum_{l=1}^{n_k}\varepsilon_{kl}\delta_{kl}}{(n-1)X_n} - 1 \qquad (5.12)$$

La expresión (5.12) que nos da la razón de concentración para datos agrupados, se puede escribir en la forma

$$R = \frac{1}{(n-1)X_n}\sum_{k=1}^{r}(N_k + N_{k-1} - 1)n_k\overline{z_k} - 1 + \frac{2}{(n-1)X_n}\sum_{k=1}^{r}\sum_{l=1}^{n_k}\delta_{kl}\varepsilon_{kl} \qquad (5.13)$$

donde $\overline{z_k}$ es la media de los valores de la variable del grupo k , es decir
$n_k\overline{z_k} = S_k$

Los dos primeros sumandos de la expresión anterior representan ***la razón de concentración entre los grupos*** que denotaremos por

$$R_E = \frac{1}{(n-1)X_n}\sum_{k=1}^{r}(N_k + N_{k-1} - 1)n_k\overline{z_k} - 1 \qquad (5.14)$$

De acuerdo con (5.9) ***la razón de concentración dentro de cada grupo*** $R_k, k = 1,2 \dots, r$ se podrá escribir:

$$R_k = \frac{1}{(n-1)\overline{z_k}} Cov(y_k, z_k)$$

donde y_k es la variable rango para el grupo k . Si se tiene en cuenta que

$$\sum_{l=1}^{n_k} \delta_{kl}\, e_{kl} = n_k Cov(y_k, z_k)$$

entonces

$$\frac{2}{(n-1)X_n} \sum_{k=1}^{r} \sum_{l=1}^{n_k} \delta_{kl}\, e_{kl} = \sum_{k=1}^{r} \frac{2n_k Cov(y_k, z_k)}{2(n-1)X_n} = \sum_{k=1}^{r} \frac{2n_k R_k (n_k - 1)\overline{z_k}}{2(n-1)X_n}$$

[redacted]

Denominaremos ***razón de concentración dentro de los grupos***, que denotaremos por R_D , a la suma ponderada de las razones de concentración dentro de cada grupo

$$R_D = \sum_{k=1}^{r} \left(\frac{S_k}{X_n}\right)\left(\frac{n_k - 1}{n-1}\right) R_k \qquad (5.15)$$

donde los correspondientes pesos son el producto de la fracción de individuos de cada grupo respecto al total de individuos por la correspondiente fracción de renta que poseen los individuos de ese grupo.

En consecuencia, la razón de concentración se puede descomponer como la suma de la razón de concentración entre los grupos más la razón de concentración dentro de los grupos, es decir

$$R = R_E + R_D \qquad (5.16)$$

$$R = \left[\frac{1}{(n-1)X_n} \sum_{k=1}^{r} (N_k + N_{k-1} - 1) n_k \overline{z_k} - 1\right] + \sum_{k=1}^{r} \left(\frac{S_k}{X_n}\right)\left(\frac{n_k - 1}{n-1}\right) R_k$$

Esta descomposición de la razón de concentración es útil como criterio para la agrupación de los datos. En efecto, puesto que R , para un conjunto de datos dado permanece invariable, se puede utilizar como criterio de agrupación: o bien hacer máxima la razón de concentración entre los grupos (R_E), o bien hacer mínima la razón de concentración dentro de los grupos (R_D).

Gini, en la obra citada (1914, § **6**, pp.438-444), comienza diciendo que "a la razón (de concentración), que aquí proponemos como medida apropiada de la concentración, se llega también perfeccionando un método gráfico que algunos autores, Lorenz (1905), Chatelain (1907, 1910, 1911), Seailles (1910), han propuesto para juzgar la mayor o menor desigualdad de la distribución de la riqueza". A pié de página cita también a Watkins (1905, 1908, 1909), Persons (1909) y King(1912). En el primer cuadrante del plano de Descartes, con ejes ortogonales, sitúa en el eje de abscisas los valores de los P_i y sobre el eje de ordenadas los valores de los Q_i. Con los puntos de coordenadas (P_i, Q_i) obtiene la ***curva de concentración***, ascendente de izquierda a derecha, desde el punto (0,0) al punto (1,1), que en virtud de (5.4) resulta ser convexa respecto al eje de abscisas. Y denomina ***recta de equidistribución*** a la recta que une los puntos (0,0) y (1,1), que se puede interpretar como la curva de menor concentración posible o de igualdad perfecta: en cada uno de sus puntos la proporción de individuos coincide con la proporción de renta que poseen. Cuánto menos pequeña sea el área comprendida entre la recta de equidistribución y la curva de concentración, *área de Lorenz*, tanto menor será la desigualdad. La medida de esta área, es para los autores citados por Gini, una medida de la concentración.

Este método gráfico de la medida de la concentración presenta para Gini dos grandes inconvenientes, que son reconocidos por Lorenz y King:

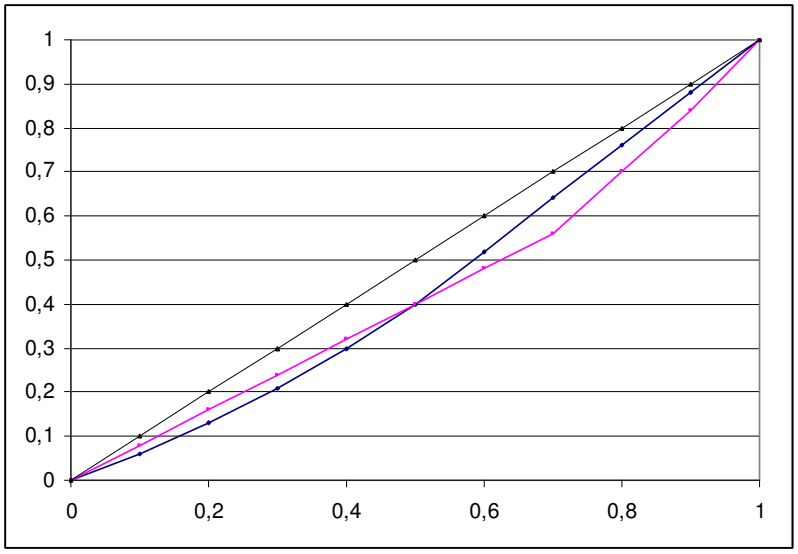
a) No aporta una medida precisa de la concentración.

b) Tampoco permite, en algunos casos, juzgar sobre la mayor o menor concentración. Así, cuando dos curvas de concentración *se cruzan* (aquí sigue el ejemplo de Lorenz que se expone a continuación), este método gráfico no permite la comparación de esas dos curvas de concentración.

Exponemos ***el ejemplo de Lorenz*** (1905), donde muestra el caso de dos curvas de Lorenz que se cruzan y no es posible la comparación de sus concentraciones. Este ejemplo fue una de las motivaciones de Gini para la introducción de su razón de concentración. Se trata de dos períodos temporales que llamamos A y B con 10 individuos cada uno y con rentas totales igual a 100 unidades monetarias. La siguiente tabla recoge los datos y los valores numéricos de las curvas de Lorenz de estos datos.

		Individuos	P	Q_A	Q_B
A	B		0	0	0
6	8	1	0,1	0,06	0,08
7	8	2	0,2	0,13	0,16
8	8	3	0,3	0,21	0,24
9	8	4	0,4	0,3	0,32
10	8	5	0,5	0,4	0,4
12	8	6	0,6	0,52	0,48
12	8	7	0,7	0,64	0,56
12	14	8	0,8	0,76	0,7
12	14	9	0,9	0,88	0,84
12	16	10	1	1	1

Las curvas de Lorenz son,



donde vemos que la curva de Lorenz de B es menos desigual al comienzo, está por encima de la curva de Lorenz de A, para pasar a ser más desigual, está por debajo de A para grandes rentas. Para estos datos, la razón R de Gini es: $R_A = 0{,}133$ y $R_B = 0{,}16$, y así para Gini el período B es más desigual que el A.

Ahora bien, en la tabla anterior se observa que los individuos 4, 5, 6 y 7 pierden 11 unidades monetarias al pasar del período A al B, de las que 3 van a los dos individuos más pobres que pasan de tener 6 y 7 u.m. a 8 y 8, respectivamente; mientras que las otras 8 u.m. van a parar a los individuos 8, 9 y 10 más ricos que pasan de tener 12, 12 y 12 a 14, 14 y 16 u.m. respectivamente. Vemos que del "centro" se toman 11 u.m. que se dan tres u.m. a los dos más pobres (transferencias progresivas) y 8 u.m. que se dan a los tres más ricos (transferencias regresivas).

De este último argumento, podemos ver que la razón de Gini, que asigna mayor desigualdad a la población B, no recoge el hecho de que los dos más pobres se han beneficiado de los individuos situados en el "centro". Si valoramos más las dos u.m. que pasan a los pobres que las 8 que pasan a los ricos, en el sentido de disminuir la desigualdad, en el primer, caso y de aumentar la desigualdad, en el segundo caso; entonces deberíamos hablar de que al pasar del período A al B la desigualdad ha disminuido.

Para tratar de evitar ambos inconvenientes (a), (b), Gini propone, por vez primera, ***medir la concentración como el cociente entre el área de Lorenz y el área de máxima concentración*** (el área del triángulo rectángulo que tiene por vértices los puntos (0,0), (1,0) y (1,1)). Es necesario indicar que en esta definición Gini está suponiendo que los *datos son continuos*. Será en el parágrafo **8** donde aplicará esta definición al caso de datos discretos. Demuestra que *el cociente entre el área de Lorenz y el área de máxima concentración es el número al cual tiende la razón de concentración* R, *cuando crece el número* n *de los casos observados, manteniéndose por igual su distribución.* En efecto, considera, en primer lugar, los $n-1$ pequeños rectángulos, que tienen la base sobre el eje de abscisas y tocan con su altura la curva de concentración, tienen por altura Q_i y como base $\frac{1}{n}$, su área será $\frac{1}{n}\sum_{i=1}^{n-1} Q_i$. Considera, en segundo lugar, los $n-1$ pequeños rectángulos que tienen como base $\frac{1}{n}$ y altura P_i, la suma de sus áreas será $\frac{1}{n}\sum_{i=1}^{n-1} P_i$. La diferencia de las dos sumas $\frac{1}{n}\sum_{i=1}^{n-1}[(P_i - Q]_i)$ será igual a la suma de los $n-1$ rectángulos limitados entre la recta de equidistribución y la curva de concentración. Ahora, al crecer n, disminuye el área de la pequeña superficie comprendida entre la recta de equidistribución y el área de la totalidad de los rectángulos de altura P_i y, análogamente disminuye el área de la pequeña superficie comprendida entre la curva de concentración y el área de la totalidad de los rectángulos de altura Q_i[11]. Con el crecimiento de n, el área del triángulo de máxima concentración se convierte en $\frac{1}{n}\sum_{i=1}^{n-1} P_i$, el área de concentración se convierte en $\frac{1}{n}\sum_{i=1}^{n-1}[(P_i - Q]_i)$ y la razón de ésta con aquella se hace igual a la razón de concentración

$$R = \frac{\sum_{i=1}^{n-1}[(P]_i - Q_i)}{\sum_{i=1}^{n-1} P_i}$$

De los razonamientos anteriores, es fácil deducir que cuando n crece la razón de concentración se aproxima a dos veces el área de Lorenz, ya que en este caso el área del triángulo de máxima concentración se aproxima a $1/2$ y el numerador de R se aproxima al área de Lorenz.

En el parágrafo ocho (pp. 444-486) demuestra para *datos discretos* que ***la razón de concentración es igual al cociente entre el área de Lorenz y el área de máxima concentración***. Pero en este caso, el área de Lorenz es el área comprendida entre la recta de equidistribución y la poligonal de Gini (5.4); y el área de máxima concentración es el área del triángulo de vértices $(0,0), \left(\frac{n-1}{n}, 0\right), (1,1)$.

Por lo que se refiere a ***la relación entre la media de las diferencias y la razón de concentración***, Gini demuestra que (Gini, 1914, pp. 447-448) la razón de concentración es igual

[11] Hacemos notar que el método utilizado por Gini en el cálculo de estas áreas es el de las sumas inferiores de Riemann. Al hacer crecer n ilimitadamente se obtienen las correspondientes integrales de Riemann.

al cociente entre la media de las diferencias sin repetición de los valores de la variable y el doble de la media aritmética de los valores de la variable, es decir:

$$R = \frac{\Delta}{2\overline{x}} \tag{5.17}$$

En la deducción de esta fórmula emplea un método bastante artificioso, cuando se puede considerar que ya tenía una demostración si hubiera considerado las expresiones (5.8) de 1914 y (4,2,1) de 1912.

A la vista de (5.17) y teniendo en cuenta que el máximo valor de la razón de concentración, R , es 1, se deduce que el máximo valor de la media de las diferencias sin repetición de los valores de la variables, Δ , es dos veces la media aritmética de los valores de la variable, $2\overline{x}$.

Si consideramos ahora que los datos están agrupados por clases, y teniendo en mente la expresión (5.16), podremos escribir:

a) Que la razón de concentración entre los grupos, R_E , se puede escribir, en virtud de (5.17), en términos de la diferencia media entre los grupos, Δ_E , tomando para cada grupo la media de los z_{kl} , $l = 1,2,\dots,n_k$, $(\overline{z_k})$:

$$R_E = \frac{\Delta_E}{2\overline{x}}$$

donde

$$\Delta_E = \frac{1}{n(n-1)}\sum_{i=1}^{r}\sum_{j=1}^{r}|\overline{z_i} - \overline{z_j}|n_i n_j , \qquad n = \sum_{k=1}^{r} n_k$$

b) Que la razón de concentración para cada grupo, R_k , se puede escribir, en virtud de (5.17), en términos de la diferencia media dentro de cada grupo, Δ_k:

$$R_k = \frac{\Delta_k}{2\overline{z_k}}$$

donde

$$\Delta_k = \frac{1}{n_k(n_k-1)}\sum_{i=1}^{n_k}\sum_{j=1}^{n_k}|z_{ki} - z_{kj}|$$

Entonces de la expresión (5.16), podemos escribir:

$$R = R_E + \sum_{k=1}^{r}\left(\frac{S_k}{X_n}\right)\left(\frac{n_k - 1}{n-1}\right)R_k$$

$$\frac{\Delta}{2\overline{x}} = \frac{\Delta_E}{2\overline{x}} + \sum_{k=1}^{r}\left(\frac{S_k}{X_n}\right)\left(\frac{n_k - 1}{n-1}\right)\frac{\Delta_k}{2\overline{z_k}}$$

En consecuencia,

$$\Delta = \Delta_E + \sum_{k=1}^{r}\frac{n_k(n_k-1)}{n(n-1)}\Delta_k$$

Esta expresión es la que aparece con la numeración (23) en (Gini, 1914, p. 448). Puede ser interpretada diciendo que la diferencia media es la suma de la diferencia media entre los grupos más la suma ponderada de las diferencias medias dentro de cada grupo (con factor de ponderación proporcional a la importancia relativa del número de individuos del grupo correspondiente).

Por otra parte,

$$(R - R_E)2\overline{x} = \frac{1}{n(n-1)}\sum_{k=1}^{r}\sum_{i=1}^{n_k}\sum_{j=1}^{n_k}|z_{ki} - z_{kj}|$$

$$(R - R_E)2\overline{x} = \frac{1}{n(n-1)}\sum_{k=1}^{r} n_k(n_k-1)\Delta_k$$

A partir de esta expresión se propone en (Agheli-Mehran, 1981), como criterio para conseguir una agrupación óptima de los datos de la distribución de la renta, minimizar la diferencia $(R - R_E)$.

Nos preguntamos ahora, ***¿qué ideas tenía en mente C. Gini para relacionar su medida de dispersión, la media de las diferencias, con el área de Lorenz***? Puesto que Gini alude a esta relación, pero no hace un razonamiento explícito la misma, pensamos que podría haber procedido como sigue.

En la memoria de 1912 Gini obtiene la fórmula (4.9):

$$\frac{\Delta_R}{2\mu} = \frac{1}{2\alpha - 1}$$

Si consideramos el índice de concentración introducido en (Gini, 1909) cuya relación con el parámetro alfa de Pareto (datos continuos), mostrada por Gini en 1910 es

$$\delta = \frac{\alpha}{\alpha - 1}$$

podemos expresar la fórmula anterior como

$$\frac{\Delta_R}{2\mu} = \frac{\delta - 1}{\delta + 1}$$

Puesto que la curva de Lorenz dual asociada al modelo de Pareto es

$$\bar{q} = \bar{p}^{\frac{1}{\delta}} \quad , \qquad \bar{q} = 1 - q, \qquad \bar{p} = 1 - p$$

Se tiene:

$$2(\acute{A}rea\ de\ lorenz) = 2\left(\int_0^1 \bar{p}^{\frac{1}{\delta}}\, d\bar{p} - \frac{1}{2}\right) = \frac{\delta - 1}{\delta + 1} = \frac{\Delta_R}{2\mu}$$

Este resultado teórico para datos continuos pensamos que podría ser el que indujo a Gini proponer el doble del área de Lorenz como una medida de desigualdad.

BIBLIOGRAFÍA

AGHEVLI B.B., MEHRAN F. (1981). Optimal Grouping of Income Distribution Data. *J. American Statistical Association,* vol. LXXVI, n.373, pp. 22-26.

BARBUT, M. (2007). *La mesure des inégalités. Ambigüités et paradoxes*. Librairie DROZ, Genève-Paris.

BARONE, E. (1942), *Principios de Economía Política*. Editorial Revista de Derecho Privado. Madrid. Traducción de su obra *Principi di economia politica* de 1920-23, incluyendo "Los réditos y su distribución" y "De los campos a los talleres" de 1936.

BENEDETTI, C. (1965). Ricordando Corrado Gini, *Rivista de Politica Economica* 54, 3-9.

BENEDETTI, C. (1984). A propósito del centenario della nascita di Corrado Gini, *Metron* **42** (1/2),3-19.

BENINI, R. (1897) Di alcune curve descritte da fenomeni economici aventi relazione con la curva del redito o con quella del patrimonio", *Giornale degli Economisti*, s. II, a. VIII, vol. XIV, n. 5 marzo, pp. 177-214.

BENINI, R. (1906), *Principii di statistica metodologica*. Unione Tipografico-Editrici, Torino.

BORTKIEWICZ, L. v. (1931), *Die Disparitätsmasse der Einkommensstatistik.* XII Session de L'Institut International de Statistique, Tokio, pp. 189-298.

BRESCIANI-TURRONI, C. (1939), Pareto's Law and the Index of Inequality of Incomes. *Econometrica*, vol. 7, No. 2 (Apr), pp. 107-133.

CASTELLANO, V. (1965). Corrado Gini: A memoir, *Metron* **24**, 3-84.

CHATELAIN, E., (1907), Les sucessions declarées en 1905*, Revue politique et parlamentaire,* Paris.

CHATELAIN, E., (1910), La trace de la courbe des sucessions en France, *Journal de la Societé de Statistique de Paris*, Paris, p. 352 y siguientes.

CHATELAIN, E., (1911), La fortune française d'aprés les sucessions en 1909, *La Democratie,* Paris, 20 Janvier, 1911.

DAVID, H. A., (1998), Early Measures of Variability, *Statiscal Science.* Vol. 13, No. 4 (Nov), pp. 368-377.

GEORGESCU-ROEGEN, N. (1966). Further thoughts on Corrado Gini's Delusioni dell'Econometria, *Metron* **25**, 265-279.

GINI, C. (1909). Il diverso accresimento delle classi sociali e la concentrazione della richezza*, Giornale degli Economiste*, anno XX (serie II), n. 1, pp. 27-83.

GINI, C. (1910). Indici di concentrazione e di dependenza*, Biblioteca dell'Economista*, serie V, vol XX, Utet, Torino.
GINI, C. (1912). Variabilità e mutabilità. *Studi económico-giuridici, Università di Cagliari III,* 2ª, en Gini (1955), 211-382.
GINI, C. (1914). Sulla misura della concentrazione e della variabilità dei caratteri. *Atti del R. Istituto Veneto di Scienze, Lettere ed Arti,* en Gini (1955), 411-459.
GINI, C. (1955). *Memorie di metodología statistica.* Vol. 1: *Variabilità e concentrazione.* Ed. E. Pizzetti y T. Salvemini, Roma: Librería Eredi Virgilio Veschi.
GINI, C. (1956). Delusioni dell'Econometria*, Giornale degli Economiste*, anno XV , pp. 174-177.
GIORGI, G.M. (1990). Bibliographic Portrait of the Gini Concentration Index, *Metron* **48,** 183-221.
GIORGI, G.M. (1996). Encounters with the Italian Statistical School: A Conversation with Carlo Benedetti, *Metron* **54** (3/4), 3-23.
JORDAN, W. (1869). Ueber die Bestimmung der Genauigkeit mehrfach wiederholter Beobachtungen einer Unbenkannten. *Astronom. Nachr.***74** 209-226.
KAKWANI, N. C. (1984), Welfare Ranking of Income Distributios. *In Advances in Econometrics*, volume 3, edited by R.L. Basmann and G.F. Rodhes, Jr. Greewich, Conn.: JAI Press.
KING, J., (1912), *The elements of statistical Methods,* The Macmillan Company, New York.
LETI, G. (1996). L'ISTAT e il Consiglio Superiore di Statistica dal 1926 al 1945, *Istituto Nazionale di Statistica, Annali di Statistica*, vol. VIII (series X), Roma.
LORENZ M.O. (1905), Methods of Measuring the Concentration of Wealth, *American Statistical Association*, vol. IX, n. 70, June, pp. 209-219.
MACCABELLI, T. (2002), La "Confusione" sull'indice di disuguaglianza di Pareto; una ricostruzione storica. Nº 85. Studi e Ricerche. Pisa. Università degli Studi di Pisa.
PARETO, V. (1895). La legge della demanda*, Giornale degli Economisti*, Serie seconda, anno VI, vol. X, 59-68.
PARETO, V. (1897). *Cours d'économie politique*, Rouge, Lausanne. [nueva edición de las obras completas de Pareto, Tomo II , publicadas bajo la dirección de G. Busino, Librairie Droz, Genève, 1964].
PARETO, V. (1897). *Aggiunta allo studio sulla curva delle entrate.* Giornale degli Economisti, Serie seconda, anno VII, vol. XIV, pp. 15-26. [nueva edición de las obras completas de Pareto, Tomo III, publicadas bajo la dirección de G. Busino, Librairie Droz, Genève, 1967]. PARETO, V. (1906), *Manuale di economia politica con una introduzione alla scienza sociale*. Padua.
SAIELLES, J. ,(1910), *La repartition des Fortunes en France,* Alcan, Paris.
SHORROCKS, A. F. (1983), Ranking Income Distributions. *Economica*. February, 3-17.
SOREL, G., (1897). La lois de revenus. *Le devenir Social*. 3e Année, Jouillet, pp. 577-607,Paris.
STIGLER, G. J. (1946). *Production and Distribution Theories.* New York.
VON ANDRAE, C. G. (1872) Ueber die Bestimmung des wahrschheinlichen Fehlers durch die gegeben Differenzen von gleich genauen Beobachtung einer Unbekaten. *Astronom. Nachr.* **79** 257-272.
WATKINS, G. P., (1905). Comment on the method of measuring concentration of wealth, *Publications of American Statistical Association*, n. 72, December.
WATKINS, G. P., (1905). An interpretation of certain statistical evidence of concentration of wealth, *Publications of American Statistical Association*, n. 81, March.
WATKINS, G. P., (1909). The measuring of concentration of wealth, *Publications of American Statistical Association*, Vol. XXIV, n. 1, November.
YITZHAKI, S. (1982), Stochastic Dominance, Mean Variance, and Gini's Mean Difference. *The American Economic Review*. Vol. 72. Num. 1. March, pp. 178-185
ZAGIER D. (1983). Inequalities for the Gini Coefficient of Composite Populations. *J. Mathematical Economics*, vol. XII, n. 2, pp. 103-118.
ZENGA, M (1987), *La distribuzione Personale del Reddito: Problemi di Formazione, di Ripartizone e di Misurazione.* Editorial: Nita e Pensiero, pp. 307-328 de "Il contributo degli italiani allo studio della cocnentrazione. Prima parte: dal 1905- 1915".

A GIS BASED MODEL TO SOLVE THE PLANAR HUFF PROBLEM WITH CONTINUOUS DEMAND

RAFAEL SUÁREZ-VEGA
E-mail: rsuarez@dmc.ulpgc.es
DOLORES R. SANTOS-PEÑATE
E-mail: drsantos@dmc.ulpgc.es
PABLO DORTA-GONZÁLEZ
E-mail: pdorta@dmc.ulpgc.es
Departamento de Métodos Cuantitativos en Economía y Gestión
Universidad de Las Palmas de Gran Canaria

Abstract

In this paper, we have used a GIS based model to solve the planar Huff problem. Most of the papers connected with the competitive location problems consider that the demand is aggregated in a finite set of points. In this case, we have considered that the demand is distributed along the plane according to a surface model of population, that is, a raster map where each pixel has associated a value corresponding to the population living in the area that it covers. Taking into account the demand distribution and the location and size of the existing facilities, we have obtained a raster map where each pixel has associated the estimated capture for a new competing fir if decides to locate on it.
A real example is presented where a new hypermarket is located on the island of Gran Canaria (Spain). Finally, for this example, the solutions for the continuous and discrete demand cases are compared.

Key-words: Competitive location, planar Huff problem, continuous demand, surface model of population
Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUCTION

In this paper we investigate the follower competitive location problem in the leader-follower model. A new firm wants to establish a facility in a market where other competing firms already exist. We suppose that each facility is characterized by its quality level that may combine aspects such as size, prices, parking areas, opening hours, accessibility, and so on. Customers make their choice according to the attraction or utility that they perceive from the facilities. The attraction perceived by the customers from the facilities has been represented mathematically by an attraction function which increases with respect to the quality level and decreases with respect to the distance. Attraction functions are used to define the customer choice rule, which represents the customer behavior and the customer flow in the market.

Assuming certain customer preferences, the firms, whose natural objective is the maximization of the market share or profit, take their location decisions, influencing with their actions the results and strategies of their competitor. This movement of individuals defines the basis of the location-spatial interaction models.

The first spatial interaction models were gravity models, which assumed analogies between human behavior and Newtonian gravity laws. The basic gravity formulation, in which the movement of individuals between two points is inversely proportional to the distance separating them, was applied by Reilly (1931) and Converse (1949) to analyze retail market areas. Later,

Huff (1964) proposed an alternative model to overcome certain limitations of the Reilly-type approach. According to this new model, the probability that a customer at *i* buys at a facility *j* is given by

$$P_{ij} = \frac{\frac{a_j}{d_{ij}^{\lambda}}}{\sum_{k=1}^{n} \frac{a_k}{d_{ik}^{\lambda}}}$$

where a_j represents the size of a service center *j*, d_{ij} is the distance (or travel time) from demand point *i* to facility *j*, and λ is a parameter which reflects the effect of the distance on the consumer's behavior and whose value is estimated empirically. The quotient $\frac{a_j}{d_{ij}^{\lambda}}$ can be interpreted as an attraction function (where that the attraction felt by a customer at point *i* towards facility *j* is directly proportional to the size of the facility and inversely proportional to a power of the distance between them).

The Huff based competitive location models have previously been studied in OR. Drezner (1994) solved the planar problem which considered Euclidean distances and each point in the plane as a potential location for the new facility. In this case, the author considered that the demand was aggregated in a finite set of points and used the Weiszfeld algorithm to obtain the best location. This is a gradient based method that yields a local optimum, so the procedure must be run several times and the best solution obtained is then selected. Later, Drezner and Drezner (1997) proposed a problem resolution for the continuous demand case. This resolution is based on the assumption that the demand density function *f(x,y)* is known, that is, that there exists a known functional relationship between the coordinates (*x,y*) of a point and its population density. Due to the complexity of the procedure for the resolution of this problem, the authors proposed an approximation approach based on the aggregation of the demand.

In this paper, the planar Huff problem has been solved using GIS based tools. Previous works combining these geographical tools and interaction models are described in Birkin et al. (1996), Birkin et al. (2004). The main use of GIS in this area has been centered on obtaining service areas or estimating market shares but only for discrete situations where the potential locations are prefixed. While in the context of the competitive location problem the demand usually is aggregated in a finite set of demand points due to the computational intractability, we have consider that the demand is continuous distributed along the feasible region. In GIS, the most real representation of the population distribution are the surface models of population, that is, raster maps where each pixel has associated a value corresponding to the population living in the area that it covers. These models have been developed, among other considerations, to reduce the effect that the aggregation level of the statistic information has over the demographic studies.

In the location literature some works focused to analyze the error due to the demand aggregation can be found. See, for example, Francis et al. (2008) for an interesting review about the aggregation measures for a large class of location models. For the particular case of competitive location problems the problem has been analyzed by Drezner and Drezner (1997). Of course, when an aggregation process is done, a trade off between computational time and accuracy occurs. For certain problems time factor is vital, for example for the choice of the shortest path for an ambulance to an accident, but in competitive location, we think, it is more important the accuracy than the computation time. The location of a new facility is a decision for a large temporary horizon and the firms need an efficient solution rather than an instant one.

The inclusion of forbidden regions, i.e. areas where the location of new facilities are not allowed, means an important increase in the complexity of the resolution method for the traditional Operation Research methods (see for example McGarvey and Cavalier 2005). The use of the GIS procedure we proposed in this paper allows us to exclude these regions in an easy way without consider any analytical assumption.

The result of the proposed method is a raster map where each pixel of the feasible region has associated the estimated Huff capture for a new facility located on it. This format of solution permits to use it into a multi-criteria decision environment in combination with other criteria (minimization of the distance to the distribution centers or to the main roads, etc.) as Suárez-Vega et al. (2008b) have proposed.

The rest of the paper is organized as follows. The model and the resolution method are explained in Section 2. In Section 3 the model is used to find the best location for a new hypermarket in the island of Gan Canaria (Spain). In Section 4, the continuous demand solution is compared with the results obtained using different aggregation levels for the population. Finally, we present some conclusions and future working lines in Section 5.

2. SOLVING THE HUFF-BASED COMPETITIVE LOCATION PROBLEM ON THE PLANE

The geographical information can be presented in two formats: vector and raster. In a vector layer the objects are represented by means of points, lines or pÿÿÿÿons. Each layer has associated a table where, for each element, the information for different attributes is stored. A raster layer is a matrix of cells (named pixels) which contain certain value and it can be represented by giving each pixel a color with respect to its value. Each raster map is determined by five parameters: number of columns (*ncol*) and rows (*nrow*), coordinates *x* and *y* of the lower left corner (*xllcorner*, *yllcorner*) and the size of the pixel (*cellsize*). Knowing these values, it is easy to associate a pixel and its coordinates in the map. If a pixel *P* is located at row *r* and column *c*, then $(x, y) = (xllcorner + (c-1) * cellsize, yllcorner + (r-1) * cellsize)$ are the coordinates of the lower left corner of *P*.

In this paper, the planar Huff problem has been solved within a raster environment. A new firm *A* wants to determine the best location for a new facility that must compete with the facilities that already exist in a market. Let x_0 and a_0 be the location and the quality level for the new facility and $X_p = \{x_1, x_2, ..., x_p\}$ and $A_p = \{a_1, a_2, ..., a_p\}$ the locations and the quality levels of the existing facilities. The attraction felt by customers at pixel *P* towards a facility *j* at x_j with quality level a_j is given by

$$a_{Pj} = \frac{a_j}{f\left(d\left(P, x_j\right)\right)},$$

where $f : \Re_0^+ \to \Re^+$ is a positive increasing function and $d\left(P, x_j\right)$ is de Euclidean distance between pixel *P* and facility at x_j.

Let *PM* be the surface model of population, that is, the raster map containing the demand distribution and let $w(P) \geq 0$ be the population at pixel $P \in PM$. Then, following the Huff model, the market share captured by a new facility with quality level a_0 located at point x_0 is given by

$$M(x_0) = \sum_{P \in PM} w(P) \frac{\dfrac{a_0}{f\left(d\left(x_0, P\right)\right)}}{\dfrac{a_0}{f\left(d\left(x_0, P\right)\right)} + \sum_{j=1}^{p} \dfrac{a_j}{f\left(d\left(x_j, P\right)\right)}}. \qquad (1)$$

Therefore, the following problem must be solved to obtain the solution for our competitive location problem:

$$\underset{x_0 \in F}{Max}\, M(x_0), \qquad (2)$$

where *F* is a raster map representing the feasible region containing all the potential locations for establishing the new facility. Note that under the assumption of a constant marginal cost, maximizing profit is equivalent to maximizing captured demand.

To solve the problem we have combined a C coded program with some ArcGis 9.3 tools. First, the feasible region and the population maps are exported to an ASCII file. Then the C program calculates $M(x_0)$ for each $x_0 \in F$ taking into account the coordinates of the existing facilities and their quality levels. The program stores this information in an ASCII file that can be imported to ArcGis into raster format. This final map reflects the estimated Huff capture for

each pixel in the feasible region. Additionally, not only the best locations can be pinpointed but we can also compare the suitability of the different zones.

The C program solve problem (1) for each pixel in *F*. The complexity of solving problem (1) is o(*nrowPM*ncolPM*) where *nrowPM* and *ncolPM* are the number of rows and columns of the population map. To solve (2) it is necessary to solve problem *nrowF*ncolF* times (*nrowF, ncolF* are number of rows and columns of the feasible region). The total complexity is o(*nrowPM*ncolPM* nrowF*ncolF)*. That means that the computational effort to solve the planar Huff model with continuous demand depends on the extent of the region analyzed and on the size of the pixels. The maximum accuracy for the resolution of problem (1) depends on the resolution of the population map, i.e., on the pixel size of this map. Smaller sizes than this one will not report any improvement in the calculation of the Huff capture. The pixel size for the feasible region may depend on the king of service is going to be located. For example, for locating a hypermarket with 5000 m^2 of sales surface with another 5000 m^2 of parking area usually it is not necessary to use a pixel size smaller than 50 *m* (four pixels to cover the total surface needed).

In order to reduce the computational effort it would be very interesting to analyze the existence of forbidden regions. For example, water bodies, natural protected areas or zones with high slope, can be discarded from the feasible region. The incorporation of forbidden regions using traditional OR methods introduces an important complexity in the resolution methods whereas that employing GIS tools this process may be quite easy if the information is available in digital maps.

3. AN APPLICATION

The previous tools have been used to find the more promising locations for a new hypermarket on the island of Gran Canaria (Spain). According to the Law of Commercial Activity Regulation in the Canaries (B.O.C., 1994) a hypermarket is a large store which is a combination of a food store and a department store with a minimum sales area of 2500 m^2. In this application, the quality level is measured by the hypermarkets sales surface (m^2) and we have supposed that the new hypermarket will have a sales surface of 5000 m^2.

Figure 1. Existing market in Gran Canaria

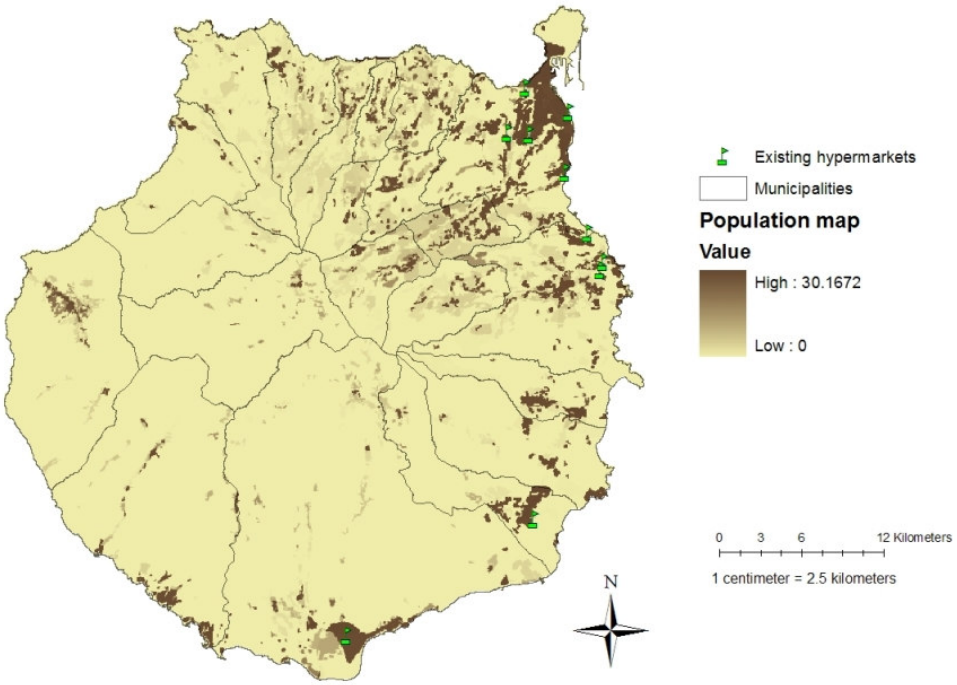


Figure 1 shows the scenario where our problem has arisen. Gran Canaria has an almost circular shape with a surface of 155824.43 Has and a population of 757821 inhabitants (Census 2006). Note that, 51.7 % of the population and 67.9 % of the sales surface of the island is concentrated in the capital (the northeastern part of the island). The rest of the hypermarkets, and an important part of the demand, are distributed on the eastern and southern sides of the island. The centre and western parts of the island are mountainous areas with a low density of population.

In 2008 there already existed in Gran Canaria thirteen hypermarkets and one in an advanced process of construction. Some very close hypermarkets were aggregated, which resulted in a set ten locations. The quality level associated to each of these locations is the sum of the sales areas of the hypermarkets allocated to it.

The population map used was elaborated by Suárez-Vega et al. (2008a). In short, a surface model of population is a raster map that stores the population of an area in a continuous model. Each pixel in the map has associated the population of the area that represents. A Land Uses map in combination with a Constructions layer was used in order to generate the ancillary map for the dasymetric method. The population densities of the polygons in the ancillary layer have been calculated taking the aggregated density of an influence area (independent of the administrative limits). Although the original population map had a cell size of 10 *m*, in this case, in order to reduce the computational effort, the data was aggregated to a pixel size of 100 *m*.

To refine the feasible region a land uses and an elevation maps were used. The land uses map (1/25000) from year 2002 provided by GRAFCAN (the official supplier of geographical information in the Canaries) was employed to discard some areas such as water bodies, beaches or protected areas. The elevation data were supplied by the Shuttle Radar Topography Mission (SRTM) and the areas with slopes higher than 30 % were discarded from the analysis to reduce the building problems.

Figure 2. Capture maps for the planar Huff competitive location problem

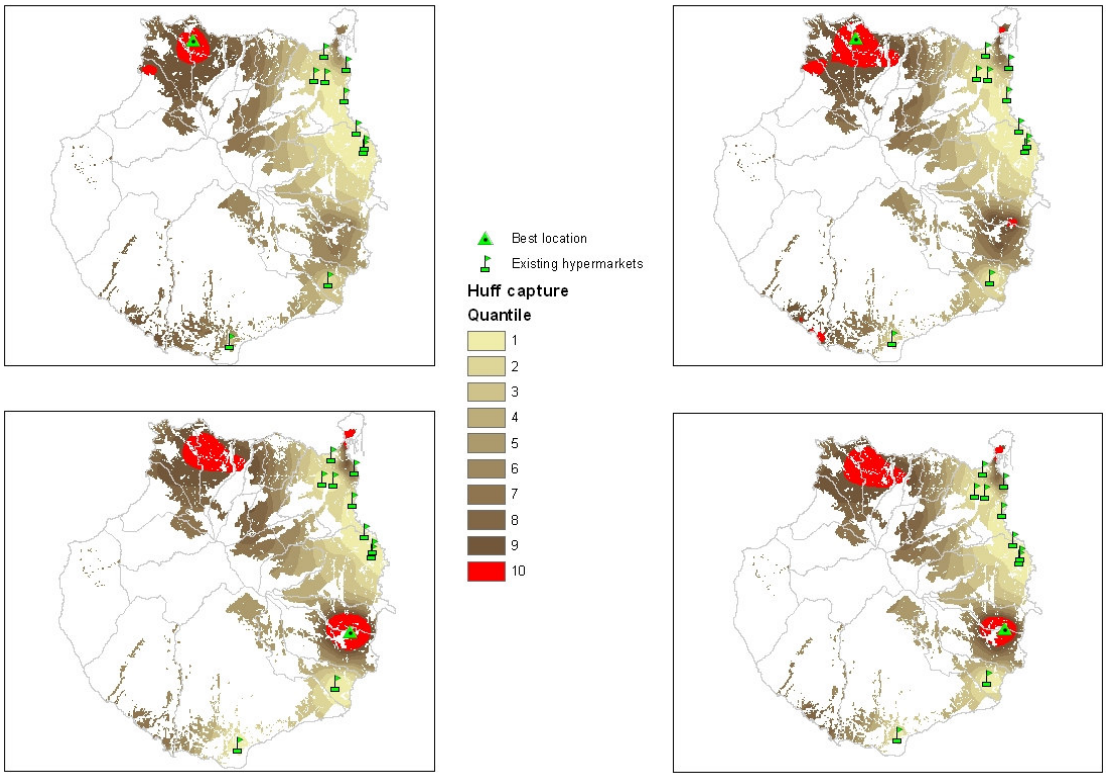


Figure 2 shows the estimated captures for the planar Huff model with continuous demand and considering that the new hypermarket has a sales area of 5000 m^2. Function $f(d)=1+d^{\lambda}$ are considered with different values of λ (0.5, 1, 2) to take into account both concave (λ = 0.5,

1) and convex ($\lambda = 2$) decay distance cost function. Note that, an increase of λ means that customers present a fewer disposition to travel. A fourth map is also presented with the average capture of the three previous cases. The maps are symbolized using the quantile method, i.e., ordering the pixels in increasing order with respect to the capture and classifying then into ten classes with the same numbers of elements each one. The most promising areas for locating a new hypermarket are painted in red. The white zones are the areas that do not belong to the feasible region. In these cases, the maximum captures for the new hypermarket are 51097.98, 57871.75, and 77520.91 for $\lambda = 0.5$, 1, and 2, respectively. Note that the optimal value is higher the greater the customers' resistance to travel for shopping.

As shows Figure 2, when $\lambda = 0.5$ or $\lambda = 1$, that is, when the distance is less penalized, the new facility is located at the North-West, in a not very populous zone, far away from the existing hypermarkets. But for $\lambda = 2$, when the distance is more penalized, the new store is located at the East, in a more populous area and between some hypermarkets. If we are not sure about the customers' perception of the distance, we may consider a combination of these scenarios. For example, the mean of the scores obtained for the different values of lambda or a weighted function of them can be used instead a particular map. In this case, we present a map with the average capture for the three scenarios as a more robust solution for the case where some uncertainty about the value of λ exists. In this case, the best location for the average map is very close to the most promising site for $\lambda = 2$.

Note that the use of these maps provides the decision maker with a broader vision of the problem and allows for comparisons between different alternatives in contrast to the traditional OR methods which only give a global optimum or, when it is not possible, a set of local optima.

4. EVALUATING THE ERROR OF CONSIDERING AGGREGATED DEMAND

In this section, the results obtained considering continuous demand have been compared with the solution resulting from different aggregation levels. The island is distributed into 21 municipalities (Level 1). Each of them is divided into entities (Level 2) and these are subdivided into nucleus of population (Level 3). Table 1 shows the description of each aggregation level. For levels 2 and 3, as the extension of the units are not very large, the population of each unit was automatically allocated to its centroid. For Level 1, in order to obtain a more accurate distribution, the aggregation points were manually allocated to the most populated area of the municipality.

Table 1. Different aggregation levels

Aggregation level	Level 1	Level 2	Level 3
Number of units	21	379	1029
Average surface (Has)	7420,21	411.15	151.43
Average population	36337.24	2013.41	741.58
Average density of population (Hab/Ha)	5. 5	6.19	21.04

Table 2 shows the results for the different demand setting and λ values. The first row shows the optimum value for the continuous model. For each aggregated model, its optimum value, the percentage of error with respect to the continuous optimum value and the distance between the best locations for the continuous and the aggregate models are presented. The continuous capture in case of locating the new store at the aggregated best location is also shown.

The aggregation levels 1 and 2 always overestimate the continuous capture whereas Level 3 underestimates it. The error decreases with the increase the number of aggregation points. When

customers present fewer objections to move to the hypermarket $(\lambda = 0.5)$ Level 2 obtains a reasonable solution. When customers' predisposition to travel decreases $(\lambda = 1,2)$, results for Level 1 and Level 2 are quite similar and very poor in contrast to the optimum obtained for Level 3. For $\lambda = 0.5$ and 1, the results obtained in Level 3 are very good with percentage of error of 0.32 and 0.21, respectively, but for $\lambda = 2$ the error presents an important increase until 16.96 %.

Table 2 also shows the distance between the best location for the new hypermarket for the different aggregation setting and the continuous best location (in Kms). The best location (for the continuous demand) is sited at the North-West of the island, when $\lambda = 0.5$ and 1 and at the East when $\lambda = 2$ (see Figure 2). For levels 1 and 2, the best locations are always at the North-East, in the capital of the island, between 19 and 24 Kms from the continuous solution. When Level 3 is considered, the results depend on parameter λ. For $\lambda = 0.5$ and 1 the best site for the new hypermarket coincides with the best location for the continuous case. However, when the distances are high penalized, the aggregated model locates the new hypermarket in the South-East, at 31 Kms of the continuous best location. Although this is the most distant alternative to the continuous location, it is the closest site to the best location resulting when the average capture between the three values of λ is considered.

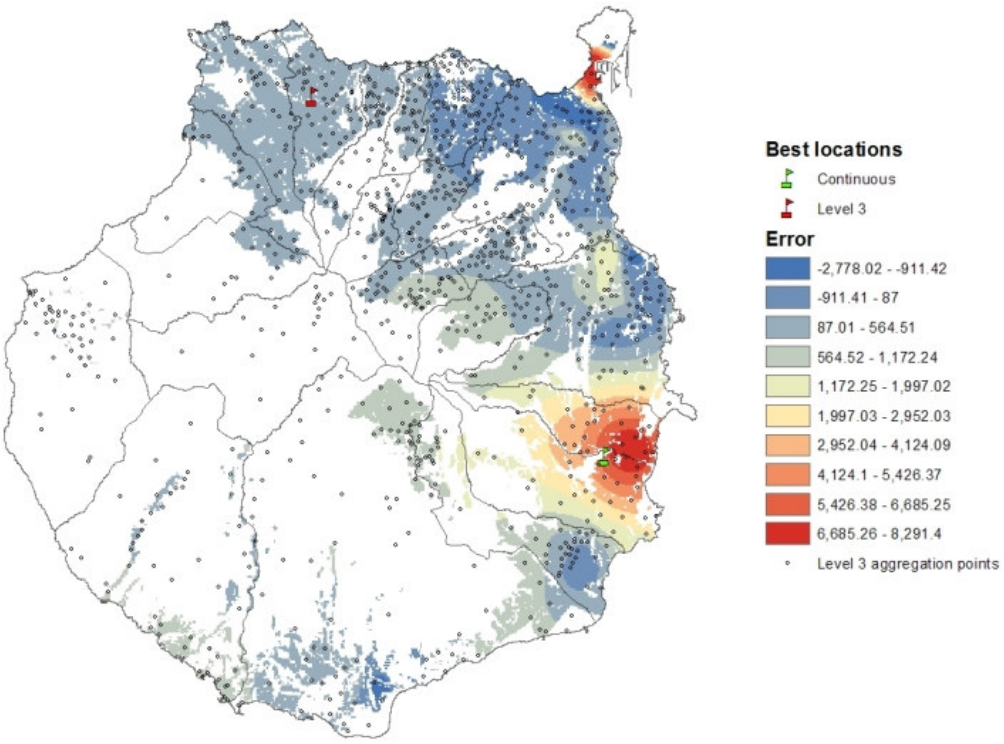
Table 2. Results for the different demand aggregation levels

	$\lambda = 0.5$	$\lambda = 1$	$\lambda = 2$
Continuous demand			
Optimum value for continuous demand	51097.98	57871.75	77520.91
Aggregation Level 1			
Optimum value	84255.78	109151.56	171986.78
% error	99.35	99.11	98.78
Distance to the best continuous location	19.61	19.72	22.68
Continuous capture at the best location	47506.80	49594.50	61078.00
Aggregation Level 2			
Optimum value	51948.16	105616.49	170795.70
% error	1.66	82.50	120.32
Distance to the best continuous location	20.07	20.01	24.02
Continuous capture at the best location	48725.58	52562.18	68754.44
Aggregation Level 3			
Optimum value	50932.46	57751.09	64379.05
% error	0.32	0.21	16.95
Distance to the best continuous location	0	0	31.39
Continuous capture at the best location	51097.98	57871.75	74716.48

If the optimum values for the aggregated models are compared with the captures resulting for the continuous model at their best locations, we can see that for levels 1 and 2 the capture for new firm is always less than the expected one. For example, for Level 2 and $\lambda = 2$, the continuous capture is 35.51 % from the expected one. For these two aggregation levels the lost of estimated capture decreases with the value of λ. For Level 3, the estimated capture is less than the continuous one. For $\lambda = 0.5$ and 1 the increase is very small but when $\lambda = 2$ the new store captures a 16 % more than the expected.

As we have previously commented, the result of this GIS tools is a map that reflects the estimated Huff capture for each pixel in the feasible region. These maps can be compared in order to obtain error maps. Figure 3 shows a map representing the error resulting from the use of aggregation Level 3 instead the continuous model. Using this map, the high underestimate zones (in red) and overestimates areas (in dark blue) can be clearly detected. In this map, all the aggregation points for Level 3 and the best location for both the continuous and aggregated models are shown.

Figure 3. Error map for aggregation Level 3



5. CONCLUSIONS

In this paper, we have used a GIS based model to solve the planar Huff problem. Most of the papers related with the competitive location problems consider that the demand is aggregated in a finite set of points. In this case, we have considered that the demand is distributed along the plane according to a surface model of population, that is, a raster map where each pixel has associated a value corresponding to the population living in the area that it covers. Taking into account the demand distribution and the location and size of the existing facilities, we have obtained a raster map where each pixel has associated the estimated capture of a new competing facility if decides to locate on it. Using these capture maps, not only the best locations can be pinpointed but we can also compare the suitability of the different zones.

A real example is presented where a new hypermarket is located on the island of Gran Canaria (Spain). Different decay distance function have been considered in order to take into account both concave and convex function. For the studied example the maximum capture is higher the greater customers' resistance to travel for shopping. When the distance is less penalized the new facility is located at the North-West, in a not very populous zone, far away from the existing hypermarkets. But when the distance is more penalized the new store is located at the East, in a more populous area and between some hypermarkets. If we are not sure about customers' perception of the distance, we may consider a combination of these scenarios. In this example the mean of the scores obtained for the different decay function was used in order to obtain a more robust solution to the problem. For the average capture map, the best location was close to the best one when the distance is more penalized.

Finally, for this example, the solutions for the continuous demand case and for the discrete demand case are compared. Three levels of aggregation were considered. The accuracy of the aggregation model increases with the number of demand aggregation points. For the less aggregated level (Level 3), when the distance is low penalized the discrete model achieves optimum values and locations very close to continuous solutions. But when the distance

represents a high weight, an error of 16.95 % for the optimum value appears and the discrete best location is sited at 31 Kms from the continuous best location.

This work can be improved making an analysis of sensibility with respect to the size of the new facility or the cell size for the feasible region or the population map. A new line of work may be to consider network distances instead the Euclidean one.

ACKNOWLEDGEMENTS

This paper has been partially supported by Ministerio de Ciencia y Tecnología (Spain) and FEDER (grant BFM2002-04525-C02-01) and by Ministerio de Ciencia e Innovación (Spain) (grant: ECO2008-05589/ECON)

REFERENCES

BIRKIN, M., G. CLARKE, M. CLARKE AND R. CULF (2004): "Using spatial models to solve difficult retail location problems in Applied GIS and Spatial Analysis", J. Stillwell and G. Clarke (eds). Willey, England.

BIRKIN, M., G. CLARKE, M. CLARKE AND A.G. WILSON (1996): *Intelligent GIS: Location decisions and strategic planning*, Geoinformation, Cambridge.

(B.O.C., 1994) Decreto 219/1994, de 28 de octubre, B.O.C. nº 140, November the 16th, 1994.

CONVERSE P.D. (1949): "New laws of retail gravitation", *Journal of Marketing*, 14: 379-384.

DREZNER, T. (1994): "Optimal continuous location of a retail facility, facility attractiveness, and market share: An interactive model", *Journal of Retailing*, 70(1): 49-64.

DREZNER T AND Z DREZNER (1997): "Replacing Continuous demand with discrete demand in a competitive location model", *Naval Research Logistics*, 44: 81-95.

FRANCIS, RL, TJ LOWE, MB RAYCO AND A TAMIR (2008): "Aggregation error for location models: survey and analysis", *Annals of Operations Research*, DOI: 10.1007/s10479-008-0344-z.

GRAFCAN (Cartográfica de Canarias S.A.): http://www.grafcan.com/

HUFF DL (1964): "Defining and estimating a trading area", *Journal of Marketing* 28: 34-38.

MCGARVEY GR AND TM CAVALIER (2005): "Constrained location of competitive facilities in the plane", *Computers & Operations Research*, 32: 359-378.

REILLY WJ (1931): *The law of retail gravitation*, *Knickerbocker Press*, New York.

SHUTTLE RADAR TOPOGRAPHY MISSION (NASA) http://www2.jpl.nasa.gov/srtm/

SUÁREZ-VEGA, R, DR SANTOS-PEÑATE, AND P DORTA-GONZÁLEZ (2008a): "Generación de un modelo superficial de la población de Gran Canaria" in the proceedings of the *XIII Congreso Nacional de Tecnologías de la Información Geográfica*, Hernández-Calvento and Parreño-Castellano (eds). Las Palmas de Gran Canaria (Spain).

SUÁREZ-VEGA, R, DR SANTOS-PEÑATE, AND P DORTA-GONZÁLEZ (2008a): "Using GIS to solve a competitive location problem" *in the proceeding of XII Congreso Nacional de Economía Aplicada*. Barcelona (Spain).

MISUSE OF MATHEMATICS IN THEORY OF PRICE IN ECONOMICS LEADING TO INCONSISTENT AND ANOMALOUS INFERENCES IN LAW OF DEMAND, INDIFFERENCE CURVES AND LOANABLE FUND THEORY

ACHYUT OAK
E-mail: achyutoak@gmail.com
Govt of India

Abstract

Demand Curve is introduced by taking some arbitrary data of price and quantity(apples/wheat etc) and a graph, generally a straight line with negative slope is drawn. In the following discussion this line takes the shape of a curve convex to the origin (called 'Demand Curve') within the first quadrant. It is then stated that it is proved ' that the demand curve slopes downwards from left to right and is convex to the origin' and this curve becomes the 'Law of Demand'. Now geometrical properties of a curve or a set of curves (convex to the origin and in the first quadrant) are stated and it becomes a demand function. Not a single mathematical equation is given (other than f(x,y)) for any demand curve. This demand curve starts with the assumption that utility measure is cardinal.

In Indifference Curves(ICs) one never finds the nature of these curves. Are they system of parallel lines, system of parabolas,concentric circles, ellipses one inside the
other ,system of hyperbolas one inside the other and all restricted to the first quadrant? No mathematical equation is given which would represent any IC or ICs. The concept of IC starts with the assumption that the measure is 'Ordinal' (no quantification).When there is no variable ,either discrete or continuous , still use of differentiation/ partial differentiation is made. . When the measure itself is considered ordinal the consumer is expected to follow mathematical conditions i)ICs are convex to the origin ii)ICs are sloping downwards to the right iii) ICs can not intersect or touch each other iv) ICs can not be tangent to each other and added to these further assumptions of i) Non-satiety ii) Transitivity and iii) Diminishing marginal rate of substitution.

In Loanable Fund Theory, Loanable fund on X axis and rate of interest Y axis. Now demand curves for Dis-savings, Investment and Hoarding are drawn but they all are strictly straight lines and are parallel to each other. Similarly supply curves for Savings, Dis-investment ,Dis-hoarding and Bank Money are drawn again all of them a set of parallel straight lines. These two sets of parallel lines are separately added to generate two lines of summations and they must intersect in the first quadrant. No longer there is a need of demand curve to be convex to the origin and a smooth curve. This is all for getting a point of equilibrium in the first quadrant. In real life situations neither the supply curve nor the demand curve in loanable fund theory are straight lines,barring few situations. If they are to be straight lines then the important factors like time,duration,risk, etc get ignored.

The author feels that this is total misuse of mathematics in price theory. Author does not suggest any alternative to the present approach but gives more importance to highlight the misuse.

Thematic Area: Quantitative Methods for Economics and Business.

I)LAW OF DEMAND

In Economics the theory of price hangs on many assumptions and on the single weak thread of assumed 'Demand Curve'. Some data of price - wheat / price - apples, etc. are taken so as to suit the straight line drawn beforehand. In most of the books thus is the demand curve defined. To look it more mathematical, within the next few lines it takes the shape of a curve (with negative slope,convex to the origin, sloping downwards from left to right) which facilitates the

researcher's predetermined inferences. Demand curve assumptions regarding price, quantity, taste, time, market conditions etc are not realistic but they are taken because of the advantage of simplicity and convenience. In further discussion it is stated that we have already proved ' that the demand curve slopes downwards from left to right and is convex to the origin' At this stage this becomes the 'Law of Demand'. Nowhere one finds any justification for calling it a law of demand or a demand curve. Now geometrical properties of a curve or a set of curves (convex to the origin and in the first quadrant) are stated and it becomes a demand function. No author of any book prefers to give a single mathematical equation (other than f(x,y)) for any demand curve. This demand curve starts with the assumption that utility measure is 'cardinal'.

Some explanations from different books are cited below . These would show that while developing the concept of demand no justification is given for inferences.

1. With the use o f arbitrary data, points are plotted , a line is drawn and this becomes the 'Demand Curve'. Similar is the case with supply . The point of intersection of these assumed demand and supply curves becomes the point of equilibrium. The author of the book does not feel it necessary to justify straight lines and also the existence of point of equilibrium in the first quadrant. On `page 97 demand curve takes the shape of rectangular hyperbola. Further the following arbitrary equation is taken for quantity X.
Quantity of Good X = Constant – A * price of good X+ B* price of Y+ C * per capita disposable income.[27]

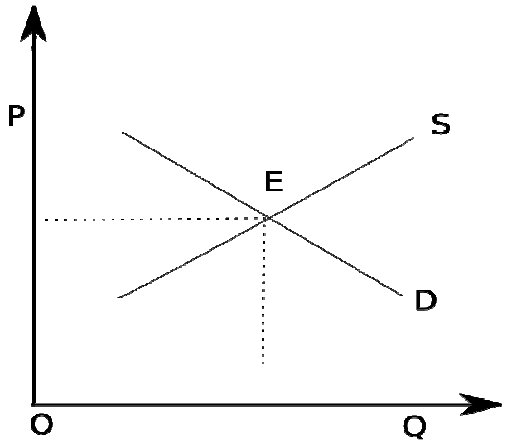


Fig 1

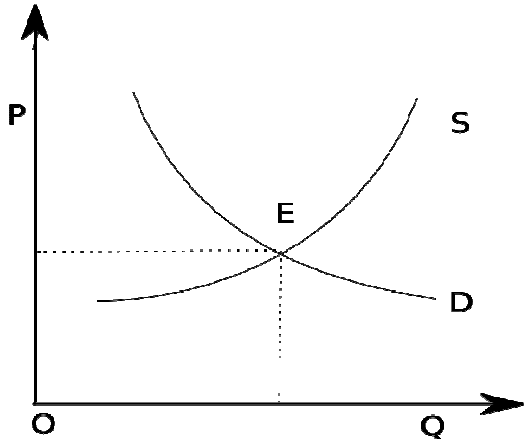


Fig 2

X1 = f(P!,P2.....Pn,Y,A,alpha) , A is advertising outlays and alpha is other influences on demand. Another equation, Demand X = Constant-A* P1 +B* P2-C* P3 +D* Y +E. These two equations are given with no supporting arguments. The derivation of these equations is not addressed at all. But still the nature of the curve remains straight lines.(7)

2. With no explanation the curves are drawn so that demand and supply curves (not straight lines) intersect in the first quadrant and the author gets point of equilibrium. Fig2 , (5)

3. QDX = f(Px,I,PY,T), QDX is quantity demanded,PX is price of commodity X, PY is price of related commodity,T is tastes of consumers. Another equation similar to equation in (1) above is given QX = Constant-A* PX + B* I +C* PY-D* PS+E ;however no elaboration is made.(32)

Some arbitrary data is considered. Nature of demand curve is assumed. Tastes, incomes,prices of other goods, new substitutes, prestige involved etc are totally ignored. No explanation is provided to justify the nature of the curves Fig 2. Both the curves drawn are not straight lines. (21)

4. Both the curves are drawn as straight lines Fig.1 The assumptions are made for no change in tastes,weather,incomes,prices of other related goods, expectations .(19)

5. Utility : 'As a metaphysical concept utility has to be defined in terms of itself. Utility is

the characteristic of commodities which makes individual want to buy them and individuals buy commodities to enjoy utility in consuming them.' This concept of utility is further used to define demand.(31)

6. Some arbitrary data is used to draw demand curves. They are not straight lines but cut curves which are convex to the origin . Fig 2 Individuals curves are added ad final demand curve is drawn. No explanation is provided for the nature of the curve. But it states that this is the demand curve. Similar treatment is adopted for the supply curve. No equation is provided .(18)

7. 'Utility is measurable .Price of a thing can be taken as a measure of its marginal utility to him'. The author is not clear about both ,Marshall's treatment and Hick's treatment about the demand theory.(30)

8. Utility is the want satisfying power of a commodity on a service which determines the demand for commodity. Jevons called this 'utility'. He further states that utility is purely a mental concept and is a measure of satisfaction derived by an individual from the consumption of a commodity. Utility is supposed to be a measure of Law of Demand.(15)

9. ' Law of Demand states that other things remaining same a household or an individual consumer will demand more of goods at a lower price than at a higher price' . 'Marshall's theory of demand is based on law of diminishing marginal utility, combined with a number of assumptions'(28)

In almost all the books dealing with Price Theory (mainly utility, demand and supply) there are many variations in the definition of the theory. Every author has his own way of interpretation, assumptions, inferences and geometrical representations . May be there are about 20/30 geometrical figures to elaborate demand curve development in every book but one does not find a single mathematical equation to satisfy even a single figure. If there was one then estimation,testing the fitness etc would have enabled to justify the nature of the lines or curves. But basically the theory and the economists do not need any mathematical involvement as that would jeopardize the entire structure of the price theory.

Geometrical figures used in this theory have only ornamental value.

If only even one concept in price theory gets quantified it has a bright future in the present days of computers and advanced softwares. In the absence of this aspect one can find there is no growth and advances in this theory. Even modern economists talk of mathematical models but one hardly sees geometrical representations and any mathematical equations which would help in estimation and testing.

II) 'INDIFFERENCE CURVES'(ICS)

First, we consider some definitions/descriptions/assumptions made by various authors in their books on Indifference Curves. In almost all the books, data to suit the curve or the curve to suit the convenience are drawn first and then assumed properties are shown as observed by the curve. To suit further assumptions, ICs are drawn so that they do not touch or intersect and have negative slopes. ICs have come through the research of Edgeworth (1881), I.Fisher (1892), Pareto (1906), Slutsky (1919), Hicks and Allan (1934) and Hicks (1939).

The theory and the concept are being studied and taught for the last 70 years or so. The researchers found the cardinal measure 'Marshallian Utility Measure' unsatisfactory and unconvincing . Two economists Hicks and Allan developed the concept of ' Indifference Curves' and considered it to be an *Ordinal Measure*. The concept has three assumptions:

i) Non Satiety

ii) Transitivity and

iii) Diminishing Marginal Rate of Substitution.

For the purpose of this paper these assumptions are accepted. Further for the ICs, the governing properties are:

a) ICs slope downwards to the right
b) ICs are convex to the origin
c) ICs can not intersect each other

d) ICs can not even meet or touch or be tangent to each other.

These governing properties are themselves all assumptions. No mathematical explanation is provided to support the same. In spite of the above assumptions geometrically and physically the measure is supposed to be 'ORDINAL' that is no quantification and only qualitative measure of grading or rating is possible.

At no stage in the development and the use of ICs, is any mathematical equation used, to represent the ICs . Only f(x,y) or some similar notations are used . This section of the paper deals with the intricacies and contrasts developed in the situation . If two or more ICs are to observe the above conditions, notional , geometrical and mathematical, then what are they actually?

Let us see what different authors have to say in the matter.

- ' IC shows various combinations of commodity X and commodity Y that yields equal utility or satisfaction to the consumer [(29)]
- Curve is shown (Jeans against CDs) with a comment that preference must be a) rational b) consistent and c) complete to define ICs[5]
- No data .Only two smooth curves are drawn (Gas against electricity units) to define ICs.The author further states that the practical problems of constructing such curves are virtually insurmountable but this does not detract us from their theoretical explanation.Accept that i) Individual ICs do exist ii) they can be meaningfully combined to yield community ICs and iii) they are convex to the origin. The curve is generally postulated as convex to the origin and the slope is declining.[(20)]
- IC is the locus of points involving combinations of rates of combinations of the two products or sets of products X and Y with all points on the curve yielding same level of satisfaction. It is ordinal in character , only preference is asked from the consumer. Example is of bread and wine. Curve is drawn first and then inferences are drawn[(17)]
- Diagram is drawn directly . Think of the consumer as sliding up and down the IC. Then he is equally well off at any point on the IC. Commodities in question are divisible into very small units. There is convenience in drawing a smooth curve. Preference means moving North-East on the map IC. Maps are drawn so that curves appear to be parallel to one another. Parallelism is devoid of economic significance.[(37)]
- Economists do not believe that utility has a cardinal measure. Only ordinal measure is considered, i.e. ' much or less' and not 'how much'. Consumer behavior in terms of preferences which in turn are defined only by behavior. The results are circular. Hypothesis places definite restrictions.[(29)]
- Shape of IC is predetermined. It is either steep or shallow.. ICs are convex to the origin.[(34)]
- ICs are negatively sloped. It shows diminishing rate of substitution . The measure is ordinal.[(32)]
- Utility is ordinal. Indifference map may have au+b, alogx, au.u. Indifference map is a system of curves u(x1,x2) = constant .I_1,I_2,I_3 ... system of curves are drawn in the OX1X2. as indifference map.)
- With no factual data ICs are drawn , first a three dimensional diagram and then a two dimensional diagram. An inference is drawn that if two points are on the same IC they have the same utility. A comment is made that 'we have come to a remarkable thing about ICs..... with wide economic significance.....given wants.....scale of preference [(9)]...

It can be seen that most of the authors have presumed the existence of the ICs in the shape of a curve

i. convex to the origin,
ii. with negative slope,
iii. not intersecting with each other
iv. not tangent to each other and
v. having an ordinal measure.

The author of this paper sees some of the possibilities. The diagrams are in the XY plane and restricted to the first quadrant. Only in the cases of hyperbola all the four quadrants are shown. This is only to show that the dichotomy that creeps in if the curves are to observe the

above imposed conditions. None of the definitions/ descriptions / assumptions of ICs by different authors of various books referred have given any significance to its presumed ordinal characteristic.

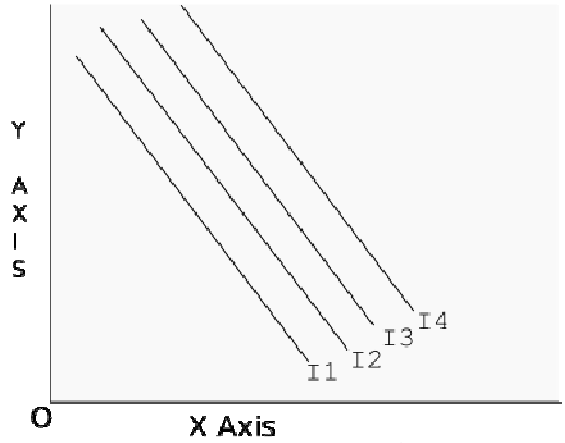


Figure 3: System of Straight Lines

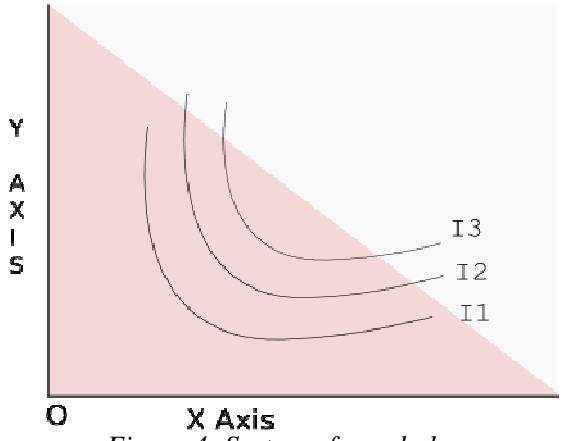


Figure 4 :System of parabolas

I) In figure No.3 the ICs are parallel to each other within the imits of goods X and Y

II) In fig.4 a system of parabolas, one inside the other, is shown. Each curve represents one IC and at any point on one curve, the same amount of satisfaction is derived by the consumer. The IC above it gives locus of points giving equal amount of satisfaction but more than the satisfaction on the curve below.

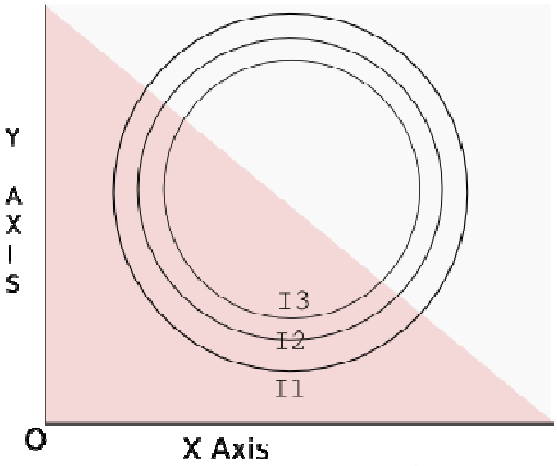


Figure 5: System of Circles

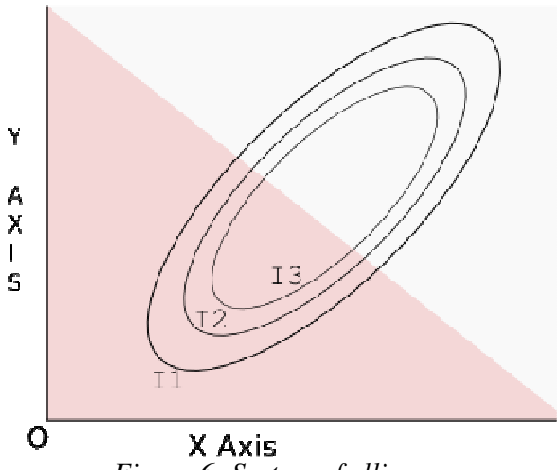


Figure 6 :System of ellipses

III) In fig 5 ICs are a system of concentric circles till the time the curve of the circle is convex to the origin.

IV) Similarly fig. 6 shows a system of ellipses one inside the other till the curves are convex to the origin.

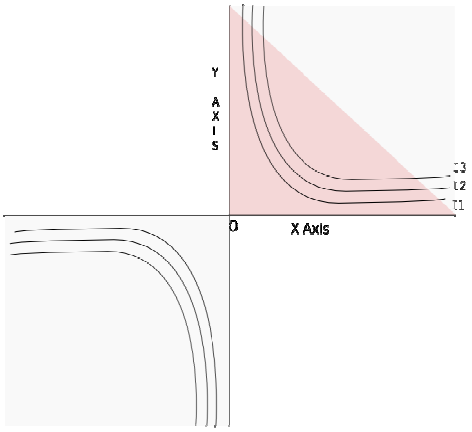


Figure 7: System of Hyperbolas

V) Fig 7 shows a system of hyperbolas restricted to the first quadrant and also that the curves are convex to the origin.

VI) ICs have to be symmetrical about a point . If this property is not observed by the ICs they are likely to intersect each other at some point or the other.

VII) Why almost all the books or papers do not give any mathematical equation(s) to represent the ICs.

VIII) The level /grade/ranking of consumer's satisfaction ,is it a discrete or a continuous variable. By very nature of assumptions it has to be 'Ordinal '

IX) Even though the commodities along the X axis and the Y axis seem to be discrete, the points on the IC (Amount of Satisfaction) are ordinal . In further analysis use of differentiation, partial differentiation is made and inferences drawn. When there is no continuous variable on the two axes,how far this use of mathematics is justified?

X) Arbitrary and suitable curves are drawn to show and justify substitution and price effects.

XI) Do ICs really support ordinal measurability?

There are ambiguities and inconsistencies in the argument and assumptions in the development of the theory of ICs.

Further once tool of ICs is ready what purpose it serves? With so many constraints and no such market prevailing any where it has become a futile theoretical exercise based on non tenable conditions. So many mathematical curves spread over the development of ICs but there is not a single mathematical equation to justify the nature of the curve. Without accounting for nature of the variable ,discrete or continuous , the tools of differentiation and partial differentiation are used freely. This shows no accountability for such uses.

It was Galileo who realized the importance of measurement and wrote,

Count what is countable,

Measure what is measurable ,

And what is not measurable ,

Make it measurable.

If the consumer is able to give his preference to particular combination of two goods (generally discrete in nature) by grade or rank only while observing all other conditions,mathematical and some assumed, the preference can not possibly be 'ORDINAL' Further he is also able to distinguish his level of satisfaction even because of very small changes in either of the goods. There is no such possibility.

The author of this paper feels that the very foundation of 'indifference curves' that the measure is 'ordinal' becomes doubtful and not tenable. No purpose is being served by the existence of ICs in the present form.

III) LOANABLE FUND THEORY IN PRICE THEORY

In Economics the word 'Interest' has been variously defined and interpreted. Initially interest was conceived as earnings of capital or rate of return on capital. This was also considered as productivity of capital. This rate was further classified as

- market rate of interest is the rate at which funds can be borrowed from the market
- natural rate of interest is the rate of return on capital investment
- when natural rate of interest is higher than the market rate of interest then there will be greater investment in capital with the result that the natural rate of interest would fall
- equilibrium will be established when the two rates would be equal.

Some economists considered interest from savings point (time preference) of view where as economists like Knight and J.B.Clark explained the interest from demand for capital (productivity of capital) point of view. Irving Fisher, Bohm Bawerk and some other economists took both factors into account while dealing with interest. These were considered classical theories.

Neo-classical economists such as, Wicksell, Ohlin, Haberler, Robertson etc developed another theory known as 'Loanable Fund Theory' . These economists thought of interaction of monetary and non-monetary forces in the determination of the rate of interest.

Keyens stressed the point that interest is purely monetary phenomenon and is determined by demand for and supply of money. According to his theory, interest is price not for the sacrifice of waiting or time preference but for parting with liquidity. His theory is known as liquidity preference theory of interest.

Later economists ,J.R.Hicks, A.P.Lerner and A.H.Hansen considered all the three approaches and developed yet another approach. The concept of interest itself generated lot many controversies and there was no consensus amongst the economists. Interest was characterized by so many factors , payment for use of capital, reward for taking , reward for inconvenience , reward for securities, credit worthiness of the borrower, period of loan, amount of loan, difference due to distance, difference in productivities, market imperfectness, etc. Classical theory generated lot of criticism and it still continues.

The present paper deals with 'Loanable Fund Theory' which falls in the category of neo-classical theory. It is presumed that the reader is aware of loanable fund theory as it exists to day. This would eliminate the need for preliminary introduction and meanings of usual terms used in economics. It is divided for the purpose of discussion under eight sub topics.

- Present Approach
- Assumptions
- Factors under Supply with geometrical representation
- Factors under Demand with geometrical representation
- Summation curves and point of equilibrium with geometrical representation
- Ambiguities and inconsistencies
- Utility of Loanable Fund Theory in further development of itself and economics in general
- Inferences.

Consider each sub topic separately.

1 Present Approach

Loanable fund is considered on the axis of X and the rate of Interest on the axis of Y.

Two main considerations are demand and supply.

Demand for funds can come from: i) funds for investment (I) At lower rate of interest the demand for funds would be higher and vice versa ii) funds for consumption or dis-investment (DS) and iii) demand for Hoarding (H). The notations are not universal.

Supply for loanable funds comes from i) savings of individuals, households, business concerns joint stock companies,corporate houses etc.(S) ii) Dis-hoarding of past accumulated savings(DH) iii) Bank Money through sale and purchase of securities,advancing loan to businessmen (BM) and iv) Dis-investment can come from depreciation reserves,withdrawal of

working capital,etc.(DI)

All the seven factors , three under demand and four under supply are represented by straight lines . Further following equations are taken

Demand for loanable funds(DL) = I + DS + H

and Supply of loanable funds (SL) = S + DH + BM + DI

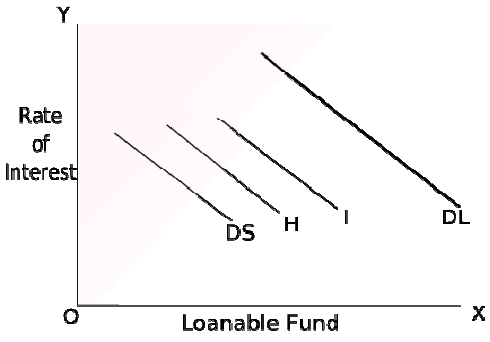
and for equilibrium DL = SL

1. Assumptions

In the present approach it is presumed that all seven contributing factors to demand and supply of loanable funds , loanable funds on the axis of X and the rate of interest on the axis of Y are represented by straight lines . All demand lines are sloping downwards to the right ,having negative slopes and all supply lines have positive slopes. Further their summations ,DL and SL, are also straight lines having negative and positive slopes respectively and that they intersect in the first quadrant to give point of equilibrium.

Comment :Any assumptions made while initiating the theory and further developing it need to have sound reasoning , need to stand simple practical situations and mathematical tests. No such thought has been given for these assumptions.

2. Geometrical Representations for Demand and
3. Geometrical Representations for Supply



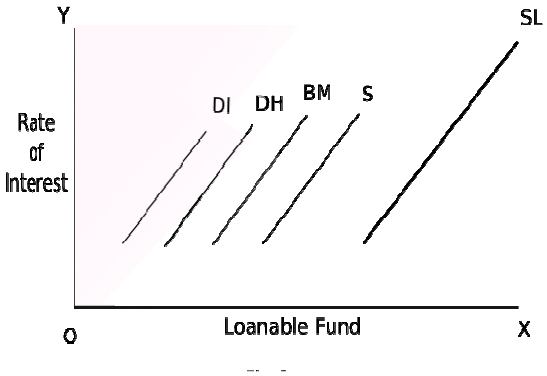


Fig 1 **Fig 2**

Demand Curves
DS - Dis-savings
I – Investment
H – Hoarding

DL= DS+I+H

Supply Curves
S- Savings
DI– Disinvestment
DH – Dis-hoarding
BM – Bank Money
SL =S+DI+DH+BM

Almost all the books covering the topic of loanable fund theory represent these seven components by straight lines .The three under demand are to be parallel to each other with negative slopes and four components under supply are parallel to each other with positive slopes.

Further the three lines are added and one gets the summation curve ,again a straight line(DL)Similarly four lines in the other set are added and one gets summation curve ,again a straight line(SL)Another assumption is made that these two DL and SL must intersect and that too in the first quadrant.

Comment : why they should be i) straight lines ii) parallel to each other separately for demand and supply iii) further must intersect in the first quadrant. No reasoning is given.

4. Summation of curves, Intersection,the point of equilibrium : Geometrical representation

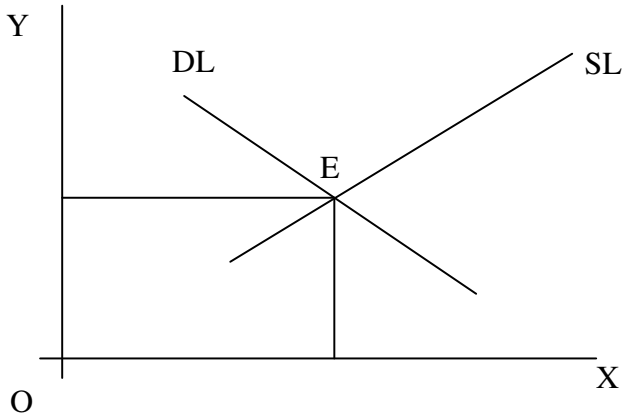


Fig 3

Fig. No.3 above shows the two summation lines DL and SL and their point of intersection E. This is point of equilibrium. We are not entering into the significance of this point E with reference to economics.

Comment : All this geometry is used with no mathematics involved . From the data ,if available whether the straight line fits the data or otherwise can be tested by statistical methods. But this is not done as initial hypothesis of the straight line for each factor has no validity or reasoning. It seems economists do not need this mathematical rigor. A decent geometrical figure gets easily appreciated by any reader. There is no accountability for such assumptions.

5. Ambiguities and Inconsistencies seen

Consider how many assumptions are made which do not exist and are not feasible in practice and are also not mathematical.

A) All the seven parameters, in demand and supply together , which are represented by straight lines are not so in practice, barring few situations. Many important factors (time, duration, risk, productivity, taxation, market standing etc.) are ignored just to get the parameter to follow a straight line.

B) Even if supposedly they are represented by straight lines then why they should be parallel to each other except to suit the convenience of the theory with no accountability for the parallel lines. .

C) The summation lines of demand and supply must intersect in the first quadrant . This is because the point of equilibrium, geometrically has to come in the first quadrant to suit the assumptions. For all assumptions there is a clause that all other factors remaining same. But the area, in which loanable funds operate, is so fluctuating that it is just not possible to fulfill this condition.

D) Summations must be straight lines. In the development of the theory , no where one finds any data, actual or even arbitrary, being used . It is all free hand drawing to the liking of the economists .There are statistical tests for testing of fitness of the curve with the help of confidence level. In the absence of any actual data the question of testing does not arise. In any science these conditions can not be accepted

E) Why data is not available? This is because in actual real life situation(s) all these assumptions of straight lines would just vanish. To extend the argument further isolation of interest rates and availability of loanable funds from so many important factors is just not possible.

F)Consider fig. No. 4

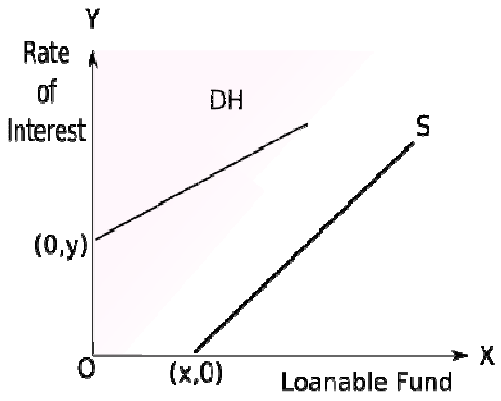


Fig 4

Under supply, savings is one component. Irrespective of any rate of interest ,high or low, there is bound to be fund available which individuals and households keep for emergency and safety. Considering the huge number of individuals and households, this fund is bound to be very large. So even for zero interest there is bound to be fund available. It shows that the line for savings would have a point with coordinates (x,0), x > 0.The line if at all there is one, it perhaps would be as drawn in Fig4 .

Under supply, dis-hoarding is another factor. This factor is generally not in terms of monetary terms. (For example , Cement, Steel, Cotton, Food Grains,Gold, Shares etc) . Hoarding/ dis-hoarding can not be in the form of monetary units as it would involve possible violation of law. But most of the books do no discuss whether hoarding/dis-hoarding is in terms of monetary or non-monetary units as they do not cite any examples while speaking of dis-hoarding. This factor would have a point with coordinates (0,y) , y > 0 fig 4. Thus it shows that the line of savings and the line of dis-hoarding under loanable funds can not possibly be parallel to each other. Further somebody's dis-hoarding is someone else's investment therefore it does not make additional fund available in the market.

Under supply, another component is bank money. Because of the very nature of the business, bank money could be in the form of monetary units, guarantees, securities, shares etc, so irrespective of rate of interest prevailing in the market the fund would always be there and a point on this line, if assumed , will have coordinates (x,0) ,x>0. The point stressed here is that the lines drawn for convenience can not be parallel to each other.

G) The summation of these assumed is again taken as straight line and that too parallel to the lines representing the three or four factors. The argument again here is that in the absence of any actual data the fitness of the line can not be tested. Still to get the point of equilibrium, these lines are freely used.

Under supply, savings, disinvestments, dis-hoarding and bank money need not be straight lines. Under some situations if they are straight lines they need not be parallel to each other.

Similar arguments are true for components under demand

H) No book refers to demand in the form of goods like,steel,cement,food grains, shares, gold, securities etc and aspects like weather, govt. policies, war,famine etc but presumes in the form of monetary values. Hoarding ,dis-hoarding, consumption, cost of hoarding, perishable goods, weather conditions, rains, floods

etc are given no weight in the analysis of the loanable fund theory.

I) Loanable Funds and rate of interest can not be studied in isolation. Therefore the Theory itself is lop sided and has no relevance to the environment outside.

J) In the past there have been objections and criticism to loanable fund theory from different Economists . The points raised in the past were, income not a constant but a variable, savings and investment schedules are not independent , indeterminate theory, neglect of sources of savings, definition of interest unrealistic assumptions of full employment etc.

6. UTILITY OF LOANABLE FUND THEORY IN THE DEVELOPMENT OF ITSELF AND ECONOMICS IN GENERAL

i) As no factors are quantifiable the theory itself has remained static for years together. Even with the availability of sophisticated computers and also the soft ware this theory could not make any progress as it fails miserably in mathematical rigor. Subjects which were considered highly subjective are making use of computer soft ware in practice.

ii) Article by Jorg Bibow, of Hamburg University " The Loanable Funds Fallacy: exercises in the analysis of disequilibrium"

Abstract :-This essay analyses the core proposition of loanable funds theory that changes in technology and time preferences directly and immediately affect interest rates . Applying what can be seen as generalized financial- buffers approach to the analysis of disequilibrium , we find that loanable fund s theory is flawed and should therefore be abandoned. A challenge to the neo-Walrasian general equilibrium approach to monetary theory remains : that of justifying the idea that -by some mechanism-inter temporal prices correctly reflect technology and time preferences."

7. INFERENCES

This is a chain of incorrect assumptions, illogical inferences and free use of geometrical figures to suit the development of theory and there are no mathematical equations. Thus loanable fund theory is poetry with little logic and no mathematical backing. One sees no use of this theory in economics. Then why teach this theory and enhance untenable and unproven way of thinking in the minds of students of Economics. Loanable funds is worth studying but not the loanable fund theory as it is taught to day. It does not deserve a place in the modern curriculum right from graduate level to post graduate level. Let it be in history of Economics.

The author therefore feels that price theory in economics has misused mathematics and further no concrete and realistic results have come forth.

REFERENCES:

1. AHUJA H.L. ' Principles of Microeconomics', S.Chand and Co. New Delhi,India,1998
2. ALLAN R.G.D.,' Mathematical Economics' Macmillan Ltd ,New York, 957
3. BOULDING K.E. , 'Economic Analysis ', Vol .I and II,Harper and Row, 1941
4. BOWDEN E.V., ' Economics the Science of Common Sense ', Federal Pub, 1977
5. COOK MARK AND COVVI FARQUHARSON,' Business Economics', Pitman Pub,1998
6. DAVID DE MEZA AND OSBORNE, ' Problems in Price Theory '
7. HEINEKE, JOHN M, ' Microeconomics for Business Decisions, Theory and Application', Prentice Hall, 1976
8. HIRSHLEIFER JACK AND AMIHAI GLAZER , ' Price Theory and Applications' ,
9. HICKS J.R. ,' Value and Capital ', Oxford Uni Press,1939
10. HICKS J.R., 'The Theory of Economic History', Oxford Uni Press, 1969
11. HICKS J.R. ,'The Social Framework,An Introduction to Economics' Oxford, 1952
12. HICKS J.R., 'A Market Theory of Money' ,Clarendon Press, 1989
13. HICKS J.R. ,'A Revision of Demand Theory' , Oxford Press,
14. JHINGAN M.L., 'Microeconomic Theory' , Vrinda Pub, New Delhi,2002
15. JORG BIBOW, 'The Loanable Funds Fallacy: exercises in the analysis of disequilibrium' University of Hamburg, ISTO, VMP20146,Hamburg,Germany
16. KUDSIA,JITENDRA,'Economics Theory', Oxford and IBH Pub, New Delhi, 1970
17. LANDSBUNG STEVEN E. ,'Price Theory and Applications', The Dryden Press, 1989
18. LESTER O BUMAS ' Intermediate Microeconomics' Prentice Hall , India , 1993
19. MCCONNELL,CAMPBELL R.,' Economics , Principles, Problems and Policies', McGraw Hill, 4th ed.,1960
20. MADDALA,GS AND MILLER,ELLEN, 'Microeconomics Theory and Applications', Tata Mcgraw Hill, 2004

21. MCCLOSKAY DONALD N., 'The Applied Theory of Price', Macmillan Pub, 1982
22. MEYERS,AL,'Elements of Modern Economics',Prentice Hall Pub, 3rd ed., 1949
23. OAK A.D. Paper on 'Indifference Curves ' at 3rd Statistics International Conference ,Nov16-18,2006 ,Bucharest,Romania, (Paper abstract accepted)
24. OAK,AD, 'Indifference Curves in Theory of Price ,a concept of Convenience', Paper presented in the All China Economics International Conference 18th Dec,2006, Hong Kong,
25. OAK A.D., Paper' Loanable Fund Theory Under Interest in Price Theory' at 34th Meeting of Academy of Economics and Finance,Jacksonville,Florida, Feb,16-17,2007 (Paper accepted for inclusion in the transactions)
26. OAK,A.D.,Paper 'Misuse of Mathematics in Theory of Price ', Hawaii University ,Honolulu, International Conference in Social Sciences,30-31 May,1-2 Jun ,2007
27. OAK,AD, Dharmadhikari and Zarapkar, 'Mathematical Techiniques for Management Studies', Himalaya Publications,Bombay, 1980
28. PASHIGIAN,BP, 'Price Theory and Applications', McGraw Hill, 1995
29. PRICE CATHERINE M , 'Welfare Economics in Theory and Practice',Macmillam,1977
30. RAY N.C. ,' An Introduction to Microeconomics' , Macmillan,New Delhi, 1975
31. REEKIE AND CROOK, 'Managerial Economics' ,Heritage Publications, 1983
32. ROBERTSON , DENNIS,' Lectures on Economics Principles', vol I, Staples Press, London, 1959
33. ROBINSON JOAN AND EATWELL,JOHN,' An Introduction to Modern Economics', McGraw Hill, England, 1973
34. SALVATORE DOMINICK, ' Managerial Economics ' McGraw Hill, 1989
35. SALVATORE,DOMINICK,'Managerial Economics in Global Economics', McGraw Hall Pub, 2nd ed, 1993
36. SAMUELSON P.A. ,' Foundation of Economic Analysis', Harvard Uni Press, NY,1965
37. SETH M.L., ' Principles of Economics', Agarwal Publications, Agra,India
38. SPEIGHT H. ,' Economics, the Science of Prices and Incomes ' Methuen, London,1964
39. WATSON DONALD S, AND GETZ, 'Price Theory and its Uses', Scientific Books,Kolkotta,1967
40. WEBSTER'S STUDENT DICTIONARY,TRIDENT PRESS INTERNATIONAL, Ed 1996

EL FRACASO EMPRESARIAL A TRAVÉS DE LA INFORMACIÓN ECONÓMICO--FINANCIERA: RATIOS EXPLICATIVOS

Mª JESÚS MURES QUINTANA
E-mail: mj.mures@unileon.es
ANA GARCÍA GALLEGO
E-mail: ana.ggallego@unileon.es
M. EVA VALLEJO PASCUAL
E-mail: eva.vallejo@unileon.es
Departamento de Economía y Estadística
Facultad de Ciencias Económicas y Empresariales
Universidad de León

Resumen

Since Beaver's and Altman's work, business failure prediction has become an important topic in corporate finance literature and numerous studies have been developed, using a variety of statistical methodologies for predictive purposes. Due to the absence of a theory of financial distress, most models have included as independent variables to predict failure a lot of financial ratios widely used in previous failure prediction models.
The purpose of this paper is to identify financial characteristics that best explain business failure. Taking a sample of failed and non-failed firms with head office in Castilla y León (Spain), we apply a principal components analysis (PCA) in order to reduce the large list of available financial ratios to a smaller number of factors with high explicative power. To confirm that the selected ratios contain the same information as the removed ones, we later apply a canonical correlation analysis (CCA).

Palabras Clave: Fracaso empresarial, ratios, análisis de componentes principales, análisis de correlaciones canónicas.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

1. INTRODUCCIÓN

Desde los trabajos de Beaver (1966) y Altman (1968), la predicción del fracaso empresarial ha sido uno de los principales objetivos que se han planteado en la investigación realizada dentro del campo financiero. A partir de la publicación de estos dos trabajos, considerados como pioneros dentro de esta línea de investigación, se han desarrollado numerosos estudios, que han utilizado diversas técnicas estadísticas con el fin de predecir el fracaso de una muestra de empresas. La mayoría de estos trabajos presenta una característica común: la utilización, como variables independientes, de ratios económico-financieros extraídos de las cuentas anuales de las empresas. Sin embargo, debido a la ausencia de una teoría sobre el fracaso empresarial, la elección de los ratios a incluir en los diferentes modelos elaborados se ha basado en su aparición y nivel de significación en estudios previos, lo que ha dado lugar a un amplio listado de ratios potencialmente explicativos del fenómeno del fracaso.

Partiendo de esta situación, este trabajo tiene como objetivo fundamental identificar aquellas características financieras, definidas en forma de ratios, que permiten explicar mejor el fracaso empresarial. Sobre una muestra de empresas con domicilio en Castilla y León (España), aplicamos un análisis de componentes principales, con el fin de reducir el inmenso conjunto de ratios financieros que se han venido utilizando en los modelos previos sobre predicción del fracaso empresarial, tomando los factores con mayor poder explicativo. Este estudio se complementa con un análisis de correlaciones canónicas, con el que se pretende confirmar que la reducción del número de ratios explicativos no significa una pérdida de información respecto

al fenómeno en estudio. Se trata de comprobar que la información contenida en los ratios rechazados aparece recogida en el conjunto de ratios finalmente seleccionados, lo que justificaría la reducción del número de ratios previamente considerados.

Para alcanzar los objetivos planteados, el trabajo se estructura en los siguientes apartados. En primer lugar, nos referiremos a los diferentes ratios de carácter económico–financiero que hemos seleccionado como explicativos del fracaso empresarial, en función de su popularidad y significación en la literatura más relevante sobre el tema. A continuación, detallamos los resultados obtenidos en el estudio empírico realizado, lo que comprende la descripción del proceso de selección de la muestra objeto de estudio y la aplicación de las técnicas estadísticas mencionadas, exponiendo a su vez, de forma breve, sus características básicas. Por último, hacemos referencia a las principales conclusiones extraídas del estudio.

2. LOS RATIOS ECONÓMICO-FINANCIEROS EN LOS ESTUDIOS SOBRE FRACASO EMPRESARIAL

Los modelos de predicción del fracaso empresarial se han elaborado con el objetivo de anticipar las dificultades financieras que puedan atravesar las empresas y que conduzcan a una situación de crisis. Para ello, han utilizado como fuente de información los estados contables de las empresas, que contienen toda la información económica y financiera que las caracteriza. Debido a la amplitud de datos que se recogen en las cuentas anuales, una de las herramientas utilizadas en el análisis financiero son los ratios, entendidos como una relación por cociente entre dos magnitudes basadas en información extraída de los estados contables (fundamentalmente Balance y Cuenta de Pérdidas y Ganancias). Su principal ventaja es que permiten reducir el gran número de conceptos a analizar a un conjunto, relativamente pequeño y de fácil comprensión, de indicadores con significado económico o financiero. Por otro lado, también ofrecen la posibilidad de efectuar comparaciones, permitiendo extraer conclusiones que no se obtendrían tomando las magnitudes en términos absolutos.

Tomando en consideración las ventajas que trae consigo la utilización de ratios, los trabajos sobre predicción del fracaso empresarial los han incluido como variables explicativas en los diferentes modelos elaborados. El problema que se ha encontrado es la falta de una base teórica sobre el fenómeno del fracaso que sirva de guía para la selección de los ratios a incluir. Por ello, estos modelos han basado la elección de los ratios en su popularidad en la literatura financiera y en su nivel de significación en los estudios previos sobre fracaso empresarial. Este criterio ha dado como resultado un conjunto numeroso de ratios económico–financieros utilizados en los diferentes trabajos, sin que, hasta el momento, se haya desarrollado una teoría que explicara el fracaso empresarial en razón de los resultados que se han ido obteniendo. Por ello, nos planteamos determinar los ratios con mayor poder explicativo del fenómeno objeto de estudio, reduciendo a un menor número el extenso conjunto de ratios utilizados.

Debido a la amplitud del listado de ratios considerados en la investigación realizada hasta el momento, hemos empleado como primer criterio para seleccionar los ratios de nuestro estudio el mismo que han seguido los trabajos anteriores, es decir, su grado de utilización y el nivel de significación obtenido en los análisis realizados.

En primer lugar, consideramos los ratios incluidos en los trabajos de Beaver (1966) y Altman (1968), dada su relevancia, no sólo por constituir el punto de partida de esta línea de investigación, sino también porque sus ratios más significativos han sido utilizados en muchos de los estudios posteriores. No obstante, por problemas de disponibilidad de información, nuestro estudio se limitó a cinco de los seis ratios más importantes de Beaver y a tres de Altman. Junto a éstos, seleccionamos otros ratios utilizados y que resultaron significativos en varios de los principales trabajos publicados[1].

[1] Un resumen de los diferentes modelos elaborados en Estados Unidos y otros países, tanto europeos como del resto del mundo, puede encontrarse en las revisiones bibliográficas que, respecto a los modelos sobre fracaso empresarial, han realizado Altman (1984), Jones (1987), Keasey y Watson (1991), Dimitras *et al*. (1996) o Cybinski (2001), entre otros. Por lo que se refiere a España, autores como Lizarraga (1996) o Laffarga y Mora (2002) recogen en sus trabajos una evolución de la investigación realizada en este campo.

Teniendo en cuenta la disponibilidad de datos y las relaciones entre los ratios, de modo que no contuvieran información redundante, resultó una selección final de 27 ratios económico–financieros, que agrupamos en los tradicionales seis grupos de liquidez, rentabilidad, endeudamiento y solvencia, rotación y actividad, generación de recursos y estructura. No obstante, no se trata de una clasificación estricta, ya que algunos ratios miden aspectos diversos de la actividad de la empresa que pueden incluirse en más de un grupo.

A continuación, definimos de forma breve los diferentes grupos considerados, recogiendo en la tabla 1 los ratios finalmente seleccionados dentro de cada uno, con las diferentes partidas contables que intervienen en su cálculo.

Tabla 1. Ratios financieros

Grupo	Etiqueta	Definición
Liquidez	RCI	Ratio de circulante o liquidez general: Activo circulante / Pasivo circulante
	PAC	Prueba ácida: (Activo circulante – Existencias) / Pasivo circulante
	LIQ	Liquidez inmediata: Disponible (Tesorería) / Pasivo circulante
	CCA	Capital circulante: Capital circulante / Activo total
	CCFO	Capital circulante: Capital circulante / Fondos propios
Rentabilidad	ROA	Rentabilidad económica: Resultado del ejercicio / Activo total
	ROE	Rentabilidad financiera: Resultado del ejercicio / Fondos propios
	REAC	Rentabilidad sobre fondos de accionistas: Resultado antes de impuestos / Fondos propios
	ROA-II	Rentabilidad económica: Resultado antes de impuestos / Activo total
Endeudamiento y solvencia	REP	Nivel de endeudamiento: Pasivo exigible / Activo total
	RECP	Endeudamiento a corto plazo: Pasivo circulante / Activo total
	RELP	Endeudamiento a largo plazo: Pasivo fijo / Activo total
	NPA	Autonomía financiera (solvencia): Fondos propios / Activo total
	FPPC	Fondos propios / Pasivo circulante
	EQUI	Cobertura de inmovilizado o equilibrio: (Fondos propios + Pasivo fijo) / Activo fijo
	CCF	Cobertura de cargas financieras: Resultado de explotación / Gastos financieros
	GFV	Cobertura de cargas financieras: Gastos financieros / Importe neto cifra de ventas
Rotación y actividad	RAC	Rotación de activo: Importe neto de la cifra de ventas (INCV) / Activo total
	Var(INCV)	Crecimiento de la cifra de ventas: $INCV_t$ / $INCV_{t-1}$
	CCV	Capital circulante / Importe neto de la cifra de ventas
	PPAG	Rotación de activo circulante: Activo circulante / Ingresos de explotación
Recursos generados	CFAT	Recursos generados sobre estructura económica: *Cash-flow* / Activo total
	CFDT	Capacidad de devolución de la deuda: *Cash-flow* / Pasivo exigible
	CFPC	Capacidad de devolución de la deuda a corto plazo: *Cash-flow* / Pasivo circulante
Estructura	AC	Activo circulante / Activo total
	AF	Activo fijo / Activo total
	TES	Tesorería / Activo total

Fuente: Elaboración propia, a partir de los principales trabajos sobre fracaso empresarial

La definición correspondiente a los grupos en que están clasificados los ratios anteriores es la siguiente:

- *Ratios de liquidez*: la liquidez de una empresa es la capacidad que tiene para pagar sus obligaciones, por lo que dependerá del grado de realización de los elementos del activo y del nivel de exigibilidad del pasivo. Por ello, los ratios de liquidez hacen referencia al grado en que los activos liquidables a corto plazo de la empresa serán suficientes para afrontar los pagos provenientes de las deudas contraídas a corto plazo.
- *Ratios de rentabilidad*: comparan el resultado obtenido, en sus distintos niveles, con las magnitudes que, según diferentes enfoques, lo han generado. Esto es, la estructura económica de la empresa y la inversión realizada por los propietarios de la misma.
- *Ratios de endeudamiento*: estos ratios, también denominados ratios de equilibrio o de solvencia a largo plazo, miden la relación existente entre los componentes del pasivo exigible (a corto y largo plazo) y los recursos propios, así como la cantidad y calidad de la deuda y el coste de la misma.

- *Ratios de rotación y actividad*: son una medida del dinamismo de la actividad empresarial en relación con la estructura de la empresa. Comparan las ventas de la empresa, en cuanto indicadores de su actividad, con elementos integrantes del activo o del pasivo, de modo que reflejan la rotación existente en los mismos.
- *Ratios de generación de recursos*: en este grupo se incluyen ratios que relacionan diversas partidas contables con el *cash–flow*, obtenido como beneficio del ejercicio más las amortizaciones y provisiones dotadas. Este concepto indica la capacidad de autofinanciación de la empresa, por cuanto que se trata de los recursos generados por la empresa en un ejercicio y que no van a suponer pagos en un periodo determinado.
- *Ratios de estructura*: miden la proporcionalidad entre las masas patrimoniales del activo y las del pasivo, así como la que existe en las partidas que componen cada una de ellas.

3. ANÁLISIS EMPÍRICO

Como hemos indicado en la introducción de este trabajo, nuestro objetivo principal es reducir el número de ratios utilizados en la investigación sobre fracaso empresarial a un conjunto más pequeño, pero con alto poder explicativo. Dada la naturaleza cuantitativa de los ratios, aplicamos el análisis de componentes principales (ACP), que es una de las técnicas estadísticas más empleadas para la reducción de este tipo de información. Asimismo, se ha realizado un análisis de correlaciones canónicas (ACC), que permite valorar la relación existente entre dos conjuntos de variables. En este caso, tratamos de comprobar si la información proporcionada por el grupo de ratios seleccionados coincide con la contenida en el conjunto de los ratios rechazados, justificando así el hecho de trabajar sólo con los primeros. Por este motivo, en primer lugar vamos a hacer una breve descripción de estos dos métodos estadísticos, indicando sus características más relevantes.

Para la aplicación de las dos técnicas, es precisa la elección previa de una muestra de empresas, por lo que también describimos el proceso de selección de la misma, haciendo referencia al tamaño muestral necesario y al método de muestreo utilizado.

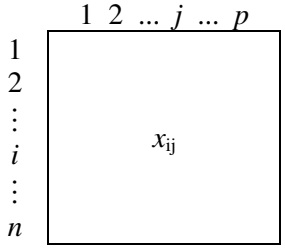
Por último, presentamos los resultados obtenidos en el estudio empírico, en cuanto a los ratios seleccionados a partir del análisis de componentes principales, así como las conclusiones extraídas del análisis de correlaciones canónicas.

3.1. Metodología

3.1.1 Análisis de componentes principales (ACP)

Partimos de una tabla de datos formada por p variables o características cuantitativas centradas y reducidas X_j (las columnas de la tabla), observadas sobre un conjunto de n individuos (las filas de la tabla). Esta tabla X se representa de la forma siguiente:

Figura 1.
Tabla de datos *X*



siendo el elemento característico de la misma x_{ij} (i= 1, 2 ... n ; j = 1, 2 ... p).

El operador $WD = X M X' D$ define la nube de individuos de la tabla, siendo M la matriz de los pesos de las variables (matriz identidad) y D la matriz de pesos de los individuos.

La descomposición de dicho operador WD en sus valores propios y vectores propios permite obtener la única reconstitución de las variables originales mediante unas nuevas variables (factores o componentes principales F_k) ortogonales y linealmente independientes, que son combinación lineal de las primeras. Así, los vectores propios de WD son las componentes principales F_k, siendo su varianza el vector propio λ_k correspondiente:

- Considerando que p es el rango de WD, la cantidad de información o inercia explicada por F_k es su valor propio y, por tanto, el porcentaje de inercia es:

$$\frac{\lambda_k}{\sum_{k=1}^{p} \lambda_k}$$

- La interpretación de las componentes principales es posible a través de sus correlaciones con las variables originales, que son precisamente las coordenadas g_{kj} de las variables X_j en el eje k de la nube de variables (t_k). Por tanto, el coeficiente de correlación de una variable con un F_k vendrá dado por:

$$\rho(F_k, Z_j) = Z_j' D\, t_k = g_{kj}$$

- En cualquier ACP se puede reconstituir las variables que han intervenido en la tabla a través de las h primeras componentes. Así, considerando la ponderación y la fórmula de reconstitución de los datos en función del vector u_k que define el eje k de la nube de individuos, se tiene que:

$$X_j = \sum_{k=1}^{h} F_k m_k^{-1} u_k$$

- El subespacio de dimensión h definido, es el que mejor se ajusta a la nube de individuos en el sentido de los mínimos cuadrados ordinarios. Por tanto, también es el que menos deforma la nube de puntos original, y el que mejor conserva la distancia entre pares de puntos y la información inicial. Teniendo en cuenta la varianza de las componentes retenidas, el porcentaje de inercia explicado por todas ellas es:

$$\frac{\sum_{k=1}^{h} \lambda_k}{\sum_{k=1}^{p} \lambda_k}$$

Por lo que respecta a la interpretación del ACP, es necesario hacer referencia a los siguientes elementos:

- *Número de factores a retener*: Cuanto más grande sea el valor propio de una componente, más información contendrá. Por lo general, los h primeros valores propios presentan un decrecimiento bastante regular, lo que indica que esos factores corresponden a características de la nube que deben ser interpretadas y sugiere que los factores siguientes sólo representan el "ruido" que acompaña a toda observación estadística. Por tanto, el número de factores que se debe retener dependerá del porcentaje de inercia extraída por el factor, puesto que es la medida de la importancia relativa de los factores que se extraen. Parece adecuado retener aquellos factores que en conjunto representan entre el 70 y el 80% de la inercia total, y que tengan además un valor propio superior a uno, siempre y cuando corresponda a componentes que tienen una interpretación clara respecto a las variables originales.
- *Coordenadas de las variables*: Dado que se obtienen mediante la correlación de las variables activas con las componentes principales, la interpretación de éstas permite dotar de significado a las h componentes retenidas.

- *Planos factoriales*: Los factores elegidos para dibujar los planos deben tener una importancia comparable, en términos de inercia, para que dichos planos sean estables. La representación de las ***variables*** en el plano permite una visión más sugestiva para interpretar los resultados numéricos. Las variables cuya coordenada esté próxima a 1, es decir, cercanas al denominado círculo de correlaciones, estarán bien representadas y los vectores que unen el origen a los puntos–variable permiten visualizar los ángulos entre variables y medir la relación entre ellas. En cuanto a la representación de los ***individuos***, interesa el movimiento de los individuos en la dirección de las variables: cada individuo estará al lado de las variables para las que tiene valores altos y estará al lado opuesto de las variables con valores bajos. Los individuos cuya representación está cerca del origen (que es el centro de gravedad) tienen una difícil interpretación, ya que no presentan valores muy extremos para el conjunto de variables, sino más bien en torno a los valores medios de todas ellas.

3.1.2 Análisis de correlaciones canónicas (ACC)

Se parte de una tabla de datos formada por n filas, correspondientes a los individuos analizados, y por $p + q$ columnas, que se particiona en dos subtablas X e Y, con p y q columnas, respectivamente, que representan dos grupos de variables o características cuantitativas. Esta tabla de datos T se representa de la forma siguiente:

Figura 2. Tabla de datos *T=X+Y*

	1 2 ... *j* ... *p*	1 2 ... *k* ... *q*
1		
2		
⋮		
i	x_{ij}	y_{ik}
⋮		
n		

Se supone, sin pérdida de generalidad, que las variables están centradas y reducidas, de modo que la suma de los elementos de cada columna de T es igual a cero.

El objetivo del análisis de correlaciones canónicas es estudiar las relaciones entre los dos conjuntos de variables, a partir de combinaciones lineales de las variables originales, denominadas variables canónicas, que presentan la máxima correlación entre sí y que, a su vez, están incorreladas con el resto de combinaciones lineales que pueden derivarse.

- La correlación entre las combinaciones lineales de cada grupo de variables originales, X e Y, que recibe la denominación de correlación canónica, se obtiene a partir del coeficiente de correlación lineal, cuyo cuadrado se corresponde con los valores propios o raíces características de las matrices cuadradas:

$$(X'X)^{-1}X'Y(Y'Y)^{-1}Y'X \text{ ó}$$

$$(Y'Y)^{-1}Y'X(X'X)^{-1}X'Y$$

  de orden p y q, respectivamente.
- El primer valor propio corresponde a la máxima correlación entre los pares de combinaciones lineales o variables canónicas, obteniéndose el resto de autovalores de las matrices en orden decreciente de correlación, hasta un máximo de raíces canónicas no nulas que coincide con el número de variables incluidas en el grupo más pequeño (mín $[p,q]$).
- Los coeficientes que definen las variables canónicas se obtienen como los vectores propios asociados a los mayores valores propios de las matrices anteriores, esto es, al primer autovalor obtenido. De este modo, los coeficientes relativos a las combinaciones lineales de las variables X se obtienen a partir de la primera de las matrices, mientras que

la segunda matriz permite obtener los vectores de coeficientes correspondientes a las diferentes combinaciones lineales de las variables *Y*.

Una vez obtenidas las variables canónicas, es preciso interpretar la información contenida en las mismas, para lo cual se consideran los siguientes elementos:

- *Coeficientes estandarizados*: Los coeficientes dados a las variables originales que forman las combinaciones lineales representan la contribución relativa de cada una a las variables canónicas.
- *Coeficientes de estructura o cargas canónicas*: Son las correlaciones entre las variables originales de cada conjunto y su respectiva variable canónica, por lo que reflejan el grado en que una variable es representada por una variable canónica. El cuadrado de estos coeficientes para cada variable original es la varianza de la variable que es explicada por una variable canónica. La media para cada conjunto de variables proporciona la proporción total de varianza del conjunto que es extraída por una determinada variable canónica.
- *Cargas canónicas cruzadas*: Miden la correlación entre las variables originales de un conjunto y las variables canónicas del otro conjunto. Por ello, se obtienen como el producto entre las cargas anteriores y el coeficiente de correlación canónica respectivo.
- *Índice de redundancia*: Este índice fue propuesto por Stewart y Love (1968), como modo de obtener una medida de la correlación entre los grupos de variables originales, ya que, como hemos señalado, las correlaciones canónicas miden la correlación entre combinaciones lineales de cada conjunto de variables. Se multiplica el cuadrado de la correlación canónica, que representa la proporción de varianza de una variable canónica en cada par obtenido que es explicada por la otra variable canónica del par, por la varianza extraída por cada variable canónica. La suma de estos productos para todas las variables canónicas es el índice de redundancia, que mide la proporción de varianza de un conjunto de variables que es explicada por el otro.

3.2. Muestra de empresas

A la hora de seleccionar la muestra de empresas hemos de tener en cuenta que, si se trata de diferenciar entre empresas fracasadas y no fracasadas, es preciso que en la muestra se incluyan empresas de ambos grupos. Sin embargo, en este punto, se plantea la cuestión de delimitar lo que se entiende por fracaso de una empresa, ya que este concepto puede englobar una variedad de situaciones con consecuencias negativas para el desenvolvimiento de la empresa.

De las revisiones bibliográficas realizadas por los autores citados, se deduce que la mayoría de trabajos han utilizado una definición legal del fracaso, en el sentido de considerar como tal la declaración de un procedimiento concursal, encontrándose distintas denominaciones según la legislación de cada país[2]. La principal razón para la consideración de este concepto es su objetividad, al tratarse de un proceso legal evidente y, por tanto, con una fecha determinada de ocurrencia. Por este mismo motivo, en nuestro trabajo vamos a definir como fracaso de una empresa la declaración de una situación concursal. Dentro de esta definición vamos a incluir, además del concurso de acreedores, actualmente vigente, las dos figuras que distinguía la legislación anterior: la suspensión de pagos y la quiebra. La razón para considerar las tres figuras concursales es que durante el periodo transitorio tras la entrada en vigor de la nueva legislación, han coexistido en el tiempo.

Para la selección de la muestra, tomamos como ámbito geográfico la comunidad de Castilla y León, identificando aquellas empresas que figuraban como fracasadas[3] en la base de datos SABI (Sistema de Análisis de Balances Ibéricos), utilizada como fuente de información en el

[2] De la legislación estadounidense surge el término *bankrupty*, ampliamente utilizado en la literatura sobre fracaso empresarial. Otros términos utilizados son *failure*, *receivership* o *financial distress*. En el caso de España, el derecho concursal vigente habla de concurso.

[3] En suspensión de pagos, quiebra o concurso, según la definición de fracaso considerada.

estudio. Debido a sus peculiaridades, excluimos los sectores financiero y asegurador, analizando el resto de sectores económicos con actividad en dicha comunidad.

El examen de la base de datos dio como resultado un total de 59 empresas fracasadas, que cumplían la condición de tener disponibles los estados financieros correspondientes, al menos, al primer año anterior al momento del fracaso. Estas empresas operaban en 24 sectores económicos, según la codificación de la Clasificación Nacional de Actividades Económicas (CNAE-93) a nivel de dos dígitos. Dado el pequeño número, se seleccionaron todas para formar la muestra de empresas fracasadas.

Del mismo modo, identificamos las empresas no fracasadas o sanas, estableciendo como condición que estuvieran disponibles sus últimos estados financieros. Al tratarse de empresas en funcionamiento, éstos se corresponden con las últimas cuentas anuales disponibles en SABI, que, a la fecha de consulta de la base de datos, estaban referidas al ejercicio cerrado a 31 de diciembre de 2006. Identificadas estas compañías, procedimos a su selección para la muestra. A diferencia de la mayoría de los estudios previos, en los que cada empresa sana es emparejada con otra fracasada, a partir de criterios como el tamaño o el sector económico, no aplicamos el procedimiento de selección por emparejamiento. En este caso, elegimos la muestra de empresas sanas en proporción al tamaño representado por éstas en la población registrada en SABI. Teniendo en cuenta el tamaño poblacional del grupo de empresas no fracasadas, a un nivel de confianza del 95% y para un error máximo de estimación del 5%, resultó un tamaño muestral de 396 empresas sanas, bajo el criterio conservador de considerar las dos proporciones poblacionales iguales. Por otro lado, para mantener la misma estructura empresarial que en la población, teniendo en cuenta, así, las características propias y peculiaridades de los diferentes sectores económicos, las empresas sanas fueron seleccionadas en los sectores en que operaban las empresas fracasadas. Así, repartimos las 396 compañías de la muestra entre los sectores en función de la proporción que las empresas sanas en cada sector representan sobre el total. Determinado el número de empresas sanas a elegir en cada sector para formar la muestra total, procedimos a seleccionar propiamente las empresas, aplicando para ello un muestreo sistemático, es decir, seleccionando las observaciones de una forma sistemática a partir de un primer elemento elegido al azar.

Como resultado del procedimiento de muestreo, se extrajo una muestra total formada por 455 empresas, entre fracasadas y sanas.

3.3. Resultados del Estudio

Una vez seleccionada la muestra y calculados los ratios definidos en el epígrafe anterior para cada una de las empresas a partir de los estados financieros correspondientes al último ejercicio, procedimos a aplicar sobre los mismos un análisis de componentes principales (ACP), utilizando el paquete estadístico SPAD. Por su parte, para la aplicación del análisis de correlaciones canónicas (ACC), utilizamos el programa STATISTICA.

A continuación, exponemos los resultados conseguidos en el estudio empírico.

3.3.1 Reducción del número de variables a través del ACP

Mediante la aplicación sucesiva del ACP se seleccionaron 15 de los 27 ratios como relevantes para caracterizar y resumir la información económico–financiera de las empresas de la muestra. El procedimiento seguido es el que a continuación se detalla:

— De la aplicación del ACP sobre los 27 ratios inicialmente considerados se deduce que con los seis primeros factores extraídos, cuyos valores propios son superiores a uno, se explica un 58,08% de la varianza de las variables originales, como muestra la tabla 2.

Tabla 2. Porcentajes de inercia (ACP con 27 variables)

Factor	Valor propio	Porcentaje	Porcentaje acumulado
1	4,3274	16,03	16,03
2	2,9956	11,09	27,12
3	2,8765	10,65	37,78
4	1,9594	7,26	45,03
5	1,7672	6,55	51,58
6	1,7544	6,50	58,08

Fuente: Elaboración propia

La tabla 3 presenta los coeficientes de correlación entre las variables originales y cada uno de los factores extraídos. De los 27 ratios, únicamente 21 muestran una fuerte correlación con alguno de los factores. En concreto, con el primer factor se correlacionan positivamente los ratios de endeudamiento, frente a los ratios que reflejan los recursos propios, tanto externos (fondos propios) como internos (*cash–flow*) de la empresa, que presentan una relación negativa con el factor. Por su parte, las variables de liquidez (RCI, LIQ, PAC), además del ratio de fondos propios sobre pasivo circulante (FPPC), presentan una alta correlación positiva con el segundo factor. El tercer factor enfrenta el ratio de activo fijo (AF) con el de activo circulante (AC) y con el de capital circulante sobre activo total (CCA). El cuarto factor viene explicado por la rentabilidad económica (ROE) y la rotación del activo (RAC), conceptos ambos que surgen del margen de beneficio sobre ventas. El siguiente factor está correlacionado negativamente con el ratio de capital circulante sobre fondos propios (CCFO) y con la rotación del activo circulante (PPAG). El último factor retenido es de rotación, en cuanto que se correlacionan con él los ratios de gastos financieros y capital circulante sobre ventas (GFV y CCV).

Tabla 3. Correlaciones entre las 27 variables y los factores del ACP

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
RAC	-0,06	-0,09	-0,28	**0,92**	0,21	0,14
ROA	-0,09	-0,10	-0,26	**0,92**	0,21	0,13
ROE	0,03	0,04	-0,03	-0,08	0,41	0,01
GFV	0,02	0,07	0,10	0,14	-0,01	**-0,77**
CCV	-0,03	0,19	0,20	0,17	0,04	**-0,77**
RCI	-0,44	**0,70**	0,02	0,07	-0,08	-0,09
LIQ	-0,54	**0,77**	-0,04	0,02	-0,04	0,04
PAC	-0,45	**0,57**	-0,03	-0,01	-0,01	0,09
TES	-0,27	0,00	0,25	-0,11	0,26	0,34
AC	0,20	0,12	**0,92**	0,13	0,16	0,15
AF	-0,20	-0,12	**-0,92**	-0,13	-0,16	-0,15
NPA	**-0,77**	-0,45	0,23	0,04	-0,02	-0,18
REP	**0,75**	0,41	-0,18	0,00	-0,02	0,15
RELP	0,42	0,26	-0,33	0,02	-0,25	0,10
RECP	**0,70**	0,39	-0,06	-0,06	0,19	0,14
EQUI	0,01	0,00	0,19	0,12	-0,33	0,03
CCF	-0,10	0,00	0,02	-0,03	0,03	0,16
PPAG	0,04	0,01	0,14	0,17	**-0,64**	0,16
CCA	0,34	0,11	**0,74**	0,21	-0,02	-0,07
CCFO	0,06	0,00	0,23	0,22	**-0,78**	0,20
CFAT	**-0,64**	-0,42	0,17	0,01	-0,07	0,03
CFDT	**-0,53**	0,00	0,06	-0,08	0,15	0,25
CFPC	**-0,60**	0,26	0,03	-0,01	0,13	0,22
FPPC	-0,51	**0,76**	-0,05	0,05	-0,08	-0,07
REAC	0,09	0,08	0,13	-0,04	0,30	0,09
ROAII	-0,44	-0,26	0,07	-0,10	-0,13	0,13
$Var(INCV)_{01}$	-0,04	-0,09	-0,16	-0,02	-0,01	0,16

Fuente: Elaboración propia

Como se observa en la tabla 3, seis variables no presentan una correlación elevada con ningún factor. Por ello, pueden eliminarse del análisis los ratios ROE, TES, EQUI, CCF, REAC y $Var(INCV)_{01}$, sin que afecte al significado de los seis factores descritos.

Los resultados anteriores se confirman en la representación de los diferentes planos factoriales que surgen de los seis factores extraídos. En los tres planos (Figuras 3, 4 y 5), se observa que las variables a eliminar no están próximas al círculo de correlaciones, sino que tienden a situarse en el origen de los ejes.

El resto de variables aparecen representadas en los diferentes planos en razón de sus correlaciones, positivas o negativas, con los respectivos factores.

Figura 3. Primer plano factorial del ACP con los 27 ratios financieros

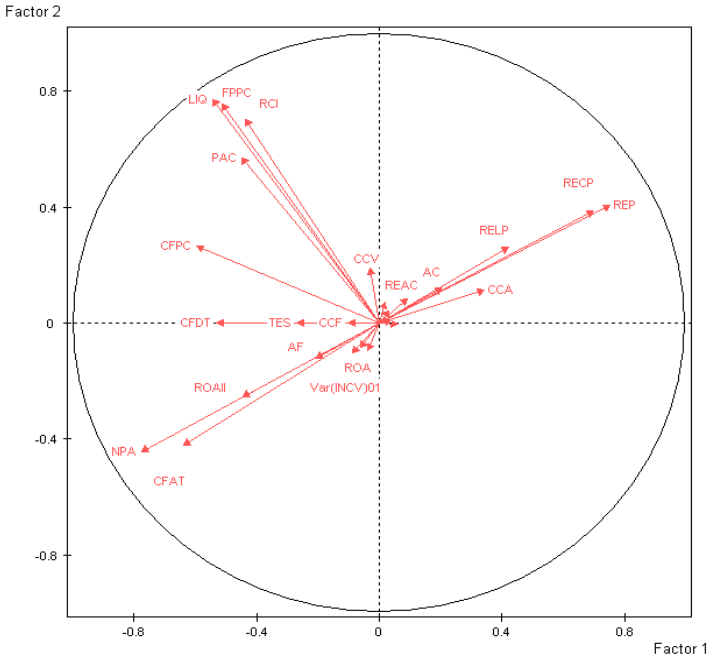


Figura 4. Segundo plano factorial del ACP con los 27 ratios financieros

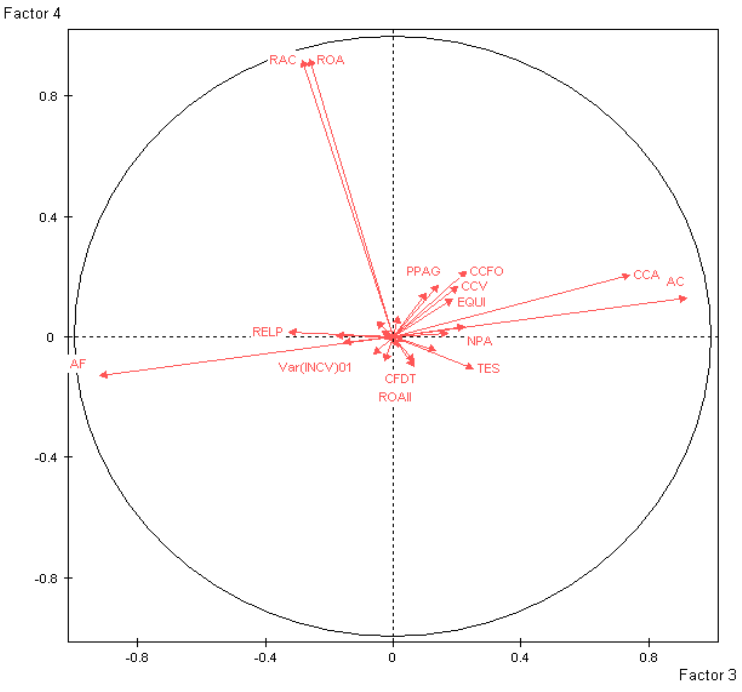
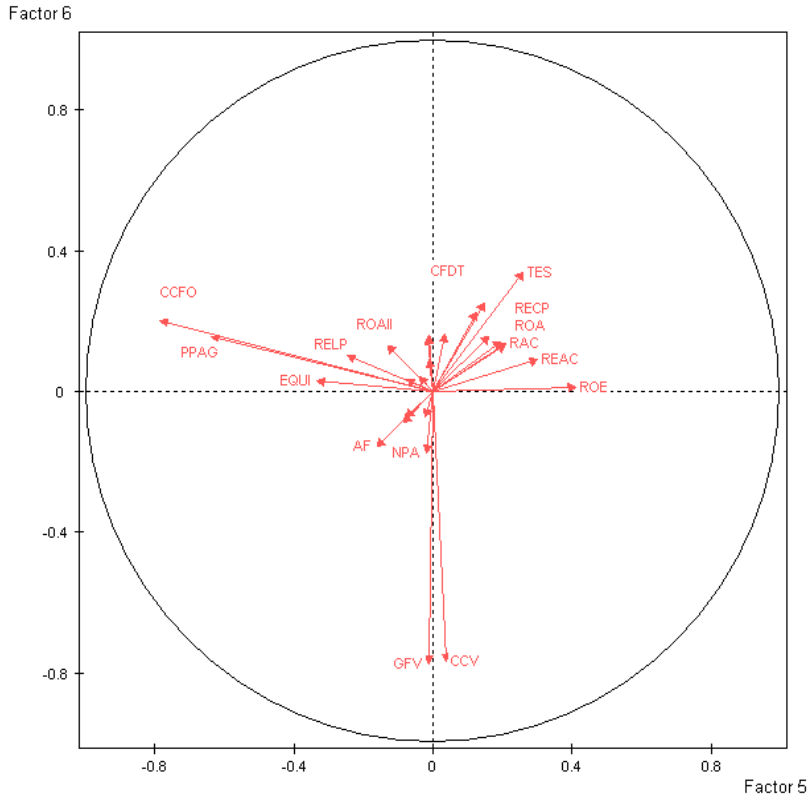


Figura 5. Tercer plano factorial del ACP con los 27 ratios financieros



— A los 21 ratios financieros seleccionados con el primer análisis, les aplicamos de nuevo un ACP, en el que el porcentaje de inercia de los seis primeros factores aumenta hasta el 71,89%. Además de este incremento en el porcentaje de varianza explicada, se observa que los ratios de rentabilidad económica (ROA y ROAII) se correlacionan con factores distintos. Esto significa que no miden el mismo aspecto de la rentabilidad de la empresa, por lo que necesario eliminar uno de los dos. Debido a la mayor aparición del ROA en la literatura sobre fracaso empresarial, eliminamos la segunda variable (ROAII).

— Del análisis exhaustivo de los 20 ratios restantes, decidimos eliminar aquellos otros que no aparecen bien explicados por los factores retenidos, al presentar una baja comunalidad. Esto ocurre con los ratios CCFO y PPAG, que además sólo aparecen en alguno de los trabajos anteriores. Una vez eliminadas estas dos variables, el ACP con las 18 variables restantes permite retener 6 factores que explican un 81,84% de la inercia.

— Los 18 ratios anteriores se examinan detenidamente puesto que puede suceder que determinados ratios, aun estando correlacionados con los factores, contengan información redundante. En este sentido, existe una fuerte correlación entre los ratios de activo fijo y de activo circulante sobre activo total, ya que, en los balances de las empresas de la muestra, no figuran otras partidas dentro del activo. Por ello y dada la mayor importancia de las cuentas de activo circulante en la capacidad de pago de una empresa, eliminamos el ratio AF.

Por otro lado, el endeudamiento de una empresa se divide en dos tipos, a corto y a largo plazo, según que las deudas venzan en el plazo máximo de un año o en un periodo superior. Debido a la relación entre los tres ratios de endeudamiento considerados y ante los mayores problemas que puede generar el pasivo circulante respecto a su devolución, sólo mantenemos en el análisis la variable que recoge la proporción de deudas a corto plazo (RECP). Por tanto, eliminamos los ratios REP y RELP.

— El ACP con los 15 ratios seleccionados permite incrementar el porcentaje de varianza explicada por los seis primeros factores, alcanzando un 85,02%, según se recoge en la tabla 4.

Tabla 4. Porcentajes de inercia (ACP con 15 variables)

Factor	Valor propio	Porcentaje	Porcentaje acumulado
1	3,6438	24,29	24,29
2	2,5104	16,74	41,03
3	2,0237	13,49	54,52
4	1,8142	12,09	66,61
5	1,5211	10,14	76,75
6	1,2396	8,26	85,02

Fuente: Elaboración propia

Los seis factores extraídos se correlacionan con las 15 variables retenidas, como se puede observar en la tabla 5.

Tabla 5. Correlaciones entre las 15 variables y los factores del ACP

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
RAC	0,06	0,28	**-0,85**	0,41	-0,18	0,02
ROA	0,08	0,31	**-0,82**	0,42	-0,20	0,01
GFV	0,01	-0,09	0,17	**0,67**	**0,61**	0,05
CCV	0,11	-0,18	0,20	**0,73**	0,49	0,06
RCI	**0,71**	-0,44	-0,05	0,07	-0,07	-0,23
LIQ	**0,84**	-0,44	-0,09	-0,05	-0,05	-0,10
PAC	**0,66**	-0,29	-0,07	-0,09	-0,05	-0,01
AC	-0,16	-0,26	0,39	0,45	**-0,65**	0,13
NPA	0,46	**0,66**	0,31	0,12	-0,09	-0,31
RECP	-0,47	**-0,69**	-0,26	-0,10	0,02	0,35
CCA	-0,25	-0,30	0,34	0,49	**-0,61**	-0,01
CFAT	0,39	**0,68**	0,30	0,10	-0,14	-0,03
CFDT	0,47	0,29	0,12	-0,05	0,00	**0,73**
CFPC	**0,66**	0,09	0,04	-0,04	-0,06	**0,61**
FPPC	**0,80**	-0,45	-0,09	0,01	-0,03	-0,20

Fuente: Elaboración propia

Estas correlaciones permiten dar una descripción a los factores obtenidos con el ACP. Así, el primer factor viene explicado por los ratios de liquidez (RCI, LIQ, PAC), aspecto de la empresa que se relaciona, a su vez, con su capacidad de devolución de la deuda a corto plazo (CFPC y FPPC). El segundo factor enfrenta los recursos propios de la empresa, tanto internos como externos, con los fondos ajenos a corto plazo, ya que se correlacionan positivamente con él los ratios que recogen la proporción que aquellos representan sobre el activo total (NPA y CFAT), mientras que la proporción de pasivo circulante (RECP) presenta una fuerte correlación negativa con el factor. Estas relaciones se muestran en la figura 6, donde se representa el primer plano factorial, resultante de los dos primeros factores del ACP. En la figura 7 se representa el segundo plano factorial, correspondiente al tercer y cuarto factor. En el gráfico se observa que el tercer factor se correlaciona positivamente con el ratio de rentabilidad económica (ROE) y la rotación del activo (RAC) que, como hemos indicado, son los dos conceptos en que puede desagregarse el margen de beneficio sobre ventas. Las variables GFV y CCV, que representan la rotación de los gastos financieros y el capital circulante, respectivamente, se correlacionan con el cuarto factor, por lo que podría definirse como un factor de rotación. Estos ratios de rotación también parecen correlacionarse con el quinto factor, que los enfrente a los ratios de activo circulante (AC) y de capital circulante (CCA) sobre activo. Por último, el sexto factor describe la capacidad de devolución de la deuda que tiene la empresa a partir de los recursos que genera

(CFPC y CFDT). En la figura 8 se muestra el último plano factorial, que surge de la representación de los dos últimos factores.

Figura 6. Primer plano factorial del ACP con 15 ratios financieros

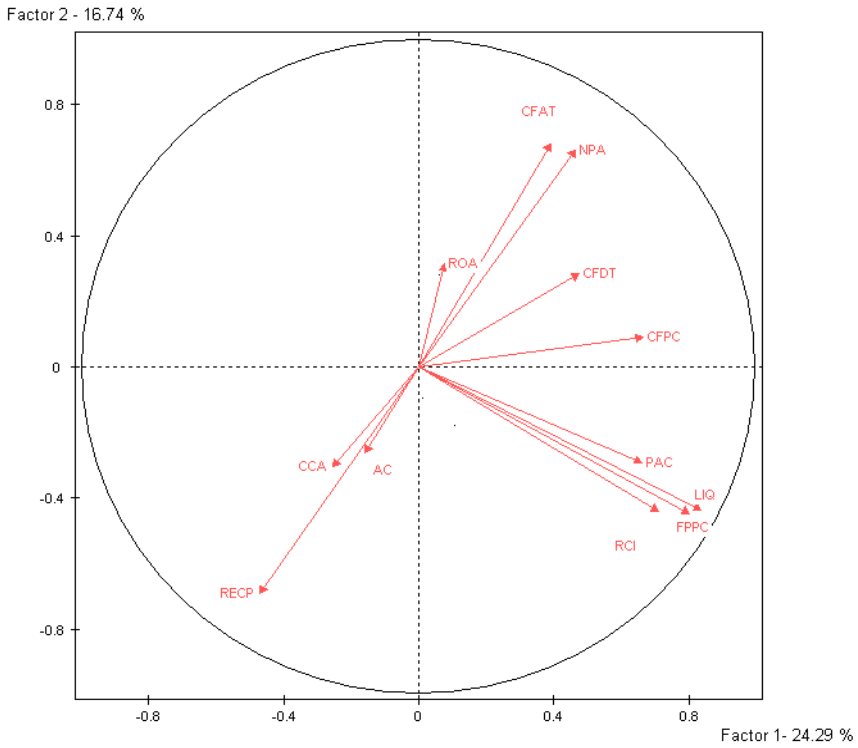


Figura 7. Segundo plano factorial del ACP con 15 ratios financieros

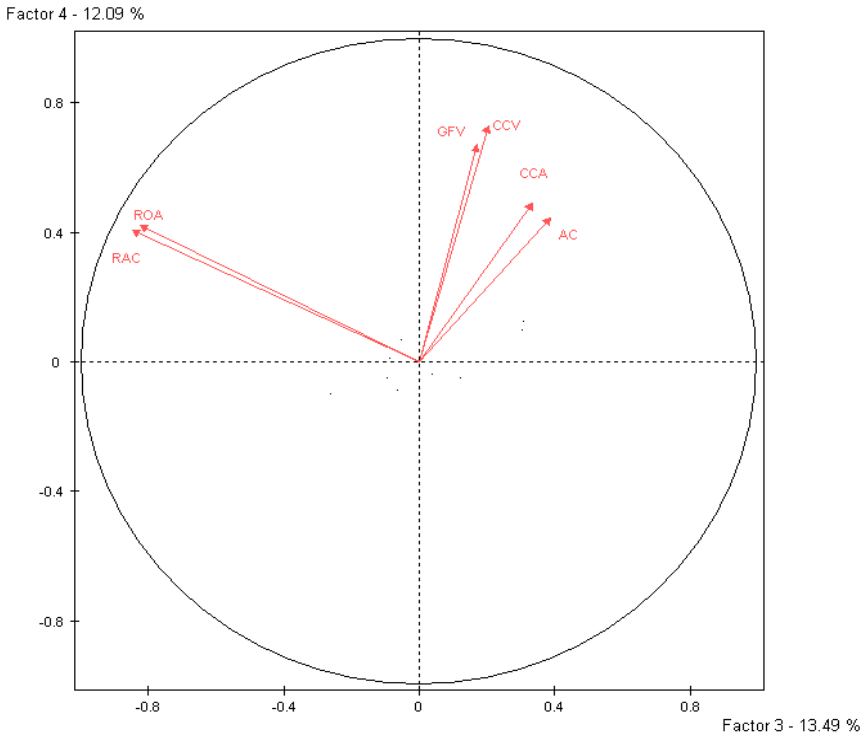
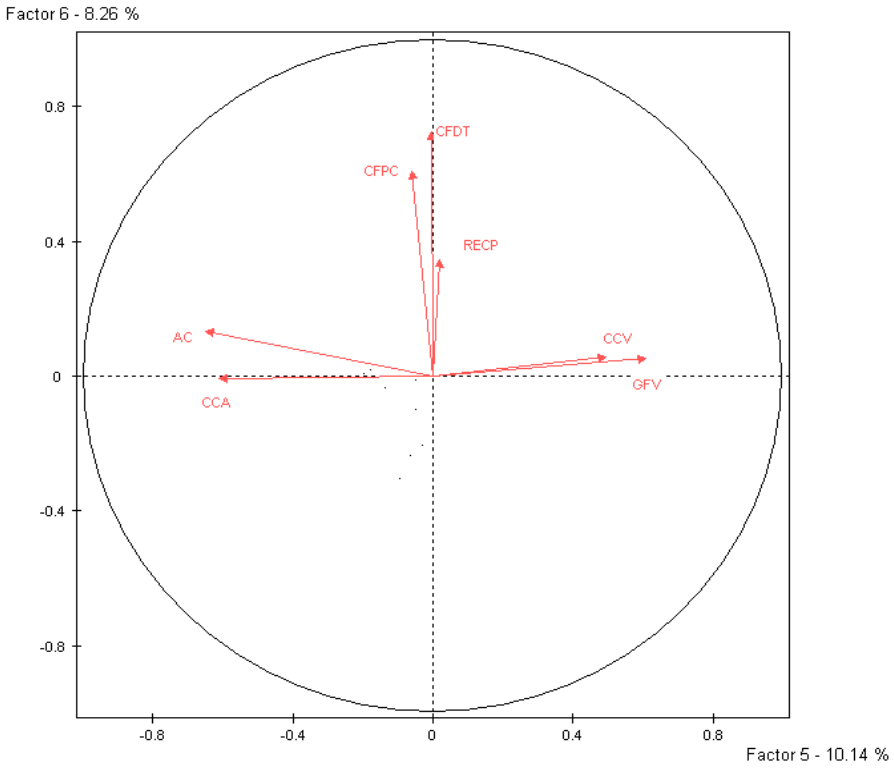


Figura 8. Tercer plano factorial del ACP con 15 ratios financieros



Por tanto, las características financieras de las empresas seleccionadas se resumen a través de los 15 ratios elegidos, explicados por los seis factores retenidos, según muestra la tabla 6.

Tabla 6. Factores explicativos y ratios que los caracterizan

Factor	Significado	Variables que lo definen
1	Liquidez	RCI LIQ PAC CFPC FPPC
2	Estructura	NPA RECP CFAT
3	Rentabilidad económica	RAC ROA
4	Rotación	GFV CCV
5	Circulante	AC CCA
6	*Cash–Flow*	CFDT CFPC

Fuente: Elaboración propia

3.3.2 Confirmación de los ratios elegidos a través de ACC

Finalizado el ACP y decidida la selección final de 15 ratios financieros, aplicamos un ACC, con el objetivo de comprobar si los 12 ratios de los que se ha prescindido proporcionan la información contenida en los ratios seleccionados, y de ahí su eliminación.

Con el ACC llevado a cabo con el programa STATISTICA se obtienen cinco correlaciones canónicas significativas, como se puede comprobar en la tabla 7, donde se muestra el valor del estadístico *chi*–cuadrado para las sucesivas correlaciones obtenidas. Como se observa, tras extraer seis correlaciones canónicas, el resto de coeficientes dejan de ser significativos, por lo que sólo son necesarias las cinco primeras y, por tanto, el mismo número de variables canónicas para explicar la relación entre los dos grupos de variables.

La misma conclusión se extrae al observar el valor de los coeficientes de correlación canónica, iguales a uno o muy próximos hasta la extracción del quinto valor propio, a partir del cual empiezan a descender, aproximándose a cero, como se muestra en la figura 9, en la que se

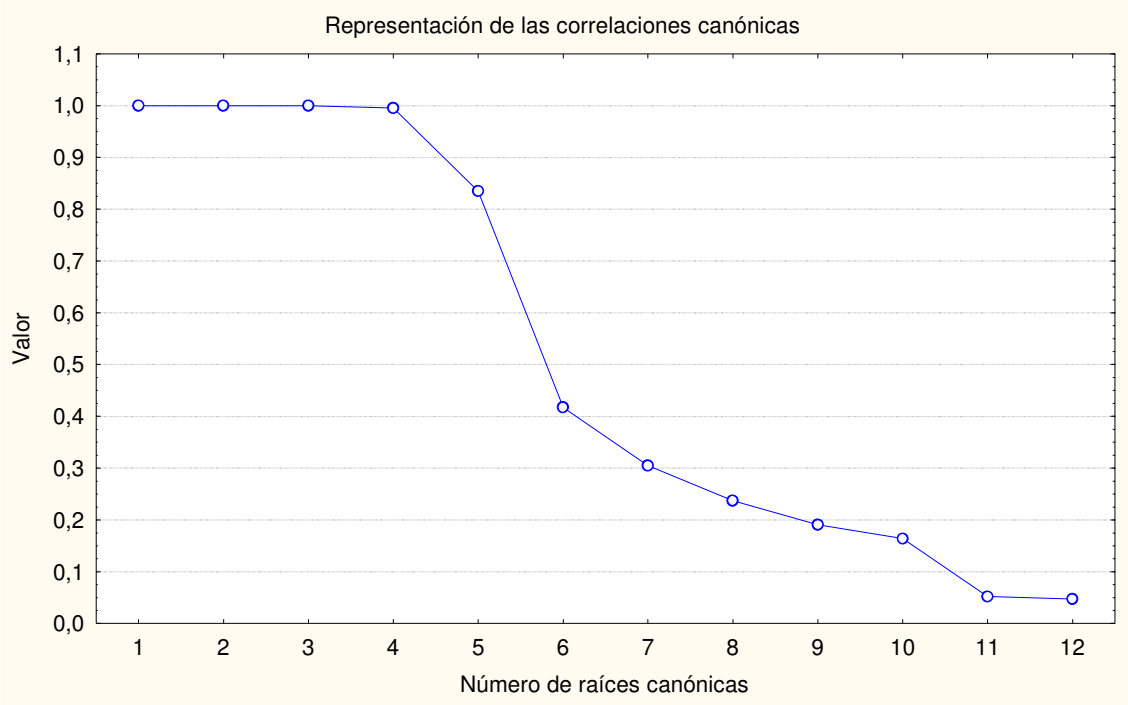
representan las doce correlaciones canónicas obtenidas. Los valores unitarios de los coeficientes indican que existe una relación prácticamente perfecta entre los tres primeros pares de variables canónicas obtenidas como combinaciones lineales de los ratios financieros. La correlación sigue siendo muy alta entre las combinaciones lineales correspondientes al cuarto valor propio o raíz canónica, descendiendo para el siguiente par de variables canónicas, que presentan una correlación de 0,835.

Tabla 7. Estadístico *chi*–cuadrado de significación de los coeficientes de correlación canónica del ACC

Raíz eliminada	Raíz canónica	Correlación canónica	Chi–cuadrado	df	p	Lambda
0	1,000000	1,000000	31190,39	180	0,000000	0,000000
1	1,000000	1,000000		154	0,000000	0,000000
2	1,000000	1,000000	9347,77	130	0,000000	0,000000
3	0,991020	0,995500	1979,25	108	0,000000	0,001794
4	0,697262	0,835022	504,17	88	0,000000	0,199736
5	0,174006	0,417141	130,17	70	0,000018	0,659764
6	0,093093	0,305111	70,33	54	0,067143	0,798752
7	0,056210	0,237085	39,75	40	0,481547	0,880743
8	0,036247	0,190387	21,64	28	0,797600	0,933198
9	0,026896	0,164000	10,08	18	0,929095	0,968296
10	0,002719	0,052143	1,55	10	0,998769	0,995059
11	0,002228	0,047205	0,70	4	0,951545	0,997772

Fuente: Elaboración propia

Figura 9. Valor de las correlaciones canónicas



A la hora de interpretar los resultados obtenidos, comenzamos por los coeficientes o ponderaciones de las variables en cada combinación lineal, lo que nos permite determinar la contribución de cada variable original a las variables canónicas. Estos coeficientes se recogen en la tabla 8, tanto para el grupo de ratios seleccionados como para el de los eliminados.

Tabla 8. Ponderaciones de los ratios en las variables canónicas (VC)

Grupo	Ratios	VC 1	VC 2	VC 3	VC 4	VC 5
Ratios seleccionados	RAC	-0,000000	0,000000	0,000033	0,00513	1,12358
	ROA	0,000000	-0,000000	-0,000033	-0,00143	-1,12986
	GFV	0,000000	-0,000000	-0,000000	0,08738	-0,01668
	CCV	-0,000000	0,000000	0,000001	-1,05736	0,09068
	RCI	0,000000	-0,000000	-0,000001	0,02781	-0,18144
	LIQ	0,000000	-0,000000	-0,000000	0,08008	0,33762
	PAC	-0,000000	0,000000	0,000001	-0,00663	0,06455
	AC	0,838141	0,613620	-0,283275	0,13803	0,97500
	NPA	-0,000101	0,000050	1,694586	-0,16224	-0,25323
	RECP	-0,817597	0,640873	1,417199	-0,16694	-0,06213
	CCA	0,000000	-0,000000	-0,000001	0,01478	-1,31826
	CFAT	-0,000000	0,000000	0,000003	0,00923	0,19468
	CFDT	0,000000	-0,000000	-0,000000	0,03502	-0,13799
	CFPC	-0,000000	0,000000	0,000000	-0,01708	0,04503
	FPPC	-0,000000	0,000000	0,000001	-0,02561	0,05671
Ratios eliminados	$Var(INCV)_{01}$	-0,000000	0,000000	0,000001	0,01606	0,032698
	ROE	0,000000	-0,000000	-0,000001	0,00707	-0,058170
	TES	-0,000000	0,000000	0,000000	0,03619	1,024536
	AF	-0,838141	-0,613620	0,283276	-0,13557	0,296403
	REP	-0,886182	0,694662	-0,158330	-0,03891	0,213939
	RELP	0,554477	-0,434626	-0,961115	0,11958	-0,007651
	EQUI	-0,000000	0,000000	0,000000	0,03677	0,026339
	CCF	0,000000	-0,000000	-0,000001	0,01239	0,044234
	PPAG	-0,000000	0,000000	0,000002	-1,00804	0,073320
	CCFO	-0,000000	0,000000	0,000001	-0,00670	-0,062832
	REAC	-0,000000	0,000000	0,000000	-0,00502	0,025180
	ROAII	0,000000	-0,000000	-0,000000	-0,00345	0,060660

Fuente: Elaboración propia

En el primer grupo, son los ratios que reflejan la proporción de activos circulantes (AC), fondos propios (NPA) y deudas a corto plazo (RECP) los de mayor peso en las primeras variables canónicas, a los que hay que añadir las variables relacionadas con el capital circulante (CCV y CCA) y los ratios RAC y ROA en las siguientes.

Por su parte, en el grupo de ratios eliminados, las variables que presentan mayor ponderación en las correspondientes variables canónicas son AF, que mide el porcentaje de activo fijo que posee la empresa, así como los ratios de endeudamiento total y a largo plazo (REP y RELP). Al comparar estas combinaciones lineales con las respectivas variables canónicas del grupo de ratios seleccionados, se comprende la alta correlación canónica obtenida entre estas primeras variables canónicas, dada la relación existente entre los ratios incluidos en las combinaciones lineales (entre AC y AF y de NPA y RECP con REP y RELP). En las siguientes variables canónicas también destacan los ratios PPAG y TES.

Sin embargo, los coeficientes estimados para los distintos ratios en las variables canónicas pueden estar influidos por la posible existencia de multicolinealidad. Por ello, se suelen utilizar, en la interpretación de las variables canónicas, los denominados coeficientes de estructura o cargas canónicas, que reflejan la correlación entre los ratios y las variables canónicas obtenidas. Estos coeficientes son similares a las correlaciones entre las variables y los factores del ACP. De este modo, igual que en este análisis, con las cargas canónicas puede resultar más fácil dar un nombre o descripción a las respectivas variables canónicas.

Las cargas canónicas para las primeras combinaciones lineales se recogen en la tabla 9. En la misma se puede observar que los coeficientes de estructura confirman la interpretación del análisis realizada a partir de las ponderaciones dadas a los ratios en las variables canónicas.

Así, las dos primeras variables canónicas vienen explicadas por los ratios AC, NPA y RECP en el primer grupo y por AF y REP en el de ratios eliminados, aunque en la segunda combinación lineal también presenta una correlación importante en el primer grupo el ratio CCA. Por su parte, la tercera variable canónica explica la relación entre la proporción de fondos propios (NPA), entre los ratios seleccionados, y la de deudas (REP), en especial las de largo plazo (RELP), en el grupo de eliminados. Con la siguiente variable presentan una alta correlación negativa el ratio CCV, en el primer grupo, y el de rotación del activo circulante sobre ventas (PPAG) en el segundo, dada la relación entre ambos, puesto que el capital circulante se obtiene a partir del activo circulante. La quinta variable canónica se correlaciona negativamente con CCA y positivamente con LIQ en el grupo de ratios seleccionados, mientras que en el segundo grupo presenta una alta correlación positiva con la variable canónica el ratio que mide la proporción que la tesorería representa en el activo (TES).

Tabla 9. Coeficientes de estructura o cargas canónicas del ACC

Grupo	Ratios	VC 1	VC 2	VC 3	VC 4	VC 5
Ratios seleccionados	RAC	-0,063720	-0,118100	0,011853	-0,004992	-0,012027
	ROA	-0,051887	-0,114302	0,022043	-0,005171	-0,012939
	GFV	-0,003906	-0,086340	-0,131630	-0,456958	-0,008951
	CCV	0,117354	0,048098	-0,089683	**-0,980905**	0,036271
	RCI	0,308282	-0,130074	0,057380	-0,018111	0,074775
	LIQ	0,218918	-0,096966	0,097151	0,026768	**0,497519**
	PAC	0,264413	-0,096117	0,109735	0,020332	0,219381
	AC	**0,616928**	**0,787020**	0,000014	-0,000000	-0,000000
	NPA	**0,597170**	**-0,543224**	**0,590166**	0,000001	-0,000000
	RECP	**-0,590740**	**0,806862**	-0,000059	-0,000000	0,000000
	CCA	0,439005	**0,594214**	-0,058858	-0,055520	**-0,650041**
	CFAT	0,414150	-0,325597	0,384145	0,003389	0,037308
	CFDT	0,269161	-0,155094	0,228899	0,031055	0,348347
	CFPC	0,209795	-0,065891	0,177931	0,007982	0,196525
	FPPC	0,106194	-0,236356	0,153059	-0,043422	0,033844
Ratios eliminados	$Var(INCV)_{01}$	-0,121775	-0,027158	0,052654	0,130427	0,066137
	ROE	-0,065311	0,057653	-0,076438	-0,025943	-0,065369
	TES	0,270016	0,129645	0,106719	0,074250	**0,939691**
	AF	**-0,616928**	**-0,787020**	-0,000014	0,000000	-0,000000
	REP	**-0,597171**	**0,543224**	**-0,590165**	-0,000001	0,000000
	RELP	-0,083456	-0,321452	**-0,943241**	-0,000002	0,000000
	EQUI	0,176391	0,032479	-0,036022	0,031522	0,208848
	CCF	0,036663	0,029763	-0,023358	0,026685	0,143633
	PPAG	0,108404	0,052834	-0,086677	**-0,987147**	0,045073
	CCFO	0,078680	0,064914	-0,057176	0,183359	-0,215424
	REAC	0,001408	0,155095	0,001787	-0,015845	-0,020817
	ROAII	0,399161	-0,327370	0,135856	0,007281	0,051392

Fuente: Elaboración propia

Una alternativa a los coeficientes de estructura son las denominadas cargas canónicas cruzadas, que miden la correlación entre las variables de un conjunto y las variables canónicas del otro conjunto. En nuestro estudio, como las primeras correlaciones canónicas son unitarias o muy próximas a uno, las cargas cruzadas son iguales o muy similares a las anteriores y, por ello, no se presentan en el trabajo.

Las cargas canónicas también proporcionan la varianza de una variable que es explicada por las variables canónicas. Como se muestra en la tabla 10, las variables canónicas obtenidas explican el 100% de la varianza del grupo de ratios eliminados, lo que se debe a que el número

de variables canónicas que se extraen coincide con el número de variables en el conjunto más pequeño. No obstante, las doce variables canónicas también explican un porcentaje elevado de la varianza de los ratios seleccionados, ya que, en total, la varianza extraída es de un 86,03%.

Tabla 10. Varianza extraída y redundancia en el ACC

Grupo	Ratios seleccionados		Ratios eliminados	
Variable canónica	Varianza extraída	Redundancia	Varianza extraída	Redundancia
1	0,087159	0,087159	0,120773	0,120773
2	0,098239	0,098239	0,145415	0,145415
3	0,107424	0,107424	0,043838	0,043838
4	0,086107	0,085334	0,078567	0,077861
5	0,083952	0,058537	0,059201	0,041278
6	0,056254	0,009788	0,048643	0,008464
7	0,085117	0,007924	0,026147	0,002434
8	0,096630	0,005432	0,023459	0,001319
9	0,080885	0,002932	0,140106	0,005078
10	0,104456	0,002809	0,052536	0,001413
11	0,068431	0,000186	0,043495	0,000118
12	0,045345	0,000101	0,078081	0,000174
TOTAL	0,860259	0,448166	1,000000	0,465865

Fuente: Elaboración propia

Una medida que suele considerarse más adecuada para interpretar los resultados del análisis es el conocido índice de redundancia, en cuanto que mide el grado de relación entre los dos conjuntos de variables. Así pues, si bien las variables canónicas obtenidas como combinaciones lineales de los ratios seleccionados y eliminados presentan una fuerte correlación entre sí, no significa que los dos grupos de variables estén correlacionados, lo que indicaría, en ese caso, que contienen información similar, por lo que quedaría justificada la eliminación de los doce ratios del segundo grupo, en favor de los seleccionados por el ACP. Puesto que éste es, precisamente, el objetivo que nos planteamos a la hora de aplicar el ACC, nos interesa obtener el índice de redundancia de cada conjunto de ratios respecto del otro grupo, a fin de determinar el porcentaje de varianza de cada grupo de variables que es explicado por el otro conjunto.

En la tabla 10 se recogen los valores obtenidos para el índice de redundancia en cada grupo de ratios. El grupo de ratios eliminados explica un 44,82% de la varianza en el conjunto de variables seleccionadas. Más importante para nuestro estudio es el segundo índice de redundancia, que indica que el conjunto de ratios seleccionados permite explicar un porcentaje de varianza del grupo de ratios eliminados próximo al 47%. De este resultado se deduce que los ratios seleccionados por el ACP explican un porcentaje importante de la variabilidad del conjunto de ratios eliminados, recogiendo, por tanto, información proporcionada por éstos. Por tanto, queda justificada la consideración de un menor número de ratios para definir las características financieras básicas de las empresas, fracasadas y no fracasadas, de la muestra y que pueden utilizarse como punto de partida para explicar el fracaso empresarial.

Un resumen de los ratios de cada grupo que definen las distintas variables canónicas obtenidas se muestra en la tabla 11. En la tabla puede observarse que los ratios que forman las combinaciones lineales del grupo de ratios seleccionados miden aspectos complementarios de los que reflejan los ratios incluidos en las variables canónicas obtenidas para el conjunto de los eliminados.

Tabla 11. Ratios que definen las variables canónicas

Variable canónica	Grupo de ratios seleccionados	Grupo de ratios eliminados
1	AC NPA RECP	AF REP
2	AC NPA RECP CCA	AF REP
3	NPA	REP RELP
4	CCV	PPAG
5	LIQ CCA	TES

Fuente: Elaboración propia

4. CONCLUSIONES

Para finalizar nuestro trabajo, exponemos las principales conclusiones extraídas del mismo.

La primera conclusión a apuntar es que la información económico–financiera expresada en forma de ratios permite caracterizar los principales aspectos de la actividad empresarial. Además, el amplio listado de ratios financieros que pueden calcularse de las cuentas anuales que presentan las empresas puede reducirse a un grupo más pequeño de factores con gran poder explicativo, a partir de la aplicación del ACP.

Los resultados del ACP permiten identificar seis componentes principales que explican un elevado porcentaje de varianza de las variables originales (85,02%). Estos seis factores se relacionan con las siguientes variables: el primer factor con los ratios de liquidez y los que reflejan la capacidad de devolución de la deuda a corto plazo; el segundo con los fondos propios, internos y externos, y los recursos ajenos que posee la empresa; el siguiente factor aparece correlacionado con los dos conceptos que componen el margen de beneficio sobre ventas (rentabilidad económica y rotación del activo); el quinto es un factor de rotación y el último se correlaciona con la capacidad de la empresa para generar recursos con los que hacer frente a sus deudas.

Por otro lado, el ACC también justifica la reducción del número de ratios a utilizar en la explicación del posible fracaso de las empresas, ya que las sucesivas variables canónicas obtenidas como combinaciones lineales de los ratios seleccionados y eliminados, respectivamente, presentan una fuerte correlación entre sí. Asimismo, el índice de redundancia del conjunto de ratios eliminados, dado el grupo de ratios seleccionados, determina que éste explica un 47% de la variabilidad de aquél, indicando que proporciona una parte importante de la información contenida en las variables eliminadas en el anterior análisis.

Por consiguiente, las técnicas estadísticas de reducción de datos, como el ACP, son adecuadas en el estudio de la información de carácter económico y financiero, por cuanto que un menor número de ratios financieros permiten explicar un alto porcentaje de la información de partida, con las ventajas que esto supone respecto a la cantidad de datos a manejar.

BIBLIOGRAFÍA

ALTMAN, E.I. (1968), "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *Journal of Finance*, vol. XXIII, no. 4, pp. 589-609.

— (1984), "The success of business failure models: An international survey", *Journal of Banking & Finance*, vol. 8, pp. 171-198.

BEAVER, W. (1966), "Financial Ratios as Predictors of Failure", Empirical Research in Accounting: Selected Studies, *Journal of Accounting Research*, supplement to no. 4, pp. 71-111.

CYBINSKI, P. (2001), "Description, Explanation, Prediction – the Evolution of Bankruptcy Studies?". *Managerial Finance*, vol. 27, no. 4, pp. 29-44.

DIMITRAS, A.I.; ZANAKIS, S.H.; ZOPOUNIDIS, C. (1996), "A survey of business failures with an emphasis on prediction methods and industrial applications", *European Journal of Operational Research*, vol. 90, no. 3, pp. 487-513.

JONES, F.L. (1987), "Current Tehniques in Bankruptcy Prediction", *Journal of Accounting Literature*, vol. 6, pp. 131-164.

KEASEY, K; WATSON, R. (1991), "Financial Distress Prediction Models: A Review of Their Usefulness", *British Journal of Management*, vol. 2, no. 2, pp. 89-102.

LAFFARGA, J.; MORA, A. (2002), "La predicción del fracaso empresarial. El estado de la cuestión en España", en Doldán Tié, F.; Rodríguez López, M. (coord.), *La Gestión del Riesgo de Crédito. Métodos y modelos de Predicción de la Insolvencia Empresarial*, Madrid: Ed. AECA Monografías, pp. 25-45.

LEBART, L.; MORINEAU, A.; PIRON, M. (1997), *Statistique exploratoire multidimensionnelle*, 2e édition, Paris: Dunod.

LIZARRAGA DALLO, F. (1996), *Modelos multivariantes de previsión del fracaso empresarial: Una aplicación a la realidad de la información contable española*, Universidad Pública de Navarra: Tesis Doctoral.

STEWART, D.; LOVE, W. (1968), "A general canonical correlation index", *Psychological Bulletin*, vol. 70, no. 3, pp. 160-163.

LA INFLUENCIA DEL TERRORISMO SOBRE LA INDUSTRIA TURÍSTICA

M. D. SOTO TORRES
E-mail: lolasoto@eco.uva.es

R. FERNÁNDEZ LECHÓN
E-mail: ramonfer@eco.uva.es

Departamento de Economía Aplicada
Facultad de Ciencias Económicas y Empresariales
Universidad de Valladolid

Resumen

Bajo la hipótesis de que las zonas turísticas tienen una mayor atracción para los terroristas, como la evidencia parece mostrar, este trabajo formula un juego dinámico, en tiempo discreto, entre los promotores turísticos y el gobierno de una zona turística, con objeto de encontrar las estrategias de equilibrio a seguir respecto a las inversiones en infraestructura y medidas de seguridad a implementar en la zona, como medio para disuadir a los terroristas. El comportamiento dinámico del turismo y la amenaza terrorista, cuando los dos jugadores siguen estrategias feedback de Nash, admite distintas posibilidades y depende de los valores que alcancen ciertos parámetros. Seleccionando dos parámetros de bifurcación, en el trabajo se genera una bifurcación de Neimark-Sacker y tanto el turismo como la amenaza terrorista siguen un comportamiento oscilatorio lo que puede interpretarse en términos de causa y efecto entre terroristas y promotores.

Palabras clave: Industria turística. Juegos dinámicos. Estrategia Feedback de Nash. Bifurcación de Neimark-Sacker.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

Considering the hypothesis according to which the tourist zones attract terrorists, as the evidence seems to show, this paper formulates a dynamic game between the promoters of a tourist zone and the government assuming discrete time. The aim is to find a equilibrium solution with regard to investments in infrastructures and security-deterrence measures to dissuade terrorists to carry out activities in the zone. If the players follow feedback Nash equilibrium, the dynamic behaviour of both tourism and terrorist threat depends on certain parametric values. Selecting two bifurcation parameters, the paper checks that both the tourism in the zone and the terrorist threat present oscillations as consequence of a Neimark-Sacker bifurcation.

Key word: Tourist industry, Dynamic games, Feedback Nash equilibrium, Neimark-Sacker bifurcation.
Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUCCION

El turismo ha sido durante los últimos años una de las principales fuentes de ingreso de un gran número de países y aunque la Organización Mundial de Turismo (OMT) pronosticó que la demanda de turismo internacional se reduciría de forma significativa durante el segundo semestre de 2008 y el año 2009 como consecuencia del avance de la crisis, para los ocho primeros meses de 2008 el crecimiento mundial fue del 3,7 % en relación con el mismo periodo de 2007. La importancia del turismo en términos económicos para los países es patente siguiendo los datos ofrecidos por el Barómetro OMT del Turismo Mundial (octubre 2008). Los ingresos por turismo internacional durante 2007 en Europa fueron 434.091 millones de dólares, destacando España con 57.795, Francia con 54.228, Italia 42.651, Reino Unido 37.669 y Alemania con 36.029. En la región de Asia y Pacífico los ingresos fueron 183.097 millones, siendo los principales países turísticos China 54.720, Malasia 20.973, Hong Kong 17.154, Tailandia 14.464 y Macao 12.942. La región de América tuvo unos ingresos de 171.343

millones repartidos de la siguiente forma: Estados Unidos 96.712, Canadá 15.486, México 12.901, Países del Caribe 22.812, Zona América Central 6.363 y América del Sur 17.069 y, por último, la región con menores ingresos fue África y Oriente Medio con 61.854 millones, correspondiendo a esta última 33.614, siendo los países con más ingresos turísticos Egipto con 9.303, Arabia Saudita con 5.228 y Jordania con 2.312 millones de dólares.

No obstante, si el turismo es para muchas regiones su industria más activa también presenta un lado negativo, ya que las zonas turísticas son bastante susceptibles de sufrir ataques terroristas debido a que la aglomeración de personas permite que una organización terrorista logre ocasionar un mayor daño con los mismos recursos. Además, las acciones terroristas en zonas turísticas propician que la prensa internacional se ocupe de los fines de la organización como consecuencia de publicar la noticia del atentado terrorista proporcionando a la organización la repercusión que la mayoría de las organizaciones terroristas buscan. Cardulch (1993, pp. 357) señala que tanto para sus actos de violencia como para sus amenazas, los terroristas necesitan acceder a los medios de comunicación masivos que existen en las sociedades contemporáneas y que poseen un alcance universal. Ejemplos de ataques terroristas en zonas de especial relevancia turística los encontramos en Bombay (India), julio de 2006, 200 muertos y 700 heridos y noviembre de 2008 con 195 muertos y 327 heridos; Madrid (España), marzo de 2004 con 191 fallecidos y más de 1800 heridos, considerado el atentado más mortífero cometido en suelo europeo; New York (Estados Unidos), septiembre de 2001, ataque a las Torres Gemelas, con una cifra oficial de muertos de 2973 personas y multitud de heridos, siendo el atentado más devastador de la historia; Egipto, 2004, 2005, 2006 con más de 150 muertos y 400 heridos; Londres (Reino Unido), julio de 2005, con 56 personas fallecidas, incluidos los cuatro terroristas, y 700 heridos; Estambul (Turquía), noviembre de 2003, con 50 muertos y 450 heridos; Jordania, noviembre de 2005, con 57 muertos; Riad (Arabia Saudí), mayo y noviembre de 2003, con 52 fallecidos y 65 heridos; Mombasa (Kenia), noviembre de 2002, con 17 muertos y 80 heridos; Casablanca (Marruecos) con más de 45 muertos, etc.

Indiscutiblemente, la amenaza terrorista es un problema para los países receptores de turismo, ya que provoca incertidumbre sobre la renta de la zona receptora pues el terrorismo ahuyenta a los turistas como ocurrió en 1986 cuando se estimó que 1,8 millones de americanos cambiaron sus planes de viajar a Europa como consecuencia de distintos ataques terroristas en aeropuertos europeos. La amenaza terrorista, por tanto, tiene un importante impacto sobre la economía de las zonas receptoras y la Administración se ve forzada a realizar inversiones en seguridad como procedimiento de disuasión del terrorismo. Por tanto, las zonas turísticas requieren una doble inversión ya que, por una parte, se necesita llevar a cabo todas aquellas inversiones en infraestructura que permitan acoger a los turistas: aeropuertos, hoteles, carreteras, etc. y por otra parte, necesitan inversiones en seguridad con objeto de persuadir a los terroristas.

Pero, ¿qué niveles de inversión en infraestructuras y seguridad son necesarios para mantener un flujo monetario proveniente del turismo? Este trabajo analiza esta cuestión considerando un juego entre la Administración y los promotores turísticos de una determinada zona geográfica. El juego, de horizonte infinito, suma no nula y formulado en tiempo discreto, tiene en cuenta los objetivos de cada uno de los jugadores y la evolución sobre el tiempo de la amenaza terrorista y flujo de turistas en función de las inversiones en infraestructura y seguridad. El juego supone que la Administración persigue maximizar el beneficio social mientras que los promotores turísticos tratan de maximizar sus beneficios.

La teoría de juegos se utiliza frecuentemente para dar soluciones a problemas ligados al terrorismo, ya que las acciones terroristas y las medidas contraterroristas adoptadas por el gobierno de la zona sometida a la amenaza terrorista pueden asimilarse a estrategias de equilibrio no cooperativo que siguen dos jugadores en el desarrollo de un juego. Siguiendo a Sandler et al. (2003) y Rosendorff et al. (2005) la teoría de juegos es un instrumento apropiado para el estudio de problemas ligados al terrorismo y enumeran diferentes motivos para justificar su posición. Por ejemplo, los autores destacan las aportaciones de la teoría para captar las interrelaciones estratégicas entre los agentes; la posibilidad de llevar a cabo amenazas por parte de algún jugador y llegar a acuerdos entre jugadores para lograr ventajas estratégicas; además, la teoría permite considerar incertidumbre y suponer que algún agente o todos pueden estar mal informados.

La teoría de juegos ha sido seleccionada para el análisis de distintos problemas de terrorismo sobre distintos aspectos. Podemos citar Feitchinger et al. (2008) que plantean un juego diferencial para estudiar la elección de medidas eficientes para luchar contra el terror. Jacobson et al. (2007) llevan a cabo un análisis del conflicto entre los israelitas y los palestinos. Siquiera et al. (2006) se centran en el análisis estratégico entre un grupo terrorista y un gobierno desde el punto de vista del soporte de sus bases. Arce et al. (2005) analizan, también considerando la teoría de juegos, las condiciones que aseguran cuando son mejores las políticas preactivas respecto a las defensivas en la lucha contra el terrorismo.

La formulación del juego entre los promotores turísticos y la administración, que estudia este trabajo, considera algunos aspectos del problema de control en tiempo continuo formulado por Feichtinger et al. (2001) donde ya se tenían en cuenta las inversiones en infraestructura y seguridad. No obstante, el juego propuesto considera que la Administración puede desarrollar un papel activo dándola capacidad de decisión sobre el grado de aplicación de las medidas de seguridad en la zona turística; las inversiones en infraestructuras, se supone siguen siendo responsabilidad de los promotores turísticos. Considerando que la Administración y los promotores turísticos no cooperan, el trabajo busca estrategias de equilibrio no cooperativas feedback a seguir por ambos jugadores durante el desarrollo del juego. De este modo, se pretenden encontrar estrategias dependientes de los valores de las variables de estado, pero independientes de los valores iniciales de estas variables**.** El trabajo de Feichtinger et al. (2001) encontraba, suponiendo que los promotores turísticos adoptaban las medidas propugnadas por las condiciones de optimalidad, que el terrorismo y el turismo seguían comportamientos cíclicos para ciertos valores de un parámetro de bifurcación. Esa misma respuesta pretende encontrarse en este trabajo considerando dos parámetros de bifurcación en el sistema dinámico que recoge las estrategias de los jugadores.

El trabajo está dividido en secciones. En la segunda se formula el juego especificando los objetivos de los jugadores y las ecuaciones de evolución de las variables de estado. En la tercera sección se encuentran las estrategias feedback no cooperativas y se obtienen los coeficientes de las funciones de valor de ambos jugadores, comprobando que dichas funciones están perfectamente determinadas. En la sección cuarta estudiamos el sistema dinámico que siguen las variables de estado para la estrategia feedback de Nash encontrada, se determina la existencia de estado de equilibrio y la evolución de las variables de estado en función de las raíces de la ecuación característica asociada al sistema. En la siguiente sección se seleccionan dos parámetros del modelo como parámetros de bifurcación y se comprueba la existencia de ciclo límite. El trabajo concluye destacando aquellos resultados considerados más importantes.

2. FORMULACIÓN DEL JUEGO

Se considera una determinada zona geográfica de especial relevancia turística por su entorno natural, monumentalidad u otros aspectos, estando su desarrollo turístico gobernado por los promotores turísticos, mientras que la Administración pública se encarga de su administración.

2.1. El problema de los Promotores

Los promotores turísticos de la zona tratan de encontrar el nivel de inversiones que han de realizar en infraestructuras de modo que se logre maximizar el valor actual de los beneficios que proporciona la actividad turística durante cada uno de los intervalos de tiempo en los que se divide el horizonte temporal, supuesto no finito. Los beneficios en cada intervalo son función del turismo y de los costes tanto de la inversión en infraestructura como de las medidas de seguridad adoptadas para contrarrestar la amenaza terrorista.

Se supone que los promotores conocen el número de turistas en el momento inicial y la evolución del turismo en la zona durante cada intervalo $(t,t+1]$, que es función de la inversión en infraestructura, de la amenaza terrorista y del número de turistas en el intervalo anterior. Los turistas crecen con las inversiones en infraestructuras y con el número de turistas del periodo anterior, aunque disminuyen con la amenaza terrorista.

Considerando un intervalo $(t,t+1]$ del horizonte temporal, si se denota por $T(t)$, el número de turistas en la zona, por $I(t)$, la inversión en infraestructura, por $E(t)$, las medidas de seguridad implementadas y por $N(t)$, la amenaza terrorista, que siguiendo a Feichtinger et al. (2001) puede considerarse como número de terroristas, el problema para los promotores turísticos puede ser formulado

$$\max_{I} \sum_{t=0}^{\infty} \rho^t \left[f(T(t)) - C(I(t)) - C(E(t)) \right],$$
$$s.a: T(t+1) = g_1(I(t)) + \delta T(t) - g_2(N(t)),$$

donde el número de turistas en la zona en el momento inicial es conocido $T(0) = T_0$. En la formulación $\rho, 0 < \rho < 1$, es el factor de descuento; δ es un parámetro positivo con $1 < \delta < 2$, pues suponemos que $\delta = 1 + \alpha, \quad \alpha \in (0,1)$ siendo α el atractivo turístico de la zona, medido como una proporción de los turistas, $f(T(t))$ son los beneficios generados por los turistas, $C(I(t))$ son los costes de una inversión de $I(t)$ unidades monetarias en infraestructuras, $C(E(t))$ son los costes para los promotores por mantener $E(t)$ medidas de seguridad, establecidas por la administración, para disuadir a los terroristas, $g_1(I(t)), \ g_2(N(t))$, son funciones que recogen el impacto de las inversiones y del terrorismo sobre los turistas.

2.2 El problema de la Administración

Pretendiendo encontrar las medidas de seguridad que deben establecerse en la zona, la administración trata de maximizar el valor actual del bienestar social durante cada uno de los intervalos en los que se divide el horizonte temporal. En cada intervalo de tiempo $(t,t+1]$, el bienestar social se valora como la diferencia entre un porcentaje del beneficio alcanzado por los promotores durante ese intervalo y un porcentaje de la repercusión de la amenaza terrorista en la zona. Un funcional objetivo con estas mismas características puede encontrarse en Yeung et al. (1994).

Supondremos que la administración conoce el grado de amenaza terrorista en la zona en el momento inicial $N(0) = N_0$ y su evolución, que crece con el incremento del turismo y disminuye con la efectividad de las medidas extraordinarias adoptadas para luchar contra ellos. Por tanto, su problema puede formularse

$$\max_{E} \sum_{t=0}^{\infty} \rho^t \left[\beta \left[f(T(t)) - C(I(t)) - C(E(t)) \right] - \omega C(N(t)) \right],$$
$$s.a: N(t+1) = h_1(T(t)) - h_2(E(t), N(t)),$$

donde hemos supuesto que la administración selecciona el mismo factor de descuento que los promotores y $\beta, \ \omega,$ son parámetros de ponderación positivos. Si el número de turistas en la zona atrae que puedan cometerse actos de violencia terrorista, tendremos que la función h_1 ha de ser creciente. Por su parte, el impacto de las medidas de disuasión implementadas por la Administración sobre los terroristas viene recogido por la función h_2 y será creciente, así tendremos que si el impacto de las medidas crece, el número de terroristas deberá de disminuir.

Realizando una calibración de las funciones que intervienen en la formulación del modelo, con el objetivo de encontrar estrategias de equilibrio para los jugadores, el juego entre los promotores y la administración puede expresarse:

$$\max_{I} \sum_{t=0}^{\infty} \rho^t \left[aT - \frac{b}{2} I^2 - dE \right], \qquad \max_{E} \sum_{t=0}^{\infty} \rho^t \left[\beta \left(aT - \frac{b}{2} I^2 - dE \right) - \omega N \right],$$

$$s.a: T(t+1) = gI + \delta T - \Omega N, \qquad s.a: N(t+1) = mT - pE^{\frac{1}{2}} N^{\frac{1}{2}}$$

$$T(0) = T_0 \qquad N(0) = N_0$$

donde se ha eliminado la dependencia de las variables con respecto al tiempo con el objeto de no crear confusión. Observemos que los costes de inversión en infraestructura son considerados convexos como en Feichtinger et al. (2001), mientras que los costes de las medidas de seguridad se suponen lineales como el resto de las funciones que recogen la evolución de las variables de estado, excepto para la función de impacto, que se ajusta mediante una función creciente y cóncava.

3. ESTRATEGIA DE EQUILIBRIO NO COOPERATIVA FEEDBACK

Para encontrar la estrategia de equilibrio de Nash no cooperativa feedback consideramos como función de valor para los promotores turísticos

$$V(T,N,t) = \rho^t \left[A_1 T + B_1 N + D_1 \right],$$

mientras que para la administración hemos seleccionado como función de valor $U(T,N,t) = \rho^t \left[A_2 T + B_2 N + D_2 \right]$, donde A_1, B_1, D_1, A_2, B_2 y D_2 son constantes, y siguiendo a Basard et al. (1995), hay que resolver la ecuación de Hamilton-Jacobi-Bellman (H-J-B) para cada jugador.

Para los promotores turísticos, la ecuación a resolver es

$$V(T,N,t) = \max_{I} \left[\rho^t \left[aT - \frac{b}{2} I^2 - dE \right] + V(T,N,t+1) \right],$$

y la Administración resolverá

$$U(T,N,t) = \max_{E} \left[\rho^t \left[\beta(aT - \frac{b}{2} I^2 - dE) - \omega N \right] + U(T,N,t+1) \right].$$

Resolviendo el problema clásico sin restricciones asociado a los promotores, cuya función objetivo es

$$aT - \frac{b}{2} I^2 - dE + \rho A_1 \left[gI + \delta T - \Omega N \right] + \rho B_1 \left[mT - pE^{\frac{1}{2}} N^{\frac{1}{2}} \right] + \rho D_1$$

encontramos que las condiciones necesarias de primer orden son también suficientes, ya que la función objetivo es cóncava respecto a las inversiones en infraestructura. Aplicando las condiciones de primer orden, se encuentra que las inversiones tienen que permanecer constantes durante todo el horizonte temporal y siguiendo la expresión:

$$I = \frac{\rho A_1 g}{b}.$$

Similarmente, resolviendo el problema clásico sin restricciones asociado a la Administración, que tiene como función objetivo la expresión:

$$\beta(aT-\frac{b}{2}I^2-dE)-\omega N+\rho A_2[gI+\delta T-\Omega N]+\rho B_2\left[mT-pE^{\frac{1}{2}}N^{\frac{1}{2}}\right]+\rho D_2$$

donde, de nuevo, las condiciones necesarias de primer orden también son suficientes al ser la función objetivo cóncava respecto a las medidas de seguridad, se obtiene

$$E^{\frac{1}{2}}=-\left(\frac{\rho pB_2}{2d\beta}\right)N^{\frac{1}{2}},$$

que requiere que B_2 sea negativo y que expresa que las medidas de seguridad a imponer en la zona, en cada intervalo, son proporcionales a la amenaza terrorista durante ese intervalo.

Una vez resueltos los problemas estáticos es necesario comprobar que se verifica la ecuación de H-J-B y de este modo encontrar los valores de los coeficientes que determinan las funciones de valor. Operando, se obtiene

$$A_2=\beta a+\delta\rho A_2+m\rho B_2\Rightarrow A_2=\frac{\beta a+m\rho B_2}{1-\delta\rho}$$

$B_2=\frac{\rho^2\ p^2}{4d\beta}{B_2}^2-\omega-\rho A_2\Omega$, luego sustituyendo A_2 por su valor, se tiene

$$B_2=\begin{cases}\frac{2d\beta}{\rho^2p^2}\left[1+\frac{m\Omega\rho^2}{1-\delta\rho}+\sqrt{(1+\frac{m\Omega\rho^2}{1-\delta\rho})^2+\frac{p^2\rho^2}{d\beta}(\omega+\frac{\beta a\rho\Omega}{1-\delta\rho})}\right]\\ \frac{2d\beta}{\rho^2p^2}\left[1+\frac{m\Omega\rho^2}{1-\delta\rho}-\sqrt{(1+\frac{m\Omega\rho^2}{1-\delta\rho})^2+\frac{p^2\rho^2}{d\beta}(\omega+\frac{\beta a\rho\Omega}{1-\delta\rho})}\right]\end{cases}$$

por tanto, B_2 proporciona dos valores, uno positivo y otro negativo, suponiendo $1>\delta\rho$. Seleccionando el valor negativo, y teniendo en cuenta los requerimientos de las medidas de seguridad, se encuentra que el resto de los coeficientes de las funciones de valor son:

$$A_1=a+\rho A_1\delta+\rho B_1 m\Rightarrow A_1=\frac{a+m\rho B_1}{1-\delta\rho},$$

$$B_1=-d(\frac{\rho pB_2}{2d\beta})^2-\rho A_1\Omega+\rho B_1\frac{\rho p^2B_2}{2d\beta}\Rightarrow B_1=-\frac{\frac{\rho^2p^2{B_2}^2}{4d\beta^2}+\frac{\rho\Omega a}{1-\delta\rho}}{\sqrt{(1+\frac{m\Omega\rho^2}{1-\delta\rho})^2+\frac{p^2\rho^2}{d\beta}(\omega+\frac{\beta a\rho\Omega}{1-\delta\rho})}},$$

$$D_1=-\frac{b}{2}(\frac{\rho A_1g}{b})^2+\frac{(\rho A_1g)^2}{b}+\rho D_1\Rightarrow D_1=\frac{(\rho A_1g)^2}{2b(1-\rho)},$$

$$D_2=\frac{\rho^2g^2A_1(2A_2-\beta A_1)}{2b(1-\rho)}.$$

Por tanto, los coeficientes de las funciones de valor están totalmente determinados y, por consiguiente, también lo están las funciones de valor de ambos jugadores. Observemos que las funciones de valor recogen los valores de los funcionales de los jugadores

$$V(T,N,t)=\sum_{z=t}^{\infty}\rho^{z}\left[aT(z)-\frac{b}{2}I(z)^{2}-dE(z)\right],$$

$$U(T,N,t)=\sum_{z=t}^{\infty}\rho^{z}\left[\beta\left[aT(z)-\frac{b}{2}I(z)^{2}-dE(z)\right]-\omega N(z)\right].$$

4. LAS SOLUCIONES DEL SISTEMA DINÁMICO

Sustituyendo los valores de los controles en las ecuaciones en diferencias que determinan la evolución de la amenaza terrorista y turistas, se encuentra el sistema dinámico lineal

$$T(t+1)=\frac{\rho A_1 g^{2}}{b}+\delta T(t)-\Omega N(t)$$

$$N(t+1)=mT(t)+\frac{\rho\ p^{2}B_2}{2d\beta}N(t)$$

y llamando

$$H_1=\frac{g^{2}\rho A_1}{b},\qquad H_2=\frac{p^{2}\rho B_2}{2\beta d}<0,$$

el estado de equilibrio (T^{*},N^{*}) ha de verificar

$$\begin{pmatrix}1-\delta & \Omega\\ -m & 1-H_2\end{pmatrix}\begin{pmatrix}T^{*}\\ N^{*}\end{pmatrix}=\begin{pmatrix}H_1\\ 0\end{pmatrix},$$

y comprobamos que sólo existe estado de equilibrio si $(1-\delta)(1-H_2)+m\Omega\neq 0,$ en cuyo caso el único estado de equilibrio verificará:

$$\begin{pmatrix}T^{*}\\ N^{*}\end{pmatrix}=\begin{pmatrix}1-\delta & \Omega\\ -m & 1-H_2\end{pmatrix}^{-1}\begin{pmatrix}H_1\\ 0\end{pmatrix};$$

operando obtenemos que los valores de equilibrio satisfacen

$$T^{*}=\frac{(1-H_2)H_1}{(1-\delta)(1-H_2)+m\Omega},\qquad N^{*}=\frac{mH_1}{(1-\delta)(1-H_2)+m\Omega}.$$

La ecuación característica asociada al sistema es:

$$\begin{vmatrix}\delta-\lambda & -\Omega\\ m & H_2-\lambda\end{vmatrix}=0\ \Rightarrow\ \lambda^{2}-(\delta+H_2)\lambda+(\delta H_2+m\Omega)=0$$

y bajo el supuesto de existencia de estado de equilibrio, esto es, que $(1-\delta)(1-H_2)+m\Omega\neq 0,$ tendremos que $\lambda=1$ no puede ser raíz de la ecuación característica.

Si denotamos por λ_1 y λ_2 las raíces de la ecuación característica, tendremos distintas expresiones de la solución general del sistema dependiendo de cómo sean las raíces.

- Si son reales y distintas la solución general del sistema será:

$$T(t) = E_1\lambda_1^t + E_2\lambda_2^t + E_3$$
$$N(t) = \frac{E_1}{\Omega}(\delta - \lambda_1)\lambda_1^t + \frac{E_2}{\Omega}(\delta - \lambda_2^t)\lambda_2^t + \frac{H_1 + E_3(\delta - 1)}{\Omega},$$

donde E_1, E_2 son constantes arbitrarias y E_3 toma el valor:

$$E_3 = \frac{(1 - H_2)H_1}{(1 - \delta)(1 - H_2) + m\Omega}.$$

- Si las raíces son iguales, la solución general del sistema será:

$$T(t) = (E_1 + E_2 t)\lambda^t + E_3$$
$$N(t) = \frac{(\delta - \lambda)}{\Omega}(E_1 + E_2 t)\lambda^t - \frac{\lambda E_2}{\Omega}\lambda^t + \frac{H_1 + E_3(\delta - 1)}{\Omega}.$$

- Si las raíces son complejas, en este caso las denotamos por:

$$\lambda_1 = r_1 + r_2 i = R(\cos\alpha + i\ sen\ \alpha)$$
$$\lambda_2 = r_1 - r_2 i = R(\cos\alpha - i\ sen\ \alpha)$$

donde $R = \sqrt{r_1^2 + r_2^2}$, $\alpha = arc\ tg\frac{r_2}{r_1}$, tendremos que la solución general del sistema adoptará la forma:

$$T(t) = R^t(M_1\cos\alpha t + M_2\ sen\ \alpha t) + E_3$$
$$N(t) = \frac{R^t}{\Omega}\left[\delta(M_1\cos\alpha t + M_2\ sen\ \alpha t) + R(M_1\cos\alpha(t+1) + M_2\ sen\ \alpha(t+1)\right] + \frac{H_1 + E_3(\delta - 1)}{\Omega}.$$

5. LA EXISTENCIA DE CICLO LÍMITE

Bajo las hipótesis de que el modelo presenta estado de equilibrio, cuando se comienza a potenciar una zona turística, los promotores turísticos invierten en infraestructuras con el objeto de mejorar las condiciones de estancia de los turistas y con ello se incrementan las visitas turísticas a la zona. Si este proceso continua, la zona albergará cierta aglomeración de turistas, pero esto supone una llamada para los terroristas. Ahora bien, considerar la zona por los terroristas como un medio para extender el terror, ahuyenta a los turistas, encontrándose los promotores en la situación inicial. Ese proceso dinámico debe estar inmerso en la dinámica del juego estudiado y para comprobarlo se seleccionan valores para los parámetros del modelo, excepto para dos que se consideran como parámetros de bifurcación.

Los parámetros seleccionados son *m* y *p*, el primero interviene en la función f_1 que recoge la influencia de la amenaza terrorista sobre el turismo y el segundo, determina la definición de la función h_2 que recoge la efectividad de las medidas que se adopten para la lucha

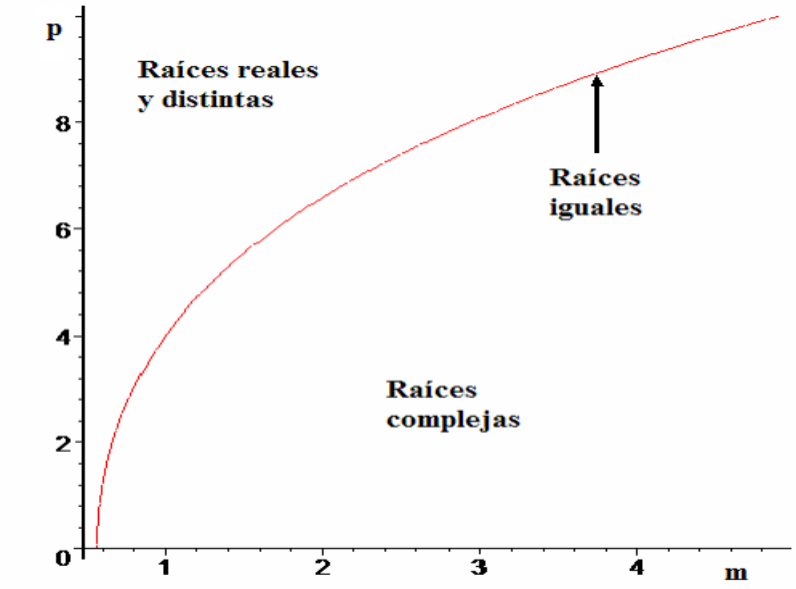
contraterrorista. El resto de los parámetros se supone toman los valores: $\rho = 0.05$, $\delta = 1.5$, $\beta = \omega = 0.8$, $a = 5$, $b = 4$, $d = 1$, $g = 0.5$, $\Omega = 1$, y puesto que el discriminante de la ecuación característica asociada al sistema es:

$$(\delta + H_2)^2 - 4(\delta H_2 + \Omega m)$$

tendremos que las raíces serán reales y distintas, reales y múltiples o complejas dependiendo de que el discriminante sea mayor, igual o menor que cero.

Las diferentes regiones se muestran en la Figura 1 considerando un plano en las variables (*m, p*).

Figura 1. Comportamiento de los valores propios en el plano de fase m-p



Para que pueda aparecer un ciclo límite, las dos raíces tiene que tener módulo unitario, pues en otro caso el estado de equilibrio sería un punto de silla, asintóticamente estable o bien inestable. Luego vamos a estudiar las distintas posibilidades dependiendo de cómo sean las raíces.

1. Si las raíces son iguales, su valor será $\lambda = \frac{1}{2}(\delta + H_2)$ y como el módulo es unitario, tendremos que $\lambda = -1$ y entonces $H_2 = -2 - \delta$. Pero para los valores seleccionados de los parámetros las curvas $(\delta + H_2)^2 - 4(\delta H_2 + \Omega m) = 0$ y $H_2 = -2 - \delta$ se cortan en *m=6.25* y *p=11.03* que son valores muy altos y con escaso o nulo sentido en el planteamiento del problema.

2. Si las raíces son reales y distintas, están en la región del plano de fase m-p donde se verifica $(\delta + H_2)^2 - 4(\delta H_2 + \Omega m) > 0$ y como han de tener módulo unitario para que aparezca un ciclo, tendremos que una es *-1* y la otra *1* que sabemos no es posible.

3. Si las raíces son complejas están en la región del plano de fase m-p donde se verifica $(\delta + H_2)^2 - 4(\delta H_2 + \Omega m) < 0$ y toman los valores

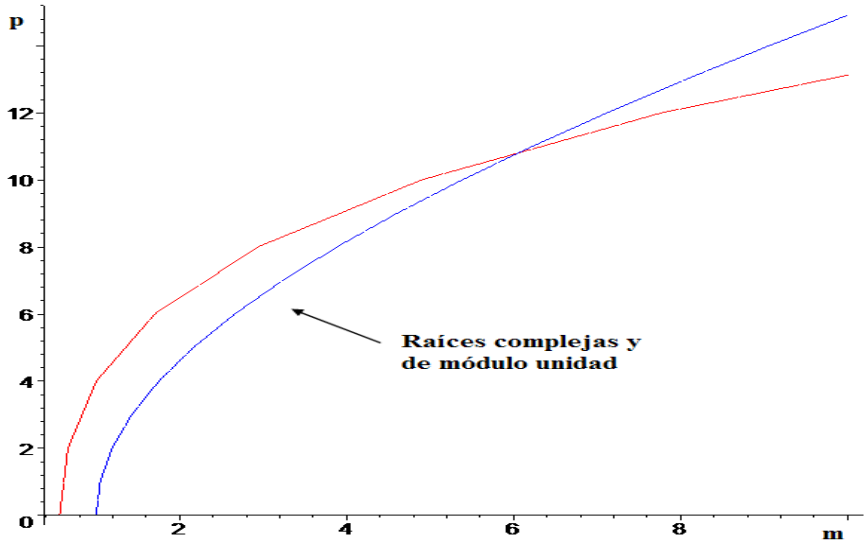
$$\lambda_1 = \frac{1}{2}(\delta + H_2) + \frac{1}{2} i \sqrt{4(\delta H_2 + m\Omega) - (\delta + H_2)^2}$$

$$\lambda_2 = \frac{1}{2}(\delta + H_2) - \frac{1}{2} i \sqrt{4(\delta H_2 + m\Omega) - (\delta + H_2)^2}$$

además, como su módulo tiene que ser unitario, se ha de verificar $\delta H_2 + m\Omega - 1 = 0$. Los valores de *m* y *p* donde esto ocurre se muestran en la Figura 2. Por tanto, es para estos valores de los parámetros donde tiene lugar la bifurcación de Neimark-Sacker, para valores cercanos a la curva, por la derecha, el punto de equilibrio es inestable, mientras que valores próximos a la curva pero por la izquierda, el estado de equilibrio es asintóticamente estable.

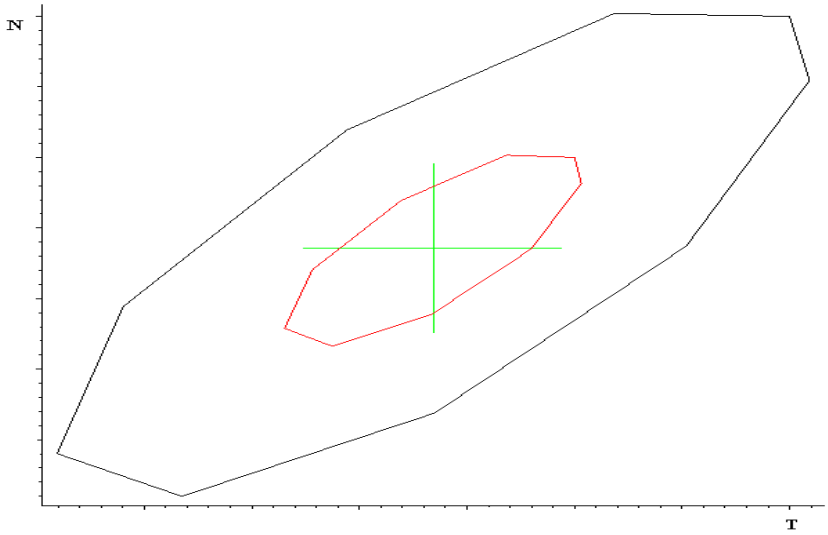
Observamos que los valores de *m* y *p* han de verificar $m < 6.25$ y $p < 11.03$, y si tomamos un valor especifico para *p*, por ejemplo $p = 0.5$, el valor de *m* es $m = 1.011873961$ y el estado de equilibrio que se obtiene es $T^* = 0.0334290502$, $N^* = 0.0335603226$.

Figura 2. Raíces complejas de módulo unitario en el plano de fase *m-p*



En la Figura 3 representamos el ciclo límite en el plano *(T,N)*, en dos situaciones distintas dependiendo de las condiciones iniciales de partida. El ciclo exterior se ha obtenido partiendo de las condiciones iniciales $T(0) = 0.05, \quad N(0) = 0.05$, mientras que para el ciclo interior las condiciones iniciales de partida han sido $T(0) = 0.04, \quad N(0) = 0.04$.

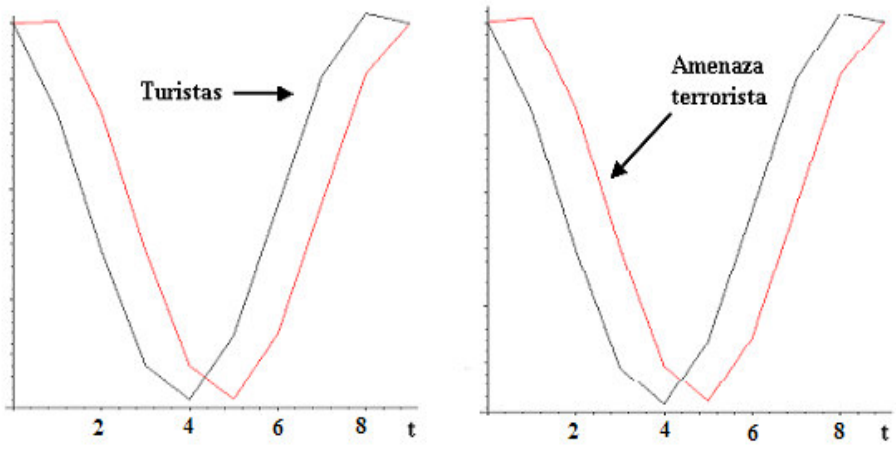
Figura 3. Ciclo límite



La evolución de las medidas de seguridad a adoptar en la zona y de los turistas en el plano *(E,T)* para ambas condiciones iniciales tiene también una representación cíclica similar a la recogida en la Figura 3.

En la Figura 4 representamos la evolución de los turistas en la zona en el plano *(t,T)* y de la amenaza terrorista en el plano *(t,N)*, para las dos condiciones iniciales de partida, mostrando la gráfica de la izquierda las correspondientes evoluciones para las primeras condiciones y la de la derecha para las segundas.

Figura 4. Evolución de los turistas y de la amenaza terrorista



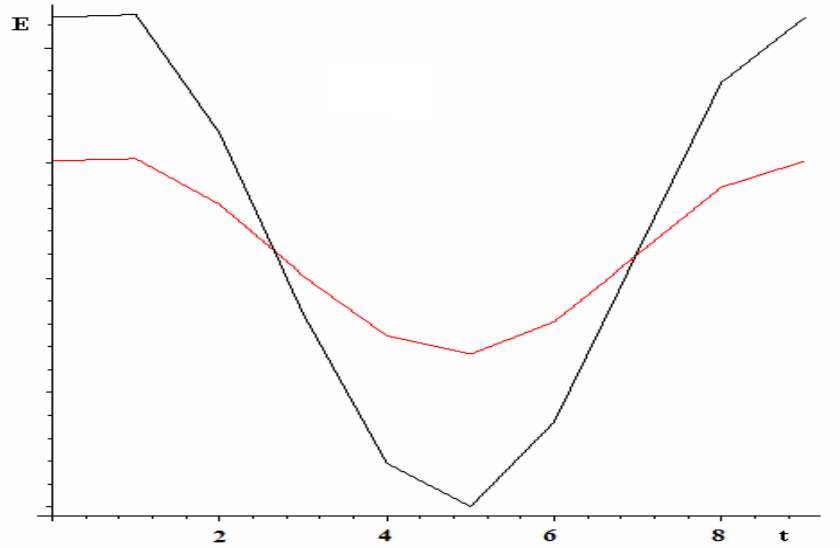
En ambas gráficas se observa la amplitud del ciclo es de nueve intervalos de tiempo; sin embargo, el comportamiento es ligeramente diferente. Así, la evolución del número de turistas en la zona durante los cuatro primeros intervalos es decreciente hasta alcanzar su nivel mínimo en $t = 4$, a partir de este instante, y durante los cuatro intervalos siguientes, crece el número de turistas y alcanza su valor máximo en $t = 8$, para de nuevo decrecer en el último intervalo del ciclo. Luego el nivel máximo de turistas en la zona se alcanza en los instantes $t = 8, 17, 26, \cdots$, mientras que el nivel mínimo se produce en $t = 4, 13, 22, \cdots$

El comportamiento de la amenaza terrorista es similar aunque ligeramente diferente, pues ahora durante el primer intervalo del ciclo, la amenaza terrorista crece, decrece en los cuatro siguientes intervalos para alcanzar en $t = 5$ su nivel mínimo y volver a crecer en los cuatro intervalos siguientes. El nivel máximo se alcanza ahora en los instantes

$t = 1, 10, 19, \cdots$ y el nivel mínimo en $t = 5, 14, 23, \cdots$ Observamos que la amenaza terrorista es alta cuando también es alta la existencia de turistas en la zona y disminuye cuando también disminuyen los turistas, como era de esperar, pero está desplazada, su nivel mínimo se alcanza en el intervalo siguiente al de los turistas, mientras que su nivel máximo se alcanza dos intervalos después de que los turistas en la zona alcancen su máximo nivel.

Por último, la Figura 5 recoge la evolución, también cíclica, de las medidas a adoptar en ambos casos; la curva más alta es para las condiciones iniciales primeras. En el primer intervalo crece ligeramente, para en el resto de los intervalos del dos al cinco decrecer, alcanzando en este instante su valor más bajo, en concordancia con la amenaza terrorista que también en este instante alcanzaba el mínimo, en los intervalos siguientes crece, hasta completar el ciclo en el noveno intervalo. Tiene, como es lógico esperar, con comportamiento totalmente análogo a la amenaza terrorista.

Figura 5. Evolución de las medidas a adoptar



6. CONCLUSIONES

El turismo es en algunas regiones una importante fuente de ingreso, pero también las zonas turísticas son más atractivas que otras regiones para sufrir ataques terroristas, precisamente por la repercusión que un acto o amenaza tendría a nivel internacional y que proporcionaría a las organizaciones terroristas una publicidad importante, además de lograr causar cuantiosos daños, con escasos recursos, debido a la aglomeración que normalmente presentan las zonas turísticas.

Por este motivo, las zonas turísticas requieren medidas adicionales de seguridad que disuadan a las organizaciones terroristas y no ahuyenten a los turistas. Pero al mismo tiempo, las zonas turísticas requieren importantes inversiones en infraestructuras que permitan acoger a los turistas y proporcionarles una agradable estancia.

Bajo estas hipótesis, el trabajo presenta un modelo que es un juego dinámico entre la Administración y los promotores turísticos de la zona, pues en los segundos se ha considerado que recaen las inversiones en infraestructuras, mientras que la Administración es responsable de las inversiones en seguridad y se pretende determinar los niveles de inversión en infraestructura y seguridad necesarios para mantener un flujo monetario proveniente del turismo.

El estudio se ha planteado en tiempo discreto y horizonte temporal finito, es un juego de suma no nula, donde cada jugador trata de maximizar el valor actual de sus beneficios, los promotores, y el bienestar social la administración y conocen la evolución en el tiempo del flujo de turistas, que depende de las inversiones en infraestructuras y de la amenaza terrorista, así como la evolución de esta última que depende del turismo y de la efectividad de las medidas de

seguridad que se adopten para luchar contra el terrorismo. En el modelo se supone que se parte de unas condiciones iniciales conocidas.

Ajustando, mediante expresiones específicas, las funciones que intervienen en la formulación del modelo, encontramos la estrategia de equilibrio de Nash no cooperativa feedback y observamos que las inversiones en infraestructuras han de permanecer constantes en todo el horizonte temporal, mientras que las medidas de seguridad que han de adoptarse en cada intervalo son proporcionales a la amenaza terrorista existente en ese intervalo.

Con los valores encontrados planteamos el sistema dinámico que siguen las ecuaciones de estado de los turistas y de la amenaza terrorista y determinamos la existencia de un único estado de equilibrio, así como la solución general de dicho sistema dinámico.

Dando valores a los parámetros del modelo, excepto a dos, que consideramos parámetros de bifurcación, hemos tratado de comprobar si se presenta un ciclo límite. Del estudio de las regiones donde están las raíces de la ecuación característica asociada al sistema dinámico, obtenemos que sólo existe ciclo límite cuando las raíces son complejas y tienen módulo unitario. Por tanto, tomando unos valores de los parámetros de bifurcación donde esto ocurra, hemos obtenido el correspondiente estado de equilibrio y la evolución cíclica de los turistas y de la amenaza terrorista, partiendo de dos condiciones iniciales distintas, así como la amplitud del ciclo que es de nueve intervalos de tiempo.

Observamos que la evolución de los turistas es decreciente durante los cuatro primeros intervalos, creciente durante los cuatro siguientes y vuelve a decrecer en el último. Sin embargo, el comportamiento de la amenaza terrorista es ligeramente diferente, pues durante el primer intervalo del ciclo, la amenaza crece, decrece en los cuatro siguientes y vuelve a crecer en los cuatro últimos. Se muestra que existe un cierto desplazamiento entre la amenaza terrorista y los turistas, su nivel más bajo se alcanza un intervalo posterior al de los turistas, mientras que el nivel máximo se consigue dos intervalos después.

BIBLIOGRAFÍA

ARCE, D.G. y T. SANDLER (2005): "Counterterrorism. A Game-Theoretic Analysis", The Journal of Conflict Resolution, Vol. 49, nº 2, April, pp. 183-200.

BASAR, T. y G.J. OLSDER (1995): "Dynamic Noncooperative Game Theory", Academic Press, Segunda Edición, London.

CALDULCH CERVERA, R. (1993): "Dinámica de la Sociedad Internacional". Ed. Ramón Areces, Madrid.

ENDER, W. y T. SandLER (1991): "Causality Between Transnational Terrorism and Tourism: The Case of Spain", Terrorism, Vol 14, nº 1, pp. 49-58.

FEICHTINGER, G., R.F. HARTL, P.M. KORT, y A.J. NOVAK (2001): "Terrorism Control in the Tourism Industry", Journal of Optimization Theory and Applications, Vol. 108, nº 2, February, pp. 283-296.

FEICHTINGER, G. y A.J. NOVAK (2008): "Terror and Counterterror Operations: Differential Game with Cyclical Nash Solution", Journal of Optimization Theory and Applications, Vol. 139, nº 3, December, pp. 541-556.

JACOBSON, D. y E.H. KAPLAN (2007): "Suicide Bombings and Targeted Killings in (Counter-) Terror Games", The Journal of Conflict Resolution, Vol. 51, nº 5, October, pp. 772-792.

ORGANIZACIÓN MUNDIAL DE TURISMO (2008): "Barómetro OMT del Turismo Mundial". Vol. 6, nº 3, octubre 08.

ROSENDORFF, B.P. y T. SandLER (2005): "The political Economy of Transnational Terrorism", The Journal of Conflict Resolution, Vol. 49, nº 2, April, pp. 171-182.

SandLER, T. y D.G. ARCE (2003): "Terrorism and Game Theory". Simulation and Gaming, Vol. 34, nº 3, pp. 319-337.

SIQUIERA, K y T. SandLER (2006): "Terrorists versus the Government: Strategic Interaction, Support and Sponsorship". The Journal of Conflict Resolution, Vol. 50, nº 6, December, pp. 878-898.

SOTO TORRES, M.D. y R. FERNÁNDEZ LECHÓN (1987): "Análisis y Estabilidad de los Sistemas Dinámicos". Ed. Mata, Valladolid, España.

YEUNG, D.W.K. and M.T. CHEUNG (1994): "Capital Accumulation Subject to Pollution Control: A Differential Game with a Feedback Nash Equilibrium", pp. 289-300. In T. Basar and A. Haurie ed., "Advances in Dynamic Games and Applications", Boston, Massachusetts.

ESTIMACIÓN DEL EVA EN LAS EMPRESAS DEL IBEX-35: UNA APLICACIÓN CON REDES NEURONALES

MERCEDES TORRES JIMÉNEZ
E-mail: mtorres@etea.com
ARACELI DE LOS RÍOS BERJILLOS
E-mail: arios@etea.com
Departamento de Gestión empresarial y Métodos Cuantitativos
Facultad de Cc. Económicas y Empresariales -ETEA
Universidad de Córdoba

Resumen

En el presente trabajo se presenta una aplicación con redes neuronales artificiales para estimar la relación entre el valor financiero generado por las organizaciones y sus grupos de interés o stakeholders, a los que se considera determinantes en la generación de dicho valor. La muestra está constituida por las empresas españolas del IBEX-35; el EVA (economic value added) es la variable empleada para representar el valor financiero y las variables independientes son indicadores de los distintos stakeholders. La metodología empleada, al tratarse de una técnica no paramétrica, permite superar determinadas restricciones que deben cumplirse en las técnicas paramétricas tradicionales como, por ejemplo, la existencia de multicolinealidad entre las variables explicativas del modelo. Además, se extraen conclusiones sobre la capacidad de predicción y generalización del modelo estimado, para lo que se han realizado distintas particiones de la muestra.

Palabras clave: EVA, stakeholders, redes neuronales, capacidad de predicción y generalización.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

The present paper shows an application with artificial neural networks to estimate the relationship between the financial value generated by the organizations and their stakeholders, who has a crucial role in this value. The sample consists of firms that were part of the Spanish IBEX-35 for five consecutive years. The economic value added (EVA) is the variable used to represent the financial value and the independent variables are indicators representing stakeholders. The methodology, as a non-parametric technique, can overcome certain data limitations like, for example, the existence of mulcolineality between the explanatory variables and heteroskedasticity.
This study analyzes both the predictive and generalizing capacities of the best artificial neural network model found.

Key words: EVA, stakeholders, artificial neural networks, predictive and generalizing capacities.
Thematic Area: Quantitative Methods for Economics and Business.

INTRODUCCIÓN

La determinación del valor de una empresa constituye un tema clásico de investigación pero de permanente actualidad como muestran los recientes trabajos de Prakash y Kumar (2008) o Castaño (2008); su interés responde a múltiples razones, internas y/o externas a la organización. Entre las razones internas señalaríamos: conocer la situación o evolución del patrimonio, evaluar la gestión del equipo directivo, determinar una política de dividendos, estudiar la capacidad de apalancamiento o ampliación de capital, así como otro tipo de motivos legales, de

seguro o por causas de herencia o sucesión. Mientras que algunas de las razones externas que justifican una valoración podrían ser la transmisión de propiedad, procesos buy-out, fusiones y adquisiciones, solicitud de préstamos, privatizaciones o estudios de viabilidad. Sin embargo, esta medición implica numerosas dificultades técnicas por lo que son múltiples los métodos que se han desarrollado con esta finalidad: descuento de flujos futuros de caja; descuento de dividendos; valor presente ajustado; el modelo del valor añadido económico (EVA); la utilización de compañías comparables; metodología de las opciones reales, etc.

Por otra parte, desde que Freeman (1984) estableciera la teoría de los grupos de interés y su papel en las organizaciones, son muchos los autores que han analizado la importancia de las relaciones que las empresas mantienen con sus stakeholders y su repercusión en la organización, por ejemplo, Clarkson (1995); Donaldson y Preston (1995), Mitchell, Agle & Wood (1997); Jones & Mechas (1999); Winstanley y Stoney (2001).

Sin embargo, este es también un tema controvertido, por ejemplo, tal y como ponen de manifiesto en sus trabajos Whysall (2000), Donaldson y Preston (1995) o Argandoña (1998), no existe un consenso sobre cuáles son los grupos de interés de una organización. Collins, Kearins y Roper (2005) han insistido en los riesgos de basar el diseño de la gestión en las relaciones con los grupos de interés. Sin embargo, otros autores como Hart y Milstein (2003) o Hart y Sharma (2004) opinan que a través de la relación con sus grupos de interés las empresas pueden anticiparse a futuras fuentes de problemas e identificar nuevas e innovadoras ideas empresariales para mantener una ventaja competitiva, por lo que contribuyen a la generación de valor y a la sostenibilidad.

Aparentemente la generación de valor financiero parece un objetivo que colisiona con la satisfacción de las demandas de los grupos de interés, es decir, se plantea el debate de si las organizaciones deberían maximizar el valor o si deberían actuar a favor de los intereses de los stakeholders. En opinión de Jensen (2001) la maximización del valor debe ser entendida en condiciones "razonables" como la maximización del bienestar social. Un bienestar social que sólo se alcanza cuando se cumplen las expectativas de los grupos de interés. Jensen afirma que no existe una solución eficiente a este problema, ya que en la maximización del valor podemos estar favoreciendo más a unos grupos que a otros.

La revisión bibliográfica efectuada nos permite introducir la siguiente hipótesis de nuestro estudio: *existe una relación positiva entre el valor financiero de las empresas del IBEX-35, medido a través del EVA y las relaciones mantenidas por la empresa con sus principales stakeholders, representadas estas relaciones a través de un conjunto de indicadores cuantitativos.*

El presente trabajo tratar de contrastar esta hipótesis empleando para ello una técnica no paramétrica las redes neuronales artificiales. En concreto, se han estimado dos modelos distintos, en el primero, denominado modelo pool, sólo se contemplan como variables explicativas los distintos indicadores de stakeholder seleccionados, en el segundo, denominado modelo de efectos fijos, se ha tenido en cuenta que cada empresa tiene una incidencia directa y diferente a la del resto de empresas en la generación de valor, es decir, cada empresa tiene un efecto fijo propio en el EVA. Además, se trata de medir la capacidad de predicción y generalización de los distintos modelos estimados.

Frente a otros estudios relacionados con esta materia, esta investigación presenta las siguientes aportaciones: 1ª) la estimación del EVA de las empresas analizadas; 2ª) la selección de los indicadores para representar a los distintos stakeholders con los que se relaciona la empresa, así como la recopilación de los datos correspondientes; 3ª) el diseño de un modelo empírico para la estimación del valor de estas empresas basado en una hipótesis hasta ahora poco analizada, según la cual el valor generado por las empresas depende de las inversiones y relaciones que la organización mantiene con todos y cada uno de sus stakeholders, considerando como una variable explicativa más las particularidades propias de cada organización; 4ª) la medición de la capacidad explicativa de cada modelo así como su capacidad de predicción y generalización.

ESTUDIO EMPÍRICO

Selección de la muestra, definición de variables y captación de datos

Se escogieron como base del estudio las empresas españolas que habían cotizado en el IBEX-35 entre los años 2000 y 2004. Las razones de selección de estas empresas son varias: en primer lugar, se partía de la hipótesis de que el valor de cotización de las acciones (como indicador del stakeholder accionistas) podía estar relacionado con el valor de las empresas, de ahí que la muestra debiera estar constituida por empresas cotizadas en bolsa, en segundo lugar, se trata de un grupo de empresas de las que se presupone que existe más información y, por tanto, que el éxito en la captación de los datos iba a ser mayor.

Variable dependiente

Desde nuestro punto de vista, el valor económico añadido (EVA) es una de las herramientas que mejor permite calcular la riqueza generada por la empresa en un período de tiempo dado, teniendo en cuenta el riesgo con el que opera, por lo que puede ser un indicador susceptible de aportar información relevante para la toma de decisiones por parte de los distintos colectivos con intereses en una empresa (stakeholders) y de los que ésta depende para su supervivencia.

Aunque son conocidas las críticas vertidas por algunos autores como Ventura, Ordóñez, García y Arias (2003), Carmona, Deyá y Gutiérrez (2001) o Brickley, Smith y Zimmerman (1997), sobre las carencias detectadas en el cálculo del EVA, por su carácter relativamente estático (al no considerar las expectativas de futuro) y la supuesta incapacidad a la hora de representar el valor del capital intelectual integrado en la empresa, éste ha sido seleccionado, ya que, como indica Mulder (1998), lo que realmente genera valor, es el grado en que los recursos (tangibles o no), se pueden convertir en un resultado financiero para la empresa.

Cálculo del EVA

A continuación se expone el cálculo del EVA según la definición de sus autores, para a continuación explicar los ajustes efectuados en la estimación de este indicador, a partir de la cual se han obtenido los valores del EVA para las empresas del IBEX-35 en el período 2000-2004. Adviértase que este indicador no es publicado por las empresas, de ahí que una parte importante del proceso de captación de datos haya consistido en su cálculo. La fórmula de cálculo del EVA propuesta evita que en su estimación quede recogido el impacto de actividades extraordinarias que, realmente, no generan valor a largo plazo en las organizaciones.

El valor económico añadido se define como el importe resultante de deducir a los ingresos la totalidad de los gastos, incluidos el coste de oportunidad del capital y los impuestos.

De aquí, que dicho valor económico añadido, pueda calcularse como la diferencia entre el resultado de las actividades ordinarias antes de intereses y después de impuestos (BAIDI) y el valor contable del activo corregido por el coste medio del pasivo (CMPC):

$$EVA = BAIDI - (\text{Valor contable activo} * CMPC) \quad [1]$$

El BAIDI se obtiene sumando los intereses al beneficio neto y eliminando los resultados extraordinarios. El valor contable del activo es el valor del activo de la empresa, de acuerdo con su balance de situación y el coste promedio del pasivo o coste medio ponderado de capital, representa el coste medio de toda la financiación que ha obtenido la empresa.

El BAIDI se ha obtenido a partir del resultado neto, incorporando los gastos financieros y restando los beneficios extraordinarios; en el caso de existencia de pérdidas extraordinarias, éstas son sumadas a dicho resultado. El motivo de no considerar los ingresos financieros se debe al hecho de que estos son generados por los activos en los que está invirtiendo la empresa y, por tanto, forman parte de su resultado ordinario. En cambio, los gastos financieros se añaden al resultado neto, para anular la deducción que se hace de los mismos en la cuenta de resultados.

Para la determinación del valor de los activos a considerar, existen varias opciones: valor de adquisición, valor de mercado, valor promedio de los activos y valor del activo neto. El criterio más extendido es el que considera el activo neto, ya que en la medida que la empresa obtiene financiación espontánea, sería menos precisa la obtención de otro tipo de financiación para el activo en el que está invirtiendo, este ha sido el criterio considerado para estimar el activo en esta investigación.

Tras el oportuno análisis de los estados financieros presentados por las empresas de la 4uestra, se detectó que en algunos años aparecían ciertas activaciones de provisiones o plusvalías generadas por determinadas desinversiones, así como ciertos ingresos o gastos extraordinarios de gran magnitud, que tenían una elevada incidencia sobre los resultados finales de dichos ejercicios. Ante el análisis de sus posibles causas (impacto fiscal, resultados anormalmente menores o mayores a su trayectoria en los últimos años, grandes indemnizaciones por despidos y otras causas difícilmente justificables), se observó que aunque, en teoría, el EVA trabaja solamente con resultados ordinarios (ya que los extraordinarios pueden desvirtuar la evaluación de la gestión), la forma de cálculo propuesta a partir del BAIDI, no garantizaba totalmente la exclusión de las causas apuntadas al comienzo de este párrafo sobre las cuentas anuales, ya que no se elimina el efecto fiscal de los resultados extraordinarios. [1]

Por estos motivos, se optó por calcular el que denominamos EVA corregido (en adelante llamado EVAc) a partir del resultado ordinario o beneficio antes de impuestos (BAT) alcanzado por las empresas en cada año analizado, sumándole los gastos financieros y detrayéndole el valor contable del activo neto corregido por el coste medio ponderado del capital. El BAT se obtiene a partir de sumarle al resultado de las actividades ordinarias de las empresas, sus gastos financieros y restarle al término obtenido, el valor contable del activo corregido por el coste medio del pasivo[2]:

EVA corregido = BAT – (Valor contable activo * CMPC) [2]

Variables independientes

Las variables independientes del modelo de regresión debían ser indicadores representativos de los distintos stakeholders de las empresas; por ello, previo a su selección se identificaron los grupos de interés más relevantes. Además, estos indicadores no deberían estar directamente implícitos en el cálculo de la variable dependiente (EVA), aunque, a priori, si deberían guardar relación con la generación de valor en las organizaciones. Por otra parte, deberían ser datos fácilmente recopilables, bien a través de fuentes de información primaria o secundaria y, por tanto, datos no comprometidos para las empresas, de lo contrario sería difícil su obtención.

Teniendo en cuenta los aspectos indicados y tras la revisión bibliográfica efectuada se identificaron los siguientes grupos de interés para las empresas del IBEX-35: clientes, proveedores, empleados, accionistas y la sociedad, como un stakeholders transversal que, a su vez, conecta a la empresa con sus stakeholders y a éstos entre sí.

Una vez identificados los grupos de interés se revisaron las distintas propuestas de indicadores, que en los últimos años han ido proliferando como consecuencia del desarrollo de los modelos de actuación social de la empresa. La accountability o rendición de cuentas, como una parte de la responsabilidad social de las empresas ha dado lugar al desarrollo de distintas propuestas de memorias de responsabilidad social, en las que se recogen baterías de indicadores,

[1] A modo de ejemplo invitamos al lector a observar detenidamente lo sucedido con las provisiones y los resultados de Acciona en el ejercicio 2003 tras realizar esta última la venta a Vodafone de su participación en Airtel, con los resultados de BBV desde 1997 hasta 1999 y BBVA desde 2000 hasta 2002, y los ingresos o quebrantos extraordinarios de SCH en los años 2001 a 2003 (a. i.) o Telefónica en los años 2002 y 2003.

[2] El coste medio del capital se ha estimado como la media ponderada de los costes de la deuda y de los fondos propios. El coste de la deuda, a su vez, se ha estimado dividiendo los gastos financieros de los créditos a corto y largo plazo entre el principal abierto de los créditos a corto y largo plazo con entidades financieras. Cabe destacar que en el caso de las entidades financieras se han considerado como parte de sus préstamos a corto los depósitos de sus impositores, teniendo en cuenta su coste de retribución. El coste de los recursos propios se estimó mediante el cociente entre el importe total de los dividendos aprobados para cada ejercicio entre el valor de los recursos propios de la respectiva anualidad a final de año.

que han sido una fuente importante de información en esta investigación. Entre las propuestas de modelos de gestión de la responsabilidad social analizadas, Global Reporting Initiative (GRI), Accountability 1000 (AA1000), SA 8000, la propuesta del GRI ha sido la base a partir de la cual se seleccionaron los indicadores que constituyen las variables independientes del modelo.

Tal y como se indicaba antes, en la selección de indicadores se ha tenido en cuenta su vinculación con stakeholders, su relación con la variable explicada y la posibilidad de captación de los datos. Muchos de los indicadores planteados al comienzo de la investigación tuvieron que ser descartados ya que la información publicada por las empresas en las memorias, cuentas anuales o informes de sostenibilidad, era incompleta o inexistente. Otro tipo de indicadores, como por ejemplo, los gastos salariales totales o los intereses de la deuda se descartaron al estar implícitos en el cálculo del EVA. Otros datos, como las horas de formación se obtuvieron directamente de las empresas mediante llamadas telefónicas a los departamentos de marketing.

En la tabla 1 se muestran los indicadores de stakeholders que, finalmente, han sido empleados en el estudio como variables independientes. Hay que destacar que los indicadores seleccionados son, a su vez, representativos de aspectos intangibles de las organizaciones, así, por ejemplo, la retribución de los miembros de los consejos de administración es un indicador cuantitativo que indirectamente está valorando aspectos intangibles como la reputación, prestigio y capacidad de los consejeros. La inversión en I+D tiene una repercusión directa sobre un aspecto difícil de cuantificar como el aumento del número de patentes y su impacto sobre la organización; adviértase que las inversiones realizadas en I+D son un indicador transversal de todos los stakeholders.

Tabla 1. Indicadores de *stakeholders*

Clientes	
1.	Las ventas netas (en millones de €).
Proveedores	
2.	Las compras y aprovisionamientos netos (en millones de €).
Empleados	
3.	La formación, medida por el número de horas anuales.
4.	La promoción de los empleados (número de directivos que existen en cada año en las distintas organizaciones).
5.	El número de empleados.
Accionistas	
6.	El dividendo repartido anualmente a lo accionistas (€ por acción).
7.	El valor añadido para el accionista constituida dicha variable por la suma del valor económico reportado por el dividendo y la variación del valor de las acciones en el período de estudio (en € por acción).
8.	La retribución total del consejo de administración en cada período (en millones de €).
Sociedad	
9.	El valor añadido a la sociedad. Dicho término ha sido tomado y adaptado del utilizado por Bankinter en sus memorias anuales, y se obtiene de sumar las partidas económicas correspondientes a retribuciones sociales e impuestos.
Indicador Transversal	
10.	Las inversiones realizadas en I+D (representadas por los gastos de investigación y desarrollo realizados en cada anualidad) (indicador transversal en relación a diferentes *stakeholders*).

Fuente: elaboración propia.

La tabla 2 muestra los valores obtenidos por los principales estadísticos descriptivos para los indicadores de stakeholders considerados y el EVA. Los indicadores en los que existen mayores diferencias entre las empresas de la muestra son: el valor añadido al accionista, la inversión en

I+D y las horas de formación anuales, además del propio EVA. Por el contrario, el indicador en que existen menos diferencias entre las empresas del IBEX-35 es el dividendo. Como se puede deducir de los elevados valores de los coeficientes de variación, la heterogeneidad de la muestra, en lo que a las variables empleadas se refiere, es muy elevada, lo que, a su vez, muestra una gran variabilidad entre las empresas de la muestra. La existencia de un comportamiento heterogéneo entre las empresas del IBEX-35 nos llevó a estimar dos modelos distintos, incorporando en el segundo como variable explicativa la identificación de la propia organización, incluyendo, por tanto, tantas variables explicativas binarias como empresas tiene la muestra (menos una). Dichas variables se interpretarían como un efecto fijo, propio de cada empresa, que determinaría un valor financiero distinto en cada organización y tomarán el valor 1 si los datos corresponden a la empresa a la que identifican y cero en el resto.

Del análisis de correlaciones (tabla 2) podemos afirmar que todas las variables presentan una correlación lineal significativa con el EVA excepto dividendos, promoción y valor añadido al accionista, siendo el valor añadido a la sociedad la variable que presenta una relación lineal más fuerte. No obstante, incluiremos todas las variables explicativas en el modelo, con independencia de que mantengan o no una relación lineal con el EVA puesto que la red neuronal artificial permite captar las relaciones no lineales entre las variables.

Tabla 2. Principales estadísticos descriptivos y correlación con el EVA

Variable	Media	Desviación típica	Coef. Variación (%)	Correlación lineal con EVA
Ventas	7565.29	9775.39	129	0.520 *
Compras	4174.91	6301.41	151	0.439*
Dividendos	0.43	0.39	91	-0.035
I+D	798.01	2146.13	269	0.207*
Formación	777636.30	1507325.80	194	0.663*
Promoción	1244.74	1875.80	151	-0.055
Retribución Consejo	5.59	6.75	121	0.664*
Valor añadido sociedad	3099.24	4941.84	159	0.959*
Nº empleados	27583.69	39286.35	142	0.652*
Valor añadido al accionista	1.58	4.54	287	-0.131
EVA(1)	1364.61	3038.20	223	1

(1) Medido en millones de euros

* Significativo para un nivel de confianza del 95%

Fuente: elaboración propia.

MODELO DE RED NEURONAL UTILIZADO PARA LA ESTIMACIÓN DEL EVA

Una red neuronal es un modelo computacional que trata de simular el comportamiento humano en el proceso de aprendizaje. A diferencia de otras técnicas que persiguen este mismo objetivo, como la Inteligencia Artificial y los Sistemas Expertos, en la red el aprendizaje no ocurre fuera del sistema, sino dentro. La red neuronal artificial trata de imitar al cerebro humano realizando un procesamiento paralelo del conocimiento y no secuencial, como las técnicas anteriormente citadas (Zahedi, 1991). Se puede afirmar que una red neuronal está constituida por un conjunto de neuronas conectadas entre sí según una determinada estructura. La estructura de una red se compone de los siguientes elementos:

- ***Capas*:** se denomina capa al conjunto de neuronas que se encuentran en un mismo nivel de la red. La estructura más común de una red neuronal contiene una capa de entrada, una de salida y una o varias capas ocultas adicionales;
- ***Nudos*** (neuronas): cada capa contiene un número de nudos;
- ***Señales de entrada $x_i(n)$:*** Constituyen las variables independientes que le son suministradas a la red para la predicción de las variables dependientes. Así por ejemplo, en el

problema de estimación del EVA las variables de entrada serían los indicadores anteriormente descritos para cada empresa y año.

- ***Conexiones***: las conexiones entre neuronas muestran las relaciones existentes entre las mismas.
- ***Pesos***: cada conexión de un nudo i a un nudo j tiene un peso (w_{ij}), el valor de estos pesos oscila entre -1 y 1;
- ***Nivel de activación*** (v_j)**:** El nivel de activación de una neurona de salida, $v_j(n)$, para el patrón n-ésimo de entrada $x_i(n)$, es igual a la suma de las señales de entrada que recibió de la capa anterior, x_i, ponderada por los pesos, W_{ij}, correspondientes, como se muestra en [2].

$$v_j(n) = \sum_{i=0}^{p} Wij(n) \cdot x_i(n) \qquad [2]$$

(Siendo p el número de entradas aplicadas a la neurona de salida, excluido el umbral)

- ***Salida de un nudo, "yj":*** un nudo "dispara" o tiene una salida cuando el nivel de sus entradas sobrepasa un umbral T. Un nudo puede tener una salida continua si se aplica a la salida una ***función de activación o transferencia.*** Dicha función podrá ser lineal o no, de hecho una de las funciones de activación más utilizada en la práctica es la función no lineal sigmoide: $f(x)=(1+ e^{-x})^{-1}$.

En el presente estudio se utilizó un modelo de red neuronal simple feedforward en la que solo es posible la conexión entre las neuronas de una capa precedente con las de una capa posterior, nunca se produce proyección en sentido contrario (hacia atrás) ni entre neuronas de una misma capa.

Se utilizó *aprendizaje supervisado* para entrenar y enseñar a la red. En dicho método un grupo de datos, denominado *grupo de ejemplos de entrenamiento*, se utiliza para ayudar a la red en la determinación de los valores apropiados de sus pesos. Cada ejemplo está constituido por una señal de entrada con su respuesta deseada correspondiente, o *target.* El proceso de aprendizaje es el siguiente: se presenta a la red un ejemplo del grupo de entrenamiento, siendo los pesos synápticos de la misma fijados libremente al comienzo. Si se llama $\hat{y}_j(n)$ a la respuesta deseada para una determinada neurona j de la capa de salida en la iteración n; e $y_j(n)$ a la verdadera respuesta obtenida por la red, el error cometido $e_j(n)$, será igual a la diferencia entre el valor deseado y el valor real como se puede observar en [3].

$$e_j(n) = \left(y_j(n) - \hat{y}_j(n)\right) \qquad [3]$$

A cada entrada en la red $x_i(n)$ que estaría constituida por el valor de los indicadores de stakeholders de cada empresa en cada año analizado, le corresponde una respuesta o salidadeseada $d_j(n)$, que sería el valor previamente calculado del EVA correspondiente, constituyendo dicho par de valores un ejemplo particular presentado a la red en el instante n.

Durante el entrenamiento un grupo de n pares de vectores de entrada y su correspondiente salida es presentada a la red ($(X(1), y(1), X(2), y(2),\ldots,X(n), y(n))$. Después se calcula la suma de las entradas ponderadas por los pesos correspondientes y se le añade el valor del umbral como se puede observar en [4].

$$NET_t = \sum_{i=1}^{N} w_{ti}\, x_i + b_t \qquad [4]$$

Cada nodo oculto usa una función (sigmoide en este caso) para generar una salida Z_t.

$$Z_t = [1 + exp(-NET_t)]^{-1} = f(NET_t) \qquad [5]$$

Las salidas de cada nodo de la capa oculta son enviadas junto con el umbral b_0 hasta el nodo de la capa de salida, y una nueva suma ponderada es calculada, como en [6]. Esta suma ponderada se convierte en la entrada para la función de transferencia sigmoide del nodo de salida, de esta manera se determina el valor estimado por la res para el EVA según [7].

$$NET = \sum_{c=1}^{h} v_c Z_c + b_0 \qquad [6]$$

$$\hat{Y} = f(NET) = [1 + exp(-NET)]^{-1} \qquad [7]$$

En este momento comienza el proceso de aprendizaje, que consiste en ir modificando los pesos que fueron valorados aleatoriamente al comienzo del entrenamiento, para minimizar la diferencia entre las respuestas deseadas (Eva real) y las respuestas reales (Eva estimado) que la red ha producido elevadas al cuadrado (denominadas Suma de los Errores al Cuadrado "SSE" en [8]). El entrenamiento de la red es repetido para un número considerable de patrones, hasta que la red alcance un estado en el que los pesos no sufran cambios significativos.

$$SSE = \sum_{j=1}^{n} (y_j - \hat{Y}_j)^2 \qquad [8]$$

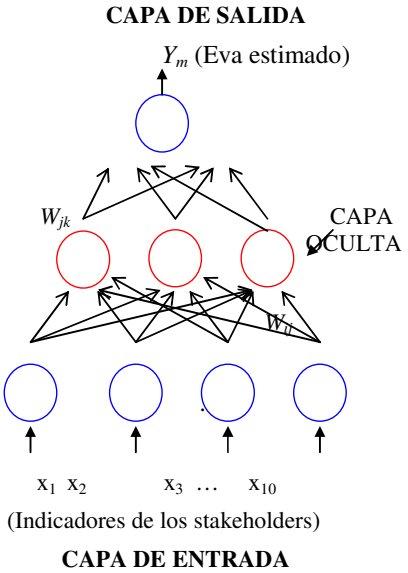
siendo $\hat{Y}_j$ la salida de la red para la j-ésima observación.

El algoritmo de retropropagación BP (Back propagation) descubierto inicialmente por Werbos (1974) y utilizado después por otros investigadores (Rumelhart *et a.l* 1986) es el método de aprendizaje supervisado más conocido y el utilizado en este trabajo.

La figura 1 representa el modelo de red neuronal aplicado en esta investigación, incluye una capa de entrada con 10 variables (correspondientes a los valores que toman los 10 indicadores de stakeholders, previamente descritos, que constituyen las variables exógenas del modelo)[3], una capa oculta cuyo número de nodos fue seleccionado experimentalmente, y una capa de salida con un nodo que representa el valor del Eva estimado por la red.

El número de nodos de la capa oculta se modificó entre un máximo igual al número de variables explicativas del modelo y un mínimo igual a 1. Los mejores resultados obtenidos en la estimación del EVA correspondieron a redes con 5 nodos en la capa oculta, con más nodos se producía sobre-entrenamiento y con menos disminuía la capacidad de aprendizaje de la red. Por lo tanto la red finalmente seleccionada tendría 5 nodos en la capa oculta (en la figura se representan 3 nodos ocultos para simplificar el gráfico).

Figura : Modelo de red neuronal utilizada en la estimación del EVA



Fuente: elaboración propia

3 Este número será mayor en el modelo que incluye variables binarias identificativas de cada empresa de la muestra.

Evaluando la capacidad predictiva y de generalización del modelo de red neuronal utilizado

La muesra total fue dividida en dos grupos: el grupo de entrenamiento, que incluye los datos empleados para estimar los parámetros de la red neuronal (es decir el valor de los pesos y umbrales) y el conjunto de teste, el cual incluye el resto de datos de la muestra que permite medir la capacidad de generalización del modelo. El número de casos incluidos en cada grupo variará en function del objetivo del análisis:

- Capacidad predictiva: si deseamos medir la *capacidad* predictiva de la RNA (es decir su capacidad para ofrecer buenas predicciones del EVA de las empresas) el conjunto de entrenamiento estará constituido por los datos corresponidnetes a todas las empresas durante el período 2000-2003 y el conjunto de teste incluirá los datos de las mismas empresas en el año 2004 . También se ha analizado la influencia de la cantidad de información disponible para el entrenamiento en los resultados, para ello hemos reducido el grupo de entrenamiento (datos de 2000-2002) y hemos incluido los datos del año 2003 en el conjunto de teste.

- Capacidad de generalización: si queremos medir la capacidad de generalización del modelo, es decir su capacidad para ofrecer buenas estimaciones del EVA para nuevas empresas que no han formado parte del proceso de entrenamiento, el conjunto de entrenamiento estará constituido por el 70% de las empresas de la muestra total y el 30% restante formará el grupo de teste. Como es muy probable que la selección de las empresas que formarán parte del grupo de teste afecte a los resultados se hicieron 10 particiones diferentes de la muestra original, dando lugar a 10 grupos distintos de entrenamiento y teste.

En ambos casos, se utilizaron dos parámetros para estimar la bondad del modelo de red neuronal propuesto:

La raíz del error cuadrático medio (RMSE):

$$RMSE = \sqrt{\sum_{i=1}^{n} (y_j - \hat{y}_j)^2 / n} \qquad [9]$$

El coeficiente de determinación R^2 calculado como el complementario del porcentaje que supone el Error Cuadrático Medio sobre el valor Medio Cuadrático Total.

RESULTADOS

Se han estimado dos modelos distintos, en el primero, denominado "modelo pool", sólo se contemplan como variables explicativas los distintos indicadores de stakeholder seleccionados, en el segundo, "modelo de efectos fijos", se ha tenido en cuenta que las particularidades de cada empresa tienen una incidencia directa y diferente a la del resto de empresa en la generación de valor, es decir, cada empresa tiene un efecto fijo propio en el EVA, por lo que en este modelo se incluyen tantas variables independientes binarias como empresas contiene la muestra, estas variables toman en valor 1 para la empresa en cuestión y cero en el resto de los casos.

La tabla 3 muestra los resultados obtenidos, para ambos modelos (Pool y Efectos Fijos), en los grupos de entrenamiento y teste con las distintas particiones efectuadas y las distintas metodologías empleadas en los modelos que incluyen todas las variables (9 explicativas, 1 dependiente y 28 variables binarias identificativas de cada empresa en el modelo de efectos fijos). Se ha resaltado en negrita los mejores resultados obtenidos en entrenamiento y teste.

Tabla 3. Capacidad de predicción

Diseño muestral	Modelo	Nº de variables independientes	$R^2_{Entrenamiento}$	RMSE Entrenamiento	R^2_{Teste}	RMSE Teste
Datos 2000-2003 para entrenamiento 2004 Teste	Pool	10	0.989	0.0254	0.980	0.0321
	EF	38	**0.998**	0.0106	**0.989**	0.0230
Datos 2000-2002 Entrenamiento 2003-2004 Teste	Pool	10	0.993	0.0316	**0.979**	0.0407
	EF	38	**0.997**	0.0112	0.934	0.0394

Como se puede observar los resultados son muy buenos y la incorporación de los Efectos Fijos (EF) no supuso cambios significativos en la capacidad predictiva.

Al disminuir el grupo de entrenamiento (datos 2000-2002) y aumentar el de predicción (2003 y 2004), comprobamos que en el modelo Pool, los resultados generados por la red neuronal fueron prácticamente iguales a los generados en la partición 1, es decir, no le afecta de forma significativa la reducción del grupo de entrenamiento y el aumento del de teste.

También se analizó la capacidad de generalización del modelo pool[4], los resultados se muestran en la tabla 4. En este caso mostramos los resultados medios obtenidos en entrenamiento y teste en las 10 particiones diferentes efectuadas sobre la muestra total para constituir los grupos de entrenamiento y teste.

Tabla 4. Modelo Pool: capacidad de generalización.

Nº variables explicativas	Modelo	Resultados medios en entrenamiento		Resultados medios en teste	
		$R^2_{entre.}$	RMSE $_{entre.}$	R^2_{Teste}	RMSE $_{Teste}$
10	Pool	0.9571 (0.092)	0.0387 (0.0031)	**0.7592** (0.061)	0.0608 (0.0030)

Como cabía esperar los resultados no son tan buenos como en los respectivos a la capacidad predictiva, lo cual es lógico pues en este caso estamos evaluando cómo se comporta la red ante datos nuevos, de empresas para los cuales no fue entrenada, por lo tanto la información de partida es considerablemente inferior. Aún así el valor del RMSE en generalización y el R^2 es bastante bueno, sobre todo si consideramos la alta variabilidad de la muestra.

CONCLUSIONES

El estudio empírico realizado ha puesto de manifiesto la relación entre el valor EVA generado por las empresas del IBEX-35 y un conjunto de indicadores representativos de los principales grupos de interés de las organizaciones, confirmando la primera hipótesis del trabajo. Los resultados obtenidos muestran que la generación de valor en las empresas no es

[4] En este caso no tiene sentido incluir los efectos fijos pues la muestra de generalización se compone de datos relativos a empresas que no han formado parte del grupo de entrenamiento, y por tanto son distintas.

incompatible con la satisfacción de las demandas de los distintos grupos de interés, por el contrario guarda relación con muchos de los indicadores representativos de dichos stakeholders.

El modelo de red neuronal propuesto para resolver el problema de predicción del EVA y de generalización del mismo para empresas diferentes obtiene muy buenos resultados. La capacidad predictiva es muy buena, con valores de R^2 en torno al 98% en los grupos de teste. La capacidad de generalización también es muy aceptable, con valores de R^2 próximos al 76% de media en las diferentes particiones de la muestra total.

Por lo tanto, podemos afirmar que las redes neuronales artificiales constituyen una herramienta eficaz para resolver problemas de predicción y de generalización en aplicaciones con datos de panel (de serie temporal y de corte transversal) en los que no se suelen cumplir las hipótesis de partida exigidas en las técnicas estadísticas paramétricas (no existencia de multicolinealidad entre las variables independientes, normalidad de las variables, no heterocedasticidad, etc.). Los resultados obtenidos muestran que el modelo de red neuronal artificial propuesto se puede utilizar para hacer predicciones y generalizaciones del valor de las empresas, medido éste por el EVA.

En futuros trabajos se estudiará la significación de variables mediante selección, aplicando algoritmos de poda. Además, se desea ampliar el estudio ampliando la muestra incorporando otras empresas diferentes y nuevos indicadores representativos de los grupos de interés de las organizaciones. Otra línea de trabajo es la comparación de los resultados obtenidos con otras técnicas clásicas de análisis, por ejemplo, los modelos de regresión con datos de panel, en los que ya se está trabajando.

BIBLIOGRAFÍA

A. KOLK & J. PINKSE (2006), "Stakeholder mismanagement and corporate social responsibility crises", European Management Journal, vol. 24(1), pp. 59-72.

ÁLVAREZ, J. (1999), "Priorización de los stakeholders en la actual sociedad del conocimiento", Técnica contable, vol. 51, nº 610, pp. 665-670.

AMAT, O. (1999), "EVA. Valor añadido económico. Un nuevo enfoque para optimizar la gestión, motivar y crear valor". Ediciones Gestión 2000, S.A. Barcelona.

AMBLER, T. Y WILSON, A. (1995), "Problems of Stakeholder Theory", Business Ethics: A European Review, vol. 4, nº 1, January, pp. 30-35.

ATKINSON, A., WATERHOUSE, J. Y WELLS R. (1997), "A stakeholder approach to strategic performance management", Sloan management review. Spring, pp. 25 – 37.

BOTH, R. (1997), "Economic value added as a management incentive", Management Accounting. Octubre, p. 48.

CARMONA, S., DEYÁ, B. Y GUTIÉRREZ, I. (2003), "El EVA como moda en control de gestión", Revista Española de Financiación y Contabilidad, vol. XXXII, núm. 118, pp. 725-753.

CASTAÑO, C. (2008) "Valoración de pequeñas empresas: una aplicación a la marca "denominación de origen Dehesa de Extremadura", Tesis doctoral. Edición electrónica gratuita. Texto completo en www.eumed.net/tesis/2008/ccg/.

COLLINS, E., KEARINS, K. & ROPER, J. (2006), "Sustainability and stakeholder engagement", Harvard Deusto Finanzas-Contabilidad, pp. 72.

COLLINS, E., KEARINS, K., & ROPER, J. (2005), "The risks in relying on stakeholder engagement for the achievement of sustainability", Electronic Journal of Radical Organization Theory, 9, 1, 1-19.

COPELAND, T. (1990), "Measuring and Managing the Value of Companies", John Wiley & Sons. New York.

FREEMAN, R. E. (1984), "Strategic Management: A stakeholder approach", Boston: Pitman Publishing.

FREEMAN, R. E. Y PHILIPS, R. A. (2002), "Stakeholders Theory: a libertarian defence", Business Ethics Quartely, vol. 12, núm. 3: 331-350.

HARRISON, J. S. Y FREEMAN, R. E. (1999), "Stakeholders, Social Responsibility, and Performance: Empirical Evidence and Theoretical Perspectives", Academy of Management Journal, vol. 42, nº5, pp. 470-485.

HART, S.L. Y MILSTEIN, M.B. (2003), "Creating sustainable value", Academy of Management Executive, núm. 17 (2), pp. 56-67.

HART, S.L. Y SHARMAN, S. (2004), "Engaging fringe stakeholders for competitive imagination", Academy of Management Executive, núm. 18 (1), pp. 7-18.

HOZ, C (2008) "Métodos más empleados para la valoración de empresas". Harvard Deusto Finanzas&Contabilidad, nº 82, Mar-Abr, pp 22-26.

JENSEN, M. C. (2001), "Value Maximization, Stakeholder Theory and the Corporate Objective Function", Journal of Applied Corporate Finance, vol. 14, nº 3, pp. 8-21.

JENSEN, M. C. (2002), "Value maximization, stakeholder theory, and the corporate objective function", Business Ethics Quarterly, vol. 12, nº 2, pp. 235-256.

LISBOA, P; VELLIDO, A (2000) "Business applications of neural networks. The State of the At of Real Word Applications, World Scientifics, 220 p.

LORCA, P. (2003), "La creación de valor en la empresa y los stakeholders". Harvard-Deusto – Finanzas & Contabilidad, nº 51, pp. 48-54.

LOZANO, J. M. (2005), "Towards the relational corporation: from managing stakeholder relationships to building stakeholder relationships (waiting for Copernicus)", Corporate Governance, vol. 5, nº 2, pp. 60-77. Emerald Group Publishing Limited.

PRAKASH, R; KUMAR, A (2008) "Forecasting Economic Growth Using an Artificial Neural Network Model" Journal of Financial Management and Analysis, Vol. 21, No. 1, January-June.

ROJO RAMÍREZ, A. Y GARCÍA PÉREZ DE LEMA, D. (2006), "La valoración de empresas en España: un estudio empírico", Revista Española de Financiación y Contabilidad, enero-marzo. núm. 132, 913-934.

Causal Effects of Alternative Participation Patterns: a labour policy application

Alcina Maria de Almeida Rodrigues Nunes, Professora Adjunta
alcina@ipb.pt
Instituto Politécnico de Bragança

Escola Superior de Tecnologia e Gestão de Bragança
Instituto Politécnico de Bragança
Campus de Sta Apolónia
Apartado 1134
5301-857 Bragança
Portugal
Phone: +351 273303108
Fax: 273 313051

Resumo
A política pública de emprego portuguesa oferece, ao conjunto de desempregados registados, não só a possibilidade de participar sucessivamente em diferentes programas do mercado de trabalho como, também, a possibilidade de adiar essa participação. Face à coexistência de padrões de participação alternativos, este trabalho utiliza uma metodologia dinâmica de *matching* para calcular e comparar o resultado gerado por cada uma das decisões de participação. Os resultados obtidos pela decisão de participar consecutivamente, em compração com a alternativa de adiar a participação, sugerem um maior grau de eficácia da política pública de intervenção no mercado de trabalho.

Palavras-chave: avaliação de políticas sociais, propensity score matching, políticas de trabalho

Abstract
The Portuguese public labour market policy offers, to the registered unemployed individuals, not only the possibility of successive participation in different labour programmes but also the possibility to postpone that participation. Dynamic matching techniques allow to understand the causal effects of the above alternative participation patterns, as this paper shows. The compute results of a consecutive participation decision compared with the alternative postponed participation suggest a higher degree of effectiveness of the public labour public policy.

Key words: evaluation of social programmes, propensity score matching, labour policy

Thematic Area: Quantitative Methods for Economics and Business

1. **INTRODUCTION**

A complete evaluation of public labour market measures should be, as much as possible, based on the empirical evaluation of the effects caused to the exposed population. The stimulus given to the public intervention in the European labour market followed, closely, by the increasing budget constrains turned crucial the need to develop more rigorous techniques of social programmes evaluation. Above all, techniques that could represent a major approximation to the diverse policy practices that, nowadays, compose the *menu* of labour market policies.

Heckman, Lalonde and Smith published, in 1999, a seminal paper discussing the genesis of programmes evaluation literature. Their discussion could be considered, by the light of the more recent developments, as relatively dated concerning the econometric methodology and the nature of the available data. Still, from the pioneer analysis of these authors had resulted two crucial conclusions - the inexistence of a single, and uncontested, programme evaluation method and the importance of using better data - that have cleared the way for the development and maturity of the actual econometric evaluation methodologies of social programmes. The present paper, aware of the Heckman and his co-authors lessons, follows what Kluve (2006) designates by *third generation* of labour policies microeconometric evaluation studies, characterized by the adoption - as the elected estimation technique - of a non-parametric matching estimation technique and tries to go even further by overcoming the limits of the static traditional evaluation models.

The reality of the public intervention in the European labour markets does not obey to close criteria of description. Several institutional frameworks, a vast range of distinct programmes and diverse target-populations demand specific adaptations of the econometric estimation techniques and of the available (non-experimental) data. For instance, outside the limits of theoretical models, it is an oversimplification to only assume that the actual choices have influence on future potential results. Expectations about future results also affect participation choices. In the particular case of the Portuguese institutional framework for the public intervention on the labour market it is possible to find dynamic selection problems related with endogenous elements not properly identify by static evaluation models. Indeed, an unemployed Portuguese individual can choose from a vast range of active labour market programmes. These programmes are available over time which allows the individual to choose the time of participation and/or the sequence of programme participation.

The latest theoretical evaluation literature developments do present more realistic models and identification assumptions, with a particular emphasis on dynamic models that can mimic the context of participation on a social programme more accurately. In particular, a dynamic approach provides a more realistic perspective of the public intervention on the labour market. It assumes that the individual faces not only a single participation decision but a sequence of participation choices over time since it is allowed the participation in different sequences of labour programmes. By developing new scientific evaluation tools the dynamic approach to programme evaluation is in principle better suited to explain the real impact of the labour market policies.

It is possible to report several developments concerning the introduction of dynamic elements in the traditional evaluation econometric models. To start, there are the contributions of Robins (1986, 1989, 1997), Gill and Robins (2001) and Abbring and Berg (2003). In the particular case of nonparametric models, where the causal effect is estimated by the

"artificial" construction of a comparison group through a matching methodology, Lechner and Miquel (2001) extended Robins' application to epidemiology and biostatistics to the analysis of participation on active labour market policies (Robin, 1986). The suggested extension of the Robins causal model presents explicitly the causal effects of dynamic sequences making use of potential results and allowing the introduction of intermediate results which determine subsequent sequences. Empirical applications of the sequential matching methodology suggested by Lechner and Miquel (2001) are scarce. One such an application is Lechner and Miquel (2001), or Lechner (2004, 2006), who present exploratory results obtained from the empirical applications of a propensity score matching estimator to the active public intervention on the German and Swiss labour markets, respectively. Other than these two studies, we are not aware of any other application of the estimator.

The application presented in this research study adjusts an administrative dataset, containing plenty individual information concerning the active public intervention on the Portuguese labour market, to the needs of microeconometric estimation. The performed adjustments show it is possible to make a quantitative analysis on the impact of an unemployed individual participation on a sequence of active measures offered by the unemployment offices. The present empirical study analysis the power of the dynamic nonparametric model as a fundamental tool of a labour market policy evaluation and allows, even if in an exploratory way, to withdraw important conclusions relating to the participation choices of unemployed individuals engaged in the Portuguese labour market public intervention.

Specifically, the study aims to evaluate if, an unemployed individual registered in time period t, who had decided to participate in a labour market programme, would have been better of continuing to participate in another available active programme, in the sequence of the first participation and consequent results, rather than to postpone his/her participation to the next period. In analysis are the potential answers for two essential questions for the register unemployed individual: (i) to participate in a labour programme, proposed by the public employment service, collecting a remuneration and keeping the link with the labour market; or, (ii) to postpone the decision to participate in a labour market programme which does not correspond to the achievement of a regular job until he/she finds that job or has the possibility to participate in another programme.

The paper is organized as follows. Section 2 presents the microeconometric framework for a sequential treatment evaluation process. Section 3, describes the Portuguese institutional context for the public labour market intervention, the empirical application strategy and the estimation results. Section 4 concludes.

2. SEQUENTIAL PROPENSITY SCORE MATCHING MODEL

In this paper a model which uses the simplest structure developed by Lechner and Miquel (2001) will be assumed. Even if the structure could be easily adapted to a bigger number of periods and to a more vast range of treatments three periods (years) and two potential treatments will be considered, to our interest. An initial moment in which each individual is in the same treatment state and posterior moments for which each individual could be in a different state of treatment will be considered. Specifically, one initial moment where all the individuals are in an unemployment state and two subsequent periods where those same individuals are engaged in a particular labour market programme will be considered. The

different T sequential time periods, after the initial period, will be present as t e t' and, consequently, $t,t' \in \{0,1,2\}$.

As referred above, in this particular research work, the model that is being develop will be applied directly since the scope of the investigation is the evaluation of a sequential participation in active labour market programmes over only two periods. Period 1 and period 2 which is assumed as a direct time sequence of period 1. Also as referred, in period 0 the individual is registered as unemployed and, by definition, does not participate in any programme. The previous assumption implies to assume that before period 1 each individual holds the same labour history concerning the participation in a labour market programme and therefore the causal effects are not conditioned by that labour history.

The complete treatment sequence an individual is subject to over three time periods is defined by a vector of Bernoulli random variables as $\underline{D_2} = (\underline{D_0}, D_1, D_2)$. $\underline{D_0}$ represents the history of participation up to time 0 and it will always take the value $\underline{d_0}$. This value is part of the conditioning set, since everything that will be estimated will be conditioned on it, but will not appear explicitly further in the notation, this is, $\underline{D_2} = (D_1, D_2)$. A particular implementation of D_t will be represented by two possible values, $d_t \in \{0,1\}$, denoting that in each particular time period only one programme will be available and so the individual will have only two possible choices – participate, or not, in the available treatment. To clarify the mathematical notation, the history of variables up to period t will be represented by a bar under the variable. For example, $\underline{d_2} = (d_1, d_2)$, where the small letters represent specific values of a random variable represented by a capital letter. The model defined by Lechner and Miquel (2001) admits the heterogeneity of treatments over time and so the potential outcomes are defined in terms of potential treatment states sequences. The difference among different treatment sequences will be represented by an index letter. $\underline{d_t^j}$, for example, identifies the sequence of treatment j until the period t.

In period 1 an individual can be observed in one of two mutually exclusive states, participation (1) or non-participation (0). In period 2 this same individual can be observed in one of four mutually exclusive sequences of states: $(1,1)$, $(1,0)$, $(0,0)$ or $(0,1)$. Therefore the potential outcomes on the second period will be subordinated to the treatment state in period 2 and on what had happened in period 1. Consequently 2 states defined by treatment status in period 1 and 4 states defined by treatment status in both period, 1 and 2, will be considered.

For each unemployed individual several potential outcomes are defined - one outcome for each treatment state sequence. They are called potential outcomes because, of course, all but one of these treatment effects with different time lengths (1 or 2 periods), are unobservable. For each individual it is only possible to follow one treatment sequence. The potential outcomes are measured at the end of each period, whereas the treatment status is measured at the beginning of those periods, and are indexed by the treatment sequences, $Y_t^{\underline{d_1}}(t \geq 1)$ or $Y_t^{\underline{d_2}}(t \geq 2)$. The observable outcome is denoted by Y_t.
To conclude the model's presentation, the attributes X_t, which are observable characteristics that could influence the participation selection process and/or the potential outcomes, must be considered. It is possible that a participation sequence, and consequently the outcomes of previous periods, may affect the values of these variables. The k-dimensional vector of

characteristics X_t may contain functions of Y_t and it is observable at the same time, this is, at the end of the time period[1].

In the dynamic model, as in the static one, the potential outcomes are used to define several average causal effects. For a specific time period t, a sequence of treatments defined up to period 1 or 2, (t,t') is compared to an alternative treatment sequence of the same or a different time length for a population defined by one of those sequences or even a third sequence. The possible average causal effects are defined by the following equation:

$$\theta_{\tilde{t}}^{d_{\underline{t}}^k, d_{\underline{t}'}^l}\left(d_{\tilde{\underline{t}}}^j\right) = E\left(Y_{\tilde{t}}^{d_{\underline{t}}^k} \middle| \underline{D_{\tilde{t}}} = d_{\tilde{\underline{t}}}^j\right) - E\left(Y_{\tilde{t}}^{d_{\underline{t}'}^l} \middle| \underline{D_{\tilde{t}}} = d_{\tilde{\underline{t}}}^j\right), \quad (1)$$
$$\tilde{t} \geq 0; \quad 1 \leq t,t' \leq 2 \text{ com } \tilde{t} \leq t',t; \quad k \neq l \text{ com } k,l \in (1,2); \quad j \in (0,1,2)$$

The first sequence, $\underline{d_t^k}$, defines the treated population; the second sequence, $\underline{d_{t'}^l}$, defines the nontreated population and the last sequence, $\underline{d_{\tilde{t}}^j}$, defines the population for which the causal effect will be computed. Whithin this framework, the observed result can the presented as follows:

$$\begin{aligned} Y_t &= D_1 Y_t^1 + (1-D_1) Y_t^0 \\ &= D_1 D_2 Y_t^{11} + D_1(1-D_2) Y_t^{10} + (1-D_1) D_2 Y_t^{01} + (1-D_1)(1-D_2) Y_t^{00}, \; t = 0,1,2 \end{aligned} \quad (2)$$

Table 1 helps to illustrate the dynamic model for sequential treatments as well as the entire mathematical notation presented in this section.

Table 1: Dynamic Model Notation and Time Line

Time Periods						
$t=0$		$t=1$		$t=2$		
Treatment Sequence	Attributes	Treatment Sequence	Attributes (Results)	Treatment Sequence	Results	Average Causal Effects
$D_0 = 0$	X_0	$D_1 = 0$	$X_1(Y_1)$	$D_2 = (0,0)$	Y_2^{00}	$\theta_2^{01,00}(.)$, $\theta_2^{10,00}(.)$, $\theta_2^{11,00}(.)$
				$D_2 = (0,1)$	Y_2^{01}	$\theta_2^{10,01}(.)$, $\theta_2^{11,01}(.)$
		$D_1 = 1$	$X_1(Y_1)$	$D_2 = (1,0)$	Y_2^{10}	$\theta_2^{11,10}(.)$
				$D_2 = (1,1)$	Y_2^{11}	

Source: Adapted from Miquel (2003).

[1] To simplify the presentation of the problem, we will assume that the variables denoted by X are not affected by the treatment states. Consequently we will explicitly add the observed outcomes in the conditioning set when the case of endogeneity (this is, when the attributes are influenced by the treatment status) is investigated. If some other variables (as the outcomes) are also influenced by the participation status all we have to do is to handle these endogenous variables in the say way as the outcomes in the conditioning set.

3. EMPIRICAL APPLICATION TO THE REGISTER UNEMPLOYEMENT

3.1. TREATMENT SEQUENCES AND INTEREST CAUSAL EFFECTS

The main concern of this research is the estimation of the participation causal effect, in successive time periods, for those individuals who are registered as unemployed in the statistical information system of the Portuguese public employment service.

As it was already discussed the adoption of a sequential dynamic model allows to deal with the endogeneity problem induced by successive labour market participations. In the particular case of the *Programas ocupacionais (POCs)* – the Public Employment Programmes – it is explicitly declared that they are programmes aiming to occupy unemployed individuals receiving an unemployment subsidy or in a situation of economical need as long as an opportunity of participation in another programme or an effective job vacancy does not appear. In this manner, in period 1, an unemployed individual participates, or not, in a *Programa Ocupacional* and the result he/she could achieve depends on the decision he/she made and of the potential result of that decision. In the second period, the same individual can participate again, or not, in one of the available active programmes offered by the unemployment office. Consequently, the participation decision, in period 2, depends not only on the observable characteristics, X_0, but also depends on the decision of participation in the first period and of the subsequent result, Y_1, giving rise to a problem of dynamic selection.

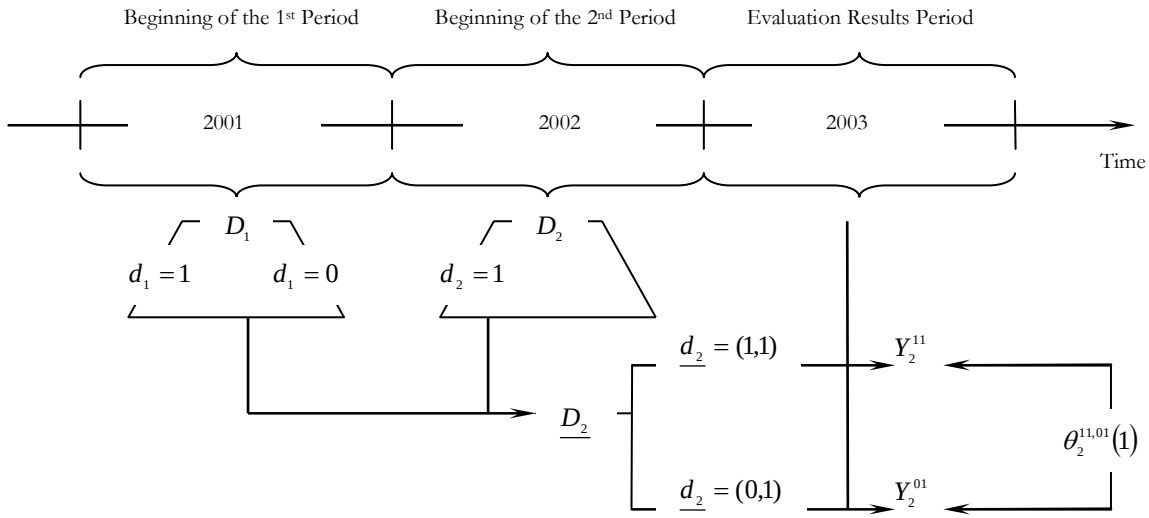
Each considered period corresponds to a year since this is the maximum extension for the participation in a POC[2]. After a year of participation the individual should, compulsorily, abandon the programme and change to another state of participation. Meanwhile if the unemployment register has not been abandoned the second state of participation could correspond to a state of participation in another active programme or to a state of non-participation in any of the interventions offered by the employment public service. The maintenance of the unemployment register and the consequent possibility of participation in another of the available active programmes for the unemployed depend, obviously, on the result obtained after the first participation. So the final impact evaluation of a sequence of participation states contains selection issues which demand a specific econometric management.

As the result of the former discussion, all the states of participation that had began over the year 2001 and were concluded over a spell of 12 months correspond to period 1: (i) participation in a POC or (ii) absence of participation in any of the active programmes available for the unemployed individual. All the states of participation that had began during the year 2002 correspond to period 2: (i) undistinguished participation in any of the available active programmes or (ii) absence of participation. Figure 1 allows to illustrate the sequences in analysis as well as the causal effects which are the object of the present evaluation process[3].

[2] It is also the maximum extension for the generality of the available active programmes offered by the public employment service.

[3] The dynamic approach is not particularly useful to some potential comparisons since their results are similar to the ones that are possible to reach with the static approach. Namely, the mentioned comparisons refer to treatment occuring only in one period and to comparisons among treatment sequences defined in two periods where the first period is equal for both sequences and for the target group. For instance, $\theta_2^{11,10}(1)$.

Figure 1: Empirical Evaluation – Time Sequences and Interest Causal Effects



With:

* $d_1 = 1$: Participation in *Programa Ocupacional* in 2001 (target population);
* $d_1 = 0$: Non-Participation in *Programa Ocupacional* in 2001 (participation decision postponed)
* $d_2 = 1$: Participation in 2002;
* $\underline{d_2} = (1,1)$: Participation in *Programa Ocupacional* in 2001 and participation in 2002;
* $\underline{d_2} = (0,1)$: Non-participation in *Programa Ocupacional* in 2001 and participation in a labour programme in 2002;
* $\theta_2^{11,01}(1)$: Causal effect of a consecutive participation compared to participation only in the second period, for the target population.

The sequence of intervention states is the result of the consecutive participation on the two selected time periods. That is, it is possible to define four distinct sequences: (i) if the unemployed individual registered in 2001 decides to participate in a POC in a first period, $(d_1 = 1)$, he/she must decide if in the next period he/she will participate again, or not, in another active programmes, which results in two different sequences, $\underline{d_2} = (1,1)$ or $\underline{d_2} = (1,0)$, respectively; (ii) if the individual decides to postpone the participation decision in the first period, $(d_1 = 0)$, in the second period he/she must decide if he/she will continue to postpone the participation or if he/she will decide to participate in one of the available active programmes, this is, $\underline{d_2} = (0,0)$ or $\underline{d_2} = (0,1)$, respectively.

This particular study aims to evaluate if, an unemployed individual registered in time period t (2001), who had decided to participate in a POC in a first period, would have been better of continuing to participate in another available active programme, in the sequence of the first participation and consequent results, rather than to postpone his/her participation to the second period. In analysis are the potential answers for two essential questions for the register unemployed individual: (i) to participate in the active programme proposed by the public employment service collecting a remuneration and keeping the link with the labour market; or, (ii) to postpone the decision to participate in a *Programa Ocupacional* which does not correspond to the achievement of a regular job until he/she finds that job or has the possibility to participate in another active programme. The interest results are measured by the maintenance of the register in the information system of the employment public service during the year 2003, the year where the conclusion of the second participation period is

observed. The perception of a larger maintenance of the unemployment register allows the conclusion of a minor effectiveness of the analysed sequences.

3.2. OBSERVABLE CHARACTERISTICS AND INTERMEDIARY RESULTS

Table 2 presents the distribution (in percentage) of the set of observable characteristics, previous to the participation sequences, as well as the intermediary and final results for each one of the selected participation sequences. The results are measured through the unemployment register in the statistical information system of the employment offices. The individual characteristics are observed in the beginning of the first treatment period and identify the individuals relating to their age, gender, formal education level, geographical location, labour market status (namely if the individual looks for a first job or if the individual has some labour market experience) and dependents.

Table 2: Pre-Treatment Characteristics and Results

Variables	**Treatment Sequences**			
	$d_1 = 1$	$d_1 = 0$	$\underline{d}_2 = (1,1)$	$\underline{d}_2 = (0,1)$
Number of Observations	3445	163855	895	15872
Pre-Treatment Characteristics (Values in %)				
Gender (Men)	21.45	38.90	17.21	28.26
Age (Years Average)	36.98	37.32	36.94	35.14
Educational level				
-None	9.26	6.42	9.72	6.51
- Primary (4 years)	37.13	33.83	39.89	31.0
- Compulsory Secondary (9 years)	35.21	34.99	33.63	37.97
- Secondary (12 years)	14.22	16.07	12.29	16.53
- Superior (>12 years)	4.18	8.69	4.47	7.98
Geographical Location				
- Norte	28.65	39.11	24.02	29.25
- Centro	20.32	12.07	18.55	17.6
- LVT	27.23	39.49	26.70	34.66
- Alentejo	19.45	5.99	25.36	12.44
- Algarve	4.35	3.33	5.36	6.05
Unemployment Category				
- First Employment	9.46	10.90	10.83	12.5
- New Employment	90.54	89.1	89.16	87.5
Dependents (Yes)	60.98	48.18	63.8	51.67
Intermediate Results (Values in %)				
Unemployment 6 months after the beginning of a POC	77.41	32.44	81.56	47.22
Unemployment 12 months after the beginning of a POC	62.41	31.21	74.86	59.07
Final Results (Values in %)				
Unemployment 30 months after the beginning of a POC	26.21	5.56	38.88	25.38

After the participation in a *Programa Ocupacional*, in the first period, the participation result can be measured through the maintenance (or not) of the unemployment register. Having the annual time spell been chosen as the participation period, the identification of the unemployment register 12 months after the beginning of the first participation period is a natural solution. However, since there are no abandonment registers it is believed important to

observe the unemployment registers some time before the end of the first analysed period and introduce them in the estimation process. In this specific case of this empirical application we chose go back 6 months, that is, half of the participation period initially predicted. The average value of unemployment registers two and a half years after the beginning of the first participation period was included in the descriptive table as an illustration and potential reference value for results discussion

When comparing the variable values among the different treatment sequences the high number of individuals who postpone the decision of participation to the second participation period should be stressed. The second aspect that should be stressed refers to the conclusion that women participate more in the programmes. Comparing the formal education level it is possible to observe that the unemployed individuals with more years of formal education are the ones who delay the first participation decision more frequently.

The variables appearing in the tables previously presented (with the exception of the values for the unemployment registers final results) represent the set of variables that was used to estimate the probabilities of participation, or propensity score, in each one of the treatment sequences relating to the target population – those who had participated in a POC during the year 2001. Finally, it must be reported that for the average values of the final result, even in a merely descriptive way, wide differences in the accounting of unemployment registers among the several treatment sequences are observed.

3.3. SEQUENTIAL PROPENSITY SCORE MATCHING ESTIMATION

The application of the estimator proposed by Lechner and Miquel (2001) follows the protocol suggested by Lechner (2004, 2006). The adoption of a propensity score matching estimator instead of just a matching estimator is due to the estimation simplicity offered by the first method. Conditioning the participation and the consequent result in a unique scalar value instead of conditioning in a wide set of variables simplifies the estimation process without losing the estimators properties (Rosenbaum e Rubin, 1983 e Lechner, 2004).

The probabilities of participation conditioned on the set of observable characteristics in the beginning of the first participation period (to sequences defined for only one period) and on the participation intermediary results (to sequences defined for the two periods) are presented in Table 3. The probabilities presented were estimated regarding the needs of each one of the matching sequences through binary probit models which should be (sequentially) subjected to specification tests. The parametric model is particularly simple and do not include a full set of variables since it could generate common support problems and therefore become counterproductive. In other words a complete set of variables could generate probabilities of participation so diverse that it would not be possible to find a set of values for which the individuals, in all the treatment sequences, had the same probability of participation. Since the matching process results from the selection of individuals with the same (or with a very similar) probability of participation the inexistence of a common support would not allow such individual selection.

The Mahalanobis distance metric technique with replacement of observations was applied to perform the selection of individuals in the matching process. The matching with replacement becomes more important with more diverse number of observations for each one of the defined sequences, as happens is this particular study.

Table 3: Probit Estimation

Variables	Treatment Sequences considering the participation in a POC in period 1 ($d_1 = 1$)		
	$d_1 = 0$	$\underline{d_2} = (1,1)$	$\underline{d_2} = (0,1)$
Pre-Treatment Characteristics (Values in %)			
Gender (Men)	0.293 (*) (0.0169)	-0.19 (**) (0.061)	-0.183 (*) (0.01)
Age	-0.067 (*) (0.005)	0.024 (0.018)	0.012 (*) (0.003)
Age2	0.001 (*) (0.000)	-0.000 (***) (0.000)	-0.000 (*) (0.000)
Educational level			
-None	-0.477 (*) (0.043)	(a)	0.156 (*) (0.026)
- Primary (4 years)	-0.352 (*) (0.037)	0.041 (0.086)	0.081 (*) (0.02)
- Compulsory Secondary (9 years)	-0.303 (*) (0.035)	-0.07 (0.091)	0.081 (*) (0.018)
- Secondary (12 years)	-0.275 (*) (0.037)	-0.156 (0.108)	0.061 (**) (0.02)
- Superior (>12 years)	(a)	0.048 (0.145)	(a)
Geographical Location			
- Norte	0.633 (*) (0.023)	-0.429 (*) (0.069)	-0.609 (*) (0.023)
- Centro	0.277 (*) (0.026)	-0.305 (*) (0.074)	-0.172 (*) (0.024)
- LVT	0.621 (*) (0.024)	-0.268 (*) (0.069)	-0.464 (*) (0.022)
- Alentejo	(a)	(a)	(a)
- Algarve	0.392 (*) (0.040)	-0.008 (0.121)	0.07 (**) (0.026)
Unemployment Category: First Employment	-0.026 (0.027)	0.182 (**) (0.085)	0.05 (**) (0.016)
Dependents (yes)	-0.068 (*) (0.017)	0.061 (0.054)	0.086 (*) (0.011)
Intermediate Results (Values in %)			
- Unemployment 6 months after the beginning of a POC	-----	-0.114 (0.071)	0.104 (*) (0.012)
- Unemployment 12 months after the beginning of a POC	-----	0.561 (*) (0.062)	0.816 (*) (0.012)
Constant	2.913 (*) (0.104)	-0.955 (**) (0.329)	-1.10 (*) (0.064)
Number of Observations	167300	3445	163855
$LR\chi^2(13) / LR\chi^2(15)$	1906.59 (*)	165.27 (*)	11816.28 (*)
Pseudo - R^2	0.0568	0.0419	0.1133
Log likelihood	- 15832.456	- 1890.8152	- 46221.031

Notes: (a) Reference Variable; (*) Statistical significance at 1%; (**) Statistical significant at 5%; (***) Statistical significant at 10%. Standard errors in parentheses.

The results of the model contain all the selection equations that are important to estimate the treatment causal effects defined in section 2. These results will not be discussed extensively however is important to refer the individual statistical significance of the generality of the estimated coefficients. It is also important to refer to the joint statistical significance of the coefficients included in the models. These statistical results point out to a good approximation to the definition of the unemployed individuals real probabilities of participation both in terms of individual characteristics and in terms of labour market characteristics.

The dynamic estimation results of the interest causal effects are presented in Table 4. In this table it is also possible to identify which sequences are being compared, the target population and the observations which were eliminated in each one of the propensity score matching sequences to generate a common support for the probability of participation.

When, for those who had participated in a POC, it is compared the consecutive participation in two active programmes – a *Programa Ocupacional* followed by another active programme not necessarily a POC – with a sequence characterized by a postponed participation the process of sequential propensity score matching is defined by a set of sequential steps which could be described as follows.

To obtain the counterfactual result for the participant population in a POC during the year 2001 it is necessary to select the participants in the sequence of non-participation who are the closest in terms of observable characteristics or, as happens in this specific case, in terms of the distribution of probability of participation conditioned in the set of observable characteristics. The counterfactual result corresponds to the result of the situation where that same population would have participated in a sequence characterized by the absence of some type of participation in the first period. As the target population, $d_1 = 1$, only participated once it is necessary, firstly, to select the most similar non-participant individuals, $d_1 = 0$, in terms of the participation probability in the first period of the sequence. Such selection process demands that the individuals in both groups share the same distribution of the participation probability which will be called $p_i^{d_1^0}$. In the propensity score matching methodology this first group of selected matched individuals would have constituted the comparison group if the selection process had been merely static and if it had happened in a single period[4].

Empirically, and for the estimation related to the entire population, the three steps explained above enable us to find within the group of non-participants in the first participation period the 3445 observations with the probability of participation most similar to the 3445 observations of the target population. The 3445 comparison observations do not mean the same number of distinct individuals since it was allowed the repeated use of observations whenever they were the ones which better fit in the propensity score matching process. Each time an observation was used more than once its weight increased by one for each application. In this particular case no observation of the target population was abandoned to construct the common support for the probability of participation.

However to estimate the result of the complete sequence it is necessary to select, from the group of comparison found in period 1, the group of individuals who have participated also in period 2. The result for the complete sequence arises from the result of the second comparison group which selection process corresponds to the following step.

[4] In order to guarantee the possibility of selection of the first comparison group, all the target population observations not comprised between the minimum and maximum limits of the participation probability of those individuals who form the potential comparison group were withdraw.

The step that follows consists in finding to the group of comparison individuals, achieved for the first period, those who, in the second period, will participate and have the same distribution of participation probability both for period 1, $\left(p_i^{d_1^0}\right)$, and period 2, $\left(p_i^{\underline{d_2^0},d_1^1}\right)$, than the individuals in the target group. This second comparison group will be the suitable comparison group. When the second comparison group was obtained two observations in the entire target population were lost. This was demanded by the need to observe the same probability of participation in both groups. The causal effect is estimated only for 3443 individuals.

With respect to the estimated values for the causal effects it is observed (Table 4) that, for the participants in a POC, exists one smaller probability to maintain the unemployment register if the programme participation is followed by the participation in other programme than if those same individuals had chose postpone participate in the labour programme offered by the public employment service. At the end of a sequence period it is estimated that for individuals who had participated in a *Programa Ocupacional*, in the initial period, the decision to participate again leads to a 4.3% lower probability of keep the unemployment register than the decision to wait and postpone the participation for a second period.

Table 4: Dynamic Causal Effects.

Sequences	**Target Population**	**Sample Size**		**Observations lost due to the construction of a common support**			$E\left(Y_t^{\underline{d_2^1}} \middle\| D_1 = d_1\right)$ (1)	$E\left(Y_t^{\underline{d_2^0}} \middle\| D_1 = d_1\right)$ (2)	$\theta_t^{\underline{d_2^1},\underline{d_2^0}}(d_1)$ (Values in %) (3)=(1)-(2)
$\underline{d_2^1}$ $\underline{d_2^0}$	d_1	$N_{\underline{d_2^1}}$ $N_{\underline{d_2^0}}$	N_{d_1}	t=1	t=2	% of the Target Population			
$\underline{d_2^1} = (1,1)$ $\underline{d_2^0} = (0,1)$	$d_1 = 1$	895 15872	3445	--- 0	2 0	0,058 0	0,283	0,326	**- 4,3**

4. CONCLUSION

The empirical application of a dynamic propensity score methodology demonstrates to be an effective tool for the estimation of dynamic causal effects where is tested the effectiveness of a sequence of participation decisions comparing to a distinct sequence of decisions. This dynamic propensity score methodology allows to introduce in the traditional static models intermediary participation results. These present a better characterization of the dynamic participation selection in the programmes offered by the public intervention in the Portuguese labour market.

The empirical application of a sequential propensity score matching methodology to evaluate the decision of a consecutive participation comparing to the option of postpone the participation decision one time period presents results favourable in terms of the effectiveness of consecutive public intervention in the labour market.

Still such results are short-run results obtained immediately after the second participation period. This time horizon limitation could, for example, hide locking-in effects- the programme participation withdraws to the unemployed individual time and availability to look for an effective job effects.

REFERENCES

Abbring, J.H. (2003), “Dynamic Econometric Program Evaluation”, IZA discussion paper, 804.

Abbring, J.H., van den Berg, G.J. (2003), “The nonparametric identification of treatment effects in duration models”, Econometrica, 71 (5), pp. 1491–1517.

Gill, R.D., Robins, J.M. (2001), “Causal inference for complex longitudinal data: the continuous case”, The Annals of Statistics, 29 (6), pp. 1785–1811.

Lechner, M. and Miquel, R. (2002), “Identification of effects of dynamic treatments by sequential conditional independence assumptions”, Discussion Paper, University of St. Gallen.

Lechner, M. (2004), “Sequential Matching Dynamic Causal Models”. IZA discussion paper, 1042.

Lechner, M. (2006), Matching Estimating of Dynamic Treatment Models: Some Practical Issues, Discussion Paper, University of St. Gallen.

Miquel, R. (2003), “Identification of Effects of Dynamic Treatments with a Difference-in Differences Approach”, Discussion Paper, University of St. Gallen.

Robins J. M. (1989), "The Analysis of Randomized and Nonrandomized AIDS Treatment Trials Using a New Approach to Causal Inference in Longitudinal Studies", in Sechrest, L., H. Freeman, A. Mulley (eds.), *Health Service Research Methodology: A Focus on Aids*, pp. 113-159, Washington, D.C., Public Health Service, National Center for Health Services Research.

Robins J. M. (1997), "Causal Inference from Complex Longitudinal Data. Latent Variable Modeling and Applications to Causality", in M. Berkane, (ed.), *Lecture Notes in Statistics* (120), New York: Springer, pp. 69-117.

Rosenbaum, P. R., e Rubin, D. B. (1983), “The Central Role of the Propensity Score in Observational Studies for Causal Effects”, Biometrika, 70(1), pp. 41-55.

Rubin, D.B., (1974), Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies, Journal of Educational Psychology, 66, pp. 688-701.

WELFARE COMPARISONS BASED ON A FAMILY OF LORENZ CURVES: AN EMPIRICAL APPLICATION

NOELIA GONZÁLEZ-PRIETO
noelia.gonzalez@unican.es
JOSÉ MARÍA SARABIA
MARTA PASCUAL
Department of Economics
University of Cantabria, Santander

Abstract

In this paper different welfare comparisons based on Lorenz curves are obtained. The relationship between income distribution and welfare is a key aspect not only for economists but also for policymakers. We proposed to use parametrical Lorenz curves to obtain rankings of welfare. Firstly, we analyse the properties of the parametric family including inequality measures such as Gini index. Obviously, Gini coefficients provide information about the changes in income inequality but do not reflect the welfare of the population. Secondly, we combine social welfare functions with these functions in order to obtain ordered utility functions. We present an empirical application based on the UPND Data. The results obtained are stable respect to the different functional forms we have proposed.

Key Words: Income Inequality, Welfare Comparisons, Lorenz Curves.
Thematic Area: Quantitative Methods for Economics and Business.

1.- INTRODUCTION

Actually, an important problem that worries economists is income inequality because there are several studies which show that from the'60s inequality has increased considerably. There are many ways to analyze inequality inherent in a given size distribution of income. The Lorenz curve is one of the most popular and intuitive methods. The Lorenz curve has been studied for a long time as a statistical instrument to measure inequality. In this paper we make a revision of different Parametric Lorenz Curves and its properties. We use them to calculate Gini index to study income inequality in different countries. The structure of the paper is as follows. Section 2 introduces the basic results, which will be used along the following paper. In Section 3, the parametric Lorenz curves will be presented. Finally, Section 4 presents the results obtained.

2.- PREVIOUS RESULTS

In this paper we consider the Lorenz curve as defined by Gastwirth (1971). Let be the class of all non-negative random variables with positive finite expectations. For a random variable X in L with cumulative distribution function F_X we define its inverse distribution function F_X^{-1} by:

$$F_X^{-1}(y) = \inf\{x : F_X(x) > y\}$$

Thus, the Lorenz curve associated with X is defined by,

$$L_X(p) = \left[\int_0^p F_X^{-1}(y)dy\right] \Big/ \left[\int_0^1 F_X^{-1}(y)dy\right],\ 0 \le p \le 1.$$

Note that $\mu_X = \int_0^1 F_X^{-1}(y)dy$ is the expectation of the random variable X .

A characterization of a Lorenz curve attributed to Gaffney and Anstis by Pakes (1994) is given by the following theorem.

Theorem 1-. Suppose that $L(p)$ is defined and continuous on $[0,1]$ with second derivative $L''(p)$. The function $L(p)$ is a Lorenz curve if and only if

$L(0) = 0$, $L(1) = 1$, $L'(0^+) \geq 0$, $L''(p) \geq 0$ in $(0,1)$.

Note that $L'(0^+)$ means the value of first derivative of function $L(p)$ for $p = 0$ to the right. The next two theorems are due to Sarabia, Castillo and Slottje (1999).

Theorem 2.- Let $L(p)$ be a Lorenz curve and consider the transformation $L_\alpha(p) = p^\alpha L(p)$, $\alpha \geq 0$. Then if $\alpha \geq 1$, $L_\alpha(p)$ is a Lorenz curve too. In addition, if $0 \leq \alpha < 1$ and $L'''(p) \geq 0$, $L_\alpha(p)$ is also a Lorenz curve.

Theorem 3.- If $L(p)$ is a Lorenz curve,

$L_\gamma(p) = L(p)^\gamma$, $\gamma \geq 1$

is also a Lorenz curve.

The above theorems suggest a method for obtaining hierarchical families of Lorenz curves. Towards this aim we start an initial Lorenz curve $L_0(p)$ and consider the following parametric hierarchy:

$L_1(p;\alpha) = p^\alpha L_0(p)$, $(\alpha \geq 1)$ or $\left(0 \leq \alpha < 1, L_0'''(p) \geq 0\right)$,

$L_2(p;\gamma) = L_0(p)^\gamma$, $\gamma \geq 1$,

$L_3(p;\alpha,\gamma) = p^\alpha L_0(p)^\gamma$, $(\alpha,\gamma \geq 1)$ or $\left(0 \leq \alpha < 1, \gamma \geq 1, L_0'''(p) \geq 0\right)$.

On the other hand, it is very interesting to study the relationship between Lorenz curves and Social Welfare Functions (SWF). A social welfare function can be defined as a function which gives a measure of the total welfare of society, given a number of economic variables as inputs. In 1970, Atkinson established the conditions under which s Lorenz inequality comparison has normative significance in terms of utility functions.

Theorem 4.- (Atkinson, 1970)

Let $F(x)$, $G(x)$ be two income distributions with equal means $\mu_F = \mu_G$. Then

$L_F(p) \geq L_G(p)$ for all $p \in [0,1] \Leftrightarrow \int U(x)f(x)dx \geq \int U(x)g(x)dx$,

for every utility function $U(x)$ such that $U'(x) > 0$ and $U''(x) < 0$.

Corollary (Atkinson, 1970)

Let $F(x)$, $G(x)$ be two income distributions. Then $\mu_F > \mu_G$ and $L_F(p) \geq L_G(p)$ for all $p \in [0,1] \Rightarrow \int U(x)f(x)dx \geq \int U(x)g(x)dx$,

for every utility function U(x) such that $U'(x) > 0$ and $U''(x) < 0$.

So if the following conditions are satisfied:

i. The Lorenz Curve of distribution Y dominates the Lorenz Curve of distribution *X*.

ii. The distributions have an equal mean income or the *Y* mean is greater than the *X* mean income

iii. The decision-maker is income seeking and inequality-averse (that is, the Social Welfare Function – SWF- has positive first derivative and negative second derivative with respect to individual incomes).

then, social welfare is higher in *Y* than in *X*.

3.- PARAMETRIC LORENZ CURVES.

The Lorenz curve is defined by points $(p, L(p))$, where p represents the cumulative proportion of income-receiving units, and $L(p)$ the cumulative proportion of incomes, when the incomes are arranged in ascending order of magnitude.

In the discrete case, if $x_{1:n} \le x_{2:n} \le ... \le x_{n:n}$ denote the ordered individual incomes in the population then:

$$L(\frac{i}{n}) = \frac{\sum_{j=1}^{i} x_{j:n}}{\sum_{j=1}^{n} x_{j:n}}. \qquad (3.1)$$

The points $(\frac{i}{n}, L(\frac{i}{n}))$ are then linearly interpolated to complete the corresponding Lorenz curve.

On the other hand, in the continuous case, if n is large the distribution of incomes can be approximated by a continuous distribution function $F(x)$, with density $f(x)$ related by $F(x) = \int_0^x f(y)dy$. For each positive x, $F(x)$ represents or approximates the proportion of individuals in the population whose income is less than or equal to x.

In this paper we are going to consider Pareto Lorenz curves which have been obtained based on the previous results. Thus, these are the curves analysed:

$$L_0(p) = 1 - (1-p)^{C_1}$$

$$L_1(p)1 = 1 - \left(1 - p^{C_1}\right)^{C_2}$$

If two distribution functions have associated Lorenz curves which do not intersect, the can be ordered without ambiguity in terms of welfare functions which are symmetric, increasing and quasiconcave (Atkindon, 1970; Dasgupta et al., 1973; Shorrocks, 1983).

Definition: Let X and Y be random variables belonging to *L* class. The Lorenz order $\le_L$ on *L* is defined by,

$$X \le_L Y \Leftrightarrow L_X(p) \ge L_Y(p), \ \forall p \in [0,1]$$

If $X \le_L Y$, then X exhibits less inequality than Y in the Lorenz sense. Note that Lorenz order is a partial order and is invariant with respect to scale transformation.

As a special case, let $X_i \approx Pa(\alpha_i, \sigma_i)$ be Pareto distribution with cdf $F_X(x) = 1 - \left(\frac{x}{\sigma_i}\right)^{\alpha_i}$, $x \ge \sigma_i, i = 1,2$, where $\sigma_i > 0$ is a scale parameter and $\alpha_i > 0$ a shape parameter. Using 3.6 we obtain:

$$L_X(p) = \frac{\alpha_i - 1}{\alpha_i \sigma_i} \int_0^p \sigma_i (1-y)^{-1/\alpha_i} \, dy = 1 - (1-p)^{1-1/\alpha_i}, 0 < y < 1, i = 1, 2$$

Then, $X_1 \leq_L X_2 \Leftrightarrow \alpha_1 \geq \alpha_2$.

3.1.- Inequality measures derived from Lorenz Curves.

The two more popular measures of inequality which are directly related to the Lorenz curve are the Gini and Pietra indices.

The Gini index is defined as twice the area between the egalitarian line and the Lorenz curve. The Gini coefficient is defined as follows (Gini, 1912):

$$G_X = 2\int_0^1 [p - L_X(p)]dp = 1 - 2\int_0^1 L_X(p)dp$$

The Gini coefficient takes on values between 0 and 1 with zero interpreted as no inequality.There are several alternative expressions of the Gini index equivalent to the equation above. One of them is:

$$G_X = 1 - \frac{E(X_{1:2})}{\mu} = 1 - \frac{1}{\mu}\int_0^\infty [1 - F_X(x)]^2 dx$$

where $X_{1:2}$ is the smaller of a sample of size 2 coming from the cdf F_X. This expression is useful when we have a closed form for the cumulative distribution function (e.g. Cronin, 1979).

The Pietra index is defined as the maximal vertical deviation between the Lorenz curve and the egalitarian line:

$$P_X = \max_{0 \leq p \leq 1} \{p - L_X(p)\}$$

In this paper, we will based our results on Gini index.

4.- EMPIRICAL RESULTS

The data used were obtained from the "Human Development Report 2004, published for the United Nations Development Program (UNDP), (Oxford University Press 2004). The Office of Human Development Report is first user, not producer, of statistics. Therefore, based on information provided to international statistical agencies have the resources and expertise needed to collect and process data on international statistical indicators specific. The Human Development Report uses to the extent possible, data that can be compared internationally collected by major international agencies and other specialized institutions. However, there are still many gaps in the data, even in core areas related to human development. At the same time calling for the data on human development, the Office of Human Development Report in principle and practical reasons does not collect information directly from the countries or make arbitrary estimates to fill those gaps.

The data correspond to cumulate income for 127 countries. These countries are divided in three groups: developing countries of high (41 countries), medium development (60 countries) and developing countries under (26 countries).

In Tables 1-6 we present the results of the estimation of the Lorenz curves for the different countries which are classified according to the level of development: high, medium and low. These tables include the estimators of the models together with the sum squared error (SCE) and Gini indices. It can be concluded that the both functions, though simple and easy to estimate, are very satisfactory in fitting data.

Table 1.

High development Country	C	SCE(%)	Gini
Argentina	0.38	0.46	0.4524
Australia	0.55	0.53	0.2876
Austria	0.59	0.31	0.2549
Belgium	0.61	0.34	0.2422
Canada	0.57	0.39	0.2776
Chile	0.31	0.22	0.5274
Costa Rica	0.43	0.41	0.4016
Croatia	0.58	0.24	0.2667
Denmark	0.63	0.43	0.2256
Estonia	0.51	0.38	0.3208
Finland	0.62	0.18	0.2346
France	0.57	0.35	0.2763
Germany	0.62	0.33	0.2361
Greece	0.52	0.24	0.3160
Hong Kong	0.43	0.26	0.3961
Hungary	0.61	0.41	0.2460
Ireland	0.53	0.27	0.3092
Israel	0.52	0.30	0.3191
Italy	0.54	0.38	0.2984
Japan	0.64	0.12	0.2201
Korea	0.61	0.39	0.2435
Latvia	0.56	0.28	0.2807
Lithuania	0.57	0.27	0.2723
Luxembourg	0.59	0.26	0.2581
Mexico	0.34	0.32	0.4901
Netherlands	0.57	0.34	0.2756
New Zealand	0.52	0.37	0.3149
Norway	0.61	0.17	0.2425
Poland	0.53	0.25	0.3029
Portugal	0.49	0.37	0.3395
R. Czech	0.63	0.14	0.2267
Singapore	0.45	0.35	0.3752
Slovakia	0.65	0.37	0.2155
Slovenia	0.64	0.28	0.2229
Spain	0.57	0.32	0.2770
Sweden	0.62	0.25	0.2331
Switzerland	0.56	0.39	0.2782
Trinidad Tobago	0.49	0.38	0.3405
UK	0.51	0.37	0.3207
Uruguay	0.44	0.37	0.3856
USA	0.49	0.41	0.3402

Table 2.

High development Country	C(1)	C(2)	SCE(%)	Gini
Argentina	1.89	0.54	0.0074	0.5253
Australia	1.57	0.73	0.0001	0.3529
Austria	1.37	0.72	0.0005	0.3010
Belgium	1.38	0.74	0.0006	0.2898
Canada	1.45	0.71	0.0003	0.3316
Chile	1.62	0.40	0.0041	0.5748
Costa Rica	1.66	0.57	0.0029	0.4654
Croatia	1.33	0.69	0.0016	0.3074
Denmark	1.42	0.79	0.0023	0.2778
Estonia	1.49	0.65	0.0005	0.3764
Finland	1.26	0.71	0.0000	0.2680
France	1.42	0.70	0.0002	0.3268
Germany	1.37	0.75	0.0004	0.2823
Greece	1.36	0.62	0.0005	0.3584
Hong Kong	1.47	0.54	0.0023	0.4441
Hungary	1.43	0.75	0.0007	0.2990
Ireland	1.39	0.64	0.0030	0.3549
Israel	1.42	0.63	0.0056	0.3673
Italy	1.47	0.68	0.0001	0.3526
Japan	1.19	0.71	0.0010	0.2458
Korea	1.41	0.75	0.0002	0.2947
Latvia	1.36	0.68	0.0009	0.3250
Lithuania	1.35	0.69	0.0009	0.3160
Luxembourg	1.33	0.70	0.0020	0.2995
Mexico	1.74	0.47	0.0020	0.5489
Netherlands	1.41	0.70	0.0001	0.3251
New Zealand	1.47	0.65	0.0022	0.3691
Norway	1.24	0.70	0.0004	0.2742
Poland	1.36	0.64	0.0004	0.3464
Portugal	1.51	0.63	0.0018	0.3952
R. Czech	1.21	0.71	0.0005	0.2546
Singapore	1.55	0.58	0.0021	0.4319
Slovakia	1.37	0.79	0.0019	0.2626
Slovenia	1.33	0.76	0.0000	0.2645
Spain	1.39	0.69	0.0002	0.3248
Sweden	1.31	0.73	0.0001	0.2728
Switzerland	1.45	0.71	0.0001	0.3318
Trinidad Tobago	1.53	0.63	0.0011	0.3979
UK	1.48	0.65	0.0001	0.3753
Uruguay	1.58	0.58	0.0047	0.4441
USA	1.54	0.64	0.0002	0.3992

Table 3.

Medium development			
Country	C	SCE%	Gini
Albania	0.84	0.0000	0.0883
Algeria	0.54	0.3095	0.3006
Armenia	0.50	0.2512	0.3325
Azerbaidzhán	0.51	0.1797	0.3263
Bangladesh	0.55	0.1162	0.2880
Bielorussia	0.59	0.2448	0.2610
Bolivia	0.46	0.5387	0.3731
Bosnia-Herzegovina	0.64	0.2356	0.2223
Botswana	0.23	0.1719	0.6221
Brazil	0.30	0.4580	0.5381
Bulgaria	0.59	0.4877	0.2618
Cambodia	0.46	0.1228	0.3698
Cameroon	0.43	0.2091	0.3994
China	0.45	0.3939	0.3829
Colombia	0.31	0.2881	0.5233
Dominican Republic	0.40	0.2015	0.4273
Ecuador	0.36	0.3517	0.4758
Egypt	0.52	0.0867	0.3187
El Salvador	0.36	0.4131	0.4660
Georgia	0.52	0.3528	0.3142
Ghana	0.49	0.3803	0.3448
Guatemala	0.29	0.2824	0.5449
Guyana	0.44	0.4032	0.3859
Honduras	0.35	0.4038	0.4845
India	0.55	0.1056	0.2940
Islamic Republic of Iran	0.44	0.3148	0.3850
Republik Indonesia	0.52	0.1217	0.3116
Jamaica	0.49	0.2334	0.3410
Jordan	0.51	0.1497	0.3271
Kazakhstan	0.58	0.2641	0.2651
Kyrgyzstan	0.60	0.2009	0.2498
Macedonia	0.62	0.3317	0.2349
Malaysia	0.39	0.2712	0.4366
Moldova	0.52	0.2467	0.3161
Mongolia	0.42	0.1726	0.4103
Morocco	0.48	0.2374	0.3483
Namibia	0.17	0.2339	0.7059
Nepal	0.50	0.1559	0.3294
Nicaragua	0.33	0.2112	0.5032
Panama	0.33	0.4327	0.4981
Papua New Guinea	0.37	0.2277	0.4600
Paraguay	0.33	0.4370	0.4996
Peru	0.40	0.5074	0.4265
Philippines	0.42	0.2188	0.4124
Laos	0.50	0.1355	0.3352
Romania	0.59	0.2890	0.2545
Russian Federation	0.42	0.2696	0.4063
Saint Lucia	0.46	0.3330	0.3693
South African Rep	0.03	0.0000	0.9429
Sri Lanka	0.53	0.1599	0.3065
Swaziland	0.28	0.2005	0.5570
Tajikistan	0.57	0.2541	0.2739
Thai	0.44	0.2124	0.3845
Tunisia	0.47	0.2765	0.3566
Turkey	0.48	0.2936	0.3485
Turkmenistan	0.47	0.2488	0.3585
Ukraine	0.60	0.2344	0.2469
Uzbekistan	0.63	0.2526	0.2299
Venezuela	0.41	0.5560	0.4224
Vietnam	0.51	0.1140	0.3273

Table 4.

Medium development Country	C(1)	C(2)	SCE%	Gini
Albania	1.00	0.84	0.0000	0.0883
Algeria	1.41	0.66	0.0032	0.3492
Armenia	1.39	0.61	0.0010	0.3770
Azerbaidzhán	1.31	0.59	0.0009	0.3630
Bangladesh	1.22	0.62	0.0012	0.3157
Bielorussia	1.32	0.69	0.0015	0.3016
Bolivia	1.75	0.64	0.0020	0.4466
Bosnia-Herzegovina	1.29	0.74	0.0004	0.2600
Botswana	1.77	0.32	0.0030	0.6655
Brazil	2.32	0.49	0.0105	0.6181
Bulgaria	1.50	0.75	0.0001	0.3218
Cambodia	1.27	0.53	0.0018	0.4001
Cameroon	1.41	0.52	0.0025	0.4418
China	1.61	0.59	0.0068	0.4442
Colombia	1.76	0.43	0.0005	0.5789
Dominican Republic	1.43	0.49	0.0025	0.4698
Ecuador	1.75	0.49	0.0035	0.5372
Egypt	1.20	0.57	0.0000	0.3431
El Salvador	1.83	0.52	0.0010	0.5332
Georgia	1.46	0.65	0.0008	0.3673
Ghana	1.53	0.62	0.0056	0.4024
Guatemala	1.84	0.41	0.0059	0.6020
Guyana	1.61	0.58	0.0002	0.4469
Honduras	1.88	0.50	0.0028	0.5524
India	1.21	0.61	0.0002	0.3206
Islamic Republic of Iran	1.52	0.56	0.0036	0.4384
Republik Indonesia	1.24	0.59	0.0015	0.3406
Jamaica	1.38	0.59	0.0034	0.3839
Jordan	1.28	0.58	0.0005	0.3601
Kazakhstan	1.34	0.69	0.0035	0.3076
Kyrgyzstan	1.28	0.70	0.0032	0.2855
Macedonia	1.37	0.75	0.0001	0.2813
Malaysia	1.54	0.50	0.0030	0.4876
Moldova	1.37	0.62	0.0009	0.3594
Mongolia	1.36	0.50	0.0011	0.4481
Morocco	1.39	0.58	0.0030	0.3919
Namibia	2.89	0.33	0.0147	0.7637
Nepal	1.28	0.58	0.0014	0.3634
Nicaragua	1.56	0.42	0.0002	0.5487
Panama	2.00	0.50	0.0036	0.5703
Papua New Guinea	1.51	0.47	0.0081	0.5065
Paraguay	1.99	0.50	0.0006	0.5710
Peru	1.82	0.57	0.0021	0.4979
Philippines	1.43	0.51	0.0071	0.4563
Laos	1.27	0.57	0.0001	0.3667
Romania	1.35	0.71	0.0003	0.2986
Russian Federation	1.49	0.53	0.0000	0.4554
Saint Lucia	1.51	0.59	0.0008	0.4234
South African Rep	2.99	0.57	0.0296	0.6489
Sri Lanka	1.28	0.61	0.0011	0.3401
Swaziland	1.65	0.37	0.0004	0.6025
Tajikistan	1.34	0.68	0.0002	0.3159
Thai	1.39	0.54	0.0101	0.4263
Tunisia	1.44	0.58	0.0021	0.4047
Turkey	1.45	0.60	0.0032	0.3979
Turkmenistan	1.41	0.57	0.0039	0.4039
Ukraine	1.30	0.71	0.0006	0.2857
Uzbekistan	1.31	0.74	0.0001	0.2693
Venezuela	1.90	0.59	0.0008	0.4999
Vietnam	1.24	0.57	0.0011	0.3558

Table 5.

Low development			
Country	C	SCE%	Gini
Burkina Faso	0.32	0.11	0.5143
Burundi	0.46	0.34	0.3697
Central African Republic	0.29	0.41	0.5466
Ethiopia	0.58	0.14	0.2694
Gambia	0.39	0.36	0.4397
Guinea	0.47	0.21	0.3581
Guinea Bissau	0.39	0.15	0.4358
Ivory Coast	0.42	0.21	0.4035
Kenya	0.42	0.19	0.4050
Lesotho	0.28	0.51	0.5579
Madagascar	0.41	0.29	0.4214
Malawi	0.36	0.13	0.4670
Mali	0.37	0.21	0.4576
Mauritania	0.50	0.31	0.3356
Mozambique	0.48	0.21	0.3524
Niger	0.41	0.64	0.4171
Nigeria	0.37	0.21	0.4577
Pakistan	0.53	0.10	0.3038
Rwanda	0.59	0.13	0.2585
Senegal	0.46	0.17	0.3723
Sierra Leone	0.32	0.73	0.5166
Tanzania	0.50	0.22	0.3368
Uganda	0.44	0.19	0.3896
Yemen	0.55	0.29	0.2863
Zambia	0.37	0.33	0.4651
Republic of Zimbabwe	0.38	0.21	0.4544

Table 6.

Low development				
Country	C(1)	C(2)	SCE%	Gini
Burkina Faso	1.38	0.38	0.0038	0.5464
Burundi	1.52	0.59	0.0007	0.4241
Central African Republic	2.25	0.47	0.0106	0.6222
Ethiopia	1.23	0.65	0.0004	0.2991
Gambia	1.68	0.52	0.0133	0.5009
Guinea	1.37	0.57	0.0008	0.3994
Guinea Bissau	1.36	0.47	0.0008	0.4720
Ivory Coast	1.41	0.51	0.0008	0.4456
Kenya	1.39	0.51	0.0009	0.4453
Lesotho	2.67	0.51	0.0091	0.6456
Madagascar	1.53	0.52	0.0157	0.4752
Malawi	1.36	0.43	0.0003	0.5008
Mali	1.49	0.46	0.0066	0.5024
Mauritania	1.45	0.62	0.0052	0.3864
Mozambique	1.36	0.57	0.0000	0.3934
Niger	2.03	0.63	0.0017	0.5039
Nigeria	1.48	0.46	0.0001	0.5019
Pakistan	1.21	0.59	0.0001	0.3296
Rwanda	1.21	0.66	0.0042	0.2863
Senegal	1.34	0.54	0.0001	0.4095
Sierra Leone	2.93	0.62	0.0051	0.6224
Tanzania	1.36	0.59	0.0018	0.3785
Uganda	1.37	0.52	0.0000	0.4293
Yemen	1.38	0.67	0.0012	0.3325
Zambia	1.68	0.49	0.0001	0.5231
Republic of Zimbabwe	1.47	0.46	0.0028	0.4983

ACKNOWLEDGMENTS

We thank the Ministerio de Ciencia y Tecnología, project SEJ2007-65818/ECON, for partial support of this work.

5.- REFERENCES

ATKINSON, A. B. (1970) "On the Measurement of Inequality", Journal of Economic Theory, Volume 2, September pp 244-263.

COWWLL F.A. (1977) "Measuremente of Inequality", Handbook of Income Distribution, Volume 1.

CREEDY, J. (1985) "Dynamics of Income Distribution", Basil Blackwell, Oxford.

GASTWIRTH, J.L. (1971) "A General Definition of Lorenz Curve", Econometrica, Volume 39, November pp 1037-1039.

HENLE, J.M., HORTON, N.J., JAKUS, S.J. "Modelling Inequality with a Single Parameter" Springer, New York.

KÁMPKE, T. (2004), "Functional inequality measurement", Empirical Economics, Volume 29, September pp 697-704.

LAMBERT, P. (1989), "The Distribution and Redistribution of Income", Basil Blackwell, Oxford.

OGWANG, T., GOURANGA RAO, U.L. (2000) "Hybrid Models of the Lorenz Curve", Economics Letters, Volume 69, October pp 39-44.

SARABIA, J.M. (2008), "Parametric Lorenz Curves: Models and Applications", Springer, New York.

SARABIA, J.M., CASTILLO, E. "A simple Method for Generating Lorenz Curves"

SARABIA, J.M., CASTILLO, E., PASCUAL, M., SARABIA, M. (2005), "Mixture Lorenz Curves", Economics Letters, Volume 89, October pp 89-94.

SARABIA, J.M., CASTILLO, E., SLOTTJE, D.J. (1999), "An ordered family of Lorenz Curves", Journal of Econometrics, Volume 91, July pp 43-60.

SARABIA, J.M., PASCUAL, M. (2002), "A Class of Lorenz Curves Based on Linear Exponential Loss Functions", Communications in statistics. Theory and methods, Volume 31, December pp 925-942 .

IMPACT OF THE EUROPEAN INTEGRATION PROCESS ON THE EXPORTS FROM CASTILLA Y LEÓN (1993-2007): AN ECONOMETRIC ANALYSIS FROM THE GRAVITY EQUATION

Mª DEL PILAR SIERRA FERNÁNDEZ
E-mail: pilar.sierra@unileon.es
ALMUDENA MARTÍNEZ CAMPILLO
E-mail: amarc@unileon.es
Direction and Business Economics Department
Faculty of Economics and Business Studies
León University

Resumen

Este trabajo analiza el impacto del proceso de integración europea sobre el volumen de exportaciones de Castilla y León hacia los países comunitarios (UE-27) durante el período 1993-2007. La metodología utilizada se basa en la estimación de una ecuación de gravedad, que constituye una herramienta útil en la modelización de los flujos de comercio bilateral. Tras aplicar un análisis de datos de panel, los resultados indican que la riqueza del país importador y de Castilla y León, la población regional, así como la distancia y una frontera común entre ellos, son las variables que mejor explican el flujo exportador analizado. Respecto a los distintos acontecimientos vinculados al proceso de integración, la pertenencia del país importador a la UE sólo produjo un impacto positivo y significativo a mediados de la década de los noventa, mientras que el hecho de compartir una moneda común inhibió la cifra de exportaciones a lo largo de todo el período considerado.

Palabras clave: Exportaciones, Proceso de Integración Europea, Ecuación de Gravedad, Castilla y León
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

The aim of this work is to develop an econometric framework to determine the impact of the European integration process on the exports from Castilla y León to the EU-27 countries during the period 1993-2007. This framework is based on the estimation of a gravity equation, which has proved a very useful tool in modelling bilateral currents of trade. After applying panel data analysis, results indicate that incomes of the importing country and Castilla y León, the regional population, the bilateral distance and a common land border are the main explanatory factors of export flows from this Spanish region. Regarding the different events associated to the integration process, the EU membership of the importer country only caused positive and significant effects between 1994 and 1996, whereas the EMU membership reduced export flows from Castilla y León to the European countries during the whole period.

Classification JEL: F13, F14, F15
Keywords: Exports, European Integration Process, Gravity Equation, Castilla y León
Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUCTION

The internationalisation process experienced by the Spanish economy during the recent years represents one of the main structural changes that have enabled a very brief transition from a slow growth model, with a very high level of protection and intervention, to a much

faster model, open to international competition. This internationalisation process has evolved in clear parallelism with the growing integration in the Community Europe, being business activity its main growth factor for the whole of productive sectors.

A number of relevant events have occurred in the process of Community integration consolidation from the beginning of the Spanish association with the European Communities. The first one would be agreed and ratified between 1992 and 1993 with the entry into force of the Maastricht Treaty, by virtue of which not only the European Union (EU) is officially created, but also the free circulation of persons, services, goods and capitals is definitely fostered among the 12 states that formed the Community space at the time (Germany, Belgium, Denmark, Spain, France, Greece, Holland, Ireland, Italy, Luxembourg, Portugal and United Kingdom). This Treaty therefore represents a crucial step on the economic integration process, creating a large internal market, free from barriers and customs, to which three new Western European members are incorporated later, in January 1995, (Austria, Finland and Sweden), forming together the so-called 15 EU.

The second milestone in the integration process took place on January 1st 1999, with the creation of the European Monetary Union (EMU), that involved the institutionalisation of a central authority responsible for defining and implementing a common monetary and exchange policy, as well as managing the new Community currency; although the unification did not become effective until 2002, when the Euro started to circulate as single legal currency in 12 of the 15 EU member countries[1]. At present, four other countries are sharing the single currency, what forms a 16 European countries[2] Monetary Union. With the definitive Euro implementation, the commercial relations among Community members have been consolidated, apart from eliminating the transaction costs linked to the exchange rate fluctuations.

Until 2004 the Community strengthening focused on the consolidation of the economic and monetary integration process among the 15 Western European member countries. But another essential event happens that year with the incorporation of new members to the EU. On May 1st 2004, 10 countries enter the EU (Czech Republic, Cyprus, Slovakia, Slovenia, Estonia, Hungary, Latvia, Lithuania, Malta and Poland, the EU-10), that have different economic, politic and social roots, as well as a very heterogeneous profile compared to the traditional EU-15, forming together the EU-25. The enlargement process continues with the incorporation of Bulgaria and Rumania on January 1st 2007, framing the current 27 member states EU (EU-27).

The chronology of these events, linked to the European integration process, offer an important potential to form a driving force of the Community bilateral trade. They explain the beginning of the study in 1993, with the Maastricht Treaty entry into force, and its end in 2007, the last year with public available data and when the joint effects of the above-mentioned events are already evidenced in commercial relations among member states[3]. In this context our work intends to contribute filling the gap in the study of Spanish external commerce behaviour during that period: the modelling of the Comunidad Autónoma de Castilla y León exports to the rest of the 27-UE countries.

The three specific objectives of this investigation are: First, the analysis of the factors that have influenced the volume of exports from this region to the rest of the Community countries during the last years. Second, to find out if the economic and monetary events linked to the European unification process have affected that export flow. Third, to detect if the economic and monetary integration impact has been constant (such as it is implicitly assumed in most empirical studies), or it has varied throughout time. The so-called "gravity equation" is used to reach these objectives, a widely used econometric tool in the estimation of the bilateral trade standard determinants.

[1] From the fifteen countries that formed the EU, Denmark, United Kingdom and Sweden availed of a voluntary exclusion clause in the adoption of the Monetary Union, which still keeps them outside the Euro.

[2] On January 1st 2007 the Euro starts to circulate in Slovenia. Later, on January 1st 2008, Cyprus and Malta adopt it, while Slovakia does it on January 1st 2009.

[3] Baldwin (2006) indicates that when it comes to study the effects of the European integration process in the commercial flows, it can be adequate to start the sampling period after 1992, since a change in the calculation process and commerce statistics presentation at Community level takes place from that year.

The report describes in the first place the Castilla y León export behaviour to the Community countries during the study period. Next, the theoretic support and the empirical background of the gravity equation are reviewed, specifying the econometric model that will explain the exports intensity from Castilla y León to the rest of the EU-27 countries. The aspects related to the data base, variables and adopted methodology are described subsequently. The fifth section summarises the results obtained, and finally, the main conclusions of the study are presented.

2. THE EXPORT BEHAVIOUR OF CASTILLA Y LEÓN TOWARDS THE EU-27 COUNTRIES (1993-2007)

The Comunidad Autónoma of Castilla y León has been object of important transformations during the past years, among which the gradual internationalisation of its economy stands out, with the commercial facet as main determinant. During the period between 1993 and 2007 in particular, this region's exports rate increased remarkably, tripling the total numbers delivered by the Castilla y León companies (Chart 1). This export dynamism is marked by a high concentration around the EU-integrated countries. Chapter 1 specifically points out that during all the period the exports to the Community countries have always involved more than 80% of international sales.

Chart 1: Temporary dynamic of the Castilla y León exports to the European Union

Exports Volume (in million Euro)															
	1993	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
EU 15	2.361	3.251	3.944	4.219	4.750	5.936	6.583	7.147	7.110	6.753	7.438	7.809	7.366	7.351	7.653
EU 10	17	26	71	56	71	128	141	173	305	323	254	314	412	396	446
EU 2	0,32	1.04	0.60	1.31	1.53	4.69	3.74	17.70	24.45	15.06	16.99	39.27	89.99	157	201
EU 27	2.378	3.278	4.015	4.275	4.821	6.064	6.724	7.320	7.415	7.076	7.692	8.123	7.778	7.904	8.300
Castilla y León Total Export	**2.912**	**3.601**	**4.114**	**4.736**	**5.370**	**6.623**	**7.427**	**8.239**	**8.254**	**7.920**	**8.653**	**9.212**	**8.949**	**9.090**	**9.769**
Representativeness (% on total exports)															
EU 15	81,08	90,28	95,87	89,08	88,45	89,63	88,64	86,75	86,14	85,27	85,96	84,77	82,31	80,87	78,34
EU 10	0,58	0,72	1,73	1,18	1,32	1,93	1,90	2,10	3,70	4,08	2,94	3,41	4,60	4,36	4,57
EU 10 + EU 2	0,59	0,75	1,73	1,18	1,32	1,93	1,90	2,10	3,70	4,08	2,94	3,41	4,60	6,08	6,62
EU 27	81,67	91,03	97,59	90,27	89,78	91,56	90,53	88,85	89,84	89,34	88,89	88,18	86,91	86,95	84,96

Chart by authors.
Source: Special Taxation and Customs Department of the Tax Administration Government Agency (AEAT).
The 2007 data are provisional.

Nevertheless, there are important disparities among the 27 member countries when observing the export dynamic during the analysed period. On the one side, we can point out the essential interpenetration with the EU-15 area of consolidated commercial relations with Castilla y León based on important exchanges volumes, that have however progressively lost representativeness. The EU-15 area involved 78,84% of the region's total exports in 2007; this means a drop of almost 3 percentage points in comparison with 1993 and more than 17 points in comparison with 1995, when that area reached the maximal representativeness of total exports. On the other side, we can observe an emerging link with the new adhesion countries from 2004 (EU-10 and EU-2), that has presented exponential growth during the reference period, even though it still carries low volumes, that at the end of 2007, did not represent more than 6,62% of the total exports dispatched from Castilla y León (Chart 1).

Three periods with different impact on the above-mentioned areas stand out, other than the indisputable disparity in the exported volumes to the traditional EU-15 and the enlarged EU (EU-10 and EU-2), as well as the evolution of its representativeness. In 1993, with the definitive

implementation of the Common Internal Market, the basis for a continuous and sustained growth of the Castilla y León exports to the EU-15 countries is provided. The elimination of barriers and obstacles to the mobility of production inputs and outputs encouraged and deepened the volumes negotiated towards these Community countries during the 1990s. Nevertheless, the volume of exports starts to stabilise from the year 2000, going however through certain corrections and subsequent volume recuperations until the year 2007. In 2002 the economic weakness situation of the main importer countries for Castilla y León (France, Germany, Italy and Portugal), determined a reduction in the exports volume to the EU-15, that would slightly recover in the two following years as the crisis was overcome. Again in 2005 the exports figures decreased, next to the EU-10 incorporation and the emerging competitiveness loss of Spanish products, even though the exports to the EU-15 increased slightly in 2007. In terms of representativeness there is a more noticeable participation of this area in the total exports of Castilla y León between 1994 and 1998, setting 1995 up as the year with a maximum Community specialisation in the exports from Castilla y León (Chart 1). But from 1999, with the official creation of the EMU, the growth observed in the consignments volume to the EU-15 (although with certain corrections), starts to contrast with the constant reductions in their representativeness. The currency unification process subsequently does not seem to have proved very relevant in the export dynamic of Castilla y León to the EU-15 area.

On the other side, the announcement of the future incorporation of 10 Eastern European countries in 2001, and its effective association in 2004 determined the base years in which the Castilla y León exports were relaunched to the EU-10 area. Thereby, in the current decade the new enlarged countries have monopolised the most outstanding growth in the Castilla y León external trade with the EU, both in the consignments volume and the representativeness of the total exported. As seen in Chart 1, the exporting interest of the Castilla y León economic agents did not focus in this area during most of the 1990s, since it did not enjoy trading profits. Nevertheless, the advantages of the quota and customs expenses elimination in trade were advanced, starting a growing export process with the possible incorporation of these 10 new members from 1998, and especially from 2001, coinciding with its officialisation. In particular it is from 2004 when that process experience a notable relaunch period, coinciding with the effective incorporation of the 10 countries to the large Community internal market.

Finally in 2005, when the link of Rumania and Bulgaria to the EU starts to show, the basis of an astonishing growth of Castilla y León exports to the EU-2 area is formed. Regarding its representativeness on the region total exports, Chart 1 shows that the participation of these two countries is much reduced until their incorporation into the Community is official. However, in 2006 and 2007, once their effective adhesion is known, these two new members attain more relevance in the total Castilla y León exports.

Chart 2 shows the progressive geographical diversification experienced during 1993-1997 by the exports consignment from Castilla y León. Thus, in 1993 the prevalence of three bordering countries with similar cultures is noticed: France, Portugal and Italy, that during that year monopolised 81,82% of the Community exports, representing 66,85% of the total Castilla y León exports for that period. However, the attraction exercised by the French market stands out. Although this is an aspect shared at national level, it is more pronounced in Castilla y León, given the relevance of the Renault multinational settled in Valladolid, which generates an important intra-industrial trade with its parent company, located in France.

Chart 2: Exports from Castilla y León to the Community countries (EU-27)

Exports Volume (in million Euro) and Representativeness (% in Community exports)

COUNTRIES	1993		1999		2007		Variation index 1993-2007	
	Million Euro	%	Million Euro	%	Million Euro	%	Million Euro	%
France	1.329,35	55,88	3.981,91	59,17	3.208,26	38,65	141,34	-30,83
Netherlands	43,92	1,85	133,34	1,98	216,85	2,61	393,74	41,51
Germany	127,36	5,35	417,54	6,20	944,32	11,38	641,46	112,51
Italy	305,24	12,83	691,36	10,27	804,61	9,69	163,60	-24,45
United Kingdom	91,49	3,85	230,11	3,42	826,93	9,96	803,85	159,05
Ireland	2,09	0,09	10,46	0,16	37,57	0,45	1.697,61	415,22
Denmark	5,67	0,24	15,88	0,24	76,64	0,92	1.251,68	287,41
Greece	6,50	0,27	111,92	1,66	97,80	1,18	1.404,62	331,24
Portugal	311,81	13,11	769,44	11,43	939,21	11,32	201,21	-13,67
Belgium Luxembourg	115,42	4,85	150,89	2,24	349,44	4,21	202,76	-13,23
Sweden	12,36	0,52	37,04	0,55	64,95	0,78	425,49	50,61
Finland	3,05	0,13	11,09	0,16	38,96	0,47	1.177,38	266,11
Austria	6,87	0,29	22,94	0,34	47,60	0,57	592,87	98,58
EU-15	**2.361,13**	**99,25**	**6.583,92**	**97,84**	**7.653,14**	**92,20**	**224,13**	**-7,10**
Malta	0,36	0,02	1,02	0,02	5,14	0,06	1.327,78	309,22
Estonia	0,00	0,00	0,34	0,01	2,68	0,03	26.700,00	7.581,20
Latvia	0,00	0,00	1,85	0,03	3,18	0,04	31.700,00	9.014,26
Lithuania	0,01	0,00	0,76	0,01	6,19	0,07	61.800,00	17.641,27
Poland	5,11	0,21	15,97	0,24	124,27	1,50	2.331,90	597,01
Czech Republic	0,26	0,01	8,39	0,12	56,88	0,69	21.776,92	6.170,18
Slovakia	0,15	0,01	3,44	0,05	30,98	0,37	20.553,33	5.819,49
Hungary	0,69	0,03	6,18	0,09	46,20	0,56	6.595,65	1.819,05
Slovenia	8,58	0,36	100,96	1,50	154,74	1,86	1.703,50	416,90
Cyprus	2,40	0,10	2,95	0,04	16,14	0,19	572,50	92,75
EU-10	**17,56**	**0,74**	**141,86**	**2,10**	**446,4**	**5,38**	**2.442,14**	**628,61**
Romania	0,28	0,01	2,05	0,03	194,30	2,34	69.292,86	19.788,81
Bulgaria	0,036	0,00	1,69	0,03	6,62	0,08	18.288,89	5.170,47
EU-2	**0,316**	**0,01**	**3,74**	**0,06**	**200,92**	**2,42**	**63.482,28**	**18.123,43**
EU-27	**2.379,01**	**100**	**6.729,52**	**100**	**8.300,46**	**100**	**248,90**	**0,00**

Chart by authors.
Source: Special Taxation and Customs Department of the Tax Administration Government Agency (AEAT).
The 2007 data are provisional.

At the end of the 1990s the exports to France are intensified (as shown in Chart 2, in 1999 they involve 59,17% of the Community exports), while Portugal and Italy lose ground in terms of representativeness in favour of other EU-15 countries such as the United Kingdom and Germany, more distant physically and "psychologically"[4], but with high income levels and large demand markets.

During the current decade, the geographical diversification keeps consolidating, so in 2007 not only a redistribution of the exports consignment to the EU-15 takes place (in which Portugal, Italy and specially France lose relevance versus the import reinforcement of Germany, United Kingdom, the Netherlands, Denmark or Finland, for example), but also EU enlarged countries (EU-10 and EU-2) that stand out because of their productive and demographic dimension, such as Slovenia, Poland, Czech Republic, Hungary, as well as Romania, move on to demand 7,32% of the Community total.

3. THE GRAVITY EQUATION: BACKGROUND REVIEW AND SPECIFICATION PROPOSAL

The gravity equation is one of the most widely used and successful tools for the study of international trade flows determining factors (Evenett y Keller, 2002; Baldwin, 2006). The main idea consists of applying to bilateral trade relations an analogue concept to the Newton Law,

which links the gravity attraction between two objects to their mass size and relative distance. In particular, this equation was originally posed by Tinbergen (1962) and Pöyhönen (1963), who independently suggest that the bilateral trade flow between two countries is positively associated to their income levels, and inversely proportional to their distance. Later, Linnemann (1966) adds the demographic variables (population of both countries) to reflect the economies of scale role. Thereby, the most basic and commonly employed gravity equation can be presented as follows:

$$X_{ijt} = \beta_1(Y_{it} * Y_{jt}) + \beta_2(N_{it} * N_{jt}) + \beta_3(D_{ij}) + \varepsilon_{ijt} \quad (1)$$

X_{ijt} is the total bilateral trade volume between i and j countries in t year; (Y_{it} * Y_{jt}) is the income product of i and j countries in the year t; (N_{it} * N_{jt}) is the populations product of i and j countries in the year t; D_{ij} is the distance between i and j countries; and, ε_{ijt} is the error term.

The income levels of i and j countries normally represent the potential demand of tradeable products from i country and the potential offer of j country, respectively. Therefore, their coefficients are expected to be positive. The distance between i and j countries, intends to evaluate the transport and time costs, as well as the market information access and the markets themselves. These costs increase with distance, so a negative coefficient for this variable is predictable. Regarding populations, several interpretations have been offered, generating an ambiguity in the expected signs of their coefficients. On the one side, the bigger the population is, the more possibilities the country has of taking advantage of the economies of scale, increasing this way its productive specialisation. Under this perspective, it might be expected that it gets more actively involved in the international trade.

The specification of the gravity equation (1) justifies the opportunity of applying natural logarithms to obtain a lineal relation between the trade flows logarithm and the logarithm of the different cause variables, which involves that the estimated coefficients have to be interpreted in terms of elasticity. After applying this transformation, the gravity equation is defined as follows:

$$L_n(X_{ijt}) = \beta_1 L_n(Y_{it}) + \beta_2 L_n(Y_{jt}) + \beta_3 L_n(N_{it}) + \beta_4 L_n(N_{jt}) + \beta_5 L_n(D_{ij}) + \varepsilon_{ijt} \quad (2)$$

The equation (2) reveals the bilateral trade volume among a group of countries during a certain period of time ("*multi-country approach*"). However, the gravity equation can also be used to estimate the unilateral trade flows of a certain country with the rest ("*sole-country approach*"). The main differences between both approaches are the following (Földvári, 2006): (1) with the "*multi-country approach*" the total bilateral trade flows between two countries are usually estimated –exports plus imports[5]-, while with the "*sole-country approach*" it is necessary to estimate two different models; one for exports (*EXP*) and another one for imports (*IMP*); (2) with the "*multi-country approach*" temporal *dummys* can be included to control the unobservable effects dependant on time, while with the "*sole-country approach*" it is necessary to omit such variables in order to read properly the impact of the explanatory variables that only vary over time, such as the income and the reference country population (Y_t y N_t). Since the objective of our work is coherent with the "*sole-country approach*", the referent gravity equation to modelize the export flow is the following:

$$L_n(EXP_{it}) = \beta_1 L_n(Y_{it}) + \beta_2 L_n(Y_t) + \beta_3 L_n(N_{it}) + \beta_4 L_n(N_t) + \beta_5 L_n(D_i) + \varepsilon_{it} \quad (3)$$

EXP_{it} is the exports volume from the reference country to the i country in the t year; Y_t is the reference country income in the t year; N_t is the reference country population in the t year; D_i is the distance between the reference country and the i country; and ε_{it} is the error term.

Despite the gravity equation explanatory power (Minondo, 2007), several authors have added other variables, more or less systematically, that determine the transaction costs and the easiness of the trade relations between two countries. Certain dummy variables stand out among

[5] A criticism to the gravity equation assigned to the "*multi-country approach*" is that the use of the total bilateral trade as a dependent variable can be inappropriate, since it does not enable the discrimination between imports and exports (Dhar and Panagariya, 1999).

them, such as the belonging to a regional trade block, the use of a common language, the border effect and the countries insular character, among others.

3.1. Theoretical justification of the gravity equation

Although the first application of the gravity equation dates back to 1962, a theoretical justification according to the conjectures about international trade has been searched for more than forty years. In this regard, the most notable attempts were those of Anderson (1979) and Bergstrand (1985, 1989). Anderson infers a functional form similar to the gravity equation on the basis of an expenditure system in which all the states have the same utility function and there is product differentiation by country of origin. Therefore, except for this last consideration, the rest of the assumptions are compatible with the Heckscher-Ohlin theory. Bergstrand determines that the gravity equation is a reduced form of a partial balance subsystem emanating from a world trade general balance model, involving in this case the differentiation of companies by country of origin. It is therefore a hybrid between the Heckscher-Ohlin model, that involves perfect competition and goods homogeneity, and the Monopolistic Competition model.

Besides, there are other justification attempts of the gravity equation, a lot more *ad-hoc* than the ones above-mentioned. Thereby, Helpman and Krugman (1985) and Krugman (1995) prove that the gravity equation can proceed from the Monopolistic Competence model with scale rising incomes. Deardorff (1998) on his part, evaluates the gravity equation utility in the trade alternative models test and proves that it can be compatible with both the Ricardian and the Heckscher-Ohlin models. Later, Evenet y Keller (2002), derive the gravity equation both from the Heckscher-Ohlin theory and the scale rising income hypothesis, under the product perfect an imperfect specialisation assumptions, while Eaton and Kortum (2002) develop a Ricardian trade model of homogeneous goods that fits the equation. This large theoretic support explains the recent opinion that the gravity equation is coherent with the most relevant international trade theories (Feenstra, 2004; Anderson and Wincoop, 2004 and Baldwin, 2006, among others). This suggests it can be considered "*one of the higher explanatory power equations of the economic science*" (Minondo, 2007), which has increased the trust in its utility to predict the bilateral trade patterns.

3.2 Empirical background of the gravity equation

An important number of empirical works has led to the suitability of the gravity equation as a tool to explain the international trade flows. While some of them have improved the econometric specification of the equation (Mátyás, 1997; Egger, 2000, 2002; Anderson and Wincoop, 2004; Chen and Wall, 2005; De Nardis *et al*., 2008), others have contributed to the refinement of the explanatory variables or the addition of new variables (Bougheas *et al*., 1999; Limao and Venables, 2007; Voicu and Horsewood, 2007).

Furthermore, recently the use of the gravity equation has become very popular as an empirical study instrument of the effect of several economic integration processes at international level within the bilateral trade pattern (e.g. Baldwin, 1997; Nilsoon, 2000; Rose, 2000; Glick and Rose, 2002; Micco *et al*., 2003; Bun and Klaasen, 2007; De Nardis *et al*., 2008). Several of these studies have been carried out in the specific field of the European Union. While some of them prove that the Community block member states conduct more trade inside than outside their borders (Nitsch, 2000; Evans, 2003; Chen, 2004; Minondo, 2007), others try to consider the European Monetary Union effects (De Souza, 2002; Flam and Nordstrom, 2003; De Nardis and Vicarelli, 2003).

In the Spanish case, the gravity equation has also been used in some works, from which Sanso *et al.* (1989) and (1990) stand out. While the first one tries to analyse if the gravity equation is compatible with the Heckscher-Ohlin theory when explaining the international trade bilateral flows, the main concern of the second one is to determine the functional form of that equation in the study of the external Spanish trade between 1960 and 1985. More recently, Martínez *et al*. (2003) focus their study on the determinants assessment of the international trade bilateral flows among 34 countries during 1980-1999, as well as the preferential agreements effects of several economic blocks, concentrating on Spanish-Mexican relations. On his part,

Gil *et al* (2003) tries to quantify with a 1988-2001 panel data, the impact of the Euro on the Spanish external trade with the EU-15 countries, while Gil *et al.* (2005) uses the gravity equation to estimate the border effect for the whole of the Spanish economy.

3.3. Proposal of our gravity equation

Our first tool to analyse the exports volume from Castilla y León to the Community countries is the gravity equation (3). Two dummy variables *(EU_{it}* and *EMU_{it})* are added to this basic equation, which reveals if the economic and monetary integration process associated to the EU has influenced that export flow during the analysed period. Also, an additional dummy variable is introduced to observe the effect of sharing a common border with the importer country (*$Front_i$*). Therefore, the extended gravity equation can be presented as follows:

$$Ln(EXP_{it}) = \beta_1 Ln(Y_{it}) + \beta_2 Ln(Y_t) + \beta_3 Ln(N_{it}) + \beta_4 Ln(N_t) + \beta_5 Ln(D_i) + \beta_6 UE_{it} + \beta_7 UME_{it} + \beta_8 Front_i + \varepsilon_{it} \quad (4)$$

being:

EXP_{it} is the Castilla y León export intensity to the "*i*" country in the *t* year
Y_{it} is the "*i*" country income in the *t* year
Y_t is the Castilla y León income in the *t* year
N_{it} is the "*i*" country population in the *t* year
N_t is the Castilla y León population in the *t* year
D_i is the distance between Valladolid and the "*i*" country capital
EU_{it} is a dummy variable that shows if the "*i*" country belongs to the EU in the *t* year
EMU_{it} is a dummy variable that shows if the "*i*" country belongs to the EMU in the *t* year
$Front_i$ is a dummy variable that shows if the Spain and the "*i*" country share a common border
ε_{it} is the error term

The extended gravity equation (4) involves that the EU_{it} and EMU_{it} dummy variables effect remains constant along the time. In order to analyse if their impact changes over time, the *EUyearit* or *EMUyearit* will be alternatively included in that equation, eliminating the original dummy variable in each case. These new variables are calculated as the product of the original dummies multiplied by the different yearly *dummy* ($_tYear$), as follows: *EUyearit* = EY_{it} * $_t$Year, where *t* covers the period from 1993 until 2007, and *EMUyearit* = EMU_{it}* $_t$Year, where *t* covers the period from 1999 until 2007. These variables coefficients can be respectively read as the economic integration (EU) and monetary integration (EMU) impact on the Castilla y León exports volume in each *t* year. These variables ratios can be read as the impact of the economic integration (EU) and the monetary integration (EMU) on the Castilla y León exports volume in the *t* year. Accordingly, the two new gravity equation specifications would be:

$$Ln(EXP_{it}) = \beta_1 Ln(Y_{it}) + \beta_2 Ln(Y_t) + \beta_3 Ln(N_{it}) + \beta_4 Ln(N_t) + \beta_5 Ln(D_i) + \beta_6 UEaño_{it} + \beta_7 UME_{it} + \beta_8 Front_i + \varepsilon_{it} \quad (5a)$$

$$Ln(EXP_{it}) = \beta_1 Ln(Y_{it}) + \beta_2 Ln(Y_t) + \beta_3 Ln(N_{it}) + \beta_4 Ln(N_t) + \beta_5 Ln(D_i) + \beta_6 UE_{it} + \beta_7 UMEaño_{it} + \beta_8 Front_i + \varepsilon_{it} \quad (5b)$$

4. INVESTIGATION DESIGN

4.1. Data base

The analysed sample includes the other 26 countries of the EU-27 (all, except for Spain) that can act as Castilla y León goods importers. These countries are, in particular, Germany, Austria, Belgium, Bulgaria, Cyprus, Denmark, Slovakia, Slovenia, Estonia, Finland, France, Greece, Holland, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Poland, Portugal, United Kingdom, Czech Republic, Rumania and Sweden. The investigation period goes from 1993 until 2007. Therefore, there are a total of 390 observations[6].

[6] In the cases of Belgium and Luxembourg, there are only separated data regarding their external trade from 1999. Therefore, there is an incomplete panel data.

4.2. Variables measure

Castilla y León export intensity (EXP_{it})

The dependant variable was quantified from the total number of exports (FOB) from Castilla y León to the respective Community countries in each studied year, in constant million Euro (base year 2000). The initial information to measure this variable was taken from the External Trade Statistics elaborated by the Junta de Castilla y León, deflating the resulting series with the EU-27 CPI, obtained from the OECD data bases.

Yearly income level (Y_{it} ,Y_t)

In order to measure the yearly income level, the Gross Domestic Product (GDP) per capita at constant prices was used, (base year 2000). In order to determine its value, the EUROSTAT yearbooks were used in the case of the Community countries, as well as the Spanish Regional Accounting yearbooks elaborated by the National Statistics Institute (INE), in the case of Castilla y León. The original series at current market prices ware deflated using the EU-27 and the national GDP; both obtained from the OECD data bases.

Population (N_{it} , N_t)

The population measure employed was the number of inhabitants, in thousands, on January 1st of the selected year. This data was obtained from the EUROSTAT yearbooks for the different Community countries and from the Spanish Regional Accounting yearbooks elaborated by INE for Castilla y León.

Distance (D_i)

The distance between the capital of Castilla y León (Valladolid) and those of the Community countries was calculated in kilometres, based on the latitude and longitude of the respective geographical centres (geodesic distance). The data was obtained from the Geodesic Utilities Service (IGN), offered by the National Geographic Institute.

European Union Membership (EU_{it})

This dummy variable considers the incentive for the Castilla y León exporters of belonging to a single market of goods, services, persons and capitals, which includes the participation of the export destination country. This variable takes value 1 when Castilla y León exports to the EU *i* country in the *t* year, or value 0 in the opposite case. So, in 1993 and 1994 this variable only equals 1 for the 12 signatory countries of the Maastricht Treaty in 1992. From 1995 until 2003 the variable takes value 1 in the 12 above-mentioned countries, as well as in the exports to Austria, Finland and Sweden, which join the European Union in 1995. In 2004 and 2006 the variable equals 1 for the 15 aforementioned countries, as well as for the 10 states that joined the European Union in 2004. Finally, in 2007 it takes value 1 for the 26 countries of the sample.

European Monetary Union Membership (EMU_{it})

This dummy variable measures the stimulating effect that the use of a common currency with the export destination country can have on the Castilla y León exports. This variable takes value 1 when Castilla y León exports to the EMU *i* country in the *t* year, and 0 value in the opposite case. Therefore, since the EMU took place in 1999, this variable is 0 for all the trade flows before that year. From 1999 until 2006, the variable takes value 1 for the exports from Castilla y León to the countries that adopted the single currency from the beginning (from 2001 in the case of Greece). Finally, in 2007 this dummy variable takes value 1 for the 12 aforementioned countries, as well as Slovenia, that joined the eurozone at the beginning of that year.

Border ($Front_i$)

This dummy variable considers if the export destination country shares a border with Spain. So it shows if the existence of a common border facilitates the trade between two market areas, by reason of geographic proximity or the potential existence of cultural or historic links. This variable takes value 1 if this is the case, and 0 value in the opposite case.

Year ($Año_t$)

It is a temporary dummy variable that allows to control the non-observable effects dependant on time (t = 1993,…2007). Since in this work the gravity equation is contemplated under the *"sole-country approach"*, this variable is not directly included in the analysis to the extent that it would prevent the correct reading of the explanatory variables impact that only vary with time (such as the Castilla y León income and population). In fact, this variable is only used for the construction of the *UEyearit* and *UMEyearit* variables, allowing to detect if the economic integration (EU) and the monetary integration (EMU) vary over time.

Chart 1 considers the different variables introduced in the gravity equation to analyse the export flow from Castilla y León to the Community countries during 1993-2007

Chart 1. Study variables

* These variables are introduced in the gravity equation in the form of logarithms.

Dependant variable	**Measure**	
$*EXP_{it}$	Export figures from Castilla y León to the *i* country in the *t* year (*constant million €*)	
Independent variables	**Measure**	**Expected effect**
$*Y_{it}$	Real income level of the *i* country in *t* year (*GDP per capita at constant prices*)	Positive
$*Y_t$	Castilla y León real income level in *t* year (*GDP per capita at constant prices*)	Positive
$*N_{it}$	Population of *i* country in *t* year (*thousands of inhabitants on January 1st)*	Positive/Negative
$*N_t$	Castilla y Leon population in *t* year (*thousands of inhabitants on January 1st*)	Positive/Negative
$*D_i$	Geodesic distance between Valladolid and *i* country capital (*kilometres*)	Negative
EU_{it}	EU Membership of *i* country in *t* year (*Dichotomous: 0,1*)	Positive
EMU_{it}	EMU Membership of *i* country in *t* year (*Dichotomous: 0,1*)	Positive
$Border_i$	*i country shares border with Spain* (*Dichotomous: 0,1*)	Positive
$Year_t$ (t =1993 until 2007)	Export flow regarding *t* year (*Dichotomous: 0,1*)	

4.3 Estimation method

The gravity equation was initially developed to address cross-cutting analysis. Nevertheless, these analysis have an essential problem, since they do not take into consideration any heterogeneous characteristic associated with the bilateral trade relation; this means that a region can export to two countries in a different way, even though they are equal in size and distance, due to certain non-observable characteristics linked to cultural, historic, politic or geographic factors, which if not considered could affect the validity of the results. For this reason it is said that the cross-cutting models suffer a heterogeneity bias.

At present, in order to solve this problem, the gravity equation estimations are normally conducted through panel data analysis techniques. It is generally accepted that the estimators based on this type of analysis are a lot more precise than those obtained through other techniques, reducing as well the problems concerned with the models identification. In this respect, static and dynamic panel data analyses are identified, as far as their specifications include or not variables belonging to different time periods[7].

[7] A static panel data analysis has been applied to estimate the gravity equation in the works of Mátyás (1997), Glick and Rose (2002), Martínez *et al.* (2003), Cheng and Wall (2005) and Bun and Klaassen (2007), among others. On the other side, the works from Bun and Klaassen (2002), De Nardis and Vicarelli (2003) and De Nardis *et al.* (2008) have used with this purpose a dynamic panel data analysis.

If static models are used, the non-observable heterogeneity can be studied including fixed or random effects in the model. With the *fixed effects estimator* (EF), the "country" and "time" effects are assumed as fixed parameters since they are estimated and correlated with the regressors. In this case, for econometric reasons, all the explanatory variables that remain constant over time (distance and border, for example), have to be excluded from the specification, and the individual differences will be captured by the constant term. On the contrary, with the *random effects estimator* (EA), the "country" and "time" effects are stochastic assumptions, non-correlated with the regressors. Therefore, the inclusion of time-invariant explanatory variables is allowed in the specification of these models. The Hausman[8] contrast allows determining which of the two estimators is more appropriate in each case.

The dynamic models are used when the variables present an autoregressive character; that is, when the retarded dependant variable is among the explanatory variables. So they can be applied to estimate the gravity equation to the extent that the external trade operations of a certain period can be related to those of previous periods. However, this circumstance generates inconsistent estimations when working with panel data. Different estimation models can be applied in order to address these inconsistencies, especially the Anderson and Hsiao model (1981), the GMM estimator (Arellano and Bond, 1991), the GMM-System estimator (Blundell and Bond, 1998) and the LSDV correction method (Bun and Kiviet, 2003).

The academic community has recognised that when the *"sole-country approach"* is used, generally involving the provision of a small sample to estimate the gravity equation, it proves more convenient from the econometric point of view to apply a panel data static analysis versus a dynamic one (Egger, 2000, 2002; Martínez *et al*., 2003; Chen and Wall, 2005). For this reason, the panel data static analysis[9] has been employed in this work, so the error term of the different models is defined as $\mu_{it} = \alpha_i + \varepsilon_{it}$, where α_i collects the "destination country" unobservable effects that do not vary over time and influence the different observations in a different way (non-observable heterogeneity) and ε_{it} is a white noise variable that fits the following properties:

$E[\varepsilon_{it}] = 0 \quad i = 1, \ldots, N; \quad t = 1, \ldots, T$
$E[\varepsilon_{it}^2] = \alpha_\varepsilon^2 \quad i = 1, \ldots, N; \quad t = 1, \ldots, T$
$E[\varepsilon_{it}\,\varepsilon_{js}] = 0 \quad i, j = 1, \ldots, N; \quad t, s = 1, \ldots, T \;$ such that $i \neq j$ ó $t \neq s$

5. ANALYSIS RESULTS

Chart 2 presents the main descriptive statistics of the data panel. It particularly shows information about the mean, median, standard deviation, maximum and minimum value of the gravity equation base variables, expressed in the original measurement unit. On the other side, chart 3 shows the frequencies of the three dummy variables introduced in the extended equation.

[8] The Hausman contrast (1978) analyses the possible correlation between the error term and the regressors in order to decide between a fixed effects or random effects estimation. Under the null hypothesis of no difference between both estimators, the Hausman statistic converges in distribution to a chi-squared distribution with K degrees of freedom, where K is the number of explanatory variables included in the model specification.

[9] Additionally, apart from applying a panel data statistical analysis, the different specifications of the gravity equation were estimated from a dynamic model – the GMM-System estimator, implanted in the STATA program with the *xtabond2* command-. After the estimation, the results obtained reveal that the retarded dependant variable does not result significant in any specification, proving the inconvenience of applying a panel data dynamic analysis. Regarding the rest of equation variables, the results generally agree with those obtained in the static analysis. It all confirms the convenience of applying a panel data static model in this case.

Chart 2. Descriptive statistics

EXP_{it} is the exports figure from Castilla y León to the *i* country in *t* the year (in constant million €); Y_{it} is the real income of the *i* country in the *t* year and Y_t is the real income of Castilla y León in the *t* year (in GDP per capita at constant prices); N_{it} is the population of the *i* country in the *t* year and N_t is the Castilla y León population in the *t* year (in thousands of inhabitants); D_i is the geodesic distance between the capital of Castilla y León -Valladolid- and the *i* country capital (in kilometres).

Variables	N	Mean	Median	S.D.	Minimum	Maximum
EXP_{it}	378	184,182	17,998	494,572	0,009	3.240,430
Y_{it}	385	12.724,832	10.558,008	9.238,584	235,185	46.598,530
Y_t	390	11.323,090	11.264,341	1.752,098	8.537,863	14.308,450
N_{it}	390	17.057,738	8.595,500	22.188,893	371,000	82.520,000
N_t	390	2.517,560	2.493,918	40,496	2.479,118	2.584,407
D_i	390	2.440,153	2.419,500	800,677	588,000	3.978,000

*p < 0.10; ** p < 0.05; *** p < 0.01

Chart 3. Chart of Frequencies

UE_{it} is a dummy variable that reflects whether or not the *i* country belongs to the EU in the *t* year; UME_{it} is a dummy variable that reflects whether or not the *i* country belongs to the EMU in the *t* year; $Border_i$ is a dummy variable that reflects if Spain and the *i* country share a common border.

	EU_{it}	EMU_{it}	$Border_i$
NO (0)	144 (36,9%)	293 (75,1%)	360 (92,3%)
YES (1)	246 (63,1%)	97 (24,9%)	30 (7,7%)
N	390	390	390

Chart 4 shows the results of the estimation of our gravity equation. Their different specifications are estimated through panel data analysis techniques; in particular, on the basis of the EF and EA estimators, among which the respective Hausman contrast is carried out.

Chart 4. Results of the gravity equation estimation

Dependant variable: Exports figure from Castilla y León to the *i* country in the *t* year (at constant prices)

Y_{it} is the *i* country real income in the *t* year and Y_t is the Castilla y León real income in the *t* year (in GDP per capita at constant prices); N_{it} is the *i* country population in the *t* year and N_t is the Castilla y León population in the *t* year (in thousands of inhabitants); D_i is the geodesic distance between Valladolid and the *i* country capital (in kilometres); EU_{it} is a dummy variable that reflects if the *i* country belongs to the EU in the *t* year; EMU_{it} is a dummy variable that reflects if the *i* country belongs to the EMU in the *t* year; $Border_i$ is a dummy variable that reflects if Spain and the *i* country share a common border; *EUyearit* refers to the 15 variables resulting from multiplying the dummy variables EU_{it} and $Year_t$, where *t* goes from 1993 until 2007; *EMUyearit* refers to the 9 variables resulting from multiplying the dummy variables EMU_{it} and $Year_t$, where *t* goes from 1999 until 2007.

	Basic gravity Equation		Extended gravity Equation		Extended Equation with EU evolution		Extended Equation with EMU evolution	
Variables	**Model 1 EF**	**Model 1´ EA**	**Model 2 EF**	**Model 2´ EA**	**Model 3 EF**	**Model 3´ EA**	**Model 4 EF**	**Model 4´ EA**
Constant	65,008*** (25,059)	53,089** (24,042)	70,459*** (25,481)	59,941** (24,524)	62,487 (51,905)	48,005 (49,042)	37,447 (32,555)	17,061 (31,541)
$Ln(Y_{it})$	2.094*** (0,196)	1,808*** (0,132)	1,910*** (0,197)	1,701*** (0,135)	1,780*** (0,203)	1,547*** (0,142)	1,877*** (0,198)	1,482*** (0,107)
$Ln(Y_t)$[a]	1,797*** (0,411)	1,973*** (0,332)	2,279*** (0,483)	2,608*** (0,394)	3,187*** (0,859)	3,866*** (0,817)	2,828*** (0,576)	3,703*** (0,475)
$Ln(N_{it})$	-2,147 (1,564)	0,897*** (0,136)	-1,051 (1,556)	0,882*** (0,132)	-0,208 (1,612)	0,874*** (0,132)	-0,542 (1,591)	0,915*** (0,076)
$Ln(N_t)$[a]	-10,047*** (2,746)	-10,679*** (2,815)	-12,348*** (2,818)	-12,974*** (2,808)	-13,225** (5,454)	-12,669** (5,457)	-9,322*** (3,379)	-8,185** (3,609)
$Ln(D_i)$	–	-1,216*** (0,498)	–	-0,410 (0,690)	–	-0,504* (0,689)	–	-0,796** (0,401)
EU_{it}[b]			0,126 (0,140)	0,964 (0,139)			0,027 (0,151)	0,073 (0,156)
EMU_{it}[c]			-0,554*** (0,147)	-0,703*** (0,137)	-0,469** (0,188)	-0,580*** (0,177)		
$Border_i$			–	2,423** (1,003)	–	2,283** (1,001)	–	2,157** (0,570)
UEyearit[a]					EU_{94} +*** EU_{95}+** EU_{96}+**	EU_{94}+*** EU_{95}+** EU_{96}+**		
EMUyearit							EMU_{99}-* EMU_{00}-* EMU_{01}-* EMU_{02}-** EMU_{03}-** EMU_{04}-*** EMU_{05}-** EMU_{06}- ** EMU_{07}-***	EMU_{99}-* EMU_{00}-** EMU_{01}-* EMU_{02}-*** EMU_{03}-*** EMU_{04}-*** EMU_{05}-*** EMU_{06}-*** EMU_{07}-***
Hausman	χ^2 (4) = 30,28***		χ^2 (6) = 11,72*		χ^2 (20) = 7,83		χ^2 (14) = 9,62	
R^2	0,66		0,68		0,83		0,84	
N	373		373		373		373	

a. Only the significant variables are shown with the respective signs and signification levels.

The values are the non-standardised ratios, with the typical mistakes in brackets.
* p < 0.1; ** p < 0.05; *** p < 0.01

The [*1*] and [*1´*] *Models* refer to the EF y EA estimators for the basic gravity equation (3). The [*2*] and [*2´*] *Models* collect the results of both estimators for the extended gravity equation (4). The [*3*] and [*3´*] *Models* are for the EF and EA estimators of the extended gravity equation with the evolution of the EU impact over time (5a) and, finally, the [*4*] y [*4´*] *Models* reflect the results of both estimators for the extended equation with the evolution of the EMU effect (5b). Before making the estimations, it is confirmed that the different models do not have autocorrelation or heterocedasticity problems. Besides, after applying the F test we can observe that in all cases, the unobservable "destination county" effects are jointly significant, indicating that there is unobservable heterogeneity and thereby, the adequacy of the panel data analysis application.

Regarding the basic gravity equation ([*1*] and [*1´*] *Models*), the Hausman contrast indicates that it is correct to consider the individual effects as fixed effects, and therefore to interpret the variables ratios that vary over time, obtained from the EF estimator ([*1*] *Model*)[10]. In this respect, the results reveal that the *importer country income level* contributes to explain a higher export flow from Castilla y León to the Community countries. The wealth of a country has a direct impact on the exports figure to that country, which proves the importance of the buying capacity of the buyer country. In fact, Western European countries with high per capita income levels and advanced infrastructures, such as France, Italy, Germany and the United Kingdom, have traditionally been the main destination of the dispatches from Castilla y León. Even within this group of countries, a reduction in the representativeness of the exports to France and Italy occurred during 1993-2007, possibly as a consequence of their slowed economic growth, while the volumes dispatched to Germany and the United Kingdom – countries with a high living standard – has increased significantly. Besides, other Western European countries, such as Finland, Sweden, Austria and Denmark, have reached a growing importance as importers during the last years, as well as other Eastern countries, like Poland, Czech Republic, Slovakia and Hungary, with higher wealth levels, and faster economic growths and purchasing power than the rest of countries.

On the other side, the positive influence of the *Castilla y León income level* proves the importance of the productive capacity of this region to foster its exports. A high regional income level shows a high production, and thereby, more export resources. The importance of the income level on the own exports rests on the modern theories that link the export volumes to the offer factors. According to these theories, those regions able to generate competitive advantages in an exogenous way, will be more successful in their openness to the external markets. So the higher export figures will concentrate on the regions where those advantages are more easily developed due to the market relative dimension, new technologies investment capacity and the availability of more human capital stock.

Regarding the *population* variables, the results show that only the number of Castilla y León inhabitants affects in a statistically significant way the exports volume to the EU-27 countries. In particular, this seems to be the most significant factor in the region export flow, with an important negative influence to the extent that a smaller number of inhabitants results in a smaller self-sufficiency tendency and therefore, a higher external trade commitment. In this sense, since Castilla y León is characterised by a constant reduction in the number of inhabitants, it is consistent to induct its growing openness to international markets.

Regarding the enlarged gravity equation ([*2*] y [*2´*] *Models*), the Hausman contrast also shows that it is convenient to interpret the coefficients obtained, based on the EF estimator ([*2*] *Model*)[11]. As it can be observed, the wealth levels of the buyer country and Castilla y León, as well as the population level of the region, still have a significant statistical effect. Besides, regarding the new dummy variables introduced in this specification, the results show that the

[10] The EF estimation of *Model* [*1*] does not present any ratio for the "distance" variable because its value remains constant over time.

[11] The EF estimation of *Model* [*2*] does not present any coefficient for the "*distance*" and "*border"* variables because their values remain constant over time.

fact that the *importer country is a member of the EU regional block* has affected in a positive but insignificant way the Castilla y León Community exports during the analysed period. In particular, the fact that the buyer country belongs to the EU, has only increased on average 13,4% the export flow of this region between 1993 and 2007[12]. This means that if the economic integration process has had any favourable impact on the Castilla y León exports, this has been unimportant. Two reasons can explain this fact: (1) the exports growing volumes to the EU-15 only kept a high rate of growth until the end of the 1990s. From then, apart from the fact that the main competitive advantages were already deducted, an economic slow down occurs in some of the main target markets of the Castilla y León exports, beginning a progressive loss of representativeness in the numbers dispatched to the EU-15; and (2) the export flow that, from 2001 directs towards the Community countries of Eastern Europe, despite its fast growth and, thus acquisition of a larger role in the final figures of Castilla y León, still assumes reduced export volumes. Therefore, despite that the international sales of Castilla y León businesses have benefited from the process of European economic integration, the results have been limited, having been unable to take advantage of the total potential derived from the larger size of the European internal market and, above all, the free circulation of goods within it.

In regard to the *membership of the import country to the EMU*, the results reveal that the process of monetary unification has not had the hoped for stimulus in the exports from Castilla y León. Thus, in contrast to what was expected, the fact of sharing a common currency with the import country has produced a significant negative effect on the sales of this region towards the Community countries in the reference period. In particular, this process has led to an average reduction of 42 percent in the export flow. Effectively, since 2002, reductions have been observed in export figures towards such countries as France, Italy, Portugal and Greece. Faced with this fact, there have been increases in the volume of dispatches towards the Western European countries that have not voluntarily joined the eurozone (United Kingdom, Denmark and Sweden) as well as the Eastern European countries that have not reached the required convergence criteria. Without considering other factors with potential to influence this unexpected result (for example, the de-acceleration of fundamental economies of the Community Europe), our finding permits to infer that Castilla y León has had growing difficulties to obtain advantages of the monetary integration, probably caused by the progressive loss of the competitiveness of Spanish products. Therefore, once the advantages of the exchange rates eliminated, the penalty that has generated a high and rising inflation on the exportable products has been revealed, together with higher labour costs. In this line of argument, Berger and Nitsch (2005) point out regarding the single currency, that the abolition of national currencies can lead to intra-Community trade flows, but also demand that the decisions about exporting within the eurozone are taken, since then, in function of the products competitiveness and quality. It is to say, with a single currency, the relative prices factor becomes the only determinant in the real exchange rate; that is, the competitiveness of external trade.

Regarding the extended gravity equations that consider that the impact of membership of the EU (*Models [3] and [3']*) and to the EMU (*Models [4] and [4']*) is not constant and can vary over time, the Hausman contrast leads to the acceptance of the null hypothesis and, therefore, the results of the EA estimator can be interpreted in both cases (*Models [3'] and [4']*). In these two specifications, the basic variables of the gravity equation continue to show the previously detected effects on the basis of the EF estimator, with the exception of the population from the export destination country, whose coefficient is now positive and significant. This result would reflect the growing importance of the "size of market" effect in international trade, with a general openness increase of foreign trade.

In addition, the results of both models indicate that, just as expected, *the distance* affects significantly and negatively the export flow from Castilla y León. Thus, the larger the route to cover between the geographic centres of this region to the import country, the lower the attraction of trade will be between them, due to, among other aspects, the larger time and

[12]The way the gravity equation is defined in *Model* [*2*], the interpretation of the coefficient matching the UE_{it} variable, has to be carried out on the basis of the following expression: $(e^{0,126}-1)*100 = 13,4\%$.

transport costs. With countries like France, Portugal and Italy, key to the regional export, this variable may have been essential to explain the export flow from Castilla y León. Regarding the *border* variable, the estimation of both models shows the importance of sharing a common border with the buying country, for the sales of this Spanish region to the EU-27 countries; this means that the Community countries tend to trade more intensely with the businesses of Castilla y León as a result of their geographic location and/or the existence of possible cultural and historic links. In fact, of all the EU countries, France and Portugal have always been the principal destination markets of the Castilla y León exports. The sign and significance of this variable is important in that it can explain, at least partially, the lack of statistical significance of the variable that reflects the membership of the import country to the EU.

Finally, when the *Model [3']* introduces the different variables that collect changes in the impact of the economic integration in the EU between the years 1993-2007, it is possible to point out that, in the large part of the studied period, the import country membership of this regional block has no significant effect on the exports of Castilla y León, so that only in 1994, 1995 and 1996 the influence appears to be positive and significant. So this Autonomous Community only benefited significantly from the advantages derived from the creation of a common market, without hindrance or barriers in trans-border sales, during the three years that followed the creation of the European Union. Regarding the variables that reflect if the effect UME membership changes over the analyzed period, the *Model (4)* results indicate that from 1993 to 2007, the fact that the importer country shares a common currency with Castilla y León, has a negative and significant impact on the export flow of this region; this means that during all that period, the EMU produces an inhibiting effect on the Castilla y León exports to the Community countries. This find indicates that Castilla y León has suffered difficulties to adapt to the challenges of the monetary unification process, inhibiting the exploitation of the comparative advantages of a European market that negotiated a single currency.

6. CONCLUSIONS

The most relevant change that the economy of Castilla y León has experienced in the past years has been, without a doubt, its progressive opening to international markets, especially to the EU countries. For this reason, the main purpose of this report has been to try to explain what the impact of the EU integration process has been on the export volume from Castilla y León to the rest of the community countries (EU-27) during the period of 1993-2007.

Specifically, there have been three objectives to reach. First, to determine what factors have influenced the external sales of this region to the community countries during the study period. Second, to determine if the events linked to the European economic and monetary integration have served as a true incentive for the Castilla y León exports. Third, to discover if the integration process has been constant or has varied over the course of time. In order to reach these objectives, a gravity equation estimated on the basis of panel data analysis, was specified. The obtained results can prove of vital importance in the identification of the countries with a higher tendency to import products from Castilla y León; that is to say, the potential export markets for the companies of this Autonomous Community. As a result, they can provide a valuable guide for the development of the internationalisation policy of this Spanish region.

With respect to the first objective, our findings indicate that most of the basic variables of the gravity equation are relevant and reveal the expected results. Thus, taking into account the type of goods exported from Castilla y León, elaborated with medium-high technology and meeting medium quality standards, a factor such as the standard of living of the buying countries becomes fundamental when it comes to explain the number of exports from this region to the Community countries. In effect, besides Western European countries with high income per capita, such as France, Italy, Germany, United Kingdom, Finland and Denmark, the consignments originating in Castilla y León have also gone towards other eastern importers, such as Poland, the Czech Republic, Slovakia and Hungary, characterized by growing purchasing power and important economic growth. The richest countries tend to import more

goods, not only because of their income volume, but also because their wealth allows them more and better infrastructure.

In reference to the characteristics of Castilla y León, its income level as much as a fall in the size of its population, appear to augment the export flow from this region towards European countries. These results show the relevance of the supply factors related with the level of development in Castilla y León, as well as the number of inhabitants of this Autonomous Community, when it comes to determine the magnitude of its commercial relations in the European sphere.

On the other hand, in accordance with economic intuition, exterior sales from this Spanish region to the remainder of the Community countries appear to increase when the distance is less between them. Furthermore, it can be qualified that among the nearest Community countries, Castilla y León tends to export more intensely towards those geographically adjacent, confirming in this manner the positive impact of the so called "border effect" (Minondo, 2007). In fact, France and Portugal have always monopolized a large portion of the total exports of Castilla y León towards European countries.

The results obtained regarding the *second and third objectives* indicate that, in general, during the period of 1993-2007, the creation of the EU only stimulated in a limited form the growth of export volumes from Castilla y León to the rest of the countries belonging to this regional block, except for the three years immediately following its officialisation, when the membership of the import country to the EU had an important favourable impact on the number of consignments originating in this region. This allows to point out that the companies located in Castilla y León rapidly deduced the commercial advantages derived from the creation of a single market. In addition, even though in the new decade the internal market grew with the incorporation of 12 new commercial partners towards which Castilla y León redirected a growing volume of exports, the trade figures reached very little representativeness within the overall Community exports and the total dispatched from this region.

In regard to the monetary unification process (EMU), contrary to expectations, the fact of sharing a common currency – the euro – with the buying country, significantly and negatively affected the external sales of Castilla y León towards the European countries during the period of 1993-2007. Besides, this adverse effect has remained constant over the course of time, manifesting itself in all the years during the analyzed period. The initial introduction of the euro became effective only among Castilla y León business partners that belonged to the traditional EU-15 (leaving the Eastern European countries out), corresponding with mature markets where the commercial advantages had already been exploited. Therefore, the introduction of a common currency may have become an obstacle for the exports of this Spanish region, due to the heightened loss of competitiveness of the products from Castilla y León, and the relative factor price (penalised by a growing and elevated inflation, together with high labour costs compared to those of our community environment) becoming the only determinant of the competitiveness of external trade. Nevertheless, it is important to note that these conclusions are only acceptable as measures of the initial impact, not of the long-term effects of the European integration process. In fact, it has been pointed out that in general, it takes a few years before a regional integration process produces a significant amplifier effect on bilateral trade between the countries involved (De Nardis *et al*., 2008).

7. BIBLIOGRAFÍA

ANDERSON, J. (1979): "A Theoretical Foundation for the Gravity Equation", *American Economic Review*, 69, pp. 106-116.

ANDERSON, J. y WINCOOP, E. (2004): "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review*, 93, pp. 170-192.

ANDERSON, T.W y HSIAO, C. (1981): "Estimation of dynamic models with error components", *Journal of the American Statistical Association*, 76, 598-606.

ARELLANO, M. y BOND, S. (1991): "Some Tests of Specification for Panel Data: MonteCarlo Evidence and an Application to Employment Equations". *Review of Economic Studies*, 58, pp. 277-297.

BALDWIN, R. (1997): "The Causes of Regionalism", *The World Economy*, 20, pp. 247-281.

BALDWIN, R. E. (2006): *The Euro's Trade Effects*. Mimeo, The Graduate Institute of International Studies, Geneva.

berger, H. y NITSCH, V. (2005): "Zooming Out: The Trade Effect of the Euro in Historical Perspective", *CESifo Working Paper*, 1435

BERGSTRAND, J.H. (1985): "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence", *The Review of Economics and Statistics*, 67, pp. 474-481.

BERGSTRAND, J.H. (1989): "The Generalised Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade", *The Review of Economics and Statistics*, 71, pp. 143-153.

BLUNDELL, R. y BOND, S. (1998): "Initial conditions an moment restrictions in dynamic panel data models", *Journal of Econometrics*, 87, pp. 115-143.

BOUGHEAS, S., DEMETRIADES, P.O. y MORGENROTH, E. (1999): "Infrastructure, Transport Costs and Trade", *Journal of International Economics*, 47, pp. 169-189.

BUN, M. y KIVIET, J.F., (2003): "On the diminishing returns of higher order terms in asymptotic expansions of bias. *Economics Letters*, 79, 145-15.

BUN, M. y KLAASSEN, F. (2002): *The Importance of Dynamics in Panel Gravity Models of Trade*. Mimeo.University of Amsterdam.

BUN, M. y KLAASSEN, F. (2007): "The Euro effect on trade is not as large as commonly though", *Oxford Bulletin of Economics and Statistics*, 69, pp. 473-496.

CHEN, N. (2004): "Intra-National Versus International Trade in the European Union: Why Do National Borders Matter?", *Journal of International Economics*, 63, pp. 93-118.

CHEN, I.H. y WALL, H.J. (2005): "Controlling for Heterogeneity in Gravity Models of Trade and Integration". *Review "Federal Reserve Bank of St. Louis"*, January, pp. 49-63

DE NARDIS, S. y VICARELLI, C (2003): "Currency Unions and Trade: The special Case of EMU", *Review of World Economics*, 139, pp. 625-649.

DE NARDIS, S.; DE SANTIS, R. y VICARELLI, C (2008): "The Euro´s effects on trade in a dynamic setting", *European Journal of Comparative Economics*, 5, pp. 73-85.

DE SOUZA, L. (2002): "Trade effects of monetary integration in large, mature economies: a primer on the European Monetaria Unión", *Kiel Working Paper*, 1137. Kiel Institute for the World Economy

DEARDORFF, A.V. (1998): "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?", in J.A. Frankel (eds) *The Regionalisation of the World Economy*. pp.7-22. Chicago: University of Chicago Press.

DHAR, S. y PANAGARIYA, A. (1999): "Is East Asia Less Open than North America and the EEC? No", in J. Piggott & A. Woodland (eds) *International Trade Policy and the Pacific Rim*. London: Macmillan.

EATON, J. y KORTUM, S. (2002): "Technology, Geography, and Trade.", *Econometrica*, 70, pp.1741-1779.

EGGER, P. (2000): "A Note on the Proper Econometric Specification of the Gravity Equation", *Economics Letters*, 66, pp. 25-31.

EGGER, P. (2002): "An Econometric View of the Estimation of Gravity Models and the Calculation of Trade Potentials", *The World Economy*, 25, pp. 297-312.

EVANS, C.L. (2003): "The Economic Significance of National Border Effects", *American Economic Review*, 93, pp. 1291-1312.

EVENETT, S. J. y KELLER, W. (2002): "On Theories Explaining the Success of the Gravity Equation", *Journal of Political Economy* 110(2).

FLAM, H. and H. NORDSTROM (2003): *Trade Volume Effects of the Euro: Aggregate and Sector Estimates*. Mimeo, IIES Stockholm.

Feas

FEENSTRA, R. (2004): *Advanced International Trade*. Princeton, Princeton University Press.

FÖLDVÁRI, P. (2006): *The Economic Impact of the European Integration on the Netherlands. A Quantitative Analysis of Foreign Trade and Foreign Direct Investments*. Tekst. - Proefschrift Universiteit Utrecht.

GIL, S., LLORCA, R. y MARTÍNEZ, J.A. (2003): "El Euro y la Integración Comercial Española. Un Análisis Comparado", *Economía Industrial*, 349/350, pp. 139-146.

GIL, S., LLORCA, R., MARTÍNEZ, J.A. y OLIVER, J. (2005): "The Border Effect in Spain", *The World Economy*, 28, pp. 1617-1631.

GLICK, R. y ROSE, A.K. (2002): "Does a Currency Union Affect Trade? The Time-Series Evidence", *European Economic Review*, 46, pp. 1125-1151.

JOHANSON, J. y WIEDERSHEIM, P. (1975): "The internationalization of the firm-four Swedish cases". Reeditado en Buckley, P. & Ghauri, P. (1993) *The Internationalization of the Firm: A Reader*. Academic Press.Cambridge, págs 16-31.

HAUSMAN, J.A. (1978): "Specification test in econometrics". *Econometrica*. 46: 1251-1271.

HELPMAN, E. y KRUGMAN, P. (1985): *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition and the International Economy*. Cambridge, MIT Press.

KRUGMAN, P. (1995): *Growing World Trade: Causes and Consequences*. Economic Studies Program, The Brookings Institution, 26, pp. 327-377

LIMAO, N. y VENABLES, A.J. (2007): "Infraestructuras, Desventajas Geográficas, Costes de Transporte y Comercio", *Información Comercial Española*, 834, pp. 23-43.

LINNEMANN, H. (1966): *An Econometry Study of Internatioanl Trade Flow*. Amsterdam, North-Holland.

MARTÍNEZ, I, CANTAVELLA, M. y FERNÁNDEZ, J.I. (2003): "Estimación y Aplicaciones de una Ecuación de Gravedad para el Comercio Atlántico de la Unión Europea", *Información Comercial Española,* 806, pp. 23-32.

MÁTYÁS, L. (1997): "Proper Econometric Specification of the Gravity Model", *The World Economy*, 20, pp. 363-368.

MICCO, A., STEIN, E. y ORDONEZ G. (2003) "The Currency Union Effect on Trade: Early Evidence From EMU", *Economic Policy*, 18, pp. 316-356.

MINONDO, A. (2007): "The Disappearance of the Border Barrier in some European Union Countries' Bilateral Trade», *Applied Economics*, 39, pp. 119-124

NILSSON, L. (2000): "Trade Integration and the EU Economic Membership Criteria", *European Journal of Political Economy*, 165, pp. 807-827.

NITSCH, V. (2000): "National Borders and International Trade: Evidence from the European Union", *Canadian Journal of Economics*, 33, pp. 1091-1105.

PÖYHÖNEN, P. (1963): "A Tentative Model for the Volume of Trade between Countries", *Weltwirtschaftliches Archiv*, 90, pp. 93-99.

ROSE, A.K. (2000): "One Money, one Market: the Effect of Common Currencies on Trade", *Economic Policy*, 30, pp. 7-46.

SANSO, M., CUAIRAN, R. y SANZ, F. (1989): "Flujos Bilaterales de Comercio Internacional, Ecuación de Gravedad y Teoría Heckscher-Ohlin", *Investigaciones Económicas*, 13, pp. 155-166.

SANSO, M., CUAIRAN, R. y SANZ, F. (1990): "Flujos Españoles de Comercio Internacional y Ecuación de Gravedad. Una Aplicación para el Período 1960-1985", *Investigaciones Económicas*, 14, pp. 425-456.

TINBERGEN, J. (1962): *Shaping the World Economy: Suggestions for an International Economic Policy*. New York, The Twentieth Century Fund.

VOICU, A.M. y HORSEWOOD, N.J. (2007): "El Comercio en Europa Central y Oriental: ¿Han Convertido en Obsoletos los Costes de Transporte a los Modelos Gravitacionales?", *Información Comercial Española*, 834, pp. 113-131.

ESTIMACIÓN DE LA DISTRIBUCIÓN DE LA RENTA EN ESPAÑA: UNA APROXIMACIÓN SEMIPARAMÉTRICA

FEDERICO PALACIOS-GONZÁLEZ
E-mail: fpalacio@ugr.es
ROSA MARÍA GARCÍA-FERNÁNDEZ
E-mail: rosamgf@ugr.es
Departamento de Métodos Cuantitativos para la Economía y la Empresa
Facultad de Ciencias Económicas y Empresariales

Resumen

En este trabajo, se analiza el comportamiento de las mixturas de distribuciones Lognormales en su aplicación al problema de la distribución de la renta y de la localización de subpoblaciones según la metodología propuesta por Flachaire y Nuñez (2007). El objetivo principal es valorar el comportamiento de tal metodología aplicándola a datos reales de ingreso de los hogares españoles en los años 2000 y 2006.

Clasificación JEL: C-14, D-31
Palabras clave: Distribución de la renta, mixtura de Lognormales, Probit.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

In this paper, a mixture of Lognormal distributions is applied to estimate the income distribution and to locate sub-populations following the methodology proposed by Flachaire y Nuñez (2007). The main purpose of this work, is to evaluate this method using Spanish household's data for the years 2000 and 2006.

JEL Classification: C-14, D-31
Key word: Income distribution, Lognormal mixture, Probit.
Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUCCIÓN Y DEFINICIÓN DEL MODELO

Los efectos que los cambios sociales y tecnológicos han ejercido sobre la distribución de la renta han despertado un creciente interés en la medición de la polarización. Publicaciones recientes (Zelli and Pittau 2007, Esteban y otros 2007, Duclos y otros 2004, entre otros) cuestionan el hecho que estos cambios no favorecen por igual a los distintos grupos de renta, aumentando la distancia entre los grupos más y menos favorecidos y disminuyendo el peso de la clase media.

El concepto de polarización está relacionado con la existencia de grupos homogéneos antagónicos entre sí (véanse Esteban y Ray, 1994 y Wolfson, 1994 para una definición de polarización). Por tanto, para analizar la polarización es necesario disponer de herramientas estadísticas que faciliten la localización de subpoblaciones en la distribución del ingreso. En esta línea se encuadran los trabajos de Flachaire y Nuñez (2007), Pittau y Zelli (2006). En estos trabajos, adoptando un enfoque semi-paramétrico, se modeliza la distribución de la renta mediante mixturas finitas de distribuciones de probabilidad (McLachlan y Peel, 2000) y se asume por construcción, que cada componente de la mixtura podría representar una

subpoblación, aunque como se expone en la sección de conclusiones, los resultados introducen un cierto grado de duda en la veracidad de tal afirmación.

El principal objetivo de este trabajo es valorar el comportamiento de la metodología desarrollada por Flachaire y Nuñez (2007) para estimar y detectar subpoblaciones en la distribución de la renta. A continuación se describe la metodología propuesta por estos autores. En la segunda sección de este trabajo se aplica a datos de renta de los hogares españoles y por último en la sección de conclusiones se realiza una revisión de la misma.

Siguiendo a Flachaire y Nuñez (2007), en este artículo se modeliza la distribución de la renta de los hogares españoles mediante la siguiente mixtura finita de distribuciones Lognormales

$$f(y) = \sum_{k=1}^{K} p_k \Lambda(y; \mu_k, \sigma_k)$$

donde *K* el número de elementos de la mixtura, cada uno de ellos representando una subpoblación homogénea de tamaño p_k y $\Lambda(.;\mu,\sigma)$ es la distribución Lognormal de parámetros μ y σ .

Sea X un vector de características o variables demográficas asociadas al sustentador principal de los hogares en estudio. Supongamos que el ingreso de los hogares con vector de características X se distribuye según la mixtura

$$f(y|\mathbf{X}) = \sum_{k=1}^{K} p_k(\mathbf{X})\, \Lambda(y; \mu_k, \sigma_k)$$

donde $p_k(X)$ es la probabilidad de que el hogar con vector de características X pertenezca a la sub-población *k*.

El modelo a estimar viene dado por la siguiente expresión

$$f(y_i|X_i) = \sum_{k=1}^{K} p_{ik}\varphi(y_i; \mu_k, \sigma_k) \quad (1)$$

donde

$$p_{ik} = \Phi(\gamma_k - X_i\beta) - \Phi(\gamma_{k-1} - X_i\beta) \quad (2)$$

siendo $\Phi(.)$ la función de distribución de una Normal de media cero y varianza 1. Por tanto, para cada hogar, la probabilidad de pertenecer al grupo k-ésimo viene dada por la probabilidad de que una variable con distribución $N(X\beta_i, 1)$ pertenezca al intervalo $(\gamma_{k-1}, \gamma_k]$.

Los parámetros del modelo, que combina una mixtura de distribuciones Lognormales con un modelo probit ordenado, se estiman por máxima verosimilitud.
El logaritmo de la función de verosimilitud asociada al modelo a estimar viene dada por

$$\ell_n(\theta, y) = \sum_{i=1}^{n} [\log f(y_i|X_i])$$

donde $\theta = (\mu, \sigma, \gamma, \beta)$ siendo $\mu = (\mu_1, \ldots, \mu_K)$, $\sigma = (\sigma_1, \ldots, \sigma_K)$, $\gamma = (\gamma_1, \ldots, \gamma_{K-1})$ y $\beta = (\beta_1, \beta_2, \ldots, \beta_l)$.

El vector de parámetros θ se estima aplicando en primer lugar el algoritmo EM (Dempster y otros, 1977) y posteriormente el algoritmo Newton-Raphson (para más detalles véase Flachaire y Nuñez, 2007).
El modelo estimado proporciona las probabilidades de pertenencia del hogar a cada grupo de ingresos en función de las características del hogar. Concretamente,

$$\hat{p}_{i,k} = \Phi\left(\hat{\gamma}_k - X_i\hat{\beta}\right) - \Phi\left(\hat{\gamma}_{k-1} - X_i\hat{\beta}\right) = P(\hat{\gamma}_{k-1} < Z < \hat{\gamma}_k) \qquad (3)$$

donde $Z \to N(X_i\hat{\beta}, 1)$. Si un beta estimado es positivo, entonces, $X_i\hat{\beta}$ aumenta cuando la variable que le acompaña crece, ceteris paribus. Supongamos que inicialmente $\gamma_l < X_i\hat{\beta} < \gamma_{l+1}$. Si $X_i\hat{\beta}$ se desplaza hacia la derecha, entonces $P(\hat{\gamma}_{k-1} < Z < \hat{\gamma}_k)$ disminuye para todo $k < l$ y aumenta para todo $k > l + 1$. La probabilidad del intervalo (γ_l, γ_{l+1}) crece mientras $X_i\hat{\beta}$ esté por debajo del centro del intervalo y comienza a disminuir cuando se sobrepasa dicho punto central. En definitiva un desplazamiento de $X_i\hat{\beta}$ hacia mayores cotas indica un aumento de la probabilidad de que el hogar pertenezca a los grupos de renta más altos. Consecuentemente un beta positivo significa un aumento de las probabilidades de los grupos de renta altos cuando la variable a la que acompaña crece. Un beta negativo se interpreta en sentido contrario. Todo ello sin olvidar que, con variables ficticias, cada beta representa el efecto diferencial entre la modalidad representada por dicha variable ficticia y la que se ha retirado para evitar la multicolinealidad exacta.

El número de componentes de la mixtura puede fijarse a priori o bien aplicando los criterios de selección de modelos existentes en la literatura especializada. En este trabajo se aplica el criterio BIC (Schwarz, 1978):

$$BIC = -2\ell_n(\hat{\theta}, y) + (3K - 1 + l)log(n)$$

seleccionándose el modelo con menor BIC.

Después de estimar los parámetros de la mixtura y seleccionar el número de componentes de la misma, se obtiene la función de densidad estimada para el total de la población como sigue:

$$\hat{f}(y) = \sum_{k=1}^{K} [\bar{p}_k \Lambda(y; \hat{\mu}_k], \hat{\sigma}_k)$$

donde

$$\bar{p}_k = \frac{1}{n}\sum_{i=1}^{n} \hat{p}_{ik}$$

2. APLICACIÓN EMPÍRICA

En esta sección, se aplica el modelo definido previamente a datos de ingresos de los hogares españoles para los años 2000 y 2006. Los datos utilizados proceden de la Encuesta Continua de Presupuestos Familiares (Base 1997, fichero longitudinal) y de la Encuesta de Presupuestos Familiares (Base 2006). Ambas encuestas son realizadas por el Instituto Nacional de Estadística (INE). La ECPF dejó de realizarse en el año 2005 siendo sustituida por la EPF, por lo que en el año 2000 se utiliza la ECPF y en el año 2006 la EPF. Estas encuestas proporcionan estimaciones a nivel microdato de los gastos e ingresos de los hogares españoles así como una amplia información relativa a las características de los hogares que forman la muestra. Concretamente, se han extraído del fichero de hogares los ingresos netos totales del hogar, el tamaño del municipio en el que reside el hogar, la región a la que pertenecen los hogares encuestados (NUTS) y las siguientes variables demográficas relativas al sustentador principal del hogar:

sexo, estudios completados y principal fuente de ingresos (véase Tabla 1). Cada una de estas variables cualitativas se introduce en el modelo utilizando una variable ficticia por cada modalidad ("valor cualitativo") de cada variable. Para evitar la multicolinealidad exacta se ha introducido una variable ficticia menos en cada categoría. Por tanto cada parámetro estimado representa el efecto diferencial entre la modalidad representada y la que se ha retirado (aparece en cursiva en la Tabla 1) para evitar la multicolinealidad. Los tamaños muestrales son de 9601 y 18734 hogares para los años 2000 y 2006 respectivamente.

Tabla 1. Variables demográficas

ESTRATO (ECPF) 1 Municipio capital de provincia 2 Municipio no capital con más de 100000 habitantes 3 Municipio no capital con más de 50000 y menos de 100000 habitantes 4 Municipio no capital con más de 20000 y menos de 50000 habitantes 5 Municipio no capital con más de 10000 y menos de 20000 habitantes 6 *Municipio no capital con menos de 10000 habitantes*
ESTRATO (EPF) 1 Municipio de 100.000 habitantes o más 2 Municipio con 50.000 o más y menos 100.000 habitantes 3 Municipio con 20.000 o más y menos de 50.000 habitantes 4 Municipio con 10.000 o más y menos de 20.000 habitantes 5 *Municipio con menos de 10.000 habitantes*
SEXO (sexo sustentador principal) 1 *Hombre* 2 Mujer
NIVEL DE ESTUDIOS (alcanzado por el sustentador principal) 1 *Educación primaria o primer ciclo de secundaria* 2 Segundo ciclo de educación secundaria 3 Educación superior
PRINCIPAL FUENTE DE INGRESOS (sustentador principal) 1 Trabajo por cuenta propia y rentas de la propiedad y del capital 2 Trabajo por cuenta ajena 3 *Pensiones, subsidios y otras prestaciones e ingresos regulares*
NUTS 1 Noroeste (Comunidades de Galicia, Asturias y Cantabria). 2 Noreste (Comunidades de País Vasco, Navarra, La Rioja y Aragón). 3 Madrid (Comunidad de Madrid). 4 Centro (Comunidades de Castilla y León, Castilla-La Mancha y Extremadura). 5 Este (Comunidades de Cataluña, Comunidad Valenciana e Islas Baleares). 6 Sur (Comunidades de Andalucía, Región de Murcia y Ceuta y Melilla). 7 *Canarias*

Los ingresos de los hogares se corrigen utilizando la escala de equivalencia modificada de la OCDE convirtiendo la renta total de los hogares en renta equivalente. Adicionalmente, en cada año se normalizan los ingresos dividiéndolos por su media aritmética para trabajar en términos relativos y posteriormente se calcula el logaritmo del ingreso real equivalente normalizado.

El modelo se ha estimado aplicando la metodología descrita en el anterior apartado para los años 2000 y 2006. Las estimaciones se han realizado para $K = 3;\ 5\ \mathbf{y}\ 8$. Una vez estimado el modelo se aplica el criterio de selección BIC (Schwarz, 1978):

$$BIC = -2\ell_n(\hat{\theta}, y) + (3K - 1 + l)Log(n)$$

y se selecciona el modelo con menor BIC que para el año 2000 y 2006 se corresponde con $K = 5$.

Los parámetros estimados de la mixtura se muestran en la Tabla 2 y las Figuras 1 y 2 muestran las estimaciones de las mixturas y de cada una de sus componentes para los años 2000 y 2006 respectivamente.

Figura 1. Distribución de la renta en 2000

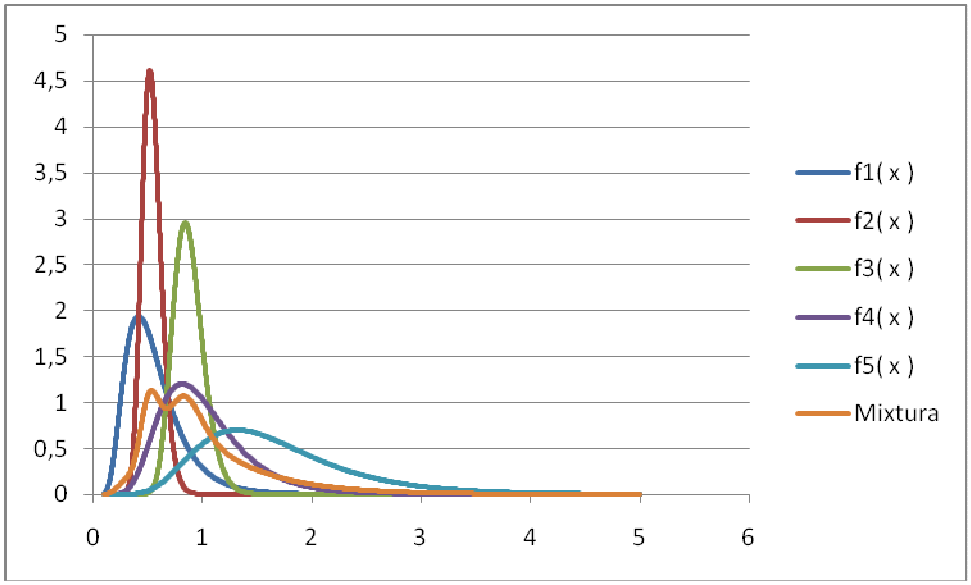


Figura 2. Distribución de la renta 2006

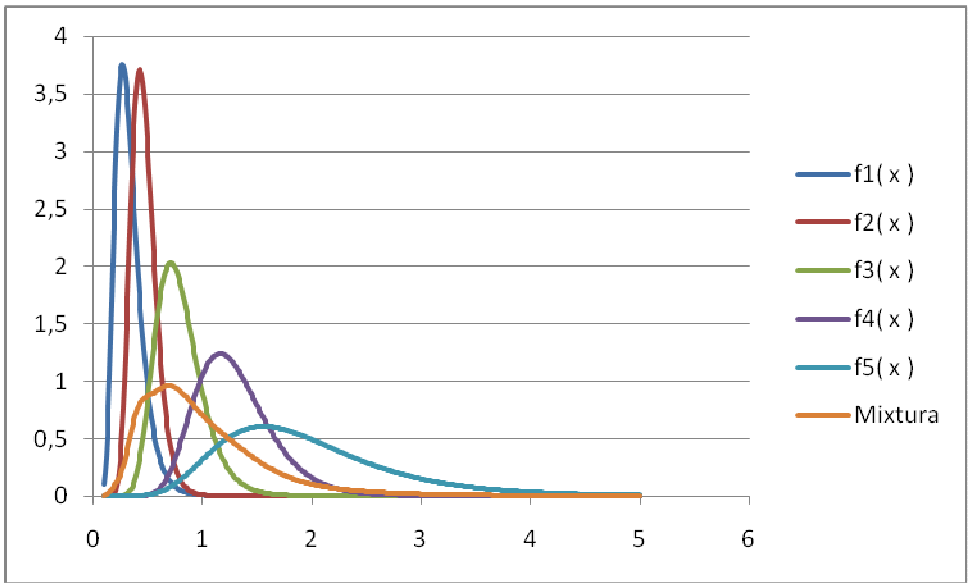


Tabla 2. Parámetros estimados, valores esperados y varianzas Lognormales

	2000		2006	
	parámetros	v. esperados y varianzas	parámetros	v. esperados y varianzas
$\hat{\mu}_1$	-0,67514522	0,56282368	-1,17666083	0,33000222
$\hat{\mu}_2$	-0,6283252	0,54074018	-0,79207751	0,4666017
$\hat{\mu}_3$	-0,14712003	0,87404838	-0,26983099	0,79124627
$\hat{\mu}_4$	-0,06108681	1,01077397	0,22216826	1,29414058
$\hat{\mu}_5$	0,43706122	1,67746363	0,5960617	1,95924655
$\hat{\sigma}_1$	0,44800971	0,26534828	0,3687952	0,12596113
$\hat{\sigma}_2$	0,16437042	0,08948543	0,24412385	0,115627
$\hat{\sigma}_3$	0,15811697	0,1390702	0,2671515	0,2152109
$\hat{\sigma}_4$	0,37895424	0,39720914	0,26712756	0,35196033
$\hat{\sigma}_5$	0,40055384	0,69978848	0,39114776	0,79662343
$\bar{p}_1$	0,10482513		0,03527939	
$\bar{p}_2$	0,14918884		0,16444325	
$\bar{p}_3$	0,14935048		0,39904314	
$\bar{p}_4$	0,40522541		0,29009176	
$\bar{p}_5$	0,19141013		0,11114247	

Nótese que los parámetros de las distribuciones Lognormales son por definición el valor esperado y la varianza de la variable transformación logarítmica de la renta relativa. Esta variable transformada, toma valores negativos cuando los datos de renta son inferiores a la media y, consecuentemente, la renta relativa es inferior a 1. Esto explica que los parámetros estimados para las distribuciones Lognormales asociadas a los grupos de menor renta sean negativos. No obstante toda variable Lognormal toma valores exclusivamente positivos aunque su parámetro μ pueda ser negativo. En las columnas segunda y cuarta de la Tabla 2 aparecen los valores esperados y la varianza de cada una de las componentes Lognormales de la mixtura. Para cada Lognormal de parámetros μ_k y σ_k^2 se calculan dichos valor esperado y varianza mediante las expresiones (véase Johnson y Kotz, 1970 pp. 115).

$$E[Y_k] = exp\left(\mu_k + \frac{1}{2}\sigma_k^2\right) (4)$$

$$Var(Y_k) = e^{2\mu_k} e^{\sigma_k^2}\left(e^{\sigma_k^2} - 1\right) (5)$$

Si comparamos las distribuciones de los años 2000 y 2006 se observa que las dos modas que aparecen en la mixtura correspondiente al año 2000 se suavizan en el año 2006. Este hecho puede explicarse por el aumento del tamaño de la segunda y tercera componente de la mixtura en el año 2006 en detrimento del tamaño del resto de componentes. Tanto las Figuras 1 y 2 como los valores $\hat{\mu}_1, \ldots \hat{\mu}_5$ nos permiten observar por un lado, un desplazamiento hacia la izquierda de las distribuciones Lognormales de las tres primeras componentes de la mixtura en el año 2006. Por otro lado, las distribuciones Lognormales de las dos últimas componentes de la mixtura se desplazaron hacia la derecha en el año 2006. Adicionalmente disminuyen los valores esperados asociados a las tres primeras componentes ocurriendo lo contrario en las dos últimas. Si esta separación de las componentes de la mixtura se interpreta como una separación de grupos de ingresos se podría concluir que la polarización aumentó del año 2000 al 2006. No obstante, este resultado debe matizarse teniendo en cuenta el incremento que experimenta el tamaño de la clase media pasando del 15% en 2000 al 40% en 2006.

En las Tablas 3 y 4 se presentan las estimaciones de los parámetros de las variables ficticias y sus respectivas desviaciones típicas. Estos resultados nos permiten analizar la posición de los hogares en la distribución de la renta según las variables que aparecen en la Tabla 1.

Tabla 3. Parámetros estimados $\hat{\beta}_j$ asociados a las variables X_j para el año 2000

2000	$\hat{\beta}_j$	Desv. típica	Estadístico t	p-valor
$\hat{\beta}_{muni1}$	0,53243905	0,0299439	17,7812192	2,5886E-69
$\hat{\beta}_{muni2}$	0,21028895	0,0471389	4,46104926	1,6498E-05
$\hat{\beta}_{muni3}$	0,30624138	0,0487095	6,28709754	6,7517E-10
$\hat{\beta}_{muni4}$	0,24019689	0,0405087	5,92951377	6,2871E-09
$\hat{\beta}_{muni5}$	0,15741096	0,0419524	3,75213238	0,00035275
$\hat{\beta}_{mujer}$	-0,14710286	0,0282757	-5,20244791	4,0151E-07
$\hat{\beta}_{edu1}$	1,35941724	0,0347126	39,1620692	0
$\hat{\beta}_{edu2}$	2,84670243	0,0349203	81,5199879	0
$\hat{\beta}_{trab1}$	0,95429454	0,0255365	37,3698253	4,025E-285
$\hat{\beta}_{trab2}$	1,10062952	0,0340711	32,3039032	4,387E-217
$\hat{\beta}_{nuts1}$	0,56524542	0,0584313	9,67367529	9,8417E-22
$\hat{\beta}_{nuts2}$	0,74979129	0,0582847	12,8642902	2,9263E-37
$\hat{\beta}_{nuts3}$	0,83862374	0,0640048	13,1025132	1,3759E-38
$\hat{\beta}_{nuts4}$	0,2652	0,0581552	4,56021138	1,0347E-05
$\hat{\beta}_{nuts5}$	1,01084845	0,055294	18,2813406	4,1101E-73
$\hat{\beta}_{nuts6}$	-0,1068933	0,0570804	-1,87267959	0,12228621
$\hat{\gamma}_1$	-0,05140342	0,0580682	-0,88522495	0,75211514
$\hat{\gamma}_2$	0,76833942	0,0582594	13,1882481	4,5193E-39
$\hat{\gamma}_3$	1,39217609	0.0588755	23,6461022	6,94E-120
$\hat{\gamma}_4$	3,52498408	0.0601907	58,5636	0

Tabla 4. Parámetros estimados $\hat{\beta}_j$ asociados a las variables X_j para el año 2006

2006	$\hat{\beta}_j$	Desv. típica	Estadístico t	p-valor
$\hat{\beta}_{poblac1}$	0,780202325	0,0396916	19,6566106	7,34446E-85
$\hat{\beta}_{poblac2}$	0,841534644	0,0347179	24,23921504	1,6036E-127
$\hat{\beta}_{poblac3}$	0,763142256	0,03423	22,29454443	2,1716E-108
$\hat{\beta}_{poblac4}$	0,297036187	0,0342724	8,666921108	9,5921E-18
$\hat{\beta}_{mujer}$	-0,19648144	0,0204233	-9,62045521	1,47292E-21
$\hat{\beta}_{edu1}$	0,682780091	0,0234529	29,11282149	5,6543E-182
$\hat{\beta}_{edu2}$	1,549008893	0,0209711	73,86397914	0
$\hat{\beta}_{trab1}$	1,012651013	0,0201923	50,15035497	0
$\hat{\beta}_{trab2}$	1,141743309	0,0261622	43,64095178	0
$\hat{\beta}_{nuts1}$	0,403928151	0,042969	9,400454997	1,20856E-20
$\hat{\beta}_{nuts2}$	0,573151786	0,049163	11,65819388	5,3489E-31
$\hat{\beta}_{nuts3}$	0,634776099	0,0493823	12,8543243	2,34852E-37
$\hat{\beta}_{nuts4}$	0,122725317	0,0415267	2,955335165	0,006254381
$\hat{\beta}_{nuts5}$	0,5832771	0,0405638	14,37925194	2,47714E-46
$\hat{\beta}_{nuts6}$	-0,02426301	0,0419503	-0,57837514	0,873964325
$\hat{\gamma}_1$	-0,36976134	0,0504011	-7.33637439	4,5686E-13
$\hat{\gamma}_2$	0,85650279	0,0507707	16,8700213	4,4013E-63
$\hat{\gamma}_3$	2,38868556	0,051403	46,4697695	0
$\hat{\gamma}_4$	3,82427416	0,0523674	73,0277647	0

A grandes rasgos, cabe destacar la alta significatividad de los parámetros estimados, exceptuando el parámetro $\hat{\beta}_{nuts}$ el resto son significativos en 2000 y 2006. Todos los parámetros significativos son positivos excepto el asociado al sexo del sustentador principal. Teniendo en cuenta que cada parámetro estimado representa el efecto diferencial entre la modalidad representada y la que se ha retirado, se tiene que los hogares cuyo sustentador principal es mujer tienen menor probabilidad de pertenecer a los grupos de renta más altos. Este resultado es coherente con la diferenciación salarial desfavorable a la mujer existente en el mercado laboral español. La situación de los hogares cuyo cabeza de familia es mujer empeora decreciendo el parámetro estimado de -0,1471 en 2000 a -0,1965 en 2006. Con respecto al resto de parámetros, en el año 2000 los hogares que residen en la Comunidades de Cataluña, Valencia e Islas Baleares (nuts5), cuyo sustentador principal tiene mayor nivel de estudios y sus ingresos proceden del trabajo por cuenta ajena tienen mayor probabilidad de pertenecer a grupos de ingresos elevados. En el año 2006, se repite este patrón pero se produce un cambio en la región de residencia siendo los hogares residentes en la Comunidad de Madrid (nuts 3) los que tienen mayor probabilidad de pertenecer a niveles de renta altos.

No obstante, se quiere hacer notar que, aunque, por construcción, cada componente de la mixtura representa una subpoblación. La existencia real de tal subpoblación puede, ocasionalmente, llegar a ser dudosa. Atendiendo a las Figuras 1 y 2, cabe destacar que una mixtura de 5 distribuciones Lognormales no significa necesariamente 5 modas (en la densidad resultante) que aglutinen claramente grupos de población homogéneos entre sí. Este hecho también se observa en trabajos como en el de Pittau y Zelli, 2006 y Flacharie y Nuñez, 2007 y es señalado por sus propios autores. Parece que la contribución principal de los componentes de la mixtura (y el número de ellas) es proporcionar (a la densidad resultante) la flexibilidad y adaptabilidad requerida por los datos. Esto no es incompatible con la localización de subpoblaciones realmente existentes pero tampoco garantiza que cada componente de la mixtura represente una subpoblación con entidad propia.

3. CONCLUSIONES

La razón de este trabajo es un análisis del comportamiento de las mixturas de distribuciones lognormales en su aplicación al problema de la distribución de la renta y de la localización de subpoblaciones. Para ello se ha aplicado tal metodología a datos reales de ingreso de los hogares españoles en los años 2000 y 2006. Los resultados principales se encuentran en el epígrafe precedente.

Se quieren destacar en esta sección las incidencias que han sido observadas, durante el proceso de estimación, respecto al comportamiento de esta metodología en el proceso de estimación. Estas son las siguientes

1.- Lentitud exasperante de los algoritmos de ajuste. Es verdad que los tamaños muestrales son elevadísimos, pero el número de parámetros del modelo no es excesivamente alto

2.- Las estimaciones de los valores esperados, de la variable "logaritmo de los ingresos relativos", para los distintos grupos, obtenidas mediante la expresión

$$\hat{m}_k = \frac{1}{n}\sum_{i=1}^{n} \hat{p}_{i,k} y_i$$

(donde y_i el logaritmo de los ingresos relativos para el hogar $i-$ésimo) no resultan ordenadas ni están próximas a los valores $\hat{\mu}_k$ como cabría esperar. Ello es fruto del propio método en el que, al ajustar simultáneamente el probit y la mixtura, el proceso de definición de grupos no concluye antes de aplicar el probit y por tanto la selección de grupos no se hace exclusivamente en función de los niveles de ingresos de los hogares sino que también están afectados por los perfiles sociológicos que definen las variables explicativas X del modelo.

3.- El coeficiente de determinación de la variable ingresos estimados

$$\hat{E}[Y_i|X_i] = P_i = \sum_{k=0}^{K} \hat{p}_{i,k}\hat{\mu}_k$$

por la variable ingresos observados (aunque significativo si se somete a un test de R^2 crítico siendo $F = \frac{R^2/k-1}{1-R^2/n-k} = 118{,}967$ y $F^{0.05}_{34.9566} = 1{,}4307$) es relativamente bajo: no llega al 30%

4.- Las componentes de las mixturas presentan un acusado solapamiento originado por valores demasiado altos de algunos parámetros $\hat{\sigma}_k^2$. Esto hace, por una parte, confusa la identificación de individuos en los distintos grupos y, por otra parte, que los valores esperados de las distintas variables Lognormales (calculados según las expresiones 4 y5) no resulten ordenados aún siéndolo sus parámetros $\hat{\mu}_k$. Esto refuerza la idea de que la función principal de tales componentes es proporcionar flexibilidad y adaptabilidad a la mixtura resultante más que definir subpoblaciones de ingresos homogéneos.

5.- Si se utiliza la parte del modelo (presentado por Flacharie y Nuñez, 2007) que exclusivamente atañe a la mixtura y se aplica el algoritmo EM usando las siguientes expresiones que se extractan del citado trabajo

$$\hat{p}_{i,k} = \frac{p_{i,k}\varphi(y_i;\mu_k,\sigma_k)}{\sum_{j=1}^{K} p_{i,j}\varphi(y_i;\mu_j,j)} \qquad N_k = \sum_{i=1}^{n} \hat{p}_{i,k}$$

$$\hat{\mu}_k = \frac{1}{N_k}\sum_{i=1}^{n} \hat{p}_{i,k}\, y_i \qquad \hat{\sigma}_k = \sqrt{\frac{1}{N_k}\sum_{i=1}^{n} \hat{p}_{i,k}\,(y_i - \hat{\mu}_k)^2}$$

entonces, se produce una clasificación de los hogares de la muestra, atendiendo exclusivamente al volumen de ingresos, (y lógicamente al margen del perfil sociológico del hogar) que es la resultante de dividir el rango de rentas mediante el correspondiente juego de cuantiles. Por ejemplo para cuatro grupos el resultado del algoritmo EM es que el grupo 1 es el de los hogares con ingresos que no superan el primer cuartil, el segundo el de ingresos entre el primer cuartil y la mediana etc. Consecuentemente la muestra es sistemáticamente partida en grupos de igual tamaño.

6.- Aunque el método es comúnmente aceptado como un instrumento válido para la obtención de subpoblaciones en el estudio de la distribución de la renta, creemos que necesita una revisión detallada para resolver y explicar las dificultades que hemos experimentado y comentado en los apartados precedentes.

4. BIBLIOGRAFÍA

DEMPSTER, A.P., LAIRD N.M., RUBIN, D.B., (1977), "Maximum likelihood from incomplete data via the EM algorithm (with discussion)", J. Roy. Statist. Soc.:Ser. B 39, pp. 1-38.

DUCLOS, J. Y., ESTEBAN J.M., y RAY D., (2004), "Polarization: Concepts, Measurement, Estimation", Econometrica, 72 (6) pp. 1737-1772.

ESTEBAN J.M., y RAY D., (1994), "On the Measurement of Polarization", Econometrica, 62 (4) pp. 819-852.

ESTEBAN J.M., GRADÍN C. y RAY D., (2007), "Extensions of a Measure of Polarization
OCDE Countries", Journal of Economic Inequality, 5, (1) , pp. 1-19.

FLACHAIRE, E. y NUÑEZ, O. (2007), "Estimation of the income distribution and detection of subpopulations: An explanatory model", Computational Statistics & Data Analysis, 51, pp. 3368-3380.

JOHNSON, N. y KOTZ, S. (1970), "Distributions in Statistics", John Wiley & Sons, New York.

MCLACHLAN, G.P. y PEEL. D. (2000), "Finite Mixture Models", Willey Series in Probability and Mathematical Statistics: Applied Probability and Statistic Section.

PITTAU G. y ZELLI R. (2006), "Trends in income distribution in Italy: A non parametric and a semi-parametric analysis", Journal of Income Distribution, 15, pp.90-118.

SCHWARZ, G. (1978), "Estimating the dimension of a model", Ann. Statist. 6, pp. 461-464.

WOLFSON, M. C. (1994) "When Inequalities Diverge?", American Economic Review, 84, pp.353-5.

ZELLI R. y PITTAU M.G., (2007), "Exploring patterns of income polarization using siZer", Journal of Quantitative Economics, 5, pp. 101-111.

MODELAÇÃO LONGITUDINAL DA SATISFAÇÃO LABORAL: *STRUCTURAL EQUATION MODELS VERSUS MULTI-PROCESS MODELS*

PATRÍCIA SERRA
E-mail: pghsa@iscte.pt
MARIA DE FÁTIMA SALGUEIRO
E-mail: fatima.salgueiro@iscte.pt
ISCTE – Instituto Superior de Ciências do Trabalho e da Empresa

Resumo

A satisfação laboral tem-se tornado, ao longo dos anos, um assunto de grande interesse por parte de Economistas. O European Community Household Panel (ECHP) é uma pesquisa longitudinal com informação indivídual e sobre o agregado familiar, desde 1994. Tem também várias medidas de satisfação laboral. Para além de uma questão sobre a satisfação laboral global, os indivíduos são inquiridos sobre o seu grau de satisfação com seis dimensões: rendimento, segurança do trabalho (factores extrínsecos), tipo de trabalho, número de horas de trabalho semanais, horário de trabalho e ambiente/condições de trabalho (factores intrínsecos). Para modelar a satisfação laboral a literatura considera, usualmente, a satisfação laboral global ou a construção de um score total. Modelos probit e regressões lineares múltiplas têm sido usados.
O presente artigo compara a modelação longitudinal da satisfação laboral através de cross-lagged SEM de multi-process models. São usados dados do ECHP sobre Portugal, das vagas 5 a 8.

Palavras-chave: ECHP, dados longitudinais, satisfação laboral, SEM
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

Job satisfaction has been an area of increasing research over the last years by Economists. The European Community Household Panel (ECHP) is a longitudinal survey which includes information about individuals and their families, since 1994. It also includes several job satisfaction measures. In addiction to a question on overall job satisfaction, respondents have been asked to rate their satisfaction levels with six domain dimensions: earnings, job security (extrinsic job facets), type of work, number of working hours, working times and working conditions/environment (intrinsic job facets). Statistical approaches adopted in the literature to model job satisfaction often consider either using the overall measure or constructing a total score. Ordered probit models and multiple regression models have been used.
The current paper compares cross-lagged SEM and multi-process model as alternative approaches to longitudinal modelling of job satisfaction. Data from waves 5 to 8 of the ECHP for Portugal are used.

Keywords: ECHP, longitudinal data, job satisfaction, SEM.
Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUÇÃO

A satisfação laboral (SL) tem-se tornado, ao longo dos anos, um assunto de grande interesse por parte de Economistas e investigadores em Ciências Sociais.

Vários estudos foram realizados para avaliar o efeito de características pessoais e características do trabalho na satisfação laboral. Vieira et al. (2005) distinguiu a qualidade do trabalho com salários baixos e com salários altos, onde concluiu que os indivíduos com salários baixos apresentam menor qualidade de trabalho, e os níveis de satisfação com os determinantes, tais como, a remuneração horária, os benefícios disponibilizados pela empresa e o subsídio de

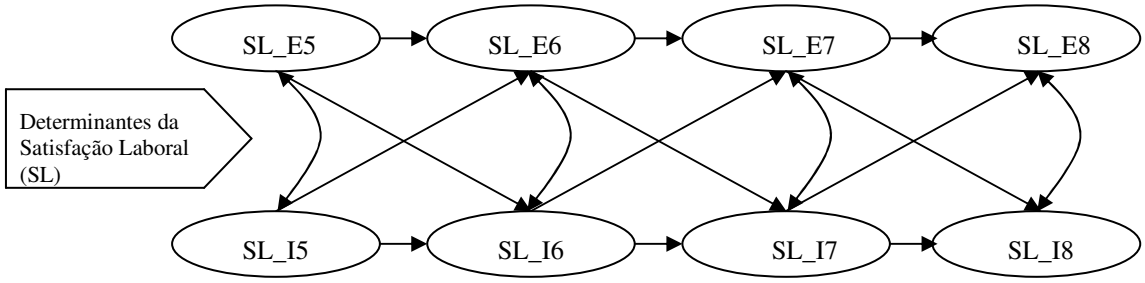
saúde e habitação, variam nestes dois segmentos. Clark e Oswald (1996) e Drakopoulos e Theodossiou (1997) mostram também que a satisfação laboral decresce com o aumento do número de horas de trabalho. Clark et al. (1996) estabeleceram haver uma relação em U entre a idade e a satisfação laboral. A relação entre o género e a satisfação laboral não é clara. Clark (1997) afirma que, no Reino Unido, as mulheres estão mais satisfeitas do que homens, no entanto, Sousa-Poza (2003) conclui que a satisfação laboral das mulheres tem vindo a diminuir, enquanto os homens se mantém constantes. Ter a percepção de um bom estado de saúde também aumenta a satisfação laboral, tal como Vieira (2005) conclui. Um estudo realizado em Espanha, por García (2006) considera que o estado civil e a vida familiar não afectam o grau de satisfação laboral, no entanto, o facto de ter filhos é considerado relevante na satisfação laboral.

Rose (2001) pressupõe haver diferenças consideráveis entre os aspectos financeiros e materiais do trabalho – os factores extrínsecos – e os aspectos que são mais simbólicos ou qualitativos – factores intrínsecos. Por exemplo, os factores extrínsecos são os determinantes como o rendimento, promoção ou segurança do trabalho, enquanto os factores intrínsecos são aspectos, tais como, número de horas de trabalho, relação com o supervisor, o tipo de trabalho ou o uso de iniciativa. Aparentemente, os indivíduos quando submetidos a uma avaliação de satisfação destes aspectos, são mais rigorosos na avaliação dos factores extrínsecos, pelo que os níveis de satisfação destes factores são geralmente mais baixos que os níveis de satisfação com os factores intrínsecos.

Para modelar a satisfação laboral, a literatura considera frequentemente a medida global de satisfação (*overall*) ou a construção de um *score* total, por exemplo, através da soma de itens, ou da média dos mesmos. Modelos *probit* e regressões lineares múltiplas têm sido usados para efeitos de modelação estatística – ver Clark et al. (1996), Rose (2001), Skalli (2008), Sousa-Poza (2003), Vieira et al. (2005) e Vieira (2005).

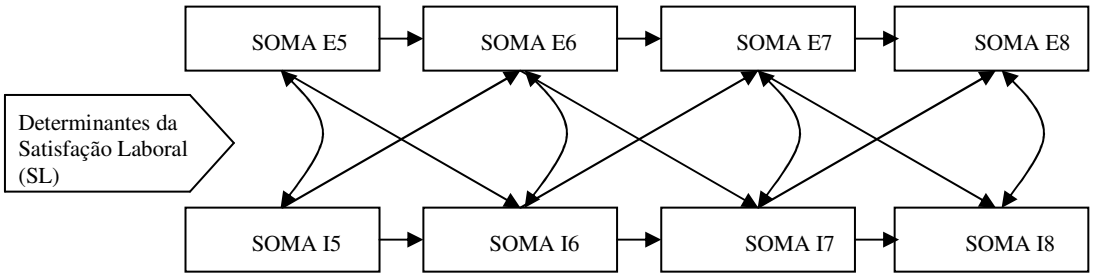
O presente artigo tem o objectivo de comparar duas metodologias para modelar a satisfação laboral e os seus determinantes. Numa primeira fase, consideram-se *cross-lagged strutural equation models* (SEM), com relações múltiplas entre variáveis independentes e dependentes, sendo que as variáveis de satisfação em cada vaga serão consideradas indicadores de medida de duas variáveis latentes, uma variável referente aos factores extrínsecos do trabalho, que estão relacionados com os aspectos materiais (rendimento e segurança do trabalho), e outra referente aos factores intrínsecos que compreendem os aspectos qualitativos (ou seja, o tipo de trabalho, o número de horas de trabalho semanais, o horário de trabalho e o ambiente/condições de trabalho). Considerar os factores extrínsecos e intrínsecos como variáveis latentes permite assim a incorporação de erros de medida. Na figura 1 está representado o modelo *cross-lagged* SEM proposto.

Figura 1: *Structural Equation Model* (SEM)



Numa segunda fase será utilizado um *multi-process model* (MPM), com a construção de dois *scores* de satisfação através da soma dos itens dos factores extrínsecos e intrínsecos do trabalho, respectivamente. Desta forma não serão incluídos no modelo os erros de medida. Na figura 2 está representado o modelo do MPM proposto.

Figura 2: *Multi-Process Model* (MPM)



O *European Community Household Panel Survey* (ECHP) é uma pesquisa longitudinal, representativa dos agregados familiares em cada país, com informação individual e sobre o agregado familiar, realizada entre 1994 e 2001. Inclui variáveis de medida da satisfação laboral, nomeadamente é pedido aos indivíduos que avaliem o seu grau de satisfação com seis dimensões, tais como o rendimento, a segurança do trabalho, o tipo de trabalho, o número de horas de trabalho, os horários de trabalho e o ambiente/condições de trabalho, medidas numa escala de *Likert* de 1 – não satisfeito a 6 – totalmente satisfeito. Isto para além de uma dimensão sobre satisfação global (*overall job satisfaction*).

Neste estudo serão usados dados do ECHP sobre Portugal, relativos aos anos de 1998 a 2001 (vagas 5 a 8). São ainda considerados possíveis determinantes da satisfação laboral, tais como o género, a idade, o nível de habilitações, o estado civil, a percepção do estado de saúde, a existência de crianças menores de 12 anos no agregado, o número de horas de trabalho por semana e o rendimento.

O software estatístico a utilizar é o LISREL 8.80 (Jöreskog e Sörbom, 2006).

O presente artigo está organizado com a seguinte estrutura, na secção 2 é descrita a amostra em estudo, na secção 3 é sumarizada a metodologia aplicada, na secção 4 são apresentados os resultados mais relevantes, terminado o artigo com uma discussão dos resultados.

2. AMOSTRA EM ESTUDO

O *European Community Household Panel* (ECHP) é uma pesquisa longitudinal, representativa dos agregados familiares em cada país, com informação individual e sobre o agregado familiar, realizada entre 1994 e 2001. A pesquisa foi aplicada a mais de 60.000 agregados e 130.000 indivíduos, sempre que possível aos mesmos indivíduos em anos consecutivos.

Os dados analisados no presente artigo são do ECHP de Portugal dos anos 1998, 1999, 2000 e 2001, respectivamente, as vagas 5,6,7 e 8.

Foram admitidos neste estudo os indivíduos que, se encontravam empregados ao longo do tempo do estudo, trabalhando mais de 15 horas por semana, sejam eles trabalhadores por conta de outrem ou por conta própria e do sector público ou privado. Os indivíduos com mais de 65 anos, e as observações com valores omissos nas variáveis de satisfação foram excluídos da amostra, resultando assim, uma amostra com 4014 observações válidas.

São usadas algumas variáveis socio-demográficas como possíveis determinantes da satisfação laboral, tais como, o género, a idade, o nível de habilitações literárias, o estado civil, a percepção do estado de saúde, a existência de crianças menores de 12 anos no agregado, o número de horas de trabalho por semana e ainda o rendimento individual.

Podemos observar que a média de idades dos indivíduos é, aproximadamente, 38 anos e em média, trabalham, aproximadamente, 44 horas semanais, ligeiramente acima das 40 horas propostas por lei. A amostra é constituída principalmente por indivíduos do sexo masculino, com cerca de 61% das observações. A nível de habilitações literárias, mais de 78% dos indivíduos não têm mais do que o nível secundário, sendo que 71% são casados e 37% tem um

ou mais filhos com menos de 12 anos. Cerca de 63% dos indivíduos consideram que, em geral, a sua saúde é boa ou muito boa, enquanto os restantes 37% a considera razoável, má ou muito má. As variáveis de rendimento individual e do agregado referem-se ao ano anterior à entrevista, ou seja, se o questionário data de 1998, então é usado o rendimento anual referente a 1997.

Os indicadores de satisfação laboral global e os seis determinantes da satisfação laboral provêm das seguintes questões:

- *Satisfaction with work or main activity?*
- *How satisfied are you with your present job in terms of earnings?*
- *How satisfied are you with your present job in terms of job security?*
- *How satisfied are you with your present job in terms of type of work?*
- *How satisfied are you with your present job in terms number of working hours?*
- *How satisfied are you with your present job in terms of working times (day time, night time, shifts, etc)?*
- *How satisfied are you with your present job in terms of working conditions/environment?*

As variáveis são medidas numa escala de tipo *Likert* de 1 – não satisfeito a 6 – totalmente satisfeito. Nas tabelas 1 e 2 pode-se observar a frequência de resposta em cada variável de satisfação, no primeiro ano da análise, 1998, e no último ano, 2001.

Tabela 1: Frequência de resposta das medidas de satisfação em 1998

	Não satisf	2	3	4	5	Total satisf
Satisfação com o rendimento	4,6%	14,7%	34,3%	39,6%	4,8%	2,0%
Satisfação com a segurança do trabalho	1,5%	4,4%	14,7%	51,9%	20,9%	6,5%
Satisfação com o tipo de trabalho	,5%	2,5%	10,2%	53,4%	27,6%	5,8%
Satisfação com o número de horas de trabalho	1,0%	3,7%	15,9%	61,8%	13,9%	3,6%
Satisfação com os horários de trabalho	,6%	2,6%	12,0%	62,7%	17,6%	4,4%
Satisfação com o ambiente/condições de trabalho	,5%	2,2%	9,9%	57,2%	24,2%	5,9%
Satisfação global	,6%	4,0%	14,0%	56,3%	20,3%	4,9%

De acordo com a tabela 1, o nível de resposta com menor frequência em todos os determinantes de satisfação é o 1 – não satisfeito, enquanto aquele com maior percentagem de respostas é o nível 4 de satisfação. É interessante verificar que a variável de satisfação com o rendimento tem uma distribuição diferente de todas as outras variantes de satisfação, havendo diferenças nas percentagens em todos os níveis de resposta, sendo que, os níveis mais baixos apresentam maior frequência de resposta e os níveis mais altos tem menor frequência. As restantes variáveis apresentam maior frequência nos níveis 4 e 5, e em relação aos extremos, é maior a percentagem de respostas nos níveis de maior satisfação do que nos níveis inferiores.

Tabela 2: Frequência de resposta das medidas de satisfação em 2001

	Não satisf	2,0	3,0	4,0	5,0	Total satisf
Satisfação com o rendimento	3,1%	12,6%	34,3%	40,4%	8,1%	1,6%
Satisfação com a segurança do trabalho	,6%	3,1%	12,3%	53,2%	25,6%	5,3%
Satisfação com o tipo de trabalho	,4%	2,2%	9,8%	52,0%	30,8%	4,8%
Satisfação com o número de horas de trabalho	,8%	3,0%	16,0%	60,5%	17,3%	2,4%
Satisfação com os horários de trabalho	,7%	2,6%	11,4%	62,6%	19,7%	3,0%
Satisfação com o ambiente/condições de trabalho	,6%	1,9%	9,8%	56,4%	26,5%	4,8%
Satisfação global	,6%	3,9%	14,4%	56,0%	21,2%	3,9%

Em 2001 (vaga 8), verifica-se novamente a diferença da distribuição da variável de satisfação com o rendimento em relação às restantes. Em geral, continua a haver maior frequência nos níveis de satisfação 4 e 5, embora, em relação ao 1998 (vaga 5), haja um decréscimo no nível 4 contrariamente ao nível 5 que admite um ligeiro aumento de frequência. Constata-se que as respostas tendem a concentrar-se nos níveis intermédios de satisfação.

Na tabela 3 pode observar-se a evolução do nível médio de resposta às variáveis de satisfação ao longo do período em análise.

Tabela 3: Médias das variáveis de satisfação laboral ao longo dos anos em estudo

	1998	1999	2000	2001
Satisfação com o rendimento	3,312	3,348	3,395	3,427
Satisfação com a segurança do trabalho	4,059	4,112	4,136	4,159
Satisfação com o tipo de trabalho	4,225	4,204	4,244	4,252
Satisfação com o número de horas de trabalho	3,951	3,950	3,955	3,979
Satisfação com os horários de trabalho	4,072	4,054	4,067	4,069
Satisfação com o ambiente/condições de trabalho	4,201	4,204	4,211	4,206
Satisfação global	4,06	4,07	4,06	4,05

Tal como foi constatado anteriormente, a variável de satisfação com o rendimento apresenta menores níveis médios de satisfação, seguida da variável de satisfação com o número de horas de trabalho semanais. De um modo geral, os níveis médios de satisfação aumentam ao longo dos anos, embora seja mais regular nos factores extrínsecos de satisfação laboral, enquanto os factores intrínsecos revelam-se ligeiramente mais inconstantes, como é o caso da satisfação com o tipo de trabalho e com os horários de trabalho que verificam um decréscimo, em termos médios, do ano 1998 (vaga 5) para 1999 (vaga 6). Já a satisfação média com o ambiente/condições de trabalho diminui de 2000 (vaga 7) para 2001 (vaga 8).

As correlações *polychoric* entre as variáveis ordinais de satisfação em 1999 estão na tabela 4.

Tabela 4: Correlação *Polychoric* em 1999 – vaga 6

	S1	S2	S3	S4	S5	S6
S1 – Satisfação com o rendimento	1					
S2 – Satisfação com a segurança do trabalho	0,429	1				
S3 – Satisfação com o tipo de trabalho	0,428	0,585	1			
S4 – Satisfação com o número de horas de trabalho	0,403	0,439	0,580	1		
S5 – Satisfação com os horários de trabalho	0,339	0,524	0,563	0,664	1	
S6 – Satisfação com o ambiente/condições de trabalho	0,334	0,429	0,627	0,491	0,572	1

Em 1999 (vaga 6), verifica-se uma forte evidência dos factores intrínsecos, isto é, valores ligeiramente superiores nas correlações entre as variáveis intrínsecas, embora a variável S2 – satisfação com a segurança do trabalho – tenha uma estrutura de correlações elevada com as variáveis denominadas extrínsecas, principalmente com a variável S3 – satisfação com o tipo de trabalho.

Relativamente aos *scores* totais da satisfação laboral, em 1998 (vaga 5) o *score* extrínseco da satisfação laboral tem média 7,3708, que aumenta ao longo dos anos até 7,5870 em 2001 (vaga 8). O *score* intrínseco da satisfação laboral em 1998 tem média 16,4448, que decresce em 1999 (vaga 6), para 16,4123, e aumenta nos dois anos consecutivos para 16,5046 em 2001.

A correlação entre o *score* dos factores extrínsecos e o *score* dos factores intrínsecos aumenta desde o primeiro momento temporal até ao terceiro, isto é, até 2000 (vaga 7), partindo de 0,566 no ano 1998 até 0.607 em 2000, sofrendo uma ligeira descida no último ano de análise (2001), para 0,600.

3. MODELOS SEM E MPM

Os modelos SEM são uma extensão de várias técnicas multivariadas, tais como, a análise factorial e a regressão linear múltipla.

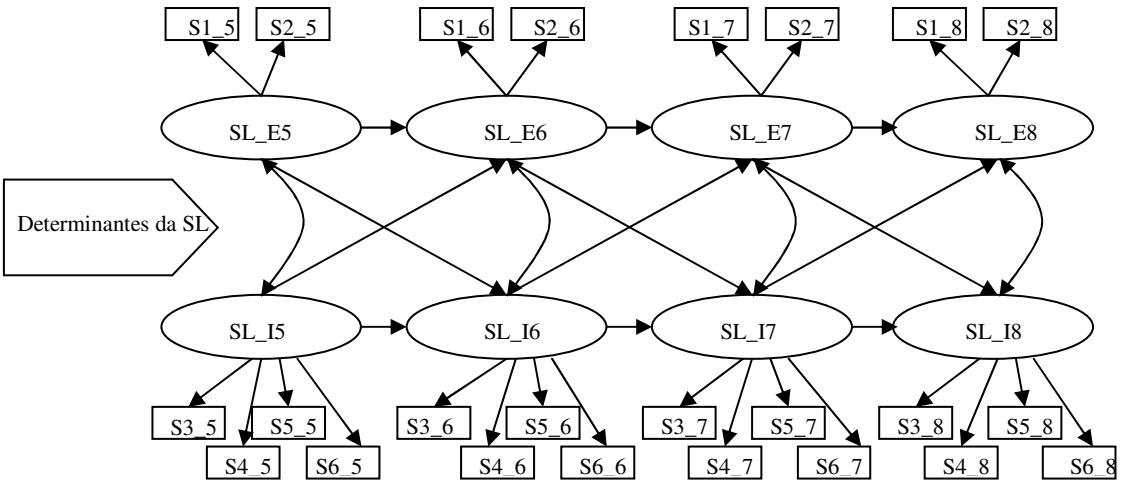
Os modelos de equações estruturais são integrados nos métodos de dependência, com relações múltiplas entre variáveis independentes e dependentes, onde algumas variáveis do modelo podem ser variáveis latentes. As variáveis latentes são variáveis que não podem ser observadas objectivamente/directamente, são representadas por várias variáveis. As variáveis latentes dependentes ou independentes são criadas através da Análise Factorial Confirmatória – a componente de medida dos SEM. Estes modelos oferecem a possibilidade de testar a relação entre estas variáveis não mensuráveis e incorporar a componente erro de medida no processo de estimação (ver, por exemplo, Salgueiro, 2008).

O modelo SEM da figura 3 inclui oito variáveis latentes, duas em cada vaga, sendo que uma respeita aos factores extrínsecos do trabalho e a outra aos factores intrínsecos. Numa primeira etapa foi efectuada uma Análise Factorial Confirmatória para cada ano de modo validar as duas dimensões latentes da satisfação em cada ano. Numa segunda etapa foi estimado um modelo estrutural, sem considerar os determinantes sócio-demográficos da satisfação laboral.

São considerados os factores extrínsecos e intrínsecos do trabalho, em cada momento temporal, como variáveis latentes, a primeira medida por dois indicadores ordinais e a segunda medida por 4 indicadores, também ordinais. Deixaram correlacionar-se os erros de medida das medidas repetidas ao longo dos quatro momentos temporais, assim como os resíduos das duas equações estruturais dos factores extrínsecos e intrínsecos, em cada momento temporal.

Numa terceira etapa incluíram-se os determinantes sócio-demográficos da satisfação laboral.

Figura 3: SEM



Este tipo de abordagem permite detectar quais os efeitos de desfasamento temporal, *lagged* e *cross-lagged*, significativos, assim como evidenciar quais os determinantes da satisfação laboral com impacto significativo sobre cada um dos factores latentes de satisfação. Os determinantes da satisfação laboral foram somente considerados na primeira vaga, isto é, invariantes no tempo, devido à dificuldade de considerar a estrutura de correlação das medidas repetidas das variáveis explicativas neste tipo de modelos.

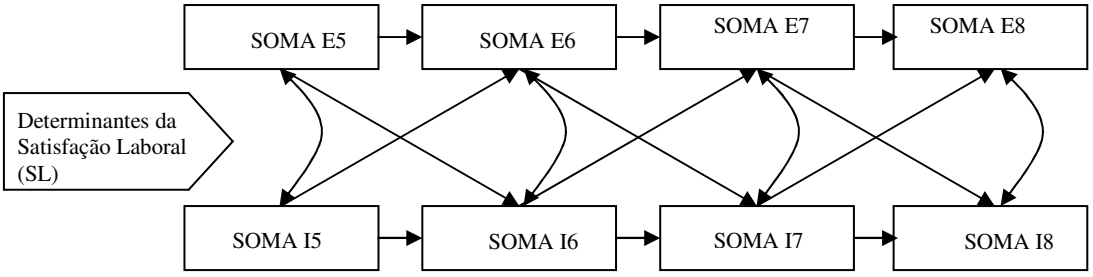
O MPM permite a análise de relações múltiplas entre as variáveis num modelo de caminhos (path analysis), sendo uma extensão de um modelo de regressão linear múltipla: várias relações entre um conjunto de variáveis podem ser simultaneamente estimadas. Num modelo de caminhos pode ainda considerar-se a correlação ou covariância entre as variáveis, que é

representada por uma seta curvilínea. Este tipo de modelos requer que todas as variáveis sejam observadas, não sendo possível a incorporação de erro de medida. Podem estabelecer-se relações entre duas variáveis com desfasamento temporal e covariâncias ou correlações entre as variáveis.

Para o modelo apresentado na figura 4 são criadas oito novas variáveis, construídas através da soma de itens dos factores extrínsecos e intrínsecos do trabalho, para cada momento temporal.

São considerados os factores extrínsecos e intrínsecos em cada momento temporal como as variáveis resposta num sistema de oito equações. Admite-se correlacionar os termos residuais das duas equações dos scores em cada momento temporal. Criaram-se variáveis *dummy* para os determinantes da satisfação categoriais.

Figura 4: MPM



Este tipo de abordagem vai permitir detectar quais os efeitos de desfasamento temporal, *lagged* e *cross-lagged*, significativos e ainda evidenciar quais os determinantes da satisfação laboral significativos. Os determinantes foram apenas considerados na primeira vaga, ou seja, invariantes no tempo.

4. RESULTADOS

Todos os modelos apresentados neste estudo foram estimados em LISREL 8.80 e utilizando o método de estimação de Máxima Verosimilhança Robusto, disponibilizado pelo *software*.

De acordo com Rose (2001), existem fortes evidências em dividir os determinantes da satisfação laboral entre factores extrínsecos e factores intrínsecos, sendo que, os primeiros dizem respeito aos aspectos materiais do trabalho, no nosso estudo são considerados a satisfação com o rendimento e com a segurança do trabalho, e os factores intrínsecos são os aspectos simbólicos e qualitativos do trabalho, sendo que neste estudo se considera a satisfação com o tipo de trabalho, com o número de horas de trabalho semanais, com os horários de trabalho e com as condições/ambiente de trabalho.

Assim, numa primeira etapa, foi efectuada uma análise factorial confirmatória para cada momento temporal deste estudo, de forma a validar a componente de medida dos factores latentes extrínsecos e intrínsecos da satisfação laboral. O diagrama do modelo e os resultados da estimação, para 2001, estão na figura 5 e na tabela 5, respectivamente.

Figura 5: Modelo de 2001

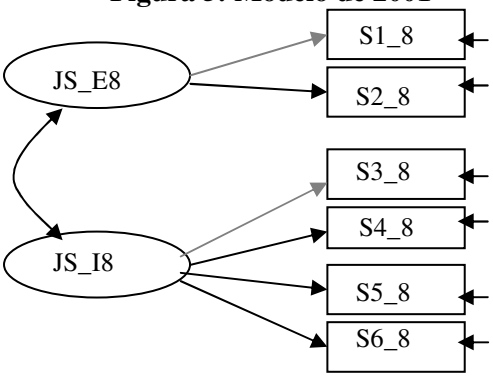


Tabela 5: Estimativas do modelo 2001

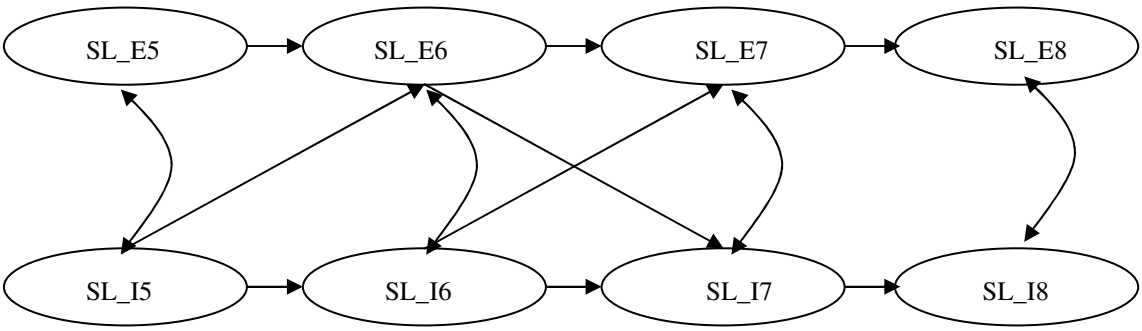
	Estimativas estandardizadas	Estimat.	t-values
S1	0,573	1	--
S2	0,817	1,425	21,503
S3	0,827	1	--
S4	0,757	0,915	12,314
S5	0,754	0,912	11,556
S6	0,741	0,895	35,498
Bondade de ajustamento			
df	Chi-Sq	RMSEA	RMR
8	9,507	0,00685	0,0279

Tendo em conta a tabela 5, este modelo é um modelo com 8 graus de liberdade, e são consideradas diversas medidas de Bondade de Ajustamento, duas medidas de ajustamento absoluto, o C*hi-Square* e o *Root Mean Square Residual* (RMR), e uma medida baseada no erro de aproximação à população, o *Root Mean Square Error of Approximation* (RMSEA). A medida RMR e a medida RMSEA encontram-se dentro dos limites aceitáveis que são 0,08 e 0,05, respectivamente. Embora a variável S1_8, relativa à satisfação laboral com a remuneração tenha um factor *loading* abaixo do valor considerado aceitável (0,7), este é considerado pela literatura um determinante importante da satisfação laboral, e no prosseguimento da análise deste estudo, quando é estimado o modelo longitudinal, com a introdução dos indicadores de satisfação laboral, este determinante assume um *loading* de dependência bastante elevado, na ordem dos 0,9.

Numa segunda etapa estimou-se um modelo longitudinal com base no modelo anterior, para cada vaga, de forma a avaliar quais os efeitos *lagged* e *cross-lagged* significativos. Deixou-se correlacionar os erros de medida das medidas de satisfação repetidas, ao longo dos quatro momentos temporais, assim como os resíduos das duas equações estruturais, em cada momento temporal.

Na figura 6 e na tabela 6, estão representados o modelo obtido e os resultados da estimação do modelo, respectivamente.

Figura 6: Modelo longitudinal com efeitos *lagged* e *cross-lagged*



Estimado o modelo *cross-lagged* SEM sem os determinantes da satisfação, não se revelaram significativos os efeitos *cross-lagged* entre os factores intrínsecos de 2000 (SL_I7) e os factores extrínsecos de 2001 (SL_E8) com um *t-value* de -0,702. No limite estavam os efeitos *cross-lagged* entre os factores extrínsecos de 1998 (SL_E5) e os factores intrínsecos de 1999 (SL_I6) e entre os factores extrínsecos de 2000 (SL_E7) com os factores intrínsecos de 2001 (SL_I8) com um *t-value* de -2,334 e 2,218, respectivamente. Optou-se por excluir estas relações porque aquando a estimação do modelo com introdução dos indicadores da satisfação laboral, estas revelaram-se não serem significativas para o modelo.

Tabela 6: *Loadings* dos efeitos *lagged* e *cross-lagged* do modelo longitudinal SEM

	JS_I5		JS_E5	
	Estimativas estandardizadas	t-value	SS	t-value
JS_I6	0,806	16,893	-0,108	-2,334
JS_E6	-0,305	-5,637	0,948	12,649

	JS_I6		JS_E6	
	Estimativas estandardizadas	t-value	SS	t-value
JS_I7	0.701	27,859	0.074	3,666
JS_E7	0.353	12,512	0.488	13,904

	JS_I7		JS_E7	
	Estimativas estandardizadas	t-value	SS	t-value
JS_I8	0.636	12,933	0.102	2,218
JS_E8	-0.037	-0,702	0.797	13,158

Bondade de ajustamento	
df	201
Chi-Sq	474,35
RMSEA	0.0184
RMR	0,102
AIC	672,35

Todos os efeitos de desfasamento temporal (*lagged*) são significativos e todos têm uma relação de dependência positiva, isto é, a satisfação laboral com os factores extrínsecos de determinado ano, tem um efeito positivo sobre a satisfação laboral com os factores extrínsecos do ano seguinte, ao longo dos quatro anos em estudo, o mesmo acontecendo com os factores intrínsecos.

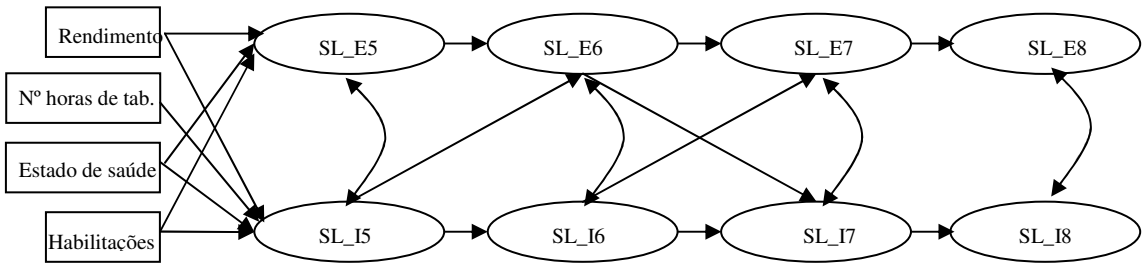
Nos efeitos cruzados (*cross-lagged*) significativos destacam-se as relações positivas entre os factores extrínsecos e intrínsecos, entre 1999 e 2000, e uma relação negativa entre os factores intrínsecos de 1998 (SL_I5) e os factores extrínsecos de 1999 (SL_E6), isto é, a satisfação laboral com os factores intrínsecos em 1998 influenciam negativamente os factores extrínsecos de 1999.

Numa terceira etapa foi testado um modelo com a introdução dos determinantes da satisfação, tendo em conta a estrutura avaliada na figura 6.

Os possíveis determinantes da satisfação laboral considerados são o género, a idade, o nível de habilitações, o estado civil, a percepção do estado de saúde, a existência de crianças menores de 12 anos no agregado, o número de horas de trabalho semanais e o rendimento. As variáveis relativas à idade, às horas de trabalho semanais e ao rendimento são tratadas como contínuas e as restantes foram tratadas como ordinais com dois níveis de resposta. As habilitações literárias estão separadas entre ter o ensino secundário completo ou superior e ter menos que o secundário. O estado civil divide-se entre ser casado e não ser casado e a percepção do estado de saúde está entre a considerar boa ou muito boa e considerar nem muito boa, nem muito má, má ou muito má.

Tal como está indicado na figura 7, os determinantes que têm impacto significativo sobre os factores extrínsecos e intrínsecos de satisfação laboral neste modelo são o rendimento, o número de horas de trabalho semanais, a percepção do estado de saúde e as habilitações.

Figura 7: Modelo longitudinal com os determinantes da satisfação



A variável rendimento tem uma relação positiva com os factores extrínsecos e intrínsecos se satisfação laboral em 1998, ou seja, de uma forma geral, rendimentos mais elevados induzem níveis mais elevados de satisfação e rendimentos mais baixos, implicam níveis de satisfação mais baixos. A variável número de horas de trabalho semanal apenas tem uma relação de dependência significativa com os factores intrínsecos do trabalho em 1998, uma relação negativa, ou seja, quanto mais horas de trabalho semanais, menores os níveis de satisfação. As variáveis percepção do estado de saúde e habilitações literárias indicam que uma boa percepção do estado de saúde e indivíduos com o ensino secundário completo ou superior implicam maiores níveis de satisfação.

Tabela 7: Loadings dos indicadores da satisfação

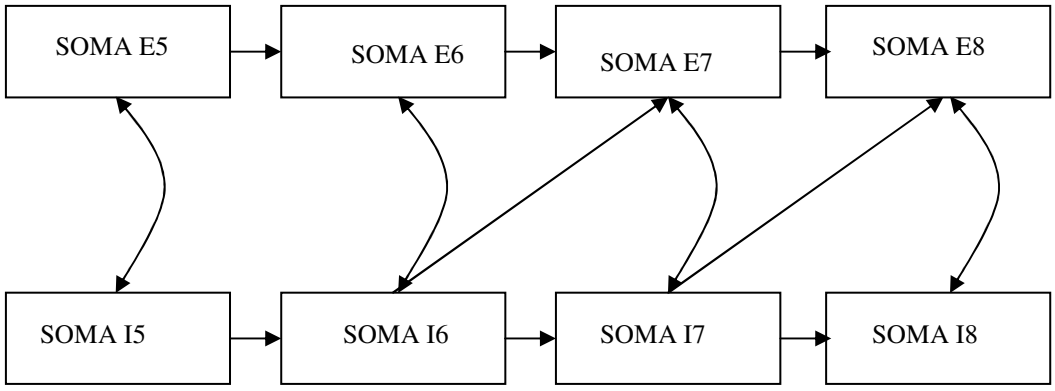
	JS_I5		JS_E5	
	Estimativa estandardi.	t-value	Estimativa estandardi.	t-value
Rendimento	0,09	3.065	0.187	3.919
Número de Horas	-0,08	-2,81	-	-
Estado de saúde	-0,09	-3,699	-0,131	-5,178
Habilitações Literárias	-0,11	-3,57	-0,114	-3,349

Bondade de ajustamento	
df	293
Chi-Sq	811,942
RMSEA	0,021
RMR	0,092
AIC	1037,942

As medidas de bondade de ajustamento calculadas revelam um ajustamento modelo-dados aceitável.

Os *scores* totais dos factores extrínsecos e intrínsecos da satisfação laboral, foram utilizados num MPM, uma variável *score* diz respeito à soma dos factores extrínsecos da satisfação laboral e a outra à soma dos factores intrínsecos de satisfação laboral. O primeiro modelo a ser testado é o modelo longitudinal sem os determinantes de satisfação, de forma a apurar quais os efeitos *lagged* e *cross-lagged* significativos. Foi ainda considerada a correlação entre os termos residuais das duas equações dos *scores* em cada momento temporal. Os resultados deste modelo apresentam-se na figura 8 e na tabela 8.

Figura 8: MPM longitudinal com efeitos *lagged* e *cross-lagged*



Tal como no modelo SEM, no MPM os efeitos *lagged* revelaram-se todos significativos. Relativamente aos efeitos *cross-lagged*, apenas são significativos os efeitos entre o *score* dos factores intrínsecos de 1999 (SOMA I6) e o *score* dos factores extrínsecos de 2000 (SOMA E7) e o efeito entre o *score* dos factores intrínsecos de 2000 (SOMA I7) e o *score* dos factores extrínsecos de 2001 (SOMA E8). Todas estas relações são positivas, então os efeitos *lagged* indicam que a satisfação laboral entre os factores extrínsecos de determinado ano, tem um efeito positivo sobre a satisfação laboral dos factores extrínsecos do ano seguinte, ao longo dos quatro anos em estudo, o mesmo acontece para os factores intrínsecos.

Tabela 8: *Loadings* dos efeitos *lagged* e *cross-lagged* do MPM longitudinal

	E5		I5	
	Estimativa estandardizada	t-value	Estimativa estandardizada	t-value
E6	0,618	7,789	0,082	0,998
I6	0,1	0,742	0,59	2,394

	E6		I6	
	Estimativa estandardizada	t-value	Estimativa estandardizada	t-value
E7	0,575	8,423	0,121	1,656
I7	0,124	0,996	0,559	2,395

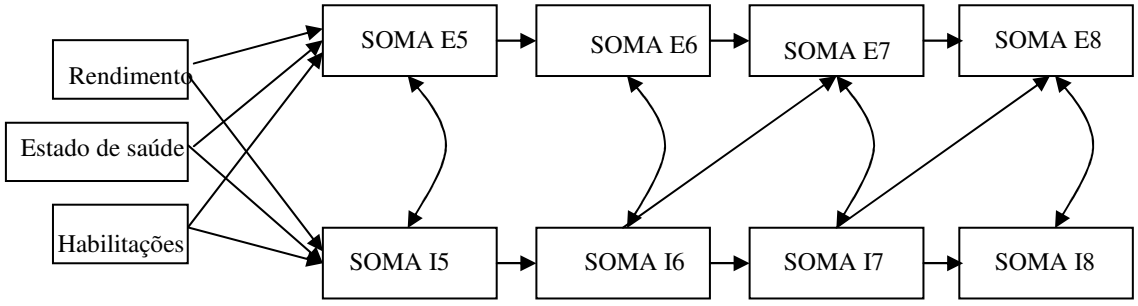
	E7		I7	
	Estimativa estandardizada	t-value	Estimativa estandardizada	t-value
E8	0,576	8,13	0,129	1,62
I8	0,108	0,764	0,564	2,301

Pode observar-se na tabela 8 que os efeitos *cross-lagged* entre 1998 e 1999 não são significativos, pois os *t-value* são muito inferiores a 2, o mesmo acontece com o efeito entre os factores intrínsecos de 1999 (SOMA E6) e os factores extrínsecos de 2000 (SOMA I7), e ainda entre os factores extrínsecos de 2000 (SOMA E7) e os factores intrínsecos de 2001 (SOMA I8), com *t-value* 0,996 e 0,764, respectivamente.

Finalmente, é testado o MPM com a estrutura da figura 9, mas com a introdução dos determinantes de satisfação.

Neste caso, as medidas ordinais foram convertidas em variáveis *dummy*, uma vez que este modelo segue o raciocínio de um modelo de regressão linear múltipla. Assim, tendo em conta a estrutura do modelo testado anteriormente, apenas são significativos para a satisfação laboral as variáveis rendimento, percepção do estado de saúde e nível de habilitações literárias.

Figura 9: MPM longitudinal com determinantes de satisfação



A variável rendimento tem uma relação positiva com os factores extrínsecos e intrínsecos da satisfação laboral, pelo que rendimentos mais elevados implicam níveis de satisfação mais elevados. A variável percepção do estado de saúde tem como referência ter a percepção de má ou muito má saúde. Os indivíduos com percepção de saúde boa ou muito boa têm significativamente maior nível médio de satisfação laboral (nos factores extrínsecos e intrínsecos) do que indivíduos com pior percepção do seu estado de saúde. A variável de habilitações literárias tem como referência ter menos que o ensino secundário, portanto indivíduos com o ensino secundário ou superior têm valor médio de satisfação laboral (extrínseco e intrínseco) significativamente superior aos indivíduos com menor nível de habilitações.

Tabela 9: Coeficientes de regressão dos determinantes de satisfação

	E5		I5	
	Estimativas estandardiz.	t-value	Estimativas estandardiz	t-value
Rendimento	0,191	13	0,068	3
Estado de saúde	0,092	6,958	0,085	6,372
Habilitações Literárias	0,081	5,828	0,101	6,786

Em comparação das duas abordagens estatísticas, destacam-se os efeitos *lagged* comuns entre os factores extrínsecos e intrínsecos da satisfação laboral nos dois modelos, ao longo dos anos em estudos. Os efeitos *cross-lagged* são distintos nos dois modelos. Os determinantes da satisfação laboral são o rendimento, a percepção do estado de saúde e as habilitações literárias em comum nos dois métodos e no modelo SEM o número de horas de trabalho semanais influencia negativamente os factores intrínsecos da satisfação laboral.

5. DISCUSSÃO

Foram considerados neste estudo dois procedimentos estatísticos distintos de forma a apurar quais os efeitos significativos entre os factores intrínsecos e extrínsecos da satisfação laboral em Portugal, assim como determinar quais os determinantes de satisfação relevantes utilizando modelação SEM e MPM, que se distinguem pela inclusão, ou não, no modelo de erros de medida.

Os modelos propostos têm por base as conclusões de Rose (2001) que afirma existirem diferenças significativas na abordagem de satisfação laboral tendo em conta os factores extrínsecos e intrínsecos.

São consideradas medidas de satisfação laboral em 4 vagas, de 1998 até 2001.

Tendo em conta as duas abordagens verificou-se que à medida que a satisfação laboral com os factores extrínsecos aumenta em determinado ano, então a satisfação laboral com os factores extrínsecos também aumenta no ano seguinte, o mesmo acontece com os factores intrínsecos. As diferenças verificam-se nos efeitos *cross-lagged.* Em SEM a satisfação laboral com os factores intrínsecos de 1998 influenciam positivamente a satisfação laboral dos factores extrínsecos de 1999. Entre 1999 e 2000, também são significativos os efeitos *cross-lagged* entre os factores extrínsecos e intrínsecos. Em MPM, os efeitos *cross-lagged* significativos são entre a satisfação laboral com os factores intrínsecos e os factores extrínsecos de 1999 com 2000 e 2000 com 2001.

Partindo de um conjuntos de 8 possíveis indicadores da satisfação laboral (género, idade, nível de habilitações, estado civil, percepção do estado de saúde, existência de crianças menores de 12 anos no agregado, número de horas de trabalho semanais e rendimento) apenas quatro se mostraram importantes na modelação SEM (rendimento, número de horas de trabalho semanais, percepção do estado de saúde e habilitações literárias) e três na modelação MPM (rendimento, percepção do estado de saúde e habilitações literárias).

Na modelação SEM, rendimentos mais elevados induzem níveis mais elevados de satisfação e rendimentos mais baixos implicam níveis de satisfação mais baixos, quer nos factores extrínsecos, quer nos factores intrínsecos. O número de horas de trabalho semanais apenas tem uma relação de dependência significativa com os factores intrínsecos do trabalho em 1998, esta relação é negativa, ou seja, quanto mais horas de trabalho, menores os níveis de satisfação laboral. A percepção de um bom ou muito bom estado de saúde implica maiores níveis de satisfação laboral em relação aos indivíduos que têm a percepção de um estado de saúde pior, e as habilitações literárias para indivíduos com o ensino secundário completo ou superior apresentam maiores níveis de satisfação laboral (nos factores extrínsecos e intrínsecos).

Na modelação MPM a variável rendimento tem uma relação positiva com os factores extrínsecos e intrínsecos da satisfação laboral, pelo que rendimentos mais elevados implicam níveis de satisfação laboral mais elevados. Os indivíduos com percepção de saúde boa ou muito

boa têm significativamente maior nível médio de satisfação laboral (nos factores extrínsecos e intrínsecos) do que indivíduos com pior percepção do seu estado de saúde. Os indivíduos com o ensino secundário completo ou superior têm valor médio de satisfação laboral (extrínseco e intrínseco) significativamente superior em relação aos indivíduos com menor nível de habilitações.

Para estudos futuros seria interessante procurar lidar com as não respostas das variáveis de satisfação nos dois tipos de modelação, bem como procurar incorporar a informação disponível sobre o desenho amostral aquando da estimação dos modelos.

Trabalhos futuros procurarão ainda modelar as trajectórias de satisfação laboral no *framework* dos *Latent Growth Curve Models* (ver, por exemplo, Salgueiro, 2009).

REFERÊNCIAS BIBLIOGRÁFICAS

CLARK, A. (1997), "Job Satisfaction and Gender: why are Women so Happy at Work?", Labour Economics, 4, pp. 341-372.

CLARK, A. e OSWALD, A. (1996), "Satisfaction and Comparison Income", Journal of Public Economics, 61, pp. 359-381.

CLARK, A., OSWALD, A. e WARR, P. (1996), "Is Job Satisfaction U-shaped in Age?", Journal of Occupational and Organizational Psycology, 69, pp. 57-81.

DRAKOPOULOS, S. e THEODOSSIOU, I. (1997), "Job satisfaction and target earnings", Journal of Economic Psycology, 18, pp. 693-704.

GARCÍA. E. e GARCÍA, J. R. (2006), "Determinants Socioeconòmics de la satisfacció Laboral a Catalunya", Nota d'Economia, nº86, pp. 99-119.

HAIR, J. F., BLACK, W. C., BABIN, B. J., ANDERSON, R. E. e TATHAM, R. L. (2005), "Multivariate Data Analysis", Prentice-Hall International, Inc., 6ª edição.

JÖRESKOG, K. G. e SÖRBOM D. (2006) LISREL 8.80, Scientific Software International, Inc. url: http://www.ssicentral.com/index.html

ROSE, M. (2001), "Disparate Measures in the Workplace: Quantifying Overall Job Satisfaction", Colchester: Institute of Social and Economics Research, 2001 British Household Panel Survey Researchs' Conference, 5 – 7 July.

SALGUEIRO, M. F. (2009), "Modelling Job Satisfaction Trajectories in the British Household Panel Survey using Latent Growth Curve Models", Asepelt'09 – XXIII International Congress of Applied Economics, Universidade da Beira Interior, Junho de 2009.

SALGUEIRO, M. F. (2008), "Modelos de Equações Estruturais – Aplicações em LISREL", (no prelo).

SCHUMACKER, R. E. e LOMAX, R. G. (2004), "A beginner's Guide to Structural Equation Modeling", Lawrence Erlbaum Associates, 2ª edição, New Jersey.

SKALLI, A., THEODOSSIOU, I. e VASILEIOU, E. (2008), "Jobs as Lancaster Goods: Facets of Job Satisfaction and Overall Job Satisfaction", The Journal of Socio-Economics, 37, pp. 1906-1920.

SOUSA-POZA, A. e SOUZA-POZA, A. A. (2003), "Gender Differences in Job Satisfaction in Great Britain, 1991-2000: Permanent or Transitory?", Applied Economics Letters, 10:11, pp. 691-694.

VIEIRA, J. C. (2005), "Skill Mismatches and Job Satisfaction", Economics Letters, 89, pp. 39-47.

VIEIRA, J. C., Menezes, A., Gabriel, P. (2005), "Low Pay, Higher Pay and Job Quality: Empirical Evidence for Portugal", Applied Economics Letters, Vol. 12, pp. 505-511.

HOSPITAL COST EFFICIENCY IN PORTUGAL: A FRONTIER ANALYSIS OF PANEL DATA

JEAN GOMES
E-mail: jg@ubi.pt
Department of Management and Economics
University of Beira Interior

Abstract

Starting off with a standard homoscedastic stochastic cost frontier approach so as to find the best practice reference technology of the mainland Portugal public hospitals, we finish up favouring cost frontier estimation with heteroscedasticity across hospitals, particularly the one we carry out by means of a corrected random effects (CRE) cost frontier procedure of our own. The CRE cost frontier procedure we bring forth may be helpful as it calls for no distributional assumptions about the time-invariant inefficiency component. Based on panel data with 63 hospitals for the period 2002 – 2006, the computed technology is thought to be one and the same for the whole sample even for hospitals that differ very much in size.

Keywords: hospital, cost, efficiency, frontier.
Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUCTION

Roughly all over the western world there have been more and more national debates on how to get better efficiency in the public health sector and specifically in public hospitals, as national health care expenditures increase progressively often with small resultant helpful effects. Particularly in what concerns Portugal, as APES (1997) – the Portuguese Health Economics Association – once put it, "Portugal does not spend little in health care, rather it spends badly the resources made available". Also Antunes (2001) argues that "putting more money in a (health care) system which wastes in this manner means creating more and more squander and does not contribute at all to fight against it". Although there is a broad national consensus on the lack of efficiency of Portugal public hospitals, there has been only a small amount of works addressing the issue.

In this work we attend to the lack of efficiency of Portugal public hospitals matter by first going after the best practice technology so as to have a reference technology cutting edge, and then appraise how far each hospital stands from the estimated reference technology frontier. Also by putting the computed reference cost and the observed expense head-to-head makes it possible to evaluate how much savings each hospital can aim to arrive at whilst producing its actual amount of output, or likewise but the other way around, how much extra output each hospital can aim to arrive at whilst spending its observed actual amount of expense.

Besides finding the reference frontier cost of treatment at the hospital is helpful for quite a few applications. Several health economics theoretical models make use of an abstract unit cost of treating a patient at the hospital which may lead to valuable policy points of orientation, once it is unveiled. For instance, Barros and Martinez-Giralt (2003) address the issue of how different organisational settings mixed up with different payment designs have impact on the incentives to supply preventive health care, and so as to do so, they do bring into play the cost of treatment at the hospital in their conceptual framework.

After gathering panel data on 63 mainland Portugal public hospitals for the period 2002 – 2006, we to begin with go for the best practice reference technology cutting edge by means of a homoscedastic stochastic cost frontier approach which we put faced up to roughly all possible alternative modelling we could think of so as to be assured enough about the reliability of our results. We end up preferring cost frontier estimation with heteroscedasticity across hospitals. In any case though, and being in line with quite a few authors' equal findings, we find that either homoscedasticity yields biased estimates when there is heteroscedasticity in the data or adding heteroscedasticity whilst there is none confounds extremely the results. Also the estimation with heteroscedasticity is undertaken by means of a corrected random effects (CRE) cost frontier procedure of our own. Although its reasoning is enormously analogous to the cross-section data context corrected ordinary least squares procedure, the CRE cost frontier procedure we bring forth may be valuable because it requires no particular assumption about the shape or the location of the time-invariant inefficiency component distribution.

The results suggest a one and the same reference technology for all considered hospitals even for hospitals which differ in size considerably. The estimated reference average cost is found to be monotonically decreasing and so there is no minimum efficient scale. This does not coincide with a number of authors' works which specify economies of scale are exploited in hospitals with either 100–200 beds or instead approximately 300 beds, plus diseconomies could come out in hospitals with 300–600 beds. We further show that whenever the average cost function is monotonically decreasing, recognising from which level of output economies of scale are fully exploited depends badly upon which criteria one resorts to.

As for policy implications, the cost efficiency rankings we present are of special interest for sure. But we also present the overall potential cost reduction plus the excess cost in terms of output if all 63 hospitals implemented the efficient behaviour. Both are quite substantial. In addition, we estimate the frontier unit cost of one specific Portugal public hospital and compute how much savings/extra output it can arrive at if it behaved fully efficiently.

The rest of this work is organised as follows. Section 2 starts with a global discussion on the methodology we use focusing in particular on the stochastic frontier approach. A discussion on the functional form and all possible alternative econometric models one would opt for is next undertaken. Section 3 presents the model specification. Section 4 provides the data and summary statistics. The stochastic frontier regression results and the estimates of the corrected random effects model with heteroscedasticity are given and discussed in Section 5 and in Section 6, respectively. Section 7 reports the overall cost efficiency estimates plus it reports the degree of output-specific and ray economies of scale. Our efficiency scores and measures of economies of scale are there evaluated against international evidence. The frontier unit cost of treating a patient at the Cova da Beira hospital centre is then computed in Section 8. The impact on its costs/production of fully efficiency is discussed too. Section 9 concludes.

2. METHODOLOGY

Several authors have argued that high hospital costs are the direct outcome of inefficient conduct. Bringing the stochastic frontier methodology into the debate on which options an individual organisation has so as to improve its efficiency is quite helpful because this approach not only determines the best practice reference technology but it also evaluates how inefficient the conduct of the organisation has been, by focusing on the difference between the cost the estimated reference technology frontier forecasts and the observed expense. The key assumption is that the regression error can be split into two terms, specifically one asymmetric component which reflects the inefficient conduct plus one as usual symmetric normally distributed random error.

Besides several authors favour the stochastic frontier methodology in order to find cost functions because it adjusts the ordinary least squares over-estimated intercept, hence moving the corresponding cost frontier towards the proper level (Zuckerman et al. 1994; Linna et al. 1998; Rosko 2001; Farsi and Filippini 2008). Plus it brings forth efficiency rankings for individual organisations which are now and then seen as its main purpose.

There is some criticism though. Incomplete output specification and the lack of suitable quality measures may have a stronger lousy impact on stochastic frontier estimates than under ordinary least squares (Newhouse 1994). In addition, estimates or rankings are often enormously sensitive to estimation decisions that have to be made, most of which are non-testable (Horrace and Schmidt 1996; Jensen 2000; Street 2003). Stochastic frontier models of panel data where organisations are observed over several periods are superior to their cross-section data equivalents though as they are less likely to produce biased estimates due to omitted variables and they call for weaker assumptions (Dor 1994).

One may think inefficiency is time-invariant. If so plus one has panel data, the time-invariant inefficiency component can be identified by standard individual fixed or random effects and inefficient conduct is no longer necessarily assumed asymmetric positive. Putting it rather differently, when inefficiency is constant across time, the best practice reference technology frontier can also be computed by standard fixed or random effects and only weaker assumptions may be required. Specifically the standard fixed effects panel data estimator makes it possible to give up on all distributional assumptions about efficiency, which is particularly welcome as there is no economic ground for choosing one distribution over another (Schmidt 1985). Besides it does not require inefficiency to be uncorrelated with the regressors. Preferring the standard random effects over the standard fixed effects estimator though comes at the price of having to impose at least some distributional assumptions on efficiency – which is as a minimum required to be independently identically distributed with constant mean and variance – as well as inefficiency is asked to be independent of both the regressors and the random component. At last and as a matter of fact, whenever inefficiency is constant across time, provided that inefficiency follows the same asymmetric positively distribution and the random component is normally distributed, the standard random effects estimator yields precisely the – what we choose to call – stochastic frontier methodology estimator.

A rather recent improvement in frontier models for panel data addresses the limitation of having to impose there is no unobserved heterogeneity. But organisations may well be exposed to different external factors which are beyond their control or are improperly measured. Take hospital location for instance. Neither locations face an equal profile of illnesses nor is health care uniformly integrated among primary care providers. Besides even if all factors are taken into account there are frequently only weak proxy measures available and it may hence be too strong to assume heterogeneity among organisations is precisely observed. Remark that if the assumption turns out to be wrong, the fixed or random effect, which up until then would be interpreted as taking place solely from inefficiency, would then be a muddle of inefficiency and unobserved heterogeneity. This new line of research specifies "true" models by adding either organisational fixed dummies or distributional assumptions on the random organisation-specific constant parameter (Polachek and Yoon 1996; Greene 2004; Farsi et al. 2005; Greene 2005).

Acknowledging the shortcomings in recognising each organisation's relative efficiency, standard or "true" stochastic frontier models are nonetheless particularly useful to assess conduct across hospital groups (Hadley and Zuckerman 1994; Folland and Hofler 2001; Farsi et al. 2005). Also going after both standard and "true" panel's individual fixed or random effects besides the stochastic frontier methodology is helpful as later sensitivity analysis enables to either strengthen or make the results unlikely to be trusty.

2.1. Functional form

None of the numerous criteria for model selection makes it possible to choose only on a fully objective ground. There is necessarily room for subjective judgement. A functional form can be preferred depending upon its parallel to the underlying theory or even to an empirical root for the correct relationship between outputs and the different mixes of inputs. Also a functional form can be preferred depending upon how undemanding its econometric estimation would be or if econometric estimation is doable in the first place. Besides neither the goodness of fit nor broad conformity to data should be omitted. Once having chosen at least two estimable functional forms with reasonable parallel to the beneath theoretical or empirical basis and suitable application properties, the applied economics researcher may afterwards favour a

strictly statistical criterion to support her final choice. See Griffin et al. (1987) for a comprehensive roll of functional forms and a discussion on form selection.

We first choose to confront two of the most commonly used functional forms so we could portray cost functions. We choose to oppose the Cobb-Douglas (log-linear) model to the translog model. Both do maintain an appropriate correspondence to the dual problem of the neo-classical theory of the firm or at least to an empirical behavioural source, and in both cases the econometric estimation is doable. The Cobb-Douglas reduced number of parameters which have to be computed makes the functional form easier both to assess and to interpret. As usual, the major weakness of this form is the presumption of a constant rate of economies of scale which may indeed be a limitation because this means smaller organisations always benefit more than larger ones from an equal relative growth in output. The translog flexible functional form allows for possible adjustments in scale elasticity with output and so this scale elasticity aspect favours translog. The second-order elements which are estimated under the translog model may though exacerbate multicollinearity and so the model's statistical performance turns out to be negatively influenced. With high multicollinearity, the variance of the parameter estimates may be so enlarged that it might not be feasible to establish how much disparity in the dependent variable is explained by different explanatory variables.

We went on with our specification analysis by recognising how each functional form does influence the results when confronting the linear homogeneous in the prices Cobb-Douglas form to its translog model equivalent. We essentially putted each parsimonious model face to face and did the exact same thing with the model we ended up proposing. While the main estimated output coefficients do not vary much when adding more explanatory variables to the functional form, the estimated output coefficients contrast dramatically when choosing between the Cobb-Douglas and the translog functional form. Moreover, putting together the unexpected negative sign of one major computed output coefficient – in particular, one used for estimating the economies of scale – with the size of all estimated key parameters suggests the presence of high multicollinearity. We checked and confirmed the existence of high multicollinearity. Additionally, each time likelihood-ratio testing did not reject the Cobb-Douglas functional form against its translog counterpart, at a 1% significance level. We thus decided to centre our work on the Cobb-Douglas form.

2.2. Theoretical models

We first focus on the stochastic frontier model, and so the analysis we to begin with go through in this paper presumes that costs are determined as follows:

$$C_{jt} = f(q_{jt}, w_{jt}) \exp(v_{jt}) \exp(u_j) \quad (1)$$

where C is the total cost, $f(.)$ is the deterministic cost frontier which depicts the optimal functional relationship between outputs and the different mixes of inputs, q and w are the vectors of multiple outputs and input prices, respectively, $\exp(v)$ stands for the random shocks and $\exp(u)$ is the efficiency indicator. Almost as a footnote, $f(.)\exp(v)$ bears the stochastic cost function and $\exp(v)\ \exp(u)$ is the composite error term. Also, the amount by which $\exp(u)$ exceeds 1 measures cost inefficiency. At last, observations are indexed by a hospital index $j = 1, 2, ..., n$ and a time index $t = 1, 2, ..., T$.

The logarithmic form of (1) comes as in expression (2) below:

$$\ln C_{jt} = \ln f(q_{jt}, w_{jt}) + v_{jt} + u_j \quad (2)$$

The random noise v is symmetric normally distributed with zero mean and variance σ_v^2 and the non-negative departures from the cost frontier u which capture the cost inefficiency in production have a truncated-normal distribution. Notice that u does not present a time index which means we believe efficiency does not vary through the five years we consider in this work.

There are two broad alternative approaches each concerning the conviction one has about whether or not efficiency fluctuates over time. We performed an exploratory analysis to identify which form would best fit the data each time confronting two models which only differ from each other in what concerns the assumption of having time-invariant or time-varying efficiency. Again both parsimonious forms were putted face to face and also the form we ended up recommending here. The time-varying efficiency assumes the Battese and Coelli (1992) parameterisation of time effects with the inefficiency term being defined as below:

$$u_{jt} = \exp[-\eta(t - T)]\, u_j, \quad (3)$$

where η is an unknown coefficient to be computed and u_j represents independent and identically distributed non-negative arbitrary variables obtained by the truncation at zero of the normal distribution with both unknown mean and variance. Both Wald and likelihood-ratio tests corroborate the parameter η is statistically non-significant, at a 5% significance level, and so, with $\eta = 0$, equation (3) turns into $u_{jt} = u_j$, for all t.

There are several alternative ways to model temporal changes in efficiency. The perhaps simplest way would be to split inefficiency into two terms, one particular to each organisation, the other being industry-wide but specific to each period (Greene 1993). One other way would let the pattern of efficiency change be equal among organisations and would also assume the magnitude of the change differs among establishments, as we supposed above when pursuing the Battese and Coelli (1992) modelling, but would differently portray the pace at which efficiency varies. Some researchers include a set of dummy variables for time interacted with the efficiency term while others prefer to impose a different distribution for how efficiency evolves (Kumbhakar 1990; Lee and Schmidt 1993). Also other researchers go one step further and specify how efficiency might be expected to evolve or which organisation-specific and environmental factors could explain efficiency over time (Cornwell et al. 1990; Battese and Coelli 1995).

STATA version 9.2 we resort to in this work makes it straightforward to compute the parameters of the time-varying decay model following Battese and Coelli (1992) but not for other parameterisations. We do not think though the assumption of a constant level of efficiency through the five years we consider here is a drawback at all. Neither the data is observed over long periods of time nor should it be expected major external pressures whose impact would significantly affect the temporal pattern of efficiency, such as deep regulatory programmes or a technological boost.

2.3. Standard fixed or random effects models

The time-invariant inefficiency component $u_{jt} = u_j$ can also be identified by standard panel's individual fixed or random effects (Pitt and Lee 1981; Schmidt and Sickles 1984). Writing down expression (2) again so as to arrive at a broad logarithmic form of the model as below definitely makes it straightforward to anticipate why.

$$\ln C_{jt} = \alpha + \sum_k \beta_k \ln x_{kjt} + \sum_l \phi_l z_{ljt} + v_{jt} + u_j \quad (4)$$

where C is the total cost, α is the intercept, x comprises the multiple outputs q and the input prices w, z stands for all additional complementary explanatory variables, β and ϕ are the corresponding coefficients, v is the error random component as always, and u is the cost inefficiency term.

We then computed the fixed effects and the random effects estimators. Choosing which one, among those standard estimators, best fits the data set, and afterwards putting the resultant estimates against the ones we find under the stochastic frontier framework, will make it possible to undertake some sensitivity analysis and next confirm results (or not) so as to being able to place enough reliance on the estimates we work out.

Picking the random effects estimator over the fixed effects estimator comes at the price of having to require distributional assumptions on *u* and to presume there are no correlation between the explanatory variables and the inefficiency term. Some researchers just require *u* to be independently identically distributed with constant mean and constant variance while others need stronger assumptions about the shape of the distribution of *u* (Pitt and Lee 1981; Battese and Coelli 1988; Greene 1993; Kumbhakar and Lovell 2000). Admitting those presumptions are true, the random effects estimator is afterwards computed either by generalised least squares or by maximum likelihood estimation. In any case, it is always more efficient than the fixed effects estimator.

None of the standard fixed effects models we ran – least squares dummy variable, within effects and between effects models – seems to satisfactorily estimate a theoretical based cost function from our data set. There are always issues concerning the major computed output coefficients or the input prices. Some of the key output parameters come either with an unexpected negative sign or seem to be statistically non-significant, as posterior Wald tests confirm, at a 1% significance level. We performed Fisher tests to determine whether there were fixed effects or not, and the large F statistics rejected the null hypothesis that all dummy variables (except one) were zero in the least squares dummy variable forms we computed, consequently favouring the fixed effects models. The presence of high multicollinearity, which we verified, may explain why the fixed effects models perform so poorly. Also we may be facing the well-known problem of attenuation bias. Common to every panel data application, this problem is intensified by the fixed effects estimator as it erases a considerable amount of variation in the data, whereby measurement error in the regressors can lead fixed effects parameters to go towards zero, and so being interpreted as having little or even no impact (McKinnish 2000).

Both standard random effects models match suitable cost functions though, either admitting *u* is simply independently identically distributed with zero mean and constant variance, which is required for the generalised least squares estimation, or by imposing inefficiency *u* is truncated-normal as asked by the maximum likelihood computations. Perhaps most surprisingly is the fact that the (generalised least squares) random effects model fails the Hausman specification test, therefore suggesting correlation between the individual effects and the explanatory variables (Hausman 1978). If *u* and the regressors are correlated, the random effects model should produce biased estimators, and as a result the corresponding fixed effects model would usually be preferred. We after that performed the Breusch and Pagan (1980) Lagrange multiplier test to assess whether or not the organisational effects could indeed be random draws from our underlying population. The null hypothesis H_0: $\sigma_u^2 = 0$ were rejected, and so were the pooled regression model in which *u* is set equal to zero.

Although both the Fisher and the Hausman specification tests point in the direction of the standard fixed effects models, no fixed effects model can actually be chosen as none seems to properly fit a cost function from our data set. In addition, not only the random effects models can match proper cost functions but also the (generalised least squares) standard random effects model passes the Breusch-Pagan Lagrange multiplier test which checks on random effects. We hence conclude the one-way random group effects form should be the most adequate, and we move ahead by taking the generalised least squares random effects estimator rather than its maximum likelihood equivalent.

2.4. Assuming heteroscedasticity across organisations

Having inferred the random effects model should be appropriate, we went a small step further and fitted the exact form admitting heteroscedasticity across groups. As opposed to the common linear regression model under which only computation efficiency is worse off, this should be here a significant issue as heteroscedasticity should produce biased estimates of both the frontier model and the cost efficiency scores. The posterior likelihood-ratio test we ran rejected the homoscedasticity null hypothesis, and so when hesitating between including or not including heteroscedasticity, the heteroscedastic one-way group random effects model should be the most suitable.

2.5. Unobserved heterogeneity

We mentioned earlier a quite recent line of research has been proposing "true" specifications of the stochastic frontier model, as an attempt to surmount the unobserved heterogeneity matter and be therefore able to separately capture efficiency as well as all other relevant factors that would not be properly taken into account (Polachek and Yoon 1996; Greene 2004; Farsi et al. 2005; Greene 2005). This line of investigation proposes to incorporate group-specific dummies as an additional set of regressors, the resultant general model specification in its logarithmised form being put as:

$$\ln C_{jt} = \alpha + \sum_k \beta_k \ln x_{kjt} + \sum_l \phi_l z_{ljt} + v_{jt} + \alpha_j + u_j \quad (5)$$

We ran the "true" fixed effects regression model as it does not require any distributional assumption on the time-invariant organisation-specific heterogeneity parameter. We afterwards chose to rule this specification out since some of the key output coefficients as well as a certain number of group-specific parameters were found to be statistically non-significant, at a 1% significance level. This outcome is independent of setting the global intercept α equal to zero or not. Also we did not run the "true" random effects regression model since STATA version 9.2 we are making use of in this work does not compute estimates as expressed in (5) and assuming at the same time the heterogeneity parameter is independently identically normally distributed (Greene 2004).

3. MODEL SPECIFICATION

The general specification of the Cobb-Douglas cost function model we employ for computing the cost frontier of Portugal public hospitals can be written as follows:

$$\begin{aligned}\ln \text{COST}_{jt} = \quad & \alpha + \beta_1 \ln \text{ID}_{jt} + \beta_2 \ln \text{LOS}_{jt} + \beta_3 \ln \text{VISITS}_{jt} + \rho_1 \ln \text{PL}_{jt} + \rho_2 \ln \text{PK}_{jt} \\ & + \theta \, \text{NPB}_{jt} + \delta \, \text{TEACHING}_{jt} + \square_1 \, \text{BEDS}_{jt} + \square_2 \, \text{POP}_{jt} + v_{jt} + u_j \end{aligned} \quad (6)$$

Our total cost – COST – dependent variable matches in fact the total variable cost, or in other words, the total expenses less both the payment in instalments and the financial year provision.

Inpatient discharges – ID – and total outpatient visits – VISITS – are the major outputs in our cost function model. Inpatient discharges seek to describe the inpatient episodes of care while outpatient attendances portray ambulatory care. Many authors prefer to use the diagnosis-related groups weighted number of hospitalisations to express the inpatient episodes of care as diagnosis-related groups weights are capable of identifying the cost intensity of each episode (Linna 1998; Rosko 2001; Farsi and Fillipini 2008). We would also have preferred to use such number of hospitalisations if it were available. The average length of hospitalisation stay – LOS – in days is incorporated into the model though as a complementary output adjuster variable so as to provide a measure of severity of the inpatient episodes of care, where longer hospitalisation stays suggesting in principle more severe illnesses treated, and so it should surmount, at least up to a certain extent, the lack of the diagnosis-related groups weights data (Vita 1990; Scuffham et al. 1996; Carey 1997; Farsi and Fillipini 2008). Additionally, some researchers use the total number of emergency room episodes instead of the total outpatient visits to depict ambulatory care (Cowing and Holtman 1983). Others authors prefer to use both variables to portray outpatient treatment (Vita 1990; Fournier and Mitchell 1992; Carreira 1999). At a preliminary stage we replaced the total number of outpatient visits by the total number of emergency room episodes as well as we introduced both outpatient and emergency room attendances to depict outpatient treatment. Each times both Wald and likelihood-ratio tests pointed out the emergency room episodes variable as being statistically non-significant. Also this outcome seems to be independent not only of which natural unit of measurement one uses to quantify the number of emergency room episodes but also of whether the number of

emergency room episodes is putted into its logarithmised form or not. We therefore chose the total number of outpatient visits only to portray ambulatory care which in fact is in line with another third group of authors (Scuffham et al. 1996). Moreover, at a preliminary stage too, we putted in the number of emergency room episodes as a percentage of the total number of outpatient visits, which would function as an additional adjuster variable so as to control for variations in outpatient case-mix (Dor and Farley 1996; Rosko 2001). The Wald and the likelihood-ratio tests nonetheless rejected the added variable.

It perhaps should be expected in this particular context that the number of emergency room episodes would finish up as being statistically non-significant. The way Portugal public hospitals are organised makes most movements in the emergency room ending up to have a straight echo in the inpatient discharges or in the outpatient attending numbers. Since neither emergency medicine is a medical specialty nor there is an emergency medicine specific accounting, a larger number of emergency room episodes for instance may aggravate the number of surgeries which should increase not only the number of inpatient discharges but also the number of follow-up outpatient visits in the subsequent stage. Besides even if there is no impact on the inpatient episodes of care, a larger number of emergency room episodes should at least affect the number of outpatient visits.

The labour factor prices – PL – were computed as the ratio of the total annual personnel expenditure divided by the number of personnel employed in the hospital each year being measured on December 31. The capital prices – PK – follow the 2002 year grounded European harmonised index of consumer prices divided by one hundred, so that the index equals 1 in year 2002. Some researchers rather approximate capital prices using the ratio of the hospital's total capital expenses divided by the number of available beds (Wagstaff and Lopez 1995; Rosko 2001; Farsi and Fillipini 2008). Aspiring to such data, we employed the payment in instalments as an approximation of each year's capital spending and divided it by the total number of beds being used in the hospital. Unfortunately, the calculated variable came out as being statistically non-significant.

The nurse per bed ratio – NPB – is then included to represent the quality of hospital care (Farsi and Filippini 2008). It perhaps may be worth reporting at this point that at an initial step we added a binary dummy variable for hospitals which had received a quality accreditation either from the United Kingdom Health Quality Service (former King's Fund) or from the United States Joint Commission International. The binary dummy variable came out as being statistically non-significant. Moreover, we put in a binary dummy variable for teaching hospitals – TEACHING – as teaching hospitals not only generate different outputs but also may well pursue different aims too (Rosko 2001). The number of available beds each year on December 31 – BEDS – which may be regarded as being fixed in the short run is as well incorporated as a proxy variable for each hospital's level of fixed factors which go beyond the control of hospital managers (Vita 1990; Scuffham et al. 1996; Jacobs et al. 2006). We at last turned to the size of the population living in the town where each hospital is settled – POP – to resemble the magnitude of the population covered by the hospital. Notice also that posterior attempts concluded that neither the proportion of resident men and women aged over 64 years old nor the proportion of female patients were found to be statistically significant.

The Portugal health authorities have undertaken a number of restructuring measures since year 2002 or so particularly in what concerns public hospitals. A group of public hospitals were first converted into incorporated companies where the State is in each case the only shareholder. The idea behind this change would be to provide those public hospitals with much more similar running and legal tools private hospitals benefit from every day so as to improve efficiency in each incorporated company converted public establishment. Also there has been a certain hospital management regrouping wherein two or more chosen hospitals were merged into one. It happened sometimes that first merged hospitals were again joined together with other hospitals. Having buildings in different places should augment costs as a quantity of human and capital resources ought to be replicated in each different location besides communication and managing troubles which should appear as a result of distance. Moreover, management integration may also increase costs at the beginning of the new only administration as managers require time acquiring all information they need to successfully deal with the adjoined health

units and their own professionals, as well as human resources need time to fully adjust to the recent merged setting.

Not studying those effects would certainly not be an option since the restructuring measures the Portugal health authorities have undertaken are regularly under crossfire. In a preliminary stage, we hence added one binary dummy variable to characterise each effect in our econometric model, and added each variable either separately, either together with one other binary dummy variable or all together. The Wald and the likelihood-ratio tests rejected the binary dummy variable for hospitals converted into incorporated public companies, at a 1% significance level, implying therefore the legal changes the recent incorporated company hospitals have experienced have roughly had no impact on costs, and so the converted hospitals should generally be as efficient as the non-converted public ones. Additionally, the binary dummy variable for hospitals facing management integration in a given period of time was also found to be statistically non-significant. At last, even if we were not able to perform likelihood-ratio testing for the binary dummy variable which stands for hospitals with buildings in different locations because the samples differ from each other in size, the Wald test found the variable to be statistically non-significant, at a 5% significance level. We therefore excluded this latter regressor from our cost function model.

Our specification omits several factors. In particular, a more precise measure of the hospitals' teaching activity such as for instance the absolute number of medical students or the number of medical student whole-time teaching equivalents per inpatient spell would be welcome. Also there is no measure concerning the research undertaken in each hospital. Those aspects may represent a serious drawback as teaching and research activities may well constitute significant secondary outputs of hospitals. The conclusions of this work should hence be taken within the boundaries of both the data on hand as well as all the simplifying assumptions we make.

One last remark remains to be made before closing the present section. First, whenever STATA version 9.2 commands do not permit imposing restrictions on the computed parameters and we are required to guarantee our input prices are linearly homogeneous, we divide the total variable costs and the labour factor prices, on the left-hand side and the right-hand side of the equation, in that order, both of which are at the start assessed at current prices, by the capital price which becomes the numeraire price, therefore making sure that the input shares add up to 1. Also since the capital price tracks the ratio of year 2002 grounded European harmonised index of consumer prices by one hundred, the total variable costs and the labour factor prices at current prices, divided by the numeraire price as defined above, turn respectively into the total variable costs and the labour factor prices evaluated at year 2002 constant prices.

4. DATA

We make use of panel data on 63 mainland Portugal public hospitals for the period 2002 – 2006 (T = 5). During this period of time, several hospitals were merged into larger managing structures, and it occasionally happened that first merged hospitals were again merged to shape even bigger organisations. We in this work consider the year 2006 existing rearranged hospitals, or putting it differently, we bring into play the 63 mainland Portugal public hospitals as they were on December 31, 2006. Whenever two or more hospitals were merged into one but for which only individual data are available, we add each individual data value to the corresponding data values of the merged hospitals that together form the same new unique running structure so as to obtain data even when each individual now assembled hospital was being run separately. Also STATA version 9.2 simply ignores the observations for which there are missing values. This well-known complete cases only practice should be econometrically satisfactory as the missing values seem to be fully arbitrary. Furthermore, we included in our data set all mainland Portugal public hospitals that we intuitively found as having a similar nature or sharing a common technology. We for that reason did not include the specialised hospitals such as maternities, paediatric, orthopaedic, oncology or psychiatric hospitals or others.

Roughly all data come up either from the Portuguese National Health Service (NHS) annual information or from each hospital's own financial reports made available by means of the Portuguese Central Administration of the Health System. The European harmonised index of consumer prices was issued by the Bank of Portugal (Portuguese central bank). The profile of the city's population where each considered hospital is settled was issued by the Portuguese National Institute of Statistics. The information on the progress of each hospital's application for a quality accreditation, either from the United Kingdom Health Quality Service (former King's Fund) or from the United States Joint Commission International, was supplied by the Portugal (former) Institute for Quality in Health.

Table 1 describes each variable we make use of. Table 2 exhibits the summary statistics.

Table 1. Variable definitions

Name	Description
Cost of treatment	
COST	Total variable cost (in Euros)
Output variables	
ID	Inpatient discharges (in units)
LOS	Average length of hospitalisation stay (in days)
VISITS	Outpatient visits (in units)
Factor prices	
PL	Ratio of the total current expenditure on personnel divided by the number of personnel employed on December 31.
PK	Follows the 2002 year grounded European harmonised index of consumer prices divided by 100.
Distinctive characteristics	
NPB	Ratio of nurses to beds in the hospital
TEACHING	Binary dummy variable (1, 0) for teaching hospitals
Environmental adjuster variables	
BEDS	Available beds on December 31
POP	Inhabitants in the town where the hospital is settled (in millions)

Table 2. Summary statistics

Variables	Number of observations	Mean	SD	Min	Max
COST	309	59,800,220	66,646,580	3,962,644	293,820,000
ID	311	12,566.73	10,250.53	6	47,851
LOS	299	7.37	1.42	2.75	13.40
VISITS	312	111,216	108,507.70	5,259	520,029
PL	240	25,920.15	4,047.58	2,077.80	48,053.26
PK	314	1.06	0.04	1	1.11
NPB	236	1.15	0.36	0.59	5.30
TEACHING	315	0.44	0.50	0	1
BEDS	306	338.21	297.49	10	1,497
POP	315	0.34	0.80	0.02	2.81

5. STOCHASTIC FRONTIER RESULTS

The estimation results of our stochastic cost frontier regression as equation (6) specifies come in the first column of Table 3. All estimates are found to be strongly statistically significant and present the expected signs. Table 4 gives the subsequent cost efficiency ranking for the 63 mainland Portugal public hospitals. Rather than adding the cost inefficiency score for each hospital, Table 4 displays the matching – what we choose to call – potential cost reduction (PCR) index in percentage which indicates the potential cost reduction each hospital can aim to arrive at if it shifts downward onto the computed cost frontier whilst still be producing its actual amount of output. Particularly for policy purposes, we find it more appealing to present the PCR index instead of the individual inefficiency scores. The PCR index is obtained through

substracting each individual efficiency score in percentage to one hundred. Also each individual efficiency score is the inverse of the sum of scalar one plus the relevant cost inefficiency estimate. And the sum of scalar one plus the relevant cost inefficiency estimate is given by the exponential of the fitted value for *u*. The measure of cost inefficiency is the amount by which the exponential of the fitted value for *u* exceeds 1.

Table 3. Estimation results

	Stochastic cost frontier model			CRE model with
	SF full sample	SF ≥ 150 beds		heteroscedasticity
Constant (α)	3.5833	2.9299	*	1.4265
	(0.5742)	(1.4655)		(0.4014)
IS	0.3261	0.2593		0.5103
	(0.0527)	(0.0886)		(0.0186)
LOS	0.4691	0.4334		0.8958
	(0.0884)	(0.1279)		(0.0317)
VISITS	0.4607	0.4157		0.4094
	(0.0474)	(0.0665)		(0.0209)
PL	0.3787 *	0.5462		0.3832
	(0.0380)	(0.1224)		(0.0386)
PK	0.6213	0.4538		0.6168
	(0.0380)	(0.1224)		(0.0386)
NPB	0.1890	0.3968		0.4513
	(0.0335)	(0.0776)		(0.0250)
TEACHING	0.1311	0.0898	*	0.0781
	(0.0486)	(0.0415)		(0.0117)
BEDS	0.0006	0.0006		0.0002
	(0.0001)	(0.0002)		(0.0000)
POP	0.0752	0.0693		0.0547
	(0.0315)	(0.0260)		(0.0066)
μ	0.3820	0.3715		
	(0.1231)	(0.1275)		
$\sigma^2 = \sigma_u^2 + \sigma_v^2$	0.0361	0.0232		
	(0.0083)	(0.0054)		
$\gamma = \sigma_u^2/(\sigma_u^2 + \sigma_v^2)$	0.8106	0.7632		
	(0.0500)	(0.0687)		
σ_u^2	0.0292	0.0177		
	(0.0084)	(0.0056)		
σ_v^2	0.0068	0.0055		
	(0.0008)	(0.0008)		
Log – likelihood	153.78	131.17		198.17
# of hospitals	63	42		63
Number of observations	221	155		221

The standard errors come in parentheses. All estimates are found to be statistically significant at 1% significance level, except the ones made noticeable by means of a star (*) which are statistically significant at 5% only. The corrected random effects feasible generalised least squares constant term and the PK parameter estimates were computed separately as explained in the main text above.

Table 4. Cost efficiency ranking and potential cost reduction (PCR) index (%) from the stochastic frontier model for the 63 mainland Portugal public hospitals with 150 beds or more (SF full sample – computing the model's coefficients by making use of the whole 63 hospitals sample)

Rank	Hospitals	PCR index	BEDS
1	Hospital Nossa Senhora da Assuncao (Seia)	3.9	< 150
2	Hospital Distrital de Pombal	4.6	< 150
3	Centro Hospitalar de Lisboa – Zona Central	7.4	> 1,000
4	Hospitais da Universidade de Coimbra	8.4	> 1,000
5	Centro Hospitalar do Nordeste, EPE	9.2	
6	Centro Hospitalar do Alto Ave, EPE	9.8	
7	Hospital Distrital de Sao Joao da Madeira	10.8	< 150

Table 4 (...)

Rank	Hospitals	PCR index	BEDS
8	Hospital Arcebispo Joao Crisostomo (Cantanhede)	11.8	< 150
9	Centro Hospitalar do Alto Minho, EPE	12.3	
10	Hospital Santa Maria Maior, EPE (Barcelos)	12.3	
11	Hospital Padre Americo Vale do Sousa, EPE (Penafiel)	13.3	
12	Hospital Sao Miguel (Oliveira de Azemeis)	13.5	< 150
13	Hospital de Nossa Senhora da Ajuda (Espinho)	14.3	< 150
14	Hospital Distrital da Figueira da Foz, EPE	14.7	
15	Hospital Nossa Senhora da Conceicao (Valongo)	15.7	< 150
16	Hospital Bernardino Lopes de Oliveira (Alcobaca)	15.7	< 150
17	Hospital Sao Goncalo (Amarante)	15.8	
18	Hospital Francisco Zagalo (Ovar)	15.9	< 150
19	Hospital Visconde Salreu (Estarreja)	16.7	< 150
20	Hospital Distrital de Agueda	17.4	< 150
21	Hospital Conde de Sao Bento (Santo Tirso)	17.4	< 150
22	Hospital Sao Sebastiao, EPE (Santa Maria da Feira)	17.7	
23	Centro Hospitalar de Lisboa Ocidental, EPE	17.7	
24	Centro Hospitalar de Setubal, EPE	20.5	
25	Hospital Sao Joao de Deus (Vila Nova de Famalicao)	20.6	
26	Hospital Candido de Figueiredo (Tondela)	20.6	< 150
27	Hospital Distrital de Lamego	22.3	< 150
28	Centro Hospitalar Povoa do Varzim – Vila do Conde, EPE	23.0	< 150
29	Hospital de Jose Luciano de Castro (Anadia)	24.5	< 150
30	Hospital Curry Cabral (Lisboa)	24.6	
31	Hospital de Sao Teotonio, EPE (Viseu)	25.0	
32	Hospital de Sousa Martins (Guarda)	25.1	
33	Hospital de Santa Marta, EPE (Lisboa)	26.0	
34	Hospital Infante Dom Pedro, EPE (Aveiro)	26.1	
35	Hospital do Montijo	26.1	< 150
36	Hospital de Santa Maria, EPE (Lisboa)	26.2	> 1,000
37	Hospital Distrital de Faro	26.4	
38	Hospital Sao Pedro Goncalves de Telmo (Peniche)	26.8	< 150
39	Hospital de Santa Luzia (Elvas)	27.5	< 150
40	Hospital de Reynaldo dos Santos (Vila Franca de Xira)	27.9	
41	Hospital Amato Lusitano (Castelo Branco)	28.0	
42	Hospital Doutor Jose Maria Grande (Portalegre)	28.1	
43	Centro Hospitalar Vila Real – Peso da Regua, EPE	28.4	
44	Centro Hospitalar de Vila Nova de Gaia	28.5	
45	Centro Hospitalar de Cascais	29.4	
46	Hospital Distrital de Chaves	29.8	
47	Hospital de Sao Marcos (Braga)	30.8	
48	Centro Hospitalar da Cova da Beira, EPE	30.8	
49	Hospital de Sao Joao, EPE (Porto)	30.8	> 1,000
50	Centro Hospitalar de Torres Vedras	31.4	
51	Centro Hospitalar do Medio Tejo, EPE	31.5	
52	Hospital Santo Andre, EPE (Leiria)	32.1	
53	Hospital Pulido Valente, EPE (Lisboa)	32.3	
54	Hospital Distrital de Santarem	33.1	
55	Centro Hospitalar de Coimbra	34.3	
56	Hospital do Espirito Santo (Evora)	35.3	
57	Hospital do Litoral Alentejano (Santiago do Cacem)	35.7	< 150
58	Centro Hospitalar de Caldas da Rainha	36.7	< 150
59	Centro Hospitalar do Barlavento Algarvio, EPE	37.6	
60	Hospital Nossa Senhora do Rosario, EPE (Barreiro)	39.1	
61	Centro Hospitalar do Baixo Alentejo, EPE	39.2	
62	Hospital Geral de Santo Antonio, EPE (Porto)	43.4	
63	Hospital Garcia de Orta, EPE (Almada)	44.5	

5.1. Size does not matter

As shown in Table 2, whilst the smallest hospital in the sample comprises 10 beds only, the biggest one counts almost 1,500 beds available. One may effortlessly anticipate that smaller and larger hospitals would not face the same technology, or in other words, that each sort would chase a different cost function. It should be after all quite intuitive that bigger university hospitals would treat a comparatively sterner patient mix, eventually also with relatively short but expensive hospitalisations.

We go one step further in our analysis and classify the hospitals according to their size. The hospitals are "small" when they have less than 150 beds available and "huge" when they count more than 1,000 beds. The remaining hospitals are put as being of "medium size". The cutting points were determined instinctively as follows. On one hand, the lowest cutting point should keep out a somewhat considerable number of hospitals but at the same time should not be too extensive or we would lose the most significant part of the sample. The lowest cutting point at 150 beds excludes roughly one-third of the hospitals. On the other hand, the main concern about the highest cutting point would be to assemble the three foremost central university hospitals, each one located in one different leading town in Portugal. Settling the highest cutting point at 1,000 beds accomplishes that goal. Also it gathers not three but the four largest hospitals.

Bearing in mind the hospital size categorisation as above, we went back to the cost efficiency ranking in Table 4. Whilst the hospitals ordering first and second are "small" hospitals, the hospitals lying in the third and the fourth positions are "huge" hospitals. Also a "medium size" hospital lies in the fifth place. When looking at the bottom of the ranking, one finds "medium size" hospitals only in the last five positions. This suggests that size does not matter after all, and in the end "small", "medium size" and "huge" hospitals hunt the same technology.

We next ran the stochastic cost frontier regression as in equation (6) on a sample which had previously excluded all "small" hospitals, and sorted out the lasting larger hospitals according to their cost efficiency scores. The regression results are in the second column of Table 3. Table 5 exhibits the consequent cost efficiency ranking for the 42 mainland Portugal public hospitals with 150 beds available or more. Table 6 also presents a cost efficiency ranking for those 42 larger hospitals but its ranking stems from running the regression on the full 63 hospitals sample, and so this latter efficiency ranking does no more than ruling out the 21 "small" hospitals in Table 4. Plus Table 5 and Table 6 give the PCR index for each hospital.

Table 5. Cost efficiency ranking and potential cost reduction (PCR) index (%) from the stochastic frontier model for the 42 mainland Portugal public hospitals with 150 beds or more (SF ≥ 150 beds – computing the model's coefficients by making use of the 42 largest hospitals sample only)

Rank	Hospitals	PCR index	BEDS
1	Hospital Sao Goncalo (Amarante)	5.9	
2	Centro Hospitalar de Lisboa – Zona Central	8.4	> 1,000
3	Centro Hospitalar do Nordeste, EPE	8.7	
4	Hospital Distrital da Figueira da Foz, EPE	9.1	
5	Hospital Santa Maria Maior, EPE (Barcelos)	9.5	
6	Hospitais da Universidade de Coimbra	11.4	> 1,000
7	Centro Hospitalar do Alto Minho, EPE	12.9	
8	Centro Hospitalar do Alto Ave, EPE	13.2	
9	Hospital Padre Americo Vale do Sousa, EPE (Penafiel)	13.9	
10	Hospital Sao Joao de Deus (Vila Nova de Famalicao)	16.3	
11	Centro Hospitalar de Lisboa Ocidental, EPE	19.5	
12	Hospital Sao Sebastiao, EPE (Santa Maria da Feira)	20.1	
13	Centro Hospitalar de Setubal, EPE	20.6	
14	Hospital de Santa Marta, EPE (Lisboa)	20.7	
15	Hospital de Sousa Martins (Guarda)	20.8	
16	Hospital Doutor Jose Maria Grande (Portalegre)	22.8	
17	Hospital Distrital de Chaves	23.0	
18	Hospital de Reynaldo dos Santos (Vila Franca de Xira)	23.5	

Table 5 (…)

Rank	Hospitals	PCR index	BEDS
19	Centro Hospitalar de Vila Nova de Gaia	24.5	
20	Hospital Amato Lusitano (Castelo Branco)	24.6	
21	Centro Hospitalar Vila Real – Peso da Regua, EPE	25.1	
22	Hospital Infante Dom Pedro, EPE (Aveiro)	25.5	
23	Hospital de Sao Teotonio, EPE (Viseu)	25.7	
24	Hospital Curry Cabral (Lisboa)	25.9	
25	Hospital Distrital de Faro	26.8	
26	Hospital de Santa Maria, EPE (Lisboa)	27.4	> 1,000
27	Centro Hospitalar de Torres Vedras	27.9	
28	Centro Hospitalar da Cova da Beira, EPE	28.0	
29	Centro Hospitalar de Cascais	28.5	
30	Centro Hospitalar do Medio Tejo, EPE	28.6	
31	Hospital de Sao Joao, EPE (Porto)	28.8	> 1,000
32	Hospital Santo Andre, EPE (Leiria)	29.5	
33	Hospital Pulido Valente, EPE (Lisboa)	29.8	
34	Hospital de Sao Marcos (Braga)	30.1	
35	Hospital do Espirito Santo (Evora)	30.8	
36	Hospital Distrital de Santarem	32.6	
37	Centro Hospitalar de Coimbra	33.1	
38	Centro Hospitalar do Barlavento Algarvio, EPE	33.1	
39	Hospital Nossa Senhora do Rosario, EPE (Barreiro)	34.1	
40	Centro Hospitalar do Baixo Alentejo, EPE	36.7	
41	Hospital Geral de Santo Antonio, EPE (Porto)	38.7	
42	Hospital Garcia de Orta, EPE (Almada)	40.7	

Table 6. Cost efficiency ranking and potential cost reduction (PCR) index (%) from the stochastic frontier model for the 42 mainland Portugal public hospitals with 150 beds or more (SF full sample – computing the model's coefficients by making use of the whole 63 hospitals sample)

Rank	Hospitals	PCR index	BEDS
1	Centro Hospitalar de Lisboa – Zona Central	7.4	> 1,000
2	Hospitais da Universidade de Coimbra	8.4	> 1,000
3	Centro Hospitalar do Nordeste, EPE	9.2	
4	Centro Hospitalar do Alto Ave, EPE	9.8	
5	Centro Hospitalar do Alto Minho, EPE	12.3	
6	Hospital Santa Maria Maior, EPE (Barcelos)	12.3	
7	Hospital Padre Americo Vale do Sousa, EPE (Penafiel)	13.3	
8	Hospital Distrital da Figueira da Foz, EPE	14.7	
9	Hospital Sao Goncalo (Amarante)	15.8	
10	Hospital Sao Sebastiao, EPE (Santa Maria da Feira)	17.7	
11	Centro Hospitalar de Lisboa Ocidental, EPE	17.7	
12	Centro Hospitalar de Setubal, EPE	20.5	
13	Hospital Sao Joao de Deus (Vila Nova de Famalicao)	20.6	
14	Hospital Curry Cabral (Lisboa)	24.6	
15	Hospital de Sao Teotonio, EPE (Viseu)	25.0	
16	Hospital de Sousa Martins (Guarda)	25.1	
17	Hospital de Santa Marta, EPE (Lisboa)	26.0	
18	Hospital Infante Dom Pedro, EPE (Aveiro)	26.1	
19	Hospital de Santa Maria, EPE (Lisboa)	26.2	> 1,000
20	Hospital Distrital de Faro	26.4	
21	Hospital de Reynaldo dos Santos (Vila Franca de Xira)	27.9	
22	Hospital Amato Lusitano (Castelo Branco)	28.0	
23	Hospital Doutor Jose Maria Grande (Portalegre)	28.1	
24	Centro Hospitalar Vila Real – Peso da Regua, EPE	28.4	
25	Centro Hospitalar de Vila Nova de Gaia	28.5	
26	Centro Hospitalar de Cascais	29.4	
27	Hospital Distrital de Chaves	29.8	
28	Hospital de Sao Marcos (Braga)	30.8	

Table 6 (…)

Rank	Hospitals	PCR index	BEDS
29	Centro Hospitalar da Cova da Beira, EPE	30.8	
30	Hospital de Sao Joao, EPE (Porto)	30.8	> 1,000
31	Centro Hospitalar de Torres Vedras	31.4	
32	Centro Hospitalar do Medio Tejo, EPE	31.5	
33	Hospital Santo Andre, EPE (Leiria)	32.1	
34	Hospital Pulido Valente, EPE (Lisboa)	32.3	
35	Hospital Distrital de Santarem	33.1	
36	Centro Hospitalar de Coimbra	34.3	
37	Hospital do Espirito Santo (Evora)	35.3	
38	Centro Hospitalar do Barlavento Algarvio, EPE	37.6	
39	Hospital Nossa Senhora do Rosario, EPE (Barreiro)	39.1	
40	Centro Hospitalar do Baixo Alentejo, EPE	39.2	
41	Hospital Geral de Santo Antonio, EPE (Porto)	43.4	
42	Hospital Garcia de Orta, EPE (Almada)	44.5	

The cost efficiency estimates and the cost efficiency rankings of Table 5 and Table 6 are particularly helpful for comparison purposes. From both one can check that the absolute variation in the individual efficiency scores which occurs as of shifting from the full 63 hospitals to the 42 larger hospitals sample does not exceed 10% and it is in average less than 3%. Also no hospital among the 42 comparable hospitals jumps more than 10 ranks when switching from one efficiency ranking to the other in Table 5 and in Table 6. Plus the hospitals which lay in the first ten and in the last five positions are the same in each efficiency ranking, and although the first ten hospitals switch places from one ranking to the other, the last five hospitals rank exactly the same in both cases. In addition, the series of individual efficiency scores are highly correlated and similar high correlation coefficients turn up as computing the correlations of the individual efficiency ranks. The Pearson product-moment correlation coefficient of the series of efficiency scores (ranks) lays at 95.5% (93.5%), the Spearman's rank correlation coefficient is at 94.5% (94.5%) and the Kendall's rank correlation coefficient is at 80.7% (80.7%).

In addition, Figure 1 and Figure 2 depict the histograms of the series of cost efficiency estimates which derive from the full sample plus the sample with no "small" hospital, in that order. The diagrams show equivalent distributions of the cost efficiency estimates which suggest that most hospitals are either comparatively efficient or are close to being comparatively efficient as in both diagrams they gather on the right-hand side of the distribution.

Figure 1. Distribution of cost efficiency in the stochastic frontier model (SF full sample)

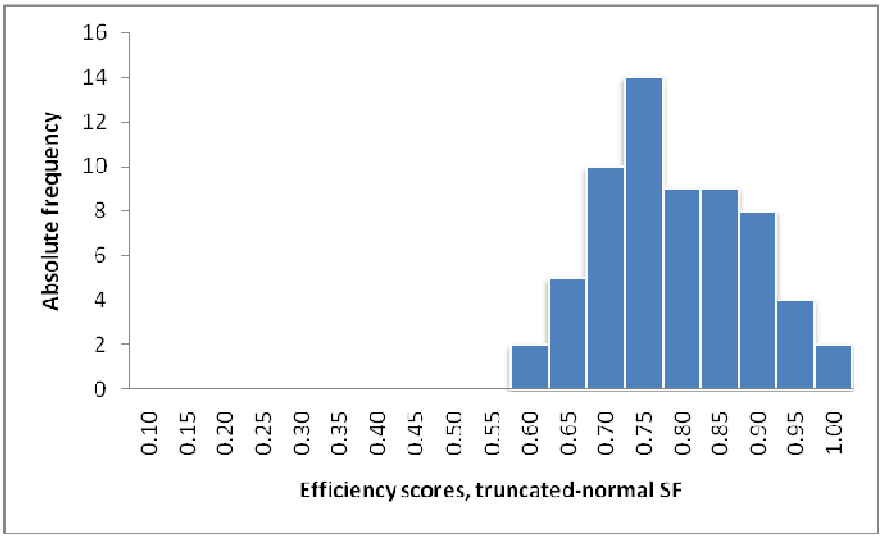
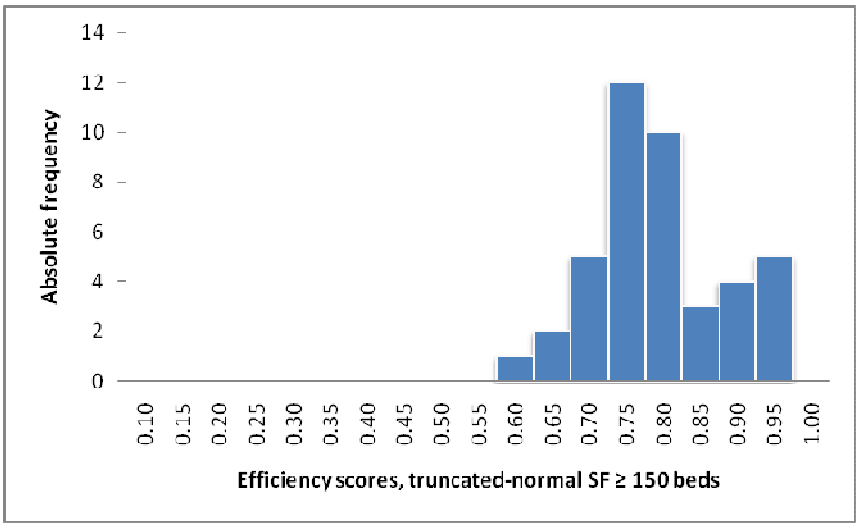


Figure 2. Distribution of cost efficiency in the stochastic frontier model (SF ≥ 150 beds)



All this should be sufficient evidence then to conclude size does not matter and all 63 hospitals in the full sample share the same technology.

5.2. Head-to-head with the standard random effects estimator results

The estimation results of the standard random effects regression as in expression (6) are extremely close to the results we discover by using the stochastic cost frontier estimator. The results are in fact so identical that we find it pointless to present them here in greater detail.

It should not be such a disclosure that the estimation results under the stochastic frontier framework match so well the results one finds when using the standard random effects estimator. In the end there are only up to two tiny differences between both settings. On one hand, while inefficiency u is required to be truncated-normal under the stochastic cost frontier setting, it is simply asked to be independently identically distributed with zero mean and constant variance for the generalised ordinary least squares standard random effects estimation to be computed. On the other hand, whilst estimation is done by maximum likelihood in the stochastic cost frontier framework, computation is performed by generalised least squares under the random effects model. Besides provided that inefficiency u follows the same truncated-normal distribution and the error random component v is equally normally distributed, there should be no difference at all when the standard random effects estimation is also done by maximum likelihood.

6. CORRECTED RANDOM EFFECTS MODEL WITH HETEROSCEDASTICITY

Up to now all outcomes emerged by assuming homoscedasticity across organisations. Earlier in our discussion though we mentioned the heteroscedastic one-way group random effects model ought to be preferred over the random effects model which does not include heteroscedasticity across organisations. Of course the likelihood-ratio test which pointed towards heteroscedasticity does not directly oppose the stochastic frontier model as defined above to the one-way group random effects model with heteroscedasticity. The test opposes two generalised ordinary least squares random effects models whose one and only difference lays in the fact that one model considers homoscedasticity whilst the other includes heteroscedasticity across organisations. But as the standard homoscedastic random effects model roughly replicates the stochastic frontier estimates, we take both models as being quite equivalent and infer perhaps abusively that our discussion should at least also take into account the heteroscedastic one-way group random effects estimates. As mentioned above, this should be a

key aspect since under the stochastic frontier framework heteroscedasticity should lead to biased estimation of both the model and the level of cost efficiency.

6.1. Corrected random effects procedure

Mostly as STATA version 9.2 we are making use of in this work does not fit stochastic frontier panel data models with heteroscedasticity across groups, we went onwards and computed ourselves a corrected random effects (CRE) cost frontier through generalised least squares and by assuming heteroscedasticity across groups. Inspired by the cross-sectional data context corrected ordinary least squares (COLS) procedure which gets the cost frontier by adding the lowest negative residual value to the constant term of the ordinary least squares regression and substracting it from the residuals, we concretely went ahead as follows.

Settling organisation *i* as being the organisation with the lowest negative residual value, the cross-sectional data context COLS adjustment procedure may be summarised as in expression (7) below.

$$\ln C_j = \left(\hat{\alpha} + \hat{e}_{i*}\right) + \sum_k \hat{\beta}_k \ln x_{kj} + \sum_l \hat{\phi}_l z_{lj} + \left(\hat{e}_j - \hat{e}_{i*}\right) \tag{7}$$

As always, the "^" symbol stands for fitted values. Besides $e_j \equiv v_j + u_j$ depicts the whole departures from the cost frontier and $\hat{e}_{i*} \equiv \min(\hat{e}_j)$ for all j is the computed residual of organisation i which we look at as being the one with the lowest negative residual value.

With panel data, as there are up to T observations for each organisation, one has to choose among different sorts of "lowest negative" residuals so as to identify the alleged most efficient organisation and depict the estimated cost frontier. Although using the lowest absolute negative residual would yield a more demanding technology, we end up preferring the lowest negative residual mean as it contains more information about the most (mean) efficient organisation. Taking organisation i as having the lowest negative residual mean, we denote the lowest residual mean as:

$$\bar{\hat{e}}_{i\bullet *} \equiv \min\left(\bar{\hat{e}}_{j\bullet}\right), \text{ for all } j. \tag{8}$$

The CRE cost frontier may hence be written as:

$$\ln C_{jt} = \underbrace{\left(\hat{\alpha} + \bar{\hat{e}}_{i\bullet *}\right)}_{\tilde{\alpha}} + \sum_k \hat{\beta}_k \ln x_{kjt} + \sum_l \hat{\phi}_l z_{ljt} + \left(\hat{e}_{jt} - \bar{\hat{e}}_{i\bullet *}\right) \tag{9}$$

Notice the cost frontier is positioned by assuming the organisation with the (computed) lowest intercept is the most efficient, down to the point where no residual is negative and only the most efficient organisation's residual is zero, as in the COLS approach.

Moreover, the COLS efficiency ranking is built with the corrected computed residuals $\hat{e}_j - \hat{e}_{i*}$. With panel data, again as there are up to T observations for each organisation, we rank organisations by means of their up to T corrected computed residuals mean, which may be put as follows:

$$\frac{\sum_t \hat{e}_{jt} - \bar{\hat{e}}_{i\bullet *}}{T} = \bar{\hat{e}}_{j\bullet} - \bar{\hat{e}}_{i\bullet *} \tag{10}$$

Remark the non-negative departures from the cost frontier in (10) may also be denoted by making use of the stochastic frontier framework notation for inefficiency as:

$$\hat{u}_j \equiv \bar{\hat{e}}_{j\bullet} - \bar{\hat{e}}_{i\bullet *} \tag{11}$$

Also the fitted random component must in that case be:

$$\hat{v}_{jt} = \hat{e}_{jt} - \bar{\hat{e}}_{j\bullet} + \bar{\hat{e}}_{i\bullet *} \tag{12}$$

6.2. Results under heteroscedasticity

The regression results of the CRE cost frontier model are in the third column of Table 3. They differ quite considerably from the homoscedastic stochastic frontier estimates, on hand in the first column of Table 3. Table 7 gives the ensuing CRE heteroscedastic cost efficiency ranking for the 63 mainland Portugal public hospitals. It adds the PCR index for each hospital too.

Table 7. Cost efficiency ranking and potential cost reduction (PCR) index (%) from the heteroscedastic corrected random effects frontier model for the 63 mainland Portugal public hospitals

Rank	Hospitals	PCR index	BEDS
1	Hospital Distrital de Pombal	0.0	< 150
2	Hospital Nossa Senhora da Assuncao (Seia)	8.5	< 150
3	Centro Hospitalar do Alto Minho, EPE	11.3	
4	Hospital Sao Goncalo (Amarante)	22.6	
5	Centro Hospitalar do Alto Ave, EPE	24.2	
6	Hospital Arcebispo Joao Crisostomo (Cantanhede)	25.3	< 150
7	Hospital de Nossa Senhora da Ajuda (Espinho)	26.3	< 150
8	Hospital Distrital da Figueira da Foz, EPE	27.0	
9	Centro Hospitalar Vila Real – Peso da Regua, EPE	27.6	
10	Hospital Nossa Senhora da Conceicao (Valongo)	28.3	< 150
11	Hospital Conde de Sao Bento (Santo Tirso)	28.4	< 150
12	Centro Hospitalar da Cova da Beira, EPE	28.7	
13	Hospital Padre Americo Vale do Sousa, EPE (Penafiel)	29.0	
14	Hospital de Sao Teotonio, EPE (Viseu)	29.0	
15	Hospital Distrital de Sao Joao da Madeira	29.4	< 150
16	Hospital Infante Dom Pedro, EPE (Aveiro)	30.3	
17	Hospitais da Universidade de Coimbra	30.6	> 1,000
18	Centro Hospitalar de Lisboa Ocidental, EPE	31.0	
19	Hospital Santo Andre, EPE (Leiria)	31.7	
20	Hospital de Jose Luciano de Castro (Anadia)	32.2	< 150
21	Hospital Distrital de Lamego	32.4	< 150
22	Hospital de Sao Marcos (Braga)	33.1	
23	Hospital Santa Maria Maior, EPE (Barcelos)	33.3	
24	Hospital Sao Miguel (Oliveira de Azemeis)	33.5	< 150
25	Hospital Sao Sebastiao, EPE (Santa Maria da Feira)	33.6	
26	Hospital Pulido Valente, EPE (Lisboa)	33.8	
27	Hospital de Sousa Martins (Guarda)	34.9	
28	Hospital de Sao Joao, EPE (Porto)	34.9	> 1,000
29	Centro Hospitalar do Nordeste, EPE	35.0	
30	Centro Hospitalar de Lisboa – Zona Central	35.0	> 1,000
31	Hospital Distrital de Agueda	35.8	< 150
32	Hospital Bernardino Lopes de Oliveira (Alcobaca)	36.0	< 150
33	Centro Hospitalar Povoa do Varzim – Vila do Conde, EPE	36.1	< 150
34	Hospital Visconde Salreu (Estarreja)	36.2	< 150
35	Hospital de Santa Maria (Lisboa)	36.3	> 1,000
36	Hospital Distrital de Chaves	36.4	
37	Hospital Distrital de Santarem	36.6	
38	Hospital Curry Cabral (Lisboa)	37.4	
39	Hospital de Santa Marta, EPE (Lisboa)	37.4	
40	Hospital Doutor Jose Maria Grande (Portalegre)	37.4	
41	Centro Hospitalar de Vila Nova de Gaia	37.8	
42	Hospital Candido de Figueiredo (Tondela)	37.8	< 150
43	Hospital Francisco Zagalo (Ovar)	37.9	< 150
44	Hospital do Espirito Santo (Evora)	38.0	
45	Hospital Nossa Senhora do Rosario, EPE (Barreiro)	38.1	
46	Centro Hospitalar do Barlavento Algarvio, EPE	39.0	
47	Hospital Distrital de Faro	39.3	

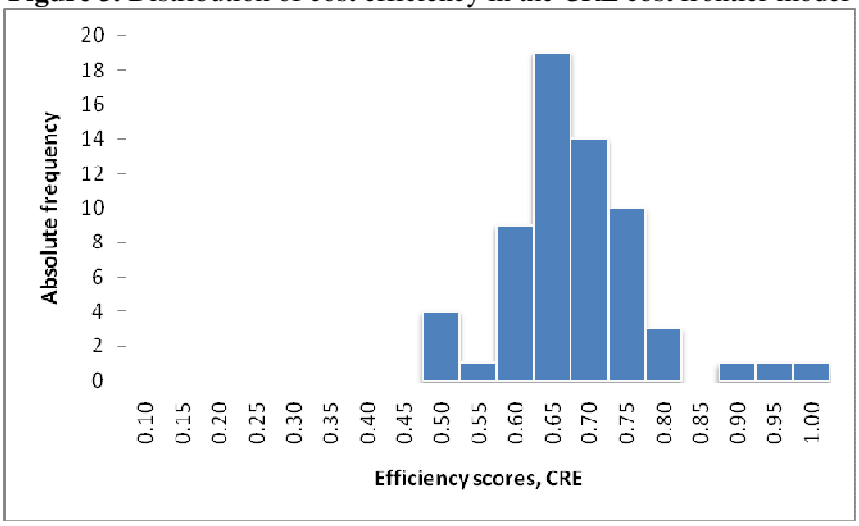
Table 7 (…)

Rank	Hospitals	PCR index	BEDS
48	Hospital Geral de Santo Antonio, EPE (Porto)	39.4	
49	Hospital de Reynaldo dos Santos (Vila Franca de Xira)	39.4	
50	Hospital Sao Joao de Deus (Vila Nova de Famalicao)	41.3	
51	Hospital de Santa Luzia (Elvas)	41.4	< 150
52	Hospital Garcia de Orta, EPE (Almada)	42.3	
53	Hospital do Montijo	43.3	< 150
54	Centro Hospitalar de Torres Vedras	43.9	
55	Centro Hospitalar do Medio Tejo, EPE	44.3	
56	Centro Hospitalar de Cascais	44.4	
57	Centro Hospitalar de Setubal, EPE	44.6	
58	Hospital Amato Lusitano (Castelo Branco)	44.8	
59	Centro Hospitalar de Coimbra	48.3	
60	Centro Hospitalar de Caldas da Rainha	51.5	< 150
61	Centro Hospitalar do Baixo Alentejo, EPE	51.8	
62	Hospital Sao Pedro Goncalves de Telmo (Peniche)	52.9	< 150
63	Hospital do Litoral Alentejano (Santiago do Cacem)	53.1	< 150

Focus on the cost efficiency scores and rankings of Table 4 and Table 7. From both one can check that even though in average the absolute variation in the individual efficiency estimates which turns up when shifting from the homoscedastic stochastic frontier to the CRE heteroscedastic specification is 12%, not only the absolute variation in the scores is up to 28% but also it exceeds 17% for a quarter of the 63 hospitals in the sample. Plus in average hospitals jump 12 positions from one efficiency ranking to the other in Table 4 and in Table 7. The biggest jump is of the Cova da Beira hospital centre which moves 36 positions from position 48 in the stochastic frontier setting up to the position 12 in the heteroscedastic specification. Bearing in mind the hospital size classification we outlined earlier, the Cova da Beira hospital centre is a "medium size" hospital. But there are also big jumps of both "huge" and "small" hospitals. The Lisboa hospital centre for its central perimeter is a "huge" hospital which shifts 27 positions from position 3 in the homoscedastic ranking down to position 30 in the heteroscedastic framework. And the Francisco Zagalo hospital is a "small" hospital which moves 25 positions from rank 18 in the stochastic frontier context down to rank 43 in the CRE ranking. Besides a quarter of the 63 hospitals in the sample shift more than 17 ranks. Additionally, the individual efficiency estimates are just weakly correlated and equivalent correlations come up for the individual efficiency ranks. The Pearson product-moment correlation coefficient of the series of individual efficiency scores (ranks) lays at 65.4% (64.8%), the Spearman's rank correlation coefficient is at 64.7% (64.6%) and the Kendall's rank correlation coefficient is at 46.9% (46.8%). Also hospitals of similar size are much more evenly distributed across the heteroscedastic ranking. While two out of 21 "small" hospitals are in the third quartile and two other "small" are in the fourth quartile of the homoscedastic ranking, five out of the same 21 "small" hospitals are in the third quartile and five other "small" are in the fourth quartile of the heteroscedastic ranking. Therefore ten "small" hospitals lay in the second half of the heteroscedastic ranking which strongly contrasts with just four "small" hospitals in the second half of the homoscedastic ranking. And while two out of four "huge" hospitals are in the first quartile of the homoscedastic ranking, there is no "huge" hospital in the first quartile of the heteroscedastic ranking.

Figure 3 gives the histogram of the series of the cost efficiency estimates which derive from the CRE heteroscedastic cost frontier specification. The diagram strongly contrasts with the diagram in Figure 1. Recall that the diagram in Figure 1 depicts the distribution of the cost efficiency estimates which come up from the homoscedastic stochastic frontier setting. Plus it implies that most hospitals either are comparatively efficient or are near to being so. On the contrary, the diagram in Figure 3 suggests that only a small number of hospitals are comparatively efficient.

Figure 3. Distribution of cost efficiency in the CRE cost frontier model



All this should be enough proof to conclude there are important differences between the homoscedastic stochastic and the CRE heteroscedastic cost frontier results. Taking both the standard homoscedastic random effects model and the homoscedastic stochastic cost frontier model as being equivalent as we discussed, either the homoscedastic (stochastic frontier) estimates are biased if the variance is not constant across groups or taking heteroscedasticity in whilst there is none in the data has a strong confusing impact on the results. This finding is in line with several authors' equivalent reasoning (Guermat and Hadri 1999; Hadri, Guermat and Whittaker 2003).

Several authors have argued that results strongly depend upon the choice of whether to compute a stochastic rather than a COLS frontier (Street 2003; Jacobs et al. 2006). The motive is that the COLS frontier yields considerably lower efficiency estimates because the COLS residuals include both the random shocks and inefficiency. As the CRE cost frontier procedure is deeply inspired by the COLS procedure, it might be worth at this point asking too whether the different results we get are more the effect of shifting from the stochastic frontier to a specification which is inspired by the COLS specification or are the different results the effect of either not putting heteroscedasticity in the model whilst it is present in the data or considering it in the model whilst in fact there is none.

There is no reason for the CRE cost frontier procedure to yield efficiency scores that are lower than the efficiency scores which stem from the stochastic frontier specification. In contrast with the COLS procedure, in the CRE procedure the "lowest negative residual" which is subtracted from the residuals is actually the lowest negative residual mean, as we pointed out earlier. If one believes the random shocks are normally distributed with zero mean, as in the stochastic frontier setting, the bias caused by the random shocks on the efficiency estimates vanishes as the number of observations for each group increases through time.

Additionally, Street (2003) reveals that although there is in average little effect on the efficiency scores when switching from a stochastic frontier to a COLS specification, the effect is substantial when looking at the individual efficiency scores. As we extensively discussed above, there is not only a substantial effect on the individual efficiency scores but also the effect is in average substantial when shifting from the homoscedastic to the CRE heteroscedastic framework.

One should then be confident enough that the important differences between the results from the homoscedastic stochastic and the CRE heteroscedastic frontier frameworks are caused by the assumption of heteroscedasticity.

But which model should one pick? The exactness of the cost efficiency scores matters a lot as cost efficiency scores usually affect policy decisions. In what concerns our data set, putting together the alleged equivalence between both homoscedastic models (standard random effects time invariant and stochastic frontier), the likelihood-ratio test which pointed towards

heteroscedasticity, also the existence of heteroscedasticity in the data which is often related to size attributes of the organisations in the sample, and the fact that hospitals we observe differ very much in size, we conclude heteroscedasticity is most likely in our data set. For that reason, the CRE cost frontier model should perhaps be the most appropriate.

7. OVERALL COST EFFICIENCY AND ECONOMIES of Scale

Table 8 describes the summary statistics of the efficiency scores. The summary statistics of the efficiency scores from the stochastic cost frontier model are in the first line. The second line presents the efficiency scores of the CRE cost frontier model. The efficiency scores give the fraction of the actual costs the hospitals would face if all hospitals in the sample operated as efficiently as the most efficient hospital. According to the efficiency scores from the stochastic frontier setting, the hospitals would in average incur in a little more of 76% of their actual costs if all hospitals shifted downward onto the computed cost frontier and with no cuts in the output produced. The mean efficiency score of the CRE cost frontier model indicates an overall average potential cost reduction of 35%, which is larger than the potential cost reduction given by the stochastic frontier model. Furthermore, according the stochastic frontier model, a quarter of the sample could accomplish a cost cut of 30%. With an efficiency score of 61% for the first quarter of the sample, the CRE cost frontier model reports a potential cost reduction which is even larger and up to 40% for that quarter of the hospitals considered.

Table 8. Summary statistics of the efficiency scores (%)

Model	Mean	SD	Min	1st Q	Median	3rd Q	Max
SF full sample	76.4	9.7	55.5	69.2	74.9	84.3	96.1
CRE	65.1	9.5	47.0	60.6	64.0	69.7	100.0

The actual costs of all hospitals put together was in average 3,696 million Euros each of the five years we look at in this work. If for policy reasons one takes the product of this number and the mean potential cost reduction which must lie between the 24% of the stochastic frontier model and the 35% of the CRE frontier model, then the overall likely cost reduction for the Portuguese NHS with no cuts whatsoever in the output produced would be up between 872 to 1,290 million Euros annually, with all 63 hospitals in the sample moving downward onto the efficient cost frontier. This corresponds to between 11% and 17% of the Portuguese NHS year 2006 initial endowment. Or put this the other way around, and admit at this point no achievable scale gains, shifting all 63 hospitals downward onto the efficient cost frontier and keeping the hospitals' costs at their actual levels would enable to discharge between 24% and 35% more inpatients each year. And so it would be possible to discharge (at least) between 184 and 273 thousands inpatients more annually.

Table 9 provides the summary statistics of the inefficiency scores. As in Table 8, the summary statistics of the inefficiency scores from the stochastic frontier model come in the first line. Likewise the second line presents the inefficiency scores of the CRE cost frontier model. Inefficiency and efficiency scores are of course closely linked. One gets the inefficiency scores of Table 9 simply by inverting the efficiency scores which are in percentages in Table 8 plus substracting scalar 1. In contrast with the interpretation of the efficiency scores, the inefficiency scores express the actual costs of each hospital in the sample in terms of the costs each would face while operating onto the computed cost frontier. From the stochastic frontier setting, the actual costs of the hospitals in our sample were in average 33% larger than the efficient frontier cost level and a quarter of the hospitals exhibit 40% or more excess costs. The inefficiency scores from the CRE cost frontier model are even higher as the actual costs of the hospitals would exceed the efficient frontier cost level in 57% in average whilst a quarter of the sample reveals 65% or more excess costs and the most inefficient hospital more than doubles the efficient frontier cost level.

Table 9. Summary statistics of the inefficiency scores (%)

Model	Mean	SD	Min	1st Q	Median	3rd Q	Max
SF full sample	33.0	17.3	4.0	18.7	33.6	44.5	80.2
CRE	56.7	21.9	0.0	43.5	56.3	65.0	113.1

Tables from Table 10 up to Table 13 give information on the efficiency scores at which a number of authors arrive. Each time in the first column is the country where the under study hospitals are located. The second and the third columns display respectively the number of hospitals and the specific years. Of course not all estimation strategies coincide. All estimates are cost efficiency estimates though and all except one come from stochastic frontier models. The lower cost efficiency score of Street (2003) is from a COLS model. Table 10 and Table 11 display information regarding public hospitals only. Table 12 and Table 13 provide information with public and private hospitals taken jointly. Plus both private "for-profit" and private "not-for-profit" are comprised.

Table 10. Comparison of efficiency scores of public hospitals (%)

Location	#	Years	Mean	Author
Finland	48	1994	86.0	Linna and Hakkinen (1997)
Germany	NS	2001 – 2003	85.0 – 86.0	Frohloff (2007)
Germany	641	2000 – 2003	86.0 – 87.0	Herr (2008)
India	40	2004	40.0	Mathiyazhgan (2006)
Spain	67	1994	78.1	Garcia-Prieto (2001)
Spain	67	1995	68.3	Garcia-Prieto (2003)
UK	226	1995 – 1996	69.4 – 90.3	Street (2003)

Table 11. Further information on efficiency scores of public hospitals (%)

Location	Mean	Min	Max	Author
Finland	86.0	63.0	97.0	Linna and Hakkinen (1997)
Spain	78.1	55.4	90.3	Garcia-Prieto (2001)
Spain	68.3	37.5	98.6	Garcia-Prieto (2003)

Table 12. Comparison of efficiency scores of both public and private ("for-" plus "not-for-profit") hospitals put together (%)

Location	#	Years	Mean	Author
Finland	43	1988 – 1994	83.0 – 93.4	Linna (1998)
Germany	1,565	2001 – 2003	83.0	Frohloff (2007)
Germany	1,594	2000 – 2003	82.0 – 89.0	Herr (2008)
India	99	2004	53.0	Mathiyazhgan (2006)
Spain	43	1988 – 1991	55.8	Wagstaff and Lopez (1996)*
Switzerland	156	1998 – 2001	79.8 – 83.8	Farsi and Filippini (2006)*
Switzerland	148	1998 – 2003	91.2 – 94.5	Farsi and Filippini (2008)*
US	140	1999 – 2001	87.3 – 87.5	Deily and McKay (2006)*
US	1,661	1985	88.1 – 89.8	Folland and Hofler (2001)*
US	791	1985	85.1 – 86.6	Folland and Hofler (2001)*
US	1,631	1990 – 1996	84.0 – 88.2	Rosko (2001)*
US	1,600	1986 – 1987	82.9 – 87.5	Zuckerman et al. (1994)*

The articles marked with a star (*) report cost inefficiency estimates. We extracted all non-reported efficiency scores through inverting the sum of 1 plus the cost inefficiency estimate reported. The sum of scalar 1 plus the cost inefficiency estimate is given by the exponential of the fitted value for u and the measure of inefficiency is the amount by which the exponential of the computed value for u exceeds 1.

Table 13. Further information on efficiency scores of both public and private ("for-" plus "not-for-profit") hospitals put together (%)

Location	Mean	Min	Max	Author
Germany	83.0	12.0	98.0	Frohloff (2007)
Germany	82.0 – 89.0	11.0	98.0	Herr (2008)
Switzerland	79.8 – 83.8	27.7 – 43.0	NS	Farsi and Filippini (2006)
Switzerland	91.2 – 94.5	72.3 – 76.9	97.1 – 98.9	Farsi and Filippini (2008)

There are apparently no boundaries within which one would anticipate to observe either the ampleness or the limits of the interval of variation of the individual efficiency scores reported by the literature. Most authors do not report the interval of variation for scores. But from what we could gather, it seems there is a bit of everything. We find efficiency scores of Finland public hospitals ranging between 63% and 97% (Linna and Hakkinen 1997). In Spain efficiency scores of public hospitals fit between 55% and 90% (Garcia-Prieto 2001) or lay in a larger interval between 38% and 99% (Garcia-Prieto 2003). Working together with both public and private hospitals, the interval of variation of the cost efficiency scores is even broader ranging 11% – 98% with Germany hospitals (Herr 2008) or much slimmer ranging 72% – 97% plus 77% – 99% with Switzerland hospitals (Farsi and Filippini 2008). Our efficiency scores from the 63 Portugal public hospitals move between 55% and 96% with the stochastic frontier framework plus between 47% and 100% with the CRE cost frontier specification. Recall our maximum value of 100% is a direct consequence of the CRE procedure of our own.

But even if there are no definite boundaries either for the ampleness or the limits of the interval of variation of the individual efficiency scores, when looking at mean values it seems hospitals from most OECD countries present mean efficiency scores above 85%. Plus this appears to be so both in studies in which only public hospitals are considered (Street 2003; Frohloff 2007; Herr 2008; Linna and Hakkinen 1997) and in studies where public and private hospitals are considered jointly (Zuckerman et al. 1994; Folland and Hofler 2001; Rosko 2001; Deily and McKay 2006; Frohloff 2007; Herr 2008; Linna 1998; Farsi and Filippini 2008). Among the articles we could gather, Spain and Portugal are the only OECD countries whose hospitals do not reach a mean efficiency estimate of 80%. Spain public hospitals have mean efficiencies of 68% – 78% (Garcia-Prieto 2001; 2003) and Catalan hospitals display a mean efficiency of 56% (Wagstaff and Lopez 1996). This is in value comparable to the mean efficiencies of Portugal public hospitals of 76% with the stochastic frontier model and 65% with the CRE cost frontier setting. Recall no additional conclusion can be drawn from this as we cannot tell whether or not the hospitals in one country and in another share the same technology. It is unlikely that they do so. Also the mean efficiency estimate for England public hospitals which is got through the COLS procedure is 69% (Street 2003) but as we mentioned earlier the COLS procedure yields significantly lower efficiencies as the COLS residuals contain both the random shocks and the inefficiency term.

Table 14 reports the output-specific economies of scale depending upon the stochastic and the CRE frontier settings. It reports ray economies of scale too. For each output, the output-specific economies of scale match the inverse of the individual cost-output elasticity and assess the response of total cost to an increase in the quantity of one output whilst everything else remains constant. Ray economies of scale are given by the inverse of the sum of the individual cost-output elasticities and evaluate the overall economies of scale for a proportionate expansion of all outputs (Baumol et al. 1988). By "output" we mean "major outputs" and specifically inpatient discharges – ID – and total outpatient visits – VISITS. The average length of hospitalisation stay – LOS – is a complementary output adjuster variable and it is not included for measuring the extent of economies of scale. Economies of scale exist if the inverse of the (individual or sum of the individual) cost-output elasticities is greater than scalar 1. Diseconomies of scale exist if that ratio is less than 1 and there are no scale effects if the ratio is equal to 1.

Table 14. Economies of scale

Model	ID	VISITS	Ray
SF full sample	3.07	2.17	1.27
CRE	1.96	2.44	1.09

As all values in Table 14 exceed 1, it seems there are output-specific and ray economies of scale. Take for instance the value of 3.07 (1.96). This indicates that a 10% increase in the total number of inpatient discharges whilst everything else remains still causes a 3.26% (5.10%) increase in total costs only under the stochastic (CRE) frontier setting. The percentage matches the inverse of 3.07 (1.96) or equally it matches the estimate for variable ID in Table 3 times scalar 10. Take now the value of 1.27 (1.09). This indicates that a 10% increase in both the total number of inpatient discharges and the total number of outpatient visits just provokes a 7.87% (9.20%) increase in total costs with the stochastic (CRE) frontier specification. The percentage matches the inverse of 1.27 (1.09) or equivalently it matches the sum of the estimate for variable ID plus the estimate for variable VISITS also in Table 3 times 10. As the values for ray economies of scale are reasonably close to scalar 1, we tested whether or not they could indeed be 1. Both Wald tests rejected that possibility at a 1% significance level.

All three tables from Table 15 up to Table 17 exhibit the degree of economies of scale which either is reported by or we computed from several articles. The countries of the under study hospitals appear in the first column. The second and the third columns give the number of hospitals and the specific years. Table 15 and Table 16 give the measure of economies of scale which sums up the impact on costs of expanding the number of in that order inpatient episodes and outpatient visits, holding everything else constant. Table 17 provides the measure of ray economies of scale which sums up the impact on costs of expanding all outputs proportionately. Evidently not all estimation strategies agree.

Table 15. Comparison of output-specific economies of scale: inpatients

Location	#	Years	ES	Author
Germany	1,565	2001 – 2003	0.98 – 1.01	Frohloff (2007)*
Germany	1,594	2000 – 2003	0.98 – 1.02	Herr (2008)*
Greece	91	1992	1.79	Aletras (1999)*
Spain	43	1988 – 1991	1.37	Wagstaff and Lopez (1996)
Switzerland	145	1989 – 1991	1.53	Crivelli et al. (2000)
Switzerland	156	1998 – 2001	1.17 – 1.27	Farsi and Filippini (2006)*
Switzerland	148	1998 – 2003	1.45 – 1.58	Farsi and Filippini (2008)
UK	226	1995 – 1996	1.03 – 1.10	Street (2003)*

Table 16. Comparison of output-specific economies of scale: outpatients

Location	#	Years	ES	Author
Greece	91	1992	3.69	Aletras (1999)*
Spain	43	1988 – 1991	0.32	Wagstaff and Lopez (1996)
Switzerland	148	1998 – 2003	6.68 – 6.87	Farsi and Filippini (2008)*
UK	226	1995 – 1996	5.50 – 6.33	Street (2003)*

Table 17. Comparison of ray economies of scale

Location	#	Years	ES	Author
Germany	1,565	2001 – 2003	0.98 – 1.01	Frohloff (2007)*
Germany	1,594	2000 – 2003	0.98 – 1.02	Herr (2008)*
Greece	91	1992	1.21	Aletras (1999)
Spain	43	1988 – 1991	1.15	Wagstaff and Lopez (1996)
Switzerland	145	1989 – 1991	1.53	Crivelli et al. (2000)
Switzerland	156	1998 – 2001	1.17 – 1.27	Farsi and Filippini (2006)*
Switzerland	148	1998 – 2003	1.19 – 1.29	Farsi and Filippini (2008)*
Vietnam	654	1996	0.20 – 1.38	Weaver and Deolalikar (2004)

Annotation for tables from Table 14 up to Table 17: The articles marked with a star (*) do not report any measure of economies of scale. We computed all values we would like to look at by making use of the coefficient estimates each article does report. For each output, the inverse of the individual cost-output elasticity gives the output-specific economies of scale. Also by inverting the sum of the individual cost-output elasticities one gets ray economies of scale.

By confronting Table 14 and Table 15, the largest impact on costs of raising the number of inpatient episodes should happen in the Portugal public hospitals as the highest value of output-specific economies of scale for inpatients is by far found in our work. It seems there are considerable returns to scale too but at a lower extent in Greece (Aletras 1999), Spain (Wagstaff and Lopez 1996) and Switzerland (Crivelli et al. 2000; Farsi and Filippini 2006; Farsi and Filippini 2008) hospitals. Returns to scale are less important in Germany (Frohloff 2007; Herr 2008) and in the UK (Street 2003) hospitals. Bringing Table 16 into the discussion, just Wagstaff and Lopez (1996) get diseconomies of scale for outpatients. All lasting articles report economies of scale, including ours. Although considerably above 1, the output-specific economies of scale for outpatients we have for Portugal public hospitals are lower than the values found for Greece (Aletras 1999), Switzerland (Farsi and Filippini 2008) or the UK (Street 2003) hospitals. In addition, the ray economies of scale we find for the Portugal public hospitals and which are in the third column of Table 17 are pretty much in line with the ray economies of scale found in Greece (Aletras 1999), Spain (Wagstaff and Lopez 1996), Switzerland (Crivelli et al. 2000; Farsi and Filippini 2006; Farsi and Filippini 2008) and in a number of hospitals in Vietnam (Weaver and Deolalikar 2004) too. It seems at first glance there are practically no scale effects in Germany hospitals (Frohloff 2007; Herr 2008) and there are ray diseconomies of scale in some hospitals in Vietnam (Weaver and Deolalikar 2004).

According to a recent sum up of the existing empirical work, and which Aletras (1999) mentions, economies of scale would be fully exploited in hospitals with 100–200 beds, whilst hospitals with 300–600 beds could display diseconomies of scale. Dranove (1998) brings up that economies of scale would be exhausted in hospitals with over 10 thousands inpatient discharges annually and these would correspond to 270 beds. Also Crivelli et al. (2000) refer that the minimum efficient scale would be 110 thousands annual hospitalisation days or equivalently 300 beds. Recall that the minimum efficient scale is the level of output which minimises the (long-run) average cost function.

We first find no minimum efficient scale and no diseconomies of scale ever as under both the stochastic and the CRE frontier alternative specifications the estimated average cost function is monotonically decreasing. And so the point of interest would then be to define what marginal or unimportant economies of scale are.

Figure 4 exhibits the estimated average cost function of producing output ID for the stochastic plus the CRE frontier settings when holding all remaining variables at their respective mean values. Figure 5 depicts the "marginal average cost" function which can roughly be defined as providing the amount of change which occurs in the average cost as a consequence of raising the number of inpatient discharges by one additional unit. Algebraically it is the first derivative of the average cost function. It is graphically not possible for sure to choose up to which point there are significant economies of scale and from which point on economies of scale become insignificant as graphs strongly depend upon the scale one previously chooses for them. Figure 5 is included for illustrational purposes only.

Figure 4. Average cost function (per inpatient discharge)

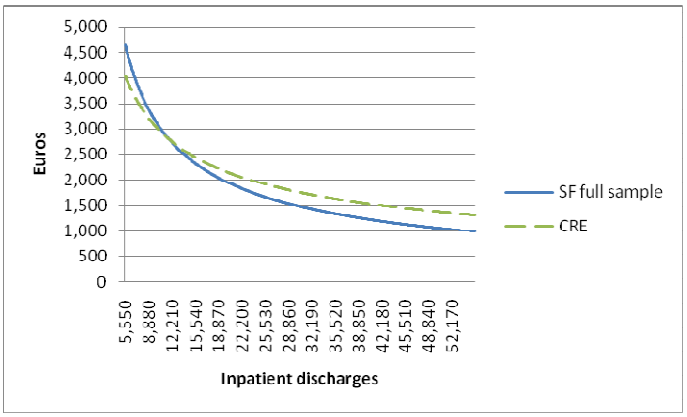


Figure 5. "Marginal average cost" function (per inpatient discharge)

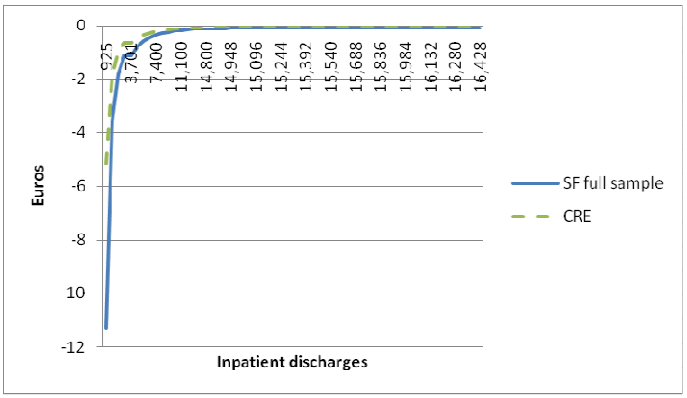


Table 18 gives the average cost of inpatient discharges when fixing all other variables at their respective mean values, and sums up the information on beds size admitting 37 ID = 1 BEDS holds. This is based upon the data in Table 2 and through computing the ratio of the mean of variable ID (which is 12,566.73) divided by the mean of BEDS (which is 338.21). Table 19 shows the percentage reduction in average cost of raising the number of inpatient discharges by 3,700 each time, or likewise by 100 beds.

Table 18. Average cost (per inpatient discharge) in Euros

ID					3,700		5,550	7,400	11,100	14,800	18,500	22,200	37,000
BEDS	100	150	200	300	400	500	600	1,000					
SF full sample					6,102		4,643	3,825	2,910	2,397	2,063	1,824	1,293
CRE					4,927		4,039	3,508	2,877	2,499	2,240	2,049	1,595

Table 19. Reduction in the estimated average cost of increasing the number of inpatient discharges by 3,700 (= 100 beds) each time (%)

ID	7,400	11,100	14,800	18,500	22,200	37,000
BEDS	200	300	400	500	600	1,000
SF full sample	37	24	18	14	12	7
CRE	29	18	13	10	9	5

One can focus on the percentage reduction in average cost of raising the number of inpatient discharges by 3,700 each time and define for instance that 10% is the cut-off point under which economies of scale are marginal or unimportant. Again it depends upon what one thinks marginal or unimportant economies of scale are.

Take a hospital with 3,700 inpatient discharges. The average cost decreases from 6,102 (4,927) down to 3,825 (3,508) Euros with the stochastic (CRE) frontier model. This is a 37% (29%) cut in average cost and so there are substantial economies of scale from doubling the number of beds from 100 up to 200. Consider now a hospital with 18,500 inpatient discharges. Through increasing the production by 3,700 inpatient discharges, the average cost diminishes from 2,063 (2,240) to 1,824 (2,049) Euros in the stochastic (CRE) frontier model. This is a 12% (9%) cut in average cost from increasing the number of beds from 500 up to 600. Applying the rule of the 10% cut-off point there are still substantial economies of scale arising from the stochastic frontier model whilst they seem to be fully (mostly) exploited under the CRE frontier framework.

Of course the results are sensitive to the somewhat arbitrary choice of the 10% cut-off point. But the results are above all sensitive both to the amount of absolute change one chooses to depend upon so as to assess the percentage reduction in average cost of each steady raise of output plus the assumption one defines for the rate of the number of inpatient discharges in terms of number of beds (we assumed 37 ID = 1 BEDS holds).

Dranove (1998) computes the percentage reduction in average cost of increasing output by 2,500 inpatient discharges each time and matches the number of inpatient discharges to the number of beds in a non-linear way. If one follows that exact procedure and still keeps the cut-off point at 10%, one finds that economies of scale are exploited in hospitals with just 17,500 (15,000) inpatient discharges. Therefore by admitting steady rises of output by 2,500 rather than 3,700 inpatient discharges each time, economies of scale seem to be fully exploited in hospitals with a notably lower number of inpatient discharges. If 37 ID = 1 BEDS holds, then economies of scale are fully exploited in hospitals with 475 (400) beds approximately. But through converting non-linearly the number of inpatient discharges into a number of beds as in Dranove (1998), it comes out that economies of scale are exploited in even smaller hospitals with 310–440 beds (most likely under 400 beds) and 310 beds in the stochastic and the CRE frontier models, respectively. In addition, either following the exact procedure of Dranove (1998) or not, the conclusion does not agree with the results in Aletras (1999), Dranove (1998) or Crivelli et al. (2000).

Perhaps there could be an alternative process that would not depend upon which amount of absolute change one chooses to rely on so as to compute the percentage reduction in average cost of each steady raise of output. One could for instance focus on the financial per unit reduction of raising the number of inpatient discharges by one additional unit plus define 25 (Euro-) cents as the cut-off point under which economies of scale are marginal or unimportant. Table 20 shows how the average cost of inpatient discharges falls in response to raising the number of inpatient discharges by one additional unit. All values express the "marginal average cost" function evaluated at the reported different quantities of variable ID. Applying the rule of the 25 cents then economies of scale are exploited in hospitals with 11,100 (7,400) inpatient discharges. Assuming as well 37 ID = 1 BEDS holds, the stochastic frontier model predicts economies of scale are exploited in hospitals with 300 beds whilst the CRE frontier model predicts they are exploited in hospitals with 200 beds. This conclusion would not only reasonably agree with Aletras (1999) but also with Dranove (1998) and Crivelli et al. (2000).

Table 20. "Marginal average cost" (per inpatient discharge) in Euros

ID	3,700	5,550	7,400	11,100	14,800	18,500	22,200	37,000
BEDS	100	150	200	300	400	500	600	1,000
SF full sample	-1.11	-0.56	-0.35	-0.18	-0.11	-0.08	-0.06	-0.02
CRE	-0.65	-0.36	-0.23	-0.13	-0.08	-0.06	-0.05	-0.02

As for a general conclusion on this matter of identifying from which level of production economies of scale are exploited, when there is no minimum efficient scale because the average cost is a monotonic decreasing function, the results depend upon the criteria one relies on in order to determine whether there are substantial economies of scale or not, plus the rapport which is assumed for the amount of produced inpatient care and the size of the hospital in beds.

8. THE UNIT COST OF TREATMENT AND LIKELY SAVINGS/EXTRA OUTPUT

Finding out the estimated frontier unit cost of providing treatment in each hospital is a rather appealing task as the computed outcome may be applied in numerous settings. As quite a few economic theoretical models use an abstract unit cost of treating a patient at the hospital in their conceptual framework, computing the estimated frontier unit cost of treatment unveils one model parameter, which by itself is a noteworthy contribution. And putting adequate parameter estimates into those models matters of course as they enable policy guidance.

But finding the estimated frontier cost of providing treatment at the hospital is not as direct as one may firstly think as hospitals provide multiple outputs of treatment. Consider for instance the following trail a patient may go through so as to receive adequate treatment at the hospital. The patient may first be given treatment through outpatient attending, go after that to surgery and turn into an inpatient, leave the hospital and go home but then revisit the hospital for follow-up attending. Trails of that sort are echoed in our data set both through the number of inpatient discharges and the number of outpatient visits. In this not so uncommon particular case, as inpatient discharges and outpatient visits reflect treatment, one shall define what one wants treatment to be so as to find the best rough computation for the theoretical models' abstract unit cost of treating a patient at the hospital. Also treatment can happen to be even more complex if for instance one adds a first patient attending in the emergency room to the trail the patient goes through at the hospital or if in addition one considers laboratory and radiology services which have a say in the patient's recovery too.

Treating inpatients is unquestionably the hospitals' key activity among the multiple outputs of treatment and non-treatment hospitals offer. Other organisations may as well handle outpatient attending or adequately respond to emergency occurrences. Several organisations can substitute hospitals in laboratory and radiology supply as it happens every day. Also organisations such as universities for instance should be able to suitably restore all hospitals' teaching and research if such activities would come to an end in the hospitals. Roughly all hospital outputs can be produced by alternative organisations except one. Only hospitals can treat inpatients. And as one must define treatment, and because no other organisation has the ability to treat inpatients, the cost of inpatient care should then be taken as matching the cost of treating a patient at the hospital.

Furthermore, opposing the estimated frontier cost to the actual observed expense is for policy reasons of interest too because it makes it possible not only to evaluate how far each organisation lies from the reference technology frontier but also how much savings the organisation should aim to arrive at whilst producing the actual amount of output, or to put this the other way around, how much extra output should the organisation aim to arrive at whilst spending the actual amount of expense.

All these aspects should be of large importance to the national public health authorities which are responsible altogether for financing public hospitals plus to assess their performance. The Portugal health authority should for instance very much welcome estimated frontier costs, particularly when there is a broad national concord that rather than requiring more money, the Portuguese NHS should instead learn and spend its budget more properly.

8.1. The Cova da Beira hospital centre

We next compute the frontier unit cost of one specific Portugal public hospital and find out, on one hand how much savings it could arrive at whilst producing its actual amount of output, and in the other hand how much extra output it could produce whilst spending as much as it has been spending. For what concerns the aims of this work, it is worthless to extend the exercise on the full sample.

We focus on the Cova da Beira hospital centre (CBHC) as it is of particular interest. First, it is one "medium size" hospital and "medium size" hospitals are altogether 60% of the sample. Second, as we mentioned earlier, it is the hospital which does the highest jump, from position 48 in the efficiency ranking of the stochastic frontier model up to position 12 in the efficiency

ranking of the CRE frontier model. It matters to check what happens to this "medium size" hospital as it does the highest jump from one efficiency ranking to the other because, up to a certain extent, this will say what impact in the field one should anticipate of either adding or substracting heteroscedasticity from the model. Besides it is of further interest as it serves the students of medicine of the University of Beira Interior, which is the university where we are.

Throughout the years 2002 – 2006 CBHC spent in average 49 million Euros annually to produce 12,271 inpatient discharges plus 95,316 outpatient visits each year. Covilha, which is the town where CBHC's most important buildings are, has a little more than 52.5 thousands inhabitants. Inpatients remained hospitalised for just about 8 days and the occupancy rate was 76%. The occupancy rate is the ratio of the inpatient stay (which was computed as the number of inpatient discharges – ID – times the average length of stay – LOS – in days) to the inpatient capacity (which is the number of available beds – BEDS – times the 365 days of the year). CBHC counted in average 349 available beds. Table 21 presents the summary statistics of CBHC.

Table 21. Summary statistics of the Cova da Beira hospital centre

Variables	Number of observations	Mean	SD	Min	Max
COST	5	49,087,030	8,719,876	36,765,020	57,443,260
ID	5	12,270.6	82.98072	12,170	12,358
LOS	5	7.94	0.15	7.70	8.10
VISITS	5	95,316.2	19,124.5	69,487	116,375
PL	5	24,966.13	1,663.3	23,390.17	27,155.76
PK	5	1.06	0.44	1	1.11
NPB	5	1.19	0.09	1.09	1.29
TEACHING	5	1	0	1	1
BEDS	5	349.2	11.34	336	360
POP	5	0.05	0	0.05	0.05

When confronted with the means of the full sample in Table 2, CBHC spent in average less 18% than the mean expense of all 63 hospitals in the sample whilst produced less 2% of inpatient discharges and less 14% of outpatient visits. The town where its most important buildings are has in average 85% less inhabitants than the mean of inhabitants per public hospital. CBHC's inpatients were hospitalised roughly through one day more than the inpatients of all 63 Portugal public hospitals taken altogether and its occupancy rate was 1% higher. It is slightly bigger in terms of number of beds than the mean of the number of beds per public hospital (103% or 11 beds).

We first compute the frontier unit cost of treatment. The computation is carried out by taking the exponential of equation (6), then by replacing each variable in the estimated frontier cost function by the corresponding CBHC observed mean for the period 2002 – 2006, with the homoscedastic stochastic frontier estimates and next the estimates of the heteroscedastic CRE frontier model, as shown in the first and in the third columns of Table 3, respectively.

Under homoscedasticity, CBHC's frontier unit cost (per inpatient discharged) is 2,759 Euros. Divide this by its average length of hospitalisation stay and the daily frontier unit cost is 348 Euros. And by adding heteroscedasticity, CBHC's frontier unit cost becomes 2,836 Euros and its daily equivalent is 357 Euros. These contrast with the observed costs. The actual expense (per inpatient discharged) is in average 4,000 Euros each year and the actual daily expense amounts to 504 Euros.

From Table 4 and Table 7, if CBHC shifts downward onto the computed cost frontier, it saves up to 31% or up to 29% of its actual expense whilst it produces its actual amount of output, depending upon one chooses the stochastic or the CRE cost frontier model. Savings would then be up to 15 or 14 million Euros in that order.

We next do the exercise the other way around and get the equivalent of the excess cost in terms of output. In other words, we calculate how much likely extra output CBHC produces if it is onto its estimated cost frontier and spends its actual annual expense. Suppose the production

of outpatient visits is fixed at its actual level, which means all excess cost is allocated to the production of extra inpatient care. Also all other variables, except the variable ID, are held at their respective CBHC mean. CBHC then produces more 26,045 or more 11,807 inpatient discharges, depending upon the model. This is a production of more 212% or more 96% inpatient discharges than throughout the period 2002 – 2006. Divide the extra production by 365 and the excess cost is equal to more 71 or 32 inpatient discharges per day.

Additionally, it might happen that the average length of hospitalisation stay of CBHC decreases. First, because a number of patients with less severe illnesses, which firstly would not be hospitalised although their illnesses could justify the hospitalisation, may next be hospitalised and become part of the extra production of inpatient care. Second, many other aspects are known to affect the length of hospitalisation stay. Among them are the organisation of hospital care such as the accessibility of operating stages and the time engaged for diagnostic explorations, and the clinical practice approach such as the surgical procedure or the anaesthesia chosen (Morgan and Beech 1990; Simpson et al. 2005). As for an example, a surgery which resorts to laparoscopy implies in general shorter hospitalisations stays as it is minimally invasive. By admitting CBHC is onto its estimated cost frontier, one can anticipate as a consequence it relies on an entirely more efficient environment which pressures its average length of stay downward.

Suppose CBHC's average length of stay decreases and that it matches the mean of the full sample. This is just a minor reduction of 0.57 days but in that case CBHC produces more 30,378.5 or 15,170.9 annual inpatient discharges than throughout the five years, under the homoscedastic or the heteroscedastic framework. This is a production of more 248% or more 124%. The daily production is raised by 83 or 42 inpatient discharges.

But in the field, none of those numbers should be attainable because the occupancy rate goes off to rather high rates. The occupancy rate is 239%–266% with the stochastic frontier model and 150%–171% with the CRE frontier model. Table 22 sums up this information.

Table 22. Cova da Beira hospital centre: extra production of inpatient discharges whilst spending the actual amount of expense and keeping the production of outpatient visits at its actual level

	LOS	Optimal ID	Δ ID	Δ ID %	Δ ID / day	OR %
SF full sample	7.94	38,315.6	26,045.0	212.3	71.4	238.7
	7.37	42,649.1	30,378.5	247.6	83.2	265.7
CRE	7.94	24,077.5	11,806.9	96.2	32.3	150.0
	7.37	27,441.5	15,170.9	123.6	41.6	170.9

Accordingly, we assume the occupancy rate is at 90% and compute the equivalent of the remaining excess cost in terms of outpatient visits. In order to do this, we take the exponential of equation (6) and change all variables in the estimated frontier function by the relevant CBHC mean, expect for the variables ID, LOS and VISITS. The variable ID is replaced by the number of inpatient discharges when the occupancy rate is at 90%. The variable LOS is either 7.94 or 7.37, which are respectively CBHC's average length of hospitalisation stay throughout the five years and the mean of the full sample.

Table 23 presents the extra production of inpatient discharges as the result of setting the occupancy rate at 90%. Raising the occupancy rate from 76% to 90% is equal to raising the number of inpatient discharges up to 2,177 or up to 3,294, depending upon LOS = 7.94 or LOS = 7.37. The rise corresponds respectively to 18% or 27% more inpatient discharges. This is 6 or 9 inpatient discharges more per day. Table 24 gives the extra production of outpatient visits the excess cost is equivalent to, at the same time as the variable ID is at the levels of production which agree with the occupancy rate of 90%. CBHC produces then annually more 94,805–99,252 or 84,848–97,942 visits, under the homoscedastic or the heteroscedastic models, which roughly signifies doubling the average annual (or daily) production of outpatient visits of the five years period.

Table 23. Cova da Beira hospital centre: extra production of inpatient discharges whilst setting the occupancy rate at 90%

LOS	Total ID	Δ ID	Δ ID %	Δ ID / day
7.94	14,447.4	2,176.8	17.7	6.0
7.37	15,564.7	3,294.1	26.8	9.0

Table 24. Cova da Beira hospital centre: extra production of outpatient visits whilst spending the actual amount of expense and setting the occupancy rate at 90%

	LOS	Optimal VISITS	Δ VISITS	Δ VISITS %	Δ VISITS / day
SF full sample	7.94	190,121.2	94,805.0	99.5	259.7
	7.37	194,567.8	99,251.6	104.1	271.9
CRE	7.94	180,163.7	84,847.5	89.0	232.5
	7.37	193,258.0	97,941.8	102.8	268.3

As a sum up one can be confident CBHC's frontier unit cost of treatment (per inpatient discharged) is 2,800 Euros. Shifting CBHC downward onto its efficient cost frontier enables savings of at least 14 million Euros annually whilst still producing its actual amount of output. Likewise shifting CBHC downward onto its efficient cost frontier and still spend its actual amount of expense enable to raise its production of inpatient care by at least 18% and to double its production of outpatient visits. This is quite a lot and it confirms the feasible enormous savings or substantial rise of production we saw earlier in Section 7.

Finally, the impact in the field of either adding or substracting heteroscedasticity from the model can be considerable. The likely savings of this "medium size" hospital differ in one million Euros from one specification to the other and the occupancy rate at 90% is one thousand inpatient discharges more annually under heteroscedasticity, which is 50% more.

9. CONCLUSIONS

We first identify a stochastic cost frontier model both to compute the mainland Portugal public hospitals' best practice reference technology as well as to measure cost efficiency across hospitals. Making use of panel data on 63 hospitals for the period 2002 – 2006, we estimate the reference technology frontier which we believe to be the same for the full sample, even for hospitals which differ quite considerably in size. It seems that smaller and larger hospitals chase the same cost function, even the bigger university hospitals for which one anticipates they treat a comparatively sterner patient mix, also with relatively short but expensive hospitalisations.

The stochastic cost frontier model takes for granted that the variance is equal among the hospitals. But a likelihood-ratio test points towards heteroscedasticity and it is indeed possible there is heteroscedasticity in our data set as the hospitals we observe differ very much in size. Inspired by the cross-section data context COLS technique, we introduce a CRE cost frontier technique of our own through generalised least squares and by assuming heteroscedasticity across hospitals. The technique may be helpful because it requires no particular assumption about the shape or the location of the time-invariant inefficiency component distribution. Also although the COLS technique provides lower efficiency estimates since its residuals include both the random shocks and inefficiency, this does not happen with the CRE frontier technique as its "lowest negative residual" which is subtracted from the residuals is actually the lowest negative residual mean, and with normally distributed random shocks with zero mean, the same as in the stochastic frontier setting, the bias caused by the random shocks on the efficiency scores vanishes as the number of observations for each group increases through time.

Efficiency rankings as well as all other outcomes which stem from one and another model are quite dissimilar. Even if we end up favouring cost frontier estimation with heteroscedasticity, either the standard homoscedastic stochastic frontier estimates are biased when the variance evolves across hospitals or including heteroscedasticity whilst there is none

in the data puzzles the results enormously. This finding is actually in line with a few authors' similar findings too.

According to a recent sum up, economies of scale would be fully exploited in hospitals with 100–200 beds, whilst hospitals with 300–600 beds could display diseconomies of scale. In contrast, alternative works refer the number of approximately 300 beds so that economies of scale could be fully exploited. We first obtain no minimum efficient scale because the estimated average cost is a monotonic decreasing function. And second, whenever this happens, identifying from which level of output economies of scale are fully exploited depends upon which criteria one relies on so as to decide whether there still are significant economies of scale or not, plus the rapport which is assumed for the amount of production and the size of the hospital.

Among the articles we could gather, Portugal and Spain are the only OECD countries whose hospitals do not reach a mean efficiency estimate of 80%. No further conclusion can however be drawn from this. We cannot tell whether Portugal or Spain hospitals hunt the same reference technology as the hospitals of the other OECD countries. And it is unlikely that they do so.

As for policy implications, first the efficiency rankings matter of course. Second, the overall potential cost reduction for the Portuguese NHS with no cuts whatsoever in the output produced would be around 872–1,290 million Euros annually, depending upon the model one chooses, if all hospitals in the sample shifted downward onto their efficient cost frontier. This is about 11%–17% of the Portuguese NHS year 2006 initial endowment. Put this the other way around, and by keeping the hospitals' costs at their actual levels, shifting all 63 hospitals downward onto their efficient cost frontier would enable 24%–35% more inpatient discharges each year, which is 184–273 thousands inpatients more annually. Third, taking the cost of inpatient care as the best proxy for the cost of treatment at the hospital, we find the frontier unit cost of treatment (per inpatient discharged) of the Cova da Beira hospital centre is 2,800 Euros. This is a "medium size" hospital which, up to a certain extent, can be seen as representative. Also shifting CBHC downward onto its efficient cost frontier enables savings of at least 14 million Euros annually whilst still producing its actual amount of output. Likewise shifting CBHC downward onto its efficient cost frontier and still spend its actual amount of expense enable to raise its production of inpatient care by at least 18% and to double its production of outpatient visits.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Anabela Boavista of the Portugal (former) Institute for Quality in Health at Oporto, Portugal for supplying the accurate information on the evolution of each hospital's request for a quality accreditation, either from the United Kingdom Health Quality Service (former King's Fund) or from the United States Joint Commission International. This research was financed by the Portugal Foundation for Science and Technology through research grant SFRH / BD / 8537 / 2002, which is gratefully acknowledged. The author is of course solely accountable for the views expressed in this work which do not necessarily reproduce the position of any institution.

REFERENCES

ALETRAS VH. 1999. A comparison of hospital scale effects in short-run and long-run cost functions. *Health Economics* **8**: 521 – 530.

ANTUNES MJ. 2001. *A Doença da Saúde* (in Portuguese). Quetzal Editores: Lisbon.

APES. 1997. *Reflexão sobre o Financiamento da Saúde em Portugal* (in Portuguese). Associação Portuguesa de Economia da Saúde: Lisbon.

BARROS PP, MARTINEZ-GIRALT X. 2003. Preventive health care and payment systems. *Topics in Economic Analysis & Policy* **3**: Issue 1, Article 10.

BATTESE GE, COELLI T. 1988. Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics* **38**: 387 – 399.

BATTESE GE, COELLI T. 1992. Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. *Journal of Productivity Analysis* **3**: 153 – 169.

BATTESE GE, COELLI T. 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* **20**: 325 – 332.

BAUMOL WJ, PANZAR JC, WILLIG RD. 1988. Contestable markets and the theory of industry structure. San Diego: Harcourt Brace Jovanovich.

BREUSCH TS, PAGAN AR. 1980. The Lagrange multiplier test and its applications to model specification in econometrics. *Review of Economic Studies* **47**: 239 – 254.

CAREY K. 1997. A panel data design for estimation of hospital cost functions. *The Review of Economics and Statistics* **79**(2): 443 – 453.

CARREIRA C. 1999. Economias de escala e de gama nos hospitais públicos portugueses: uma aplicação da função de custo variável translog (in Portuguese). *Estudos do Grupo de Estudos Monetários e Financeiros* **1**, Department of Economics, University of Coimbra, Coimbra. Available at http://gemf.fe.uc.pt/workingpapers/pdf/1999/gemf99_01.pdf [accessed February 2009].

CORNWELL C, SCHMIDT P, SICKLES RC. 1990. Production frontiers with cross-sectional and time-series variation in efficiency levels. *Journal of Econometrics* **46**: 185 – 200.

COWING TG, HOLTMAN AG. 1983. Multiproduct short-run hospital cost functions: empirical evidence and policy implications from cross-section data. *Southern Economic Journal* **49**: 637 – 653.

CRIVELLI L, FILIPPINI M, LUNATI D. 2001. Dimensione ottima degli ospedali in uno Stato federale (in Italian). *Economia Pubblica* **5** : 97 – 119.

DEILY ME, MCKAY NL. 2006. Cost inefficiency and mortality rates in Florida hospitals. *Health Economics* **15**: 419 – 431.

DOR A. 1994. Non-minimum cost functions and the stochastic frontier: on applications to health care providers. *Journal of Health Economics* **13**: 329 – 334.

DOR A, FARLEY D. 1996. Payment source and the cost of hospital care: evidence from a multiproduct cost function with multiple players. *Journal of Health Economics* **15**: 1 – 22.

FARSI M, FILIPPINI M, KUENZLE M. 2005. Unobserved heterogeneity in stochastic cost frontier models: an application to Swiss nursing homes. *Applied Economics* **37**: 2127 – 2141.

FARSI M, FILIPPINI M. 2006. An analysis of efficiency and productivity in Swiss hospitals. *Schweizerische Zeitschrift Volkswirtschaft und Statistik* **142**(1): 1 – 37.

FARSI M, FILIPPINI M. 2008. Effects of ownership, subsidization and teaching activities on hospital costs in Switzerland. *Health Economics* **17**: 335 – 350.

FOLLAND ST, HOFLER RA. 2001. How reliable are hospital efficiency estimates? Exploiting the dual to homothetic production. *Health Economics* **10**: 683 – 698.

FOURNIER GM, MITCHELL JM. 1992. Hospital costs and competition for services: a multiproduct analysis. *The Review of Economics and Statistics* **74**(4): 627 – 634.

FROHLOFF A. 2007. Cost and technical efficiency of German hospitals: a stochastic frontier analysis. *Ruhr Economic Papers* **2**, Graduate School in Economics and Ruhr-Universität Bochum, Germany.

GARCIA-PRIETO C. 2001. Ineficiencia tecnica y asignativa en los hospitales de Castilla y Leon (in Spanish). *Actas del 7º Congreso de Economia Regional de Castilla y Leon*, Spain: 384.

GARCIA-PRIETO C. 2003. La ineficiencia en costes de los hospitales del INSALUD: un estudio detallado (in Spanish). *Estudios de Economía Aplicada* **21**(3): 467 – 484.

GAYNOR M, ANDERSON G. 1995. Uncertain demand, the structure of hospital costs, and the cost of empty hospital beds. *Journal of Health Economics* **14**: 291 – 317.

GREENE WH. 1993. The econometric approach to efficiency analysis, In *The Measurement of Productive Efficiency*, Fried HO, Lovell CAK, Schmidt S (eds). Oxford University Press: Oxford; 68 – 119.

GREENE WH. 2004. Distinguishing between heterogeneity and inefficiency: stochastic frontier analysis of the World Health Organization's panel data on national health care systems. *Health Economics* **13**: 959 – 980.

GREENE WH. 2005. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics* **126**: 269 – 303.

GRIFFIN RC, MONTGOMERY JM, RISTER ME. 1987. Selecting functional form in production function analysis. *Western Journal of Agricultural Economics* **12**(2): 216 – 227.

GUERMAT C, HADRI K. 1999. Heteroscedasticity in stochastic frontier models: a Monte Carlo analysis. *Discussion Papers* **14**, School of Business and Economics, University of Exeter.

HADLEY J. ZUCKERMAN S. 1994. The role of efficiency measurement in hospital rate setting. *Journal of Health Economics* **13**: 335 – 340.

HADRI K, GUERMAT C, WHITTAKER J. 2003. Estimating farm efficiency in the presence of double heteroscedasticity using panel data. *Journal of Applied Economics* **6**(2): 255 – 268.

HAUSMAN J. 1978. Specification tests in econometrics. *Econometrica* **46**: 1251 – 1271.

HERR A. 2008. Cost and technical efficiency of German hospitals: does ownership matter? *Health Economics* **17**: 1057 – 1071.

HORRACE WC, SCHMIDT P. 1996. Confidence statements for efficiency estimates from stochastic frontier models. *Journal of Productivity Analysis* **7**: 257 – 282.

JACOBS R, SMITH P, STREET A. 2006. *Measuring Efficiency in Health Care*. Cambridge University Press: Cambridge.

JENSEN U. 2000. Is it efficient to analyse efficiency rankings? *Empirical Economics* **25**: 189 – 208.

KUMBHAKAR SC. 1990. Production frontiers, panel data, and time-varying technical efficiency. *Journal of Econometrics* **46**: 201 – 211.

KUMBHAKAR SC, LOVELL CAK. 2000. *Stochastic Frontier Analysis*. Cambridge University Press: Cambridge.

LEE YH, SCHMIDT P. 1993. A production function model with flexible temporal variation in technical efficiency. In *The Measurement of Productive Efficiency*, Fried HO, Lovell CAK, Schmidt SS (eds). Oxford University Press: New York; 237 – 255.

LINNA M. 1998. Measuring hospital cost efficiency with panel data models. *Health Economics* **7**: 415 – 427.

LINNA M, HAKKINEN U. 1997. Determinants of cost efficiency of Finnish hospitals: a comparison of DEA and SFA. *Working Paper*, National Research and Development Centre for Welfare and Health, Finland.

LINNA M, HAKKINEN U, LINNAKKO E. 1998. An econometric study of costs of teaching and research in Finnish hospitals. *Health economics* **7**: 291 – 305.

MATHIYAZHGAN MK. 2006. Cost efficiency of public and private hospitals: evidence from Karnataka State in India. *Working Paper* **8**, Institute of South Asian Studies, National University of Singapore, Singapore.

MCKINNISH TG. 2000. Model sensitivity in panel data analysis: some caveats about the interpretation of fixed effects and differences estimators. *Unpublished manuscript*, Department of Economics, University of Colorado, Boulder. Available at http://spot.colorado.edu/-mckinnis/fe053100.pdf [accessed February 2009].

MORGAN M, BEECH R. 1990. Variations in length of stay and rates of day case surgery: Implications for the efficiency of surgical management. *Journal of Epidemiology and Community Health* **44**: 90 – 105.

NEWHOUSE J. 1994. Frontier estimation: how useful a tool for health economics? *Journal of Health Economics* **13**: 317 – 322.

PITT MM, Lee LF. 1981. The measurement and sources of technical inefficiency in Indonesian weaving industry. *Journal of Development Economics* **9**: 43 – 64.

POLACHEK S, Yoon B. 1996. Panel estimates of a two-tiered earnings frontier. *Journal of Applied Econometrics* **11**: 169 – 178.

ROSKO MD. 2001. Cost efficiency of US hospitals: a stochastic frontier approach. *Health Economics* **10**: 539 – 551.

SCHMIDT P. 1985. Frontier production functions. *Econometric Reviews* **4**: 289 – 328.

SCHMIDT P, Sickles RC. 1984. Production frontiers and panel data. *Journal of Business and Economic Studies* **2**: 299 – 326.

SCUFFHAM PA, DEVLIN NJ, JAFORULLAH M. 1996. The structure of costs and production in New Zealand public hospitals: an application on the transcendental logarithmic variable cost function. *Applied Economics* **28**(1): 75 – 85.

SIMPSON S, PACKER C, STEVENS A, RAFTERY J. 2005. Predicting the impact of new health technologies on average length of stay: Development of a prediction framework. *International Journal of Technology Assessment in Health Care* **21**(4): 487 – 491.

STREET A. 2003. How much confidence should we place in efficiency estimates? *Health Economics* **12**: 895 – 907.

VITA MG. 1990. Exploring hospital production relationships with flexible functional forms. *Journal of Health Economics* **9**: 1 – 21.

WAGSTAFF A, LOPEZ G. 1995. Hospital costs in Catalonia: a stochastic frontier analysis. *Applied Economic Letters* **3**: 471 – 474.

WEAVER M, DEOLALIKAR A. 2004. Economies of scale and scope in Vietnamese hospitals. *Social Science & Medicine* **59**: 199 – 208.

ZUCKERMAN S, HADLEY J, IEZZONI L. 1994. Measuring hospital efficiency with frontier cost functions. *Journal of Health Economics* **13**: 255 – 280.

A MODEL FOR INTEGRATING POSITIONAL VOTING SYSTEMS

BONIFACIO LLAMAZARES RODRÍGUEZ
E-mail: boni@eco.uva.es
MARÍA TERESA PEÑA GARCÍA
E-mail: maitepe@eco.uva.es
Departamento de Economía Aplicada
Universidad de Valladolid

Abstract

In many voting systems, voters' preferences on a set of candidates are represented by linear orderings. A well-known procedure to aggregate these ordinal preferences and to obtain the winning candidates consists in assigning fixed scores to the different ranks and selecting the candidates with the highest score (scoring rules). This procedure is a particular case of positional voting systems, where the score achieved by each candidate depends only on the number of first, second,..., *m*-th ranks obtained by her/him. In this paper we analyze some positional voting systems proposed in the literature and show that these systems can be integrated in a model based on cumulative standings functions. The proposed model allows us to define new positional voting systems and to obtain in a straightforward manner some results appeared in the literature.

Key-words: Positional voting systems, cumulative standings functions.

1 INTRODUCTION

In many voting systems, voters' preferences on a set of candidates are represented by linear orderings. A well-known procedure to aggregate the ordinal preferences of n voters and to obtain the winning candidates consists in assigning fixed scores to the different ranks and selecting the candidates with the highest score (*scoring rules*). Thus, given a set of candidates $\{A_1,\ldots,A_m\}$, the score obtained by candidate A_i is

$$Z(A_i)=\sum_{k=1}^{m} s_k v_{ik},$$

where v_{ik} is the number of k-th place ranks that candidate A_i occupies and $(s_1,\ldots,s_m)$, with $s_1 \geq s_2 \geq \cdots \geq s_m$ and $s_1 > s_m$, is the scoring vector used. Plurality rule, where $s_1 = 1$ and $s_k = 0$ for all $k \in \{2,\ldots,m\}$, and Borda rule, where $s_k = m-k$ for all $k \in \{1,\ldots,m\}$, are the best-known examples of scoring rules.

Besides scoring rules, there exist other methods (which are less known than the previous ones) where the score obtained by each candidate depends only on the number of first, second, . . . , *m*-th ranks obtained by her/him. We will call all these procedures positional voting systems. In this paper, we set up a model that allows us to unify the study of the different positional voting systems. This model facilitates the analysis and the characterizations of these voting systems.

Throughout the paper, we will use the following notation: Given two vectors $\mathbf{x},\mathbf{y} \in \Re^k$, $\mathbf{x} \geq \mathbf{y}$ denotes $x_i \geq y_i$ for all $i \in \{1,\ldots,k\}$; $\mathbf{x} > \mathbf{y}$ means that $\mathbf{x} \geq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$; given $b \in \Re$, $\mathbf{b}$ stands for the vector $(b,\ldots,b)$.

The paper is organized as follows. In Section 2 we propose a model where the candidates score is obtained, by means of certain functions, from the number of cumulative standings that

they have achieved. In Section 3 we express through this model some of the positional voting systems proposed in the literature and define new positional voting systems. Finally, Section 4 is devoted to conclusions and future lines of research.

2 THE MODEL

In positional voting systems the winning candidates are selected utilizing the number of first, second, . . . , *m*-th ranks obtained by each candidate. In this paper we propose using for each candidate A_i the cumulative standings

$$V_{ik} = \sum_{j=1}^{k} v_{ij}, \quad k = 1,\ldots,m,$$

instead the number of ranks v_{ik}, $k = 1,\ldots,m$ (it is worth noting that cumulative standings have been used by Green, Doyle and Cook (1996) and Stein, Mizzi and Pfaffenberger (1994), among others). The score of each candidate in every positional voting system can be obtained using a monotonic function *F* of that cumulative standings. Since $V_{im} = n$ for all $i \in \{1,\ldots,m\}$, this function, called *cumulative standings function*, will be defined without taking into account the last cumulative standings, i.e., $F : \mathbf{X} \longrightarrow \Re$, where

$$\mathbf{X} = \left\{(x_1,\ldots,x_{m-1}) \in \mathbf{N}^{m-1} \,\middle|\, 0 \le x_1 \le \cdots \le x_{m-1} \le n\right\}.$$

In this way, given $\mathbf{V}$ an $m \times (m-1)$ matrix whose rows $\mathbf{V}_i$ represent the candidate A_i's cumulative standings, the score obtained by candidate A_i is $F(\mathbf{V}_i)$ and the candidates can be ordered of the following form:

$$A_i \succ A_j \quad \Leftrightarrow \quad F(\mathbf{V}_i) > F(\mathbf{V}_j).$$

The monotonicity property imposed on *F* can be considered a natural condition: if $\mathbf{V}_i \ge \mathbf{V}_j$, it seems logical that $A_i \succ A_j$ or $A_i \sim A_j$, i.e.,

$$\mathbf{V}_i \ge \mathbf{V}_j \;\Rightarrow\; F(\mathbf{V}_i) \ge F(\mathbf{V}_j).$$

It is necessary to emphasize that different cumulative standings functions can yield the same positional voting systems. In this case we will say that the cumulative standings functions are *equivalent*.

3 POSITIONAL VOTING SYSTEMS AND CUMULATIVE STANDINGS FUNCTIONS

In this section we analyze some positional voting systems proposed in the literature and show that these systems can be included in a model based on cumulative standings functions. On the other hand, we use the cumulative standings functions to define new positional voting systems.

3.1 Scoring rules

In the scoring rules, as we comment in the introduction, the score obtained by A_i is

$$Z(A_i) = \sum_{k=1}^{m} s_k v_{ik},$$

where v_{ik} is the number of *k*-th place ranks that candidate A_i occupies and $(s_1,\ldots,s_m)$, with $s_1 \ge s_2 \ge \cdots \ge s_m$ and $s_1 > s_m$, is the scoring vector used.

If we consider cumulative standings, the score obtained by candidate A_i is

$$Z(A_i) = \sum_{k=1}^{m} S_k V_{ik},$$

where $S_k = s_k - s_{k+1}$ for all $k \in \{1,\ldots,m-1\}$ and $S_m = s_m$. In this case, the condition $s_1 \geq s_2 \geq \cdots \geq s_m$ is equivalent to $S_k \geq 0$ for all $k \in \{1,\ldots,m-1\}$, and the condition $s_1 > s_m$ is equivalent to $S_k > 0$ for some $k \in \{1,\ldots,m-1\}$. Therefore, the score achieved by each candidate can be obtained through the following affine cumulative standings function:

$$F(x_1,\ldots,x_{m-1}) = S_1 x_1 + \cdots + S_{m-1} x_{m-1} + S_m n.$$

Two scoring rules defined by the vectors **s** and **s'** are the same positional voting system if and only if there exist $a, b \in \Re$, $a > 0$, such that $\mathbf{s}' = a \cdot \mathbf{s} + \mathbf{b}$. This relation can be written using the cumulative standings as $S'_k = a S_k$ for all $k \in \{1,\ldots,m-1\}$ and $S'_m = a S_m + b$. Taking into account the previous one, it is sufficient to consider the vectors with L_1-norm equals to the unity and $S_m = 0$. Therefore, the different scoring rules can be defined using the following set of vectors:

$$\mathbf{\Lambda} = \left\{ \mathbf{S} \in \Re^{m-1} \mid \mathbf{S} > \mathbf{0}, \sum_{k=1}^{m-1} S_k = 1 \right\}.$$

Next, we show some of the best-known scoring rules:

– *Plurality*: The candidates with the largest number of first ranks win. In the set of scoring vectors $\mathbf{\Lambda}$, plurality rule is obtained when $S_1 = 1$ and $S_k = 0$ for all $k \in \{2,\ldots,m-1\}$. The cumulative standings function is

$$F(x_1,\ldots,x_{m-1}) = x_1.$$

– *Antiplurality*: The candidates with the minor number of votes in the last position win. In the set of scoring vectors $\mathbf{\Lambda}$, antiplurality rule is obtained when $S_k = 0$ for all $k \in \{1,\ldots,m-2\}$ and $S_{m-1} = 1$. The cumulative standings function is

$$F(x_1,\ldots,x_{m-1}) = x_{m-1}.$$

– *Borda rule*: A candidate receives no points for being ranked last, one point for being ranked next to last, up to $m-1$ points for being ranked first. In the set of scoring vectors $\mathbf{\Lambda}$, Borda rule is obtained when $S_k = 1/(m-1)$ for all $k \in \{1,\ldots,m-1\}$. The cumulative standings function is

$$F(x_1,\ldots,x_{m-1}) = \frac{x_1 + \cdots + x_{m-1}}{m-1},$$

i.e., the score obtained by each candidate is the arithmetic mean of her/his cumulative standings.

– *k-approval voting*: The candidates with the largest number of votes in the *k* first positions win. In the set of scoring vectors $\mathbf{\Lambda}$, *k*-approval voting is obtained when $S_k = 1$ and $S_i = 0$ for all $i \neq k$. The cumulative standings function is

$$F(x_1,\ldots,x_{m-1}) = x_k.$$

3.2 Lexicographic rule

In the lexicographic voting rule, the candidate with the highest number of first ranks wins, if different candidates have the same number of first positions, the candidate with the highest number of second ranks is chosen and so on. The winning candidates with this model can be obtained using the following cumulative standings function

$$F(x_1,\ldots,x_{m-1}) = x_1 (n+1)^{m-2} + \cdots + x_{m-2}(n+1) + x_{m-1}.$$

3.3 Brams and Kilgour's model

Brams and Kilgour (2001) proposed a procedure where the winning candidates are selected in the following manner: Given q, $1 \le q \le n$, if a candidate is ranked first by at least q voters, that candidate wins; otherwise, if this threshold is achieved by some candidate when the first and second ranks are considered, that candidate is chosen; if not, lower rankings must be added until some candidate achieves the threshold established. The winning candidates can be obtained using the following cumulative standings function:

$$F(x_1,\ldots,x_{m-1}) = \mathrm{sg}^+(x_1 - q) + \cdots + \mathrm{sg}^+(x_{m-1} - q),$$

where

$$\mathrm{sg}^+(x) = \begin{cases} 1, & \text{if } x \ge 0, \\ 0, & \text{if } x < 0. \end{cases}$$

When $q > n/2$, the winning candidates can also be obtained using the following cumulative standings function:

$$F(x_1,\ldots,x_{m-1}) = \left[\frac{x_1}{q}\right] + \cdots + \left[\frac{x_{m-1}}{q}\right],$$

where $[x]$ is the integer part of x, i.e., the largest integer smaller than or equal to x.

3.4 Sertel and Yilmaz's model

Sertel and Yilmaz (1999) proposed a model similar to the previous one with $q = n/2$, but when the previous procedure is applied to the k first positions and various winning candidates are obtained, Sertel and Yilmaz propose to select those with the biggest sum of first, second, . . . , k-th ranks. The winning candidates in this model can be obtained with the following cumulative standings function:

$$\begin{aligned} F(x_1,\ldots,x_{m-1}) \;=\; & \frac{x_1}{n}\mathrm{sg}^+(x_1 - q) + \left(\frac{x_2}{n} + \left(1 - \frac{x_2}{n}\right)\mathrm{sg}^+(x_1 - q)\right)\mathrm{sg}^+(x_2 - q) + \cdots + \\ & + \left(\frac{x_{m-1}}{n} + \left(1 - \frac{x_{m-1}}{n}\right)\mathrm{sg}^+(x_{m-2} - q)\right)\mathrm{sg}^+(x_{m-1} - q). \end{aligned}$$

3.5 Obata and Ishii's model

One of the most important issues in the context of scoring rules is the choice of the scoring vector, since the winning candidates can vary according to the scores used. To avoid this problem, Obata and Ishii (2003), following the idea of Cook and Kress (1990), proposed to assess each candidate with the most favourable scoring vector for her/him using a DEA/AR model. In a particular case of this model (when L_1-norm is used and the discrimination intensity functions are zero), the score achieved by each candidate is:

$$\begin{aligned} Z(A_i) \;=\; & \max \sum_{k=1}^{m} s_k v_{ik} \\ & s.a. \quad \mathbf{s} \in \boldsymbol{\Delta}, \end{aligned}$$

where

$$\boldsymbol{\Delta} = \left\{ \mathbf{s} \in \Re^m \;|\; \mathbf{s} \ge \mathbf{0},\ \sum_{k=1}^{m} s_k = 1,\ s_k \ge s_{k+1}\ (k = 1,\ldots,m-1),\ \sum_{k=1}^{m} s_k v_{jk} \le \sum_{k=1}^{m} s_k v_{ik}\ (\forall j \ne i) \right\}.$$

The winning candidates with this model have been characterized by Llamazares and Peña (2009). A candidate A_o is a winner if and only if

$$\max_{k=1,\ldots,m} \frac{V_{ok}}{k} \geq \max_{k=1,\ldots,m} \frac{V_{ik}}{k}$$

for all $i \in \{1,\ldots,m\}$. Moreover, in this case,

$$Z(A_o) = \max_{k=1,\ldots,m} \frac{V_{ok}}{k}.$$

Therefore, the winning candidates can be obtained using the following cumulative standings function:

$$F(x_1,\ldots,x_{m-1}) = \max\left\{x_1, \frac{x_2}{2}, \ldots, \frac{x_{m-1}}{m-1}, \frac{n}{m}\right\}.$$

3.6 Contreras, Hinojosa and Mármol's model

To obtain the candidates' scores, Contreras, Hinojosa and Mármol (2005) proposed to use the following model:

$$\begin{aligned} Z(A_i) = \quad & \max \sum_{k=1}^{m} s_k v_{ik} \\ & s.a. \quad \mathbf{s} \in \boldsymbol{\Omega}, \end{aligned}$$

where

$$\boldsymbol{\Omega} = \left\{\mathbf{s} \in \Re^m \;|\; \mathbf{s} \geq \mathbf{0},\ \sum_{k=1}^{m} s_k = 1,\ s_{m-1} - s_m \geq s_m,\ s_k - s_{k+1} \geq s_{k+1} - s_{k+2}\ (k = 1,\ldots,m-2)\right\}$$

These authors showed that with this model the candidate A_i's score is

$$Z(A_i) = \max_{k=1,\ldots,m} \left\{ \sum_{j=1}^{k} \frac{k-j+1}{\frac{k(k+1)}{2}} v_{ij} \right\}.$$

This score could also be obtained using the following cumulative standings function:

$$F(x_1,\ldots,x_{m-1}) = \max\left\{ x_1, \frac{x_1+x_2}{3}, \ldots, \frac{x_1+\cdots+x_{m-1}+n}{\frac{m(m+1)}{2}} \right\}.$$

3.7 Positional voting systems based on means

In addition to the positional voting systems proposed in the literature, others can be established. We have seen above that scoring rules are obtained when weighted mean are used as cumulative standings functions. For this reason, it would be interesting to know that positional voting systems appear when we use another class of means. A specially interesting case arises when the following cumulative standings function, equivalent to geometric mean, is used:

$$F_1(x_1,\ldots,x_{m-1}) = x_1 \cdot x_2 \cdots x_{m-1}.$$

The score that provides this function coincides with the measure of the compact interval $[0,x_1]\times\cdots\times[0,x_{m-1}]$. In this way, if $m = 3$, the area of the above mentioned interval is obtained and, if $m = 4$, the volume is found.

4 CONCLUSIONS AND FUTURE LINES OF RESEARCH

In this paper we have established a general framework based on the cumulative standings functions for the analysis of positional voting systems. The proposed model allows us to extend some of the results appeared in the literature. For example, given that the cumulative standings functions are monotonic, it is immediate to verify the following result:

Given $\mathbf{V}$ *a cumulative standings matrix,* A_i *is a winning candidate for all positional voting systems if only if* $\mathbf{V}_i \geq \mathbf{V}_j$ *for all* $j \neq i$.

This result is an extension of the one obtained by Stein, Mizzi and Pfaffenberger (1994) for the particular case of scoring rules. Furthermore, some known results can be obtained in a straightforward manner. For example, it is easy to verify that $\mathbf{\Lambda}$ is a convex set and that its vertices are the scoring vectors that originate the *k*-approval voting systems. From this result it is possible to deduce the one obtained by Saari (1992):

A candidate is the winner with every scoring rule if only if is the winning candidate with every k-approval voting system, $k = 1, \ldots, m - 1$.

On the other hand, the proposed model is going to allow us the analysis of these systems through various properties that can be established for the cumulative standings functions. In this regard, there exist a great variety of properties that have been used to characterize scoring rules (see, among others, Gärdenfors (1973), Smith (1973), Young (1974, 1975), Richelson (1978), Nitzan and Rubinstein (1981), Roberts (1991), Myerson (1995), Ching (1996), Chebotarev and Shamis (1998) and Merlin (2003)) and that we will try to translate to the area of cumulative standings functions. This will probably enable us to obtain new characterizations for some positional voting systems.

ACKNOWLEDGEMENTS

This work is partially financed by the Junta de Castilla y León (Project VA002B08), the Spanish Ministry of Education and Science (Projects SEJ2006-04267 and ECO2008-02358) and ERDF.

REFERENCES

BRAMS, SJ, KILGOUR, DM (2001) "Fallback bargaining", Group Decision and Negotiation, 10, pp 287–316.

CHEVOTAREV, PY, SHAMIS, E (1998), "Characterizations of scoring methods for preference aggregation", Annals of Operations Research, 80, pp 299–332.

CHING, S (1996), "A simple characterization of Plurality rule", Journal of Economic Theory, 71, pp 298–302.

CONTRERAS, I, HINOJOSA, MA, MÁRMOL, AM (2005), "A class of flexible weight indices for ranking alternatives", IMA Journal of Management Mathematics, 16, pp 71–85.

COOK, WD, KRESS, M (1990), "A data envelopment model for aggregating preference rankings", Management Science, 36, pp 1302–1310.

GÄRDENFORS, P (1973), "Positionalist voting function", Theory and Decision, 4, pp 1–24.

GREEN, RH, DOYLE, JR, COOK, WD (1996), "Preference voting and project ranking using DEA and cross-evaluation", European Journal of Operational Research, 90, pp 461–472.

LLAMAZARES, B, PEÑA, T (2009) "Preference aggregation and DEA: An analysis of the methods proposed to discriminate efficient candidates", European Journal of Operational Research, 197, pp 714-721.

MERLIN, V (2003), "The axiomatic characterizations of majority voting and scoring rules", Mathématiques et Sciences Humaines, 163, pp 87–109.

MYERSON, R (1995), "Axiomatic derivation of scoring rules without the ordering assumption", Social Choice and Welfare, 12, pp 59–74.
NITZAN, T, RUBINSTEIN, A (1981), "A further characterization of the Borda ranking method", Public Choice, 36, pp 153–158.
OBATA, T, ISHII, H (2003), "A method for discriminating efficient candidates with ranked voting data", European Journal of Operational Research, 151, pp 233–237.
RICHELSON, JT (1978), "A characterization result for Plurality rule", Journal of Economic Theory, 19, pp 548–550.
ROBERTS, FS (1991), "Characterizations of the Plurality function", Mathematical Social Sciences, 21, pp 101–127.
SAARI, DG (1992), "Millions of election outcomes from a single profile", Social Choice and Welfare, 9, pp 277–306.
SERTEL, MR, YILMAZ, B (1999), "The majoritarian compromise is majoritarian-optimal and subgame-perfect implementable", Social Choice and Welfare, 16, pp 615–627.
STEIN, WE, MIZZI, PJ, PFAFFENBERGER, RC (1994), "A stochastic dominance analysis of ranked voting systems with scoring", European Journal of Operational Research, 74, pp 78–85.
SMITH, JH (1973), "Aggregation of preferences with variable electorate", Econometrica, 41, pp 1027–1041.
YOUNG, HP (1974), "An axiomatization of Borda's rule", Journal of Economic Theory, 9, pp 43–52.
YOUNG, HP (1975), "Social choice scoring functions", SIAM Journal of Applied Mathematics, 28, pp 824–838.

STRUCTURAL EQUIVALENCE IN INPUT-OUTPUT FIELD. A SCOPE FROM NETWORK THEORY

ANA SALOMÉ GARCÍA MUÑIZ
E-mail: asgarcia@uniovi.es
Facultad de C.C. Económicas y Empresariales
Departamento de Economía Aplicada
Universidad de Oviedo

Resumen

El estudio de la organización de las transacciones intersectoriales es una fuente de información enriquecedora sobre el esquema de relaciones económicas de un país que permite su comparación con otra economía de referencia En éste trabajo, proponemos la aplicación y estudio de una nueva herramienta en el contexto input-ouput que permite mostrar el funcionamiento de la estructura económica a través de la determinación de grupos de sectores equivalentes estructuralmente. Este concepto se encuentra próximo a la Teoría de Redes y Grafos. Su desarrollo es muy flexible y puede ser aplicado tanto a grafos valuados como Booleanos superando las críticas habituales en los estudios sobre grafos en el campo input-output.

Palabras Clave: Teoría de redes, análisis input-output, análisis estructural
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

The systematic study of the way in which intersectoral transactions are organized is an interesting source of information about the existing scheme of economic relations of one country that allows comparison of its structural features with those of another economy of reference. In this work, we propose the application and the study of a new tool in the input-output context that allows us to reveal the functioning of the economic structure via the determination of structural equivalent sectoral groups. The concept is close to Network Theory and Graph Theory. Its implementation is very flexible and can be applied on Boolean or valued graphs overcoming the usual criticisms of graph input-output studies.

Key words: Network theory, input-output field, structural analysis
Thematic Area: Quantitative methods for economics and business.

1. INTRODUCCIÓN

El estudio sistemático de la forma en la cual las relaciones sectoriales son establecidas, constituye una fuente de información básica sobre el esquema de relaciones de un país y sus características estructurales.

Desde sus orígenes, el modelo input-output (IO) ha estado ligado al análisis económico de estructuras. "Sin embargo, en general, el análisis cuantitativo ha fallado al describir las conexiones reales entre sectores, de tal forma que la interacción industrial permanece aún no claramente determinada. En contraste, el análisis qualitativo input-output (QIOA) revela los rasgos básicos de la estructura económica" (Aroche-Reyes, 2003). El objetivo de este tipo de análisis es descubrir el conjunto de relaciones industrials que caracterizan la red económica en su conjunto (Schnabl, 1994; Ghos and Roy, 1998; Aroche-Reyes, 1996 and 2006). Construida explícitamente sobre la Teoría de Grafos, QIOA binariza los datos input-output disponibles y trabaja, por tanto, sobre grafos booleanos.

En este trabajo, se propone la aplicación y el estudio de una nueva herramienta en el contexto input-output que permita desvelar el funcioamiento de la estructura económica a través de la determinación de grupos de sectores equivalentes estructuralmente El concepto de equivalencia estructural, derivado de la Teoría de Redes, adquiere pleno sentido en la búsqueda de posicione idénticas en una estructura, lo cual lo convierte en un término de interés y enorme aplicabilidad en el campo del análisis estructural. Este concepto permite caracterizar grupos de sectores con roles similares. que comparten una posición común desde la cual expandir sus efectos sobre la red económica. Su aplicación en el contexto input-output se puede encuadrar en la línea de los trabajos basados en la teoría de grafos y redes. Asimismo, el resultado de grupos de sectores equivalentes estructuralmente permite mostrar la técnica propuesta como complementaria a los estudios de cluster industriales. Los trabajos pioneros de Czamanski (1974) propiciaron un interés creciente en este tipo de estudios. Los análisis de cluster industriales son numerosos y diversos en sus consideraciones, desde entonces. En el análisis input-output, bajo esta línea de investigación han surgido trabajos aplicando técnicas estadísticas multivariantes y más recientemente, análisis de conglomerados difuso (Czamanski, 1974; Roepke et. al., 1974; Ó hUallacháin, 1984, Díaz, Moniche y Morillas, 2006).

El resto del documento se organiza de tal forma que en el siguiente epígrafe se presenta el concepto de equivalencia estructural. En la siguiente sección se detalla la metodología empleada en la búsqueda de sectores equivalentes estructuralmente. A través de esta propuesta, se realiza un estudio comparativo de las estructuras productivas de España y Europa en 1995 a partir del cual se pueden extraer algunas conclusiones sobre el comportamiento de ambas economías.

2. EL CONCEPTO DE EQUIVALENCIA ESTRUCTURAL

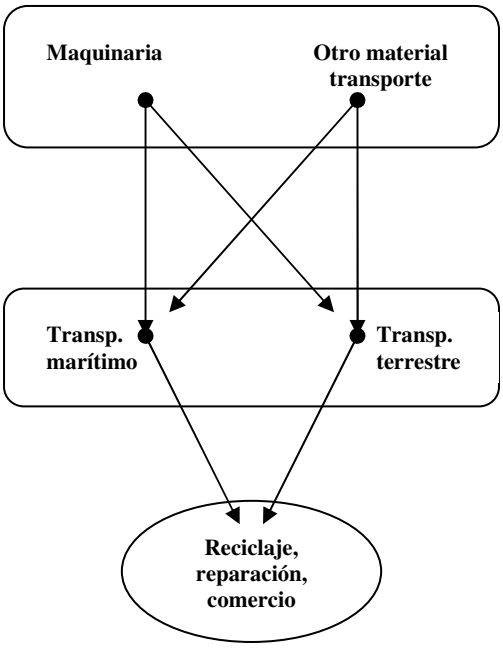
En el campo de la Teoría de Redes, se han propuesto diversas concepciones del término equivalencia[1]. El concepto inicial de clases equivalentes fue propuesto por Lorrain and White (1971) a través de la idea de equivalencia estructural.

En el análisis input-output, la clasificación de ramas productivas en agrupaciones equivalentes estructuralmente posibilita la comparación de un sector con el resto, en términos de la equivalencia de sus intercambios con el conjunto de la economía, lo que permite formar bloques que reúnen ramas con semejanzas en sus interrelaciones. Se determinan así, posiciones estructurales comunes de los agentes económicos a partir de las similitudes de las transacciones entre sectores. Determinando los sectores a través de posiciones equivalentes que se relacionan de forma análoga con otras posiciones, la red económica se transforma en un modelo simplificado donde los sectores están agrupados en posiciones y las relaciones entre sectores se convierten en relaciones entre posiciones.

Así pues, consideremos dos sectores X y Z que compran productos (bienes y servicios) a un sector Y, las ramas X y Z son estructuralmente equivalentes, desde la perspectiva de las compras, puesto que muestran un patrón relacional común. Considérese, a modo de ilustración, la siguiente red sectorial presentada en la figura nº 1.

[1] Ver Wasserman y Faust (1994).

Figura nº 1. Sectores equivalentes estructuralmente



Bajo esta óptica, los sectores servicios de transporte marítimo y aéreo y servicios de transporte terrestre resultarían estructuralmente equivalentes, ya que, por un lado, ambos ofrecen sus servicios a la rama reciclaje, reparación y comercio y por otro, compran a los sectores maquinaria industrial y otro material de transporte, respectivamente. Su posición puede ser catalogada como intermedia dentro del grafo planteado. Análogamente, maquinaria industrial y otro material de transporte compartirían una misma posición estructural. Dicho con otras palabras, aquellos sectores productivos que presentan análogas conexiones tanto en compras como en ventas con el resto de ramas de la red, son estructuralmente equivalentes.

En este sentido, las ramas que muestran similares relaciones son estructuralmente equivalentes cuando ocupan la misma posición en la red (White, Boorman y Breiger, 1976). Dichos sectores se verán afectados, por tanto, por los mismos riesgos y oportunidades asociados a lo posición que comparten. Bajo esta hipótesis, el comportamiento de sectores equivalentes estructuralmente ante cambios el entorno tenderá a ser equivalente. Un rasgo a tener en cuenta en la implementación de políticas económicas.

La perspectiva propuesta en el presente trabajo, complementa la visión tradicional de otros estudios basados en el grado de cohesión existen entre sectores determinada por las interacciones directas entre sectores e indirectas a través de intermediarios. Sin embargo, el concepto de equivalencia estructural propuesto asume que los sectores interactúan (directa e indirectamente) con otros de forma similar. Dos sectores son equivalentes estructuralmente independientemente tanto, de si mantienen o no relaciones directas, tanto de la longitud de los caminos que los conecta indirectamente. La equivalencia estructural está basada en la configuración de las relaciones sectoriales.

3. DETERMINACIÓN DE SECTORES EQUIVALENTES ESTRUCTURALMENTE

En la práctica, es difícil localizar sectores que sean exactamente equivalentes estructuralmente. Se requiere, más bien, identificar y localizar conjuntos de ramas productivas que sean "aproximadamente" equivalentes desde un punto de vista estructural de acuerdo con alguna medida de similitud.

Un procedimiento ampliamente utilizado en la literatura de redes sociales con este fin es el denominado CONCOR, un algoritmo iterativo basado en el cálculo de coeficientes de correlación[2]. Esta medida de dependencia forma parte de un extenso abanico de indicadores

[2] Planteado de forma paralela por los investigadores McQuitty y Clark (1968), Breiger et al. (1975) y White et al. (1976).

posibles para medir el grado de equivalencia estructural, entre las cuales destaca la distancia euclídea empleada en el método STRUCTURE (Burt, 1989). La selección de una u otra, depende del objeto de estudio. En concreto, las similitudes de las transacciones productivas se logran determinar a través de divergencias o distancias; mientras que el análisis de patrones análogos de relaciones intersectoriales -concepto de equivalencia estructural- se perfila a través del coeficiente de correlación[3].

El propio desarrollo metodológico del algoritmo CONCOR ha servido de inspiración para técnicas más recientes en otros ámbitos científicos, las cuales no aluden directamente al concepto de equivalencia estructural, aunque determinan unos resultados análogos. Dubnov et. al. (2002) presentan un algoritmo cluster no paramétrico, directamente comparable con el algoritmo CONCOR, basado en la estimación iterativa de perfiles de distancias entre los agentes medidas a través de la divergencia de Jensen-Shannon. Constituye ésta una técnica con unas características y propiedades análogas a las del algoritmo CONCOR, el cual se viene aplicando tradicionalmente en la teoría de redes sociales de forma eficiente.

En este sentido, la utilización del método CONCOR nos ha parecido adecuada, dado no sólo los buenos resultados que proporciona, sino también "la lógica del proceso" empleada, ya que conduce a la creación de bloques de sectores que presentan análogas relaciones utilizando como instrumento de medida el coeficiente de correlación.

3.1. Formalización del algoritmo CONCOR

El método CONCOR, de cuyas siglas se desprende su filosofía (CONvergence of iterated CORrelations), es un algoritmo iterativo desarrollado a partir de un proceso de re-estimación de coeficientes de correlación entre columnas (filas). Para su aplicación se parte de una matriz cuadrada de datos donde cada columna (fila) representa un sector, posteriormente se calcula el coeficiente de correlación de cada una de dichas columnas (filas) relacionándolas con el resto, y se reitera el proceso[4].

Formalmente, sea una tabla $\mathbf{X}$ de orden (nxn) cuyos elementos pueden ser flujos o coeficientes. Se comienza por calcular los coeficientes de correlación lineal de cada una de las columnas con el resto, de esa forma se recogen las relaciones existentes entre los distintos sectores. Se reitera el proceso repetidamente. De forma genérica, la t-*ésima* iteración del algoritmo CONCOR da lugar a los siguientes resultados:

$$\mathbf{R}^{(t)} = \hat{\mathbf{d}}_S^{(t)} \mathbf{S}^{(t)} \hat{\mathbf{d}}_S^{(t)} \qquad (1)$$

donde el superíndice "t" hace referencia a que se trata t-esima iteración, $\hat{\mathbf{d}}_S^{(t)}$ representa una matriz diagonal cuyos elementos son los inversos de las desviaciones típicas respectivas, $\mathbf{S}^{(t)}$ corresponde a la matriz de varianzas-covarianzas tal que:

$$\mathbf{S}^{(t)} = \frac{1}{n} \mathbf{R}^{(t-1)} \boldsymbol{\Psi} \mathbf{R}^{(t-1)} \qquad (2)$$

y $\boldsymbol{\Psi}$ corresponde a una matriz idempotente de orden (nxn) definida como: $\boldsymbol{\Psi} = \left(\mathbf{I} - \frac{1}{n} \mathbf{i}\mathbf{i}' \right)$

donde $\mathbf{I}$ es la matriz identidad.

De tal forma que, en el cálculo del coeficiente de la iteración t-*ésima* intervendrán los coeficientes de correlación de la iteración anterior (t-1) asociados a la posición que ocupe el coeficiente.

La secuencia de matrices $\mathbf{R}$ converge (Breiger et. al., 1975; White et. al., 1976) a una matriz $\mathbf{R}^{(\infty)}$, que muestra la estructura interna existente en la matriz de datos original ($\mathbf{X}$). Si los datos no presentan un patrón simétrico perfecto, un caso habitual según muestran los trabajos

[3] Ver Faust y Romney (1985) sobre una visión crítica del método STRUCTURE para el análisis de la equivalencia estructural.

[4] Según se desee analizar las posiciones de los sectores desde el punto de vista de las compras o ventas, pudiéndose combinar ambas perspectivas a partir de la consideración conjunta de la matriz de datos inicial $\mathbf{X}$ y su traspuesta.

aplicados de esta metodología[5], $\mathbf{R}^{(\infty)}$ converge a una matriz cuyos elementos son +1,-1, subdivida en dos grupos claramente diferenciados y de rango k=1. Es decir, la matriz obtenida tras efectuar iteraciones queda dividida en dos bloques que permiten dividir el conjunto de ramas productivas en dos subgrupos claramente diferenciados. Sucesivas aplicaciones de dicho método sobre los grupos precedentes permiten subdividirlos progresivamente.

4. DETERMINACIÓN DE SECTORES ESTRUCTURALMENTE EQUIVALENTES EN ESPAÑA Y EUROPA

En este trabajo, se determinan los sectores equivalentes estructuralmente de España y Europa. Este ejercicio empírico permite extraer algunas conclusiones sobre las similitudes y diferencias entre España y el área económica y social a la cual pertenece. La información empleada como punto de partida es la suministrada por las tablas input-output de España (TIOE-95) y Unión Europa (TIOEU-95) de 1995 con un nivel de desagregación de 25 sectores , dado que a nivel europeo constituye la base de datos más reciente. La identificación de sectores equivalentes estructuralmente puede realizarse a partir de cualquier tipo de grafo- booleano o valuado- asociado con la estructura estudiada. En este artículo, se usan grafos valuados asociados a los coeficientes input-output internos. El empleo de este tipo de grafos, permite superar las críticas habituales en el análisis cualitativo input-output[6]. En cualquier caso, los resultados asociados al grafo Booelano son similares a los presentados.

El resultado de aplicar esta metodología mediante la utilización del algoritmo CONCOR, implementado en Netminer[7], aparece en el cuadro nº 1 conjuntamente para la economía europea y española. Dicho cuadro condensa en catorce y quince bloques, respectivamente las ramas de actividad que presentan similares interrelaciones. Hemos denotado cada agrupación con una letra.

Cuadro nº 1. Bloques equivalentes estructuralmente. Grafo valuado

<table>
<tr><th>Bloques</th><th>Sectores UE</th><th>Sectores ESPAÑA</th><th>Bloques</th></tr>
<tr><td>A</td><td colspan="2">11. Alimentación, bebida y tabaco
1. Agricultura, silvicultura, pesca
18. Hostelería</td><td>A</td></tr>
<tr><td>B</td><td>9. Material y accesorios eléctricos
8. Máquinas de oficina, material informático
7. Maquinaria industrial y agrícola</td><td>9. Material y accesorios eléctricos
8. Máquinas de oficina, material informático</td><td>B1</td></tr>
<tr><td>C</td><td>10. Otro material de transporte</td><td>7. Maquinaria industrial y agrícola
10. Otro material de transporte</td><td>C1</td></tr>
<tr><td>D</td><td colspan="2">15. Otras manufacturas</td><td>D</td></tr>
<tr><td rowspan="3">E</td><td rowspan="3">3. Minerales y metales férreos y no férreos
4. Productos minerales no metálicos
6. Productos metálicos excepto maquinaria</td><td>3. Minerales y metales férreos y no férreos</td><td>E1</td></tr>
<tr><td>4. Productos minerales no metálicos</td><td>E2</td></tr>
<tr><td>6. Productos metálicos excepto maquinaria</td><td>E3</td></tr>
</table>

[5] Clark y McQuitty (1970) y Chen (2002) recopilan una serie de supuestos bajo los cuales el estado estacionario no responde a la matriz particionada en -1, +1. Sin embargo, se tratan de condiciones tan restrictivas en nuestro ámbito de estudio, que apenas presentan una repercusión práctica (Ver García, 2006).

[6] Normalmente, los teoremas y aplicaciones de la Teoría de Grafos están basados en estructuras booleanas. En el análisis input-output, las aplicaciones de la Teoría de Grafos convierten la matriz de transacciones en una matriz binaria. Todos los valores por debajo de cierto filtro son sustituidos por cero y aquellos mayores o iguales a dicho filtro por uno. Por ello, cierta información se pierde en el proceso de binarización (de Mesnard, 2001).

[7] Ver http://www.netminer.com.

Cuadro nº 1. (cont)

Bloques	Sectores UE	Sectores ESPAÑA	Bloques
F	13. Papel y edición 5. Química	13. Papel y edición 5. Química 12. Textil, vestido, cuero y calzado	F1
G	14. Caucho y plástico 12. Textil, vestido, cuero y calzado	14. Caucho y plástico	G1
H	17. Reciclaje, servicios reparación, comercio 2. Productos energéticos	17. Reciclaje, servicios reparación, comercio 2. Productos energéticos 22. Comunicaciones	HI
I	22. Comunicaciones		
J	16. Construcción		J
K	19. Servicios de transporte terrestre 20. Servicios de transporte marítimo y aéreo 21. Servicios auxiliares de transporte		K
L	23. Servicios de crédito e instituciones de seguro		L
M	25. Servicios de no mercado		M
N	24. Otros servicios de mercado		N

Fuente: Elaboración propia a partir de TIOEU-95 y TIOE-95.

La estructura española recogida en el cuadro nº 1 muestra similitudes con la economía europea, pero asimismo diferencias notables. Por un lado, existe una serie de sectores que comparten una misma posición equivalente estructuralmente en ambos territorios. En concreto, los bloques A, D, J, K, L, M y N son comunes a ambas economías. El conglomerado A está formado por los sectores de alimentación, bebida y tabaco, agricultura, silvicultura, pesca y hostelería. Los patrones de integración vertical mostrados por estos sectores pueden haber repercutido en la posición compartida entre ambos.

El bloque D está constituido en torno a una rama industrial, otras manufacturas, de baja tecnología, cuya posición difiere del resto de actividades manufactureras. Asimismo, el conglomerado J está formado por un único sector, la construcción, clave actualmente en el grado de dinamismo que muestran las redes económicas.

Los últimos bloques (K, L, M y N) se refieren en su totalidad a un sector de servicios, que muestra una posición análoga en ambas economías, lo cual puede indicar un grado de desarrollo de las actividades terciarias similar en ambos territorios. Sus conexiones productivas, tanto en compras como en ventas, permiten establecer comportamientos diferenciados que guardan una clara relación con el grado de intensidad tecnológica. Por un lado, las ramas de servicios de no mercado y otros servicios de mercado, que constituyen de forma independiente los conglomerados M y N. Por otro, los servicios de alta tecnología entre los que se encuentran las instituciones de seguro (L) y el conglomerado de actividades relacionadas con el transporte (K). En este último es interesante destacar que los diferentes tipos de transporte –terrestre, marítimo y aéreo- junto con sus servicios auxiliares comparten una misma posición desde la cual expandir sus efectos, un rasgo a considerar en la planificación y desarrollo de políticas en infraestructuras.

Genéricamente, las diferencias fundamentales entre ambas economías se sitúan en el sector industrial recogido en los bloques B y C. La distinta composición tecnológica de las manufacturas españolas respecto a las europeas creemos que influye decisivamente en este aspecto. El menor desarrollo relativo en España de las ramas manufactureras de mayor nivel tecnológico (Buesa, 1996) puede ser la causa de los diferentes patrones determinados.

Los sectores relacionados con la transformación, fundición y fabricación de productos metalúrgicos recogidos a nivel europeo en el bloque E resultan equivalentes estructuralmente, es decir, presentan similares interrelaciones. Sin embargo, la posición de estos en España se muestra diferenciada (E1, E2, E3). No es de extrañar puesto que el sector siderúrgico presenta unas peculiares características como resultado de los períodos autárquicos y de liberación

progresiva de la economía española durante el régimen político de Franco. Su relevancia ha resultado clave a lo largo de décadas en muchas de las regiones españolas

De forma análoga, el comportamiento de los sectores recogidos en los bloques F G y H es diferente en ambas economías mostrando unas posiciones más diferenciadas a nivel europeo.

5. CONCLUSIONES

El concepto de equivalencia estructural, habitual en el campo de la Teoría de Redes, resulta de enorme utilidad en el análisis estructural input-output. Su aplicación permite perfilar los sectores que comparten una misma posición y mostrar sintéticamente los rasgos estructurales más importantes de una red económica. Este objetivo es común a las técnicas desarrolladas en el campo input-output bajo el nombre de análisis cualitativo input-output. Sin embargo, una crítica crucial de éste ultimo enfoque es la pérdida de información valiosa durante el proceso de binarización (De Mesnard, 2001). Reintroducir la información cuantitativa como se ha planteado en este trabajo y, en otros tales como Aroche-Reyes (2006), puede resolver esta crítica (Schnabl, 1994).

Los resultados obtenidos para España y la Unión Europea muestran la semejanza cualitativa entre ambas economías. Los sectores primario y terciario forman, en general, agrupaciones equivalentes estructuralmente tanto en Europa como en España. Una característica que apunta a una vertebración similar de los flujos económicos de dichos sectores en ambos territorios. En contraposición, la situación de las manufacturas en las respectivas redes económicas se muestra diferenciada. Si bien es cierto, que en cualquier país europeo el sector industrial sigue siendo fundamental para el conjunto de la actividad económica, también lo es, que muchos de ellos mantienen sus propias características. En este sentido, la manufactura española continúa mostrando una estructura tradicional, con un papel destacado de las actividades siderúrgicas y una fuerte presencia de ramas productivas de intensidad tecnológica baja.

6. BIBLIOGRAFÍA

AROCHE-REYES, F. (1996): "Important coefficients and structural change: a multi-layer approach", *Economic Systems Research,* 8*,* pp. 235-246.

AROCHE-REYES, F. (2002): "Structural transformations and important coefficients in the North American Economies", *Economic Systems Research,* 14, pp. 257-273.

AROCHE-REYES, F. (2003): "A qualitative input-output method to find basic economic structures" *Papers in Regional Science*, 82, pp. 581-590.

AROCHE-REYES, F. (2006): "Trees of the essential economic structures: a qualitative input-output method", *Journal of Regional Science*, 46, pp. 333-353.

BREIGER, R.L., BOORMAN, S.A., ARABIE, P. (1975): "An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling", *Journal of Mathematical Psychology*, 12, pp. 328-383.

BUESA, M. (1996): "La industria española en el marco europeo. Un análisis en la perspectiva sectorial y empresarial", Documento de Trabajo 2, Instituto de Análisis Económico y Financiero.

BURT, R.S. (1989): *Structure, Version 4.0*., Research Program in Structural Analysis, Center for the Social Sciences, Columbia University.

CHEN, C-H. (2002): "Generalized association plots: information visualization via iteratively generated correlation matrices", *Statistica Sinica*, 12, pp. 7-29.

CLARK, J., McQUITTY, L. (1970): "Some problems and elaborations of iterative, intercolumnar correlational analysis", *Educational and Psychological Measurement*, 30, pp. 773-784.

CZAMANSKI, S. (1974): *Study of Clustering of Industries*, Institute of Public Affairs, Dalhousie University, Halifax, Canada.

DE MESNARD, L. (1995):

DE MESNARD, L. (2001): "On Boolean topological methods of structural analysis", en Lahr, M.L., Dietzenbacher, E. (eds.): *Input-Output Analysis: Frontiers and Extensions*. Palgrave Macmillan, Basingstoke, UK.

DIAZ, B., MONICHE, L., MORILLAS, A. (2006): "A Fuzzy clustering approach to the key sectors of the Spanish Economy", *Economic Systems Research,* 18, pp. 299-318.

DUBNOV, S., EL-YANIV, R., GDALYAHU, Y., SCHNEIDMAN, E., TISHBY, N., YONA, G. (2002): "A new nonparametric pairwise clustering algorithm based on iterative estimation of distances profiles", *Machine Learning*, 47, pp. 35-61.

FAUST, K., ROMNEY, A. (1985): "Does STRUCTURE find structure?: A critique of Burt's use of distance as a measure of structural equivalence", *Social Networks*, 7, pp. 77-103.

GHOSH, S., ROY, J. (1998): "Qualitative Input-Output Analysis of the Indian Economic Structure", *Economic Systems Research*, 10, pp. 263-273.

LORRAIN, F., WHITE, H. (1971): "Structural Equivalence of Individuals in Social Networks", *Journal of Mathematical Sociology*, 1, pp.49-80.

McQUITTY, L., CLARK, J. (1968): "Clusters from iterative, intercolumnar correlational analysis", *Educational and Psychological Measurement*, 28, pp. 211-238.

Ó hUALLACHAIN, B. (1984): "The Identification of Industrial Complexes", *Annals of the Association of American Geographers,* 74, pp. 420-436.

ROEPKE, H., ADAMS D., WISEMAN, R. (1974): "A New Approach to the Identification of Industrial Complexes Using Input-Output Data", *Journal of Regional Science,* 14, pp. 15-29.

SCHNABL, H. (1994): "The evolution of production structures analyzed by a Multi-layer procedure", *Economic Systems Research,* 6, pp. 51-68.

WASSERMAN, S., FAUST, K. (1994): *Social Network Analysis. Methods and Applications, Structural Analysis in the Social Sciences*, Cambridge University Press, New York.

WHITE, H., BOORMAN, S., BREIGER, R. (1976): "Social structure from multiple networks I. Blockmodels of roles and positions", *American Journal of Sociology*, 81, pp. 730-780.

LOCALIZACIÓN COMPETITIVA EN REDES CON REGLAS DE ELECCIÓN QUE INCORPORAN UMBRALES DE PREFERENCIA

DOLORES R. SANTOS-PEÑATE
E-mail: drsantos@dmc.ulpgc.es
RAFAEL SUÁREZ-VEGA
PABLO DORTA-GONZÁLEZ
E-mail: pdorta@dmc.ulpgc.es
Departamento de Métodos Cuantitativos en Economía y Gestión
Universidad de Las Palmas de Gran Canaria (España)

Resumen

En este trabajo se estudia un problema de localización competitiva en redes donde los bienes o servicios son considerados esenciales y la regla de elección de los consumidores se basa únicamente en la distancia. Dicha regla de elección incorpora unos umbrales de preferencia, de manera que todo el poder de compra de un cliente es captado por una firma solamente si la diferencia de las distancias entre el cliente y esta firma y entre el cliente y las firmas competidoras supera cierto umbral. Este modelo de comportamiento constituye una alternativa al modelo binario donde el consumidor visita el centro de servicio más próximo, a la vez que aporta un marco general que incluye los modelos binario, parcialmente binario y proporcional como casos particulares. Se analiza el modelo propuesto, se presentan algunos resultados relacionados con la transformación del problema en redes en otro discreto y se proponen algunas extensiones.

Palabras Clave: Localización competitiva - redes - reglas de elección de los consumidores - umbrales de preferencia.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

In this work, we study a competitive location problem on networks. Goods are assumed to be essential and the customer choice rule is defined in terms of the travel distance. We introduce a customer choice rule which incorporates preference thresholds, a customer patronizes exclusively a firm only if the distance from the customer and this firm and the distance from the customer and the competitor firms are different enough. Our model constitutes an alternative to the binary choice rule according to which a customer visits the closest facility. On the other hand, this model provides a general frame where the binary, partially binary and proportional choice rules are included. We analyse the model and give some *discretization* results.

Thematic Area: Quantitative methods for economics and business.

1 INTRODUCCIÓN

Los modelos de localización competitiva representan situaciones en las que dos o más firmas compiten para ofrecer bienes o servicios en un mercado. El objetivo natural de cada firma es maximizar el beneficio, el cual se sustituye a veces por la demanda total captada o cuota de mercado. El comportamiento del consumidor es incorporado en el modelo mediante una regla de elección. Entre las reglas de elección del consumidor más utilizadas encontramos la regla binaria, la parcialmente binaria y la proporcional. En el modelo binario se supone que el individuo acude al centro más próximo; en el modelo parcialmente binario el consumidor acude al centro más próximo de cada firma; en el modelo proporcional el individuo utiliza su poder de compra en todos los centros en una proporción decreciente de la distancia que lo separa de ellos. Si los bienes o servicios ofrecidos se consideran esenciales, el individuo emplea todo su poder de compra, es decir, la demanda es constante (también denominada demanda inelástica), no depende de la distancia que tenga que recorrer el individuo para alcanzar el centro donde será servido. Cuando los bienes no son esenciales el consumidor puede decidir no utilizar todo su poder de compra si considera que los puntos de servicio están demasiado alejados. Los problemas de localización competitiva han sido tratados en varios trabajos en el marco de la Investigación Operativa, Economía Regional y Geografía Económica, entre otras áreas, véase por ejemplo Benati y Hansen (2000), Craig et al. (1984), Drezner (1995), Eiselt y Laporte (1989,1996), Friesz et al. (1988), Hakimi (1990), Plastria (2001), Santos-Peñate et al. (2007), Spaulding y Cromley (2007), y Till (1992, 2000).

En localización competitiva podemos considerar escenarios distintos: 1) una firma desea entrar en un mercado sin competidores con un cierto número de centros de servicio; 2) existen firmas operando en el mercado y una nueva firma quiere entrar en el mercado con un número dado de centros; 3) existen varias firmas competidoras operando en el mercado y una de ellas desea abrir unos centros y cerrar otros. En cualquier caso, con el fin de determinar las localizaciones mejores para los centros nuevos así como aquellos centros abiertos que conviene cerrar, es fundamental representar adecuadamente el comportamiento de los consumidores. La regla de elección binaria según la cual el consumidor acude al centro más próximo implica que el individuo es sensible a cualquier diferencia de las distancias a los centros lo cual no es muy realista. En este trabajo consideraremos una regla de elección donde un consumidor visita exclusivamente una firma solamente si la diferencia entre las distancias que separan al consumidor de los distintos competidores es mayor (mayor o igual) que cierto umbral.

Entre los modelos de localización competitiva, el modelo de Stackelberg representa un problema de localización secuencial en el que dos competidores, el líder y el seguidor, compiten con el objetivo de maximizar la demanda captada. El problema del seguidor es determinar las localizaciones óptimas de *r* centros suponiendo que la firma líder ya está operando en el mercado con *p* centros. El problema del líder es encontrar las *p* mejores localizaciones para instalar centros en un mercado sin competidores suponiendo que en el futuro entrará otra firma comportándose como un seguidor. Para el caso de demanda constante (inelástica) y espacios discretos, los problemas del consumidor, del líder y del seguidor, pueden ser formulados como problemas de programación lineal (ver Campos-Rodríguez et al. (2009)).

En este trabajo consideraremos el problema del seguidor en redes. El modelo es introducido en la sección 2. En la sección 3 se muestran algunos casos particulares y ejemplos. La sección 4 contiene varios resultados relacionados con la transformación del problema en redes en un problema discreto. En la sección 5 se mencionan algunas conclusiones y extensiones del trabajo.

2. EL MODELO

Consideramos una red $N(V,E)$ no dirigida con conjunto de nodos $V=\{v_i\}_{i=1}^{n}$ y conjunto de aristas *E*. Cada arista $e\in E$ tiene asociada una longitud $l(e)\geq 0$. Si $x_1, x_2 \in e$ con $e\in E$, denotamos $[x_1,x_2]$ el segmento cerrado que une x_1 y x_2, esto es, el conjunto de puntos

de la arista e situados entre x_1 y x_2, incluyendo x_1 y x_2. La distancia entre dos puntos de la red x_1 y x_2 es $d(x_1,x_2)$. La distancia entre un punto $z \in N$ y un conjunto $X \subset N$ es $d(z,X)=\min\{d(z,x): x \in X\}$. La red N representa un área donde varias firmas ofrecen bienes o servicios esenciales a consumidores distribuidos en el conjunto de nodos, siendo $w(v) \geq 0$ la demanda en el nodo v. La demanda total es $W=\sum_{v \in V} w(v)$.

Existen p centros en las localizaciones $X_p=\{x_1,\ldots,x_p\}$ pertenecientes a la firma líder, la firma *A*, que ya opera en el mercado. Una firma seguidora, la firma *B*, desea entrar en el mercado con *r* centros y desea determinar el conjunto de localizaciones que maximiza la demanda captada. Cualquier punto de la red es una localización candidata para abrir un centro. Considerando que *B* abre sus centros en el conjunto de puntos $Y_r=\{y_1,\ldots,y_r\}$, un consumidor en el nodo *v* visita únicamente centros de una firma si las distancias $d(v,X_p)$ y $d(v,Y_r)$ son *suficientemente diferentes*, en otro caso la demanda en este punto es compartida de alguna forma por las dos firmas. Dados los valores $\underline{\alpha}_v$ y $\overline{\alpha}_v$, los umbrales de sensibilidad para *v*, los consumidores en este nodo visitan únicamente la firma más próxima si $d(v,Y_r) < d(v,X_p)-\underline{\alpha}_v$ o $d(v,Y_r) > d(v,X_p)+\overline{\alpha}_v$. La demanda en v captada por la firma B es $w^B(v)=w(v)H_v(X_p,Y_r)$ donde $H_v(X_p,Y_r)$ es una función real no negativa, decreciente respecto de $d(v,Y_r)$, tal que $H_v(X_p,Y_r)=1$ si $d(v,Y_r) < d(v,X_p)-\underline{\alpha}_v$ y $H_v(X_p,Y_r)=0$ si $d(v,Y_r) > d(v,X_p)+\overline{\alpha}_v$. Utilizaremos funciones $H_v(X_p,Y_r)$ de la forma

$$H_v(X_p,Y_r)=\begin{cases} 1 & si \quad d(v,Y_r) < d(v,X_p)-\underline{\alpha}_v \\ h_v(Y_r) & si \quad d(v,X_p)-\underline{\alpha}_v \leq d(v,Y_r) \leq d(v,X_p)+\overline{\alpha}_v \\ 0 & si \quad d(v,Y_r) > d(v,X_p)+\overline{\alpha}_v \end{cases}$$

donde h_v es una función real no negativa decreciente respecto de $d(v,Y_r)$.

Dado X_p, la demanda total captada por la firma B es $W^B(X_p,Y_r)=\sum_{v \in V} w^B(v)$. Como la demanda es esencial, la demanda captada por la firma A es $W^A(X_p,Y_r)=W-W^B(X_p,Y_r)$. El problema de la firma B es encontrar el conjunto $Y_r^*=Y_r(X_p)=\{y_1^*,\ldots,y_r^*\}$ que maximiza la demanda captada por B, esto es,

$$W^B(X_p,Y_r^*)=\max_{Y_r \subset N, |Y_r|=r} W^B(X_p,Y_r).$$

Usando la denominación utilizada por Hakimi (1990), la solución del problema de B es un $(r \mid X_p)-$medianoide; la solución del problema de la firma líder es un $(r \mid p)-$centroide.

3. CASOS PARTICULARES

Ciertas funciones particulares $H_v(X_p,Y_r)$ en el modelo propuesto en la sección anterior dan lugar a las reglas de elección siguientes.

- Regla de elección binaria

Si $\underline{\alpha}_v = \overline{\alpha}_v = 0$ para todo nodo v, considerando $h_v(Y_r) = \mu, 0 \le \mu \le 1,$ se obtiene la regla de elección binaria. Si $\mu = 0$ resulta la regla binaria orientada a la firma A (el líder), si $\mu = 1$ se obtiene la regla binaria orientada a la firma B (el seguidor), si $\mu = \frac{1}{2}$ cada firma capta la mitad de la demanda en el nodo v cuando $d(v, Y_r) = d(v, X_p)$.

- Regla parcialmente binaria

Si para todo nodo v se tiene $\underline{\alpha}_v = d(v, X_p)$, $\overline{\alpha}_v = \max\{d(x, y) : x, y \in N\}$ y $h_v(Y_r) = \dfrac{f_v(d(v, X_p))}{f_v(d(v, X_p)) + f_v(d(v, Y_r))}$, siendo f_v una función creciente con $f_v(0) \ge 0$ y $h_v(Y_r) = \frac{1}{2}$ si $d(v, Y_r) = d(v, X_p) = 0,$ se obtiene la regla parcialmente binaria.

- La regla de elección proporcional

Si para todo nodo v se tiene $\underline{\alpha}_v = d(v, X_p)$, $\overline{\alpha}_v = \max\{d(x, y) : x, y \in N\}$ y $h_v(Y_r) = \dfrac{\sum_{1 \le i \le r} \frac{1}{f_v(d(v, y_i))}}{\sum_{1 \le i \le p} \frac{1}{f_v(d(v, x_i))} + \sum_{1 \le i \le r} \frac{1}{f_v(d(v, y_i))}}$, siendo f_v una función creciente con $f_v(0) > 0,$ se obtiene la regla de elección proporcional.

- Regla de elección lineal a trozos

Esta regla se obtiene cuando para todo nodo v, $h_v(Y_r) = h_v(d(v, Y_r))$ siendo h_v una función lineal a trozos.

Presentamos un ejemplo de la función de demanda captada para diferentes reglas de elección considerando para ello la red representada en la figura 1. La longitud de las aristas de la red figura junto a éstas, se considera *p=1* y $X_1 = \{v_1\}$. La tabla 1 contiene las distancias entre los nodos.

Figura 1

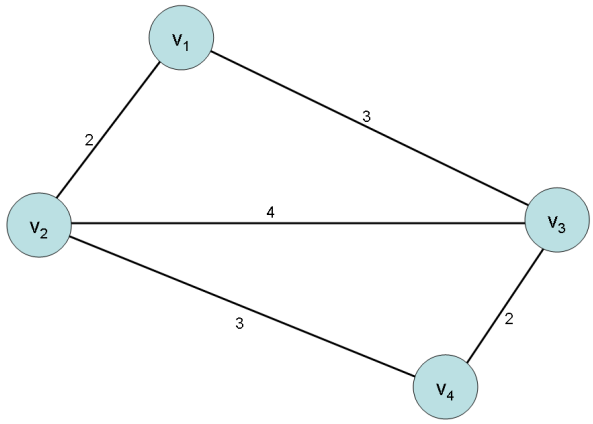


Tabla 1

Nodos	v_1	v_2	v_3	v_4
v_1	0	2	3	5
v_2	2	0	4	3
v_3	3	4	0	2
v_4	5	3	2	0

Las figuras 2 y 3 muestran la demanda captada por la firma B en los puntos de la arista $[v_2, v_3]$ para las reglas binaria (orientada al líder) y parcialmente binaria. Las figuras 4, 5 y 6, muestran la demanda captada en la misma arista cuando se aplica la regla de elección incorporando umbrales, con $\underline{\alpha}_v, \overline{\alpha}_v \in \{1,2,3\}, \forall v$. Cada punto en la arista viene dado por la distancia entre el punto y el nodo v_2.

Figura 2

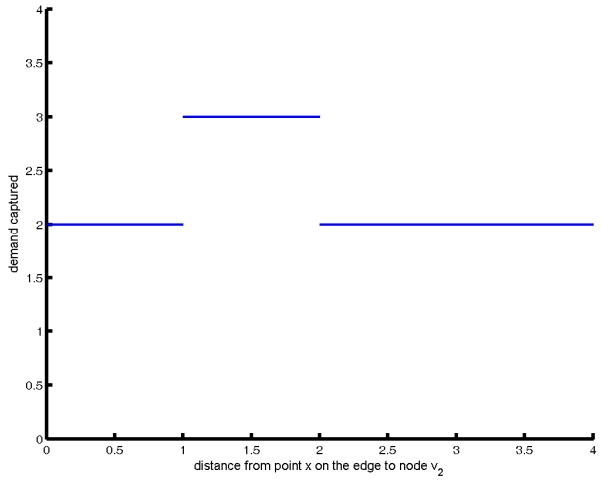


Figura 3

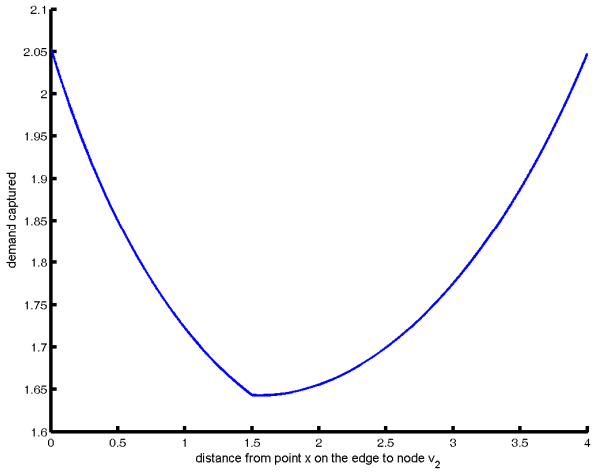


Figura 4. $\underline{\alpha}_v = 1, \overline{\alpha}_v \in \{1,2,3\}, \forall v.$

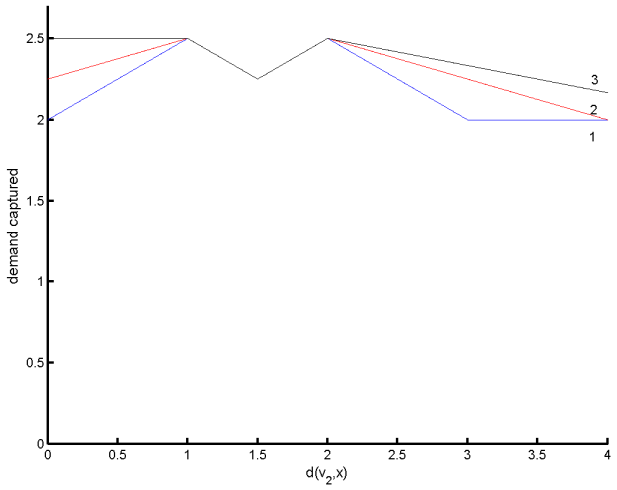


Figura 5. $\underline{\alpha}_v = 2, \overline{\alpha}_v \in \{1,2,3\}, \forall v.$

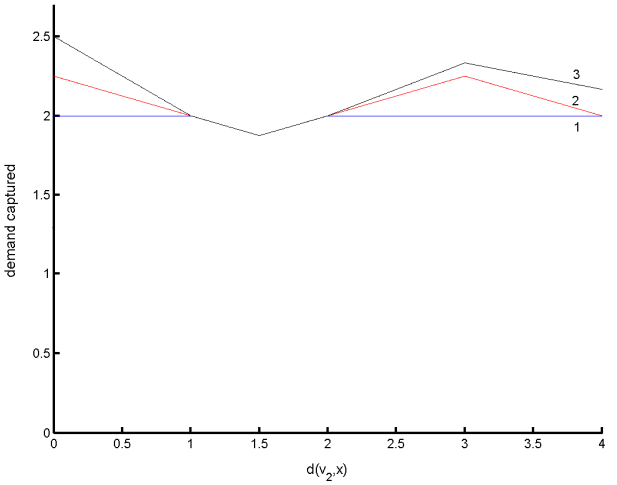
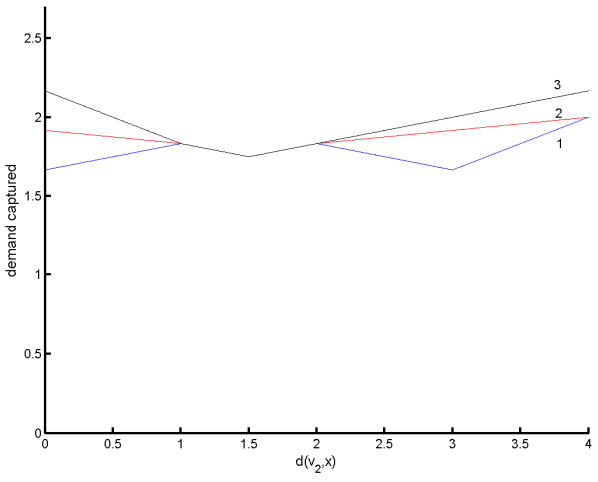


Figura 6. $\underline{\alpha}_v = 3, \overline{\alpha}_v \in \{1,2,3\}, \forall v.$



4. CONVERSIÓN DEL PROBLEMA DE LOCALIZACIÓN EN REDES EN UN PROBLEMA DISCRETO

Consideramos el problema de la firma seguidora, la firma B, y presentamos unos resultados que permiten transformar el problema de localización en redes en un problema donde el conjunto de candidatos para abrir un centro de la firma B es un conjunto finito. Existen varios resultados en la literatura sobre la transformación del problema del seguidor en redes en un problema discreto mediante la aplicación de teoremas de optimalidad en los nodos de la red o en un conjunto finito de puntos caracterizados de alguna forma. Así, para la regla binaria, puede obtenerse una solución óptima del problema del seguidor en un conjunto formado por vértices y puntos denominados *isodistantes* (Meggido et al. (1983), Peeters y Plastria (1997), Suárez Vega et al. (2004)); para la regla parcialmente binaria y la proporcional, si las funciones f_v cumplen ciertas condiciones, se demuestra que existe una solución del problema formada por vértices (Hakimi (1990)).

Los resultados siguientes garantizan la existencia de una solución del problema de localización del seguidor $Y_r^* = \left\{ y_i^* \right\}_{i=1}^r$ donde, para cada *i*, el punto y_i^* es un nodo de la red o satisface la propiedad introducida en la proposición 1.

Proposición 1.

Sea X_p el conjunto de localizaciones de los centros existentes pertenecientes a la firma A. Para todo nodo *v*, la demanda en *v* captada por la firma seguidora B con centros en las localizaciones $Y_r = \left\{ y_i \right\}_{i=1}^r$ es $w^B(v) = w(v) H_v\left(d(v, y_1), \ldots, d(v, y_r)\right)$ donde H_v es una función real, no negativa, continua y decreciente, definida en el conjunto $\left\{ \vec{z} = (z_1, \ldots, z_r) \in \mathbb{R}^r : z_i \geq 0 \, \forall i \right\}$. Para cada nodo *v*, $H_v(\vec{z})$ es una función convexa a trozos con respecto de cada coordenada, esto es, la función

$$H_{vi/z_1,\ldots,z_{i-1},z_{i+1},\ldots z_r}(z) = H_v(z_1, \ldots, z_{i-1}, z, z_{i+1}, \ldots, z_r)$$

es convexa a trozos.

Entonces existe una solución óptima del problema del seguidor, $Y_r^* = \left\{ y_i^* \right\}_{i=1}^r$, tal que para cada *i*, y_i^* es un nodo o $y_i^* = x$ satisface la siguiente propiedad:

Propiedad 1. Existen $u, u' \in V$ tales que $d(u, x) = z_{u'i}$ y $z_{u'i}$ es un *punto de corte* de la función $H_{u'i/y_1^*,\ldots,y_{i-1}^*,y_{i+1}^*,\ldots y_r^*}(z)$ (punto donde esta función no es convexa y que limita alguno de los trozos donde la función es convexa).

Corolario 1. Sea X_p el conjunto de localizaciones de los centros existentes pertenecientes a la firma A. Para todo nodo *v*, la demanda en *v* captada por la firma seguidora B con centros en las localizaciones $Y_r = \left\{ y_i \right\}_{i=1}^r$ es $w^B(v) = w(v) H_v\left(d(v, y_1), \ldots, d(v, y_r)\right)$ donde H_v es una función real no negativa, continua y decreciente, definida en el conjunto $\left\{ \vec{z} = (z_1, \ldots, z_r) \in \mathbb{R}^r : z_i \geq 0 \, \forall i \right\}$. Si, para cada nodo *v*, $H_v(\vec{z})$ es una función convexa, entonces existe una solución óptima del problema del seguidor, $Y_r^* = \left\{ y_i^* \right\}_{i=1}^r$, donde cada localización es un nodo de la red, es decir, $Y_r^* \subset V$.

Corolario 2. En las condiciones de la proposición 1, si $r=1$, existe una solución óptima del problema del seguidor formada por puntos del conjunto $F=V\cup C$ donde $C=\{x\in N:\exists u,u'\in V$ y $d(x,u)=z$ donde z es un punto de corte de $H_{u'}(z)\}$.

5. CONCLUSIONES Y EXTENSIONES

En este trabajo se propone un modelo de localización donde la regla de elección del consumidor está definida utilizando unos umbrales de preferencia. Para ciertos valores de estos umbrales se obtienen reglas de uso frecuente en la literatura como las reglas de elección binaria, parcialmente binaria y proporcional, de manera que puede considerarse que se ha creado un marco general que contiene estas reglas de decisión como casos particulares. Esta formulación permite formas de decisión intermedias, entre la regla de decisión binaria y la proporcional, permitiendo también introducir mecanismos para representar la información imprecisa sobre el comportamiento del individuo. En esta línea estaría la aplicación de los conjuntos difusos a la localización competitiva.

Los resultados que garantizan la existencia de una solución óptima del problema de localización en redes en un conjunto finito de puntos de la red permiten resolver el problema en redes mediante procedimientos usados en problemas discretos. En este trabajo se ha obtenido un resultado de esta naturaleza que permitiría resolver el problema del seguidor determinando este conjunto finito de puntos candidatos y aplicando un procedimiento de solución para el caso discreto. Así, podrían aplicarse algoritmos exactos o bien búsquedas de tipo heurístico tales como los algoritmos tabú o de búsqueda mediante entornos variables (véase por ejemplo Benati y Laporte (1994), y Campos et al. (2009), entre otros).

Ciertas estructuras de red como los árboles satisfacen propiedades particulares que podrían aprovecharse para resolver estos problemas, estos escenarios así como los casos más simples, por ejemplo el caso en que la firma seguidora abre un único centro, y el análisis en espacios discretos, pueden ser puntos de partida para investigar las reglas de elección propuestas. Otro caso de estudio es aquel donde el espacio es continuo, para este caso la aplicación de funciones de decisión con umbrales o enfoques de carácter difuso pueden encontrarse en algunos trabajos como Devletoglou (1965) y Osumi y Shiode (2003).

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por el Gobierno Español y FEDER mediante los proyectos MTM2005-09362-C03-03 y ECO2008-05589.

REFERENCIAS

BENATI, S, G. LAPORTE (1994), "Tabu search algorithms for the (r|X_p)-medianoid and (r|p)-centroid problems. Loc. Science 2(4), pp 193-204.

BENATI, S., P. HANSEN (2000), "Competitive Location with Random Utilities: Problem Formulation and Algorithms". *Cahiers du GERAD*

CAMPOS-RODRÍGUEZ, C., J.A. MORENO-PÉREZ, D.R. SANTOS PEÑATE (2009), "An exact procedure and LP formulation for the leader-follower location problem", TOP (por aparecer).

CRAIG, S., A. GHOSH, S. MCLAFFERTY (1984), "Models of the retail location process: A review", Journal of Retailing 60(1), pp 5-36.

DEVLETOGLOU, N.E. (1965), "A dissenting view of duopoly and spatial competition", Economica 32(126), pp 140-160.

DREZNER, T. (1995), "Competitive facility location in the plane", In Facility Location. A survey of applications and methods, Z. Drezner (Ed.), Springer, Berlin, pp 285-300.

EISELT H.A., G. LAPORTE (1989), "Competitive spatial models", European Journal of Operational Research 39, pp 231-242.

EISELT, H.A. G. LAPORTE (1996), "Sequential location problems", European Journal of Operational Research 96, pp 217-231.

FRIESZ, T.L., T. MILLER, R.L. TOBIN (1988), "Competitive network facility location models: A survey", Papers of the Regional Science Association 65, pp 47-57.

HAKIMI, S.L. (1990), "Location with spatial interactions: Competitive locations and games", in Discrete Location Theory, Mirchandani, P.B. and R.L. Francis (eds.), John Wiley & Sons, New York, pp 439-478.

MEGGIDO, N., E. ZEMEL, S.L. HAKIMI (1983), "The maximum coverage location problem", SIAM J Algebr Discrete Methods 4, pp 253-261.

OSUMI, S., S. SHIODE (2003), "Competitive facility location problem with fuzzy relative distance", Journal of Nonlinear Convex Analysis 4(1), pp141-149.

PEETERS P.H., F. PLASTRIA (1998), "Discretization results for the Huff and Pareto-Huff competitive location models on networks", TOP 6(2), pp 247-260.

PLASTRIA, F. (2001), "Static competitive facility location: An overview of optimisation approaches", European Journal of Operational Research 129, pp 461-470.

SANTOS-PEÑATE, D.R., R. SUÁREZ-VEGA, P. DORTA-GONZÁLEZ (2007), "The leader-follower location model, Networks and Spatial Economics 7, pp 5-61.

SPAULDING, B.D., R.G. CROMLEY (2007), "Integrating the maximum capture problem into a GIS framework", Journal of Geographical Systems. Geographical Information, Analysis, Theory and Decision 9(3), pp 267-288.

TILL, J-C. (1992), "Competitive strategies for multi-establishment firms", Economic Geography 68(3), pp 290-309.

TILL, J-C. (2000), "Network competition and branch differentiation with consumer heterogeneity", The Annals of Regional Science 34, pp 451-468.

VIABILIDAD INTEGRAL DE PROYECTOS DE INVERSIÓN PÚBLICA (I). VALORACIÓN ECONÓMICA DE LOS ASPECTOS SOCIALES

JOSÉ MARÍA MORENO JIMÉNEZ
E-mail: moreno@unizar.es
CARLOS GÓMEZ BAHILLO
E-mail: cgomez@unizar.es
JAIME SANAÚ VILLARROYA
E-mail: jsanau@unizar.es
Facultad de Económicas
Universidad de Zaragoza

Resumen

La evaluación de proyectos públicos se ha basado, casi exclusivamente, en la consideración de los aspectos económicos evaluados a través del valor actual neto o de la tasa interna de retorno. En los últimos años, la valoración económica de los proyectos se está completando con su valoración social y ambiental. Este trabajo propone una nueva metodología para estudiar la viabilidad integral de los proyectos de inversión pública que contempla los valores de uso (directo, indirecto y potencial) y los de no uso (existencia y legado). A partir de la estimación monetaria del valor de uso directo, obtenida mediante el modelo *input-output*, y la importancia relativa de los cinco valores, obtenida –como se verá en un próximo trabajo– mediante el Proceso Analítico Jerárquico (AHP), se determina el valor integral. Con la metodología se ha valorado un proyecto en Aragón, si bien -por razones de confidencialidad- se presentan exclusivamente algunos resultados retocados ligeramente.

Palabras Clave: Inversión pública, viabilidad integral, modelo *input-output*, Proceso Analítico Jerárquico (AHP).
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

The evaluation of public projects has been nearly exclusively based on economic aspects measured either through the net present value or through the internal return rate. The economic evaluation of projects has been recently complemented with their social and environmental evaluation. This paper proposes a new methodology to evaluate the integral viability of public investment projects, which combines the use values (direct, indirect and potential) and the non-use values (existence and legacy). The integral value will be determined from the monetary estimation of the direct use value, obtained by means of the input-output model, and the relative importance of the five values, obtained through the analytic hierarchy process (AHP). The proposed methodology has been applied to a real problem in Aragón, although, for reasons of confidentiality, only some slightly modified results are presented.

Keywords: Public investment, integral viability, input-output model, Analytic Hierarchy Process (AHP).
Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUCCIÓN

Con frecuencia, las administraciones públicas deben tomar decisiones sobre proyectos de inversión que si bien se localizan en un punto concreto, generan elevadas externalidades (efectos *spillovers*) en las economías de su entorno.

En tal caso, es conveniente completar las técnicas tradicionales de análisis de inversión centradas en los costes y beneficios más relevantes desde el punto de vista del inversor (VAN o

TIR, por ejemplo) y, por ende, en la viabilidad económico-financiera del proyecto, con un análisis de la viabilidad social e integral de la actuación pública considerada.

En este trabajo, como paso previo para determinar la *viabilidad integral* del proyecto público en estudio, se propone analizar su viabilidad social utilizando dos de los indicadores habituales en la literatura empírica: la generación de empleo y la cifra de negocios, calculados ambos a partir del *modelo Input-Output* (MIO).

El estudio incluye, asimismo, una nueva metodología que junto a los *valores de uso directo* anteriormente estimados tiene en cuenta otros aspectos asociados a un proyecto, de carácter más intangible, capturados a través de los *valores de uso indirecto*, de *uso potencial*, *existencia* y *legado* (Moreno *et al.*, 2009).

La viabilidad integral del proyecto se obtiene combinando los valores de uso (directo, indirecto y potencial) con los valores de no uso (existencia y legado), a través del proceso analítico jerárquico (AHP), una de las técnicas multicriterio más conocidas.

El trabajo se estructura de la siguiente forma. En el apartado segundo se sintetiza la metodología propuesta. En el tercero, se presentan los resultados obtenidos al calcular la viabilidad económico-financiera, así como la viabilidad social e integral de un proyecto de inversión pública en Aragón. El último de los epígrafes sintetiza el estudio y subraya las conclusiones más relevantes.

2. ANTECEDENTES

En la evaluación de proyectos, la consideración de aspectos sociales y ambientales –muchos de ellos intangibles– junto a los tradicionales aspectos económicos, obliga a utilizar nuevas propuestas metodológicas que, entre otras cosas, permitan combinar lo tangible y lo intangible y puedan emplearse en contextos con múltiples actores.

La metodología propuesta (Sección 3) consta de tres etapas. En la primera, siguiendo el enfoque tradicional, se evalúa la *viabilidad económica* del proyecto por cualquiera de los procedimientos habitualmente seguidos (VAN o TIR, por ejemplo). En la segunda, se calcula la *valoración económica de los aspectos sociales*. Para ello, se obtienen el empleo generado y la cifra de negocio mediante el modelo *input-output*. Esta estimación económica, correspondiente al valor de uso directo asociado al proyecto, se utiliza en la tercera etapa de la metodología como elemento pivote a partir del cual estimar económicamente los demás valores (uso y no uso) considerados al analizar la *viabilidad integral* del proyecto (Moreno-Jiménez y otros, 2009).

2.1. Uso del modelo *input-output* para el cálculo del empleo y la cifra de negocio

Para el cálculo del empleo y la cifra de negocios que puede generar un proyecto de inversión con elevadas externalidades es recomendable utilizar el *modelo input-output* (MIO) que, como es sabido, sigue las adquisiciones de *inputs* intermedios entre los diferentes sectores de una economía, utilizando la técnica matemática de inversión de matrices. Hoy en día, el MIO distingue tres grandes tablas: la de origen, la de destino y la simétrica. Esta última es una reordenación de los datos contenidos en las tablas de origen y destino.

En las tablas de origen y destino se incluyen otro tipo de informaciones que sirven para relacionar la valoración a precios básicos con la valoración a precios de adquisición: son las tablas de márgenes de distribución y de impuestos y subvenciones. También se incorpora información que diferencia el origen interior e importado de los productos.

En las tablas de destino y simétrica, a cada rama le corresponde una columna y una fila, leyéndose en columnas los insumos (*inputs* o entradas) para el proceso productivo de la rama a la que corresponde cada columna y, en filas, los *outputs* (salidas o destinos) de las producciones de cada rama (bloque de consumos intermedios).

Por definición, el total de consumos intermedios utilizados por todas las ramas (suma de los totales de las columnas) coincide con el total de salidas (suma por filas) de productos para uso intermedio del resto de ramas. Sin embargo, esta igualdad no tiene por qué producirse rama a rama.

Los factores de producción primarios, trabajo y capital, se remuneran, respectivamente, con el salario y el excedente bruto de explotación. Las rentas generadas integran el valor añadido de cada rama (bloque de *inputs* primarios). Por otra parte, la suma de los consumos intermedios y la del valor añadido en de cada una de las ramas da la producción total de toda la economía.

Las importaciones realizadas para los procesos productivos (como consumos intermedios) o por los consumidores finales se recogen en el MIO añadiendo en la parte inferior de la columna correspondiente a cada rama una fila de importaciones similares a los productos de dicha rama. Por filas, además de la parte de la producción de cada rama que es utilizada por ella y por el resto de ramas como consumo intermedio en sus respectivos procesos productivos, otra parte de la producción de cada rama se destina a usos finales, es decir, a consumo individual, consumo colectivo, formación bruta de capital y exportaciones (bloque de demandas finales).

Para exponer el funcionamiento del modelo puede usarse la siguiente notación:

x_{ij} = Cantidad registrada en la casilla de cruce de la fila i y la columna j del bloque de consumos intermedios del MIO (utilización que la rama j hace de los productos de la rama i).

X_i = Producción de la rama *i* (suma de los consumos intermedios de la rama más el valor añadido).

VA_i = Valor añadido por la rama *i* = remuneración de asalariados (W_i) + excedente bruto de explotación (EBE_i).

D = Demanda final = Consumo individual ($CInd$) + consumo colectivo ($CCol$) + formación bruta de capital (FBK) + exportaciones (E).

M = Importaciones.

REC = Oferta total de recursos = Producción (*X*) + importaciones (*M*).

ET = total de empleos, *output* o destinos = Consumos intermedios (CI) + Demanda final (D).

Los estudios basados en el MIO calculan sistemáticamente algunas relaciones cuya utilidad es manifiesta, si las mismas se consideran relaciones estructurales de una economía. De un modo sumario, puede representarse una estructura de relaciones estructurales entre las ramas de una economía calculando, para cada rama, la proporción de cada *input* en su producción. Tales proporciones se denominan coeficientes de *input* y pueden representarse en una tabla o matriz en la que cada elemento es el correspondiente al que ocupa el mismo lugar en la tabla original. Los coeficientes técnicos son un tipo especial de coeficientes de *input* que expresan la utilización que cualquier rama hace de productos de otra por unidad de producción $a_{ij} = x_{ij} / X_j$.

El cálculo de los coeficientes técnicos puede hacerse dividiendo cada x_{ij} por la X_j correspondiente (cada elemento del bloque de transacciones *interindustriales* por la producción de la rama correspondiente). Así, por ejemplo, $a_{12} = x_{12}/X_2$ es la utilización que la rama *2* hace de productos de la rama *1* por unidad de producción.

En forma matricial sería $A = Z \hat{} X^{-1}$, siendo *A* la matriz de coeficientes técnicos resultante de efectuar la multiplicación de las matrices de las x_{ij} (*Z*) por la matriz diagonal de las recíprocas de las producciones de cada rama (X^{-1}). La suma por columnas de los elementos de la matriz de coeficientes técnicos (*A*), $\sum_i a_{ij}$, dará el total de consumos intermedios que la rama *j* efectúa para producir una unidad.

Aunque la denominación de los coeficientes alude a la expresión de ciertas características tecnológicas de los procesos productivos. Sin embargo, en la práctica los coeficientes que se calculan son monetarios, esto es, el MIO proporciona el valor de los productos de una rama utilizados para obtener una unidad de producto en la rama utilizadora de los mismos.

La relación estará influida por los precios y hay que tener en cuenta la alteración de éstos a lo largo del tiempo para extraer conclusiones apropiadas respecto a la utilización de *inputs* así como a su sustitución o ahorro.

En la confección del MIO se introducen supuestos como: (a) Todos los productos de una rama se han generado con la misma estructura de inputs y (b) Todos los productos de un grupo

de productos tienen la misma estructura de inputs cualquiera que sea la rama que los produce. La opción de una u otra hipótesis dependerá de los datos y conocimientos disponibles.

Ha de resaltarse que los ingresos de las instituciones financieras derivan de los intereses de sus operaciones de activo y de los ingresos por la prestación de determinados servicios (cobro de recibos, comisiones de cambio...) son, desde el punto de vista de la Contabilidad Nacional, sólo los únicos servicios prestados por estos agentes que forman parte de la producción. Ello lleva a no incluir entre las actividades productivas operaciones importantes como la intermediación financiera y, a su vez, hay una importante generación de consumos intermedios, salarios y empleos del sector financiero que no responden a una actividad productiva (o expresado, de otra forma, la producción de servicios financieros en sentido estricto sería insuficiente para pagar por la utilización de tales *inputs*).

La solución para superar esta anomalía es la imputación de una cantidad global para la intermediación financiera (para la recogida, transformación y distribución de disponibilidades financieras). Tal cantidad resulta de la diferencia que las instituciones financieras obtienen de las rentas de propiedad y de la empresa recibidas por intereses y los intereses pagados por sus operaciones de pasivo y se denomina servicios de intermediación financiera medidos indirectamente (*SIFMI*).

En consecuencia, y dado que el MIO pretende medir relaciones técnicas, se crea una rama ficticia –servicios de intermediación financiera medidos indirectamente- que consume tal producción, como un consumo intermedio global de la economía y cuya producción es cero.

Con estos antecedentes, y dado que las filas de la tabla representan los destinos de la producción de cada rama al resto de ramas y a la demanda final (rama a rama coinciden con el valor de su producción X_i), dos filas de una tabla con dos ramas pueden expresarse:

$$x_{11} + x_{12} + D_1 = X_1$$
$$x_{21} + x_{22} + D_2 = X_2 \qquad (1)$$

o como:

$$a_{11} X_1 + a_{12} X_2 + D_1 = X_1$$
$$a_{21} X_1 + a_{22} X_2 + D_2 = X_2 \qquad (2)$$

que en expresión matricial puede representarse como

$$AX + D = X \qquad (3)$$

siendo $A = (a_{ij})$ la matriz de los coeficientes técnicos; X el vector columna de las producciones a precios básicos (X_i a la producción a precios básicos de i); D el vector columna de las demandas finales (D_i a lo que i destina a la demanda final) y x_{ij} a los destinos que desde i van a j (usos que j hace de los productos de i). La rama i destina $\sum_j x_{ij}$ a usos intermedios y D_i a destinos finales.

Sabiendo que a_{ij} es un dato y suponiendo que la demanda final se determina de forma exógena, quedan dos incógnitas y dos ecuaciones. El modelo de demanda relaciona las demandas finales autónomas y los niveles de producción necesarios para satisfacerlas. Si se realizan enunciados alternativos sobre la demanda final a conseguir, pueden conocerse las producciones necesarias de cada rama para satisfacer dicha demanda final. Lógicamente, cada rama ha de producir tal objetivo de demanda final más los productos intermedios que de ella requieran el resto de ramas para producir también una unidad.

Cuando el número de incógnitas es de cierto tamaño, el problema puede resolverse mediante el algoritmo conocido como el método de la inversa de la matriz de Leontief[1]:

$$X = (1 - A)^{-1} D$$

[1] El punto de partida de Leontief es el esquema de equilibrio general walrasiano. Walras elaboró un sistema de ecuaciones representativo del equilibrio general que se produce en los mercados de bienes y de factores. La competencia produciría unos precios de equilibrio. Era un sistema adecuado para conocer cómo se establece un sistema de precios mediante las demandas competitivas que las actividades hacen de factores de producción, así como la capacidad de sustitución que de sus productos existe en el consumo final. Leontief restringe y simplifica el esquema de Walras, eliminando los efectos-precio de la sustitución de *inputs* (no hay factores limitados y los coeficientes de producción son fijos por lo que los *inputs* intermedios no pueden reemplazarse). Las hipótesis centrales del modelo *input-output* son:

a) Cada mercancía (o grupo) es suministrada únicamente por una rama (hipótesis de homogeneidad).

b) Los *inputs* utilizados son función del nivel de producción de cada rama. Esta hipótesis excluye la existencia de economías de escala (hipótesis de proporcionalidad).

c) El efecto total de llevar a cabo varios tipos de producción constituye la suma de los efectos separados, no hay economías ni deseconomías externas (hipótesis de aditividad).

siendo X el vector columna de las producciones a precios básicos de cada rama (incógnitas del problema), $(1- A)^{-1}$ la inversa de la matriz de Leontief $(1- A)$ y D, el vector columna de las demandas finales.

Obtenida la inversa se pueden simular diversas situaciones como:

a) Producción de cada rama para que se produzca una unidad para la demanda final (consumo, exportaciones...).

b) Puede calcularse el empleo a que dará lugar un determinado nivel de demanda (o de cualquiera de sus componentes) como el generado a través de los consumos de productos intermedios de otras ramas.

c) Las necesidades de importaciones intermedias.

La contribución de una rama a la creación de empleo viene determinada por la importancia de sus productos en la demanda final y por los requerimientos que otras ramas hacen de sus productos como *inputs* intermedios. Puede realizarse esta estimación calculando la creación de empleo en cada rama cuando la demanda final para todas las ramas aumenta en una unidad. Puede hacerse mediante la expresión:

$L = \hat{l}\,(1-A)^{-1}\,D$

en la que $\hat{l}$ es la matriz diagonalizada de los coeficientes directos de empleo, D es el vector columna de las demandas finales y las L son los requerimientos de empleo de cada rama necesarios para satisfacer demandas finales para todas las ramas.

Generalizando para cualquier componente del valor añadido, distinto del trabajo, llamando F_V a los coeficientes del valor añadido y, en particular, F_l al coeficiente de trabajo por unidad de producto (N_j/X_j) y F_k al coeficiente de capital por unidad de producción (N_k/X_k), si se desea conocer la generación de empleo que llevará consigo un incremento de la demanda final se procede de la siguiente forma

$L = \hat{F}_l\,(1-A)^{-1}\,D_F$

siendo D_F el vector de demanda final y $\hat{F}_l$ la matriz diagonal de los coeficientes directos del trabajo.

Calculado el empleo, la cifra de negocios generada por el aumento de la demanda final que genere un proyecto de inversión puede aproximarse bien por el propio MIO bien introduciendo supuestos sobre la productividad aparente del factor trabajo de los empleos creados o mantenidos (o sobre el VAB o PIB que estos empleados vayan a producir).

2.2. Valores de Uso y de No Uso en Proyectos de Inversión Pública

A la hora de evaluar la viabilidad integral de los proyectos de inversión pública se sugiere considerar -junto a los tradicionales aspectos económicos- los sociales y los ambientales. Todos estos aspectos se han estructurado, en la metodología propuesta, conforme a los cinco valores considerados en valoración ambiental por Barbier y otros (1997) y Moreno-Jiménez y otros (2009): Valores de Uso (Directo, Indirecto y Potencial) y Valores de No Uso (Existencia y Legado).

En la bibliografía económica y valorativa, se pueden encontrar diferentes tipos, o clases de valor, como son: *valor de mercado*, *valor probable de mercado, valor en uso*, *valor de sustitución o renovación*, *valor objetivo,* etc. En todos estos casos, la definición de valor hace referencia a una actividad de mercado. Sin embargo, la finalidad de la valoración económica de un activo no es, en general, obtener un precio de mercado, ya que no se puede considerar como una mercancía, sino contar con un indicador de la importancia que el "producto" tiene en el bienestar de la sociedad. Se trata, en definitiva, de valorar lo que supone para el bienestar social la existencia de un determinado "bien" o "entorno", y obtener así un valor que permita compararlo con otros activos.

Barbier y otros (1997) distinguen cinco tipos de valores al hablar del *valor económico total* (VET): uso directo, uso indirecto, opción/cuasiopción, existencia y legado. Adaptados al contexto valorativo considerado (Moreno Jiménez y otros, 2009), estos componentes del VET se interpretan, en el caso del proyecto en estudio (denominado en lo sucesivo como "CV"), como:

- *Valor de uso directo* (VUD): valor que tienen los bienes y servicios asociados al proyecto en la satisfacción de las necesidades humanas y que se valoran por el propio mercado (ingresos de la explotación o cifra de negocios).
- *Valor de uso indirecto* (VUI): valor que tienen los bienes y servicios por sus usos no retribuidos, a veces difícilmente observables y cuantificables, que no se valoran directamente por el mercado pero que derivan de las funciones que desempeñan (mantenimiento y fijación de la población al territorio, potencial formativo, capacidad para aflorar el patrimonio cultural, creación de una seña identitaria de la zona, el disfrute de zonas recreativas, etc.).
- *Valor de uso potencial* (VUP), denominado de opción/cuasiopción por Barbier: refleja el valor que una persona da a la posibilidad o garantía de disponer en el futuro de unos bienes y servicios, cuyos efectos directos e indirectos son desconocidos o difíciles de establecer y prever con los conocimientos actuales.
- *Valor de existencia* (VE): valor de un activo por el hecho de ser un recurso esencial para la conservación y desarrollo de la vida en el entorno, de diversas especies animales, microhábitats florísticos, sistemas naturales únicos, valores culturales, paisaje, etc.
- *Valor de legado* (VL): valor de legar los beneficios del activo a futuras generaciones. Es decir, es el valor que se le asigna por el hecho de que las futuras generaciones tengan la oportunidad de usarlo y disfrutarlo.

De todas los componentes del VET, el único que puede ser valorado directamente en términos monetarios es el VUD. Por eso, se empleará en la metodología propuesta como el elemento "pivote" a partir del cual estimar las valoraciones monetarias de los restantes componentes. La utilización de esta vía indirecta a la hora de estimar el resto de los valores es necesaria por la inexistencia de un mercado para los mismos.

En valoración de activos ambientales, cuyo marco se utiliza como referencia, existen una serie de métodos utilizados en gran número de trabajos, que partiendo de diferentes enfoques pretenden llegar a determinar un valor monetario, bien del activo considerado o de los servicios que ese activo produce (en la Tabla 1 aparecen los métodos más utilizados y el valor que pretenden encontrar).

Tabla 1. Valores de un bien (Moreno-Jiménez y otros, 2009)

VALOR ECONÓMICO TOTAL		Método a utilizar
Valores de Uso	Valor de Uso Directo (VUD)	Precio del Mercado
		Flujo de Caja
		Renta de los Factores
		Coste Reemplazamiento
		Coste del Viaje
	Valor de Uso Indirecto (VUI)	Costes Evitados
		Renta de los Factores
		Precios Hedónicos
		Coste del Viaje
	Valor de Uso Potencial (VUP)	Valoración Contingente
		Precios Hedónicos
Valores de NO Uso	Valores de existencia (VE)	Valoración Contingente
	Valores de legado (VL)	Valoración Contingente

Fuente: Adaptado de Barbier y otros (1997).

Para calcular el VUD, al ser éste un valor que el mercado refleja, cualquiera de los métodos utilizados en la valoración de bienes económicos puede servir. La elección del mismo vendrá supeditada a la información de partida. En la determinación del resto de valores, al no estar éstos conectados con el mercado, se suele recurrir a metodologías indirectas (costos evitados, coste del viaje, precios hedónicos, etc.) en el caso del VUI y a metodologías basadas principalmente en la determinación de la disposición a pagar o a percibir de los ciudadanos afectados para el resto de valores (VUP, VE y VL).

Desde un punto de vista operativo, el estudio de la viabilidad de cualquier proyecto con elevadas externalidades incluye diferentes aspectos que corresponden a los tres grandes criterios considerados: económicos, sociales y ambientales, y que a su vez reflejan (Frey, 2000; Moreno-Jiménez y otros, 2009) los cinco valores anteriormente citados (VUD, VUI, VUP, VE y VL).

Identificados los aspectos económicos, sociales y ambientales que incluye cada uno de los cinco valores considerados, y los actores implicados en el proceso de toma de decisiones, se calcula la importancia relativa que estos actores asignan a cada uno de los cinco componentes del VET. Para ello, se propone utilizar una de las técnicas multicriterio más extendida en la literatura científica: el Proceso Analítico Jerárquico, AHP (Saaty, 1980; Moreno, 2002).

En lo que se refiere a la obtención del VUD, habría que contemplar los valores monetarios asociados al proyecto. Por un lado, se consideraría la inversión, la cifra de negocios y el empleo generado. Estos dos últimos indicadores (empleo y cifra de negocio) son los empleados habitualmente para analizar la viabilidad social de las actuaciones públicas y, en lo que sigue, se van a obtener utilizando el modelo *input-output*.

Obviamente, se persigue que el proyecto sea viable desde el punto de vista económico y que el flujo de ingresos futuros permita garantizar la buena marcha del negocio y el mantenimiento y renovación del mismo. No obstante, puede ocurrir que los valores de uso indirecto y potencial, así como los valores de no uso pudieran aconsejar realizar una inversión que no siendo rentable económicamente, fuera rentable social y ambientalmente.

3. VIABILIDAD ECONÓMICA Y SOCIAL DEL PROYECTO "CV"

3.1. Viabilidad Económico-Financiera del Proyecto CV

El valor actual neto (VAN) –también conocido como valor presente neto– es un procedimiento clásico de valoración de inversiones en activos fijos que proporciona una valoración financiera en el momento actual de los flujos de caja netos (diferencia entre cobros y pagos) que produce la inversión, descontándolos a una determinada tasa o tipo de interés.

Las principales ventajas del VAN son que tiene en cuenta el valor del dinero en cada momento y que es muy flexible, puesto que permite introducir cualquier criterio variable que pueda afectar a la inversión como la inflación, la incertidumbre o la fiscalidad. La fórmula empleada ha sido:

$$VAN = -A + \frac{Q_1}{(1+k_1)} + \frac{Q_2}{(1+k_1)(1+k_2)} + ... + \frac{Q_n}{(1+k_1)...(1+k_n)}$$

donde A es el valor de la inversión inicial; Q_i el valor del flujo neto anual de caja (diferencia entre cobros y pagos) a efectuar en el periodo i-ésimo; n es el número de períodos considerado y k_i la tasa de descuento.

Si el proyecto no tiene riesgo, suele tomarse como referencia el tipo de la renta fija, de tal manera que con el *VAN* se estimará si invertir en el proyecto es mejor que hacerlo en algo seguro, sin riesgo especifico. En otros casos, se utiliza el coste de oportunidad. De acuerdo con este criterio:

Si $VAN > 0$, la inversión produciría ganancias por encima de la rentabilidad exigida (k_i) y, en consecuencia, el proyecto puede aceptarse. Si $VAN < 0$, la inversión produciría pérdidas por encima de la rentabilidad exigida (k_i) y el proyecto debería rechazarse. Si $VAN = 0$, la inversión no produciría ni ganancias ni pérdidas. Dado que el proyecto no agrega valor monetario por encima de la rentabilidad exigida (k_i), la decisión debería basarse en otros criterios o factores.

Para la aplicación del VAN del proyecto CV se partió de las cifras de inversión, ingresos y gastos, transformadas en flujos de tesorería y proyectándolas en un horizonte temporal de 25 años, suponiendo un crecimiento nominal de los ingresos y gastos del 2% anual (tasa que coincide con el objetivo de inflación establecido por el Banco Central Europeo).

Los cálculos efectuados en dos escenarios diferentes, según el periodo de amortización considerado fuera mínimo o máximo, respectivamente[2]. Suponiendo que las inversiones se amortizan aplicando los coeficientes de amortización lineal establecidos en las tablas de amortización oficialmente aprobadas y que esta circunstancia da lugar a la consecuente reposición de los inmovilizados, en los dos escenarios el flujo de caja acumulado es siempre negativo, si bien la magnitud en el segundo caso es sensiblemente inferior al del primero. En consecuencia, para cualquier tasa de descuento k_i razonable –como las de la renta fija emitida por el Reino de España– el VAN de la inversión a realizar es negativo y, según este criterio, el proyecto no debería llevarse a cabo.

Hay que resaltar que para el cálculo del VAN sólo se consideran los efectos económicos directamente asociados al proyecto sin tener en cuenta los impactos de la intervención en la zona de influencia, esto es, los efectos sociales referidos cuantificados por el número de empleos que se generarán y por el valor añadido bruto o PIB que podrá crearse (que aproximan la cifra de negocio asociada).

Para abordar el análisis de la viabilidad integral de una intervención pública de envergadura han de considerarse, además del valor social de la intervención cuantificado en términos económicos (VUD), otro tipo de valores asociados a la actuación como son los valores de uso indirecto, de uso potencial, de existencia y de legado. La estimación monetaria de estos cuatro valores (VUI, VUP, VE y VL) mediante el Procesos Analítico Jerárquico se incluye en un segundo trabajo. En lo que sigue, utilizando el modelo *input-output* se calcula el VUD asociado a la intervención y no sólo al proyecto.

3.2. Viabilidad Social del Proyecto *CV*

La valoración económico-financiera del proyecto, a través del VAN o de algún otro indicador, es el procedimiento habitualmente seguido al analizar la viabilidad de las inversiones públicas. No obstante, a partir del decenio de 1990 y basado en los enfoques valorativo ambientales (Moreno-Jiménez y otros, 1999), cada vez se da más importancia a los aspectos sociales. A continuación, siguiendo el enfoque valorativo tradicional, para analizar la viabilidad social del proyecto se ha recurrido al análisis de dos de los indicadores más utilizados en la literatura científica: el empleo y la cifra de negocio, que serán estimados utilizando el MIO de la economía aragonesa.

El MIO puede utilizarse tanto para evaluar el empleo generado a corto plazo, asociado a las inversiones que conlleva el proyecto analizado, como para aproximar la ocupación que habrá, una vez que entre en funcionamiento, por la demanda final que se cree. Para evaluar los efectos de las inversiones que se pretende efectuar se partió del proceso inversor previsto y se concretó el desarrollo del mismo a lo largo de los años 2009, 2010 y 2011.

Cuantificadas las inversiones y el desarrollo temporal del proyecto, para aplicar el MIO hay que asignar las inversiones a las distintas ramas o sectores y determinar el porcentaje de esas inversiones que podría satisfacerse con producción de regional. La primera empresa no es compleja pues las ramas que se ven inicialmente impulsadas son: *Construcción de maquinaria y equipo mecánico*, *Fabricación de máquinas de oficina y equipos informáticos*, *Fabricación de maquinaria y material eléctrico*, *Fabricación de vehículos de motor*, *Fabricación de muebles*, *Construcción*, *Actividades informáticas* y *Otras actividades empresariales*.

La segunda tarea es, en cambio, más ardua puesto que las empresas regionales podrían beneficiarse de esta actuación por ser designadas como contratistas, participar en uniones temporales de empresas (*UTE*) con empresas que reciban adjudicaciones, por actuar como subcontratistas o simplemente por suministrar *inputs* intermedios o/y factores productivos a las empresas anteriores.

Ante la ausencia de información precisa, se decidió utilizar los porcentajes que figuran para la formación bruta de capital fijo de la Tabla Simétrica del MIO de la economía aragonesa (72,58%), criterio que, además, siguen estudios similares como el del Instituto Valenciano de

[2] En este caso, instalaciones técnicas, otras instalaciones y mobiliario se renovarían o actualizarían cada 20 años; maquinaria cada 18 años; patentes y marcas en 20 años; utillaje, equipos para procesos de información y otro inmovilizado material en 8 años; y las aplicaciones informáticas en 6 años.

Investigaciones Económicas (2005). Finalmente, ha de precisarse que para aplicar el modelo de demanda del MIO, los gastos a precios de adquisición se transformaron a precios básicos.

Teniendo en cuentas estas premisas, se aplicó el MIO de Aragón –concretamente la versión de Pérez y Pérez y Parra Rodríguez (2009)–, obteniendo los resultados que aparecen en la Tabla 2. Obsérvese que las inversiones a efectuar en el trienio 2009-2011 podrían llegar a ejercer un impacto inicial en la demanda de casi 14 millones de € (precios de adquisición), prácticamente concentrado en los años 2010 y 2011.

Tabla 2. Impacto potencial sobre el empleo de de las inversiones a efectuar en *CV*(€ de 2008, precios de adquisición ligeramente modificados por cuestiones de confidencialidad)

Generación de empleo	2009	2010	2011	Total
Inversiones consideradas. Impacto inicial	400.000	7.000.000	6.600.000	14.000.000
Inversiones consideradas. Impacto total	500.000	10.500.000	9.000.000	20.000.000
Empleo potencial (trabajadores equival. a tiempo completo)	**6**	**86**	**66**	**158**

Fuente: Elaboración propia a partir MIO.

El impacto inicial de las obras podría elevar la demanda de las restantes ramas productivas de la economía aragonesa, en particular en el bienio 2010-2011. Agregando los impactos directos e indirectos, las inversiones podrían aumentar la demanda total del trienio a 20 millones de €.

Con un volumen de inversión tan elevado se podría generar/mantener el equivalente a 158 puestos de trabajo a tiempo completo entre 2009 y 2011. Se trata de empleos tanto directos como indirectos, cuya creación, lógicamente, se concentrará en los años 2010 y 2011, con 86 y 66 puestos, respectivamente.

Lógicamente, el sector construcción sería el más beneficiado con 91 empleos. A considerable distancia se situarían los servicios con 38 puestos y la industria manufacturera con 25. Por último, se generaría actividad para mantener/generar 3 empleos en el sector primario y 1 en el sector energético.

Estos efectos desaparecerán al terminar la construcción por lo que –a los efectos de este trabajo– es más interesante cuantificar los efectos asociados a la explotación del negocio, ya que perdurarán en el tiempo.

La entrada en funcionamiento del proyecto puede generar importantes efectos en la actividad económica de Aragón como la mejora de su proyección nacional e internacional, la futura atracción de inversiones foráneas; la generación de nuevos flujos de turismo, la celebración de congresos; el uso de las marcas registradas … de difícil cuantificación *a priori*.

De ahí que se haya optado por evaluar exclusivamente los efectos que generará la puesta en funcionamiento a partir de la demanda turística que efectuarán sus visitantes. Dicho gasto turístico puede estimarse suponiendo que los dos tipos de visitantes considerados: regulares y rápidos (viajeros en tránsito) paguen, respectivamente, sus entradas al recinto turístico, realicen compras de recuerdos y efectúen gastos en restauración y hostelería.

Al igual que en el caso de las inversiones, para aplicar el MIO es preciso asignar el gasto turístico a las distintas ramas o sectores. En este caso, se supone que las ramas que se verán inicialmente impulsadas son: *Comercio de vehículos y carburantes; talleres de reparación*, *Comercio al por menor, Hostelería* y *Actividades recreativas, culturales y deportivas de mercado* y puede razonablemente suponerse que toda la demanda de los turistas se satisface por empresas regionales.

Una vez expresados los gastos a precios de adquisición en precios básicos se aplicó el MIO, obteniendo como principal resultado que la demanda potencial de los visitantes podría generar o mantener el equivalente a 194 empleos en dedicación a tiempo completo.

Si se supone que en 2012 los 194 empleos que potencialmente podrían crearse o mantenerse como consecuencia de la demanda turística de los visitantes generarán un VAB similar al promedio regional, la contribución de los visitantes podría aproximarse a 10.000.000 de € (unos 9.200.000 de € en 2008).

Si a esas cifras de *VAB* se añaden los impuestos netos sobre productos (estimados), el producto interior bruto (PIB) a precios de mercado originado como consecuencia de la puesta en funcionamiento del proyecto se situará por encima de los 11.000.000 de € en el año 2012 (unos 10.200.000 de € en 2008).

Estas últimas magnitudes son importantes, puesto que las Administraciones Públicas aumentarán sus ingresos como consecuencia del PIB generado. Baste en este sentido señalar, que según IGAE (2008), la presión fiscal en España se situó en 2007, en el 37,11% del PIB a precios de mercado.

Como conceptos más importantes de la misma, debe destacarse que en dicho año (2007): el 12,90% de los recursos no financieros de las Administraciones Públicas españolas procedían de los impuestos corrientes sobre la renta y el patrimonio (7,53% del Impuesto sobre la Renta de las Personas Físicas, un 4,73% del Impuesto sobre Sociedades y un 0,62% de otros impuestos); el 12,16% de las cotizaciones sociales; el 11,8% de impuestos sobre productos (6 puntos porcentuales del IVA, y el resto de impuestos sobre productos y otros impuestos sobre la producción) y un 0,48% de los impuestos sobre el capital.

En síntesis, con una inversión cercana a los 20 millones de €, el impacto directo evaluado como la cifra de negocio (ingresos) rondaría el 10% de la inversión, unos 2 millones y crearía 15 empleos directos. Si a este impacto directo se añade el impacto indirecto, estimado con la última versión disponible del MIO de la economía aragonesa, la demanda de los visitantes podría crear o mantener más de 190 empleos en los diferentes sectores de la actividad. Suponiendo que estos trabajadores generen un VAB similar al del promedio regional, su contribución a esta macromagnitud podría superar en 2012 los 10 millones de € y, expresada en términos de PIB a precios de mercado, los 11 millones de €.

Esta cifra de 11 millones de € es la considerado como VUD asociado al proyecto CV. A partir de la misma, y utilizando la importancia relativa (AHP) que los actores implicados proporcionan a los restantes valores (VUI, VUP, VE y VL) se obtendrá el valor social o institucional de la intervención y con él su valor integral.

4. CONCLUSIONES

Tradicionalmente la viabilidad de proyectos de inversión pública se ha realizado basándose en el estudio de su rentabilidad económica, evaluada a través de indicadores como el VAN o el TIR. Este tipo de análisis contemplaba, en la mayoría de los casos, sólamente los efectos económicos directos de la actuación en el entorno inmediato, sin tener en cuenta los efectos económicos en la zona de influencia ni otra serie de valores, muchos de ellos intangibles, hasta ahora no contemplados.

Para solventar estas limitaciones, este trabajo presenta una nueva metodología para analizar la viabilidad integral de los proyectos de inversión pública que contempla, basándose en los estudios de valoración ambiental (Barbier y otros, 1997; Aznar y otros, 2009; Moreno-Jiménez y otros, 2009), cinco tipos de valores: Valor de Uso Directo (VUD); Valor de Uso Indirecto (VUI); Valor de Uso Potencial (VUP); Valor de Existencia (VE) y Valor de Legado (VL). El VUD es el único para el que existe mercado y puede valorarse económicamente de forma directa. Para los otros cuatro no existe mercado y deben valorarse económicamente de forma indirecta.

La metodología consta, básicamente, de tres etapas. En la primera, conforme al enfoque tradicional, se estudia la viabilidad económica del proyecto en sentido estricto (VAN o TIR). En la segunda, se obtiene el VUD del proyecto (impacto en el entorno próximo) mediante el modelo *input-output* (esta parte es la desarrollada en este trabajo). Finalmente, en la tercera, se obtiene el valor integral del proyecto, combinando su VUD con el valor "institucional" o "social" del mismo (VUI+VUP+VE+VL) calculado mediante AHP (esta parte es objeto de otro trabajo).

La metodología propuesta se ha aplicado al estudio de la viabilidad integral de un proyecto turístico (denominado CV) basado en la recuperación del patrimonio histórico-cultural. Por cuestiones de confidencialidad, los resultados incluidos en el trabajo para ilustrar la metodología

han sido ligeramente modificados. Se espera que, en un futuro inmediato, los resultados alcanzados y sus valores reales puedan presentarse con todo tipo de detalle.

Como ideas más relevantes del estudio cabe mencionar as siguiente. El análisis de los retornos económicos de una actuación pública ha llegado a ser una de las mayores justificaciones de inversiones en la puesta en valor del patrimonio. La valoración de los retornos económicos que produce el movimiento turístico puede realizarse a través de un conjunto de indicadores, como son el número de visitantes, el indicador de empleo, o la cifra de negocios.

El número de visitantes es un dato muy significativo, pero muy sensible a numerosas variables que no dependen solamente del modo de puesta en valor (poder de compra, coste de transportes, problemas de seguridad) sino también de la tarea de dinamización realizada y su posterior explotación en el plan de *marketing* del proyecto.

El indicador de empleo permite medir el impacto local especialmente en áreas deprimidas o fuertemente sensibilizadas por el desempleo, aunque es necesario advertir la existencia de numerosos matices que obligan a ser prudentes con las conclusiones extraídas. En el caso concreto de *CV* se ha estimado, utilizando el MIO de la economía aragonesa, que podrían crearse algo más de 190 puestos de trabajos.

Respecto al tercer indicador mencionado para la valoración de los retornos económicos de una intervención turística, la cifra de negocio, conviene señalar que ésta se basa fundamentalmente en los ingresos de las entradas y de los servicios anexos. En la mayoría de los lugares patrimoniales, los ingresos de la billetería no siempre representan el coste de los servicios y funcionamiento.

Conforme a los datos proporcionados, la inversión necesaria para el proyecto *CV* sería de unos de 20 millones de euros. La cifra de negocio anual del proyecto (ingresos estimados) estaría muy próxima al 10% de la inversión (unos 2 millones de euros) y el resultado de la explotación oscilaría, según el periodo de amortización considerado, entre el –0,4% y el –4% de la inversión.

Utilizando el MIO para estimar el impacto en la economía de la zona se alcanza una cifra de negocios (VAB), incluyendo todos los conceptos y no sólo el gasto de los visitantes, de unos 10.000.000 € para el 2012 (VUD).

Asimismo, se ha estimado (MIO) que se crearán 196 empleos, cifra ésta ligeramente superior, pero muy próxima (10%), a las estimaciones realizadas en Francia por Greffé (Greffé, 2002), a la hora de calcular los empleos asociados a la gestión del patrimonio histórico cultural.

En síntesis, y a pesar de que económicamente el proyecto puede no resultar rentable para la cifra de visitantes estimada, si sólo se contempla la rentabilidad económica del mismo en sentido estricto, *CV* puede considerarse viable desde una perspectiva integral en la que se contemplen tanto los valores sociales como el impacto económico en la región y la generación de empleo a que da lugar.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado (primer autor) por los proyectos: "E-participación, Seguridad y Democratización del Conocimiento" (Ref. PM2007-034), del Gobierno de Aragón, y "Herramientas Colaborativas para la e-cognocracia" (Ref. TNI2008-06796-C04-04), del Ministerio de Ciencia e Innovación.

BIBLIOGRAFÍA

ACZEL, J.; SAATY, T.L. (1983): Procedures for synthesizing ratio judgments. *Journal of Mathematical Psychology* 27, 93-102.

AGUARÓN, J.; MORENO-JIMÉNEZ, J.M. (2003): The Geometric Consistency Index. Approximated Thresholds. *European Journal of Operational Research*; **147** (1), 137-145.

AZNAR, J., GUIJARRO, F.; MORENO-JIMÉNEZ, J.M. (2009): Mixed valuation methods: a combined AHP-GP procedure for individual and group multicriteria agriculture valuation. *Annals of Operations Research* (DOI 10.1007/s10479-009-0527-2).

BARBIER, E. B.; ACREMAN, M.; KNOWLER, D. (1997): *Economic Valuation of Wetlands: A guide for policy makers and planers*. Oficina de la Convención Ramsar, Gland, Suiza.

FREY, B.S. (2000): *Arts and Economics: Analysis and Cultural Policy*. Springer-Verlag.

INSTITUTO VALENCIANO DE INVESTIGACIONES ECONÓMICAS (2005): *Impacto económico y análisis coste-beneficio de la Ciudad de las Artes y las Ciencias de Valencia*, IVIE: mimeo.

GREFFE, X. (2002): *La valorisation économique du patrimoine*. Rapport au Département des etudes et de la prospective, Paris. Ministère de la Culture et de Communication.

MORENO JIMÉNEZ, J.M. (2002): El Proceso Analítico Jerárquico. Fundamentos, Metodología y Aplicaciones. En Caballero, R. y Fernández, G.M. *Toma de decisiones con criterios múltiples*. RECT@ Serie Monografías nº 1, 21-53.

MORENO-JIMÉNEZ, J.M.; AGUARÓN, J.; ESCOBAR, M.T.; TURÓN, A. (1999). The multicriteria procedural rationality on SISDEMA. *European Journal of Operational Research* 119(2): 388-403.

MORENO JIMÉNEZ, J.M.; AZNAR, J.; ESTRUCH, V. (2009): *Multicriteria Economic Valuation of Wetlands*. (en evaluación).

PÉREZ Y PÉREZ, L. PARRA RODRÍGUEZ, F. (2009): *Estructura productiva y actualización del marco input-output de Aragón 2005.* Fundear-Consejo Económico y Social de Aragón, Zaragoza (en prensa).

SAATY, T. (1980). *The Analytic Hierarchy Process*. New York: McGraw-Hill.

UTILIZACIÓN DE LA CONSISTENCIA EN DECISIONES CON MÚLTIPLES ACTORES EN AHP

JUAN AGUARÓN JOVEN
E-mail: aguaron@unizar.es
Mª TERESA ESCOBAR URMENETA
E-mail: mescobar@unizar.es
JOSÉ MARÍA MORENO JIMÉNEZ
E-mail: moreno@unizar.es
Facultad de Económicas
Universidad de Zaragoza

Resumen

Este trabajo aborda la utilización de la inconsistencia, evaluada mediante el Índice de Consistencia Geométrico –GCI– (Aguarón y Moreno-Jiménez, 2003), en la toma de decisiones con múltiples actores planteada en el contexto del Proceso Analítico Jerárquico (AHP). Empleando los intervalos de estabilidad en consistencia definidos en Aguarón, Escobar y Moreno-Jiménez (2003), se establece un procedimiento iterativo que permite acercar las posturas entre los diferentes actores implicados en la resolución del problema. Para ello, se identifican los valores precisos de los intervalos de estabilidad en consistencia comunes a los actores implicados, que maximizan la menor holgura de consistencia disponible para los diferentes decisores.

Palabras clave: Proceso Analítico Jerárquico (AHP), Decisión en Grupo, Consistencia, Intervalos de Estabilidad en Consistencia.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

This paper copes with the use of the inconsistency measured named the Geometric Consistency Index –GCI– (Aguarón and Moreno-Jiménez, 2003), in multiactor decision making with the Analytic Hierarchy Process (AHP). By means of the consistency stability intervals defined in Aguarón, Escobar and Moreno-Jiménez (2003), we establish an iterative procedure that allows us to approach the position of the different actors involved in the resolution process. In order to do this, we identify the precise values from the common consistency stability intervals, which maximize the minimum slack of consistency available for the different decision makers.

Keywords: Analytic Hierarchy Process (AHP), Group Decision Making, Consistency, Consistency Stability Intervals, Core of Consistency.
Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUCCIÓN

La complejidad de los problemas decisionales planteados en la conocida como sociedad del conocimiento[1], viene determinada, fundamentalmente, por la ambigüedad a la hora de establecer las alternativas y por la consideración de múltiples escenarios, actores y criterios. En este contexto, es preciso que las metodologías utilizadas en la resolución científica de los problemas complejos, permitan la incorporación a los modelos, junto a los tradicionales aspectos tangibles, de los aspectos intangibles asociados al factor humano y tengan un buen comportamiento en el

[1] Entendida ésta como un espacio para el talento, la creatividad y el ingenio del ser humano.

caso de múltiples actores, pues ésta es la situación que se presentará en la mayoría de las ocasiones.

Una de las técnicas de decisión multicriterio que mejor responden a estos dos retos (incorporación de intangibles y toma de decisiones con múltiples actores), es el Proceso Analítico Jerárquico (AHP), propuesto por Thomas L. Saaty a mediados de los setenta (Saaty, 1977, 1980; Moreno-Jiménez, 2002). Su metodología (AHP tradicional), consiste en cuatro etapas: modelización, valoración, priorización y síntesis.

No obstante, conforme al constructivismo cognitivo característico de la racionalidad procedimental multicriterio (Moreno-Jiménez y otros, 1999; Moreno-Jiménez, Aguarón y Escobar, 2001), se han incorporado otras tres etapas en su aplicación práctica. Una, al comienzo, orientada hacia una correcta formulación del problema (hacer lo correcto y evitar el error Tipo III) y otras dos, al final, dirigidas hacía la extracción del conocimiento asociado a la resolución científica del problema, mediante la explotación del modelo matemático utilizado en la misma.

Por su parte, y a diferencia de otras técnicas multicriterio, AHP permite evaluar la inconsistencia de los decisores (transitividad cardinal) a la hora de emitir sus juicios (Saaty, 1980; Moreno-Jiménez, 2002). Para ello, se dispone de diferentes indicadores, asociados a los métodos de priorización más extendidos. En concreto, la Razón de Consistencia (CR) de Saaty, para el método de priorización del autovector principal (Saaty, 1980), y el Índice de Consistencia Geométrico (GCI), para el método de priorización de la media geométrica por filas (Aguarón y Moreno-Jiménez, 2003).

Moreno-Jiménez (2003) distingue tres situaciones en la toma de decisiones con múltiples actores: (i) *Decisión en Grupo*, donde los actores operan conjuntamente en la búsqueda de un fin común, bajo la perspectiva del consenso; (ii) *Decisión Negociada*, donde los actores actúan de forma individualizada, buscando las zonas de acuerdo y desacuerdo y (iii) *Decisión Sistémica*, donde los actores operan de forma individualizad persiguiendo una solución holística bajo el principio de tolerancia.

La literatura científica clásica sobre AHP (Saaty, 1980; Ramanatham y Ganesh, 1994; Forman y Peniwati, 1998) contempla dos aproximaciones diferentes en la toma de decisiones en grupo: la Agregación de los Juicios Individuales (AIJ) y la Agregación de Prioridades Individuales (API). Recientemente, el Grupo Decisión Multicriterio Zaragoza (GDMZ) ha propuesto otras tres aproximaciones para esta situación (decisión en grupo): una Bayesiana (Altuzarra, Moreno-Jiménez, Salvador, 2007); la Agregación de Estructuras de Preferencia Individuales, AIPS (Escobar y Moreno-Jiménez, 2007), y el Corazón de Consistencia (Moreno-Jiménez, Aguarón y Escobar, 2008).

Esta última herramienta, basada en los intervalos de estabilidad en consistencia (Aguarón, Escobar y Moreno-Jiménez, 2003), construye la matriz de consenso en consistencia y, a partir de ella, analiza la evolución de las diferentes estructuras de preferencia para diferentes umbrales de inconsistencia. Las entradas de esta matriz indican el rango de valores en el que todos los juicios individuales pueden oscilar simultáneamente dentro del límite de inconsistencia en sus respectivas matrices iniciales.

En lo que sigue, se propone un nuevo procedimiento iterativo que permite acercar las posturas entre los diferentes actores implicados en la resolución del problema. Para ello, se identifican los valores precisos de los intervalos de estabilidad en consistencia comunes a los actores implicados, que maximizan la menor holgura de consistencia disponible para los diferentes decisores.

El trabajo ha quedado estructurado como sigue. La Sección 2 recoge los contenidos de AHP necesarios para la nueva propuesta. La Sección 3 presenta el procedimiento de cálculo. La Sección 4 incluye el caso práctico utilizado para ilustrarlo y, finalmente, la Sección 5 resalta las conclusiones más destacadas del trabajo.

2. AHP. DECISIÓN EN GRUPO Y CONSISTENCIA

El *Proceso Analítico Jerárquico* –AHP– (Saaty, 1977, 1980) es una técnica que permite la resolución de problemas con múltiples escenarios, actores y criterios, incorporando en el

modelo los aspectos subjetivos asociados al factor humano. Las características principales de este método son: la modelización del problema mediante la construcción de una jerarquía en la que se recojan los aspectos relevantes del mismo (escenarios, actores, criterios, subcriterios, alternativas, etc.); la incorporación de las preferencias mediante comparaciones pareadas; y la obtención de una escala derivada de prioridades relativas medidas en una escala de razón.

Básicamente, el método original de AHP consta de cuatro etapas:

1. *Modelización* del problema o construcción de la jerarquía.
2. *Valoración* o emisión de juicios.
3. *Priorización* u obtención de las prioridades locales y globales.
4. *Síntesis* u obtención de las prioridades totales.

En la primera etapa se construye un modelo, en este caso una jerarquía, que represente el problema de decisión. En la segunda etapa, el decisor incorpora sus preferencias mediante los juicios asociados a las comparaciones pareadas de los elementos considerados. Estos juicios, medidos en la escala fundamental de Saaty (1977, 1980), reflejan la importancia relativa respecto al elemento en común del que cuelgan en la jerarquía. El resultado es una matriz cuadrada, recíproca y positiva denominada matriz de comparaciones pareadas. En la tercera etapa, a partir de estas matrices, se calculan las prioridades locales utilizando para ello alguno de los métodos de priorización propuestos en la literatura. Los dos más utilizados son el método del autovector por la derecha (EGV) (Saaty, 1980) y el método de la media geométrica por filas (RGM) (Saaty, 1980; Crawford y Williams, 1985). A partir de las prioridades locales, aplicando el principio de composición jerárquica, se obtienen las prioridades globales (prioridades respecto a la meta o misión). En la cuarta y última etapa, se sintetizan las prioridades globales calculadas en la etapa anterior, obteniendo las prioridades totales de las alternativas.

El proceso analítico jerárquico (AHP) ha sido aplicado en numerosas ocasiones en problemas de decisión con múltiples actores y son numerosas las propuestas metodológicas al respecto (Saaty, 1989; Iz y Gardiner, 1993; Bryson, 1996; Condon et al. 2003; Moreno-Jiménez et al. 2005; Escobar y Moreno-Jiménez, 2007).

A la hora de obtener las prioridades para el grupo o colectivo a partir de las opiniones individuales, se suelen seguir diferentes procedimientos (Saaty, 1980; Aczel y Saaty, 1983; Dyer y Forman, 1992; Ramanathan y Ganesh 1994; Van den Honert y Lootsma, 1997; Forman y Peniwati, 1998; Van den Honert, 2001; Altuzarra y otros, 2007; Gargallo y otros, 2006), siendo los dos más extendidos en la práctica los siguientes: (i) la agregación de juicios individuales (AIJ) y (ii) la agregación de prioridades individuales (AIP), que se describen a continuación.

Suponiendo un contexto local (un único criterio en la jerarquía) con n alternativas ($A_1,...,A_n$) y r decisores ($D_1,..,D_r$), y denotando por $A^{[k]} = (a_{ij}^{[k]})$ la matriz de comparaciones pareadas del k-ésimo decisor (k=1,...,r; i, j = 1,...,n) y por β_k su importancia relativa en el grupo ($\beta_k \geq 0$, $\sum_{k=1}^{r} \beta_k = 1$), las prioridades de las alternativas comparadas según los dos procedimientos AIJ y AIP se obtienen como:

(i) *Agregación de Juicios Individuales* (AIJ): A partir de las matrices de juicios individuales, $A^{[k]}$ k=1,...,r, se construye una matriz de juicios para el grupo $A^{[G]} = (a_{ij}^{[G]})$ utilizando cualquier procedimiento de agregación de juicios individuales y, a partir de ella, se obtienen las prioridades de las alternativas, $w^{[G/J]} = (w_i^{[G/J]})$ mediante cualquier procedimiento de priorización.

(ii) *Agregación de Prioridades Individuales* (AIP): A partir de las matrices de juicios individuales, $A^{[k]}$ k=1,...,r, se obtienen las prioridades individuales mediante cualquier procedimiento de priorización, $w^{[k]} = (w_i^{[k]})$, k = 1,...,r y, a partir de ellas, utilizando cualquier procedimiento de agregación de prioridades individuales se obtienen las prioridades de las alternativas, $w^{[G/P]} = (w_i^{[G/P]})$.

En cuanto a los procedimientos de agregación, tanto de juicios como de prioridades, se han propuesto diferentes procedimientos en la literatura, siendo el más utilizado el de la media geométrica, que es la única función de síntesis separable que satisface las condiciones de unanimidad, homogeneidad y reciprocidad (Aczel y Saaty, 1983).

Una de las características más destacadas de AHP es la posibilidad de evaluar la consistencia del decisor en la emisión de juicios, no siendo necesario que dichos juicios sean perfectamente consistentes o cardinalmente transitivos. Dada una matriz de comparaciones pareadas $A = (a_{ij})$, $i,j = 1,\ldots,n$, se dice que dicha matriz es *consistente* cuando hay transitividad cardinal entre los juicios, esto es, cuando $a_{ij} \cdot a_{jk} = a_{ik}\ \forall i,j,k$.

Saaty sugiere para el método AHP convencional (en el que se utiliza el método del autovector principal para obtener las prioridades), que la inconsistencia sea capturada por un único valor, $\lambda_{max}-n$, que refleja la desviación de los juicios a_{ij} con respecto al cociente estimado entre las prioridades, w_i/w_j. La medida de la consistencia propuesta se define a continuación.

Definición 2.1. Dada una matriz de comparaciones pareadas $A = (a_{ij})$, $i,j = 1,\ldots,n$, λ_{max} su valor propio máximo, y $w = (w_j)$, $j = 1,\ldots,n$, el correspondiente vector de prioridades obtenido aplicando el método EGVM, se define el *Índice de Consistencia* (*CI*) como:

$$CI = \frac{\lambda_{max} - n}{n-1} \tag{2.1}$$

Este Índice de Consistencia puede rescribirse expresándolo en términos de los errores $e_{ij} = a_{ij} \cdot \frac{w_j}{w_i}$ como:

$$CI = \frac{1}{n(n-1)} \sum_{i \neq j} (e_{ij} - 1) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} (e_{ij} - 1) \tag{2.2}$$

Al depender esta medida del orden de la matriz (n), Saaty propuso la utilización de la *Razón de Consistencia* (*CR*) que se obtiene dividiendo el *CI* por su valor esperado, calculado a partir de un gran número de matrices recíprocas positivas de orden n generadas aleatoriamente. Saaty (1980, 1994) considera que el vector de prioridades tiene una inconsistencia aceptable cuando el *CR* es menor del 10% (5% y 8% para $n = 3$ y $n = 4$ respectivamente).

Otra medida que permite evaluar la inconsistencia del decisor al emitir sus juicios es la propuesta por Crawford y Williams (1985) y formalizada, bajo el nombre de Índice de Consistencia Geométrico (GCI), por Aguarón y Moreno-Jiménez (2003).

Definición 2.2. Dada la matriz de comparaciones pareadas $A = (a_{ij})$, $i,j = 1,\ldots,n$, y $w = (w_j)$, $j = 1,\ldots,n$, el correspondiente vector de prioridades obtenido mediante el método RGMM, se define el *Índice de Consistencia Geométrico* (*GCI*) como:

$$GCI = \frac{2}{(n-1)(n-2)} \sum_{1 \leq i < j \leq n} \log^2 e_{ij} \ \text{ con } \ e_{ij} = a_{ij} \cdot \frac{w_j}{w_i} \tag{2.3}$$

Aguarón y Moreno-Jiménez (2003) establecieron los umbrales, dependiendo del orden de la matriz, que permiten una interpretación de la inconsistencia análoga a la del 10% para la Razón de Consistencia de Saaty. Estos valores son 0.31 para $n = 3$; 0.35 para $n = 4$ y 0.37 para $n > 4$.

Con respecto a la consistencia en la toma de decisión en grupo con AHP, se ha demostrado (Xu, 2000; Escobar y otros, 2004) que, cuando se usa la media geométrica como procedimiento de agregación, si los decisores individualmente tienen una inconsistencia aceptable, entonces también la tiene el grupo, independientemente del método de priorización utilizado (EGV o RGM).

Otros trabajos en los que también se utiliza la consistencia para abordar problemas de toma de decisiones en grupo son los de Moreno-Jiménez y otros (2005) y Moreno-Jiménez y otros (2008) en los que se busca el consenso entre los diferentes individuos.

3. PROCEDIMIENTO DE CÁLCULO

El procedimiento que se propone consiste en la transformación gradual de las matrices de juicio de los diferentes actores participantes en el Proceso de Toma de Decisiones buscando un acercamiento progresivo de sus posturas. En ocasiones, será posible alcanzar una matriz común de juicios para todos los individuos, mientras que en otras se alcanzará un acuerdo parcial para algunos de los juicios. Este procedimiento utiliza los intervalos de estabilidad en consistencia como herramienta para acercar las posturas entre los decisores. A continuación se definen estos intervalos a partir de la noción de consistencia introducida en la sección anterior.

Definición 3.1. Dada una matriz de comparaciones pareadas $A = (a_{ij})$, $i,j = 1,\ldots,n$, y una variación Δ ($\Delta>0$) para el GCI, se definen:

El *Intervalo de Estabilidad Relativa en Consistencia* (*CRSI*) para el juicio a_{rs}, es el intervalo $\left[\underline{\delta}_{rs}(\Delta), \overline{\delta}_{rs}(\Delta)\right]$ en el que pueden oscilar sus variaciones relativas sin que el incremento del GCI exceda el valor de Δ.

El *Intervalo de Estabilidad en Consistencia* (*CSI*) para el juicio a_{rs} dado Δ, es el intervalo $\left[\underline{a}_{rs}, \overline{a}_{rs}\right]$, en el cuál este puede oscilar sin que el incremento del GCI exceda el valor de Δ. Su valor viene dado por: $\underline{a}_{rs} = a_{rs}\underline{\delta}_{rs}$ y $\overline{a}_{rs} = a_{rs}\overline{\delta}_{rs}$.

Notación: Sea $A=(a_{ij})$ la matriz de comparaciones pareadas, se denota por *GCI(A)* al valor del índice de consistencia geométrico de la matriz A y por $A'=(a'_{ij})$ a la matriz obtenida al modificar uno o varios de los juicios de la matriz original. La variación relativa del juicio a_{ij} se representa por $t_{ij} = a'_{ij} / a_{ij}$ y su logaritmo por τ_{ij} ($\tau_{ij} = \log t_{ij}$). Finalmente, e_{ij} ($e_{ij} = a_{ij}\omega_j / \omega_i$) representan los errores y ε_{ij} sus logaritmos ($\varepsilon_{ij} = \log e_{ij}$).

Teorema 3.1. Sea $A=(a_{ij})$ una matriz de de comparaciones pareadas. El Índice de Consistencia Geométrico (GCI) de la matriz A', obtenida modificar el juicio a_{rs}, viene dado por:

$$GCI(A') = GCI(A) + \frac{2}{n(n-1)}\left(\tau_{rs}^2 + \frac{2n}{n-2}\tau_{rs}\varepsilon_{rs}\right) \tag{3.1}$$

Corolario 3.1. Dada una matriz de comparaciones pareadas, el intervalo de estabilidad relativa en consistencia del juicio a_{rs} para una variación Δ, se puede obtener a partir de las soluciones de la ecuación:

$$\tau_{rs}^2 + \frac{2n}{n-2}\tau_{rs}\varepsilon_{rs} - \frac{n(n-1)\Delta}{2} = 0 \tag{3.2}$$

Se puede observar que la expresión anterior corresponde con una ecuación de segundo grado en la que el término independiente es negativo (recordar que $\Delta>0$), por lo que siempre obtendremos dos soluciones reales, una negativa y otra positiva ($\underline{\tau}_{rs}$ y $\overline{\tau}_{rs}$).

Esas soluciones corresponden con los valores "críticos" de τ_{rs} y, a partir de ellos, podemos obtener los de t_{rs} teniendo en cuenta que $t_{ij} = \exp(\tau_{ij})$. Por lo tanto, obtendremos dos valores para la variación relativa del juicio, uno mayor y el otro menor que la unidad. Estos son los valores del correspondiente intervalo buscado. Por último, a partir de los valores extremos de t_{rs} podemos obtener los intervalos de estabilidad en consistencia para el juicio a_{rs}, teniendo en cuenta que $a'_{ij} = a_{ij}t_{ij}$. Resumiendo, el juicio a_{rs} puede oscilar entre los valores $a_{rs}e^{\underline{\tau}_{rs}}$ y $a_{rs}e^{\overline{\tau}_{rs}}$ sin que el aumento del GCI supere el valor Δ.

El procedimiento que se propone en este trabajo comienza identificando aquél juicio en el que existe una menor variabilidad entre los diferentes individuos. Para dicho juicio se calculan los intervalos de estabilidad en consistencia de los diferentes individuos y se obtiene su intersección. En este intervalo común está asegurado que los juicios de los individuos pueden variar sin que el GCI sobrepase un determinado nivel de inconsistencia fijado de antemano. En lugar de considerar este intervalo u otro contenido en él como sustituto de los juicios de los decisores (Moreno-Jiménez, Aguarón y Escobar, 2008), se propone seleccionar un valor preciso perteneciente a dicho intervalo común.

Cualquier valor considerado dentro de este intervalo proporciona matrices con inconsistencias admisibles. Sin embargo, alguna de las matrices tendrá mayor inconsistencia que las demás, por lo que permitirá menos holgura para las iteraciones siguientes. Parece lógico elegir aquel juicio que deje la mayor holgura posible para la matriz más inconsistente o, dicho de otra manera, que minimice el GCI de la más inconsistente. Por lo tanto el problema considerado es:

$$\begin{aligned} &\underset{a_{rs}}{Min} \quad \underset{k}{Max} \quad GCI\left(A'^{(k)}\right) \\ &a_{rs} \in \left[\underline{a}_{rs}, \overline{a}_{rs}\right] \end{aligned} \tag{3.3}$$

Que puede ser transformado en el problema:

$$\begin{aligned} &\underset{a_{rs}}{Min} \; d \\ &GCI\left(A'^{(k)}\right) \le d \quad \text{para todo } k \\ &a_{rs} \in \left[\underline{a}_{rs}, \overline{a}_{rs}\right] \end{aligned} \tag{3.4}$$

A partir de la expresión del Teorema 3.1, se puede expresar como:

$$\begin{aligned} &\underset{a_{rs}}{Min} \; d \\ &GCI(A^{(k)}) + \frac{2}{n(n-1)}\left(\tau_{rs}^{(k)2} + \frac{2n}{n-2}\tau_{rs}^{(k)}\varepsilon_{rs}^{(k)}\right) \le d \quad \text{para todo } k \\ &a_{rs} \in \left[\underline{a}_{rs}, \overline{a}_{rs}\right] \end{aligned} \tag{3.5}$$

La variable original del problema era a_{rs}, pero se plantea en término de log a_{rs}:

$$\begin{aligned} &\underset{a_{rs}}{Min} \; d \\ &GCI(A^{(k)}) + \frac{2}{n(n-1)}\left(\left(\log a_{rs} - \log a_{rs}^{(k)}\right)^2 + \frac{2n}{n-2}\left(\log a_{rs} - \log a_{rs}^{(k)}\right)\varepsilon_{rs}^{(k)}\right) \le d \quad \text{para todo } k \\ &a_{rs} \in \left[\underline{a}_{rs}, \overline{a}_{rs}\right] \end{aligned} \tag{3.6}$$

Mediante el cambio de variable, $\alpha_{rs} = \log a_{rs}$ se obtiene:

$$\begin{aligned} &\underset{\alpha_{rs}}{Min} \; d \\ &GCI(A^{(k)}) + \frac{2}{n(n-1)}\left(\left(\alpha_{rs} - \alpha_{rs}^{(k)}\right)^2 + \frac{2n}{n-2}\left(\alpha_{rs} - \alpha_{rs}^{(k)}\right)\varepsilon_{rs}^{(k)}\right) \le d \quad \text{para todo } k \\ &\alpha_{rs} \in \left[\log \underline{a}_{rs}, \log \overline{a}_{rs}\right] \end{aligned} \tag{3.7}$$

Algoritmo:

Paso 1. Para cada comparación pareada (*i,j*), calcular la varianza de los logaritmos de los correspondientes juicios: $\sigma_{ij}^2 = Var\left(\log a_{ij}^{(k)}\right)$.

Paso 2. Tomar (*r,s*) como los índices del juicio con menor varianza.

Paso 3. Para cada decisor *k*, calcular el intervalo de estabilidad en consistencia: $\left\lfloor \underline{a}_{rs}^{(k)}, \overline{a}_{rs}^{(k)} \right\rfloor$.

Paso 4. Determinar la intersección de dichos intervalos de consistencia. Si la intersección es vacía, ir al Paso 8.

Paso 5. Obtener la intersección del anterior intervalo con el intervalo $\left[\underset{k}{Min}\{a_{rs}^{(k)}\}, \ \underset{k}{Max}\{a_{rs}^{(k)}\} \right]$. Con ello evitamos salirnos del recorrido de los juicios emitidos por los decisores. Denotaremos por $\left\lfloor \underline{a}_{rs}, \overline{a}_{rs} \right\rfloor$ a este intervalo. Es decir:

$$\left\lfloor \underline{a}_{rs}, \overline{a}_{rs} \right\rfloor = \bigcap_{k} \left\lfloor \underline{a}_{rs}^{(k)}, \overline{a}_{rs}^{(k)} \right\rfloor \bigcap \left[\underset{k}{Min}\{a_{rs}^{(k)}\}, \ \underset{k}{Max}\{a_{rs}^{(k)}\} \right].$$

Paso 6. Obtener el valor preciso $b_{rs} \in \left\lfloor \underline{a}_{rs}, \overline{a}_{rs} \right\rfloor$ que minimiza la inconsistencia de la matriz más inconsistente. Para ello resolveremos el problema (3.7).

Paso 7. Hacer $a_{rs}^{(k)} = b_{rs}$ *y* $a_{sr}^{(k)} = 1/b_{rs}$ para todo *k.*

Paso 8. Si aún quedan juicios por analizar, elegir aquel que presenta menor valor para σ_{ij}^2. Tomar (*r,s*) como los índices de dicho juicio e ir al Paso 3.
Si no quedan juicios por analizar, fin.

Como resultado de la aplicación de este procedimiento, en cada iteración se va produciendo un acercamiento entre las posturas de los diferentes actores participantes en el proceso de toma de decisión. Al utilizar la consistencia se garantiza que las matrices de juicio que se van construyendo en las diferentes iteraciones tienen una inconsistencia admisible, al permitir variar sus juicios dentro de los correspondientes intervalos de estabilidad en consistencia.

El resultado final del procedimiento puede ser una matriz común de juicios para todos los decisores. Esta matriz se utilizará como la matriz de juicios del grupo y a partir de la misma se obtendrá el vector de prioridades asociado al grupo. En otras ocasiones, no será posible alcanzar una matriz de juicios común si para algún juicio los intervalos de estabilidad en consistencia de los diferentes individuos no tienen intersección común. En este caso el procedimiento proporciona un acuerdo parcial sobre los juicios que supondrá un acercamiento en las posturas de los decisores desde sus posiciones iniciales.

4. EJEMPLO PRÁCTICO

A continuación se presenta una aplicación práctica del procedimiento propuesto. Para ello se ha utilizado el ejemplo considerado en Moreno-Jiménez y Polasek (2003) para dos decisores y cuatro alternativas.

Las matrices de comparaciones pareadas correspondientes a los dos individuos son las siguientes:

$A^{[I]} =$

	A	B	C	D
A	1	3	2	5
B	1/3	1	1/2	1
C	1/2	2	1	3
D	1/5	1	1/3	1

$A^{[II]} =$

	A	B	C	D
A	1	1/3	1/2	2
B	3	1	1	4
C	2	1	1	3
D	1/2	1/4	1/3	1

Los vectores de prioridades correspondientes a cada una de estas matrices aplicando el RGM así como el GCI son los dados en la Tabla 1.

Tabla 1. Vectores de prioridades individuales y GCI

Prioridades	Decisor I	Decisor II
A	0,487217	0,163831
B	0,133016	0,401302
C	0,273982	0,337453
D	0,105784	0,097414
GCI	0,045112	0.041194

El procedimiento comienza identificando el juicio en el que los dos decisores presentan una menor variabilidad. En este caso se trata del juicio (3,4) para el que los intervalos de estabilidad en consistencia de las dos matrices son: I1 = (0.56, 8.92) e I2 = (1, 16.08). La intersección de estos intervalos con el rango inicial de juicios es el intervalo (3, 3).

Siguiendo la propuesta realizada se debe seleccionar el valor de este intervalo que maximiza la menor holgura de consistencia disponible para las dos matrices. Como en este caso el intervalo está formado por un único punto, el valor es 3, que además es el valor original para los dos decisores, por lo que en esta iteración no se produce ningún cambio en los juicios de los decisores.

En la segunda iteración se considera el juicio (2,3). Los intervalos de estabilidad en consistencia de las matrices de los dos decisores son I1=(0.12, 1.83) e I2=(0.35, 5.76). La intersección con el rango inicial es (0.5, 1) y el valor seleccionado dentro de este intervalo, obtenido resolviendo el problema de optimización que maximiza la menor holgura de consistencia disponible para las dos matrices, es 0.77. En las dos matrices de juicio se sustituyen los valores que inicialmente se habían emitido para el juicio (2,3) por el nuevo valor 0.77.

El procedimiento se reitera para el resto de juicios que van siendo seleccionados de menor a mayor variabilidad. En la Tabla 2 se muestran de forma resumida los pasos seguidos en el algoritmo, indicando para cada iteración el juicio sobre el que se actúa, los dos intervalos de estabilidad en consistencia, el intervalo común a los dos intervalos de estabilidad en consistencia y al rango inicial de valores, así como el valor preciso seleccionado de este intervalo, que es el que maximiza la menor holgura de consistencia disponible para las tres matrices.

Tabla 2. Iteraciones del procedimiento propuesto

Iteración	Juicio	I1	I2	I común	Valor
1	(3,4)	(0.56, 8.92)	(1, 16.08)	(3, 3)	3
2	(2,3)	(0.12, 1.83)	(0.35, 5.76)	(0.5, 1)	0.77
3	(1,4)	(1.19, 15.17)	(0.38, 5.24)	(2, 5)	2.56
4	(2,4)	(0.41, 4.72)	(1.30, 13.53)	(1.30, 4)	2.55
5	(1,3)	(0.46, 4.28)	(0.16, 1.35)	(0.5, 1.35)	0.76
6	(1,2)	(0.24, 4.24)	(0.24, 4.24)	(0.33, 3)	1

En esta aplicación práctica ha sido posible encontrar para todos los juicios un intervalo común de estabilidad en consistencia y por tanto, seleccionar un valor dentro del mismo, por lo que al finalizar la aplicación del algoritmo se ha alcanzado una matriz común de juicios, que es la siguiente:

Grupo	A	B	C	D
A	1	1	0,76	2,56
B	1	1	0,77	2,55
C	1,32	1,30	1	3
D	0,39	0,39	0,33	1

Para esta matriz común el vector de prioridades aplicando el RGM y el GCI son los mostrados en la Tabla 3.

Tabla 3. Prioridades del grupo

	Grupo
A	0,271824
B	0,272447
C	0,346321
D	0,109408
GCI	0,001881

Como se puede observar, el valor del GCI se ha reducido para los dos actores participantes, de forma que en la búsqueda de un acercamiento entre las posturas de los mismos se ha producido también una mejora en los niveles de inconsistencia.

5. CONCLUSIONES

En este trabajo se ha presentado un nuevo procedimiento para la toma de decisiones en grupo con el Proceso Analítico Jerárquico (AHP). La contribución más importante del mismo es la utilización de la consistencia, a través de los intervalos de estabilidad en consistencia, para el acercamiento entre las posturas de los decisores.

El procedimiento va modificando los juicios originales de los decisores, pero siempre manteniendo los mismos dentro de sus correspondientes intervalos de estabilidad en consistencia, y por tanto, a la vez que se va produciendo un acercamiento entre las posturas de los individuos, los decisores se mantienen dentro de los niveles admisibles de inconsistencia.

En ocasiones, el procedimiento es capaz de proporcionar una matriz común de juicios, de forma que ésta será la matriz utilizada como matriz de juicios del grupo y de la misma se obtendrán las prioridades del grupo.

En estos momentos, los autores están trabajando en una extensión de este trabajo (Escobar, Moreno-Jiménez, Aguarón, 2009) para el caso en el que no es posible alcanzar una matriz común para todos los individuos. En dicho trabajo, proponen la incorporación de la actitud de los decisores ante la colaboración.

BIBLIOGRAFÍA

ACZÉL, J.; SAATY, T.L. (1983): Procedures for Synthesizing Ratio Judgements. *Journal of Mathematical Psychology*, 27 (1), 93–102.

AGUARÓN, J.; ESCOBAR, M.T.; MORENO-JIMÉNEZ, J.M. (2003): Consistency Stability Intervals for a Judgement in AHP Decision Support Systems. *European Journal of Operational Research*, 145 (2), 382-393.

AGUARÓN, J.; MORENO-JIMÉNEZ, J.M. (2003): The geometric consistency index: approximate thresholds. *European Journal of Operational Research*, 147 (1), 137-145.

ALTUZARRA, A.; MORENO-JIMÉNEZ, J.M.; SALVADOR, M. (2007): A Bayesian Priorization Procedure for AHP-Group Decision Making. *European Journal of Operational Research*, 182, 367-382.

BRYSON, N. (1996): Group Decision Making and the analytic hierarchy process: exploring the consensus-relevant information content. *Computers and Operations Research*, 23, 27-35.

CONDON, E.; GOLDEN, B.; WASIL, E. (2003): Visualizing group decisions in the analytic hierarchy process. *Computers and Operations Research*, 30 (10), 1435-1445.

CRAWFORD G.; WILLIAMS C. (1985): A Note on the Analysis of Subjective Judgment Matrices, *Journal of Mathematical Psychology*, 29, 387–405.

DYER, R.F.; FORMAN, E.H. (1992): Group Decision Support with the Analytic Hierarchy Process. *Decision Support Systems*, 8, 99-124.

ESCOBAR, M.T.; MORENO-JIMÉNEZ, J.M. (2007): Aggregation of Individual Preference Structures in AHP-Group Decision Making. *Group Decision and Negotiation*, 16(4), 287-301.

ESCOBAR, M.T., J. AGUARÓN, J.M. MORENO-JIMÉNEZ (2004): "A Note on AHP Group Consistency for the Row Geometric Mean Priorization Procedure". *European Journal of Operational Research* 153, 318-322.

ESCOBAR, M.T.; MORENO-JIMÉNEZ, J.M.; AGUARÓN, J. (2009): Consistency in AHP-Collaborative Decision Making. Aceptado para su comunicación en 23rd European Conference on Operational Research, Bonn, Julio 5 - 8, 2009.

FORMAN, E.; PENIWATI, K. (1998): Aggregating individual judgments and priorities with the Analytic Hierarchy Process. *European Journal of Operational Research*, 108, 165−169.

GARGALLO, M.P.; MORENO-JIMÉNEZ, J.M.; SALVADOR, M. (2006): AHP-Group Decision Making. A Bayesian approach based on mixtures for group pattern identification. *Group Decision and Negotiation*, 16 (6), 485-506.

IZ, P.H.; GARDINER, L.R. (1993): Analysis of Multiple Criteria Decision Support Systems for Cooperative Groups. *Group Decision and Negotiation,* 2, 61-79.

MORENO JIMÉNEZ, J.M. (2002): El Proceso Analítico Jerárquico. Fundamentos, Metodología y Aplicaciones. En Caballero, R. y Fernández, G.M. *Toma de decisiones con criterios múltiples*. RECT@ Serie Monografías nº 1, 21-53.

MORENO JIMÉNEZ, J.M. (2003): Los Métodos Estadísticos en el Nuevo Método Científico. En CASAS, J.M. y PULIDO, A.*: Información económica y técnicas de análisis en el siglo XXI* . INE, 331-348.

MORENO-JIMÉNEZ, J.M.; AGUARÓN, J.; ESCOBAR, M.T. (2001): Metodología científica en valoración y selección ambiental. *Pesquisa Operacional* 21, 3-18 (Revista de la Sociedad Brasileña de Investigación Operativa).

MORENO-JIMÉNEZ, J.M.; AGUARÓN, J.; ESCOBAR, M.T. (2008): The Core of Consistency in AHP-Group Decision Making. *Group Decision and Negotiation* 17(3), 249-265.

MORENO-JIMÉNEZ, J.M.; AGUARÓN, J.; ESCOBAR, M.T.; TURÓN, A. (1999). The multicriteria procedural rationality on SISDEMA. *European Journal of Operational Research* 119(2): 388-403.

MORENO-JIMÉNEZ, J.M.; AGUARÓN, J.; RALUY, A.; TURÓN, A. (2005): A Spreadsheet Module for consistent consensus building in AHP-Group Decision Making. *Group Decision and Negotiation*, 14 (2), 89-108.

MORENO-JIMÉNEZ, J.M.; W. POLASEK (2003): "e-Democracy and Knowledge. A Multicriteria Framework for the New Democratic Era". *Journal of Multi-Criteria Decision Analysis* 12, 163-176.

RAMANATHAN, R.; GANESH, L.S. (1994): Group Preference Aggregation Methods employed in AHP: An Evaluation and Intrinsic Process for Deriving Members' Weightages. *European Journal of Operational Research*, 79, 249−265.

SAATY, T.L. (1977): A Scaling Method for Priorities in Hierarchical Structures. *Journal of Mathematical Psychology*, 15 (3), 234−281.

SAATY, T.L. (1980): *Multicriteria Decision Making: The Analytic Hierarchy Process*. Mc Graw-Hill, New York. (2ª impresión 1990, RSW Pub. Pittsburgh)

SAATY, T.L. (1989): Group Decision Making and the AHP. En: Golden, B.L., Wasil, E.A., Harker, P.T. (eds) *The Analytic Hierarchy Process: Application and Studies*, 59-67.

SAATY, T.L. (1994a): How to Make a Decision: The Analytic Hierarchy Process, *Interfaces*, 24 (6), 19−43.

VAN DEN HONERT, R. C. (2001): Decisional Power in Group Decision Making: A Note on the Allocation of Group Members' Weights in the Multiplicative AHP and SMART. *Group Decision and Negotiation*, 10(3), 275-286.

VAN DEN HONERT, R.C.; LOOTSMA, F. A. (1997): Group preference aggregation in the multiplicative AHP The model of the group decision process and Pareto optimality. *European Journal of Operational Research*, 96(2), 363-370.

XU, Z. (2000): "On consistency of the weighted geometric mean complex judgement matrix in AHP". *European Journal of Operational Research* 126, 683-687.

PROFILING MOBILE PHONE USERS IN EUROPE

PAULA VICENTE
E-mail: paula.vicente@iscte.pt
ELIZABETH REIS
E-mail: elizabeth.reis@iscte.pt
SCTE-Lisbon University Institute

Resumo

O telefone tornou-se no final dos anos 80 do século XX um meio importante de recolha de informação nos contextos da investigação empresarial e económica, nomeadamente nos países da América do Norte e da Europa Ocidental.
Dados recentes sobre acessos telefónicos mostram um crescimento massivo e consistente dos telemóveis, ao mesmo tempo que os telefone fixos vão perdendo importância. Entre os países da Europa encontram-se diferentes perfis de acesso telefónico.
O artigo propõe uma segmentação dos países europeus de acordo com o tipo de acesso telefónico nos agregados. Adicionalmente, explora as diferenças entre utilizadores e não utilizadores de telemóveis e identifica as principais razões para a utilização dos telemóveis.

Palavras-Chave: telemóveis, telefones fixos, análise de *clusters*, regressão logística.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

The telephone has turned out to be, since mid 80s of the last century, an important tool for data collection in the context of business and economic research, namely in North America and Western European countries.
Recent statistics on phone access show a massive and consistent growth of mobile phones usage, simultaneously with a loose of importance of fixed phones. Across European countries different patterns of telephone ownership can be found.
This article suggests a segmentation of European countries according the type of phone access in households. Additionally, explores the differences between mobile phone users and non-users and uncovers the main reasons for mobile phone usage.

Key words: mobile phones, fixed phones, cluster analysis, logistic regression.
Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUCTION

The telephone can be a valuable tool for data collection purposes due to the important time and cost savings that allows in comparison to other modes. In the last quarter of the 20th century, fixed phone has achieved good coverage rates, namely across North America and Western Europe countries, which contributed for telephone to become a relevant mode of data collection in business and economic fields.

More recently, statistics on telephone access show that in some European countries the coverage provided by fixed phone is already below 60% and with a tendency to decrease (EC 2006, 2007, 2008). At the same time, is impressive the penetration rate of mobile phones – in some European countries is already above 80%. Under this scenario, mobile phone appears as an alternative to fixed phone.

Although across Europe the type of phone access varies greatly from country to country common patterns of phone access can be found among sub-groups of countries.

The article will segment European countries according to type of phone access in households; additionally, describes the basic demographic differences between those who have and those who doesn't have a mobile phone and identifies major reasons for mobile phone usage.

The article is organized as follows. The next section presents the data used in the analyses. Then it goes on to countries segmentation. This is followed by results presentation regarding mobile phone owners profile and reasons for mobile phone usage. Finally, we present the main conclusion of our research.

2. DATA

Eurobarometer is a major study carried out five or six times a year for the EC across the European Union member states. It covers the population of the respective nationalities of the States, resident in each of the Member States and aged 15 years and over.

The basic sample design applied in all states is a multi-stage, random (probability) one. In each country, a number of sampling points is drawn with probability proportional to population size (for a total coverage of the country) and to population density. In order to do so, the sampling points are drawn systematically from each of the "administrative regional units", after stratification by individual unit and type of area. In each of the selected sampling points, a starting address is drawn, at random. Further addresses (every N^{th} address) are selected by standard "random route" procedures, from the initial address. In each household, the respondent is drawn, at random (following the "closest birthday rule"). All interviews are conducted face-to-face in people's homes and in the appropriate national language.

As far as the data capture is concerned, CAPI (Computer Assisted Personal Interview) is used in those countries where this technique was available.

Data shown in these analysis is drawn from *Eurobarometer 66.3* (2007) and from *Special E-communications household survey 293* (2008). Values underlying each country are based on approximately 1000 interviews. The questions selected for analysis are contained in a module relating to access to telecommunications.

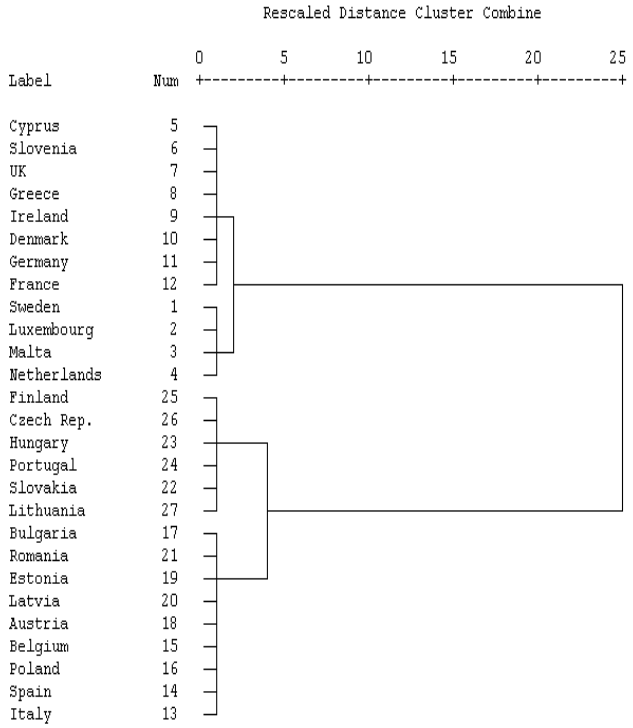
3. PHONE ACESS IN EUROPE

Historically telephone coverage was resumed to "the number of lines in the home" but the development of wireless connections has changed the research landscape and today both fixed and mobile phones must be considered when we look at telephony coverage.

Taking data for household telephony access for all the EU27 member countries regarding: percentage of households with fixed telephone access and mobile phone access (fixed and mobile), percentage of households with fixed telephone access but no mobile telephone access (fixed only), percentage of households with mobile telephone access but no fixed telephone access (mobile-only) and percentage of households without fixed telephone access or mobile telephone access (no phone), we performed a cluster analysis in order to identify clusters of countries with similar profiles in terms of phone access. We adopted Squared Euclidean Distance as a dissimilarity measure due to the type of variables – metric – and also because it is one of the most used in management sciences. Since no single method can be considered the best for classifying the countries, we have tried two hierarchical methods - the Furthest Neighbour method and the Ward method – and one non-hierarchical method – K-means – and confront the solutions obtained from each.

Figure 1 presents the dendrogram obtained with a hierarchical cluster analysis using Ward method.

Figure 1 – Dendrogram using Ward Method and Squared Euclidean Distance



Given the Dendrogram and the graphical representation of the Coefficient of Fusion (not presented) obtained under both hierarchical methods, we chose 3 clusters. We then applied the K-means method to the previous 3 cluster solution. The contingency table constructed with the solution from the hierarchical Ward method and the non-hierarchical method revealed a strong association between the two solutions (Contingency coefficient = 0.677, p-value = 0.000); this indicates that our data has a segmentation structure that is unanimously revealed by Cluster Analysis, according to the methods adopted.

Table 1 resumes the clusters and the countries classified in each one. Cluster 1 contains 12 countries, being 9 of them "old" European Union member states; cluster 2 contains 6 countries, most of them are new European Union member states and cluster 3 contains 9 countries, most of them from "new" Europe.

Table 1 – Clusters of countries according to type of phone access in the households

Cluster 1	Cluster 2	Cluster 3
Cyprus	Finland	Bulgaria
Slovenia	Czech Republic	Romania
United Kingdom	Hungary	Estonia
Greece	Portugal	Latvia
Ireland	Slovakia	Austria
Denmark	Lithuania	Belgium
Germany		Poland
France		Spain
Sweden		Italy
Luxembourg		
Malta		
Netherlands		

The description of the clusters in order to assess differences between the groups complements the segmentation. Although the groups may not be considered as independent, we have used significance tests to evaluate the differences; the p-values of the significance tests are not to be interpreted literally, but are merely standard values that say how large the difference between the realities under comparison needs to be before we can take note of it.

Table 2 presents the mean values for the percentage of household having each type of phone access in each cluster; mean values for the EU27 member states are also presented.

The One-way ANOVA test revealed significant differences between groups (*p-values*≤0.01) in fixed and mobile access, mobile only access and no phone access variables; clusters were not significantly different in terms of the percentage of fixed only households. Cluster 1 includes countries with high shares of dual-connections; cluster 2 includes the mobile only countries: the proportion of households that have only mobile phone access is higher than the proportion of households that have either dual-access or fixed only access; finally, cluster 3 has a "median" profile with similar rates between fixed and mobile access and mobile only access.

Table 2 – Mean values for the type of phone access in the households by cluster

	Fixed and mobile	Fixed only	Mobile only	No phone
Cluster 1	74.8%	11.0%	12.4%	0.8%
Cluster 2	28.7%	8.2%	53.8%	6.3%
Cluster 3	43.8%	12.7%	34.2%	6.2%
ANOVA p-value	*0.000*	*0.163*	*0.000*	*0.002*
EU27 [(a)]	57.0%	14.0%	24.0%	3.0%

[(a)] Source: Special E-communication household survey 293 (EC 2008)

On average for the EU27 countries the penetration rate of the mobile phone is higher than for fixed phone (81% against 71%, respectively). The phenomenon of mobile only access is more visible among the "new" Europe – cluster 2 and cluster 3.

4. DEMOGRAPHIC CHARACTERISTICS OF MOBILE PHONE USERS

A binary logistic regression model for the probability of having a mobile phone connection is estimated. This model allows to investigate within a multivariate framework whether, and to what extent, mobile phone users are different from mobile phone non-users.

The logistic model was constructed by testing the main effects of the independent variables: gender, age, age of education, occupation, marital status and type of community, and by taking the *mobile phone status*, coded as 1 = has a mobile phone and 0 = does not have a mobile phone, as the dependent variable.

Estimates from the Binary Logistic Model on the probability of having a mobile phone are presented in Table 3. The model suggests that the probability of being a mobile phone user varies by age (significant effect at $p < 0.000$), age of education ($p < 0.000$), occupation ($p < 0.000$), marital status ($p < 0.000$) and type of community. Gender ($p > 0.900$) is not a significant predictor of having a mobile phone.

Table 3 – Binary logistic model for predicting mobile phone status

Predictor of mobile phone status	Estimate	SE	Odds Ratio
Gender (reference: female)	–0.01	0.04	0.99
Age group (reference: 15 – 24 years)			
25 – 34 years	+2.07***	0.11	7.92
35 – 44 years	+1.82***	0.08	6.19
45 – 54 years	+1.32***	0.07	3.75
55 – 64 years	+0.80***	0.06	2.22
65 years or more	+0.78***	0.05	2.20
Age of education (reference: 15 –)			
16 – 19	+0.42***	0.04	1.53
20 +	+0.82***	0.06	2.28
Occupation (reference: self-employed)			
Employed	+0.03	0.08	1.03
Not-working	–0.63***	0.08	0.53
Marital status (reference: married or living with a partner)			
Single	–0.42***	0.06	0.66
Divorced or separated	+0.01	0.07	1.00
Widow	–0.79***	0.05	0.45
Type of community (reference: rural area or village)			
Small or middle sized town	+0.41***	0.04	1.51
Large town	+0.56***	0.05	1.74
Constant	+0.32***	0.09	

*** $p < 0.001$

The likelihood of having a mobile phone is higher among the youngest age groups. In odds metric, holding all other variables constant, for every mobile phone user aged 15 – 24 years (the baseline category) we estimate nearly eight individuals aged 25 to 34 years (odds ratio = 7.92:1) and more than six (6.19:1) individuals aged 35 – 44 years in the same condition, while for the elder the estimate is only two individuals for each individual aged 15 – 24 years. Regarding age the pattern is approximately linear: the odds of having a mobile phone decrease as age increases.

Terminal age of education has a significant effect on mobile phone status. The odds of being a mobile phone user are higher for those who left school at the age of 20 or more in comparison with those who left school at 15 or less (the reference category). Regarding this variable it can be said that the likelihood of finding a mobile phone user increases as the terminal age of education increases.

Occupation is also a significant predictor of the mobile phone status: being a non-worker (unemployed, housewife, retired) has a significant and negative association with mobile phone status. The odds of being a mobile phone user in the not-working group are nearly two times lower (odds ratio = 0.53:1) than in the self-employed group (our baseline category).

Regarding marital status the odds of single people and widow people being mobile phone users are lower than for married or living with a partner people.

Finally, a significant effect is found for the type of community. The baseline category is rural areas or village and it is found that it is less likely to find mobile phone users in those type of communities than in larger towns.

On balance, the logistic regression results indicate that except for gender all the variables under analysis are correlated with mobile phone status, but age, and more specifically being young, is the strongest correlate of mobile phone ownership.

5. MOTIVES FOR MOBILE PHONE USAGE

Table 4 shows the main benefits mobile phones users recognize to mobile phones. Among mobile phone users 38.6% highlight the "possibility to be contacted at any place, any time"

while 34.6% of mobile phone users chose "the security of knowing that a call can be made from anywhere if things go wrong" as the major benefit of owning a mobile phone.

Table 4 – Main benefits of mobile phones

Benefit	% of responses
The possibility to be contacted at any place, any time	38.6
The freedom to make calls when you are out and about	18.3
The security of knowing that you can make a call from anywhere if things go wrong	34.6
None of the above	8.5

Source: Eurobarometer 66.3 (EC 2007)

Regarding eight specific features allowed by the mobile phone service (see Table 5), seven of them had a mean value of agreement that reflect a positive feeling towards mobile phone service (mean values between 1.73 and 2.14). The possibility to verify the mobile telephone service consumption in a simple and consumer friendly way was the feature to which respondents tended to agree more (mean value of 1.73). Comparing mobile tariffs published by third parties was the feature to which respondents tended to disagree more (mean value of 2.88).

Table 5 – Positive features of mobile phone service [(a)]

Feature	Mean
You are able to verify your mobile telephone service consumption in a simple and consumer friendly way	1.73
You can keep the mobile telephone number when changing from one operator to another	1.80
You are always able to connect to the mobile network to make a phone call	1.85
You can block or unblock access to certain costly mobile services	1.86
Your mobile communication never cuts-off while on a call	1.98
You can easily compare your current mobile tariff scheme with other offers	1.99
The use of a mobile phone does not put your health at risk	2.14
You regularly read comparisons of mobile tariffs published by third parties	2.88

[(a)] Ordinal scale, 4 categories, ranging from 1="totally agree" to 4="totally disagree"
Source: Eurobarometer 66.3 (EC 2007)

Table 6 presents several reasons for not having a mobile phone. Nearly 40% of mobile phone non-users pointed out the fact that no other member in the households has a mobile phone as the main reason for not having a mobile phone. The fact that fixed telephone is enough for calling needs was chosen by nearly one quarter of mobile phone non-users.

Table 6 – Reasons for not owning a mobile phone

Reason	% of responses
No one in your household wants a mobile phone	39.7
The fixed telephone line(s) is(are) sufficient for current needs	24.8
Mobile telephony is too expensive	18.6
Someone in the household plans to get a mobile telephone in the next 6 months	6.0
There is good access to phones elsewhere, when outside the home (e.g. public payphones)	5.8
Coverage in the area is not very good, therefore it is not worth getting a mobile phone	1.3

Source: Eurobarometer 66.3 (EC 2007)

6. CONCLUSION

Taking the data on type of phone access in the households across the EU27 countries three distinct segments of countries could be found. The most notorious difference lays on the fact that there is a group of countries where dual phone access, i.e., fixed and mobile phone prevails,

and another group where mobile-exclusive access prevails. We can however expect these differences to smooth, giving the phenomenon, transversal to most countries, of a decrease in the percentage of households equipped with fixed phone and a growth in the percentage of households equipped with mobile phone.

Young people are more involved in the "mobile phone revolution" than the elder ones which justify to a great extent the fact that the likelihood of having a mobile phone is higher among the youngest age groups than the oldest age groups. Mobile phone users are also more likely to be found among people that leave school later and among those living in non-rural areas.

Regarding the motives to own a mobile phone it was patent that people see the mobile phone as a device that makes life easier, since it allows the individual to contact and be contacted permanently, anywhere, which surely is a factor that favours the continuous spread of mobile phone usage.

7. REFERENCES

EUROPEAN COMMISSION (2006), "Eurobarometer 269", European Commission, Brussels.
EUROPEAN COMMISSION (2007), "Eurobarometer 274", European Commission, Brussels.
EUROPEAN COMMISSION (2007), "Eurobarometer 66.3", European Commission, Brussels.
EUROPEAN COMMISSION (2008), "Special E-communications household survey 293", European Commission, Brussels.

UNA NUEVA DISTRIBUCIÓN CON COLAS PESADAS PARA EL AJUSTE DE DATOS FINANCIEROS

MARIA DEL MAR LÓPEZ MARTÍN
E-mail: mariadelmarlopez@ugr.es
CATALINA GARCÍA GARCÍA
E-mail: cbgarcia@ugr.es
Facultad de Ciencias Económicas y Empresariales
Universidad de Granada
JOSÉ GARCÍA PÉREZ
E-mail: jgarcia@ual.es
Departamento de Economía Aplicada
Universidad de Almería

Resumen

La necesidad de obtener distribuciones con un alto rango de curtosis para el tratamiento de los datos financieros nos conduce a la presentación de la distribución asintótica cuya construcción se basa en el sistema generador de van Dorp y Kotz (2003). Esta distribución puede alcanzar elevados rangos de curtosis superando a las distribuciones elementales utilizadas hasta el momento lo que la convierte en una adecuada candidata para el ajuste con datos empíricos de alta curtosis. Por otro lado, la necesidad de obtener distribuciones con colas pesadas, puesta de manifiesto en el trabajo de Hann (2008), nos sugiere la idea de realizar una mixtura con la distribución uniforme. Como era de esperar, esto ha permitido obtener una distribución con colas pesadas y que puede alcanzar altos valores de curtosis previamente establecidos.

Palabras Claves: Curtosis, Rendimientos Financieros, Colas Pesadas, mixtura.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

The necessity to obtain distributions with a high rank of curtosis for the treatment of the financial data leads us to the presentation of the asintotic distribution which can reach higher ranks of curtosis than the simple distributions used until now. Its construction is based on the generating system of van Dorp and Kotz (2003). On the other hand, the necessity to obtain distributions with heavy tails, make clear in the work of Hann (2008), suggests us the idea to make a mixture with the uniform distribution. Like was expected, this has allowed us to obtain a distribution with heavy tails and that can reach high value of curtosis previously established. Therefore, this paper presents the asintotic distribution and its mixture with the uniform distribution as appropriate candidates for the fit of empirical financial data with high curtosis and heavy tails.

Thematic Area: Quantitative Methods for Economics and Business.

INTRODUCCIÓN

Tradicionalmente la literatura ha asumido que los rendimientos financieros son variables aleatorias independientes e idénticamente distribuidas mediante la distribución normal.

Fama en 1976 señala que la distribución empírica de los rendimientos de los activos, es una distribución leptocúrtica y con unas colas más gruesas que la distribución normal. Por tanto, la distribución normal no logra ajustar de forma adecuada los rendimientos.

Se pueden citar numerosas distribuciones que han sido aplicadas como por ejemplo la distribucion log-normal, la distribución t de Student, la cual presenta unas colas más gruesas que la distribución normal y distintas mixturas de familias de distribuciones (Clark, 1973 y Praetz, 1972).

En este trabajo se presenta la distribución asintótica cuya principal característica es que permite un amplio rango de curtosis. Es importante destacar que cuando consideramos el caso simétrico (M=0,5) la distribución asintótica presenta curtosis 15 superando a la distribución normal, la distribución asintótica de Laplace, la distribución two-sided power y la distribución biparabólica.

Además, con el objeto de combinar este apuntamiento con unas colas pesadas se construye la mixtura de la distribución asintótica con la distribución uniforme, siguiendo la idea de Hahn (2008).

El trabajo se estructura como sigue: la Sección 2 presenta la distribución asintótica estudiando sus principales característica de forma. En la Sección 3, se construye la mixtura entre la distribución asintótica y la distribución rectangular. La Sección 4 presenta el procedimiento de elicitación de la mixtura asintótica-rectangular. Por ultimo se resumen las conclusiones y las futuras líneas de investigación.

LA DISTRIBUCIÓN ASINTÓTICA

Según van Dorp y Kotz, (2003), si $p(\cdot/\psi)$ es una apropiada función de densidad definida en $[0,1]$ con parámetros, o vector de parámetros, ψ, entonces, se puede construir la siguiente función de densidad unimodal en M:

$$g\{t/M, p(\cdot/\psi)\} = \begin{cases} p\left(\dfrac{t}{M}/\psi\right) & \text{si} \quad 0 \le t \le M \\ p\left(\dfrac{1-t}{1-M}/\psi\right) & \text{si} \quad M \le t \le 1 \end{cases} \tag{1}$$

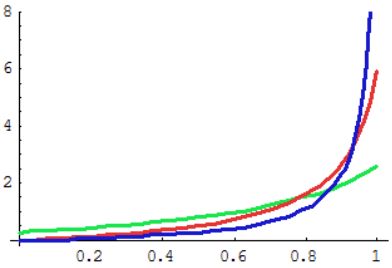
Donde $p(\cdot/\psi)$ es la función generadora de $g\{t/M, p(\cdot/\psi)\}$.

Se puede demostrar que la expresión de la función generadora que describe la distribución asintótica es:

$$p[y,n] = \frac{1}{\left(1-\frac{1}{n}\right)^{\frac{1}{n}} + L\left(1-\left(1+\frac{1}{n}\right)^{\frac{-1}{n}}\right)}\left(\frac{-1}{\left(1+\frac{1}{n-1}\right)^{\frac{1}{n}}} + \frac{-1}{y-\left(1+\frac{1}{n}\right)^{\frac{1}{n}}}\right) \tag{2}$$

La expresión (2) se obtiene a partir de las propiedades exigidas a toda función de densidad. Una de las principales características que presenta dicha distribución, y por eso recibe el nombre de distribución asíntotica, es que a medida que aumenta el valor del parámetro n el polinomio generador tiende a tomar valores muy próximos al eje de abscisas y cuando y alcanza su valor extremo ésta presenta un crecimiento muy pronunciado. Luego, teniendo en cuenta las características presentadas por dicha función de densidad generadora y utilizando las investigaciones desarrolladas por de van Dorp y Kotz (2003) se consigue describir una distribución cuya principal diferencia es que presenta un apuntamiento muy notable.

Figura 1. Función de densidad generadora
n=1.5(verde), n=3 (roja) y n=8 (azul)



Sustituyendo (2) en (1) se obtiene la función de densidad de la distribución asintótica estandarizada en $[0,1]$**:**

$$f\left(t \mid M, p(y \mid n)\right) = \frac{1}{A(n)+L(C(n))} \times \begin{cases} A(n)+\dfrac{1}{\dfrac{t}{M}-B(n)} & 0 \le t \le M \\ A(n)+\dfrac{1}{\dfrac{1-t}{1-M}-B(n)} & M \le t \le 1 \end{cases} \tag{3}$$

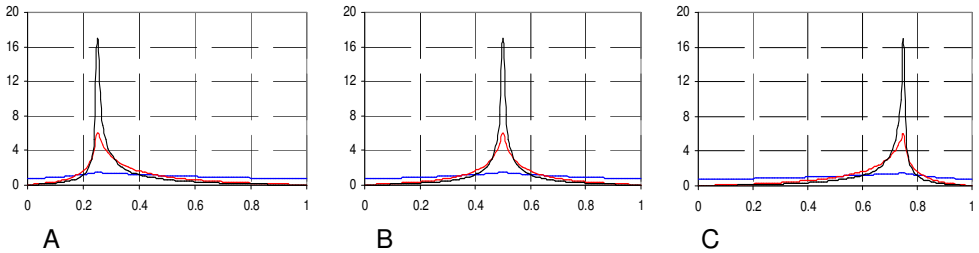
donde $A(n)=\left(1-\frac{1}{n}\right)^{\frac{1}{n}}$; $B(n)=\left(1+\frac{1}{n}\right)^{\frac{1}{n}}$; $C(n)=1-\left(1+\frac{1}{n}\right)^{\frac{-1}{n}}$ y $L[C(n)]$ indica el logaritmo neperiano de $C(n)$.

La función de distribución asociada a la función de densidad (3) corresponde a:

$$F\left(t \mid M\right) = \begin{cases} \dfrac{A(n)\dfrac{t}{M}+L\left(1-\dfrac{t}{M}\dfrac{1}{B(n)}\right)}{\dfrac{1}{M}\left(A(n)+L(C(n))\right)} & 0 \le t \le M \\ 1-\dfrac{A(n)\left(\dfrac{1-t}{1-M}\right)+L\left(1-\dfrac{1-t}{1-M}\dfrac{1}{B(n)}\right)}{\dfrac{1}{1-M}\left(A(n)+L(C(n))\right)} & M \le t \le 1 \end{cases} \tag{4}$$

Realizando el estudio gráfico de la función de densidad (2) se puede comprobar que la distribución puede presentar asimetría positiva (Figura 2.A), asimetría negativa (Figura 2.C) así como simetría cuando la moda estandarizada (M) toma el valor 0.5. Además, se observa que a medida que aumenta el valor del parámetro n, la gráfica de la función de densidad presenta una forma más puntiaguda.

Figura 2. Función de densidad para n=1.01 (azul), n=3 (roja) y n=7 (negra) con
M=0.25 (A), M=0.5 (B) y M=0.75 (C)



A continuación se analizan las principales características estocásticas que describen de una forma más precisa las características principales que hacen considerar a la distribución asintótica como una buena candidata en el estudio de los rendimientos dentro del ámbito financiero.

CARACTERÍSTICAS ESTOCÁSTICAS.

El valor esperado de la distribución asintótica estandarizada tiene la siguiente expresión:

$$E[t]=\frac{-2+4M+A(n)+2\left(1-M+\left(2M-1\right)B(n)\right)L[C(n)]}{2\left(A(n)+L[C(n)]\right)} \tag{5}$$

donde $A(n)=\left(1-\frac{1}{n}\right)^{\frac{1}{n}}$; $B(n)=\left(1+\frac{1}{n}\right)^{\frac{1}{n}}$ y $C(n)=1-\left(1+\frac{1}{n}\right)^{\frac{-1}{n}}$ son funciones dependientes únicamente del parámetro *n*.

La expresión correspondiente a la varianza resulta ser:

$$Var[t]=\frac{-3D(M,n)^2+2\left[A(n)+L\left[C(n)\right]\right]\left[-9+3ME(M,n)+6B(n)+2A(n)+6F(M,n)L\left[C(n)\right]\right]}{12\left[A(n)+L\left[C(n)\right]\right]^2} \tag{6}$$

siendo

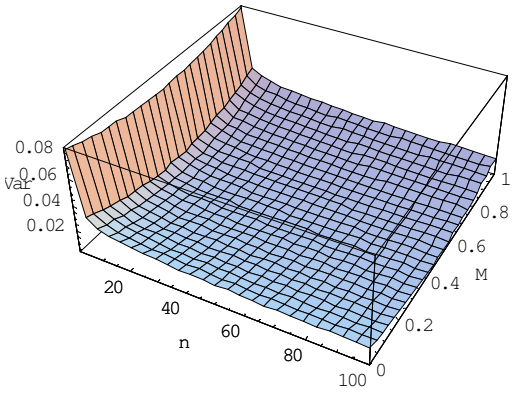
$$D(M,n)=-2+4M+A(n)+2\left(1-M+\left(2M-1\right)B(n)\right)L\left[C(n)\right]$$

$$E(M,n)=5-M+6\left(M-1\right)B(n)$$

$$F(M,n)=1-M+2\left(1-M\right)^2B(n)+\left(1-3\left(M-M^2\right)\right)B(n)^2$$

Se observa que cuando *M* toma los valores extremos y *n*=1.01, la varianza alcanza su máximo valor 0.0825894, sin embargo si se evalúa la expresión de la varianza en el caso de la distribución simétrica (*M*=0,5) entonces disminuye hasta 0.0692359. Además, tal y como se deduce de la Figura 3, la varianza decrece llegando a alcanzar valores muy próximos a cero a medida que el parámetro *n* aumenta.

Figura 3. Varianza en función de *n* y *M*



Con el objeto de profundizar en el estudio de la distribución asintótica se analizan a continuación el coeficiente de asimetría y su coeficiente de curtosis. Dada la complejidad que presentan las expresiones asociadas a ambos coeficientes se analizaran gráficamente. En las Figuras 4 y 5 se puede comprobar que la distribución presenta distintos tipos de asimetrías dependiendo del valor que tome la moda estandarizada, si ésta es menor que 0.5 entonces la distribución presenta una asimetría positiva pero a medida que el valor de *n* aumenta, la

distribución pasa a tener una asimetría negativa. En el caso particular de que *M* tome el valor central del intervalo $[0,1]$ se comprueba en la Figura 4 que la distribución es simétrica.

La simetría de la distribución normal puede considerarse en ocasiones como un inconveniente que imposibilita el apropiado ajuste de valores con asimetrías. La distribución asintótica salvaguarda este problema.

Figura 4. Coeficiente de Asimetría de la distribución estandarizada en función de *n* y *M*

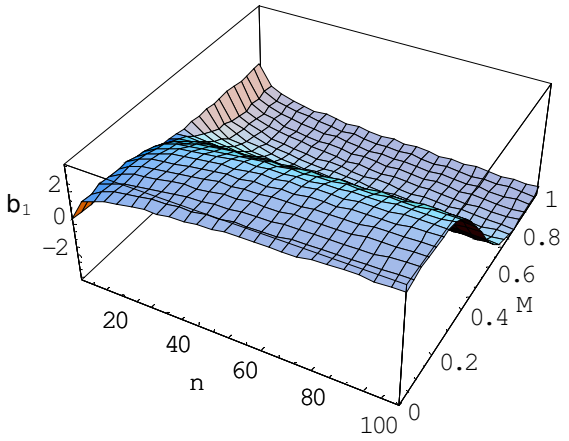
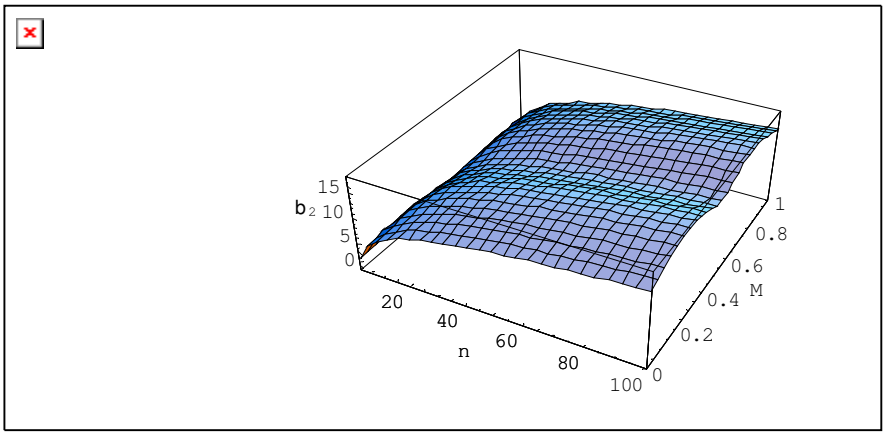


Figura 5. Coeficiente de curtosis de la distribución asintótica estandarizada (A) *M*=0.75 (azul), *M*=0.1 (verde), *M*=0.96 (negro), *M*=0.5 (rojo) (B) en función de *n* y *M*.



En la Figura 5 (A) se comprueba que cuando *n* toma sus valores mínimos el coeficiente que presenta la distribución es inferior al de la distribución normal, pero a medida que aumenta el valor del parámetro se tiene que el coeficiente presenta un crecimiento considerable, llegando a obtener unos valores de aproximadamente 17.5 para $n=100$. Teniendo en cuenta la Figura 5 (B) cuando *M* alcanza sus valores extremos el coeficiente disminuye aunque no de forma demasiado llamativa cuando *M* toma el valor 0.5.

Este hecho conlleva a la realización del estudio con más profundidad sobre el coeficiente haciendo hincapié sobre los límites en los que se mueve cuando se centra la atención los parámetros que describen la distribución. Teniendo en cuenta la expresión que describe dicho coeficiente se ha comprobado que éste puede llegar a alcanzar valores muy elevados cuando el valor del parámetro *n* tiende a infinito.

Es conocido que uno de los grandes problemas dentro del campo financiero es conseguir describir una distribución que presente un coeficiente de curtosis tan elevado como el que contiene la distribución empírica de los datos. La distribución asintótica se presenta como una

buena candidata para solventar dicho inconveniente ya que puede llegar a alcanzar elevados valores de curtosis superando a las distribuciones aplicadas tradicionalmente.

LA DISTRIBUCIÓN MIXTURA ASINTÓTICA RECTANGULAR

Partiendo de los estudios realizados sobre el desarrollo de mixturas entre distribuciones, Grant et *al.* (2006), Atkinson et al., (2006), Steele y Huber, (2004); McCullagh y Nelder, (1989), Mitchell y Zmud (1999) y Gelman et al, (2003), que permiten una mayor flexibilidad a la hora de ajustar los datos, Hahn (2008) presenta la mixtura entre las distribuciones beta y rectangular. Dicha mixtura consigue describir una distribución con unas colas más pesadas lo que la distingue de otras distribuciones tradicionalmente empleadas.

A continuación, se presenta, siguiendo la idea de Hahn, la generación de la distribución asintótica rectangular (A-R). El objetivo principal de dicha mixtura es combinar el gran apuntamiento que presenta la distribución asintótica con la elevación de las colas de ésta con la ayuda de la distribución rectangular. Así se consigue describir una distribución que solventa los problemas de apuntamiento y outlier presentados por la mayoría de las distribuciones utilizadas en el ajuste de series de rendimientos financieros.

Para introducir la expresión de la función de densidad asociada a la distribución A-R partimos del sistema generador de van Dorp y Kotz, para ello a partir de la función de densidad generadora de dicha distribución se puede describir la función de densidad (7) y la función de distribución (8) asociada por:

$$f\left(t \mid M\right)=\begin{cases}\dfrac{A(n)+\dfrac{1}{\dfrac{t}{M}-B(n)}}{A(n)+L(C(n))}\theta+(1-\theta) & 0\le t\le M\\[2ex] \dfrac{A(n)+\dfrac{1}{\dfrac{1-t}{1-M}-B(n)}}{A(n)+L(C(n))}\theta+(1-\theta) & M\le t\le 1\end{cases} \tag{7}$$

donde $A(n)=\left(1-\frac{1}{n}\right)^{\frac{1}{n}}$; $B(n)=\left(1+\frac{1}{n}\right)^{\frac{1}{n}}$; $C(n)=1-\left(1+\frac{1}{n}\right)^{\frac{-1}{n}}$ y θ es el parámetro que nos ayuda a describir la mixtura entre ambas distribuciones y el cuál solamente puede tomar valores entre 0 y 1. En el caso particular de que dicho parámetro tome el valor mínimo, es decir, 0 entonces se obtiene la expresión de la función de densidad correspondiente a la distribución rectangular. Sin embargo, si $\theta=1$ se tiene la función de densidad de la distribución presentada en el punto 1 es este trabajo, la distribución asintótica.

Con respecto a la función de distribución se tiene:

$$F\left(t \mid M\right)=\begin{cases}\dfrac{A(n)\dfrac{t}{M}+\left[\left(1-\theta\right)\dfrac{t}{M}L(C(n))+\theta L\left(1-\dfrac{t}{M}\dfrac{1}{B(n)}\right)\right]}{\dfrac{1}{M}\left(A(n)+L(C(n))\right)} & 0\le t\le M\\[2ex] 1-\dfrac{A(n)\left(\dfrac{1-t}{1-M}\right)+\left[\left(1-\theta\right)\left(\dfrac{1-t}{1-M}\right)L(C(n))+\theta L\left(1-\dfrac{1-t}{1-M}\dfrac{1}{B(n)}\right)\right]}{\dfrac{1}{1-M}\left(A(n)+L(C(n))\right)} & M\le t\le 1\end{cases} \tag{8}$$

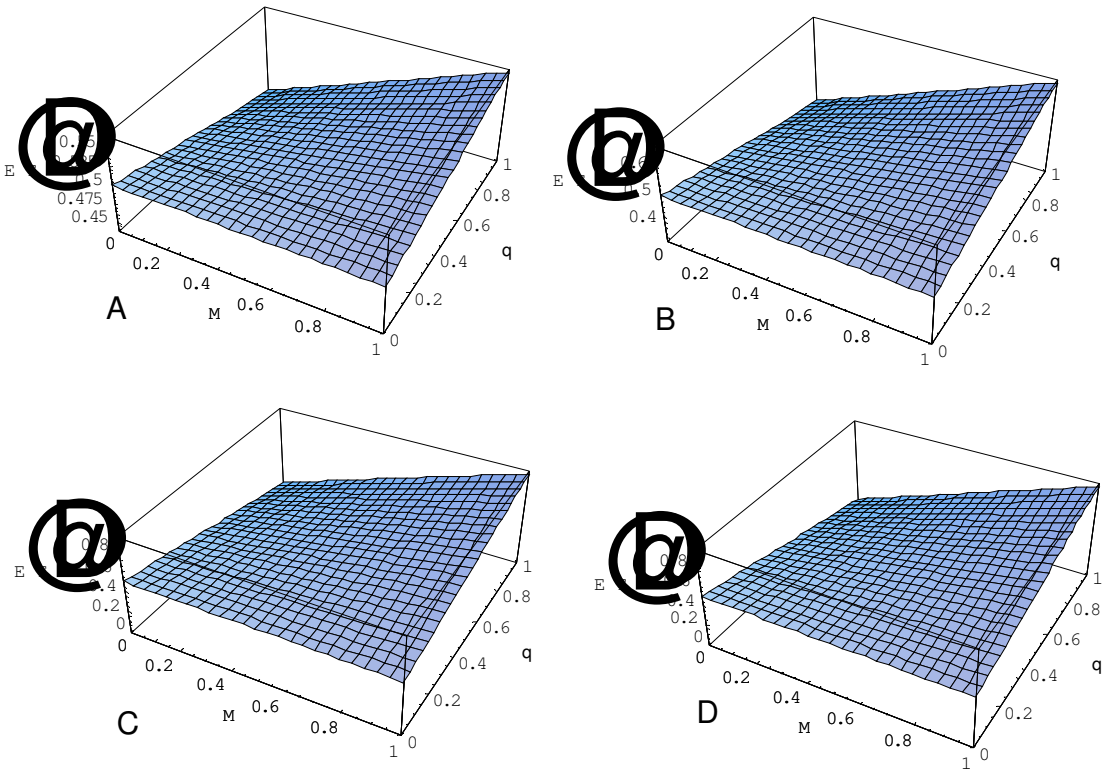
Teniendo en cuenta que en realizar la construcción de la distribución mixtura surge de combinar la distribución rectangular y la distribución asintótica, se verifica que los momentos de la mixtura A-R vendrán dados por:

$$E[T_{AR}] = \theta E[T_A] + (1-\theta) E[T_R] \tag{9}$$

$$Var[T_{AR}] = \theta^2 Var[T_A] + (1-\theta)^2 Var[T_R] \tag{10}$$

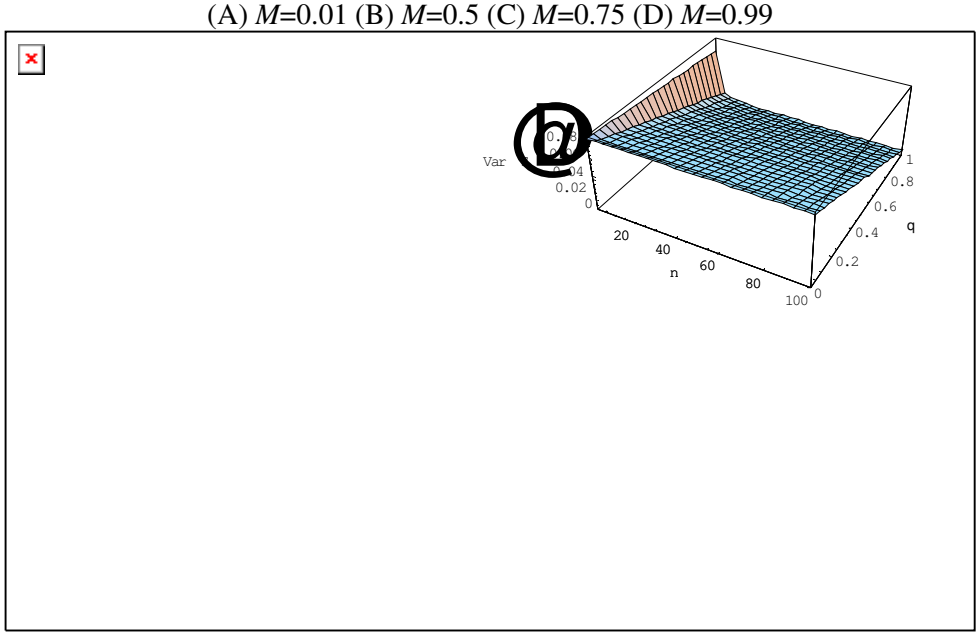
y cuyas representaciones gráficas corresponden a las Figuras 6 y 7 respectivamente

Figura 6. Valor Esperado de la distribución A-R en función M y θ
(A) *n*=1.1 (B) *n*=1.5 (C) *n*=15 (D) *n*=100
(B)



Los resultados obtenidos sobre el valor esperado de la distribución indican que la media se mantiene constante cuando θ toma valores próximos a cero, pero a medida que el parámetro se aproxima a 1 el valor esperado sufre un crecimiento aunque no demasiado elevado ya que el máximo valor que alcanza es de aproximadamente 0.8.

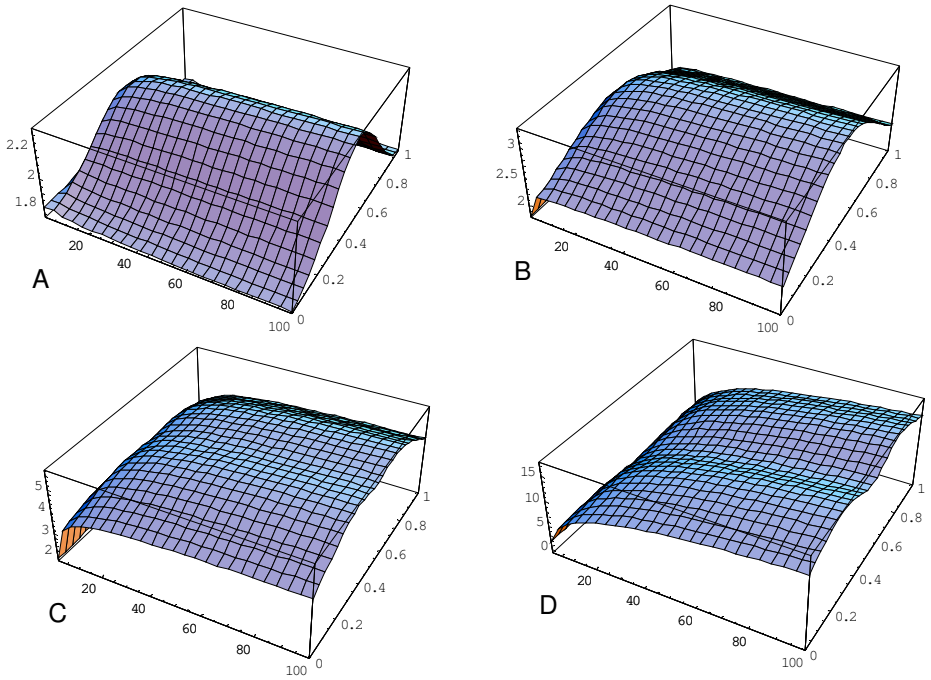
Figura 7. Varianza distribución A-R en función *n* y θ
(A) *M*=0.01 (B) *M*=0.5 (C) *M*=0.75 (D) *M*=0.99



Según la información que se puede obtener de la Figura 7, se afirmar que el rango de valores que toma la varianza de la distribución mixtura es mayor cuando la moda estandarizada toma los valores extremos, pero sin embargo, si se consideran los casos particulares en el que $M = 0.5$ o $M = 0.75$, por ejemplo, el rango de movilidad de la varianza disminuye hasta alcanzar el valor 0.08, como valor máximo.

Centrando nuestra atención en el estudio del coeficiente de curtosis se observa en la Figura 8 que a medida que aumentamos el valor del parámetro θ, el coeficiente crece hasta alcanzar aproximadamente 15. En el caso particular en que θ tome los valores extremos entonces se describen las figuras correspondientes a los coeficientes de las distribuciones utilizadas para realizar la mixtura.

Figura 8. Coeficiente de curtosis de la distribución mixtura A-R
(A) $\theta = 0.25$ (B) $\theta = 0.5$ (C) $\theta = 0.75$ (D) $\theta = 0.99$



ELICITACION MIXTURA TSP-RECTANGULAR

La elicitación de la mixtura A-R, presenta algunas ventajas respecto a la elicitación de la mixtura beta-rectangular (Hahn, 2008), por una parte le es aplicable todo el proceso de elicitación descrito en el trabajo de Hahn y por otra parte la mixtura A-R, al contrario que la mixtura beta-R, tiene función de distribución explicita G. Esto permite que dados dos cuantiles aportados por el experto y utilizando la función de distribución se obtendrán las estimaciones de los parámetros n y θ.

En definitiva el uso de los percentiles, desde el punto de vista del trabajo requerido a los expertos ha sido resaltado por numerosos autores; Selvidge (1980) propone los percentiles 0,10 y 0,9 como mas fiables que los extremo 0,01 y 0,99 y otros autores como Davidson y Cooper (1980), proponen unos percentiles intermedios como el 0,05 y el 0,95, Kotz y van Dorp (2006) desarrollan un procedimiento para elicitar la distribución TSP utilizando cuantiles aportados por el experto. Van Dorp et al. (2007) han desarrollado un procedimiento para estimar una distribución trapezoidal con un estado central uniforme en base a fractiles.

Por otra parte, la interpretación del parámetro θ en nuestras distribuciones mixtas no resulta en absoluto intuitivo. En este sentido autores como Moskowitz et al. (1979) defienden que la distribución beta puede ser ajustada de forma más apropiada con sus cuantiles en la colas que con sus cuantiles centrales, Lau y Somarajan (1995), defienden el uso de los cuantiles y proponen un métodos en dos etapas poniendo de manifiesto que existen muchas distribuciones que puede ajustarse a unos cuantiles dados, además de la distribución beta cuatriparametrica Lau et al. (1996), Lau et al. (1998) defiende el uso de distribuciones alternativas a la beta y de mas de tres cuantiles Lau y Lau (1998) presentan una nota mejorando las estimaciones de Keefer y Verdini (1993), Johnson (1998) insiste en el uso de los fractiles y en la mayor acuracidad de la estimación.

En definitiva podemos afirmar que desde los inicios de la metodología PERT, el uso de los fractiles ha sido considerado como una alternativa mejor que la estimación directa de los parámetros (véase por ejemplo Hampton et al. (1973), Chesley (1975)), considerándose el método de los fractiles como el método más directo y más utilizado para estimar los parámetros de la distribución.

A continuaciones se comprueba que el parámetro θ se puede obtener a partir de un percentil (p, a_p), siendo $G(a_p) = p$.

Como sabemos, la función de distribución estandarizada esta recogida en la expresión (8), en base a ello podemos enunciar las siguientes proposiciones.

Proposición.- Si (p, a_p) y (q, a_q) son dos percentiles de tal modo que $a_p \le M \le a_q$, entonces deberá de cumplirse que $0 \le G(a_p) = p \le a_p \le M \le a_q \le G(a_q) = q \le 1$

Para comprobar la anterior proposición bastará con sustituir en la ecuación (8) anterior para obtener:

$$p = \frac{A(n)\frac{a_p}{M} + \left[(1-\theta)\frac{a_p}{M}L(C(n)) + \theta L\left(1 - \frac{a_p}{M}\frac{1}{B(n)}\right)\right]}{\frac{1}{M}\left(A(n) + L(C(n))\right)} = \frac{a_p}{M}\frac{A(n) + (1-\theta)L(C(n)) + \theta\frac{M}{a_p}L\left(1 - \frac{a_p}{M}\frac{1}{B(n)}\right)}{\frac{1}{M}\left(A(n) + L(C(n))\right)}$$

$$\le \frac{a_p}{M}\frac{A(n) + (1-\theta)L(C(n)) + \theta\frac{M}{M}L\left(1 - \frac{M}{M}\frac{1}{B(n)}\right)}{\frac{1}{M}\left(A(n) + L(C(n))\right)} = a_p$$

La desigualdad de la derecha se puede obtener del mismo modo.

En base al resultado anterior será posible elicitar la distribución mixtura a partir de un percentil, bastara para ello que se cumpla la condición de la proposición anterior y podremos despejar fácilmente θ, de la expresión (8), en cualquiera de sus ramas:

$$\theta = \begin{cases} \dfrac{\left(a_p - p\right)\left[A(n) + L\left[C(n)\right]\right]}{a_p L\left[C(n)\right] - ML\left[1 - \dfrac{a_p}{M}\dfrac{1}{B(n)}\right]} & p \le a_p \le M \\ \dfrac{\left(a_q - q\right)\left[A(n) + L\left[C(n)\right]\right]}{\left(a_q - 1\right)L\left[C(n)\right] + \left(1 - M\right)L\left[1 - \dfrac{1-a_q}{1-M}\dfrac{1}{B(n)}\right]} & M \le a_q \le q \end{cases} \tag{11}$$

Por lo tanto, siempre que tengamos un percentil a la izquierda o a la derecha de *M*, que cumpla las condiciones establecidas en la proposición, y resolviendo el sistema:

$$\frac{\left(a_p - p\right)\left[A(n) + L\left[C(n)\right]\right]}{a_p L\left[C(n)\right] - ML\left[1 - \dfrac{a_p}{M}\dfrac{1}{B(n)}\right]} = \frac{\left(a_q - q\right)\left[A(n) + L\left[C(n)\right]\right]}{\left(a_q - 1\right)L\left[C(n)\right] + \left(1 - M\right)L\left[1 - \dfrac{1-a_q}{1-M}\dfrac{1}{B(n)}\right]} \tag{12}$$

podremos obtener la estimación de los parámetros *n* y θ.

Se ha de señalar que desafortunadamente este sistema no siempre tiene solución, pero cuando es posible resolverlo, permite obtener una elicitación de la distribución subyacente a partir de percentiles, situación que en la literatura del PERT esta considerada como más cercana a la realidad en este sentido puede citarse los trabajos de Moskowitz et al. (1979), Selvidge (1980), Davinson y Cooper (1980), Alpert y Raihfa (1982) Keefer y Verdini (1993), Lau y Somarajan (1995), Lau et al. (1996). Lau et al. (1998), Johnson (1998), Kotz y van Dorp (2006) y van Dorp et al. (2007).

CONCLUSIONES

Se ha presentado una nueva distribución con un amplio rango de curtosis que puede ser una candidata adecuada para determinados casos empíricos en el ámbito financiero. Además, se ha construido una mixtura de dicha distribución con la distribución rectangular consiguiendo combinar asi altos valores de curtosis con colas pesadas. Se ha presentado el procedimiento de elicitación concluyendo que en el caso de la mixtura A-R se ha ganado en capacidad de elicitación directa de la distribución, ya que a partir de dos fractiles es posible obtener el parámetro θ, que mide la presencia de la distribución asintótica en la mixtura, y el parámetro *n*, que esta relacionado con la concentración y el apuntamiento de la distribución.

Actualmente se esta trabajando para encontrar unos datos empíricos que presenten un adecuado ajuste y se espera que próximamente puedan presentarse resultados y conclusiones satisfactorios.

BIBLIOGRAFÍA

ALPERT, M. AND RAIFFA, H. (1982) A progress report on the training of probability assessors, in Judgment Under Uncertainty: Heuristics and Biases, Kahneman, D., Slovic, P. and Tversky, A. (eds.), Cambridge University Press, New York, NY, pp. 294–305.

CHESLEY, G. R.. (1975) Elicitation of Subjective Probabilities.*The Accounting Review.*. Vol. 50, Iss. 2; pg. 325

CLARK. P.K. (1973) A subordinate stochastic process model with finite variance for speculative prices. Econometrica, 41:135–155,

DAVIDSON, L.B. AND COOPER,D.O. (1980) Implementing effective risk análisis at Getty Oil Company. *Interfaces,* 10, 62–75.

HAMPTON J.M, MOORE P.G. AND THOMAS H. (1973) Subjective probability and its measurement *Journal of the Royal Statistical Society Ser A Vol. 136, (1973) pp 21-42*

HAHN, E.D. (2008). *Mixture Densities for Project Management Activity Times: A Robust Approach to PERT.* European Journal of Operational Research,188, pp. 450-459.

JOHNSON D (1997) The triangular distributions as a proxy for the beta distribution in risk analysis. *TheStatistician* 46(3): 387–389

JOHNSON,.D. (1998) The robustness of mean and variance approximations in risk analysis. *The Journal of the Operational Research Society*. Oxford: Mar 1998. Vol. 49, Iss. 3; pg. 253, 10 pgs

KEEFER, D. AND VERDINI, W. (1993) Better estimation of PERT activity time parameters, *Management Science*, 39(9), 1086-1091

KOTZ S. VAN DORP J.R. (2006) A novel method for fitting unimodal continuous distributions on a bounded domain utilizing expert judgment estimates. *IIE Transactions* 38, 421–436

LAU, A., LAU, H. AND ZHANG, Y. (1996) A simple and logical alternative for making PERT time estimates. *IIE Transactions,* 28(3), 183-192.

LAU, H., LAU, A.H. AND HO, C. (1998) Improved moment-estimation formulas using more than three subjective fractiles. *Management Science*, 44(3), 346–350

LAU, H. AND SOMARAJAN, C. (1995) A proposal on improved procedures for estimating task±time distributions in PERT. *European Journal of Operational Research*, 85, 39-52.

MOSKOWITZ, HERBERT, BULLERS, WILLIAM I. (1979) Modified PERT versus Fractile Assessment of Subjective Probability Distributions.*Organizational Behavior and Human Performance*. Vol. 24, Iss. 2; pg. 167

SELVIDGE, J.E. (1980) Assessing the extremes of probability distributions by the fractile method. *Decision Sciences,* 11, 493–502.

THOMAS S. WALLSTEN; DAVID V BUDESCU. (1983) State of the Art Encoding Subjective Probabilities: a Psychological and psych.Management Science; 29, 2; pg. 151

TZUNG-CHENG HUANA,B, JAY BEAMANC, LIANG-HAN CHANGD,E, SHIH-YUN HSUF, (2008) Robust and alternative estimators for ''better'' estimates for expenditures and other ''long tail'' distributions *Tourism Management, 29, 795–806*

VAN DORP, J.R. AMITA SINGH AND THOMAS A. MAZZUCHI. (2006) The Doubly-Pareto Uniform Distribution with Applications in Uncertainty Análisis and Econometrics *Mediterranean Journal of Mathematics*. 3, 205–225

VAN DORP, RAMBAUD, PEREZ, PLEGUEZUELO. (2007) An elicitation procedure for generalized trapezoidal distribution with a uniform central stage.*Decision Análisis* Vol. 4, No. 3, pp. 156–166

VAN DORP, J.R. Y KOTZ, S. (2003). *Generalizations of Two-Sided Power Distributions and their Convolution*. Communications in Statistics: Theory and Methods, Vol. 32, Nº 9, pp. 1703-1723.

DETERMINACION DEL LUCRO CESANTE EN AMBIENTE DE INCERTIDUMBRE: ASPECTOS JURÍDICOS Y METODOLÓGICOS. UN CASO PRÁCTICO

JOSÉ GARCÍA PÉREZ
E-mail: jgarcia@ual.es
Departamento de Economía Aplicada
Universidad de Almería
JOSÉ MARTÍN GARCÍA GARCÍA
Universidad de Almería
CARMEN ROMERO MARTÍNEZ
I.E.S. Los Angeles

Resumen

La determinación del lucro cesante constituye un hito ampliamente desarrollado en países como Estados Unidos, Francia, Inglaterra, etc. En nuestro país son escasas las referencias bibliográficas que abordan este tema. El objeto fundamental de nuestro trabajo es elaborar y aplicar una metodología que permita obtener un valor para el lucro cesante en ausencia de datos que permitan la aplicación de otros modelos estadísticos. Las Cortes Generales de los Estados Unidos tienen publicado un manual de procedimiento donde figuran los métodos estadísticos y econométricos aplicables cuando se tienen suficientes datos. Sin embargo, en numerosas ocasiones no existen datos de la empresa al ser de reciente creación, ni datos del sector por tratarse de un sector emergente. Cuando se dan ambas circunstancias es habitual que no se disponga de la suficiente información para aplicar las técnicas estadísticas conocidas. El objetivo de este trabajo es resolver esta problemática presentando una metodología y una aplicación practica.

Palabras clave: Lucro cesante, Incertidumbre, riesgo, sectores emergentes.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

1. INTRODUCCIÓN

El daño resarcible no puede ser reducido a una única idea si no que esta integrado por una pluralidad de conceptos, Garnica Martín J.F.(2005), sin embargo ha constituido –en palabras del citado autor- "*una práctica muy generalizada en nuestra tradición jurídica el que nuestros tribunales no exigiesen , una verdadera distinción entre los diferentes conceptos"*.

El lucro cesante es, siguiendo la terminología del art. 106 del Código civil, la ganancia que se ha dejado de obtener como consecuencia del hecho del que se es responsable. En principio la jurisprudencia exige respecto de la determinación del lucro cesante dos cuestiones : certeza y restrictividad, respecto de la primera cuestión deseamos citar la sentencia STS 1ª de 8 de julio de 1996 *:" Las ganancias que pueden reclamarse son aquellas en las que concurre verosimilitud suficiente para poder ser reputadas como muy probables , en la mayor aproximación a su certeza definitiva*". Debemos apreciar que se utilizan dos términos habituales en estadística VEROSIMILITUD y MUY PROBABLES , y queremos recordar en este momento que las Cortes de Justicia de los EEUU tienen publicado un memorandum sobre las técnicas estadísticas que pueden ser consideradas por lo tribunales. Pero es que en una sentencia mucho mas reciente del SSTS 26 se septiembre de 2002 se afirma abiertamente que la determinación del lucro cesante se funda en un *juicio de probabilidad* y en la sentencia STS de 29 de diciembre de 2001 (RJ 2001/1474) se precisa que debe de tratarse de una " *cierta probabilidad objetiva*", en conclusión la jurisprudencia no exige certeza absoluta sobre la determinación de las ganancias futuras frustradas .

El profesor Garnica, citado anteriormente, presenta en su artículo un trabajo en el que trata con mucho acierto aspectos tales como la prueba de la existencia del lucro, la prueba de la cuantía de la ganancia frustrada, el lucro cesante derivado de la paralización de vehículos industriales y las ganancias frutadas derivadas de la necesidad de cierre de un negocio.

A pesar de que el trabajo citado recoge una bibliografía muy seleccionada, no encontramos referencias a la metodología a emplear. A pesar de ello resulta evidente que las técnicas estadísticas, y econometritas están apuntadas en las diferentes sentencias que hemos mencionado anteriormente

La determinación del lucro cesante (ganancias frustradas) es un tópico en la literatura empresarial y uno de los cometidos de los economistas forense, cuya actuación esta regulada por el Consejo Superior de Colegios de Economistas a través del REFOR, no obstante no resulta sencillo encontrar publicaciones, tanto en el ámbito científico como en el ámbito profesional, sobre la metodología a emplear, la búsqueda de artículos y contribuciones profesionales devuelve pocos resultados en las revistas y trabajos nacionales .

En el ámbito internacional , podemos citar entre otros muchos los siguientes trabajos: Stewart B. (2005) donde se analiza el papel del mercado en la determinación del lucro cesante, Trout R. (2000) utiliza técnicas propias del análisis de inversiones , Anabela Botelho (2001) utiliza la Teoría de juegos para la evaluación de evidencias complejas; las técnicas macroeconómicas son utilizadas por Antón and Yao (2006) , y finalmente queremos resaltar la critica que realizan Bowles and Lewis (1999) sobre la inestabilidad de los procedimientos y usos. Podiamos elegir algunas citas pero no podemos de jar de citar el trabajo de Gaughan y Thornton (1993) que en su libro *Litigation in Economics* presentan catorce contribuciones , de diferentes autores, que en nuestra opinión vienen a situar las bases de la economía forense. Por otra parte la diferencia de la situación de la economía forense entre nuestro pais y los países "avanzados" se puede poner de manifiesto con la lectura del manual: *Referente Manual on Scientific Evidence* (2000) editado **Federal Judicial Center** , en el que establecen lo métodos estadísticos que pueden ser utilizados para obtener evidencia, es decir las Cortes Federales de Justicia de los EEUU, establecen (como hemos señalado antes) en un manual cuales han de ser lo mecanismos estadísticos que deben de emplear los economistas forenses para establecer evidencia , tanto el la determinación de daños *"damages"* , como en la determinación de *"lost profit*" o lucro cesante.

Hemos de señalar que en los trabajos recogidos en el texto de Gaughan y Thornton (1993) las técnicas cuantitativas van desde las técnicas financieras que permiten obtener el valor actual de una renta futura, las técnicas econométricas con el uso de modelos hedònicos, o el uso de variables ficticias, así como técnicas estadísticas relacionadas con el análisis de supervivencia, o métodos de valoración de empresas como el VAN, métodos basados en análisis de series temporales y tecnicas econométricas de predicción mediante modelos lineales, los modelos de regresión , los modelos probit y logit son también utilizados en otros trabajos.

Queremos resaltar que la diferencia entre los trabajos comentados hasta ahora y el trabajo que aquí se presenta la podemos preciar diciendo que todos los trabajos anteriores se han realizado en ambiente riesgo, es decir con la existencia de datos que permiten por tecnicas estadísticas o econométricas obtener estimaciones en base a las cuales se pueden realizar predicciones con las cuales se va a determinar el lucro cesante, mientras que en nuestro caso, la ausencia de datos en la característica que diferencia los métodos presentados en este trabajo de los métodos conocidos hasta ahora .

Para que podamos decir que no existen datos (es decir que estamos en ambiente de incertidumbre) deben de darse simultáneamente tres condiciones:

a) No existen datos históricos de la empresa
b) No existen datos históricos o empresas similares a la empresa en cuestión
c) No existen datos del sector en el que podemos ubicar nuestra empresa.

Cuando afirmamos que no existen datos queremos decir que estos aunque se puedan obtener no permiten por si solos realizar estimaciones y predicciones que resulten determinantes

El objeto del presente trabajo va a consistir en la valoración del lucro cesante o perjuicio económico, en concepto de rendimientos dejados de percibir por la *empresa PERTURBACIONES INESPERADAS SL.*, como consecuencia de la rescisión del contrato que ésta mantenía con una de las principales distribuidoras de telefonía móvil de nuestro país **DAVOFENA.** Por este contrato, la empresa **PERTURBACIONES INESPERADAS SL**, se había convertido en distribuidor de los productos de **DAVOFENA**, integrada en la red de distribuidores de la misma que operaban conjuntamente bajo el nombre comercial *COMRED.* Con fecha 7 de diciembre de 2005, **DAVOFENA, S.A.,** dejó sin efecto dicho contrato, lo que ha supuesto, en principio un ***LUCRO CESANTE*** o perjuicio económico para la empresa, que es el objeto de valoración en este trabajo cuyo único alcance pretende ser académico.

Se parte de la base, de que la valoración del *lucro cesante* a través de la **valoración de la empresa**, es un acto básicamente subjetivo. La Asociación Española de Contabilidad y Administración de Empresas (AECA) que advierte esta premisa, le ha dedicado cinco Documentos (Publicaciones de Monografías), a la valoración de PYMES.

AECA, es el único Organismo que, formando parte de la Unión Europea de Expertos Contables, Económicos y Financieros, emite normas en España, normas que son tomadas como propias, tanto por el Instituto de Contabilidad y Auditoria de Cuentas (ICAC), dependiente del Ministerio de Economía y Hacienda), como por el REGISTRO DE ECONOMISTAS AUDITORES CONTABLES (REAC) y por REFOR (Registro de Economistas Forenses) y las restantes Corporaciones y Colegios Oficiales.

El Documento nº 3, fechado en julio de 1983 (Métodos Prácticos de Valoración de Empresas) y, que a pesar de su antigüedad no ha sufrido variación hasta la fecha, manteniendo integra su vigencia, desarrolla cinco métodos simples, cinco maneras distintas de calcular el valor de una empresa y otros cuatro compuestos, recomendando que el PERITO, calcule varios de ellos e incluso que los combine asignándoles un coeficiente de ponderación, basándose en que se trate de empresa en funcionamiento o empresa en liquidación, en el valor contable de los activos, en los posibles cambios de legislación por el sector, en la posible entrada en vigor de nuevos contratos, y en cualquier otro factor que pueda hacer variar los beneficios esperados, según su criterio en la valoración final .

Por otra parte, y claramente dirigidos a los profesionales que efectúen valoración de empresas, AECA detalla entre otras las siguientes "Reglas Generales a Observar":

a) Evaluación Global.- La empresa debe considerarse un todo y no la suma de diferentes valores.

b) Consideración de futuro.- ***El lucro cesante de una empresa, esta siempre en función de sus beneficios futuros*** así como de su previsible estructura de activo y pasivo. Los resultados obtenidos en períodos anteriores y los activos netos en la fecha de valoración son importantes elementos orientadores para el cálculo.

c) Definición de criterios de valoración claros y Precisos.- Éstos han de ser claramente expuestos en el informe de valoración.

d) Análisis de un cierto número de cuentas de resultados anteriores.- Para determinar la evolución futura, así como para la mejor información del lector del informe, es conveniente en principio analizar los resultados, al menos, de los tres últimos ejercicios. Solo podrá prescindirse de este análisis, a título excepcional, cuando el futuro no tenga ninguna relación con los resultados pasados .Si se renuncia al mencionado análisis, convendrá explicar las razones en el informe.

e) Presentación del Informe de valoración.- El informe debe incluir entre otras la siguiente información:

- Origen de sus informaciones.
- Especificación de la naturaleza de las informaciones recibidas.
- Análisis – resumen de la información empleada.
- Indicación de lo Métodos de Valoración Empleados y de sus limitaciones.
- Puesta a disposición y calidad de los datos utilizados.

Este será el índice que seguiremos en el presente trabajo, de manera que en el apartado 2 presentaremos el Origen de las informaciones a utilizar, en el apartado 3 comentaremos y analizaremos la información empleadas, en el apartado 4 presentaremos los métodos utilizados y comentaremos sus limitaciones

2.-ORIGEN DE LAS INFORMACIONES, NATURALEZA DE LAS MISMAS Y CALIDAD.-

Las informaciones que hemos utilizado para la confección de este informe son básicamente las siguientes:

A) Cuentas anuales y Memoria de Gestión de los ejercicio 2002 a 2005de la empresa "PERTURBACIONES INESPERADAS SL."
Entendemos que esta información, en lo que se refiere a los beneficios pasados tiene tendencia a la baja, ya que la empresa durante los primeros ejercicios de su actividad, ha incurrido en gastos extraordinarios de comienzo, y por otro lado, ha estado creando una cartera de clientes que han generando ingresos a los largo del tiempo, y que lo hubieran seguido haciendo de no producirse el cese del contrato.

B) Relación de empresas incluidas en COMRED, con especificación de la cuota de participación en el volumen global de facturación. Se trata de empresas que tenían y tienen, en algunos casos, la misma actividad que "PERTURBACIONES INESPERADAS SL.", y que por lo tanto, nos pueden servir de referencia para estimar la evolución del negocio.

C) Informe de las cuentas anuales obtenidos a través de AXESOR, de las empresas elegidas como más significativas de entre las del listado anterior, por tener una cuota de participación más próxima a la de la empresa objeto de estudio.

D) Fotocopia de la tabla de módulos correspondiente a la actividad ejercida por la empresa para el ejercicio 2005. La empresa ha venido ejerciendo su actividad bajo el epígrafe 653.2, que la Agencia Estatal de Administración Tributaria, tiene catalogado como posible estimación objetiva y cuya aplicación vamos a considerar como uno de los posibles escenarios.

E) Copia de las facturas de electricidad de los dos últimos ejercicios en los que la empresa a sido distribuidor de VODAFONE (al objeto de poder aplicar los coeficientes de módulos).

F) Documentos TC2 en los que aparecen los trabajadores de la empresa, así como informe de ocupación de la misma para cada ejercicio del periodo 2000 – 2005 (Para aplicar módulos).

D) Documentación justificativa de los locales abiertos al público por la empresa en los últimos ejercicios (con objeto de aplicar módulos).

3.- ANÁLISIS – RESUMEN DE LA INFORMACIÓN EMPLEADA.

Se trata de obtener diferentes escenarios de futuro a partir de la información anterior.

APARTADO I.- ESTIMACIÓN OBJETIVA.

La empresa venía ejerciendo su actividad bajo el epígrafe 653.2, dicho epígrafe está incluido en el sistema de estimación objetiva, para el que la Agencia Estatal de Administración Tributaria, tiene previstos unos rendimientos en base a datos o índices objetivos que pueden servir de referencia y que van a servir para establecer un posible escenario.

I.1. Tablas de estimación objetiva. (Escenarios I y II).

Actividad: 653.2 Comercio al por menor de aparatos eléctricos.
Tablas aprobadas para el ejercicio 2005. La utilización de estas tablas nos evita tener que homogeneizar los datos para su comparación en términos reales.

TABLA 1

Módulo	Unidad	Importe
Personal asalariado	Persona TC	2.884,77
Personal no asalariado	Autónomos TC	14.656,86
Consumo energía eléctrica	100 kw	100,78
Sup. Local Independiente m2		36,53
Sup. Local no independiente	m2	13,37

I.2. Datos relativos a la empresa.

La empresa ha evolucionado de forma muy contundente en los años 2002 - 2005, alcanzando una posición de liderazgo en el sector de la telefonía en la provincia donde ejerce, y situándose en una buena posición en todas aquellas plazas en las que ha entrado.

Los datos económicos y técnicos de la empresa de los dos últimos años, en lo relativo a las unidades de módulo, arrojan los siguientes resultados:

TABLA 2

	2004	2005
PERSONAL ASALARIADO	28,5	37,5
PERSONAL NO ASALARIADO	1	1
CONSUMO DE ENERGÍA ELÉCTRICA	1287,5	1788,93
SUPERFICIE DE LOCAL INDEPENDIENTE	190,94	190,94

I.3. Resultados estimados para el Escenario I:

A la vista de los dos puntos anteriores, a la empresa se le estima un resultado neto de 233.602,09 € para el ejercicio 2004 y de 310.099,14 € para 2005, lo que supone un incremento del 32,74 % de un ejercicio a otro, pero vamos a considerar que la empresa ya ha alcanzado una estructura estable y que únicamente va a seguir creciendo al ritmo del IPC, el cual vamos a considerar un 3%.

Con estas estimaciones, podemos proyectar la evolución del rendimiento neto estimado para los próximos cinco años, tal y como se muestra en el siguiente cuadro. Así mismo, se determinan los importes en términos reales de 2006, usando una tasa del 5%

TABLA 3

AÑO	2006	2007	2008	2009	2010
RDTO. NETO	319.402,09 €	328.984,15 €	338.853,68 €	349.019,29 €	359.489,87 €
RDTO. NETO'	319.402,09 €	313.318,24 €	307.350,28 €	301.495,99 €	295.753,21 €

Con esta primera estimación, considerando que el crecimiento de la empresa, medido en módulos, se mantiene constante en términos reales, obtenemos que la suma del Rendimiento Neto Actualizado para el horizonte temporal de cinco años marcado, asciende a **1.537.319,80 €**. Este escenario supondría, en nuestra opinión la mejor situación de las situaciones más probables.

I.4. Resultados estimados para el Escenario II:

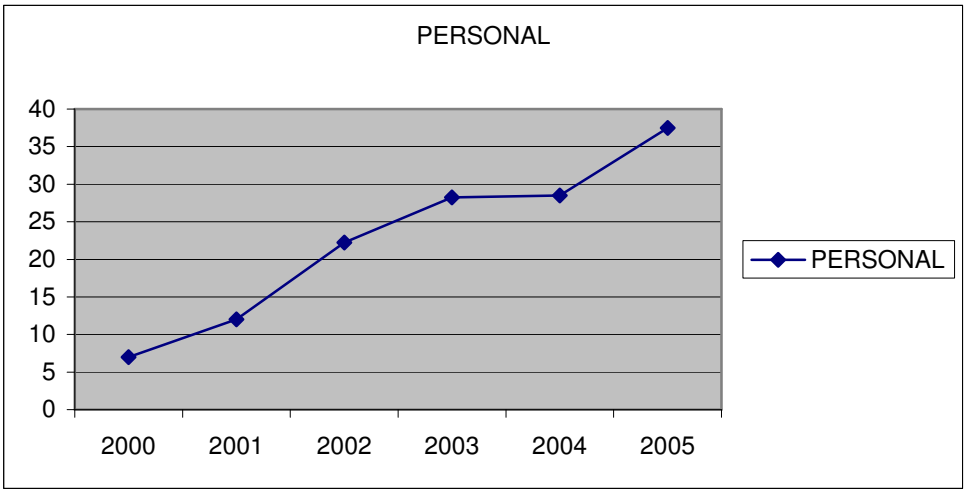
Por otro lado, podríamos considerar que el crecimiento ha sido proporcional al experimentado por el número medio de empleados, ya que disponemos de esta información

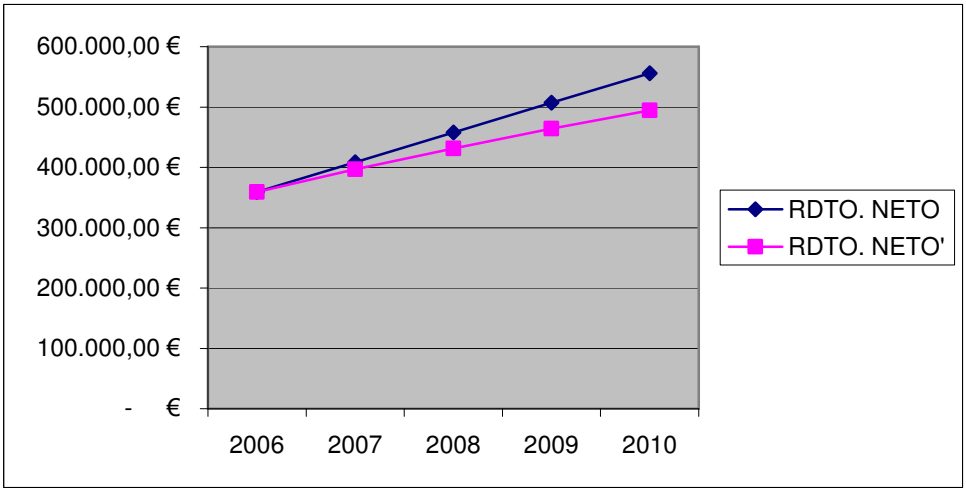
para el periodo 2000 – 2005, y vemos que el crecimiento se ajusta linealmente con un coeficiente de determinación lineal muy bueno (95'88 %). En este caso, el crecimiento esperado de la empresa para el horizonte temporal marcado, suponiendo que el crecimiento de los beneficios sigue la misma trayectoria que el incremento del personal, es el siguiente:

TABLA 4

EJERCICIO	2006	2007	2008	2009	2010
CRECIMIENTO	15,87 %	13,7 %	12,05 %	10,75 %	9,71 %
RDTO. NETO	359.311,87 €	408.537,60 €	457.766,38 €	506.976,26 €	556.203,66 €
RDTO. NETO'	359.311,87 €	389.083,43 €	415.207,60 €	437.945,15 €	457.590,13 €

Del cuadro anterior, se deduce que, en base a las condiciones expuestas, la suma del Rendimiento Neto Actualizado para el horizonte temporal de cinco años marcado, asciende a **2.059.138,18€.** La situación descrita es la OPTIMISTA, es decir, la mejor de todas las situaciones.





APARTADO II.- BENEFICIOS DEL SECTOR. ESCENARIO III.

La empresa se encontraba dentro de un conjunto de empresas colaboradoras de DAVOFANE que actuaban bajo la denominación COMRED. Cada una de estas empresas, con

una cuota de participación, según consta en el anexo 2. De estas empresas, hemos elegido aquellas con un porcentaje de participación en ingresos similar a PERTURBACIONES INESPERADAS SL, con objeto de analizar cuales han sido los beneficios medios declarados por las mismas.

Se ha elegido una muestra, eliminando de la misma aquellas empresas cuya información no se encontraba disponible en el Registro Mercantil. A continuación mostramos un resumen de la información manejada, adjuntando en el anexo III, los informes completos obtenidos del Registro Mercantil.

EMPRESA	2004		
	VENTAS	BENEFICIO	%
C M PHONE ACCESORIES	4.743.035,00 €	79.095,00 €	1,67%
MARC OPERATOR, S.L.	2.330.896,00 €	44.260,00 €	1,90%
MITEXNO, S.L.			
NIZA MOVILES, S.L.	1.475.756,00 €	26.131,00 €	1,77%
O BREOGAN COMUNICACIONS, S.L.	1.587.997,00 €	7.946,00 €	0,50%
REDFREECOM, S.L.	4.298.550,00 €	60.140,00 €	1,40%
TELEOPERATOR COMUNICACIONES, S.L.			

EMPRESA	2003		
	VENTAS	BENEFICIO	%
C M PHONE ACCESORIES	2.901.097,00 €	75.376,00 €	2,60%
MARC OPERATOR, S.L.	1.541.474,00 €	8.410,00 €	0,55%
MITEXNO, S.L.	2.132.251,00 €	104.609,00 €	4,91%
NIZA MOVILES, S.L.	1.051.620,00 €	11.154,00 €	1,06%
O BREOGAN COMUNICACIONS, S.L.	729.755,00 €	2.629,00 €	0,36%
REDFREECOM, S.L.	2.938.031,00 €	66.113,00 €	2,25%
TELEOPERATOR COMUNICACIONES, S.L.			

EMPRESA	2002		
	VENTAS	BENEFICIO	%
C M PHONE ACCESORIES	1.738.574,00 €	55.009,00 €	3,16%
MARC OPERATOR, S.L.	485.120,00 €	1.573,00 €	0,32%
MITEXNO, S.L.	1.571.824,00 €	89.665,00 €	5,70%
NIZA MOVILES, S.L.	1.015.901,00 €	14.170,00 €	1,39%
O BREOGAN COMUNICACIONS, S.L.			
REDFREECOM, S.L.	1.853.438,00 €	84.156,00 €	4,54%
TELEOPERATOR COMUNICACIONES, S.L.	1.005.356,00 €	39.473,00 €	3,93%

EMPRESA	2001		
	VENTAS	BENEFICIO	%
C M PHONE ACCESORIES	1.274.010,00 €	49.922,00 €	3,92%
MARC OPERATOR, S.L.	289.864,00 €	11.606,00 €	4,00%
MITEXNO, S.L.			
NIZA MOVILES, S.L.			
O BREOGAN COMUNICACIONS, S.L.			
REDFREECOM, S.L.	1.585.049,00 €	87.783,00 €	5,54%
TELEOPERATOR COMUNICACIONES, S.L.	893.918,00 €	67.183,00 €	7,52%

EMPRESA	2000		
	VENTAS	BENEFICIO	%
C M PHONE ACCESORIES	793.398,00 €	18.050,00 €	2,28%
MARC OPERATOR, S.L.			
MITEXNO, S.L.	2.616.200,00 €	- 64.315,00 €	-2,46%
NIZA MOVILES, S.L.			
O BREOGAN COMUNICACIONS, S.L.			
REDFREECOM, S.L.	2.016.209,00 €	115.098,00 €	5,71%
TELEOPERATOR COMUNICACIONES, S.L.	884.349,00 €	36.133,00 €	4,09%

EMPRESA	1999		
	VENTAS	BENEFICIO	%
C M PHONE ACCESORIES			
MARC OPERATOR, S.L.			
MITEXNO, S.L.	2.817.365,00 €	65.100,00 €	2,31%
NIZA MOVILES, S.L.			
O BREOGAN COMUNICACIONS, S.L.			
REDFREECOM, S.L.			
TELEOPERATOR COMUNICACIONES, S.L.	798.796,00 €	35.684,00 €	4,47%

A tenor de los datos anteriores, el beneficio medio obtenido por las empresas descritas anteriormente, durante el periodo 1999-2004, ha sido del 2,86 % de los ingresos.

Teniendo en consideración esta cifra, para determinar cual hubiera sido el beneficio obtenido por la empresa en los próximo cinco años, vamos a determinar cual ha sido el crecimiento de las ventas en los últimos años con objeto de proyectar los ingresos y obtener, de esta manera las ventas y beneficio esperados para el horizonte temporal objeto de estudio. Para calcular los beneficio en términos reales, utilizaremos, al igual que en el caso anterior, una inflación del 3 %.

TABLA 5

EJERCICIO	VENTAS
2001	2.733.159,46 €
2002	4.404.061,76 €
2003	3.909.431,35 €
2004	4.640.447,39 €
2005	5.596.834,82 €

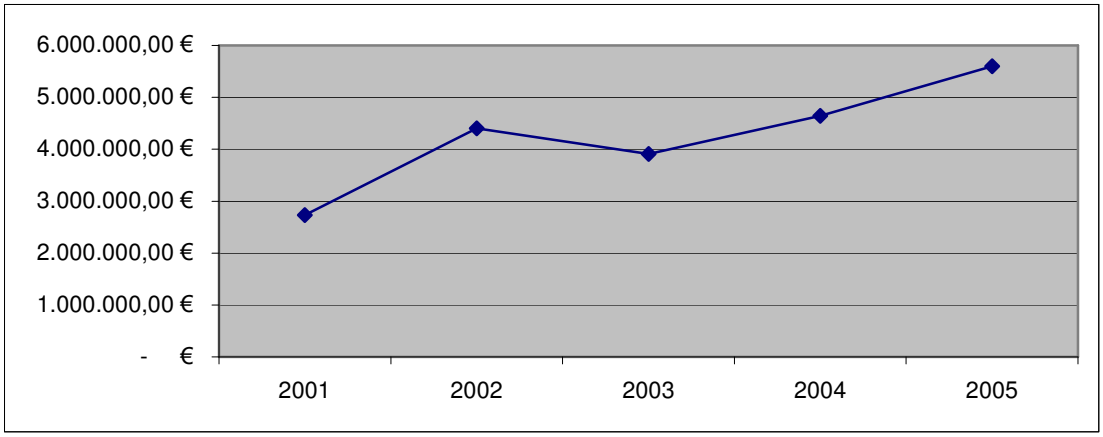


TABLA 6

EJERCICIO	VENTAS	RENDIMIETNO NETO	RENDIMIENTO NETO'
2006	6.045.907,86 €	172.912,96 €	172.912,96 €
2007	6.642.281,50 €	189.969,25 €	184.436,17 €
2008	7.238.655,13 €	207.025,54 €	195.141,42 €
2009	7.835.028,77 €	224.081,82 €	205.066,61 €
2010	8.431.402,40 €	241.138,11 €	214.248,09 €
		TOTAL AJUSTADO	971.805,25 €

A la vista de los cálculos realizados y utilizando una tasa de inflación del 5%, el rendimiento neto ajustado es de 971.805.25 €.

APARTADO III.- BENEFICIOS DE LA EMPRESA. ESCENARIO IV

Por último, vamos a suponer que la empresa en los próximos cinco años, obtenga el mismo margen de beneficio que ha obtenido en los últimos cuatro ejercicios, para ello, en primer lugar, vamos a ver cual ha sido el beneficio medio de los últimos cuatro ejercicios. Una vez obtenido este dato, lo aplicaremos a las ventas proyectadas en el apartado anterior, para determinar de este modo, el beneficio neto esperado para los próximos cinco años, así como su valor actualizado, llegando a un resultado neto ajustado de **462,117,18 €** para el conjunto de los cinco ejercicios económicos posteriores al cese del contrato. Todo ellos, sitúa a este escenario como el PESIMISTA. La empresa estaba realizando un esfuerzo considerable incurriendo en importantes gastos, para corregir las caídas de cartera que se producían por decisiones unilaterales de la compañía DAVOFANE S.A.

TABLA 7

	VOZ Y DATOS ALMERIA		
EJERCICIO	VENTAS	BENEFICIO	%
2001	2.733.159,46 €	55.523,17 €	2,03%
2002	4.404.061,76 €	55.944,11 €	1,27%
2003	3.909.431,35 €	49.102,29 €	1,26%
2004	4.640.447,39 €	66.123,19 €	1,42%
2005	5.596.834,82 €	46.254,95 €	0,83 %
BENEFICIO MEDIO			1,36%

TABLA 8

EJERCICIO	VENTAS	BENEFICIOS	BENEFICIOS'
2006	6.045.907,86 €	82.224,35 €	82.224,35 €
2007	6.642.281,50 €	90.335,03 €	87.703,91 €
2008	7.238.655,13 €	98.445,71 €	92.794,52 €
2009	7.835.028,77 €	106.556,39 €	97.514,19 €
2010	8.431.402,40 €	114.667,07 €	101.880,21 €
		TOTAL AJUSTADO	462.117,18 €

APARTADO IV.- RESUMEN DE ESCENARIOS Y RESULTADOS

Escenario I: Mejor de los más probables.- Estimación objetiva con los módulos de la Agencia Tributaria, suponiendo que el crecimiento mantiene la misma trayectoria que el personal en los últimos ejercicios.

Escenario II: OPTIMISTA, mejor de todos los posibles.- Estimación objetiva con los módulos de la Agencia Tributaria, suponiendo un crecimiento anual igual al del último ejercicio.

Escenario III: Peor de los más probables.- Se supone que la empresa hubiera evolucionado igual que las de su sector, con el crecimiento de facturación que tenía hasta el momento de corte y con un beneficio medio igual al del conjunto de empresas REDCOM.

Escenario IV: PESIMISTA, peor de los posibles.- Se supone que la empresa mantiene tanto el ritmo de crecimiento como el margen de benéficos que ha tenido hasta el momento.

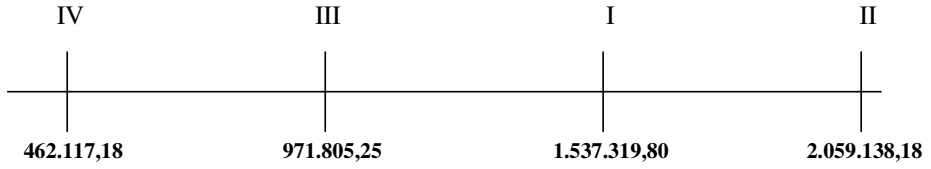


TABLA 9: RESUMEN DE RENDIMIENTOS NETOS

EJERCICIO / ESCENARIO	2006	2007	2008	2009	2010	RENDIMIENTO NETO AJUSTADO
OPTIMISTA	359.311,87 €	408.537,60 €	457.766,38 €	506.976,26 €	556.203,66 €	2.059.138,18 €
MEJOR PROBABLE	319.402,09 €	328.984,15 €	338.853,68 €	349.019,29 €	359.489,87 €	1.537.319,80 €
PEOR PROBABLE	172.912,96 €	189.969,25 €	207.025,54 €	224.081,82 €	241.138,11 €	971.805,25 €
PESIMISTA	82.224,35 €	90.335,03 €	98.445,71 €	106.556,39 €	114.667,07 €	462.117,18 €

Como puede verse, el lucro cesante o perjuicio económico actualizado a 1 de enero de 2006 conduce a cifras distintas dependiendo del escenario que se considere para el futuro de la empresa. Queremos recordar que establecer el lucro cesante siempre pasa por realizar una predicción en el futuro y esa predicción debe descansar en algunas hipótesis. Nuestra intención por lo tanto es la aplicación de una metodología estadística que contemple simultáneamente los cuatro escenarios y establezca entre ellos una ponderación.

4.- INDICACIÓN DE LOS METODOS DE VALORACION EMPLEADOS Y DE SUS LIMITACIONES.-

Por tratarse de una empresa de comercio básicamente minorista, tenemos la posibilidad de utilizar una amplia gama de métodos de valoración, ya que disponemos de cuatro estimaciones distintas y determinada en base a principios muy dispares. La consideración del futuro que hubiese tenido esta empresa, si hubiese seguido comercializando los productos que tenía hasta el momento, más aún, la cartera de clientes creada durante el ejercicio de su actividad en los últimos años, es el camino mas adecuado para establecer el valor de la empresa, en el momento en el que se rescindió el contrato por parte de la operadora (7 de diciembre de 2005), actualizando a dicha fecha los beneficios futuros estimados en los próximos cinco años .

Para actualizar los beneficios futuros, lo primero que debemos de hacer es fijar el horizonte temporal, en toda la literatura sobre valoración de empresas el horizonte temporal que mas se repite es el de los cinco años siguientes, que es lo que vamos a hacer en este informe.

Para la consideración de todos los aspectos técnicos, nos remitimos de forma general a la bibliografía que se aporta al final del trabajo.

Para establecer el valor razonable del lucro cesante, procederemos a actualizar los rendimientos netos esperados de los próximos cinco años. Por motivos técnicos, y considerando los últimos días de mes de diciembre de escasa importancia a efectos de cálculos, vamos a partir del 1 de enero de 2006 para actualizar los flujos, y desde ese momento empezarán los cinco años del horizonte temporal fijado para el presente estudio.

$$LC = \frac{Q_1}{1+K_1} + \frac{Q_2}{(1+K_1)(1+K_2)} + \frac{Q_3}{(1+K_1)(1+K_2)(1+K_3)} + \\ + \frac{Q_4}{(1+K_1)(1+K_2)(1+K_3)(1+K_4)} + \frac{Q_5}{(1+K_1)(1+K_2)(1+K_3)(1+K_4)(1+K_5)} \quad (1)$$

Siendo:

Q_1 = RENDIMIENTO NETO DEL EJERCICIO 1 (2006)
Q_2 = RENDIMIENTO NETO DEL EJERCICIO 2 (2007)
Q_3 = RENDIMIENTO NETO DEL EJERCICIO 3 (2008)
Q_4 = RENDIMIENTO NETO DEL EJERCICIO 4 (2009)
Q_5 = RENDIMIENTO NETO DEL EJERCICIO 5 (2010)

K_1 = TIPO DE ACTUALIZACION O DESCUENTO EJERCICIO 1 = 5%
K_2 = TIPO DE ACTUALIZACION O DESCUENTO EJERCICIO 2 = 5%
K_3 = TIPO DE ACTUALIZACION O DESCUENTO EJERCICIO 3 = 5%
K_4 = TIPO DE ACTUALIZACION O DESCUENTO EJERCICIO 4 = 5%
K_5 = TIPO DE ACTUALIZACION O DESCUENTO EJERCICIO 5 = 5%

Consideramos que el tipo de interés a lo largo del periodo puede subir, por lo que, aunque ahora esté más bajo, consideramos prudente manejar este tipo de actualización para todo el periodo.

Respecto a los Rendimientos netos futuros, trataremos de obtener su valor esperado y su varianza en base a los datos de la tabla 9, para ello hemos de proceder a una modelización estadística que nos permita pasar de un ambiente de incertidumbre a un ambiente de riesgo, determinando el valor esperado así como su varianza de los rendimientos netos para cada periodo futuro. Esto nos permitirá, posteriormente, usar la Ley de los Grandes Números y el Teorema Central del Límite, y en consecuencia, establecer un intervalo en el cual, con una probabilidad de un 95%, este contenido el lucro cesante o rendimiento actualizado dejado de percibir.

Los datos que utilizaremos para estimar el valor esperado y la varianza, serán los descritos en la tabla 9.

Los rendimientos netos futuros, estamos obligados a tratarlos en un ambiente de incertidumbre ya que es imposible conocer ahora lo que hubiese pasado si la empresa hubiese seguido funcionando, por lo tanto, podemos precisar, siguiendo el trabajo: "Valuations of the firm: Efects of uncertainty in a market contex" A.A. Robichek y S.C. Myers, Weston y Woods .Edditorial Ariel Barcelona 1974:

a) Lo importante en estos casos, es conocer tanto la forma como la dispersión en la distribución de probabilidades subyacentes.
b) A mayor dispersión de la distribución de probabilidades, mayor será el intervalo de determinación de daño causado en este caso.

Estamos interesados en determinar un intervalo donde se encuentre el valor del lucro cesante actualizado, para ello, vamos a elegir como distribuciones subyacentes de los rendimientos netos anuales, cuatro modelos estadísticos: la distribución Uniforme, la Beta, la Trapezoidal y la Triangular, que son las habituales (Vease SUAREZ SUAREZ A. (1980) "DECISIONES

OPTIMAS DE INVERSION Y FINANCIACION EN LA EMPRESA" ED. PIRAMIDE). De este modo podremos determinar completamente estas distribuciones, utilizando los valores habituales: pesimista, optimista y más probable, o los extremos del intervalo modal en otros casos.

Si aceptamos como hipótesis de trabajo que la distribución subyacente al comportamiento de los rendimientos netos futuros es una de las distribuciones mencionadas en este estudio, de acuerdo con la literatura especializada, debemos establecer para cada período el rendimiento neto máximo (b) y el rendimiento neto mínimo (a), en términos propios el rendimiento neto optimista y pesimista, y en algunos de los modelos, serán necesarios hasta cuatro rendimientos netos por ejercicio (a, b, m_1 y m_2), es decir:

Q_a^i = Rendimiento neto pesimista del año i = a_i

$Q_{m_1}^i$ = Rendimiento neto mínimo entre los más probables del año i = m_1^i

Q_m^i = Rendimiento neto más probables del año i = $m_i = \dfrac{m_1^i + m_2^i}{2}$

$Q_{m_2}^i$ = Rendimiento neto máximo entre los más probables del año i = m_2^i

Q_b^i = Rendimiento neto optimista del año i = b_i

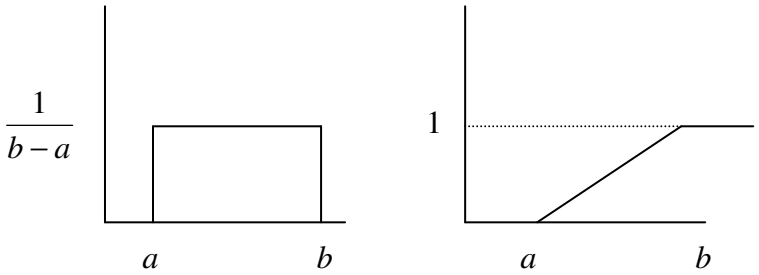
Dependiendo de la distribución subyacente que queramos utilizar, será necesario un determinado número de observaciones para cada ejercicio, con un mínimo de dos (distribución Uniforme) y un máximo de cuatro (distribución Trapezoidal).

4.1 Distribuciones Estadísticas Subyacentes: Utilizaremos como posibles distribuciones estadísticas subyacentes las siguientes:

4.1.1 Distribución Uniforme:

La distribución uniforme tiene las siguientes funciones de densidad y distribución y supone que todos los valores entre el mínimo y el máximo tienen la misma probabilidad.

En cuanto a sus principales características son:

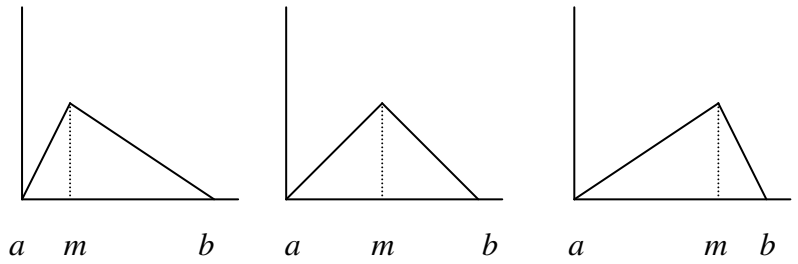


Función de densidad y función de distribución del modelo probabilístico rectangular

Función generatriz de momentos	Esperanza matemática	Varianza	Coeficiente de asimetría de Fisher	Coeficiente de curtosis de Fisher
$G(t) = \dfrac{e^{bt} - e^{at}}{t(b-a)}$	$E(x) = \dfrac{a+b}{2}$	$Var(x) = \dfrac{(b-a)^2}{12}$	$g_1 = 0$	$g_2 = -\dfrac{6}{5}$

4.1.2 Distribución Triangular:

En esta distribución, se considera que *m* es el flujo mas probable y que la densidad se distribuye entorno de *m*. Su función de densidad y distribución tienen la siguiente gráfica y sus principales características son las siguientes:



Distintas representaciones de la función de densidad de la distribución triangular

Función generatriz de momentos	$G(t) = 2\dfrac{(b-m)e^{at} + (m-a)e^{bt} - (b-a)e^{mt}}{t^2(b-a)(m-a)(b-m)}$
Esperanza matemática	$E(X) = \dfrac{a+b+m}{3}$
Varianza	$Var(x) = \dfrac{(b-m)^2 + (m-a)^2 + (b-m)(m-a)}{18}$
Coeficiente de asimetría de Fisher	$g_1 = \dfrac{\sqrt{2}(a+b-2m)(b+m-2a)(2b-m-a)}{5\left((b-a)^2 - (m-a)(b-m)\right)^{3/2}}$
Coeficiente de curtosis de Fisher	$g_2 = -\dfrac{3}{5}$

4.1.3. ***Distribución Beta:***

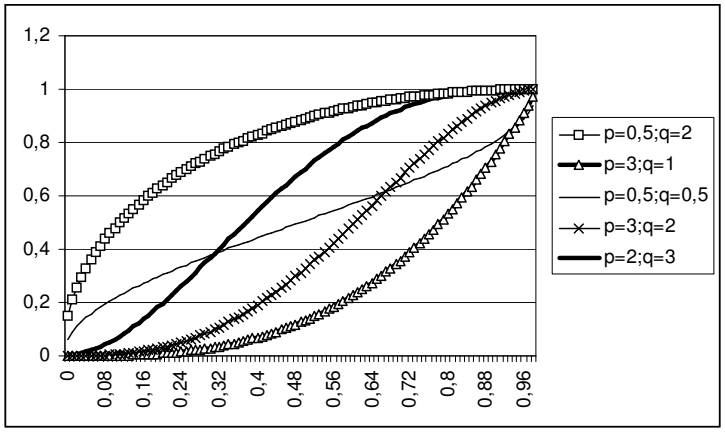
La función de densidad puede tener forma de campana, de "U" o de "J", m se considera el flujo más probable de cada periodo, las funciones tienen la siguiente forma y sus características estocásticas son:

Esperanza matemática	$E(t) = \dfrac{p}{p+q}$
Varianza	$Var(t) = \dfrac{pq}{(p+q+1)(p+q)^2}$
Coeficiente de asimetría de Fisher	$g_1 = \dfrac{2(q-p)\sqrt{p+q+1}}{\sqrt{pq}(p+q+2)}$
Coeficiente de curtosis de Fisher	$g_2 = 6\dfrac{p(p+1)(p-2q) + q(q+1)(q-2p)}{pq(p+q+2)(p+q+3)}$

Y las expresiones de la media y la varianza quedan de la siguiente forma:

$$E(x) = \frac{a + 4m + b}{6}$$

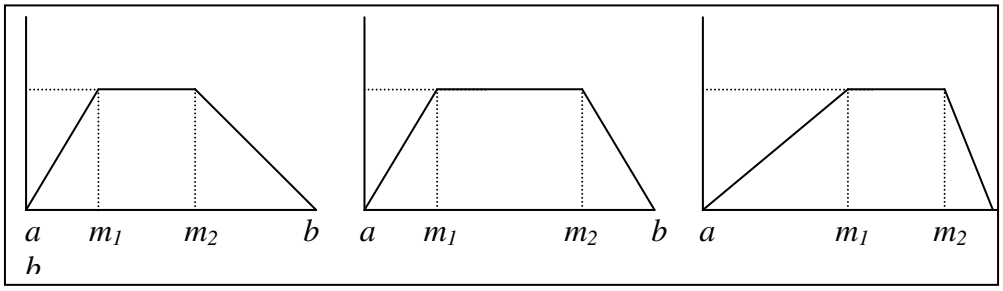
$$Var(x) = \frac{(b-a)^2}{36}$$



Diferentes formas de la función de densidad de la distribución beta

4.1.4 *Distribución Trapezoidal:*

Esta distribución, que pretende una mayor aproximación y utiliza mejor la información, ya que en lugar de un valor medio, utiliza un intervalo modal m_1, m_2, salen los flujos de caja más probables, son la mitad entre m_1 y m_2. Su figura es la siguiente:



Y sus principales características, las siguientes:

Distintas representaciones de la función de densidad de la distribución trapezoidal

Función generatriz de momentos	$G(x) = \frac{2\left[(m_1 - a)\left(e^{bt} - e^{m_2 t}\right) + (b - m_2)(e^{at} - e^{m_1 t}\right]}{(m_1 - a)(b - m_2)(b - a + m_2 - m_1)t^2}$
Esperanza matemática	$E(x) = \frac{1}{3}\left[b + m_1 + m_2 + a - \frac{bm_2 - am_1}{b - a + m_2 - m_1}\right]$
Varianza	$Var(x) = \frac{1}{18}\left[(b - m_1)^2 + (m_2 - a)^2 + (b - m_2)(m_1 - a) - \frac{2(b-a)(m_2 - m_1)(m_2 - a)(b - m_1)}{(b - a + m_2 - m_1)^2}\right]$

4.2.- Valoración Con Las Distribuciones Subyacentes.

Utilizando la información del apartado tres, en concreto, la tabla 9, y haciendo uso de las distribuciones descritas en los puntos anteriores, podemos obtener los valores esperados, $E[Q_i]$, y $Var[Q_i]$, para cada año y distribución empleada, cuyos resultados se muestran en la siguiente tabla:

TABLA 10

EJERCICIO / DISTRIB.	$E[Q_1]$	$E[Q_2]$	$E[Q_3]$	$E[Q_4]$	$E[Q_5]$
UNIFORME	220.768,11 €	249.436,32 €	278.106,05 €	306.766,33 €	335.435,37 €
TRIANGULAR	229.231,25 €	252.783,11 €	276.383,90 €	300.027,73 €	323.728,24 €
BETA	237.694,39 €	256.129,91 €	274.661,75 €	293.289,14 €	312.021,11 €
TRAPEZOIDAL	311.283,57 €	339.275,16 €	367.363,58 €	395.544,40 €	423.832,72 €

Y también podemos tabular las varianzas:

TABLA 11

EJERCICIO / DISTRIB.	$Var[Q_1]$	$Var[Q_2]$	$Var[Q_3]$	$Var[Q_4]$	$Var[Q_5]$
UNIFORME	6398124478	8437739630	10759278657	13361339358	16246213359
TRIANGULAR	3326290871	4415332694	5660330339	7060487866	8617105141
BETA	2132708159	2812579877	3586426219	4453779786	5415404453
TRAPEZOIDAL	4145206274	5032054794	6105815338	7361995131	8798113355

4.3 *Determinación Del Intervalo De Confianza:*

En definitiva, una vez que hemos obtenido los valores esperados y las varianzas correspondientes a cada uno de los cinco ejercicios objeto de estudio, podremos determinar cual es el valor esperado del lucro cesante y su varianza, utilizando las siguientes expresiones:

Lucro cesante:

$$E[LC]= \frac{E[Q_1]}{1+K_1}+\frac{E[Q_2]}{(1+K_1)(1+K_2)}+\frac{E[Q_3]}{(1+K_1)(1+K_2)(1+K_3)}+ \\ +\frac{E[Q_4]}{(1+K_1)(1+K_2)(1+K_3)(1+K_4)}+\frac{E[Q_5]}{(1+K_1)(1+K_2)(1+K_3)(1+K_4)(1+K_5)} \quad (2)$$

Varianza:

$$Vr[LC]= \frac{Vr[Q_1]}{1+K_1}+\frac{Vr[Q_2]}{(1+K_1)(1+K_2)}+\frac{Vr[Q_3]}{(1+K_1)(1+K_2)(1+K_3)}+ \\ +\frac{Vr[Q_4]}{(1+K_1)(1+K_2)(1+K_3)(1+K_4)}+\frac{Vr[Q_5]}{(1+K_1)(1+K_2)(1+K_3)(1+K_4)(1+K_5)} \quad (3)$$

Si tenemos en cuenta la expresión (1) y pasamos a una situación de riesgo, el lucro cesante, tendrá como valor esperado la expresión (2) y como varianza la expresión (3). Por otra parte,

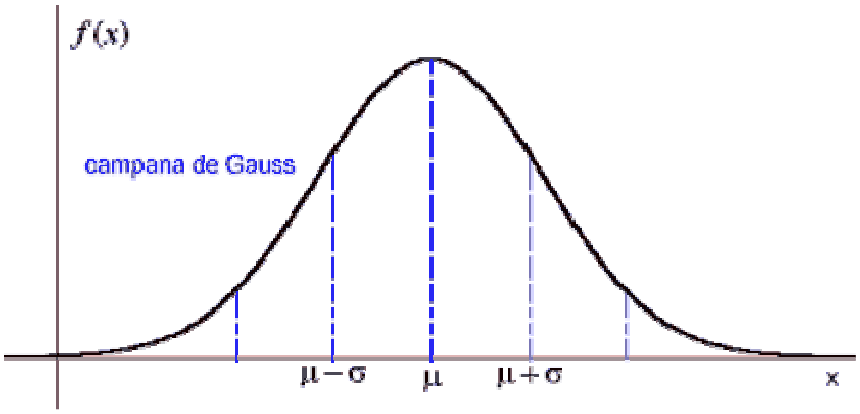
podemos considerar que el lucro cesante, sea una variable aleatoria normal, por ser la suma de variables aleatorias y por tanto:

$$LC \to N\left(E(LC), \sqrt{Var(LC)}\right) \tag{4}$$

Esto nos permite estimar un intervalo de confianza en cada uno de los casos analizados.

a) D.S. Uniforme
b) D.S. Triangular
c) D.S. Beta
d) D.S. Trapezoidal

Si utilizamos la $N\left(E(LC), \sqrt{Var(LC)}\right)$ y un intervalo de confianza del 95 %, los intervalos de confianza son:



Intervalo de confianza al 95%: $E(L) \pm \sqrt{Var(LC)}$

4.4. Limitaciones del Método:

El método empleado, se apoya en el Teorema Central del Límite del cálculo de probabilidades. Este Teorema dice que la suma de variables aleatorias, distribuyéndose todas ellas de una misma manera independientemente las unas de las otras converge, cuando n tiende a infinito, a una variable aleatoria que supone una distribución normal y que tiene por media (valor esperado) y varianza, la suma de las medias y las varianzas de n variables aleatorias respectivamente. Esta condición exigiría en nuestro caso que el número de flujos de caja y por tanto la vida de la empresa fuese muy grande. Desde un punto de vista práctico puede admitirse que el número de valores actuales (es decir de ejercicios futuros) sea 10, para que la suma converja con cierta función a una ley normal, pero indudablemente predecir los flujos de caja de una empresa a 10 años tiene una consistencia objetiva muy débil.

4.5. La Desigualdad de Tchebichev.

Para resolver el problema que resulta del punto anterior, consistente en la hipótesis de normalidad de las distribuciones, podemos bajar el nivel de confianza hasta el 50 % y aplicar la desigualdad de Tchebichev, de lo que resultan nuevos intervalos de confianza que con un 50 % de probabilidad, contendrán el lucro cesante de la empresa.

5. CONCLUSIONES.

A la visa de todos los datos contenidos en este informe y la cantidad de información manejada con objeto de llegar la una situación que nos permita pronunciarnos acerca de un

valor y un intervalo de confianza que contenga el lucro cesante de la empresa objeto de estudio, podemos establecer distintos niveles de seguridad, cada uno de los cuales nos arrojará un valor distinto.

5.1. Media de los rendimientos netos actualizados.

Una primera medida del lucro cesante, puede ser la media de los rendimientos netos obtenidos en cada ejercicio, de lo que resulta lo siguiente:

ESCENARIO	RENDIMIENTO NETO ACTUALIZADO
OPTIMISTA	2.059.138,18 €
MEJOR PROBABLE	1.537.319,80 €
PEOR PROBABLE	971.805,25 €
PESIMISTA	462.117,18 €

Para el intervalo de confianza, escogemos como extremos los valores optimista y pesimista, por tanto, diremos que el valor está contenido en el intervalo [462.117,18 , 2.059.138,18].

En cuanto al valor puntual, la media de los escenarios, que es: 1.260.627,68 €.

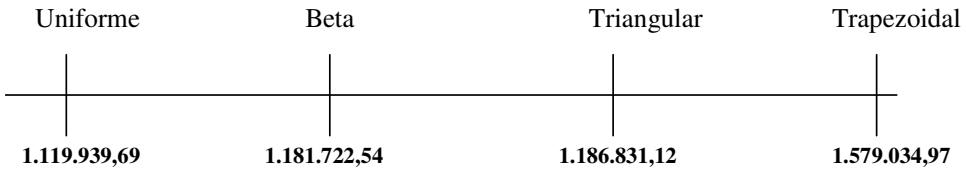
5.2. Paso de la incertidumbre al riesgo.

Entendemos que podemos obtener una mejor información, si el lugar de trabajar en un ambiente de incertidumbre, lo hacemos en ambiente de riesgo, lo que conseguimos aplicando las distribuciones subyacentes descritas en el punto 4.1 de este informe. Insistimos en que estos resultados implican un supuesto de normalidad en las distribuciones de a suma de variables aleatorias.

Aplicando las distintas distribuciones, obtenemos que los intervalos de confianza en los que, con un 95% de probabilidad se va a encontrar el valor del lucro cesante de la empresa, son los siguientes:

BETA CLÁSICA: [955.993,21 , 1.407.451,87]
TRIANGULAR: [903.190,30 , 1.470.471,94]
UNIFORME: [800.965,02, 1.582.914,36]
TRAPEZOIDAL: [1.282.821,13, 1.875.248,81]

Y los valores esperados:



De entre los resultados anteriores, nos inclinamos por el intervalo de confianza ofrecido por la distribución Trapezoidal por ser la que más datos utiliza.

5.3. Desigualdad de Tchebichev.

La desigualdad de Tchebichev, adopta la siguiente expresión:

$$P\left(\left(x - E\left(x\right)\right) \leq a\right) = 1 - \frac{E\left(x - E\left(x\right)\right)^2}{a^2}$$

Como el método anterior tiene una debilidad, podemos rebajar la confianza en el intervalo por lo que no necesitamos suponer la normalidad de las distribuciones, tal y como vimos en el punto 4.5. Con este último método, los intervalos de confianza pasan a ser los siguientes:

BETA CLÁSICA: [793.096,43 , 1.118.890,00]
TRIANGULAR: [698.501,89 , 1.107.878,70]
UNIFORME: [518.819,52 , 1.083.110,53]
TRAPEZOIDAL: [1.069.059,45 , 1.496.582,81]



La tendencia a la baja, se debe a la propia pérdida de confianza, de nuevo la distribución trapezoidal es la recomendada.

6. VALORACIÓN A TRAVÉS DE LA CARTERA DE CLIENTES.

La empresa Voz y Datos Almería, S.L., dedicada como ha sido comentado en distintas ocasiones en este informe, a la comercialización de productos de la compañía de telefonía móvil VODAFONE ESPAÑA, tenía constituida una cartera de clientes de los que recibía unos ingresos, que describimos a continuación:

- Paquete de Bienvenida
- Comisión por alta
- **CARTERA**
- Permanencia / Valor
- Ajuste por incidencias
- Apoyo de terminal
- Penalización por fraude
- Comisión prepago
- Cambio de puntos
- Otros servicios

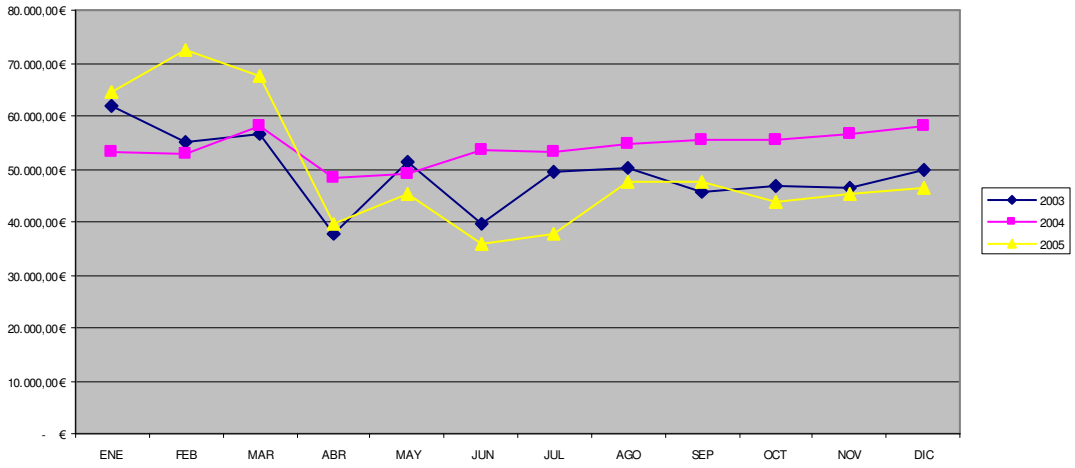
Los ingresos por cartera se corresponden únicamente al consumo de clientes captados por la empresa Voz y Datos Almería, S.L., con motivo del corte de los servicios, se han dejado de producir los citados ingresos. En primer lugar veremos cual ha sido la evolución de los ingresos por cartera en los últimos años, haciendo además un estudio de la relación existente entre la cartera de los meses de cada periodo, mediante la utilización del estadístico r de Sperman. En segundo lugar, analizaremos cual ha sido la relación existente entre la evolución de los gastos de personal y los ingresos por cartera, para comprobar si el esfuerzo en los gastos de personal se corresponde con los ingresos por cartera. En el caso de que exista una gran correlación en los ingresos por cartera de los últimos años, deberemos tomar un promedio de ésta, en otro caso, deberemos utilizar sólo la cartera del ejercicio 2005.

6.1. Estudio de la evolución de la cartera.

	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1997	2,61 €	812,88 €	1.178,61 €	1.219,84 €	1.149,24 €	1.393,77 €	1.719,18 €	1.659,69 €	2.405,18 €	3.150,67 €	3.165,90 €	4.441,63 €
1998	5.737,01 €	6.074,93 €	7.225,69 €	7.791,75 €	10.276,46 €	10.708,84 €	11.513,77 €	12.474,02 €	12.619,69 €	12.390,93 €	14.681,34 €	17.090,81 €
1999	20.010,37 €	20.010,37 €	20.468,65 €	20.850,76 €	22.494,57 €	24.305,19 €	27.841,16 €	25.628,45 €	24.686,91 €	25.834,46 €	26.066,06 €	26.364,29 €
2000	27.886,07 €	27.046,43 €	27.152,05 €	25.046,60 €	26.763,21 €	27.965,69 €	28.419,22 €	28.781,80 €	29.833,62 €	29.750,30 €	29.069,50 €	32.767,26 €
2001	33.597,30 €	33.823,92 €	33.319,89 €	34.648,92 €	36.771,19 €	35.456,23 €	36.680,49 €	41.130,82 €	42.970,29 €	43.683,42 €	68.709,21 €	33.092,38 €
2002	36.419,22 €	43.886,98 €	50.696,76 €	44.350,32 €	47.662,14 €	53.676,93 €	51.132,71 €	51.962,86 €	50.259,19 €	51.254,94 €	52.625,93 €	53.501,32 €
2003	61.888,39 €	54.997,48 €	56.450,96 €	37.676,94 €	51.480,20 €	39.491,05 €	49.483,49 €	50.259,99 €	45.765,22 €	46.946,69 €	46.306,60 €	49.670,87 €
2004	53.087,57 €	52.753,91 €	58.211,46 €	48.195,38 €	49.209,65 €	53.453,90 €	53.275,60 €	54.900,55 €	55.360,49 €	55.483,21 €	56.533,21 €	58.241,69 €
2005	64.365,50 €	72.549,88 €	67.557,96 €	39.770,44 €	45.270,09 €	35.884,79 €	37.863,34 €	47.545,82 €	47.615,15 €	43.792,79 €	45.438,39 €	46.548,29 €

En la tabla anterior, tenemos los ingresos por cartera expresados en euros mes a mes desde enero de 1997, cuando se inició la actividad, hasta diciembre de 2005, mes en el que se cortó el servicio. Una visión gráfica de la misma:

INGRESOS POR CARTERA PERIODO 2003 – 2005.



Podemos ver que la evolución de la cartera no es muy homogénea, esto quizá pueda ser debido a anulaciones de cartera por parte de la compañía, ya que existen quejas de la empresa **PERTURBACIONES INESPERADAS SL**, referente a esta situación.

Queremos saber si para calcular el lucro cesante, hemos de tener en consideración la cartera del último año, o una media de los tres últimos. Para esto, vamos a utilizar el estadístico r de Sperman, siguiendo los trabajos realizados por Grossman, Mingot, Granados y asociados.

El resultado es el siguiente:

AÑOS	r_{SP}
97/98	0,92
98/99	0,87
99/00	0,68
00/01	0,21
01/02	0,34
02/03	-0,49
03/04	0,26
04/05	0,22

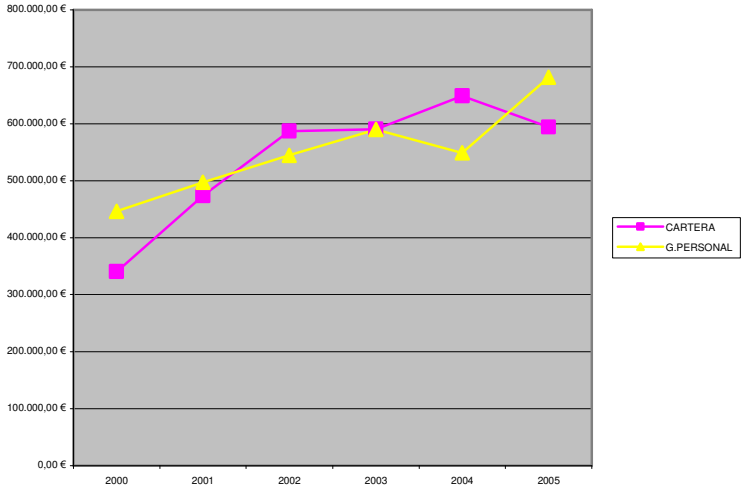
Hemos de comentar que el estadístico r de Sperman, puede tomar valores entre -1 y 1, el extremo inferior del intervalo indica que la relación existente es negativa, el cero la ausencia total de relación, y el superior una relación positiva absoluta, extremos que no se dan en la práctica, si bien, cuando el estadístico se acerca a 1, podemos inferir la que relación es muy

estrecha. En nuestro caso, vemos que la relación es muy grande inicialmente, pero que se va perdiendo con el paso de los años.

A continuación, vamos a ver la relación existente entre el gasto de personal y los ingresos por cartera.

AÑO	CARTERA	G.PERSONAL
2000	340.481,76 €	446.436,37 €
2001	473.884,08 €	497.076,09 €
2002	587.429,30 €	544.547,21 €
2003	590.417,88 €	589.746,32 €
2004	648.706,62 €	548.986,01 €
2005	594.202,44 €	681.480,86 €

Cuya gráfica es:



Como podemos ver, el esfuerzo en los gastos de personal, no se corresponde en los últimos años con la evolución de la cartera, como venía pasando en los primeros años. Es posible que se estén produciendo caídas de cartera por anulaciones de la compañía Vodafone, y de esto existe constancia en las reiteradas quejas que la empresa venía manifestando.

En nuestra opinión, deberíamos de tomar un promedio de las carteras de los últimos tres ejercicios para realizar la valoración, ya que aunque la valoración entre ellas no es muy alta, entendemos que esto se debe a actuaciones irregulares de la empresa Vodafone, anulando carteras que habían sido generadas por la empresa VOZ Y DATOS, S.L.

	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC	TOTAL
2003	61.888,39 €	54.997,48 €	56.450,96 €	37.676,94 €	51.480,20 €	39.491,05 €	49.483,49 €	50.259,99 €	45.765,22 €	46.946,69 €	46.306,60 €	49.670,87 €	**590.417,88 €**
2004	53.087,57 €	52.753,91 €	58.211,46 €	48.195,38 €	49.209,65 €	53.453,90 €	53.275,60 €	54.900,55 €	55.360,49 €	55.483,21 €	56.533,21 €	58.241,69 €	**648.706,62 €**
2005	64.365,50 €	72.549,88 €	67.557,96 €	39.770,44 €	45.270,09 €	35.884,79 €	37.863,34 €	47.545,82 €	47.615,15 €	43.792,79 €	45.438,39 €	46.548,29 €	**594.202,44 €**
PROMEDIO	**59.780,49 €**	**60.100,42 €**	**60.740,13 €**	**41.880,92 €**	**48.653,31 €**	**42.943,25 €**	**46.874,14 €**	**50.902,12 €**	**49.580,29 €**	**48.740,90 €**	**49.426,07 €**	**51.486,95 €**	**611.108,98 €**

6.2. Duración de la cartera.

A la hora de valorar la cartera de una empresa de telefonía, la mayor dificultad que podemos encontrar, puesto que conocemos la evolución de la cartera en los últimos años, es la duración que hubieran tenido estos ingresos y más aún, cual hubiera sido el gasto realizado en consumo telefónico por los clientes, ya que gran parte de los clientes no continúan en la compañía una vez que el distribuidor cambia, ya que la confianza que les ofrece el distribuidor con quien comercializan sus productos no puede ser transferida de una empresa a otra.

En la nota mensual de la Comisión del Mercado de las Telecomunicaciones, aparece un estudio en el que se pone de manifiesto el número de portabilidades que se han producido en los últimos años, sabiendo además, por un estudio del Instituto nacional de estadística cual es el

número de líneas móviles que hay en el mercado español, podemos conocer cual es el índice de bajas de una compañía, ya que desde la aparición de la legislación que permite "portar" el numero de la línea, pocos son los usuarios que cambian de número y, por tanto, las bajas de una compañía vienen producidas por las portabilidades a otras empresas.

De los estudios citados en el párrafo anterior y que se adjuntan al final de este informe, extraemos la siguiente información:

Número de líneas móviles de Vodafone 1T 2006: 13,5 millones
Número medio de líneas periodo 2000 – 2006: 13 millones.
Número de líneas perdidas en el periodo 2000-2006: 2.637.051
Número de pérdidas por año: 20,18 %
% de pérdida anual: 2,88 %

De los datos anteriores y, suponiendo un número de bajas igual al número de portabilidades, lo que supondría una pérdida total del líneas de un 5,76 %.

Por otro lado, según información de www.elpais, el cambio de compañía está relacionado con el cambio de móvil y la portabilidad y, por término medio, los usuarios suelen cambiar de móvil entre 15 y 18 meses.

Si tenemos en consideración los razonamientos anteriores, nos parece, en nuestra opinión, valorar el daño causado, multiplicando el promedio de la cartera anual, por 2,5, lo que supone una vida de la cartera superior a los 18 meses, pero muy inferior a la vida media de las carteras de Vodafone que estarían entorno a los 8-9 años y porque viene siendo lo habitual.

Por lo tanto, nuestra valoración sería la siguiente: cartera media de los tres últimos años = 611.108,98 x 2,5 = 1.527.772,45 €

7. VALORACIÓN FINAL.

En primer lugar, si hiciéramos uso solamente de los escenarios planteados en la primera parte del trabajo y realizáramos un promedio de los mismos, el perjuicio económico se podría valorar en 1.260.627,68 €, oscilando en un intervalo muy amplio que iría desde los 462.117,18 € a los 2.059.138,18 €.

La consideración conjunta de todos los escenarios mediante el tratamiento estadístico desarrollado en la primera parte del trabajo, nos conduciría a aceptar como más razonable, el supuesto de distribuciones trapezoidales subyacentes, lo que nos daría un intervalo entre 1.282.821,13 € y 1.875.248,81 €, con un valor esperado de 1.579.034,97 €, entendemos que esta distribución es la más acertada, porque es la única que utiliza conjuntamente todos los datos disponibles. El análisis de la cartera, nos conduce a un resultado puntual de 1.527.772,45 €, utilizando las técnicas habituales.

En base a todo lo anterior, proponemos, como valoración del perjuicio económico causado a la empresa Voz y Datos Almería, S.L., la cantidad de 1.527.772,45 €, siendo la cantidad mínima a considerar la de 1.282.821,13 €.

BIBLIOGRAFÍA

ANTÓN J.J. YAO D.A. (2006) :Finding "lost" profit : An equilibrium Análysis of Patent Infrigement Damages :" *The journal of Laws .Economics & Organization* " .Vol 23 nº 1

BOWLES T.J. LEWIS W.C. (1999): Unsettled Issues in Measuring Lost Profit. *Journal of Legal Economic*. p 19

BOTELHO ANABELA (2001). *Strategic behaviour at trial .the production , reporting, and evaluation of complex evidenc*e. Working Papers Nº 14

Garnica Martín J.F- La prueba del lucro cesante. *Revista de responsabilidad civil y seguro* (2005) pp 45-64

GAUGHAN, P. y THORNTON, R. J., (1994) : "*Litigations In Economics*- Editors Tornton R.J., Aronson J.R.- JAI PRESS INC.

HERRERÍAS, R.(ed.). 2001. "*Programación, selección y control de proyectos en ambiente de incertidumbre*".Editorial Monográfica. Granada.

LORING, J. 1997. "*La Gestión Financiera*". Ediciones Deusto. Córdoba.

TROUT ROBERT (2000) : Valuing Busines Goodwill in eminent Domain Cases". *The Appraisal Journal* Apr 2000 68-2 p-171

STEWART MARION B. (2005). Calculating Economics Damages in Intellectual Property Disputes .The Role o Market Definition. *Computer and Internet Lawyer* Aug 2005 .pp22 8

SÚAREZ, A.S. 1980. "Decisiones óptimas de inversión y financiación en la empresa". Pirámide. Madrid.

WILLIAM A. SUPR. Y CHARLES P. BONINI. 1978. "Toma de decisiones en Administración mediante métodos estadísticos" Editorial Limusa. Mexico.

DETERMINING THE NUMBER OF MIXTURE COMPONENTS USING AN EXPERIMENTAL DESIGN

ANA OLIVEITA-BROCHADO
E-mail: abrochado@concorrencia.pt
Autoridade da Concorrência
F. VITORINO MARTINS
E-mail: vmartins@concorrencia.pt
Faculdade de Economia, Universidade do Porto

Abstract

The aim of this work is to determine how well criteria designed to help the selection of the adequate number of mixture components perform in mixture regressions of normal data. We address this research question based on results of an extensive experimental design. The simulation experiment compares several criteria (26), including information criteria and classification-based criteria. In this full factorial design we manipulate 9 factors and 22 levels, namely: true number of segments (2 or 3), mean separation between segments (low, medium or high), number of consumers (100 or 300), number of observations per consumer (5 or 10), number of predictors (2, 6 or 10), measurement level of predictors (binary, metric or mixed), error variance (20% or 60%), minimum segment size (5-10%, 10-20% or 20-30%) and error distribution (normal versus uniform). The performance of the segment retention criteria is evaluated by their success rates; we also investigate the influence of experimental factors and their levels on success rates. The best results were obtained for the criteria AIC_3, AIC_4, HQ, ICLBIC and ICOMPLBIC. BIC and CAIC also perform well with large samples and a large number of market segments.

Key-Worlds: Market segmentation, information criteria, classification criteria, experimental design, simulation.
Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUCTION

Models for mixture regressions of normal data (DeSarbo & Cron, 1988; Wedel & DeSarbo, 1995) are currently a very popular statistical-model base to clustering. They have been of considerable interest in recent years, leading to a vast number of both methodological developments and applications in different areas. However, despite the widespread use of these models, the decision of how many mixture components are required to adequately represent the data is, according to many authors, an important problem, but without a satisfactory statistical solution. The aim of this work is to determine how well criteria designed to help the selection of the adequate number of mixture components perform in mixture regressions of normal data. As the true number of mixture components in real-world data is unknown, the evaluation of the effectiveness of segment retention criteria is usually accomplished through an experimental design. Few studies have been published focusing on the segment retention problem in mixture regression models of normal data. The first work by Hawkins *et al.* (2001), examines the performance of 12 base criteria, namely AIC (Akaike, 1973), AIC_3 (Bozdogan, 1994), MDL (Rissanen, 1987), ICOMP (Bozdogan, 1994), CL, NEC (Celeux & Soromenho, 1996), PC (Bezdek, 1981), AWE (Banfield & Raftery, 1993), MIR (Windham & Cutler, 1992), ALL, ANC, WID (Cutler & Windham, 1994) by varying the number of mixture components, the degree of separation between components and the mixing proportions. The authors conclude that PC was the least successful criterion and report good results for MDL and AWE. The study

by Andrews and Currim (2003a) compares de performance of AIC (Akaike, 1993), AIC3 (Bozdogan, 1994), BIC (Schwartz, 1978), CAIC (Bozdogan, 1987), ICOMP (Bozdogan, 1994), NEC (Celeux & Soromenho, 1996) and the validation sample log likelihood (Andrews & Currim, 2003a) manipulating eight data characteristics, namely: true number of segments, mean separation between segment coefficients, number of individuals, number of predictors, error variance, minimum segment size and measurement level of predictors. The authors find that AIC_3 is the best criterion to use with mixture regression models of normal data. Sarstedt (2008) compared the performance of AIC (Akaike, 1993), AIC3 (Bozdogan, 1994), CAIC and BIC by varying the number of mixture components and segment size. The author shows that BIC and CAIC indicate the appropriate solution more frequently.

This work extends previous studies in two ways. Fist, we intend to include an experimental condition not explored yet in previous studies, namely the distributional misspecification. According to Andrews e Currim (2003a: 242), "*for analysts, another potentially important factor that was not studied directly in these simulations was distributional misspecification, that is, when the distribution of the data does not match that of the model*". Furthermore, as a large number of criteria were not considered before, we aim at comparing the performance of 26 criteria.

The plan of this work is as follows: we start describing the criteria and our experimental design used to generate the simulated data, then discuss the findings of the study and finish with a conclusion.

2 METHODOLOGY

2.1 The criteria

We intend to compare the effectiveness of 11 information criteria and 15 classification-based criteria, identified subsequently, through a simulation. As the likelihood increases with the addition of a component to a mixture model, some heuristics, called Information Criteria, attempt to balance the increase in fit obtained against the larger number of parameters estimated to models with more clusters. Information Criteria are a general family (Brochado and Martins, 2005), including criteria that are estimates of (relative) Kullback-Leibler distance – Akaike Information Criteria, AIC (Akaike, 1973), Modified AIC, AIC_3 (Bozdogan, 1994), AIC_4 (Bozdogan, 1994), Small Sample AIC, AIC_c (Hurvich & Tsai, 1995), approaches who derived within a Bayesian framework for model selection – Bayesian Information Criteria and those named consistent criteria – Consistent AIC, CAIC (Bozdogan, 1987), CAIC with Fisher Information, CAICF (Bozdogan, 1987), Information Complexity Criterion, ICOMP (Bozdogan, 1994), Hannan-Quinn Criteria, HQ (Hannan & Quinn, 1979) and Minimum Description Length, MDL (Rissanen, 1987), MDL_2 and MDL_3 (Liang *et al.*, 1992). Although Information Criteria account for over-parameterization, as a large number of clusters is derived, is also important to ensure that the segments are sufficiently separated to the selected solution. To access the ability of a mixture model, providing well-separated clusters, an entropy statistic can be used to evaluate the degree of separation in the estimated posterior probabilities. This approach yields the Classification Criteria. Some measures in this general group are derived in the context of mixture models, as Entropy Measure, E_s (DeSarbo *et al.*, 1992), Logarithm of the Partition Probability, LP (Biernacki, 1997), Entropy, E (Biernacki, 1987), Normalized Entropy Criterion, NEC (Celeux and Soromenho, 1996), Classification Criterion, C (Biernacki and Govaert, 1996), Classification Likelihood Criterion, CLC (Biernacki and Govaert, 1996), Approximate Weight of Evidence, AWE (Banfield and Raftery, 1993), Integrated Completed Likelihood with BIC Approximation, ICLBIC (Biernacki and Celeux, 1998), ICL with BIC Approximation, ICOMPLBIC (Dias, 2004) - while others are "imported" from the fuzzy literature, as Partition Coefficient, PC (Bezdek, 1981), Partition Entropy, PE (Bezdek, 1981), Normalized Partition Entropy, NPE (Bezdek, 1981), Nonfuzzy Index, NFI (Roubens, 1978), Minimum Hard Tendency, Min_{ht} (Rivera *et al.*, 1990), Mean Hard tendency, $Mean_{ht}$ (Rivera *et al.*, 1990). These criteria are named respectively as probabilistic indices and fuzzy indices. The reader is referred

to the author's previous work (Brochado and Martins, 2005) for detailed discussions of the theoretical underpinnings of these criteria and references.

2.2 The data

In this experimental design we manipulate 9 factors and 22 levels, namely: (1) true number of mixture components (2 or 3); (2) mean separation between segments (low, medium or high), (3) number of consumers (100 or 300), (4) number of observations per consumer (5 or 10), (5) number of predictors (2, 6 or 10), (6) measurement level of predictors (binary, metric or mixed), (7) error variance (20% or 60%), (8) minimum segment size (5-10%, 10-20% or 20-30%) and (9) error distribution (normal versus uniform). As the design is factorial, with tree replications (datasets) per cell, a total of $7776=2^5 3^5$ is generated.

The variance of the error term was obtained through the approach proposed by Vriens *et al.* (1996). Vector of parameters for coefficients were generated based on Andrews and Currim (2003) proposal.

3 SUMMARY OF RESULTS

We evaluate the performance of segment retention criteria by their hit rate, or percentage of datasets in which the criteria identify the correct number of segments; we also considered the over fitting rate and the under fitting rate; given two criteria with similar success rates, we prefer under fitting to over fitting; this argument is due to two empirical arguments presented in literature: first, empirical results show that over fitting produces larger parameters bias than under fitting does (Andrews and Currim, 2003); second over fitting sometimes produce very small segments with large or unstable parameter values. Table 1 and Table 2 show the success rates (S) and rates of over fitting (O) by experimental condition and by criteria.

Table 1. Rates of Success (S) and Overfitting by information criteria and experimental condition

	AIC		AIC_3		AIC_4		AIC_c		BIC		CAIC		CAICF		ICOMP		MDL_2		MDL_5		HQ	
	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S	O
Factor 1																						
2	63%	31%	81%	5%	81%	1%	66%	26%	71%	0%	68%	0%	60%	0%	66%	26%	57%	0%	38%	0%	81%	1%
3	64%	21%	60%	8%	55%	5%	64%	18%	45%	3%	42%	3%	34%	2%	64%	17%	34%	2%	20%	1%	56%	5%
Factor 2																						
0,5	53%	27%	56%	5%	51%	2%	54%	23%	39%	1%	36%	0%	25%	0%	57%	20%	25%	0%	12%	0%	52%	2%
1	67%	24%	76%	6%	73%	3%	69%	21%	64%	2%	61%	1%	54%	1%	67%	21%	52%	1%	32%	0%	73%	3%
1,5	69%	26%	81%	8%	80%	4%	72%	22%	71%	3%	69%	3%	62%	2%	71%	22%	59%	2%	42%	1%	81%	5%
Factor 3																						
100	61%	24%	65%	6%	61%	3%	63%	19%	58%	2%	47%	1%	38%	1%	62%	21%	37%	1%	20%	0%	62%	3%
300	66%	27%	77%	6%	75%	3%	67%	26%	50%	2%	64%	2%	55%	2%	69%	22%	54%	1%	37%	1%	75%	3%
Factor 4																						
5	54%	33%	63%	8%	60%	3%	56%	28%	49%	1%	48%	1%	39%	1%	58%	27%	38%	1%	21%	0%	61%	3%
10	73%	19%	79%	4%	76%	3%	74%	16%	66%	2%	63%	2%	55%	2%	73%	15%	53%	1%	36%	1%	76%	3%
Factor 5																						
2	67%	22%	76%	7%	77%	4%	68%	21%	69%	2%	67%	2%	60%	2%	74%	13%	59%	2%	44%	1%	77%	4%
6	63%	26%	71%	6%	66%	3%	64%	22%	56%	2%	55%	2%	46%	1%	63%	25%	44%	1%	25%	0%	67%	3%
10	60%	29%	66%	6%	61%	3%	62%	23%	49%	1%	45%	1%	35%	1%	58%	25%	34%	1%	17%	0%	62%	3%
Factor 6																						
metric	64%	22%	68%	6%	64%	3%	65%	19%	53%	2%	51%	2%	41%	2%	66%	17%	41%	2%	25%	1%	64%	3%
binary	63%	30%	75%	7%	72%	3%	65%	26%	64%	2%	61%	1%	53%	1%	64%	24%	50%	1%	34%	0%	73%	4%
mixed	64%	24%	70%	6%	68%	2%	65%	21%	56%	1%	54%	1%	47%	1%	65%	22%	45%	1%	27%	0%	69%	3%
Factor 7																						
20%	69%	29%	87%	8%	88%	5%	72%	26%	83%	3%	81%	3%	72%	2%	75%	21%	72%	2%	51%	1%	88%	5%
60%	58%	22%	55%	4%	48%	1%	58%	18%	33%	0%	29%	0%	22%	0%	55%	21%	19%	0%	7%	0%	49%	1%
Factor 8																						
5-10%	55%	31%	61%	10%	58%	7%	56%	27%	47%	5%	45%	4%	38%	3%	57%	25%	36%	3%	22%	1%	59%	7%
10-20%	68%	22%	75%	5%	72%	1%	70%	19%	61%	1%	58%	1%	49%	1%	67%	21%	48%	1%	30%	0%	72%	1%
20-30%	67%	24%	77%	4%	74%	1%	69%	20%	66%	0%	63%	0%	54%	0%	72%	18%	53%	0%	35%	0%	75%	1%
Factor 9																						
normal	65%	25%	72%	6%	69%	3%	67%	21%	58%	2%	55%	2%	46%	1%	67%	20%	46%	1%	29%	1%	69%	4%
uniform	61%	26%	70%	6%	67%	3%	63%	23%	58%	2%	56%	1%	47%	1%	63%	22%	45%	1%	28%	1%	68%	3%

As example, AIC correctly identified the true number of segments in 63% of data sets with two components (Factor 1), over fitted the number of components in 31% of data sets and under

fitted the number of components in 6% of these data sets. The best overall success rates were obtained for both information criteria, as AIC_3 (71%), HQ (69%) and AIC_4 (68%), and classification criteria, as ICLBIC (70%) and ICOMPLBIC (68%). This result is consistent with (Hawkins, 1999: 70) who stated that "*augmented complete log likelihood functions may be the next generation of measures for investigation*". Although some information criteria - AIC, AIC_c and ICOMP - registered a tendency to over estimate the number of mixture components, others, as MDL_5, MDL_2, CAICF, CAIC and BIC presented higher rates of under fitting than over fitting. Almost all classification criteria (E_s, E, LP, AWE, NEC, ICL, ICLBIC, ICOMPLBIC, PC, PE, NPE, NFI and $MEAN_{ht}$) presented high rates of under fitting. When we analyse success rates by experimental condition, observe that no one criterion is best for all experimental conditions; for example AWE registered the best success rate for a 2 mixture component solution (Factor 1), while AIC, AIC_c and ICOMP presented the best results for a 3 segment solution. The criteria generally have higher success rates when there are 2 mixture components rather than 3, there are less explanatory variables, there are binary variables relatively to metric and mixed, larger separation between segments, larger samples sizes and smaller error variance. We also found that uniform distributed errors have a small negative effect on segment retention criteria success rates.

Table 2. Rates of Success (S) and Overfitting by classification criteria and experimental condition

	E_s		E		LP		AWE		NEC		CL		CLC		ICLBIC		ICOMPL		PC		PE		NPE		NFI		$MEAN_{ht}$		MIN_{ht}	
	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S	O	S	O
Factor 1																														
2	74%	26%	93%	7%	92%	8%	99%	1%	83%	17%	61%	39%	74%	26%	98%	2%	98%	2%	92%	8%	96%	4%	96%	4%	61%	39%	90%	10%	76%	24%
3	23%	23%	12%	26%	14%	27%	30%	2%	25%	30%	47%	36%	44%	31%	42%	3%	39%	4%	12%	31%	5%	15%	6%	12%	25%	16%	24%	9%	19%	49%
Factor 2																														
0,5	42%	42%	48%	16%	48%	18%	57%	1%	45%	29%	47%	44%	53%	34%	61%	3%	60%	4%	47%	19%	48%	9%	48%	8%	31%	11%	50%	15%	45%	36%
1	50%	40%	54%	15%	55%	17%	68%	1%	57%	21%	55%	37%	60%	28%	73%	2%	72%	2%	54%	18%	52%	9%	52%	8%	32%	8%	58%	8%	49%	34%
1,5	54%	39%	57%	18%	57%	18%	69%	2%	62%	19%	61%	32%	65%	24%	75%	3%	73%	3%	55%	21%	52%	10%	53%	7%	33%	5%	62%	6%	48%	39%
Factor 3																														
100	48%	11%	53%	19%	53%	20%	61%	1%	54%	26%	50%	44%	56%	34%	67%	2%	66%	3%	52%	22%	51%	10%	51%	9%	32%	8%	57%	12%	46%	42%
300	49%	11%	53%	14%	53%	15%	68%	2%	55%	21%	58%	31%	62%	24%	73%	3%	71%	4%	52%	17%	51%	8%	51%	7%	32%	9%	56%	7%	49%	31%
Factor 4																														
5	46%	11%	51%	17%	51%	18%	61%	1%	51%	27%	50%	39%	55%	30%	65%	2%	65%	4%	50%	20%	49%	10%	49%	9%	29%	8%	54%	13%	46%	37%
10	52%	12%	55%	16%	56%	17%	68%	1%	58%	20%	59%	36%	64%	27%	74%	2%	72%	3%	54%	19%	52%	9%	53%	7%	35%	8%	60%	7%	49%	35%
Factor 5																														
2	52%	16%	56%	8%	56%	9%	70%	1%	58%	14%	62%	27%	66%	19%	74%	2%	72%	3%	54%	11%	52%	5%	53%	4%	39%	11%	58%	8%	51%	28%
6	47%	9%	51%	19%	52%	20%	63%	1%	53%	25%	53%	39%	59%	29%	69%	3%	68%	4%	51%	21%	49%	11%	50%	9%	30%	8%	56%	10%	44%	40%
10	46%	8%	52%	22%	52%	24%	60%	1%	51%	31%	47%	47%	53%	38%	66%	2%	65%	3%	52%	25%	50%	13%	50%	10%	28%	6%	57%	11%	48%	41%
Factor 6																														
metric	42%	12%	50%	21%	50%	22%	62%	2%	49%	28%	54%	37%	56%	32%	68%	3%	66%	5%	50%	23%	50%	12%	50%	10%	29%	8%	54%	11%	47%	38%
binary	53%	11%	56%	10%	56%	11%	67%	1%	60%	18%	53%	40%	60%	29%	73%	2%	72%	2%	54%	14%	52%	6%	53%	4%	35%	9%	59%	9%	48%	35%
mixed	50%	11%	53%	18%	53%	20%	64%	1%	54%	24%	55%	35%	61%	26%	68%	2%	68%	3%	52%	21%	50%	11%	50%	9%	32%	7%	57%	9%	47%	36%
Factor 7																														
20%	61%	14%	60%	6%	61%	7%	78%	2%	67%	9%	59%	39%	68%	28%	86%	3%	85%	3%	59%	9%	54%	4%	55%	2%	40%	11%	66%	5%	53%	29%
60%	36%	8%	46%	27%	45%	28%	51%	1%	41%	37%	49%	36%	50%	30%	54%	2%	52%	3%	45%	30%	47%	15%	47%	14%	24%	6%	48%	14%	42%	43%
Factor 8																														
5-10%	53%	11%	54%	16%	54%	16%	60%	3%	55%	23%	52%	40%	57%	31%	64%	5%	63%	6%	53%	18%	52%	9%	52%	7%	35%	7%	58%	11%	48%	37%
10-20%	51%	9%	54%	16%	55%	17%	66%	1%	56%	23%	57%	36%	62%	27%	72%	2%	70%	3%	53%	19%	51%	9%	51%	7%	33%	8%	58%	9%	48%	36%
20-30%	42%	14%	51%	18%	51%	20%	68%	0%	51%	24%	53%	37%	58%	29%	73%	0%	72%	1%	50%	20%	50%	10%	50%	9%	28%	10%	54%	9%	45%	37%
Factor 9																														
normal	51%	11%	54%	14%	55%	16%	65%	1%	55%	23%	56%	36%	60%	28%	70%	3%	69%	3%	53%	18%	50%	0%	50%	0%	17%	0%	58%	10%	48%	35%
uniform	47%	11%	51%	19%	51%	20%	64%	1%	54%	24%	52%	39%	58%	30%	69%	2%	68%	3%	51%	21%	51%	19%	52%	15%	47%	16%	56%	9%	47%	38%

4 CONCLUSION

As we rely on heuristics as information and classification based criteria to guide us on the selection of the model to pick, it is important to understand how segment retention criteria behave. We address this question with an experimental design by manipulating nine factors. We conclude that the number of mixture, the mean separation between segments, number of consumers, number of observations per consumer, number of predictors , measurement level of predictors, error variance, minimum segment size and error distribution influence the performance of both information and classification criteria. The best overall success rates were obtained for both information criteria, as AIC_3 (71%), HQ (69%) and AIC_4 (68%), and classification criteria, as ICLBIC (70%) and ICOMPLBIC (68%). We also found that data distributional misspecification has a small negative effect on segment retention criteria success rates; however, we intend to extend this work by considering different scenarios for distribution misspecification.

REFERÊNCIAS

AKAIKE, H. (1973). "Information Theory as an Extension of the Maximum Likelihood Principle". In: B.N. Petrov e F. Csaki, Akademiai Kiado (Eds.), Second International Symposium on Information Theory, Budapest, pp. 267-281.

ANDREWS, R. L. AND CURRIM, I. S. (2003a). "Retention of Latent Segments in Regression-Based Marketing Models". *International Journal of Research in Marketing*, Vol. 20, pp. 315-321.

ANDREWS, R. L. AND CURRIM, I. S. (2003b). "A Comparison of Segment Retention Criteria for Finite Mixture Logit Models". *Journal of Marketing Research, Vol.* 40, pp. 235-243.

BANFIELD, J. D. AND RAFTERY, A. E. (1993). "Model-Based Gaussian and Non-Gaussian Clustering". *Biometrics*, Vol. 49, pp. 803-821.

BEZDEK, J. C. (1981). "Pattern Recognition with Fuzzy Objective Function Algorithms", Plenum Press.

BEZDEK, J. C., LI, W. Q., ATTIKIOUZEL, Y. AND WINDHAM, M. (1997). "A Geometric Approach to Cluster Validity for Normal Mixtures". *Soft Computing*, Vol. 1, pp. 166-179.

BIERNACKI, C. AND GOVAERT, G. (1997). "Using the Classification Likelihood to Choose the Number of Clusters". *Computing Science and Statistics*, Vol. 29, pp. 451-457.

BOZDOGAN, H. (1987). "Model selection and Akaike´s Information Criterion (AIC): the General Theory and its Analytical Extensions". *Psychometrika*, Vol. 52, pp. 345-370.

BOZDOGAN, H. (1994). "Mixture-Model Cluster Analysis Using Model Selection Criteria and a New Information Measure of Complexity". In: H. Bozdogan (Eds.), Proceedings of the Fist US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach, Vol. 2, Boston: Kluwer Academic Publishers, pp. 69-113.

BROCHADO, ANA AND MARTINS, F. V. (2005). "Assessing the Number of Components in Mixture Models: an Overview". FEP Working Paper Nº 194, Faculdade de Economia da Universidade do Porto, Porto.

CELEUX, G. AND SOROMENHO, G. (1996). "An Entropy Criterion for Assessing the Number of Clusters in a Mixture Model". *Journal of Classification*, Vol. 13, pp. 195-212.

CUTLER, A. AND WINDHAM, M. P. (1994). "Information-Based Validity Functionals for Mixture Analysis". In: H. Bozdogan (Eds.) Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modelling: An Informational Approach, pp. 149-170.

DESARBO, W. S. E CRON, WILLIAM L. (1988). "A Maximum Likelihood Methodology for Clusterwise Linear Regression". *Journal of Classification*, Vol. 5, pp. 249-282.

DESARBO, W.S., WEDEL, M., VRIENS, M. AND RAMASWAMY, V. (1992). "Latent Class Metric Conjoint Analysis". *Marketing Letters*, Vol. 3, pp. 273-288.

DIAS, J. G. (2004). "Finite Mixture Models. Review, Applications, and Computer-intensive Methods". Research School Systems, Organization and Management, Doctoral Dissertation, Groningen University.

HANNAN, E. J. AND QUINN, B.G. (1979). "The Determination of the Order of an Autoregression". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* Vol. 41, pp. 190-195.

HAWKINS, D. S., ALLEN, D. M. AND STROMBERG, A. J. (2001). "Determining the number of components in mixtures of linear models". *Computational Statistics and Data Analysis*, Vol. 38, pp. 15-48.

HURVICH, C. M. AND TSAI, C. L. (1995). "Model Selection for Extended Quasi-Likelihood Models in Small Samples". *Biometrics*, Vol. 51, pp. 1077-1084.

LIANG, Z., JASZCZAK, R. J. AND COLEMAN, R. E. (1992). "Parameter Estimation of Finite Mixture Models Using the EM Algorithm and Information Criteria with Applications to Medical Image Processing". *IEEE Transactions on Nuclear Science*, Vol. 39, pp. 1126-1133.

RAMASWAMY, V., DESARBO, W. S., REIBSTEIN, D. J. AND ROBINSON, WILLIAM T. (1993). "An Empirical Pooling Approach for Estimating Marketing Mix Elasticities with PIMS Data". *Marketing Science*, Vol. 12, pp. 103-124.

RISSANEN, J. (1986). "Stochastic Complexity and Modelling". *The Annals of Statistics*, Vol. 14, pp. 1080-1100.

RISSANEN, J. (1987). "Stochastic Complexity". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 49, pp. 223-239.

RIVERA, F. F., ZAPATA, E. L. AND CARAZO, J. M. (1990). "Cluster Validity Based on the Hard Tendency of the Fuzzy Classification". *Pattern Recognition Letters*, Vol. 11, pp. 7-12.

ROUBENS, M. (1978). "Pattern Classification Problems and Fuzzy Sets", *Fuzzy Sets and Systems* Vol. 1, pp. 239-253.

SARSTEDT, M. (2008). "Market segmentation with mixture regression models: understanding measures tat guide models selection". *Journal of Targeting, Measurement and Analysis*, Vol. 16, pp. 228-246.

SCHWARZ, G. (1978). "Estimating the Dimension of a Model". *The Annals of Statistics*, Vol. 6, pp. 461-464.

VRIENS, M., WEDEL, M. AND WILMS, T. (1996). "Metric Conjoint Segmentation Methods: A Monte Carlo Comparison". *Journal of Marketing Research*, Vol. 33, pp. 73-85.

WEDEL, M. AND DESARBO, W. S. (1994). "A Review of Recent Developments in Latent Class Regression Models". In: Bagozzi, Richard P. (Eds.), Advanced Methods of Marketing Research, Blackwell, pp. 352-388.

WEDEL, M. AND DESARBO, W. S. (1995). "A Mixture Likelihood Approach for Generalized Linear Models". *Journal of Classification*, Vol. 12, pp. 21-55.

WINDHAM, M. P. AND CUTLER, A. (1992). "Information Ratios for Validating Mixture Analysis", *Journal of the American Statistical Association*, Vol. 87, pp. 1188-1192.

LA CRISIS ECONOMICA ACTUAL

NELSON ÁLVAREZ VÁZQUEZ
E-mail: nalvarez@cee.uned.es
PABLO RAYEGO SERIÑAN
E-mail: prayego@cee.uned.es
JULIÁN RODRÍGUEZ RUIZ
E-mail: julián21@cee.uned.es
Departamento de Economía Aplicada Cuantitativa I
Universidad Nacional de Educación a Distancia

Resumen

Actualmente se habla de crisis imputando a los economistas su incapacidad para explicarla y predecir su final. La comunicación se basa en una definición de *E. Working* ciencia que ha de aplicarse a la economía, definiendo de paso lo que ha de entenderse por metodología.
La razón se atribuye a la agregación de los datos, como tiene lugar en otros fenómenos de actualidad como el cambio climático.
Se ofrece una explicación cualitativa de la existencia de las crisis económicas distinta, a las habituales basadas en las posiciones habituales de culpar al Sistema Bancario, o a razones similares, ligadas con la discusión de liberalismo o intervención. Es un hecho cierto que los comentaristas no saben que decir aunque alguno ya ha diagnosticado que se conocen las causas de la actual crisis. Se utiliza para la medición datos de paro y precios en USA, recordando la Gran Depresión de 1929, porque es un antecedente invocado.

Palabras Clave: Crisis, predecir, metodología, y medición.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

1. INTRODUCCIÓN

En esta Comunicación, se ofrecen consideraciones aparentemente genéricas, que buscan la sustitución en econometría de la hipótesis de la probabilidad por una hipótesis basada en el estudio de los movimientos sistemáticos en el tiempo, es decir, la sustitución de lo irregular por lo regular, como ocurre en otras ciencias. Ya la economía matemática de *Walras*, era concebida como análisis funcional, y determinístico.

Actualmente se habla de crisis económica, y de su explicación, imputando por lo general a los economistas, una incapacidad esencial para explicarla y para predecir su final, e imputando a determinados grupos sociales, la razón de ser de la crisis actual, que no parece cierto. Mas bien, la crisis ni es medible ni es predecible por los métodos matemáticos de la estadística, pudiendo ser que una explicación y predicción, de una crisis no sean temas científicos. La Comunicación, se basa en una definición de ciencia como la determinación de regularidades, que ha de ser aplicable a economía, definiendo de paso el significado de metodología. La principal a razón de las dificultades de la medición de una crisis económica se atribuye a la naturaleza agregada de los datos económicos, característica que puede afirmarse de otros fenómenos actuales como el cambio climático,. Una segunda razón para la imposibilidad de medición, estaría en la dimensión temporal del dinero. Estas serían las dos razones principales.

Con estas dos razones, se ofrece una explicación cualitativa de la existencia de las crisis económicas, distinta a las habituales y demagógicas, basadas en las disculpas habituales de echar la culpa al *Sistema Bancario*, o a grupos similares, relacionadas con la discusión del dilema, entre liberalismo (hoy se habla de neoliberalismo) e intervencionismo. Ha de tenerse por cierto que los comentaristas de los periódicos, radio y televisión, no saben que decir al respecto, si bien alguno ya ha diagnosticado conocer las causas de la actual crisis, afirmación,

sorprendente. La explicación de las crisis, se basa en la agregación de los datos, de difícil solución, dada la dificultad de disponer de datos referidos a unidades individuales, por razón del secreto estadístico. La agregación, posiblemente genera heterogeneidades que la propia Estadística no puede resolver.

En los ejemplos, *Moore* (1908), en su artículo: "*The Statistical Complement of Pure Economics",* habla de un solo atributo como las estaturas o las crisis económicas fechadas por Juglar, es decir, se habla de entes homogéneos como las estaturas. La estadística, la etapa final de las matemáticas, como la teoría de Juegos y la Informática, considera un atributo, como las estaturas o los pesos, no los dos, no siendo comparables, por heterogeneidad o por la diferencia entre las estaturas y los pesos.

La teoría económica racional seguía un criterio causal, basado en la utilidad. *Moore* escribe sus artículos de 1895 como si fueran de economía racional. Su libro de 1911 ha de verse como extensión de la economía racional,

Siendo de difícil aplicación en la economía, como en toda ciencia de observación los métodos elaborados para ciencias experimentales. *Moore* adoptaría de la *Biometría,* experimental, su interés por la correlación y la regresión múltiple, (ambos atemporales) criticada la regresión múltiple por *Friedman (1991)*, quien solo ve en ella un método de obtener hipótesis tentativas. Moore cuestionaría la generalidad de las leyes económicas en el espacio y tiempo.

La historicidad, contraria a la generalidad, un antecedente invocado por la estadística descriptiva era un obstáculo para la generalidad de las leyes económicas. Eliminados el espacio y el tiempo por los economistas clásicos y neoclásicos, pueden llegar a formular leyes generales.

Los lenguajes estadísticos, basados en conceptos sintéticos y formales, como media y correlación, que son atemporales. Lo atemporal permite hablar del papel del tiempo en economía teórica, vinculado a la sustitución de una economía de trueque, por una economía monetaria, basada en el oro, con dimensión temporal, teniendo el oro, con un valor permanente, en el tiempo, la temporalidad, trasladada al dinero fiduciario, Dado que los bienes tienen en el tiempo un valor caduco, no permanente se explicarían las crisis económicas como la actual por la diferencia en el valor entre el dinero y los bienes. Los agentes económicos acumularían (generarían tendencias crecientes) valor en los tiempos sin crisis, que en determinado momento, ha de detenerse. Los objetos de consumo y producción tienen fecha de caducidad, el dinero, tiene un valor permanente. No ocurre algo similar con los fenómenos carentes de valor en el tiempo, por ejemplo, en un teatro o en un cine, las entradas corresponden a los asientos disponibles.

Moore, debiera haber utilizado el término "sustituir" en vez del término "complementar" la economía pura, en su artículo de 1908, dado que se refería a una economía causal, basada en el criterio de la utilidad, que el buscaba sustituir por una aproximación estadística.

La necesidad de desarrollar las teorías económicas en el tiempo, afecta por igual a teoría económica y a las teorías las probabilidades, cuyos conceptos básicos como la regresión, la correlación o a la misma probabilidad, son atemporales. En el caso concreto de teoría económica se pone de manifiesto con la apelación por los matemáticos a elaborar una teoría económica dinámica, todavía inexistente, una teoría que resolviera los problemas imputables a la estática, buscando un sustituto de la teoría económica estática (o atemporal), siendo ejemplos de ello *Schultz* y *Roos,* ambos matemáticos, y coetáneos. En probabilidades, se habla de la regularidad o de la estabilización de las frecuencias relativas en torno a una constante, denominada probabilidad. Las frecuencias relativas del suceso "cara" en el lanzamiento de una moneda, debieran estar alrededor de 0.5, supuesto que la probabilidad sea aproximada por las frecuencias relativas, no por el grado de creencia racional de *Keynes* en su *Treatise on Probability.,* subjetivo y no medible. Las crisis económicas se deberían a que en etapas anteriores a la crisis, el dinero posibilitaría un crecimiento, incompatible con las condiciones reales de la economía, como el ahorro en términos reales, no el nominal y ni el deflactado por un índice agregado de precios. Esta sería nuestra explicación. Las crisis surgirían como un fenómeno cualitativo, causal, no teniendo sentido recomendar políticas fiscales.

La predicción del final de la crisis, algo que los comentaristas aseguran tener, sin fundamento sólido, podría basarse en estudios como el de *Oyen,* La atención prestada a la predicción en economía se compagina mal con su objeto. La comunicación trata del problema

de la imposibilidad de predecir en economía, en particular, las crisis económicas, analizando un ejemplo histórico, como la Gran Depresión de 1929. No se trata de una imposibilidad metafísica, sino de una imposibilidad basada en la heterogeneidad de los datos agregados. El hecho de que se trate de datos agregados, explicaría que sea posible obtener predicciones ajustadas. Pero si cambian los componentes del agregado, la predicción se ajustará mal a los hechos.

Como todo estudio empírico, dedicado a analizar un problema general, requiere establecer las coordenadas espacio temporales, en este caso, los años 1930, y los Estados Unidos, teniendo el inconveniente de que las conclusiones no son de valor general y necesario. Las coordenadas espacio-temporales derivan de la naturaleza de datos de series de tiempo. Si el estudio empírico no permite generalizaciones, es posible establecer las coordenadas en las que ha de moverse el análisis. Cabe apuntar que si en cuestiones como las mediciones de teorías económicas establecidas racionalmente ha existido una desviación empirista, cabe intuir lo ocurrido con problemas concretos en los que la aplicación de una teoría racional es controvertible.

La inmensa literatura existente, no facilita una interpretación sencilla. Buscando no crear un planteamiento nuevo, se sigue a *Oyen* (pp.313 - 353, 1991). En la presentación comenta que la predicción de la Gran Depresión centrada en 1929, suscita todavía controversias considerables 60 años más tarde, y la presentación actual, podría justificar hablar de ella 70 años más tarde. *Oyen* menciona dos posibles causas de la Gran Depresión: la falta de demanda y el intervencionismo gubernamental (*Oyen,* p. 313, 1991). Ilustra el problema con dos citas de afirmaciones vertidas en octubre de 1929: por una parte, *I. Fisher*, consideraba improbable una caída bursátil, es decir, un descenso de las cotizaciones, y por otra parte, la *Sociedad de Harvard*, afirmaba ser improbable una crisis como la de 1920 - 21. Concluye *Oyen* que la Gran Depresión, habría venido a refutar semejantes predicciones:"El hecho de que la Depresión sucedió refutaba la creencia de que un suceso tal fuera imposible" (*Oyen* p. 315, 1991). No es el lugar para creencias, y respecto a las proposiciones científicas, convendría recordar la afirmación de *Stigler y Becker* (1977), de que no se puede probar nada respecto a hechos concretos. Denominaremos a las predicciones anteriores, profecías. Lo cual, vuelve sobre la naturaleza empírica del estudio (o de los hechos).

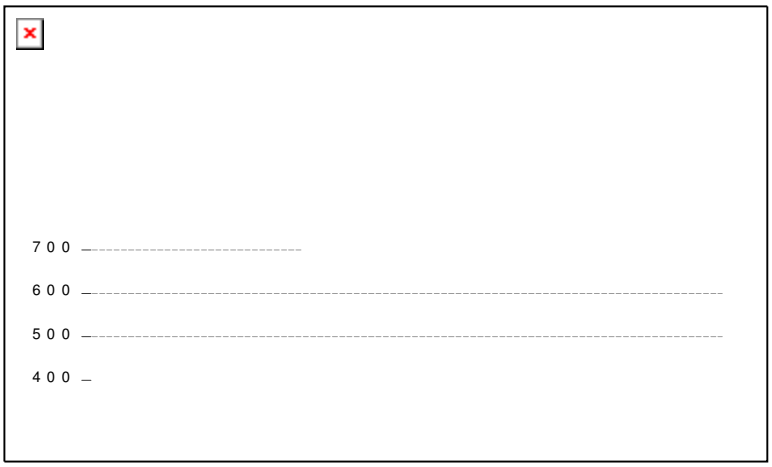
Oyen (p. 316, 1991) parece atribuir los fallos predictivos a la ausencia de datos, concluyendo en la imposibilidad de la predicción, revelando el papel importante de los historiadores:"... los economistas constataron que los datos disponibles en los años 1930 eran manifiestamente inadecuados para mostrar lo que estaba sucediendo", no pareciendo sostenible una afirmación relativa a la inadecuación de los datos, salvo por su naturaleza agregada.

2. ANÁLISIS DE LOS HECHOS (O UNA DESCRIPCIÓN CUANTITATIVA DE LA GRAN DEPRESIÓN).

Oyen, como historiador económico, revisa las series de tiempo en los años de la Gran Depresión de 1929: el PNB, la producción industrial, las cotizaciones bursátiles, el desempleo, etc., reproducidas en las tablas que siguen.

2.1 PNB, observado y potencial.

Su tabla de datos, la 12.1, reproduce dos series históricas del PNB (USA), el observado (consumo, inversión, gasto público y exportaciones) y el potencial, siendo ambos agregados: "El PNB potencial se ha calculado en el supuesto de que la economía potencial estaba fechada en 1929, y que en ausencia de la Gran Depresión hubiera experimentado un crecimiento anual del 3%. La figura... muestra un descenso en la producción de $211 billones en los cuatro años de 1929 a 1933, seguida de una recuperación de $ 196 billones durante los cuatro años siguientes. El PNB observado en 1936 estaba por debajo del de 1929... El gap entre el PNB observado y el Potencial, en 1930 no habría desaparecido hasta 1942...el coste de la Depresión es el área entre las líneas potencial y observada" (p. 317, 1991). EL supuesto es propio de un historiador económico.



La **figura II.1** representa las dos serie originales, de PNB: el potencial y el observado, con mínimo en 1933.

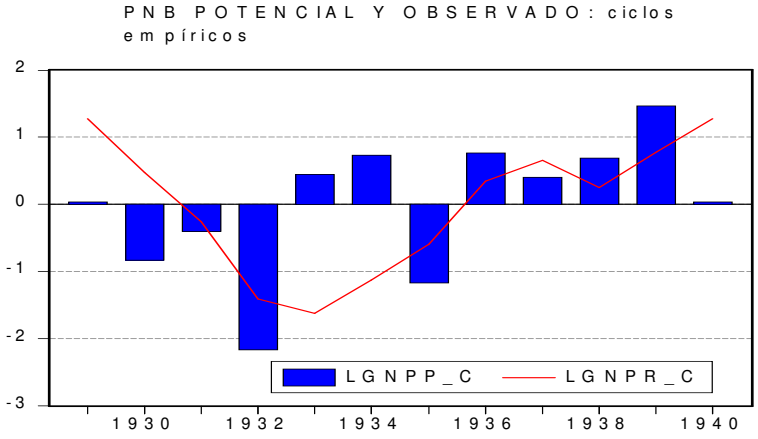


figura II.2

La figura II.2 representa en el tiempo los ciclos empíricos del PNB, potencial y observado. La eliminación de la tendencia sugiere una depresión en el PNB potencial, registrando un mínimo absoluto en 1932, seguido de otro relativo en 1935. Los ciclos modificarían la cronología temporal. El PNB observado tendría sus mínimos en 1932 y en 1935, con una explicación: "Lo que revela... es que el descenso en el PNB real de 1929 a 1933 no se extendió a los cuatro componentes principales del PNB". La tabla II.1 reproduce la descomposición periódica de los ciclos empíricos, representados los periodogramas en la figura II.3.

Tabla II.1: Periodogramas de los ciclos empíricos (1929-1940).

Períodos	PNB observado	PNB potencial
12.00000	**83.55215**	**36.66603**
6.000000	14.01481	8.636022
4.000000	0.817097	7.270824
3.000000	**1.209158**	**24.68145**
2.400000	0.277869	20.87820
2.000000	0.128918	1.867468

Oyen como un historiador se refiere tan solo a los agregados en los años de la Gran Depresión de 1929.

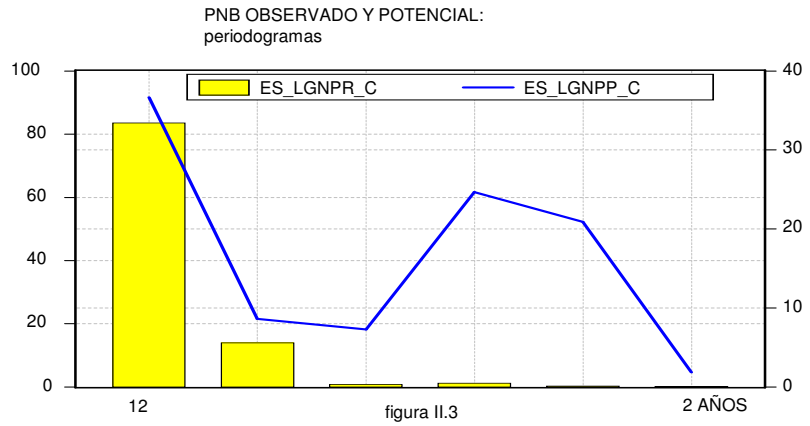


figura II.3

Tabla II.1a: datos básicos.

Obs	GNPP (PNB potencial)	GNPR (PNB real)	LGNPP_C (ciclo empírico del PNB potencial)	LGNPR_C (ciclo empírico del PNB real)
1929	710.0000	710.0000	0.000000	0.000000
1930	731.0000	643.0000	-0.000336	-0.106849
1931	753.0000	588.0000	-0.000168	-0.203995
1932	775.0000	509.0000	-0.000855	-0.356003
1933	799.0000	499.0000	0.000159	-0.383573
1934	823.0000	537.0000	0.000270	-0.317910
1935	847.0000	580.0000	-0.000470	-0.248608
1936	873.0000	662.0000	0.000281	-0.124099
1937	899.0000	695.0000	0.000144	-0.083182
1938	926.0000	664.0000	0.000251	-0.136540
1939	954.0000	717.0000	0.000557	-0.067475
1940	982.0000	773.0000	0.000000	0.000000

La tabla II.1a, reproduce los datos básicos.

Oyen (pp. 317 - 318, 1991) analiza los cuatro componentes principales del PNB: consumo, inversión, gasto público y exportaciones netas.

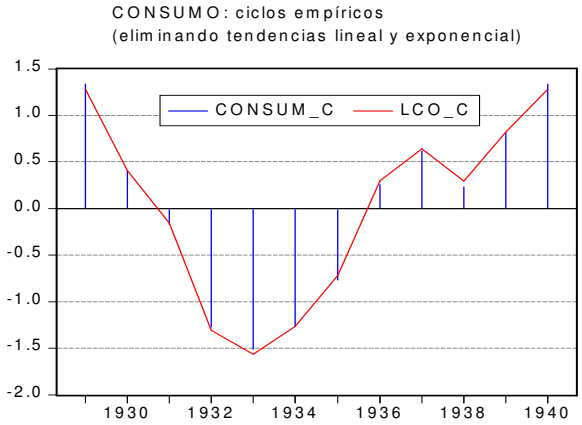


figura II.4

La figura II.4 mide el consumo, representando las dos alternativas a la medición del ciclo de consumo, basadas en eliminar dos tendencias, la una lineal, y la otra exponencial, viendo que no

hay diferencia en los ciclos calculados, un hecho frecuente, cuando la historia es corta, permitiendo suponer que el desconocimiento de la forma de la tendencia, puede no ser grave, revelando la inutilidad de la discusión entre lineal y no lineal. La tendencia es un agregado desconocido.

PNB OBSERVADO Y COMPONENTES:
ciclos empíricos

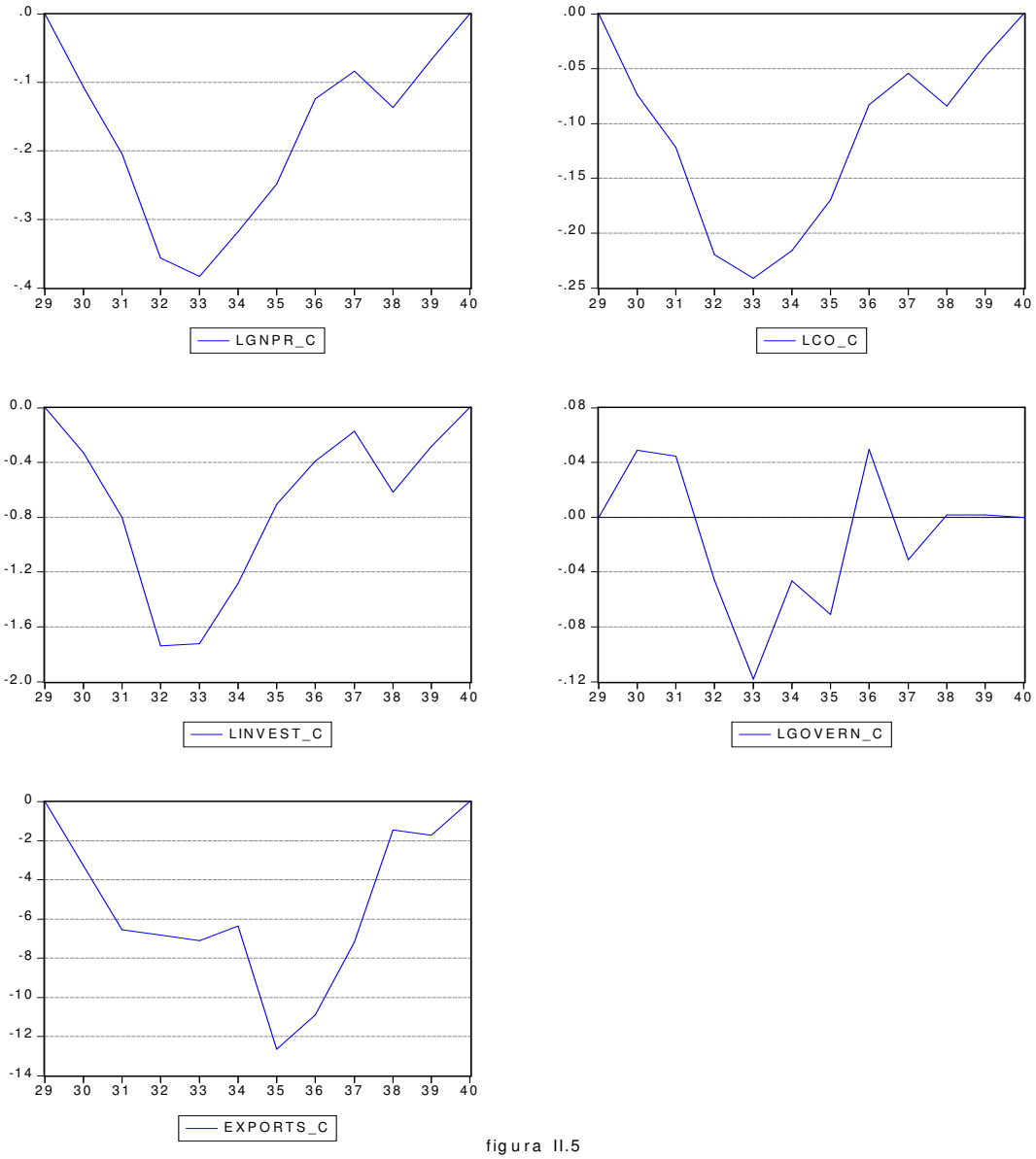


figura II.5

La figura II.5 compara el ciclo empírico del PNB con los ciclos de sus componentes. En las exportaciones se ha eliminado tendencia lineal, dado que una exponencial, chocaría con los números negativos. Todos los ciclos indican su mínimo en 1933, salvo las inversiones en 1932 y las exportaciones, en 1936. Es decir, los componentes del *PNB* concuerdan en la trayectoria temporal de la *Gran Depresión*. Los datos se reproducen en la tabla II.2a.

Tabla II.2a: datos básicos.

obs	INVEST	CONSUM	EXPORTS	GOVERN
1929	139.0000	471.0000	5.000000	94.00000
1930	98.00000	440.0000	2.000000	103.0000
1931	60.00000	422.0000	-1.000000	107.0000
1932	23.00000	385.0000	-1.000000	102.0000
1933	23.00000	379.0000	-1.000000	99.00000
1934	35.00000	391.0000	0.000000	111.0000
1935	61.00000	412.0000	-6.000000	113.0000
1936	82.00000	452.0000	-4.000000	133.0000
1937	100.0000	468.0000	0.000000	128.0000
1938	63.00000	457.0000	6.000000	138.0000
1939	86.00000	481.0000	6.000000	144.0000
1940	112.0000	503.0000	8.000000	150.0000

La semejanza en el tiempo de los ciclos empíricos torna un sin sentido considerar el PNB (o la renta) endógeno, y la inversión como exógena, como hace *Haavelmo* (1947). En su tabla 12.2 desglosa *Oyen* los cuatro componentes para tres años, 1929, 1933 y 1940, por lo que no se analizan, al no disponer de los datos para todos los años. En todos los casos, salvo en el gasto público, el mínimo se sitúa en 1933.

2.2 Precios, empleo y salarios.

La tabla 12.3.de *Oyen* reproduce el deflactor del PNB, sus tasas de variación anual de inflación, de empleo, y desempleo y de los salarios por hora reales.

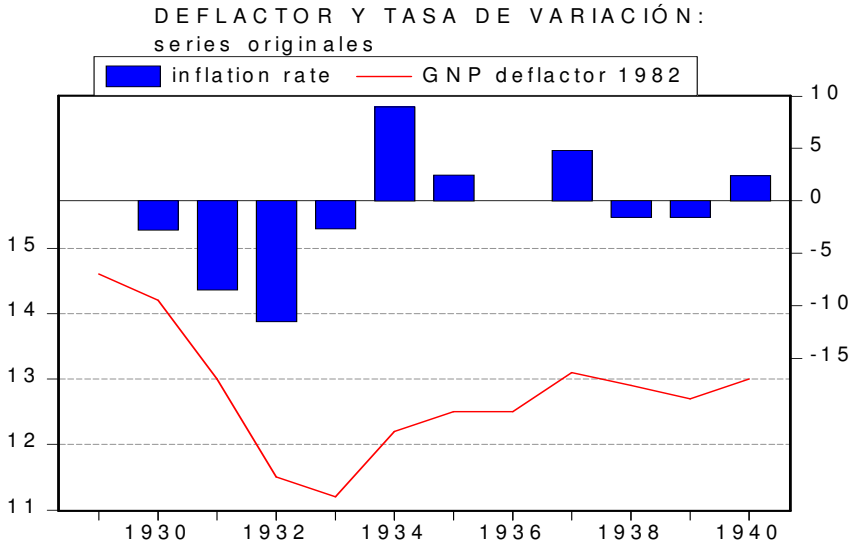


figura II.6

Representados en la figura II.6, las series originales del deflactor del PNB, así como de la tasa de variación anual, interpretada, como medición de la inflación., midiendo antes movimientos de los precios reales que de los nominales, que basada en el ciclo empírico tendría una trayectoria temporal similar a la de la tasa.

El precio del dinero, particularizado en la inflación, representaría un movimiento nominal.

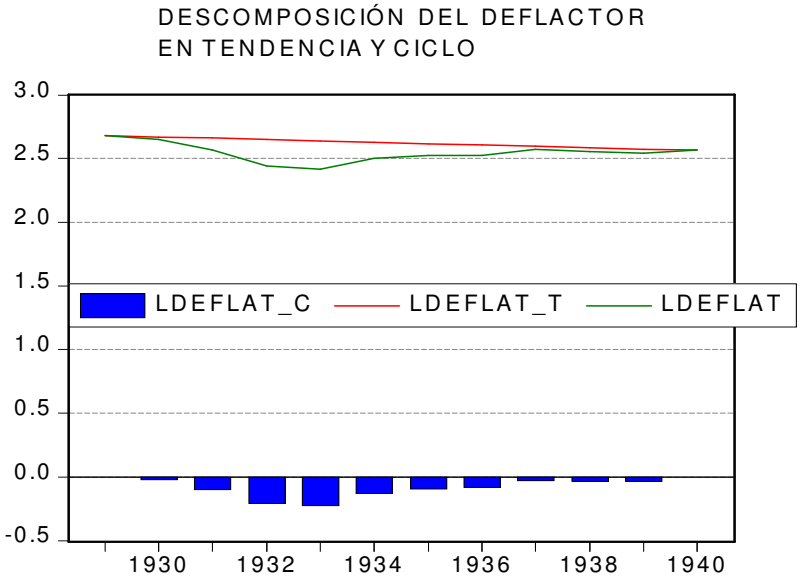


figura II.7

La figura II.7 representa la descomposición del deflactor en su tendencia (interpretada como deflación por ser decreciente) y en ciclo empírico (movimiento de los precios reales).

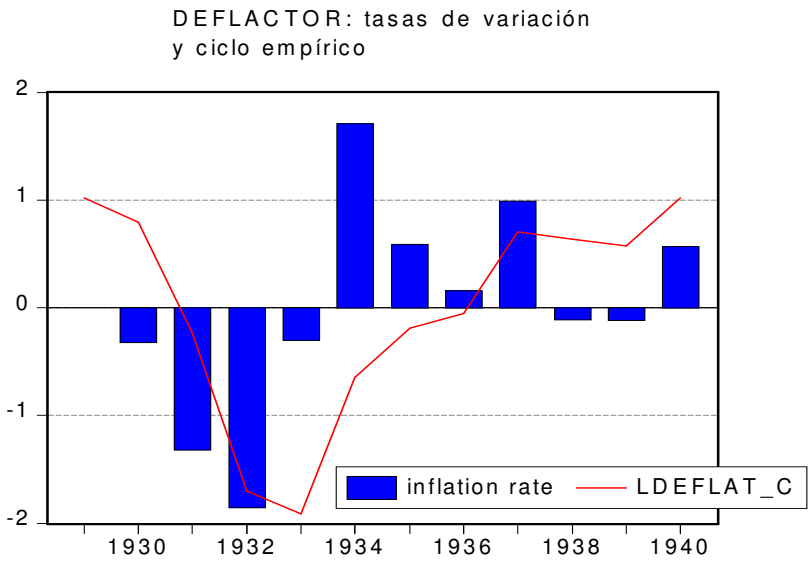


figura II.8

Interpretado el ciclo empírico como un movimiento de los precios reales, la figura II.8 sugiere que estos y la inflación un movimiento en el tiempo similar, confirmando la medición por las tasas del movimiento real.

Tabla II.2 b: deflactor, ciclo empírico y tasas de variación anual

DEFGNP	LDEFLAT_C	TDEFGNP
14.60000	0.000000	NA
14.20000	-0.017228	-2.739726
13.00000	-0.094968	-8.450704
11.50000	-0.207018	-11.53846
11.20000	-0.222900	-2.608696
12.20000	-0.126825	8.928571
12.50000	-0.091981	2.459016
12.50000	-0.081429	0.000000
13.10000	-0.023993	4.800000
12.90000	-0.028826	-1.526718
12.70000	-0.033899	-1.550388
13.00000	0.000000	2.362205

.

La tabla II.2 b reproduce los datos del deflactor, del ciclo y de las tasas.

Tabla II.2c: periodogramas del ciclo empírico del deflactor y de la tasa de variación (1929 - 1940).

Períodos	Deflactor (ciclo).	Tasas de variación
12.00000	**75.67732**	23.54218
6.000000	17.96877	**45.75440**
4.000000	3.674419	19.74465
3.000000	2.555456	8.086149
2.400000	0.118645	2.243849
2.000000	0.005387	0.628776

La tabla II.2 reproduce la descomposición periódica de las tasas de variación y del ciclo empírico del deflactor. Mientras el ciclo empírico registra la mayor contribución en los 12 años, las tasas indicarían su mayor contribución en los 6 años, no dando la misma importancia a todos los movimientos. La cuestión estaría en conocer cual es importante, si el movimiento de 12 o de 6 años, una cuestión en economía siempre desconocida, y que no cabe llegue a conocerse. Los economistas suelen suponer entre PIB y empleo, una relación estrecha. Otros datos de la tabla II.9 se refieren al mercado de trabajo. Es frecuente utilizar como indicador el ratio de desempleo.

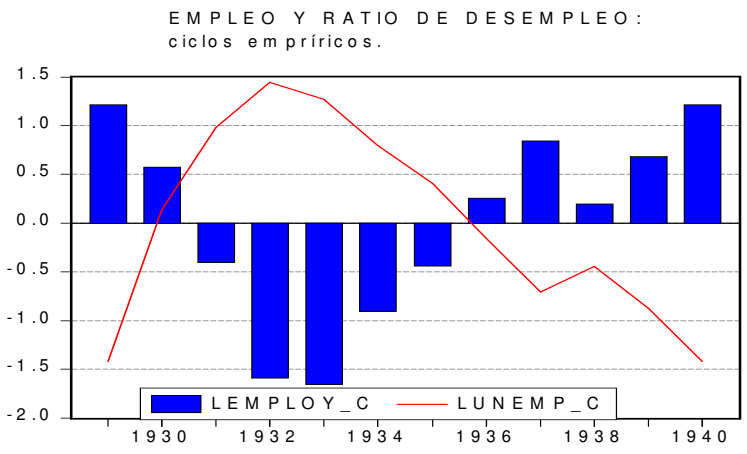


figura II.9

La figura II.9 representa en el tiempo, los ciclos empíricos de empleo y del ratio de desempleo. La evidencia sugiere movimientos contrarios, con el máximo de desempleo en 1932, y el mínimo en desempleo en 1933:"La caída en la producción de 1923 a 1933 redujo la demanda de empleo" (*Oyen,* p. 318, 1991). El ciclo parece coherente con esta conclusión, no si el ratio, una evidencia contraria al empleo de ratios en el tiempo.

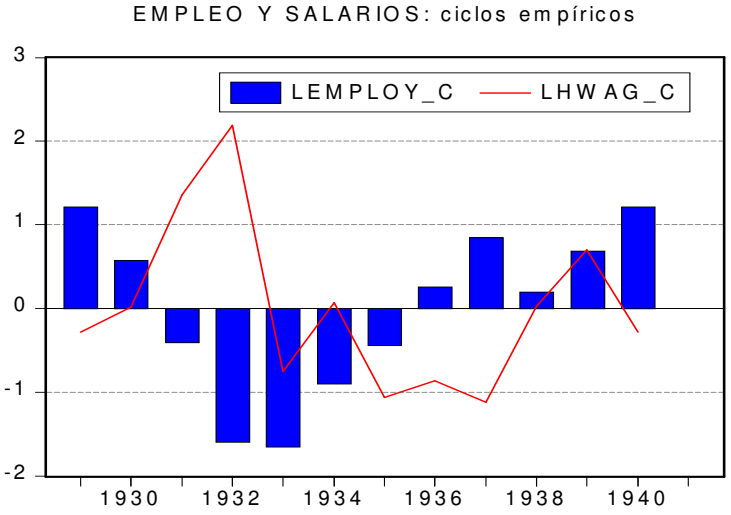


figura II.10

En cambio, la relación normal, parece estar entre los dos elementos del mercado de trabajo, las cantidades (o el empleo) y precio (o salarios). En la figura II.10 se relacionan, los ciclos empíricos de empleo y salarios, representación de los movimientos reales, en una medición de oferta y demanda de este producto. Para obviar la deflación, estando las series originales deflactadas, se obtienen las series nominales. De hecho, la serie de salarios, en dólares de 1982, es convertida en una serie de salarios nominales, multiplicando los salarios de 1982, por el deflactor.

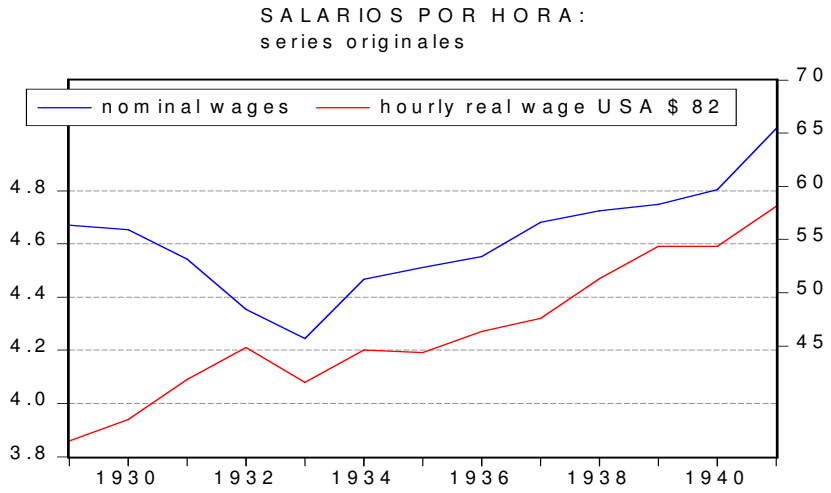


figura II.11

Este comportamiento se representa en el tiempo en la figura II.11, en las dos series de salarios (la nominal y la deflactada) con diferencias entre 1929 y 1932. Puede decirse que el ascenso de los salarios deflactados, se debería a que el deflactor habría decrecido en dichos años, siendo verosímil pensar que con la *Gran Depresión*, los salarios reales habrían descendido... Parecería que el ciclo empírico aproximase mejor los movimientos reales.

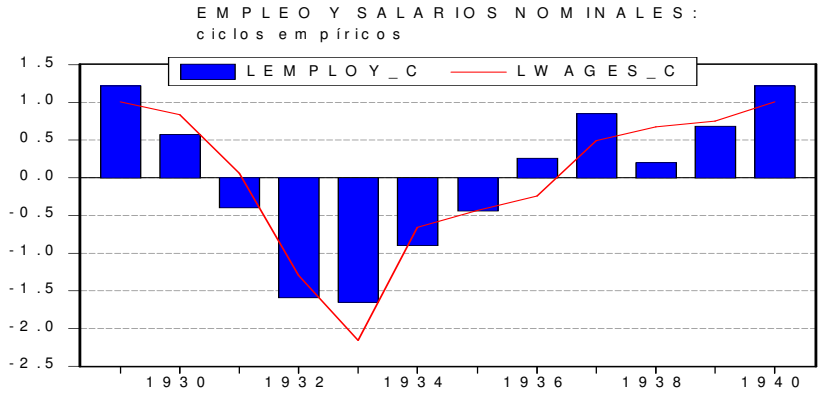


figura II.12

Se calcula el ciclo empírico de los salarios nominales, revelando la misma trayectoria temporal que el empleo, por lo que cabe esperar una ley de pendiente positiva. El descenso de los salarios en términos reales en los primeros años de la *Gran Depresión*, parece más razonable que la descripción de las figuras II.10 y II.11. Se trataría de medir la elasticidad.

Tabla II.3: correlación de los ciclos empíricos de empleo y salarios nominales.

Sample: 1929 1940

Included observations: 12

Correlations are asymptotically consistent approximations

LEMPLOY_C,LWAGES_C(-i)	LEMPLOY_C,LWAGES_C(+i)	Lag	lead
. |********* |	. |********* |	**+ 0.9370**	**+0.9370**
. |*****. |	. |******* |	0.5167	0.7387
. |* . |	. |*** . |	0.0786	0.2712
. ***| . |	. *| . |	-0.2931	-0.1050
.*****| . |	. ****| . |	-0.4829	-0.3573
. ***| . |	. ****| . |	-0.3412	-0.3600
. ***| . |	. ****| . |	-0.2467	-0.3708
. **| . |	. ***| . |	-0.2315	-0.2853
. *| . |	. *| . |	-0.0482	-0.0833
. |* . |	. |* . |	0.0760	0.0769
. |* . |	. |* . |	0.1538	0.1352

La tabla II.3, reproduce la correlación entre los ciclos empíricos. Como el máximo se registra en el retardo nulo, la evidencia numérica, no permite determinar la dirección de la influencia, si procede del empleo o de los salarios.

Tabla II.4: Elasticidad de pendiente positiva del empleo a salarios.

Dependent Variable: LEMPLOY_C
Method: Least Squares
Sample: 1929 1940
Included observations: 12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LWAGES_C	**+ 0.920203**	0.108532	8.478608	0.0000
C	-0.019932	0.010948	-1.820500	0.0987
R-squared	**+ 0.877880**	Mean dependent var		-0.086999
Adjusted R-squared	0.865668	S.D. dependent var		0.071544
S.E. of regression	0.026222	Akaike info criterion		-4.293439
Sum squared resid	0.006876	Schwarz criterion		-4.212621
Log likelihood	27.76063	F-statistic		71.88680
Durbin-Watson stat	1.912686	Prob(F-statistic)		0.000007

Solo se calcula elasticidad (relación de empleo a salarios), (tabla II, 4, figura II.12 A). Es + 0.92, con un valor cercano a la unidad.

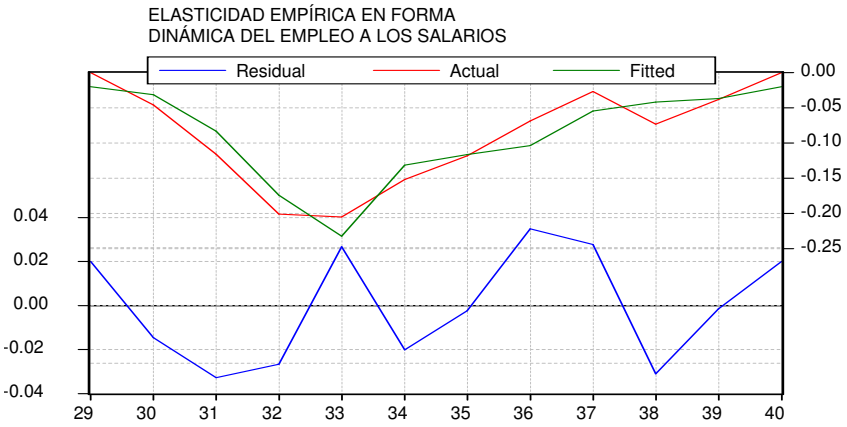


figura II.13

La regresión representada en el tiempo en la figura II.13, en la parte inferior tiene las discrepancias de la regresión, indicando sin explicación otras causas del empleo.

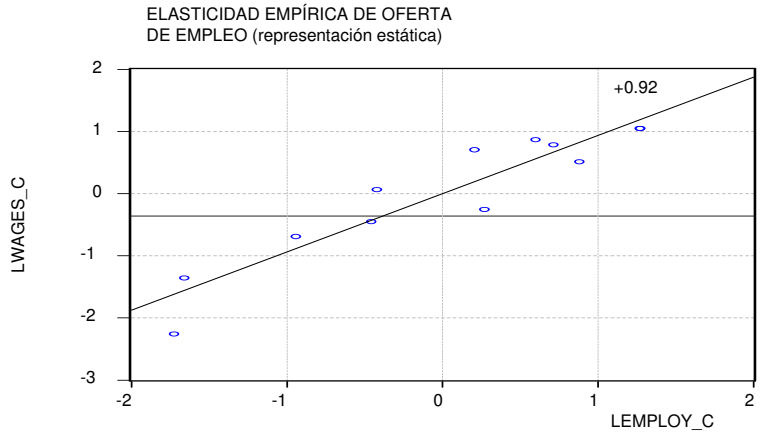


figura II.14

En forma atemporal, la elasticidad de pendiente positiva se representa en la figura II.14.

Tabla II.5: descomposición periódica de los ciclos empíricos (1929-1940).

Periodos	Empleo	Salarios
12.00000	**78.92538**	**78.95344**
6.000000	17.93576	12.12051
4.000000	0.487068	4.576368
3.000000	**2.195785**	**3.339865**
2.400000	0.270601	0.740915
2.000000	0.185404	0.268900

Los periodogramas (tabla II.5), registran picos comunes en los 12 y 3 años.

Tabla II.6: correlación entre los ciclos periódicos de 12 años.

Sample: 1929 1940

Included observations: 12

Correlations are asymptotically consistent approximations

LEMPLO12,LWAGE12(-i)	LEMPLO12,LWAGE12(+i)	i	lag	lead
. |**********	. |**********	**0**	**+0.9909**	**+0.9909**
. |******* |	. |******** |	1	0.6951	**0.8185**
. |*** . |	. |*****. |	2	0.2640	**0.4584**
. **| . |	. | . |	3	-0.1607	0.0413
.*****| . |	. ***| . |	4	-0.4597	- 0.3042
******| . |	.*****| . |	5	-0.5697	- 0.4911
.*****| . |	.*****| . |	6	-0.4954	- 0.4954
. ***| . |	. ****| . |	7	-0.2997	- 0.3558
. *| . |	. **| . |	8	-0.0746	- 0.1523
. |* . |	. | . |	9	0.0933	+ 0.0260
. |* . |	. |* . |	10	0.1537	+ 0.1148

El pico común en los 12 años y el análisis de sus ciclos periódicos, con la correlación máxima en el retardo nulo, se reproduce en la tabla II. 6. En el primer retardo, la dirección de la influencia parecería ir de los salarios al empleo.

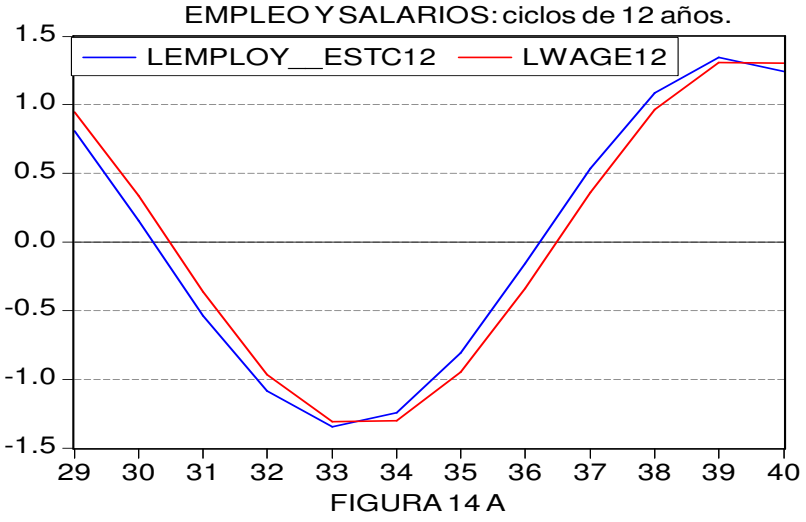


FIGURA 14 A

La figura II.14A, indicaría prevalecer flexibilidad (o influencia del empleo sobre los salarios), pareciendo existir una contradicción entre la correlación atemporal y los movimientos en el tiempo. La flexibilidad tomarías un valor de + 1.009.

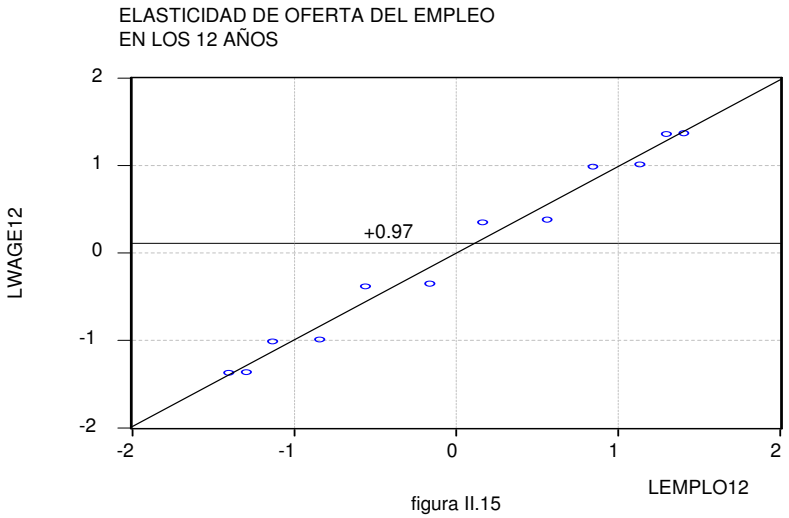


figura II.15

La elasticidad de pendiente positiva del empleo a los salarios en los 12 años es similar a la empírica,+ 0.97, confirmando la hipótesis de que el ciclo empírico resulte determinado por los ciclos largos.

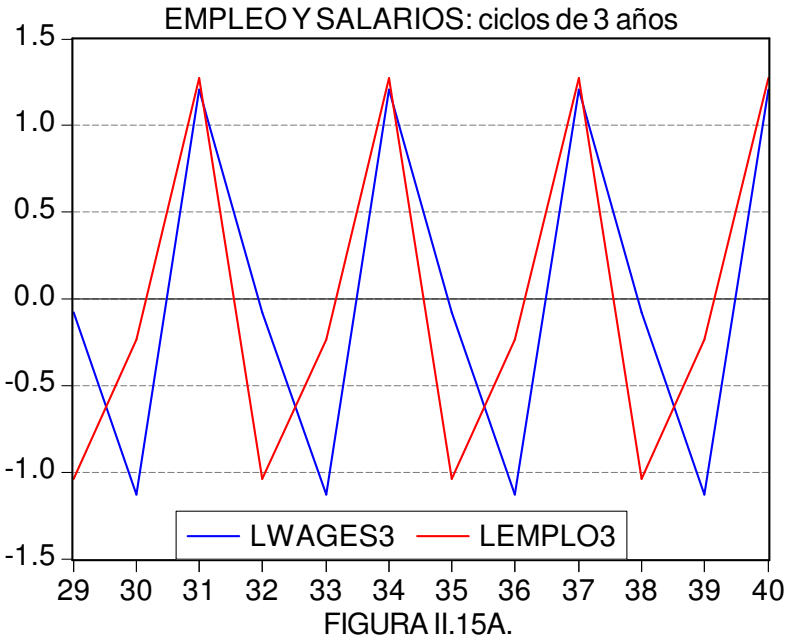


FIGURA II.15A.

Nuevamente se aprecia una contradicción similar, al menos aparente, entre la correlación (atemporal) y los movimientos en el tiempo (figura II.15A), en los ciclos de 3 años con una elasticidad del empleo a salarios en los 12 años es similar a la empírica, + 0.97.

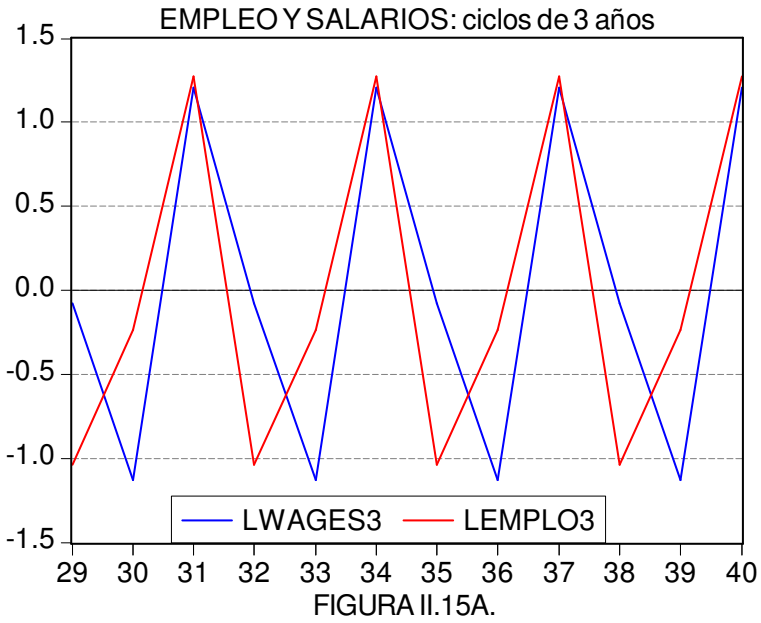


FIGURA II.15A.

Tabla II.7: correlación entre los ciclos periódicos de 3 años de empleo y salarios.
Sample: 1929 1940

Included observations: 12

Correlations are asymptotically consistent approximations

LEMPLO3,LWAGES3(-i)	LEMPLO3,LWAGES3(+i)	i	Lag	Lead
. |******* |	. |******* |	0	0.6854	0.6854
*********| . |	. |*** . |	**1**	**-0.8591**	0.2970
. |** . |	*********| . |	2	0.2068	-0.8442
. |*****. |	. |*****. |	3	0.5141	0.5141
******| . |	. |** . |	4	-0.6158	0.2250
. |* . |	******| . |	5	0.1349	-0.6008
. |*** . |	. |*** . |	6	0.3427	0.3427
. ****| . |	. |* . |	7	-0.3725	0.1530
. |* . |	. ****| . |	8	0.0629	-0.3575
. |** . |	. |** . |	9	0.1714	0.1714
. *| . |	. |* . |	10	-0.1291	0.0811

La tabla II.7, indicaría la correlación máxima, en los 3 años, en el desfase de un año, con dirección de la influencia de los salarios al empleo.

Tabla II.8: elasticidad de pendiente negativa del empleo a los salarios en los 3 años.

Dependent Variable: LEMPLO3
Method: Least Squares
Sample(adjusted): 1930 1940
Included observations: 11 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LWAGES3(-1)	**- 0.790674**	0.064463	-12.26555	0.0000
C	-0.144783	0.004858	-29.80367	0.0000

R-squared	0.943554	Mean dependent var	-0.085998
Adjusted R squared	0.937282	S.D. dependent var	0.010508
S.E. of regression	0.002632	Akaike info criterion	-8.879500
Sum squared resid	6.23E-05	Schwarz criterion	-8.807155
Log likelihood	50.83725	F-statistic	150.4437
Durbin-Watson stat	2.750000	Prob(F-statistic)	0.000001

Tabla II.8: elasticidad de pendiente negativa del empleo a los salarios en los 3 años.

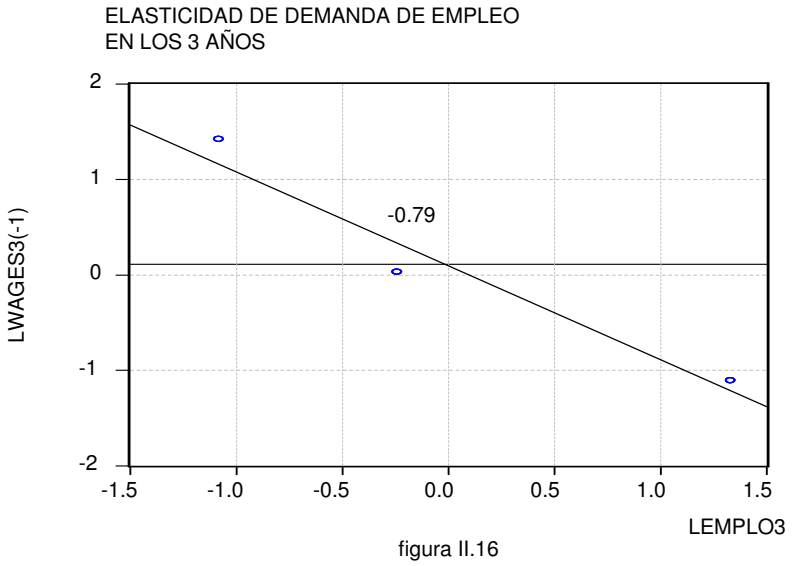


figura II.16

La elasticidad se representa en forma atemporal en la figura II.16, con su valor en la tabla II.8.

2.3 El análisis del mercado bursátil y monetario.

La tabla 12.4 de *Oyen* contiene las series históricas de cotizaciones bursátiles, de quiebras bancarias y de dinero, reproducidas en la tabla II.8A, junto con los ciclos empíricos de cotizaciones de títulos y dinero.

Tabla II.8A: datos primarios y ciclos empíricos de cotizaciones y dinero.

obs	STOCKS	M1	LSTOCKS_C	LM1_C
1929	260.0000	26.60000	0.000000	0.000000
1930	210.0000	25.80000	-0.135374	-0.066940
1931	137.0000	24.10000	-0.484300	-0.171507
1932	69.00000	21.10000	-1.091975	-0.340849
1933	90.00000	19.90000	-0.748072	-0.435806
1934	98.00000	21.90000	-0.584714	-0.376443
1935	106.0000	25.90000	-0.428042	-0.245090
1936	155.0000	29.50000	0.030144	-0.151346
1937	154.0000	30.90000	0.101872	-0.141384
1938	115.0000	30.50000	-0.111948	-0.190817
1939	121.0000	34.20000	0.017110	-0.112722
1940	110.0000	39.70000	0.000000	0.000000

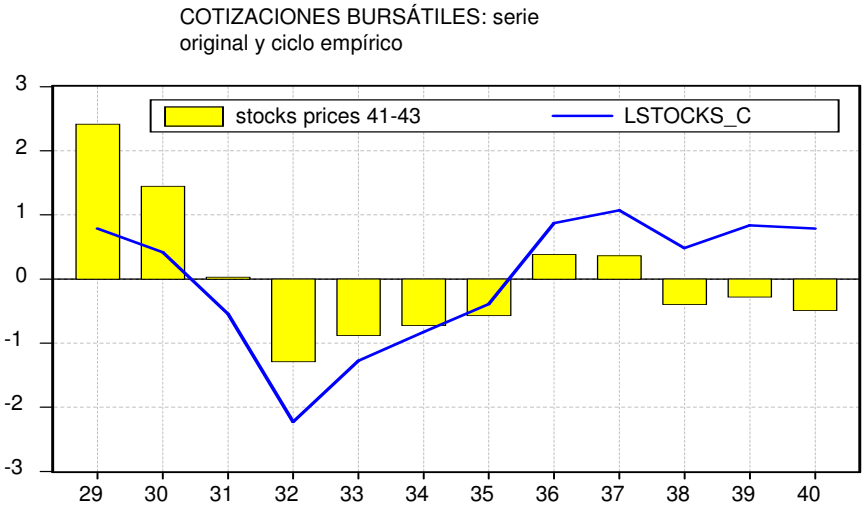


figura II.17

Eliminando la tendencia, se fecharía el mínimo en 1932: "El descenso de las cotizaciones en octubre de 1929 fue grande, un 33%. Los lectores deben tener en cuenta que el deterioro de las condiciones económicas contribuye al descenso de las cotizaciones: cuando decrece la actividad económica, descienden los beneficios de las corporaciones, y con ellos, el valor de las cotizaciones que representa un derecho sobre los mismos" (*Oyen*, pp. 320 - 321, 1991). La explicación de este hecho, atestiguado por la representación gráfica y los datos, no menciona el precio del dinero (o el interés), y según estudios basados en agregados como oferta de dinero y renta, como los de *Friedman y Schwartz*, sobre la Gran Depresión de 1929, mantenidos por algún comentarista la semejanza de la crisis actual a la Gran Depresión de 1929, criticando *Friedman y Schwartz*, la postura de la *R. F USA,* por su comportamiento con la cantidad de dinero u oferta monetaria.

En consecuencia, se considera la imposibilidad de predicción de agregados y del final de la crisis, por ser considerada imposible la medición en economía con los métodos estadísticos, con datos agregados (posiblemente heterogéneos), que pueden revelar su inexistencia, como la del PIB, de la Inflación, o de su relación con los precios al consumo, en la medición económica, buscando sin conseguirlo, sustituir una economía racional y causal, por otra economía de contenido empírico (experimental).

3. EXPLICACIÓN DE LA ACTUAL CRISIS.

En la figura I se representa un índice agregado de los precios de los cuatro productos inicialmente considerados por *Moore (1914*). El índice agregado coincide en su trayectoria temporal con el de maíz, sirviendo esta evidencia para negar la existencia de agregado en los precios al consumo, El argumento sería válido si lo productos fueran más de 4.

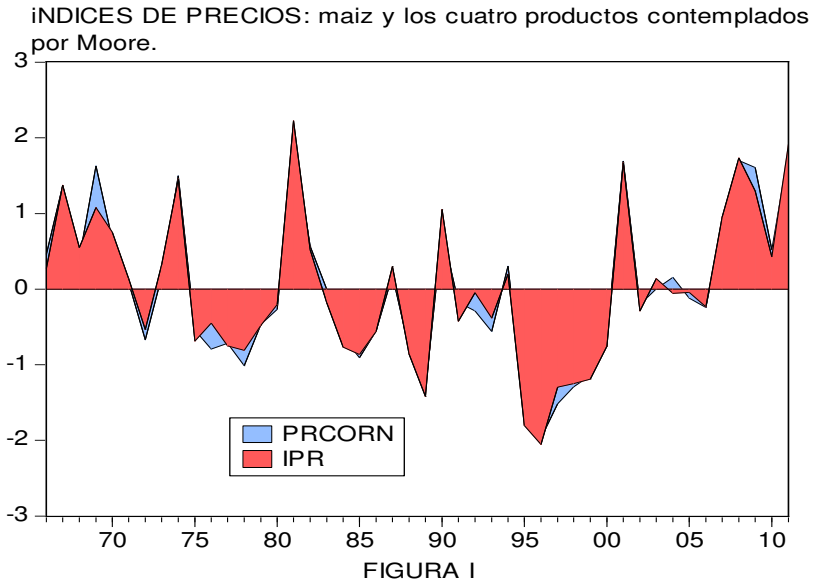


FIGURA I

Por tanto, la explicación de la crisis se basa en dos razones ya apuntadas: la agregación de los datos económicos, cuya existencia, defendida por *Vining* (la existencia de un agregado), por parte del *NBER*, parece cuestionable, a juzgar por la evidencia de la figura y en una segunda razón, en la naturaleza temporal del dinero, una explicación que siendo causal y cualitativa, puede ser novedosa.

Que los datos económicos son agregados, es un hecho ya barruntado en la aplicación del análisis de *Fourier*, tratando de encontrar una explicación real en economía, a los distintos ciclos periódicos, obtenidos en las frecuencias ortogonales de Fourier, resultando imposible encontrar una explicación real, a las periodicidades calculadas de este modo. No hay ciencia que resuelva los problemas de la heterogeneidad, ni la Estadística, ni la Teoría de Juegos, ni la Informática, exigiendo la noción de ciencia, no pareciendo razonable, y aceptable la idea de *Bacon*, de reducir la ciencia a experimental, si bien sigue siendo útil en las ciencias experimentales, de dudoso valor en las ciencias no experimentales, como la economía, o ciencia de observación.

Un problema clave para definir lo que se entienda por ciencia en una disciplina no experimental, parece ser la validez del criterio de ajuste, entre las observaciones y teoría, no pareciendo justificado. En estadística, se suele usar el criterio MCO, que en un principio alega a su favor la de ser objetivo. El criterio de ajuste en la ciencia, debido a *Bacon,* no se ha discutido, simplemente se ha aceptado. Pero existiendo ya una ciencia basada en la causalidad, "el conocimiento por la causas" entra en juego el criterio causal. La causalidad, siendo una regularidad atemporal, no en el tiempo, es un ente lógico. Con buena razón se objetaba a la economía racional, la posibilidad de deducir sus conclusiones económicas a partir de un número reducido de premisas.

Por ciencia, se entiende la determinación de las regularidades, sinónima de una ley o de una teoría. Como regularidad atemporal estaría la causalidad, y como regularidad en el tiempo, estarían la tendencia y los ciclos, o movimientos sistemáticos. Las tendencias, son confundidas a veces con los ciclos, como al hablar de tendencias cíclicas, de ciclo - tendencia, o de ciclos de

crecimiento, (en este caso, con tendencias crecientes) las tendencias, son movimientos sistemáticos no recurrentes, como los ciclos son movimientos sistemáticos y recurrentes. La tendencia se distingue del ciclo, por la recurrencia o repetitividad. En ambos movimientos sistemáticos se basa la aproximación propuesta para la economía, como ciencia, siendo una aproximación alternativa a la probabilidad.

No todo ha de ser científico en el saber, como no lo es la música clásica o el arte, si bien lo científico da prestigio, aconsejando la división del trabajo que sean las Universidades, las instituciones encargadas de impartir la ciencia.

Los ciclos han encontrado dificultades de aceptación en economía por la irregularidad de los movimientos económicos medidos en las series de tiempo (o históricas) irregularidad postulada por la Probabilidad y por la historia, influyendo en Econometría, o Estadística, la historia mediante la Escuela histórica alemana, como ilustran las alusiones de *Giannetti*, en su estudio de las crisis en el siglo XIX, a S*chumpeter,* a *Kitchin*, a *Kondratieff* y a otros historiadores económicos. El peso de los historiadores, enorme en Econometría, que han de verse como una mera prolongación de la Estadística. *Descosieres* (1998), se refiere a los estadísticos de la escuela histórica alemana. De modo que la historia ejercería su influencia en la medición económica a través de la estadística. *Louca* (2001) afirma que en un principio, en el origen de la econometría, estaba una visión matemática, debida a *R. Frisch*, y otra, una visión histórica debida a S*chumpeter,* imponiéndose la visión matemática sobre la histórica. En consecuencia contempla, la historia, mal a la idea de ciencia, como determinación de regularidades, pues un historiador buscaría la singularidad de los hechos, no su regularidad.

La idea de tendencia, como movimiento agregado, parece una generalización al tiempo de la idea estadística de media aritmética, atemporal, y como revelan los promedios móviles, medias en el tiempo, siendo criticados en su empleo a raíz del efecto *Slutzky*., quien solo afirma la posibilidad de generar los ciclos económicos en la aleatoriedad, algo que los exégetas han tomado como un hecho probado, favoreciendo el probabilismo. Otro problema de ajuste es la naturaleza matemática de la función que aproximando la tendencia, en último extremo, puede conducir al dilema entre lo lineal y lo no lineal. *Persons* solo aceptaba las tendencias lineales, no encontrando significado económico a las tendencias no lineales. ¿Cómo se puede saber matemáticamente el carácter lineal o no lineal de una tendencia?, interrogante sin respuesta. En otros trabajos se define un ciclo como una desviación de la linealidad. En otros estudios se han considerado tendencias, a las funciones matemáticas exponenciales, dada la finalidad y creencia en poder medir las elasticidades, hoy siendo ello una duda, por defender la no medición económica por los métodos estadísticos. El empleo de una función exponencial, se justificaba por la aceleración de las trayectorias temporales al final de las series de tiempo.

Por Metodología, se entiende la Lógica de la Filosofía, aplicada a la Estadística. Así se entiende que dentro de metodología, se incluyan en algunos Congresos, los desarrollos informáticos. Si un día existió la Lógica, hoy existe la Metodología, de la que se habla mucho, sin definirla, como la ciencia, siendo esta la primera vez que se critica en una Comunicación la idea de ciencia de *Bacon.*

En periodos anteriores se discutía entre racionalismo y empirismo, o entre deducción e inducción como entre métodos de hacer ciencia. Basta con citar a *J. S. Mill o a Jevons,* existiendo una evolución en el tiempo desde el transito del racionalismo al empirismo.

La causalidad, criterio ligado a la utilidad de los agentes económicos, consumidores y productores, no parece aplicable a funciones como las teorías macro – económicas de consumo que suponen una relación entre consumo y renta, relaciones de demanda – renta, cuestión que no han podido resolver las diferentes aproximaciones tanto las deductivas como las inductivas, siendo una prueba la inexistencia de una teoría económica dinámica (en el tiempo), o de la aleatoriedad en el tiempo, dado que la noción de proceso estocástico, considerada fácilmente generalizable al tiempo, en términos matemáticos, no se ha probado su existencia.

Los económetras buscaban sustituir las teorías basadas en la utilidad (en las unidades individuales) por otras aproximaciones estadísticas, (basadas en unidades agregadas), pensando en sustituir las leyes de la economía pura por leyes estadísticas, establecidas con base en datos de series de tiempo o históricas (el papel influyente de los historiadores!), no con el propósito como dice *Moore (1908)* de complementar la economía pura con una medición basada en los

métodos estadísticos. Un ejemplo de influencia nociva de las matemáticas en economía puede estar en identificar el signo matemático de una pendiente con una ley económica, causal, basada en la utilidad como la ley de demanda con una pendiente negativa, la de oferta con pendiente positiva.

Se consideraba la economía basada en la Ética (*Aristóteles* en la *Ética a Nicómaco* menciona la palabra economía por primera vez), o en la *Psicología*, disciplinas ambas, basadas en unidades individuales, sin agregados, sean homogéneos o heterogéneos, invalidando considerar la economía como ciencia, salvo que se demuestre la existencia de regularidades en los agregados y su homogeneidad. No se entiende que *Keynes* afirme que la ley económica de macro – economía de consumo – renta es una ley psicológica fundamental, cuando el se esta refiriendo a una ley predicada de agregados, y toda ley psicológica está basada en unidades individuales, llevando a muchos economistas a la confusión de individual y agregado. Cuestión la de las unidades, relacionada con la división de la ciencia económica en micro – economía y macro - economía.

Los filósofos encontraron que la ciencia de la economía se planteaba en términos de unidades individuales, de modo que lo agregado (léase macroeconomía, como las crisis macroeconómicas bien podrían no ser ciencia), algo que los comentaristas citados consideran inaceptable. En otras palabras: ¿Existe la p m c en USA o en cualquier otro país, calculada por *T. Haavelmo* o por otro estadístico*?* El que no se haya obtenido todavía, lleva a la duda sobre la existencia de regularidades lógicas en los agregados de consumo y renta, no a las dudas de la existencia de las regularidades en el tiempo. *Charemza y Deadman* (1997) han cuestionado, la existencia de la p m c por la inestabilidad del ratio consumo/renta en el tiempo, siendo ahora el problema, lógico, y no temporal. La cuestión ofrece mayor enjundia como es la existencia de la economía como ciencia.

La Econometría, en su intento de sustituir causalidad por correlación, ha de entenderse no como un complemento de la economía pura, sino como un sustituto, discrepando de *Moore* (1908). Como indicaba *King* (1917), la correlación, siendo atemporal podría reducirse a la causalidad, categoría lógica: la correlación sería la misma causalidad lógica, aplicada a grupos de individuos o a unidades agregadas, posiblemente heterogéneas entre si. Cuando se confiaba en los métodos estadísticos, se pensaba en sustituir lo causal por algo semejante, se hablaba de las perturbaciones aleatorias como englobando otras causas. Ahora bien, ni esta resuelto el problema de la heterogeneidad de las unidades individuales, sin precisar el significado de una influencia económica, determinante pero no causal. En micro - economía ya se contemplaba el paso de las unidades individuales a otras más agregadas como industrias, pudiendo verse en *Lipsey* un ejemplo, en su libro de economía.

Pensadores como *Cournot* y *Marshall,* habrían dedicado más reducido espacio al tema estadístico de la medición, que al de la agregación en economía.

Otra de las razones que abonan la importancia del tiempo en economía es la necesidad de establecer supuestos como el mantenimiento en el futuro de las condiciones actuales. Un supuesto de esta naturaleza, careciendo de argumentos a su favor no podría explicar una crisis económica, como la actual. La crisis actual no podría estar generada en una etapa de crecimiento. Se ha afirmado la inexistencia de una teoría económica ni deductiva ni inductiva que contemple el tiempo, añadiéndolo a las formulaciones estáticas, en el lenguaje continuo como mediante un cociente de diferenciales, o en el discreto, como una tasa de variación, lo cual parece relevante hablar de tasas de variación por ser los datos de series de tiempo (series históricas) y la aplicación de las tasas, tiende a amplificar los movimientos cortos, los que parecen irregulares, atenuando los movimientos medios y largos, con mayor apariencia de regularidad.

No existe una teoría económica dinámica, o en el tiempo. Solo existe una teoría económica estática (atemporal) causal, basada en la utilidad, que parece de difícil sustitución por una teoría económica en el tiempo (o dinámica), dificultad que puede llevarse a la idea de regularidad, o a una ciencia causal.

4. CONCLUSIONES

a) Se niega le existencia de agregados como el PIB o el IPC, siendo la evidencia gráfica la razón principal, contra la aserción de Vining. De ello se sigue la inutilidad de la modelización econométrica, no siendo de interés, los valores de los coeficientes, sean estimadores o parámetros.

b) Se afirma la inexistencia de una teoría económica en el tiempo, siendo sorprendente que *Schultz* abogue por ella, publicando muy próximo en el tiempo a *Roos.*

c) La inexistencia de la macro - economía como ciencia.

5. BIBLIOGRAFÍA

CHAREMZA, W. W. and DEADMAN. D. F. (1997): New Directions in Econometric Practice. General to Specific Modelling, Co integration, and Vector Auto regression. Edward Elgar, U.K.

DESROSIERES, A. (1998): The Politics of Large Numbers. A History of Statistical Reasoning. Harvard University Press. England.

FRIEDMAN, M. (1991): Alternatives Approaches to Analyzing Economic Data. The American Economic Review, vol. 81, pp. 39- 49.

GIANNETTI, R. (1988):Crisis Económicas: El siglo XIX. Oikos – Tau. Barcelona.

HAAVELMO, T. (1947)::Methods of Measuring the Marginal Propensity to consume. Journal of the American Statistical Association, XLII, 237, pp. 105-122.

KING, G. A. (1963): An Appraisal of the Strengths and Weakness of the Econometric Approach. Journal of Farm Economics, XLV, 5, pp. 1408 - 1419.

KING, W. I. (1917): The Correlation of Historical Economic Variables and the Misuse of Coefficients in Connection. Publications of the American Statistical Association, XV, 120, pp. 847 - 853.

KLEIN, J. L. (1997): Statistical Visions in Time. A History of Time Series Analysis, 1662-1938.Cambridge University Press. U.S.A.

LOUCA, f. (2001): Intriguing Pendula: Founding Metaphors in the Analysis of Economic Fluctuations. Cambridge Journal of Economics.

MOORE, H. L. (1895): Von Thünen´s Theory of Natural Wages. The Quarterly Journal of Economics, IX, 3, pp. 291-304.

- (1895a): Von Thünen´s Theory of Natural Wages. The Quarterly Journal of Economics, IX, 4, pp. 388-408.

- (1907): The differential Law of Wages. Journal of the Royal Statistical Society, LXX, pp. 638-651.

- (1908): Statistical Complement of Pure Economics. Quarterly Journal of Economics, XXIII, pp. 1 - 33.

- (1911): Laws of Wages. An Essay in Statistical Economics. Reprinted of Economic Classics. Augustus M. Kelley. Publishers.(1967). New York.

- (1912): A Reply to professor Edgeworth´s Review of Professor H. L. Moore, "Law of Wages". The Economic Journal, XXII, 86, pp. 317-323.

- (1914): Economic Cycles: Their Law and Cause. Reprinted of Economic Classics. Augustus M. Kelley. Publishers.(1967). New York.

- (1917): Forecasting the Yield and the Price of Cotton. Reprints of Economic Classics. Augustus M. Kelly Publishers. New York (1967).

- (1921): Generating Cycles of Products and Prices. Quarterly Journal of Economics, XXXV, pp. 215 - 239.

- (1921a): Generating Cycles Reflected in a Century of Prices. Quarterly Journal of Economics, XXXV, pp. 502-523.

- (1921b): The Origin of the Eight - Year Generating Cycle. Quarterly Journal of Economics, XXXVI, pp. 1-29.

- (1922): Elasticity of Demand and Flexibility of Prices. Journal of the American Statistical Society. March.

- (1923): Generating Economic Cycles. New York. Macmillan.

-(1925): A Moving Equilibrium of Demand and Supply. Quarterly Journal of Economics, XXXIX, 3, pp. 351-373.

- (1926): Partial Elasticity of Demand. The Quarterly Journal of Economics, XL, 3, pp. 393-401.

- (1926a): Pantalleoni´s Problem in the Oscillation of Prices. Quarterly Journal of Economics, XL, 4, pp. 586-596.

- (1926b): A Theory of Economic Oscillations. Quarterly Journal of Economics, XLI, 1, pp. 1-29.

- (1929): Synthetic Economics. Reprinted of Economic Classics. Augustus M. Kelley. Publishers.(1967). New York.
MORGAN, M. S. (1990): The History of Econometric Ideas. Cambridge University Press. Cambridge.
OYEN, D. B. (1991): Business Fluctuations and Forecasting. Deaborn Financial Publishing, Inc. U.S.A.
PERSONS, W. M. (1910): The Correlation of Economic Statistics. Journal of the American Statistical Association, Vol. 12, pp. 289-294.
- (1917): On the Variate Difference Correlation Method and Curve Fitting. Publications of the American Statistical Association, XV.
- (1919): An Index of General Business Conditions. Review of Economic and Statistics, 1, pp. 111 - 212.
- (1925): Statistics and Economic Theory. Review of Economic Statistics. Vol. VII, Num. 3, pp. 179 - 197.
- (1926): Theories of Business Fluctuations. Quarterly Journal of Economics, XLI, pp. 94-128. Del libro, The Foundations of Business Cycle Theory, Edited by D. P. O'Brien, pp. 3 - 37. Great Britain.
ROOS, C. F. (1927): A Dynamical Theory of Economics. Journal of Political Economy, XXXV, pp. 632-656.
- (1934): Theoretical Studies of Demand. Econometrica, Vol. II, pp. 73-90.
- (1948): A Future Role for the Econometric Society in International Statistics. Econometrica, XVI, N. 2, pp. 127-134.
SCHULTZ, H. L. (1924): The Statistical Measurement of the Elasticity of Demand for Beef. Journal of Farm Economics, VI, 3, pp 254-278.
- (1925): The Statistical Law of Demand as Illustrated by the Demand of Sugar, I. Journal of Political Economy, XXXIII, pp. 481-504 (Washington, Children's Bureau); II; Journal of Political Economy, XXXIII, pp.577-631.
- (1925a): An Extension of the Methods of Moments. Journal of the American Statistical Association, XX, 150, pp. 242-244.
- (1927): Theoretical Considerations Relating to Supply. Journal of Political Economy, XXXV, pp. 437-464.
- (1927a): Mathematical Economics and the Quantitative Method. Journal of Political Economy, XXXV, pp. 702-706.
- (1927b): Cost of Production, Supply and Demand, and the Tariff. Journal of Farm Economics, IX, 2, pp, 192-209.
- (1927c): Theoretical Considerations Relating to Supply. Journal of Political Economy, XXXV, pp. 437-464.Chicago.
- (1928): Statistical Laws of Demand and Supply with Especial Application to Sugar. University of Chicago Press. Chicago.
- (1928a): Rational Economics. The American Economic Review, Vol. XVIII, pp. 643-648.
- (1929) Marginal Productivity and the General Pricing Process Journal of Political Economy, XXXVII, pp. 505-551.
- (1931): The Italian School of Mathematical Economics. Journal of Political Economy, XXXIX, pp. 76-85.
- (1931a): The Shifting Demand for Selected Agricultural Commodities, 1875-1929. Journal of Farm Economics, XIV, nº 2, pp. 201-227.
- (1932): The Shifting Demand for Selected Agricultural Commodities, 1875-1929. Journal of Farm Economics, XIV, 2. pp. 201-227.
-(1933): A Comparison of Elasticities of Demand Obtained by Different Methods. Econometrica, Vol. 1, pp.274 - 308.
- (1933a):Interrelations of Demand. Journal of Political Economy, XL, pp. 468-512.
- (1935):Interrelations of Demand, Price and Income. Journal of Political Economy, XLIII, pp. 433-481 (data).
- (1938): The Theory and Measurement of Demand. The University of Chicago Press. Chicago.
SCHULTZ, T. (1939): Scope and Method in Agricultural Economics Research. Journal of Political Economy, XLVII, pp. 705-717.
SCHULTZ, H. and VAILEY, R. S. (1924): Elasticities of Demand and Coefficient of Correlation. Quarterly Journal of Economics, XXXVIII, pp. 169-171.
SLUTZKY, E. (1937): The Summation of Random Causes as the Source of Cyclic Processes. Econometrica, V, pp. 105 - 146.
STIGLER, G., and BECKER, G. (1977): De Gustibus Non Est Disputandum. The American Economic Review, LXVII, 2, pp. 76-90.
VINING, R. (1949): Methodological Issues in Quantitative Economics. Koopmans on the Choice of Variables to be Studied and of Methods of Measurement. A Rejoinder. The Review of Economics and Statistics, XXXI, pp. 77-94.

THE CONTINUOUS ALGEBRAIC REPRESENTABILITY PROPERTY

JUAN C. CANDEAL
e-mail:candeal@unizar.es
Deparatmento de Análisis Económico
Universidad de Zaragoza

Resumen

En el presente trabajo introducimos la propiedad de representación continua en un contexto algebraico. Presentamos una caracterización de dicha propiedad en el caso de álgebras reales. Mostramos también una aplicación de estas técnicas a un problema de elección social.

Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

The continuous representability property in algebraic enviroments is introduced. We provide a characterization of this property for the case of real algebras. An application of this algebraic approach to the context of social choice theory is also offered.

Keywords: Algebraic utility/Social choice.
Thematic Area: Quantitative Methods for Economics and Business.

1 INTRODUCTION

In this paper we study a topological property, called the continuous representability property (briefly, CRP), in the context of totally preordered algebraic structures. Let (X,τ) be a topological space. The topology τ (or the toplogical space (X,τ)) has CRP if any continuous total preorder defined on X has a continuous real-valued order-monomorphism. This property was introduced by Herden in [14], under the name of a useful topology, and was suggested by the statements of the two major results in topological utility theory: namely, Eilenberg's theorem and Debreu's theorem, respectively (see the introduction of Section 3 below). Utility theory is a part of Decision Sciences that mainly has to do with numerical representations of binary relations defined on a nonempty set. The continuous representability property has focused the attention of some researchers since then (see, e.g., [4,5,14,15]). CRP is discussed in Section 3 and some new results are presented, including a general characterization in terms of countability properties of related topologies to τ. In the current paper an extension of CRP in the algebraic setting, called continuous algebraic representability property (shortly, CARP), is introduced. Basically, whereas CRP asks for special types of numerical representations (the continuous ones), CARP asks for more specific ones. Indeed, those which are continuous and, in addition, preserve the algebraic structure (i.e., they are algebraic-homomorphisms). This requires some kind of compatibility among order, topology, and algebra involved. In Section 4, we pay attention to the case of totally preordered real algebras and then CARP is discussed in this context. Although we could have introduced CARP in a more general framework (such as totally preordered semigroups or groups) we have decided to focus on totally preordered real algebras, mainly for the applications in decision theory. In Section 5, we develop an application of the previous approach to social choice theory. In particular, we prove an Arrow-like theorem which characterizes strongly dictatorial social welfare functionals.

2 PRELIMINARIES

A *preorder* $\P$ on an arbitrary nonempty set X is a binary relation on X which is reflexive and transitive. An antisymmetric preorder is said to be an *order*. A *total preorder* $\P$ on a set X is a preorder such that if $x, y \in X$ then $[x \P y]$ or $[y \P x]$.

If $\P$ is a preorder on X, its *asymmetric* relation $\prec$ is defined by $[x \prec y \Leftrightarrow (x \P y) \wedge \neg(y \P x)]$, and its *equivalence* relation $\approx$ is given by $[x \approx y \Leftrightarrow (x \P y) \wedge (y \P x)]$. Let $(X, \P)$ be a totally preordered set and let $X/\approx$ be the set of equivalence classes. If $x \in X$ we denote the equivalence class of x by $[x]$. The total preorder $\P$ on X induces a total order on $X/\approx$ defined by $[x] \lhd [y] \Leftrightarrow x \P y$. Given $A \subseteq X$, the $\approx$-*saturated* set of A is defined as $sat(A) = \bigcup_{x \in A} [x]$. A subset $A \subseteq X$ is said to be $\approx$-*saturated* if $A = sat(A)$.

Let $(X, \P)$ be a totally preordered set. A subset Z of X is said to be *order-dense in* X *with respect to* $\P$, if $x, y \in X$ and $x \prec y$ imply that there exists $z \in Z$ such that $x \P z \P y$. $(X, \P)$ is said to be *order-separable* if it has a countable order-dense subset.

If $(X, \P)$ is a preordered set then a real-valued function $u : X \to R$ is said to be an *order-monomorphism* if or every $x, y \in X$, $[x \P y \Rightarrow u(x) \leq u(y)]$ and $[x \prec y \Rightarrow u(x) < u(y)]$. If such a function u does exist then $\P$ is said to be *representable*. This function is also referred to as a *utility function* for $\P$.

If a nonempty set X is endowed with a topology τ then a *total* preorder $\P$ on X is said to be *continuously representable* if there exists a utility function that is continuous with respect to the topology τ and the usual topology on the real line R.

Let $(X, \P)$ be a totally preordered set. The family of all sets of the form $L(x) = \{a \in X; a \prec x\}$ and $G(x) = \{a \in X; x \prec a\}$, where $x \in X$, is a subbasis for a topology $\tau_{\P}$ on X, called the *order topology.*

If $(X, \P)$ is a preordered set and τ is a topology on X, then the preorder $\P$ is said to be τ*-continuous* on X if for each $x \in X$ the sets $L(x)$ and $G(x)$ are τ-open in X. A topology τ on X is said to be *(total) preorderable* if $\tau \equiv \tau_{\P}$, for some total preorder $\P$ defined on X.

3 THE CONTINUOUS REPRESENTABILITY PROPERTY

The two most important results in *topological utility theory* are the Eilenberg theorem ([10]) and the Debreu theorem ([9]), respectively. Eilenberg's theorem states that any continuous total preorder defined on a connected and separable topological space is continuously representable. Debreu's theorem gives the same conclusion as Eilenberg's theorem provided that connected and separability are replaced by the single condition of second countability.

A way of viewing these two important statements is as follows. Both results provide sufficient conditions for a topology to satisfy the ordinal property that every continuous total preorder is continuously representable. This point of view was stated by Herden ([14]) who defined the following concept.

Definition 1. Let (X, τ) be a topological space. Then the topology τ satisfies the *continuous representability property* (CRP) if every continuous total preorder $\P$ defined on X has a continuous utility function.

Remark 2. (i) A CRP topology was formerly called by Herden a *useful topology.*

(ii) CRP topologies have been studied in [4,5,6,14,15].

(iii) According to the definition of CRP it turns out that both Eilenberg's theorem and Debreu's theorem can be re-stated by saying that a connected and separable topology and a second countable topology, respectively, have CRP.

The next theorem provides information about familiar topologies that satisfy CRP.

Theorem 3. *Let* (X,τ) *be a topological space. Then CRP holds in the following cases:*

(i) If τ *is connected and separable.*

(ii) If τ *is second countable.*

(iii) If τ *is path-connected and* σ*-compact.*

(iv) If τ *is locally connected and separable.*

Proof. See [6,9,10,17].

The fact that τ is second countable is also a necessary condition for CRP to have in two important cases: namely, the case of metric spaces and the case of totally preordered spaces.

Theorem 4. *Let* (X,τ) *be a metric space. Then* τ *satisfies CRP iff it is second countable (*$\equiv$*separable).*

Proof. See [5,13].

For totally preordered spaces we reach the following characterization.

Theorem 5. *Let* $(X,\preccurlyeq)$ *be a totally preordered space. Then* $\tau_{\preccurlyeq}$ *satisfies CRP iff it is second countable.*

Proof. See [4].

The previous result has an interesting consequence: namely, the representability, which is an ordinal property, can be characterized in purely topological terms.

Corollary 6. *Let* $(X,\preccurlyeq)$ *be a totally preordered space. Then* $\preccurlyeq$ *is representable iff* $\tau_{\preccurlyeq}$ *is second countable.*

Remark 7. The representability property for a total order was studied early by Cantor [7,8]. An ordinal characterization of this property was obtained, among others, by Milgram [16] and Birkhoff [1] in terms of the notion of order-separability.

We conclude this section with a result that takes advantage of the previous corollary to obtain a intrinsic characterization of CRP.

Theorem 8. *Let* (X,τ) *be a topological space. The topology* τ *satisfies CRP if and only if all its preorderable subtopologies are second countable.*

Proof. Let us first observe the following simple fact. If τ' is a preorderable subtopology of τ, $\tau' = \tau_{\preccurlyeq}$, for some total preorder $\preccurlyeq$ defined on X. Obviously $\preccurlyeq$ is τ-continuous. Conversely, if $\preccurlyeq$ is a τ-continuous total preorder on X, then its order topology $\tau_{\preccurlyeq}$ is a preorderable subtopology of τ. Thus we see that there is a bijection between τ-continuous total preorders defined on X and preorderable subtopologies of τ.

Suppose now that the topology τ has CRP. In particular, every continuous total preorder $\preccurlyeq$ is representable through an order-monomorphism, so that the topology $\tau_{\preccurlyeq}$ is second countable by the previous Corollary. Because of the aforementioned bijection, this implies that all the preorderable subtopologies of τ are second countable.

Conversely, suppose that $\preccurlyeq$ is a τ-continuous total preorder defined on X. Thus, $\tau_{\preccurlyeq} \subseteq \tau$. By hypothesis, $\tau_{\preccurlyeq}$ is second countable. Now, by Corollary 6 again, $\preccurlyeq$ is representable by a utility function $f : X \to R$ that is continuous if we consider on X the *order* topology $\tau_{\preccurlyeq}$ and on R the usual topology. Since $\tau_{\preccurlyeq} \subseteq \tau$, f is also τ-continuous.

Q.E.D.

4 THE CONTINUOUS ALGEBRAIC REPRESENTABILITY PROPERTY

In this section we explore the algebraic version of CRP. In the topological set-up CRP appears whenever we look for ordinal representations of total preorders that, in addition, preserve a nice topological property: namely, the continuity. In the algebraic context, in addition to the order and continuity properties for a utility function, we will ask for a new demanding requirement; that of being an algebraic-homomorphism. Of course, this imposes some kind of compatibility among order, topology, and algebra involved. Although we could begin with a very simple algebraic ordered sructure (e.g., think of a totally preordered semigroup), we will focus on richer algebraic systems. In particular, we will pay attention to *totally preordered real algebras.* The reason for this choice will be made clear later on, in the section of applications, when all of this algebraic machinery will be applied to characterize *strongly dictatorial social welfare functionals.* We now introduce some basic definitions.

A *real algebra* $(X,+,\cdot_{\mathbb{R}},*)$ is a set X endowed with three binary operations such that:

(i) $(X,+,\cdot_{\mathbb{R}})$ is a real vector space.

(ii) $(X,+,*)$ is a ring.

(iii) $\lambda\cdot(x*y)=(\lambda\cdot x)*y=x*(\lambda\cdot y)$, $\forall x,y\in X$, $\forall\lambda\in\mathbb{R}$.

Let $(X,+,\cdot_{\mathbb{R}},*)$ be a real algebra. A function $v:X\to\mathbb{R}$ is said to be an *algebra-homomorphism* if it is linear and multiplicative, i.e.:

(a) $v(\lambda\cdot x+\beta\cdot y)=\lambda v(x)+\beta v(y)$, $\forall x,y\in X$, $\forall\lambda,\beta\in\mathbb{R}$.

(b) $v(x*y)=v(x)v(y)$, $\forall x,y\in X$.

In the next definition some usual compatibility conditions between order and algebra are introduced. This leads to the concept of a *totally preordered real algebra*.

Definition 10. A *totally preordered real algebra* $(X,\preccurlyeq,+,\cdot_{\mathbb{R}},*)$ is a real algebra equipped with a total preorder $\preccurlyeq$ which is compatible with the operations $+,\cdot_{\mathbb{R}}$ and $*$, i.e.:

(1) $x\preccurlyeq y$ implies $x+z\preccurlyeq y+z$, $\forall z\in X$ (*translation-invariance*).

(2) $x\preccurlyeq y$, $0\le\lambda$ implies $\lambda\cdot x\preccurlyeq\lambda\cdot y$ (*homotheticity*).

(3) $x\preccurlyeq y$, $\mathbf{0}\preccurlyeq z$ imply $z*x\preccurlyeq z*y$ and $x*z\preccurlyeq y*z$ (*multiplicativeness*. **0** denotes the null element with respect to +).

Next we define the notion of a *topological real algebra.*

Definition 11. A real algebra $(X,+,\cdot_{\mathbb{R}},*)$ equipped with a topology τ is said to be a *topological real algebra* if $(X,\tau,+,\cdot_{\mathbb{R}})$ is a topological vector space and $*$ is a τ-continuous binary operation on $X\times X$.

We now introduce the concept of CRP in the algebraic enviroment.

Definition 12. Let $(X,+,\cdot_{\mathbb{R}},*)$ be a real algebra and τ a topology on X. Then τ satisfies the *continuous algebraic representability property* (shortly, CARP) if every τ-continuous total preorder on X, for which $(X,\preccurlyeq,+,\cdot_{\mathbb{R}},*)$ becomes a totally preordered real algebra, admits a continuous utility function which is an algebra-homomorphism (shortly, a continuous algebraic utility function).

Before proving the main result in this setting a preliminary lemma is needed.

Lemma 13. *Suppose that* $\mathbb{R}$ *is endowed with both a binary operation* $*$ *and a non-trivial continuous total order* $\preccurlyeq$*, in such a way that* $\preccurlyeq$ *satisfies translation-invariance (with respect to the usual operation* $+$*) and multiplicative invariance (with respect to* $*$*). Then* $\preccurlyeq$ *is representable by a continuous utility function which, in addition, is additive and multiplicative. Moreover, there exists* $a\in\mathbb{R}\setminus 0$ *such that* $x*y=axy$*, for all* $x,y\in\mathbb{R}$.

Proof. Since $\preccurlyeq$ is continuous on $\mathbb{R}$, it is very easy to see that there exists an (injective) continuous utility function $u:\mathbb{R}\to\mathbb{R}$ for $\preccurlyeq$. Actually, u is a homeomorphism. Thus, without loss of generality, we can assume that either $u(x)=x$ or $u(x)=-x$ $(x\in\mathbb{R})$. Suppose that

$u(x) = x$; the other case being analyzed in a similar way.

Now, by a result by Pickert and Hion (see Birkhoff [1,p.398-399], $\P$ is representable by a utility function $v : \mathbb{R} \to \mathbb{R}$ which is additive and multiplicative (with respect to $*$). Since u and v are utility functions for $\P$, v is also increasing. Remember that v is an additive function and it is well known that if v is discontinuous at some point, then it is discontinuous at every point of $\mathbb{R}$. This last fact, v being increasing, leads to $v(x) = ax$ (with $a > 0$). Since $a(x * y) = v(x * y) = v(x)v(y) = a^2 xy$, this type of representation gives $x * y = axy$ for every $x, y \in \mathbb{R}$.

Q.E.D.

Now, we reach the following characterization of CARP.

Theorem 14. *Let $(X, +, \cdot_{\mathbb{R}}, *)$ be a real algebra and τ a topology on X. The the following assertions are equivalent:*

(i) τ has CARP.

(ii) $(X, \tau_{\P}, +, \cdot_{\mathbb{R}}, *)$ is a topological real algebra, for every preorderable subtopology $\tau_{\P} \subseteq \tau$ that makes $(X, \P, +, \cdot_{\mathbb{R}}, *)$ to be a totally preordered real algebra.

Proof. $(i) \Rightarrow (ii)$. Let $\P$ be a τ-continuous total preorder on X which satisfies the invariance properties (1) to (3) of Definition 10 above. Let us see that, endowed with the order topology $\tau_{\P}$, $(X, \tau_{\P}, +, \cdot_{\mathbb{R}}, *)$ is a topological real algebra. Since, by hypothesis, τ has CARP there is a τ-continuous algebraic utility function, say $\psi : X \to \mathbb{R}$, for $\P$. It is then simple to see that $X/\approx$, endowed with the total order induced by $\P$, is order-isomorphic to $\mathbb{R}$. Moreover, because ψ is linear and multiplicative, $(X/\approx, +, \cdot_{\mathbb{R}}, *)$ is algebraically-isomorphic to the reals with the usual operations. But this means that $(X, \tau_{\P}, +, \cdot_{\mathbb{R}}, *)$ is a topological real algebra.

For the converse, $(ii) \Rightarrow (i)$, we argue as follows. Let $\P$ be a τ-continuous total preorder on X which satisfies the invariance properties (1) to (3) of Definition 10 above. Let us prove that there is a τ-continuous algebraic utility function which represents $\P$. Consider the order topology $\tau_{\P}$ on X, so that $(X, \tau_{\P}, +, \cdot_{\mathbb{R}}, *)$ is a topological real algebra, by hypothesis. Define the set $I(0) = \{x \in X; x \approx 0\}$. First, let us see that $I(0)$ is both a vector subspace and an ideal of X, for which we need to prove the following two properties:

(a) $I(0)$ is a real vector subspace of X.

(b) For every $x \in I(0)$, $y \in X$, it holds that $x * y \in I(0)$.

Let $x, y \in I(0)$. Since $\P$ is translation-invariant, it follows that $x + y \approx x + 0 = x \approx 0$. So, in order to prove (a), it is sufficient to see that, given $x \in I(0)$ and $\lambda \in \mathbb{R}$, then $\lambda \cdot x \approx 0$. If $\lambda \geq 0$ then, by homotheticity, $\lambda \cdot x \approx 0$. If $\lambda < 0$, then $(-\lambda) \cdot x \approx 0$. But $(-\lambda) \cdot x = -\lambda \cdot x$ and, by translation invariance of $\P$, $\lambda \cdot x \approx 0$.

To prove (b), let $y \in X$ and $x \in I(0)$. If $0 \P y$ then, because $\P$ is multiplicative, $x * y \P 0 * y = 0$ and $0 = 0 * y \P x * y$. Therefore, $x * y \approx 0$. If $y \P 0$, then $0 \P -y$ and so $-x * y \approx 0$. Since $x * y = -(-x) * y$, and $I(0)$ is a vector subspace of X, it holds that $x * y \approx 0$. Thus (b) holds and therefore $I(0)$ is an ideal of X.

Now, consider the quotient space $X/I(0)$. Since $X/I(0)$ coincides with the quotient space $X/\approx$, $X/I(0)$ is a totally ordered set. We denote the natural order on $X/I(0)$ by $\lhd$. Since $I(0)$ is both a vector subspace and an ideal of X, the operations $+, \cdot_{\mathbb{R}}$ and $*$ pass to the quotient space in such a way that $\lhd$, defined on $X/I(0)$, satisfies the invariance properties of

Definition 10 too. Moreover, $I(0)$ is a $\tau_{\P}$-closed subset of X because $\P$ is $\tau_{\P}$-continuous. By using a standard connectedness argument it follows that $X/I(0)$ is homeomorphic and isomorphic as a vector space to $\mathbb{R}$. Now, by Lemma 13, there is a $\tau_{\P}$-continuous algebraic utility function $u: X/I(0) \to \mathbb{R}$ that represents $\lhd$. Let us denote by $p: X \to X/I(0)$ the projection map. It is easy to see that p is linear, multiplicative and continuous. Then, by considering the composition $\psi = u \circ p: X \to \mathbb{R}$ we obtain a $\tau_{\P}$-continuous (hence τ-continuous), linear and multiplicative utility function that represents $\P$. Thus τ has CARP.

Q.E.D.

If τ makes $(X,+,\cdot_{\mathbb{R}},*)$ to be a topological real algebra then we have the following interesting consequence.

Corollary 15. *Let $(X,+,\cdot_{\mathbb{R}},*)$ be a real algebra. Then any topology τ on X that makes $(X,\tau,+,\cdot_{\mathbb{R}},*)$ to be a topological real algebra satisfies CARP.*

Proof. According to the statement of Theorem 14 we only have to prove that, given a preorderable subtopology $\tau_{\P} \subseteq \tau$ for which $(X,\P,+,\cdot_{\mathbb{R}},*)$ is a totally preordered real algebra, it holds that $(X,\tau_{\P},+,\cdot_{\mathbb{R}},*)$ is a topological real algebra. Let us show that $+: X \times X \to X$ is a $\tau_{\P}$-continuous operation; the proof for the remaining operations ("$\cdot_{\mathbb{R}}$" and "$*$") being similar. Let then $C \subseteq X$ be a $\tau_{\P}$-open subset. Since $\tau_{\P} \subseteq \tau$ and "$+$" is τ-continuous, there are τ-open subsets $A,B \subseteq X$ such that $A+B = \{a+b; a \in A, b \in B\} \subseteq C$. Now, because $\P$ is translation-invariant, it easily follows that $sat(A)+sat(B) \subseteq sat(C) = C$. But $sat(A)$ and $sat(B)$ are both $\tau_{\P}$-open subsets of X. Thus, "$+$" is a $\tau_{\P}$-continuous map.

Q.E.D.

Remark 16. In particular, Corollary 15 applies to the n-dimensional Euclidean space $\mathbb{R}^n$ endowed with the usual binary operations defined componentwise. Moreover, in this case it can be easily seen that the continuous algebraic utility functions are of the form $\psi(x_1,\ldots,x_j,\ldots,x_n) = x_j$, for some $j \in \{1,\ldots,n\}$. In other words, any continuous total preorder defined on $\mathbb{R}^n$ which is translation-invariant, homothetic and multiplicative is *projective*.

5 APPLICATION TO SOCIAL CHOICE THEORY

In this section we develop an application of the algebraic approach presented in the previous section to the context of utility theory in the social choice framework. A social welfare functional is a map that assigns a preference relation (total preorder) to any profile of individual utilities. We characterize strongly dictatorial social welfare functionals in terms of the invariance and continuity properties of such rules, in the spirit of the literature of utility measurability and personal comparability (see, e.g.,[2]). Before presenting the application some definitions and notations are needed.

Let X be a nonempty set (alternatives). Let us denote by $\mathfrak{R}$ the class of all total preorders defined on X. Let $n > 1$ be a natural number (number of agents). A (utility) function for the agent i will be denoted by u_i. A *profile* of (utility) functions will be denoted by $(u_1,\ldots,u_n)$ and U^n will stand for the set of all possible profiles. The set which consists of the n first natural numbers will be denoted by N.

A *social welfare functional* is a rule $F: U^n \to \mathfrak{R}$ that assigns a preference relation $F(u_1,\ldots,u_n) \in \mathfrak{R}$, interpreted as the social preference relation, to any profile $(u_1,\ldots,u_n)$ in the

domain U^n. In order to shorten the notation, for a profile $(u_1,\ldots,u_n)$ we will use the notation (u_i). Also, $F(u_i)_s$ stands for the strict preference (asymmetric relation) associated with $F(u_i)$.

A social welfare functional $F:U^n\to\Re$ is *(weak) Paretian* if, for any pair of alternatives $x,y\in X$ and any profile $(u_i)\in U^n$, we have that $u_i(x)\le u_i(y)$ for all $i\in N$ implies $xF(u_i)y$ and also that $u_i(x)<u_i(y)$ for all i implies $xF(u_i)_s y$.

A social welfare functional $F:U^n\to\Re$ satisfies the *binary independence of irrelevant alternatives* condition if, for any pair of alternatives $x,y\in X$ and any pair of profiles $(u_i),(u_i^{'})\in U^n$ with the property that $u_i(x)=u_i^{'}(x)$ and $u_i(y)=u_i^{'}(y)$ for all $i\in N$, we have that $xF(u_i)y$ if and only if $xF(u_i^{'})y$.

A social welfare functional $F:U^n\to\Re$ is called *strongly dictatorial* if there is an agent $j\in N$ such that $F(u_i)$ coincides with $\succcurlyeq_j$, for every $(u_i)\in U^n$, where $\succcurlyeq_j$ denotes the preference relation on X induced by the utility function of the j-agent, u_j.

A social welfare functional $F:U^n\to\Re$ is called *anonymous* if for any pair of profiles $(u_i),(v_i)\in U^n$ such that (v_i) is derived from (u_i) by permuting the individuals' utility functions, one has that $F(u_i)$ and $F(v_i)$ coincide.

A social welfare functional $F:U^n\to\Re$ is called *continuous* if $\{((a_i),(b_i))\in\mathbb{R}^n\times\mathbb{R}^n;\exists(u_i)\in U^n, x,y\in X$ with $u_i(x)=a_i, u_i(y)=b_i$ and $xF(u_i)y\}$ is an Euclidean-closed subset of $\mathbb{R}^n\times\mathbb{R}^n$.

Finally, a social welfare functional $F:U^n\to\Re$ satisfies *information invariance with respect to cardinal measurability* if, for any profile $(u_i)\in U^n$, and any tuple of functions $(\phi_1,\ldots,\phi_n)$ such that, for each $i\in N$, the map $\phi_i:\mathbb{R}\to\mathbb{R}$ is of the form $\phi_i(t)=a_i+b_i t$ $(t\in\mathbb{R})$ with $a_i\in\mathbb{R}$ and $b_i>0$, it holds that $F(u_i)$ and $F(\phi_i\circ u_i)$ coincide.

We reach the following result.

Theorem 17. *Suppose that X contains at least three elements and let $F:U^n\to\Re$ be a social welfare functional. Then the following assertions are equivalent:*

(i) F is (weak) Paretian, satisfies binary independence of irrelevant alternatives, continuity, and information invariance with respect to cardinal measurability.

(ii) F is strongly dictatorial.

Proof. It is straightforward to show that (ii) implies (i). To prove the converse, we employ the welfarist approach (see [2] for a survey of welfarist social choice). Consider the binary relation $\succcurlyeq^*$ defined on $\mathbb{R}^n$ as follows: Given $a=(a_1,\ldots,a_n)$ and $b=(b_1,\ldots,b_n)\in\mathbb{R}^n$, then $a\succcurlyeq^* b$ if and only if there exist $x,y\in X$ and a profile $(u_i)\in U^n$ such that for every $i\in\{1,\ldots,n\}$ it holds that $u_i(x)=a_i$, $u_i(y)=b_i$ and $xF(u_i)y$. Since F satisfies Pareto and binary independence of irrelevant alternatives, and X contains at least three elements, it can be shown that $\succcurlyeq^*$ is a well-defined total preorder that "generates" F. It should be noted that $\succcurlyeq^*$ is non-trivial because F is Paretian. These are the contents of the so-called *welfarism theorem* (see, e.g., [2,p.1107]).

Notice that, because F is Paretian, $\succcurlyeq^*$ is increasing[1] and, moreover, for every $a,b\in\mathbb{R}^n$

[1] A total preorder $\succcurlyeq$ defined on $\mathbb{R}^n$ is said to be *increasing* if, for every $x=(x_1,\ldots,x_n)$, $y=(y_1,\ldots,y_n)\in\mathbb{R}^n$ such that $x_i\le y_i$ for all $i\in N$, it holds that $x\succcurlyeq y$. "$\prec$" will stand for the

such that $a \prec b$ it holds that $a \prec^* b$. Let us show that $\P^*$ satisfies all the conditions stated in Remark 16.

Continuity of $\P^*$, referred to the Euclidean topology in $\mathbb{R}^n$, follows directly from the continuity of F. Let us prove now that $\P^*$ is translation-invariant. To that end, let $a=(a_i), b=(b_i), c=(c_i) \in \mathbb{R}^n$ such that $a \P^* b$. Then there are $x, y \in X$ and a profile of utility functions $(u_i) \in U^n$ such that $u_i(x)=a_i$, $u_i(y)=b_i$, for all i, and $xF(u_i)y$. For every i, consider the functions $u_i'(z)=u_i(z)+c_i$ for all $z \in X$. By information invariance with respect to cardinal measurability, it holds that $F(u_i')$ coincides with $F(u_i)$. In particular, $xF(u_i')y$ or, equivalently, $(u_i'(x)) \P^* (u_i'(y))$. Thus, $a+c=(a_i+c_i) \P^* b+c=(b_i+c_i)$.

Similarly, let us prove that $\P^*$ is homothetic. To see this, let $a=(a_i), b=(b_i), c=(c_i) \in \mathbb{R}^n$, with $0 \le \lambda$, such that $a \P^* b$. If $\lambda = 0$, obviously $\lambda \cdot a \P^* \lambda \cdot b$. So, suppose that $0 < \lambda$. Then, there are $x, y \in X$ and a profile of utility functions $(u_i) \in U^n$, such that $u_i(x)=a_i$, $u_i(y)=b_i$, for all i, and $xF(u_i)y$. For each i, consider the (utility) function $u_i' = \lambda u_i$. By information invariance with respect to cardinal measurability, it holds that $F(u_i')$ and $F(u_i)$ coincide. In particular, $xF(u_i')y$ or, equivalently, $(u_i'(x)) \P^* (u_i'(y))$. Thus, $\lambda \cdot a = (\lambda a_i) \P^* \lambda \cdot b = (\lambda b_i)$. Notice that, by translation invariance and homotheticity of $\P^*$ it follows, as in the proof of Theorem 14, that $I(0_n)=\{z \in \mathbb{R}^n ; z \approx^* 0_n\}$ is a vector subspace of $\mathbb{R}^n$.

The fact that $\P^*$ is multiplicative is a little more tricky. Since $\P^*$ is translation-invariant it should be noted that, in order to prove that $\P^*$ is multiplicative, it is sufficient to show that for any pair $a, b \in \mathbb{R}^n$ such that $0_n \P^* a, b$ it holds that $0_n \P^* a * b$. Now, observe that, in a similar way to what we have done for proving the homotheticity property, it can be seen that $0_n \P^* a=(a_i)$, implies $0_n \P^* a * c$, $\forall\, 0_n = c = (c_i)$. In particular, $0_n \approx^* a$ implies $0_n \approx^* a * c$, $\forall\, 0_n \prec c$. Indeed, because $0_n \P^* a$, there are $x, y \in X$ and a profile of utility functions $(u_i) \in U^n$, such that $u_i(x)=0$, $u_i(y)=a_i$, for all i, and $xF(u_i)y$. For each i, consider the (utility) function $u_i' = c_i u_i$. By information invariance with respect to cardinal measurability, it holds that $F(u_i')$ and $F(u_i)$ coincide. In particular, $xF(u_i')y$ or, equivalently, $(u_i'(x)) \P^* (u_i'(y))$. Thus, $0_n) \P^* a * c = (a_i c_i)$.

Let then $a=(a_i) \in \mathbb{R}^n$ such that $0_n \P^* a$ and denote by $(e^i)_{i \in N}$ the canonical basis in $\mathbb{R}^n$. Let us see that $0_n \P^* a_i e^i$ and, in addition, if $a_i < 0$ then $0_n \approx^* a_i e^i$. Indeed, for every $m \in \mathbb{N}$, consider the vector $c^m=(c_k^m) \in \mathbb{R}^n$ defined as follows: $c_k^m = 1/m$, if $k \neq i$ and $c_i^m = 1$. Clearly, for every $m \in \mathbb{N}$, $0_n \prec c^m$ and the sequence c^m goes to e^i, as m goes to ∞. Since $\P^*$ is continuous, the set $G(0_n)=\{d \in \mathbb{R}^n ; 0_n \P^* d\}$ is a closed subset of $\mathbb{R}^n$. Thus, as m goes to ∞, $a * c^m \in G(0_n)$ converges to $a_i e^i \in G(0_n)$. Moreover if $a_i < 0$ then, because $\P^*$ is increasing and total, it follows that $0_n \approx^* a_i e^i$.

---

strict partial order in $\mathbb{R}^n$ (i.e., $x \prec y$ iff $x_i < y_i$, $\forall i \in N$). The zero vector in $\mathbb{R}^n$ will be denoted by 0_n.

Let then $a=(a_i), b=(b_i)\in \mathbb{R}^n$ such that $0_n \precsim^* a,b$. Notice that $a*b=\sum a_i b_i e^i$. Now, each of the vectors $a_i b_i e^i$ belongs to $G(0_n)$. Indeed, if $0\le a_i, b_i$ then $a_i b_i e^i \in G(0_n)$ because $\precsim^*$ is increasing and total. If either $a_i<0$ or $b_i<0$ then, by the argument above, either $0_n \approx^* a_i e^i$ or $0_n \approx^* b_i e^i$. But, because $I(0_n)$ is a vector subspace, it follows that $a_i b_i e^i \in I(0_n)\subset G(0_n)$. By translation invariance of $\precsim^*$ again it follows that $a*b\in G(0_n)$. Thus, $\precsim^*$ is multiplicative. Hence all the conditions of Remark 16 are satisfied and therefore $\precsim^*$ is a projective total preorder on $\mathbb{R}^n$. This clearly implies that F is a strongly dictatorial social welfare functional.

Q.E.D.

As a direct consequence of the previous theorem, we obtain the following impossibility result.

Corollary 18. *Suppose that X contains at least three elements. Then there is no continuous social welfare functional $F:U^n\to\Re$ that is Paretian, satisfies binary independence of irrelevant alternatives, information invariance with respect to cardinal measurability, and anonymity.*

REFERENCES

[1] BIRKHOFF G, LATTICE THEORY, American M. Society, Providence RI, 1948.

[2] BOSSERT W, WEYMARK JA (2004) Utility in social choice, Chapter **20** of the Handbook of Utility Theory (Vol.**2**), Edited by Barberà S, Hammond PJ and Seidl Ch, Kluwer Academic Publishers, Amsterdam, 2004.

[3] BRIDGES DS, MEHTA GB, Representation of preference orderings, Springer, Berlin, 1995.

[4] CAMPIÓN MJ, CANDEAL JC, INDURÁIN E, MEHTA GB (2009) Continuous order-representability of topological spaces. Preprint.

[5] CANDEAL JC, HERVÉS C, INDURÁIN E (1998) Some results on representation and extension of preferences. Journal of Mathematical Economics 29, 75-81.

[6] CANDEAL JC, INDURÁIN E, MEHTA GB (2004) Utility functions on locally connected spaces. Journal of Mathematical Economics 40, 701-711.

[7] CANTOR G (1985) Beiträge zur Begründung der transfinite Mengenlehre (I), Mathematische Annalen 46, 481-512.

[8] CANTOR G (1897) Beiträge zur Begründung der transfinite Mengenlehre (II), Mathematische Annalen 49, 207-246.

[9] DEBREU G (1964) Continuity properties of Paretian utility. International Economic Review 5, 285-293.

[10] DUGUNDJI J, Topology, Allyn and Bacon, Boston, 1966.

[11] EILENBERG S (1941) Ordered topological spaces. American Journal of Mathematics 63, 39-45.

[12] ENGELKING R, General Topology, Sigma Series in Pure Mathematics **6**, Heldermann Verlag Berlin, 1989.

[13] ESTÉVEZ M, HERVÉS C (1995) On the existence of continuous preference orderings without utility representations. Journal of Mathematical Economics 24, 305-309.

[14] HERDEN G (1991) Topological spaces for which every continuous total preorder can be represented by a continuous utility function. Mathematical Social Sciences 22, 123-136.

[15] HERDEN G, PALLACK A (2000) Useful topologies and separable systems. Applied General Topology 1 (1), 61-82.

[16] MILGRAM A (1940) Partially ordered sets and topology. Reports of a Mathematical Colloquium (Second Series). University of Notre Dame, 3-9.

[17] MONTEIRO PK (1987) Some results on the existence of utility functions on path connected spaces. Journal of Mathematical Economics 16, 147-156.

ESTRUCTURA DE LOS SISTEMAS DE VOTACIÓN NEUTRALES, MONÓTONOS Y QUE RESPETAN EL ANONIMATO

BONIFACIO LLAMAZARES RODRÍGUEZ
E-mail: boni@eco.uva.es
Departamento de Economía Aplicada
Universidad de Valladolid

Abstract

Social welfare functions between two alternatives have been widely studied in the literature of Social Choice. One of the results given by Fishburn [*The Theory of Social Choice*, Princeton University Press, 1973] allows us to characterize anonymous, neutral and monotonic social welfare functions by means of functions satisfying adequate conditions. From among all kinds of functions, the class of affine functions is highly interesting because from these functions it is possible to obtain the social welfare functions most used in practice. In this paper we analyze the structure of the set of these functions and we show that this set is convex and its extreme points are the functions that generate the following social welfare functions: simple majority, absolute majority, unanimous majority and Pareto majority.

Key words: Social choice functions; convex sets.
Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUCCIÓN

Un problema clásico que aparece en diversos ámbitos científicos consiste en determinar una preferencia colectiva a partir de las preferencias de un grupo de individuos sobre un conjunto de opciones. La situación más simple se produce cuando los individuos tienen que elegir únicamente entre dos alternativas. En este caso no existen problemas de transitividad en la preferencia colectiva, y ciertas cuestiones espinosas pueden ser evitadas (por ejemplo, la paradoja de la votación o el Teorema de Imposibilidad de Arrow). Sin embargo, a pesar de ser el caso más sencillo, no es trivial, y el análisis de diversos sistemas de votación entre dos alternativas (que en este trabajo denominaremos funciones de elección social) ha generado una abundante literatura sobre el tema.

Uno de los trabajos más relevantes y exhaustivos sobre las funciones de elección social se debe a Fishburn (1973). Uno de sus resultados, Fishburn (1973, p. 56), puede ser formulado como una caracterización de las funciones de elección social que satisfacen las propiedades de neutralidad, monotonía y respeto al anonimato. A partir de este resultado se pueden caracterizar las funciones de elección social que satisfacen las propiedades anteriores a través de unas determinadas funciones monótonas.

Dado que las funciones de elección social más utilizadas en la práctica se pueden generar cuando las funciones monótonas consideradas son afines, este trabajo tiene por objeto analizar la estructura del conjunto formado por dichas funciones. En este sentido se muestra que dicho conjunto es convexo y que sus vértices son las funciones que generan las siguientes mayorías: la mayoría simple, la mayoría absoluta, la mayoría unánime y la mayoría de Pareto.

El trabajo está organizado en dos secciones. En la primera de ellas, Sección 2, se introducen los conceptos y la notación necesaria relacionada con las funciones de elección social, así como

el resultado de caracterización dado por Fishburn. En la segunda, Sección 3, se establecen los resultados del trabajo.

2. PRELIMINARES

Sean $\boldsymbol{N}=\{1,\ldots,n\}$ un conjunto de votantes, con $n\geq 3$, y x, y dos alternativas. Las preferencias individuales entre x e y se describen a través de un perfil $D=(d_1,\ldots,d_n)$, donde d_i toma los valores 1, -1 y 0 según que el individuo i prefiera x a y, y a x o sea indiferente entre ambas opciones. El conjunto de perfiles de preferencia será denotado por $\boldsymbol{D}$. Dados $D,D'\in \boldsymbol{D}$ y σ una permutación de $\boldsymbol{N}$, a lo largo del trabajo se utilizará la siguiente notación: $D_\sigma=(d_{\sigma(1)},\ldots,d_{\sigma(n)})$, $-D=(-d_1,\ldots,-d_n)$, $n^+(D)=\#\{i\in \boldsymbol{N}\mid d_i=1\}$, $n^-(D)=\#\{i\in \boldsymbol{N}\mid d_i=-1\}$ y $D'\geq D$ significará $d'_i\geq d_i$ para todo $i\in \boldsymbol{N}$.

Para cada perfil de preferencias, la preferencia colectiva será obtenida a través de una función de elección social.

Definición 1. Una *función de elección social (FES)* es una aplicación $F:\boldsymbol{D}\longrightarrow\{-1,0,1\}$.

Los posibles valores de F, 1, 0 y -1, tienen una interpretación similar a la dada en el caso de las preferencias individuales. Así, F toma los valores 1, -1 y 0 según que x sea preferida colectivamente a y, y lo sea a x o ambas opciones sean indiferentes para el colectivo.

A continuación se muestran algunas propiedades de las FES que son sobradamente conocidas en la literatura: neutralidad, monotonía y respeto al anonimato. La neutralidad expresa un tratamiento igualitario entre las alternativas: si todos los individuos invierten sus preferencias entre x e y entonces la preferencia colectiva también resulta invertida. La monotonía significa que cuando aumenta el respaldo de algunos individuos a una alternativa, entonces la preferencia colectiva no puede empeorar para esa alternativa. El respeto al anonimato refleja un tratamiento igualitario entre los individuos: la preferencia colectiva sólo depende del conjunto de preferencias individuales, y no de qué individuos tienen esas preferencias.

Definición 2. Sea F una FES.

1. F es *neutral* si para cualquier perfil de preferencias $D\in \boldsymbol{D}$ se verifica $F(-D)=-F(D)$.
2. F es *monótona* si para cualesquiera $D,D'\in \boldsymbol{D}$ tales que $D'\geq D$ se verifica $F(D')\geq F(D)$.
3. F *respeta el anonimato* si para cualquier permutación σ de $\boldsymbol{N}$ y cualquier perfil de preferencias $D\in \boldsymbol{D}$ se verifica $F(D_\sigma)=F(D)$.

Cabe destacar que si F es una FES que respeta el anonimato, entonces el valor $F(D)$ queda determinado por $n^+(D)$ y $n^-(D)$. Asimismo, cabe resaltar también que cualquier FES neutral F queda caracterizada por el conjunto $F^{-1}(1)$, dado que

$$F^{-1}(-1)=\{D\in \boldsymbol{D}\mid -D\in F^{-1}(1)\},\ F^{-1}(0)=\boldsymbol{D}-\left(F^{-1}(1)\cup F^{-1}(-1)\right).$$

Las FES neutrales, monótonas y que respetan el anonimato han sido estudiadas en profundidad por Fishburn (1973). Uno de sus resultados, Fishburn (1973, p. 56) puede ser dado como una caracterización de estas funciones. Dado $a\in\Re$, $[a]$ denotará la parte entera de a, es decir, el mayor entero menor o igual que a.

Teorema 1. Sea *F* una FES. Entonces las siguientes condiciones son equivalentes:

1. *F* es neutral, monótona y respeta el anonimato.
2. Existe una secuencia de números enteros $r_0 \le r_1 \le \cdots \le r_{[n/2]}$ que verifican $r_i \ge i$ para todo $i \in \{0,1,\ldots,[n/2]\}$ y tal que

$$F(D) = \begin{cases} 1, & \text{si } n^+(D) > r_{n^-(D)}, \\ -1, & \text{si } n^-(D) > r_{n^+(D)}, \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

3. LOS RESULTADOS

El Teorema 1 de la sección anterior permite caracterizar las FES neutrales, monótonas y que respetan el anonimato a través de funciones monótonas que satisfacen una determinada propiedad. Para ello basta tener en cuenta que los números enteros $r_0, r_1, \ldots, r_{[n/2]}$ pueden ser vistos como la parte entera de los valores que toma una función. Además, dado que cualquier FES neutral *F* queda determinada a través del conjunto $F^{-1}(1)$, el resultado de caracterización anterior puede ser dado de la siguiente forma.

Corolario 1. Sea *F* una FES neutral. Entonces las siguientes condiciones son equivalentes:

1. *F* es monótona y respeta el anonimato.
2. Existe una función monótona $f : [0, n/2] \longrightarrow [0, \infty)$ que satisface $f(t) \ge t$ para todo $t \in [0, n/2]$ y tal que

$$F(D) = 1 \Leftrightarrow n^+(D) > f\left(n^-(D)\right).$$

Sea $\boldsymbol{C}$ el conjunto de funciones monótonas $f : [0, n/2] \longrightarrow [0, \infty)$ que satisfacen $f(t) \ge t$ para todo $t \in [0, n/2]$. De acuerdo con el Corolario 1, dada $f \in \boldsymbol{C}$, se puede asociar con *f* una FES neutral, monótona y que respete el anonimato, F_f, definida por

$$F_f(D) = 1 \Leftrightarrow n^+(D) > f\left(n^-(D)\right).$$

Es importante destacar que dos funciones distintas, $f, g \in \boldsymbol{C}$, pueden generar la misma FES. Para caracterizar esta clase de funciones se utilizará el conjunto N_f, definido por: $\boldsymbol{N}_f = \{t \in \{0,1,\ldots,[n/2]\} \mid t + [f(t)] \le n-1\}$. Este conjunto está estrechamente relacionado con la FES F_f, como se muestra en el siguiente lema.

Lema 1. Sea $f \in \boldsymbol{C}$. Entonces las siguientes condiciones son equivalentes:

1. $t \in N_f$.
2. Existe $D \in \boldsymbol{D}$ tal que $n^-(D) = t$ y $F_f(D) = 1$.

Comÿÿse hÿÿcoÿÿntado anteriorÿÿnte, en la siguiente proposición se caracterizan las funciones que generan una misma FES.

Proposición 1. Sean $f, g \in \boldsymbol{C}$. Entonces las siguientes condiciones son equivalentes:

1. $F_f = F_g$.
2. $\boldsymbol{N}_f = \boldsymbol{N}_g$ y $[f(t)] = [g(t)]$ para todo $t \in \boldsymbol{N}_f$.

Entre las FES neutrales, monótonas y que respetan el anonimato se encuentra la FES nula, definida por $F(D)=0$ para todo $D\in \boldsymbol{D}$. Se puede comprobar fácilmente que una función $f\in \boldsymbol{C}$ genera la FES nula si y sólo si $\boldsymbol{N}_f=\emptyset$, que a su vez es equivalente a $f(0)\geq n$.

Cabe destacar que las FES que habitualmente se emplean en los procedimientos de toma de decisiones se pueden generar mediante funciones afines. Por esta razón, a continuación vamos a analizar el conjunto formado por las funciones afines que generan FES neutrales, monótonas y que respetan el anonimato. Dada una función afín $f(t)=a\,t+b$, en primer lugar se muestra la repercusión que sobre los valores *a* y *b* tienen algunas de las condiciones impuestas a la función $f(t)$.

Observación 1.

1. La monotonía de *f* equivale a $a\geq 0$.
2. La condición $f(t)\geq t$ para todo $t\in[0,n/2]$ equivale a $b\geq 0$ y $b\geq(1-a)(n/2)$.
3. Para no tener en cuenta la FES nula, se considera $f(0)=b<n$.
4. Si $a+b>n-1$ (y $b<n$), entonces la función $g(t)=(n-1-[b])t+[b]$ verifica $\boldsymbol{N}_g=\boldsymbol{N}_f=\{0\}$ y $[g(0)]=[f(0)]$; es decir, por la Proposición 1 *g* genera la misma FES que *f*.

Por tanto, podemos considerar únicamente funciones afines que verifiquen la condición $a+b\leq n-1$.

Así pues, podemos restringir nuestro estudio al siguiente conjunto de funciones:

$$\boldsymbol{L}=\{a\,t+b\mid a\geq 0, b\geq 0, (1-a)(n/2)\leq b\leq n-1-a\}.$$

Como se ha comentado anteriormente, las FES que habitualmente se utilizan en los procedimientos de toma de decisiones se pueden obtener a partir de funciones afines. En el siguiente ejemplo se presentan algunas de estas funciones y las FES que generan.

Ejemplo 1.

1. Cuando $f(t)=t$ se obtiene la *mayoría simple*:

$$F(D)=1\Leftrightarrow n^+(D)>n^-(D).$$

La mayoría simple ha sido caracterizada por diversos autores (véanse, entre otros, May (1952), Fishburn (1973, p. 58, 1983), Campbell (1982, 1988), Maskin (1995), Campbell and Kelly (2000, 2003), Aşan y Sanver (2002), Woeginger (2003, 2005), Miroiu (2004), Yi (2005) y Llamazares (2006)).

2. Cuando $f(t)=n/2$ se obtiene la *mayoría absoluta*:

$$F(D)=1\Leftrightarrow n^+(D)>\frac{n}{2}.$$

La mayoría absoluta ha sido caracterizada por Fishburn (1973, p. 60).

3. Cuando $f(t)=n-1$ se obtiene la *mayoría unánime*:

$$F(D)=1\Leftrightarrow n^+(D)=n.$$

La mayoría unánime ha sido caracterizada por Woeginger (2003) y Llamazares (2006).

4. Cuando $f(t)=(n-1)t$ se obtiene la *mayoría de Pareto*:

$$F(D)=1\Leftrightarrow n^+(D)>0 \text{ y } n^-(D)=0\ .$$

La mayoría de Pareto ha sido caracterizada por Sen (1970, p. 76).

5. Dado $\alpha\in[1/2,1)$, si $f(t)=\alpha n$, entonces se obtienen las *mayorías cualificadas* A_α:

$$A_\alpha(D)=1\Leftrightarrow n^+(D)>\alpha n.$$

Estas FES se encuentran entre la mayoría absoluta (para $\alpha=1/2$) y la mayoría unánime (para $\alpha\geq(n-1)/n$). Las mayorías cualificadas han sido estudiadas por Ferejohn y Grether (1974), y caracterizadas por Fishburn (1973, p. 67) (sin la suposición de neutralidad), Austen-

Smith y Banks (1999, p. 80), Aşan y Sanver (2006) y Houy (2007b). Por otra parte, de acuerdo con el apartado 4 de la Observación 1, si $\alpha \in ((n-1)/n, 1)$, entonces las funciones $f(t) = \alpha n$ y $g(t) = n - 1 = (n-1)/n \cdot n$ generan la misma FES; es decir, $A_\alpha = A_{(n-1)/n}$ para todo $\alpha \in ((n-1)/n, 1)$. Por tanto, podemos considerar únicamente $\alpha \in [1/2, (n-1)/n]$.

6. Dado $\beta \geq 1$, si $f(t) = \beta t$, entonces se obtienen las *mayorías relativas* R_β:

$$R_\beta(D) = 1 \Leftrightarrow n^+(D) > \beta n^-(D).$$

Estas FES se encuentran entre la mayoría simple (para $\beta = 1$) y la mayoría de Pareto (para $\beta \geq n-1$). Las mayorías relativas han sido estudiadas por Craven (1971), Fishburn (1973, p. 68) (sin la suposición de neutralidad), Ferejohn y Grether (1974) y Jain (1983, 1986a). Por otra parte, de acuerdo con el apartado 4 de la Observación 1, si $\beta > n-1$, entonces las funciones $f(t) = \beta t$ y $g(t) = (n-1)t$ generan la misma FES; es decir, $R_\beta = R_{n-1}$ para todo $\beta > n-1$. Por tanto, podemos considerar únicamente $\beta \in [1, n-1]$.

7. Dado $\delta \in [0,1]$, si $f(t) = (1-\delta)/(1+\delta) \cdot t + \delta/(1+\delta) \cdot n$, entonces se obtienen las *mayorías semi-estrictas* S_δ:

$$S_\delta(D) = 1 \Leftrightarrow n^+(D) > \frac{1-\delta}{1+\delta} n^-(D) + \frac{\delta}{1+\delta} n.$$

Estas FES se encuentran entre la mayoría simple (para $\delta = 0$) y la mayoría absoluta (para $\delta = 1$). Las mayorías semi-estrictas han sido estudiadas por Pattanaik (1971, p. 54) y Jain (1986b).

8. Dado $k \in \{0,1,\ldots,n-1\}$, si $f(t) = t + k$, entonces se obtienen las *mayorías por diferencia de votos* M_k:

$$M_k(D) = 1 \Leftrightarrow n^+(D) > n^-(D) + k.$$

Estas FES se encuentran entre la mayoría simple (para $k = 0$) y la mayoría unánime (para $k = n-1$). Las mayorías por diferencia de votos han sido caracterizadas por Llamazares (2006) y Houy (2007a). Por otra parte, de acuerdo con el apartado 4 de la Observación 1, las funciones $f(t) = t + n - 1$ y $g(t) = n - 1$ generan ambas la misma FES, la mayoría unánime. Por tanto, podemos considerar únicamente $k \in \{0,1,\ldots,n-2\}$.

En el siguiente teorema se muestra que el conjunto ***L*** es convexo y que los vértices de dicho conjunto corresponden a las funciones que generan las siguientes mayorías: la mayoría simple, la mayoría absoluta, la mayoría unánime y la mayoría de Pareto.

Teorema 2. El conjunto ***L*** es convexo y sus vértices son las funciones $f(t) = t$, $f(t) = n/2$, $f(t) = n-1$ y $f(t) = (n-1)t$.

Las funciones afines que generan algunas de las FES presentadas anteriormente son combinaciones convexas de dos vértices del conjunto ***L***. Por ejemplo:

1. Las funciones que generan las mayorías cualificadas son combinaciones convexas de las funciones que generan la mayoría absoluta y la mayoría unánime:

$$\alpha n = (1-\theta)\frac{n}{2} + \theta(n-1), \text{ donde } \theta = \frac{\alpha n - n/2}{n/2 - 1}.$$

2. Las funciones que generan las mayorías relativas son combinaciones convexas de las funciones que generan la mayoría simple y la mayoría de Pareto:

$$\beta t = (1-\theta)t + \theta(n-1)t, \text{ donde } \theta = \frac{\beta - 1}{n-2}.$$

3. Las funciones que generan las mayorías semi-estrictas son combinaciones convexas de las funciones que generan la mayoría simple y la mayoría absoluta:

$\frac{1-\delta}{1+\delta}t+\frac{\delta}{1+\delta}n=(1-\theta)t+\theta\frac{n}{2}$, donde $\theta=\frac{2\delta}{1+\delta}$.

Se pueden obtener otras tres familias de FES generadas por funciones que sean combinación convexa de dos vértices del conjunto ***L***. Así:

1. Cuando se consideran combinaciones convexas entre las funciones que generan la mayoría de Pareto y la mayoría unánime, se obtiene $f(t)=(1-\theta)(n-1)t+\theta(n-1)$, con $\theta\in[0,1]$. Si $l=[\theta(n-1)]$, entonces $l\in\{0,1,\ldots,n-1\}$ y la mayoría generada por *f* es

$P_l(D)=1\Leftrightarrow n^+(D)>l$ y $n^-(D)=0$.

2. Cuando se consideran combinaciones convexas entre las funciones que generan la mayoría de Pareto y la mayoría absoluta, se obtiene $f(t)=(1-\theta)(n-1)t+\theta(n/2)$, con $\theta\in[0,1]$. En este caso, la mayoría generada por *f* es

$Q_\theta(D)=1\Leftrightarrow n^+(D)>(1-\theta)(n-1)n^-(D)+\theta\frac{n}{2}$.

3. Cuando se consideran combinaciones convexas entre las funciones que generan la mayoría simple y la mayoría unánime, se obtiene $f(t)=(1-\theta)t+\theta(n-1)$, con $\theta\in[0,1]$. En este caso, la mayoría generada por *f* es

$C_\theta(D)=1\Leftrightarrow n^+(D)>(1-\theta)n^-(D)+\theta(n-1)$.

Por su parte, las funciones que generan las mayorías por diferencia de votos, $f(t)=t+k$, $k\in\{0,1,\ldots,n-2\}$, son combinación convexa de las funciones que generan la mayoría simple, la mayoría de Pareto y la mayoría unánime:

$t+k=\left(1-\frac{k}{n-2}\right)t+\frac{k}{(n-2)(n-1)}(n-1)t+\frac{k}{n-1}(n-1)$.

Por último, cabe destacar que las mayorías generadas a través de las funciones afines del conjunto ***L*** quedan determinadas por los valores que toman los parámetros *a* y *b*. El primero de ellos muestra el apoyo que necesita una opción por cada voto a favor que tenga la otra alternativa, mientras que el segundo refleja el apoyo que necesita una opción para ser elegida cuando la otra no cuenta con ningún respaldo.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por la Consejería de Educación y Cultura de la Junta de Castilla y León (Proyecto VA002B08), el Ministerio de Educación y Ciencia (Proyecto SEJ2006-04267/ECON) y la Unión Europea (FEDER).

BIBLIOGRAFÍA

AŞAN, G Y SANVER, MR (2002), "Another characterization of the majority rule", Economics Letters, 75, pp 409-413.

AŞAN, G Y SANVER, MR (2006), "Maskin monotonic aggregation rules", Economics Letters, 91, pp 179-183.

AUSTEN-SMITH, D Y BANKS, JS (1999), "Positive Political Theory I: Collective Preference", The University of Michigan Press, Ann Arbor.

CAMPBELL, DE (1982), "On the derivation of majority rule", Theory and Decision, 14, pp 133-140.

CAMPBELL, DE (1988), "A characterization of simple majority rule for restricted domains", Economics Letters, 28, pp 307-310.
CAMPBELL, DE Y KELLY, JS (2000), "A simple characterization of majority rule", Economic Theory, 15, pp 689-700.
CAMPBELL, DE Y KELLY, JS (2003), "A strategy-proofness characterization of majority rule", Economic Theory, 22, pp 557-568.
CRAVEN, J (1971), "Majority voting and social choice", Review of Economic Studies, 38, pp 265-267.
FEREJOHN, JA Y GRETHER, DM (1974), "On a class of rational social decisions procedures", Journal of Economic Theory, 8, pp 471-482.
FISHBURN, PC (1973), "The Theory of Social Choice", Princeton University Press, Princeton.
FISHBURN, PC (1983), "A new characterization of simple majority", Economics Letters, 13, pp 31-35.
HOUY, N (2007a), "Some further characterizations for the forgotten voting rules", Mathematical Social Sciences, 53, pp 111-121.
HOUY, N (2007b), "A new characterization of absolute qualified majority voting", Economics Bulletin, 4(3), pp 1-8.
JAIN, SK (1983), "Necessary and sufficient conditions for quasi-transitivity and transitivity of special majority rules", Keio Economic Studies, 20, pp 55-63.
JAIN, SK (1986a), "Special majority rules: Necessary and sufficient condition for quasi-transitivity with quasi-transitive individual preferences", Social Choice and Welfare, 3, pp 99-106.
JAIN, SK (1986b), "Semi-strict majority rules: Necessary and sufficient conditions for quasi-transitivity and transitivity", Mimeo, Jawaharlal Nehru University, New Delhi.
LLAMAZARES, B (2006), "The forgotten decision rules: Majority rules based on difference of votes", Mathematical Social Sciences, 51, pp 311-326.
MASKIN, ES (1995), "Majority rule, social welfare functions and game forms", en: Basu, K, Pattanaik, PK y Suzumura, K (Eds.), Choice, Welfare and Development, The Clarendon Press, Oxford, pp 100-109.
MAY, KO (1952), "A set of independent necessary and sufficient conditions for simple majority decisions", Econometrica, 20, pp 680-684.
MIROIU, A (2004), "Characterizing majority rule: from profiles to societies", Economics Letters, 85, pp 359-363.
PATTANAIK, PK (1971),"Voting and Collective Choice", Cambridge University Press, Cambridge.
SEN, AK (1970), "Collective Choice and Social Welfare", Holden-Day, San Francisco.
WOEGINGER, GJ (2003), "A new characterization of the majority rule", Economics Letters, 81, pp 89-94.
WOEGINGER, GJ (2005), "More on the majority rule: Profiles, societies, and responsiveness", Economics Letters, 88, pp 7-11.
YI, J (2005), "A complete characterization of majority rules", Economics Letters, 87, pp 109-112.

ALGORITMOS GENÉTICOS Y ELECCIÓN DE ESTUDIOS UNIVERSITARIOS

MONTSERRAT HERNÁNDEZ LÓPEZ
E-mail: mhdezl@ull.es
JOSÉ JUAN CÁCERES HERNÁNDEZ
E-mail: jcaceres@ull.es
Departamento de Economía de las Instituciones, Estadística Económica y Econometría
Universidad de La Laguna

Resumen

A procedure based on genetic algorithms is developed to forecast the relative presence of different university studies in higher education demand. The increase in the future demand for a degree is assumed to be related to the job opportunities associated with such a specific university study. In this sense, a selection operator is defined. Furthermore, a transition matrix is introduced to take other explanatory factors into account. The proposed algorithm is applied to forecast the demand composition on 2002/2003 to 2005/2006 courses.

Palabras clave: genetic algorithms, higher education demand.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

1. INTRODUCCIÓN

Un alumno que ha finalizado la Educación Secundaria Obligatoria y se plantea recibir formación universitaria, tiene que elegir entre una amplia oferta de titulaciones. Sus preferencias le permitirán desechar buena parte de ellas, pero, posiblemente, su orientación vocacional no sea tan fuerte como para determinar completamente la decisión final, en la que intervienen múltiples factores. En cualquier caso, en términos agregados y no individuales, cabe suponer que las vocaciones no son susceptibles de cambio a corto plazo y, por tanto, las transformaciones operadas en el número de alumnos que se matricula en las diferentes licenciaturas parecen responder más a las perspectivas de conseguir empleo, que también dependen de otras circunstancias o características del propio alumno.

Este es el punto de vista adoptado en este trabajo de investigación, que pretende, haciendo uso de un algoritmo genético (véase Holland, 1975, Goldberg, 1989 y Hernández-López y Cáceres-Hernández, 2007), predecir la composición de la demanda de un conjunto específico de titulaciones en función de las expectativas laborales que se asocien a cada una de ellas. En concreto, las titulaciones seleccionadas son la Diplomatura en Ciencias Empresariales, la Licenciatura en Administración y Dirección de Empresas y la Licenciatura en Economía, que se imparten en la Facultad de Ciencias Económicas y Empresariales de la Universidad de La Laguna. Nótese que se trata de titulaciones suficientemente homogéneas para que pueda asumirse que las posibilidades de acceso al mercado laboral sean más relevantes que los cambios vocacionales a la hora de obtener las predicciones deseadas. En cualquier caso, debe tenerse en cuenta que la Diplomatura está diseñada en tres cursos académicos, mientras que las licenciaturas consideradas poseen una duración de 4 cursos (hasta el curso 2001/2002) o 5 cursos (desde el 2002/2003). En este sentido, la incorporación al mercado laboral uno o dos años antes es un factor que debe ser ponderado a la hora de evaluar las posibilidades laborales que ofrece cada titulación.

El algoritmo empleado proporciona características individuales de cada uno de los individuos de una población final de alumnos de nuevo ingreso en la que puede, por tanto,

observarse la composición de la demanda de estudios universitarios. La definición de la función de calidad, que determina la actuación del operador selección, se efectúa a partir de la evaluación de las expectativas de empleo del futuro egresado en función de sus características individuales. De esta evaluación se ocupa la siguiente sección. En las secciones tercera y cuarta se exponen las herramientas que componen el algoritmo y se presentan los resultados obtenidos. Finalmente se recogen algunas conclusiones.

2. EVALUACIÓN DE LAS EXPECTATIVAS DE EMPLEO INDIVIDUALES

El punto de partida básico para aplicar la herramienta de predicción planteada es que las titulaciones con mejores expectativas de empleo irán incrementando su peso relativo. La cuestión es cómo definir las expectativas de empleo en términos de la titulación del individuo, así como de otras características. Por ello, y de modo exploratorio, se procedió a examinar las características de los individuos que terminaron sus estudios en los cursos 98/99 a 01/02. Para un grupo de estos egresados, el observatorio canario de empleo registró su actividad laboral en un horizonte de 2 años después del egreso.

A partir de la base de datos original, y para cada uno de los cursos y titulaciones señalados, se construyó una base de datos que recogía, para cada individuo, las características que se indican a continuación: Y:"situación laboral", con las modalidades 1 (parado), 2 (contrato temporal) y 3 (contrato indefinido); XTIT:"titulación", con las modalidades 1 (diplomado en Ciencias Empresariales), 2 (licenciado en Administración y Dirección de Empresas) y 3 (licenciatura en Economía); XEDAD:"edad en el curso de egreso", definida como la diferencia entre el año de egreso y el año de nacimiento y agrupada en las modalidades 1 (año de nacimiento posterior a 1978, 1977, 1976 o 1975 para los egresados del curso 01/02, 00/01, 99/00 o 98/99, respectivamente), 2 (nacimiento en 1977 o 1978, 1976 o 1977, 1975 o 1976, o 1974 o 1975 para los egresados del curso 01/02, 00/01, 99/00 o 98/99, respectivamente) y 3 (año de nacimiento anterior a 1977, 1976, 1975 o 1974 para los egresados del curso 01/02, 00/01, 99/00 o 98/99, respectivamente); XGEN:"género", con las modalidades 1 (femenino) y 2 (masculino).

De la población de ocupados se han eliminado aquellos alumnos que tenían empleo antes del egreso. De la población de parados, se han eliminado todos aquéllos que figuraban también como ocupados. Ahora bien, se han considerado como parados aquéllos que figuran en esta base y que han conseguido empleos coyunturales. Y también fueron eliminados del análisis aquellos egresados que permanecen inactivos y que por tanto no buscan empleo. Por otra parte, si un individuo ha conseguido un trabajo indefinido y luego ha optado por un contrato temporal, se ha incluido en el grupo de los que consiguen trabajo indefinido.

Considerando que los cambios en la demanda de estudios universitarios están relacionados con las expectativas laborales que el alumno matriculado asocia a cada titulación, parece razonable considerar que el alumno configura su idea al respecto en función de lo ocurrido en los últimos años. Es decir, quizás el alumno tome en consideración, más que la información que posea en el último año de educación secundaria obligatoria, un horizonte temporal pasado más amplio. Por esta razón, se ha optado por utilizar los parámetros del logit estimado para las observaciones del conjunto de los cuatro cursos disponibles (del 98/99 al 2001/02), como elemento sobre el que construir las probabilidades de encontrar empleo con contrato indefinido, con contrato temporal o de estar parado. Los resultados de la estimación conjunta del logit con las observaciones correspondientes a los cursos 1998/1999 a 2001/2002 se resumen a continuación (cuadros 1 a 3).

Cuadro 1. Estimaciones de los parámetros del logit multinomial

Regresor	Temporal	Indefinido
Constante	0.7834*	-1.1700*
Empresariales	-0.5299*	-0.6178*
Economía	-0.6234*	-1.2080*
Edad intermedia	0.4094**	0.7305*
Edad más alta	0.0454**	1.1598*
Género masculino	-0.0065	0.3167***

* Significativo al 1%. ** Significativo al 5%. *** Significativo al 10%.

Cuadro 2. Estimación de probabilidades de inserción laboral

Probabilidades predichas medias	Parado	Temporal	Indefinido
Empresariales	0.3149	0.5488	0.1363
ADE	0.2111	0.6216	0.1673
Economía	0.3528	0.5617	0.0855
Edad menor	0.3955	0.5318	0.0728
Edad intermedia	0.2951	0.5937	0.1113
Edad mayor	0.2722	0.5713	0.1566
Género femenino	0.3217	0.5737	0.1047
Género masculino	0.3217	0.5359	0.1427
Total	0.3064	0.5719	0.1217
Cambio discreto medio	Parado	Temporal	Indefinido
Empresariales/ADE	-0.1038	0.0728	0.0311
Empresariales/Economía	0.0379	0.0129	-0.0508
ADE/Economía	0.1417	-0.0599	-0.0819
Edad menor/Edad intermedia	-0.1004	0.0619	0.0385
Edad menor/Edad mayor	-0.1233	0.0395	0.0838
Edad intermedia/Edad mayor	-0.0229	-0.0224	0.0453
Género femenino/masculino	-0.0003	-0.0377	0.0381

Atendiendo a dichos resultados, la licenciatura en ADE ofrece mejores expectativas laborales, con una probabilidad media de encontrar empleo indefinido que duplica a esta misma probabilidad en el caso de la licenciatura en Economía, que es a su vez la que presenta la probabilidad más alta de no encontrar empleo. Los cocientes de odds-ratios revelan que el ratio entre la probabilidad de encontrar empleo indefinido y la probabilidad de no encontrar empleo es para ADE casi el doble que para la diplomatura de Empresariales y más del triple que en Economía. De todas formas, las expectativas laborales no pueden medirse sólo en términos de las probabilidades anteriores si la incorporación al mercado se produce antes como consecuencia de la menor duración de los estudios. De ahí que las expectativas laborales asociadas a la diplomatura de Empresariales deban considerarse mejores que las que se derivan directamente de las estimaciones del modelo logit.

Cuadro 3. Cocientes de odds-ratios. $\dfrac{\Omega_{j/k}\left(x_{i,m}=p\right)}{\Omega_{j/k}\left(x_{i,m}=q\right)}$.

$x_{i,m}$	p/q			p/q			p/q		
Titulación	Empresariales/ADE			Empresariales/Economía			ADE/Economía		
$j^{\cdot}\cdot.k$	1	2	3	1	2	3	1	2	3
1	1	1.6987	1.8549	1	0.9107	0.5543	1	0.5361	0.2988
2	0.5887	1	1.0919	1.0980	1	0.6086	1.8652	1	0.5573
3	0.5391	0.9158	1	1.8042	1.6432	1	3.3466	1.7942	1
Edad	Menor/Intermedia			Menor/Mayor			Intermedia/Mayor		
$j^{\cdot}\cdot.k$	1	2	3	1	2	3	1	2	3
1	1	1.5058	2.0761	1	1.5749	3.1894	1	1.0459	1.5363
2	0.6641	1	1.3787	0.6350	1	2.0251	0.9561	1	1.4689
3	0.4817	0.7253	1	0.3135	0.4938	1	0.6509	0.6808	1
Género	Masculino/Femenino								
$j^{\cdot}\cdot.k$	1	2	3						
1	1	0.9370	1.3725						
2	1.0673	1	1.4649						
3	0.7286	0.6827	1						

Por otro lado, la edad parece actuar positivamente en las expectativas laborales. Así, la probabilidad de encontrar empleo indefinido entre los de más edad más que duplica esta probabilidad entre los más jóvenes, mientras que la probabilidad de no encontrar empleo para los primeros está más de 10 puntos porcentuales por debajo de la probabilidad estimada para los

segundos. Esta circunstancia queda reflejada en los cocientes de odds-ratios entre las dos probabilidades comentadas para los diferentes grupos de edad.

Y con respecto al género, no existen diferencias en las probabilidades de no encontrar empleo. Ahora bien, existe una propensión hacia el empleo temporal ligeramente mayor entre las mujeres, mientras que la probabilidad de encontrar empleo indefinido es mayor entre los hombres.

3. FUNCIÓN DE CALIDAD, OPERADOR SELECCIÓN Y MATRICES DE TRANSICIÓN

En términos del algoritmo genético, las probabilidades estimadas a través del logit constituyen el elemento a partir del cual se define una medida de facilidad de inserción laboral como función de las características del individuo. En otras palabras, se obtiene una función de calidad que permite definir el valor de adecuación del individuo al contexto, es decir, una medida de la calidad con la que un individuo con determinadas características accede al mercado laboral (parado, con contrato temporal o con contrato indefinido). Es lo que se llama en términos del algoritmo genético, *fitness function*, cuyo valor medio para la población tiende a incrementarse de generación en generación, como resultado de la aplicación del principio básico de "supervivencia de los mejores", concepto fundamental en la construcción y definición de un algoritmo genético.

En concreto, se ha asumido que la calidad del individuo i es:

$$C_i^* = 0P(Y_i = 1) + 3P(Y_i = 2) + 10P(Y_i = 3)$$

donde $P(Y_i = j)$, j = 1, 2, 3, son las probabilidades estimadas de que el individuo i esté en la situación laboral j. Los valores 0, 3 y 10 se han elegido de forma arbitraria, pero teniendo en cuenta que tiene más peso específico tener un contrato indefinido (valor 10) que uno temporal (valor 3) o que no tener empleo (valor 0), y que el paso de no tener empleo a tener uno de carácter temporal es de menor importancia que el paso de tener empleo temporal a una situación laboral indefinida. Por otra parte, dado que los estudios de la diplomatura están diseñados para tres años, mientras que la duración prevista de los estudios correspondientes a las dos licenciaturas consideradas es de 4 años (hasta el curso 2001/2002) o 5 años (desde el curso 2002/2003), se ha aplicado un factor de ponderación de 4/3 al valor de la calidad para los alumnos de la diplomatura.

Con esta función de calidad, el algoritmo genético se ha ejecutado en dos fases sobre la población de individuos del curso 01/02. En la primera, una vez ordenadas las estructuras observadas en la población original se ha aplicado el operador selección proporcional para generar una nueva población de 1000 individuos. En la segunda fase, a cada una de las 1000 copias anteriores se la ha aplicado la fila oportuna de la matriz de transición, es decir, la correspondiente a la estructura copiada en la fase de selección. Una vez identificada la copia final en que se transforma cada estructura copiada en la fase de selección, se tiene una población final de 1000 nuevos estudiantes definidos a través de estructuras que permiten la comparación con la población futura 2002/03 con el fin de evaluar la capacidad predictiva del algoritmo en términos del peso de las diferentes titulaciones. Nótese que se está obteniendo una predicción para 2002/2003 con información sobre posibilidades de empleo durante los dos años posteriores al egreso en el curso 2001/2002. Por tanto, esta predicción se utiliza a efectos de valoración del funcionamiento del algoritmo y mejora de su diseño para predicciones posteriores al periodo observado.

El operador selección incorporado se ha definido de dos maneras distintas: 1) Proporcional a la calidad; 2) Proporcional al cuadrado de la calidad. La explicación para el planteamiento de esta segunda alternativa radica en que la primera parecía incapaz de captar la velocidad de cambio de la población original.

La matriz de transición puede ser definida de dos formas alternativas. 1) Se puede asumir que la probabilidad de que una estructura copiada en la fase de selección no se altere es suficientemente alta y las transformaciones más probables son aquéllas que implican cambios

menos bruscos, es decir, cambios en menor número de características (matriz A, cuadro 4). Así, el número de diferencias entre dos estructuras se ha calculado como se indica a continuación: a) titulación: de diplomado a licenciado (2), de licenciado a licenciado (1), no cambio de titulación (0); b) edad: de edad 1 a edad 3 (2), de edad 1 a edad 2 o de edad 2 a edad 3 (1), no cambio de edad (0); c) género: cambio de género (1), no cambio de género (0), con lo que el número máximo de diferencias es 5. Una vez evaluado el número de diferencias d entre estructuras, para cada estructura observada se calcula el inverso del número de diferencias existente entre tal estructura y las demás (valor i). Así se recoge la idea de que es más factible un cambio de una estructura a otra con la que presente menos diferencias que a otra con mayor número de diferencias. Entonces para transformar estos valores en probabilidades, y teniendo en cuenta que cada estructura se transformará en cualquiera de las otras observadas (incluyendo la posibilidad de que se transforme en sí misma), los valores i se multiplicarán por una constante de tal forma que, evidentemente, la suma por fila sea igual a 1. De esta primera forma, los parámetros de la matriz de transición han sido construidos sobre la idea de que las probabilidades de transformación de una estructura en otra son inversamente proporcionales al número de diferencias entre ambas. 2) Otra posibilidad es que los parámetros de la matriz de transición se deduzcan a partir de los cambios en la calidad. De este modo, sería más factible un cambio hacia una estructura que implique menos cambio de calidad que un cambio hacia otra que signifique un cambio más brusco (matriz B, cuadro 5).

Cuadro 4. Matriz de transición A

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0.80	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
2	0.03	0.80	0.01	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
3	0.03	0.01	0.80	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
4	0.01	0.03	0.03	0.80	0.01	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
5	0.01	0.01	0.03	0.01	0.80	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
6	0.01	0.01	0.01	0.03	0.03	0.80	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
7	0.01	0.01	0.01	0.01	0.01	0.00	0.80	0.02	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01
8	0.01	0.01	0.01	0.01	0.00	0.01	0.02	0.80	0.01	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01
9	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.80	0.02	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01
10	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.80	0.01	0.02	0.01	0.01	0.01	0.02	0.01	0.01
11	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.80	0.02	0.01	0.01	0.01	0.01	0.02	0.01
12	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.80	0.01	0.01	0.01	0.01	0.01	0.02
13	0.01	0.01	0.01	0.01	0.01	0.00	0.02	0.01	0.01	0.01	0.01	0.01	0.80	0.02	0.02	0.01	0.01	0.01
14	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.02	0.80	0.01	0.02	0.01	0.01
15	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.02	0.01	0.80	0.02	0.02	0.01
16	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.02	0.02	0.80	0.01	0.02
17	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.02	0.01	0.80	0.02
18	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.02	0.02	0.80

Cuadro 5. Matriz de transición B

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0.80	0.01	0.00	0.00	0.00	0.00	0.01	0.13	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03
2	0.02	0.80	0.01	0.00	0.00	0.00	0.01	0.02	0.07	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.02
3	0.01	0.01	0.80	0.01	0.01	0.00	0.00	0.01	0.01	0.03	0.09	0.01	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	0.01	0.80	0.02	0.01	0.00	0.00	0.00	0.01	0.01	0.12	0.00	0.00	0.00	0.00	0.00	0.00
5	0.01	0.01	0.01	0.05	0.80	0.01	0.00	0.00	0.01	0.01	0.01	0.06	0.00	0.00	0.00	0.00	0.00	0.00
6	0.01	0.01	0.01	0.02	0.03	0.80	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01
7	0.02	0.01	0.00	0.00	0.00	0.00	0.80	0.02	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.06	0.02
8	0.13	0.01	0.00	0.00	0.00	0.00	0.01	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03
9	0.02	0.07	0.01	0.01	0.00	0.00	0.01	0.01	0.80	0.01	0.01	0.01	0.00	0.00	0.01	0.01	0.01	0.01
10	0.01	0.01	0.04	0.01	0.01	0.00	0.01	0.01	0.01	0.80	0.06	0.01	0.00	0.00	0.00	0.00	0.01	0.01
11	0.01	0.01	0.08	0.01	0.01	0.00	0.00	0.01	0.01	0.05	0.80	0.01	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.01	0.12	0.03	0.01	0.00	0.00	0.00	0.01	0.01	0.80	0.00	0.00	0.00	0.00	0.00	0.00
13	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.80	0.09	0.02	0.01	0.01	0.01
14	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.08	0.80	0.02	0.01	0.01	0.01
15	0.01	0.01	0.01	0.00	0.00	0.00	0.02	0.01	0.01	0.01	0.01	0.00	0.01	0.02	0.80	0.04	0.03	0.01
16	0.01	0.01	0.00	0.00	0.00	0.00	0.03	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.03	0.80	0.05	0.01
17	0.01	0.01	0.00	0.00	0.00	0.00	0.06	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.04	0.80	0.02
18	0.05	0.01	0.00	0.00	0.00	0.00	0.02	0.06	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.80

Los parámetros de estas matrices de transición pueden ajustarse en función de la bondad predictiva del algoritmo para años anteriores y, de hecho, parece razonable utilizar los parámetros de las matrices de transición derivadas para un curso académico como herramienta predictiva para los cursos académicos siguientes en los que no se conozcan las proporciones de demanda de titulaciones. En este sentido, la demanda de matrícula en junio de 2002 en primera opción en las titulaciones consideradas se distribuyó como se indica: Empresariales, 61.55%; ADE, 30.08%; Economía, 8.38%. Y aunque los resultados no se muestran aquí por motivos de espacio, la aplicación del algoritmo genético construido con las dos matrices de transición A (probabilidad de transición inversamente proporcional al cambio en el número de características) y B (probabilidad de transición inversamente proporcional al cambio en la calidad) a la población de alumnos del curso 01/02 para predecir la demanda en 2003 proporcionó resultados peores que la predicción obtenida con un algoritmo genético en el que sólo operase el operador selección A (selección proporcional a la calidad) o B (selección proporcional al cuadrado de la calidad). Esta circunstancia indica que las matrices de transición están construidas de forma que sesgan la predicción y llegan a distorsionar la dirección de cambio indicada por el operador selección. En este sentido, parece lógico plantearse: a) ¿cómo debe construirse la matriz de transición de modo que la predicción sea más acertada?; b) ¿qué transformaciones está implicando cada una de las matrices de transición antes definidas?; c) ¿cómo corregir estas matrices para que actúen generando predicciones acertadas? Por supuesto, la corrección de las matrices de transición en función del error de predicción para 02/03 sólo se sostiene en el caso de que se vayan a realizar predicciones posteriores. A continuación se aborda cada una de las tres cuestiones anteriores.

4. AJUSTE DEL ALGORITMO Y PREDICCIÓN DE DEMANDA EN CURSOS 2003/2004 A 2005/2006

En el algoritmo construido, la combinación de las modalidades de los tres atributos considerados (titulación, edad y género) da lugar a 18 estructuras distintas $\{E_i\}_{i=1,\ldots 18}$. La probabilidad de seleccionar la estructura E_i en la fase de selección, $p_S^A(i)$ o $p_S^B(i)$, viene determinada por el peso de dicha estructura en la población original, así como por su calidad y a partir de las probabilidades anteriores, pueden deducirse las probabilidades de selección para cada una de las tres titulaciones. Cuando se aplica el operador selección a la población del curso 2001/2002, estas probabilidades de selección por titulación son las que se muestran en el cuadro 6.

Cuadro 6. Probabilidades de selección y probabilidades finales predichas (2002/2003)

Titulación	$p_S^A(T_i)$	$p_S^B(T_i)$	$p_F^{AA}(T_i)$	$p_F^{AB}(T_i)$	$p_F^{BA}(T_i)$	$p_F^{BB}(T_i)$
Empresariales	0.5282	0.5833	0.4949	0.4922	0.5417	0.5391
ADE	0.2962	0.2864	0.3017	0.3294	0.2926	0.3221
Economía	0.1756	0.1303	0.2034	0.1785	0.1656	0.1389

Si se define como T_i a una de las titulaciones consideradas (Empresariales, A; ADE, B; Economía, C), se denota la probabilidad de que en cualquiera de las estructuras seleccionadas a través del operador selección se elija la titulación T_i por $p_S(T_i)$, se define $p_T(T_i, T_j)$ como la probabilidad de transición de la titulación T_i a la titulación T_j, y se denota por $p_F(T_i)$ a la probabilidad de que aparezca la titulación T_i en cualquiera de las estructuras de la población final, entonces se tiene que

$$p_F(A) = p_S(A)p_T(A,A) + p_S(B)p_T(B,A) + p_S(C)p_T(C,A)$$
$$p_F(B) = p_S(A)p_T(A,B) + p_S(B)p_T(B,B) + p_S(C)p_T(C,B)$$

$$p_F(C) = p_S(A)p_T(A,C) + p_S(B)p_T(B,C) + p_S(C)p_T(C,C).$$

Las probabilidades $p_S(A)$, $p_S(B)$ y $p_S(C)$ constituyen parámetros que vienen dados por la población original y el operador selección aplicado. Y si se desea una predicción insesgada (que se espere acertar), las probabilidades $p_F(A)$, $p_F(B)$ y $p_F(C)$ deben coincidir con la proporciones de individuos que demandan estas titulaciones en junio de 2002 en primera opción. En cualquier caso, las tres ecuaciones anteriores no permiten determinar los elementos de la matriz de transición

$$\begin{pmatrix} p_T(A,A) & p_T(A,B) & p_T(A,C) \\ p_T(B,A) & p_T(B,B) & p_T(B,C) \\ p_T(C,A) & p_T(C,B) & p_T(C,C) \end{pmatrix}$$

De hecho, las tres ecuaciones son linealmente dependientes y, por tanto, sería necesario definir siete restricciones para configurar un sistema determinado. En este sentido, es claro que

$$p_T(A,A) + p_T(A,B) + p_T(A,C) = 1$$
$$p_T(B,A) + p_T(B,B) + p_T(B,C) = 1$$
$$p_T(C,A) + p_T(C,B) + p_T(C,C) = 1$$

Pero es necesario incorporar restricciones adicionales. Una posibilidad consiste en hacer coincidir las probabilidades que implican transición en un sentido y en el contrario y fijar de antemano una de las probabilidades de transición, de modo que todos los demás elementos de la matriz quedan determinados por el sistema. Sin embargo, esta opción implica una simetría que no se corresponde con la dinámica de cambio observada y, de hecho, si se imponen estas restricciones, las probabilidades resultantes son siempre, en algún caso, inconsistentes con los axiomas. Por ello, se optó por fijar las tres probabilidades de no transición y otro elemento adicional de la matriz. Es decir, las restricciones son

$$p_T(A,A) = k_A,\ p_T(B,B) = k_B,\ p_T(C,C) = k_C,\ p_T(A,C) = 0.$$

Dado que el operador selección funciona relativamente bien, la matriz de transición distorsionará menos el efecto del operador anterior cuanto más alta sea la probabilidad de no transición. Por ello, se optó por fijar $p_T(A,A) = 0.9$, $p_T(B,B) = 0.8$ y $p_T(C,C) = 0.4$. La probabilidad de transición de la titulación A a la titulación C se consideró nula porque implica la transición desde la titulación más frecuente (infraestimada por el algoritmo) a la menos frecuente (sobreestimada por el algoritmo). De acuerdo con este segundo planteamiento, y si se consideran los parámetros propios del operador selección proporcional a la calidad, los parámetros apropiados de la matriz de transición son

$$\begin{pmatrix} p_T(A,A) & p_T(A,B) & p_T(A,C) \\ p_T(B,A) & p_T(B,B) & p_T(B,C) \\ p_T(C,A) & p_T(C,B) & p_T(C,C) \end{pmatrix} = \begin{pmatrix} 0.90 & 0.10 & 0.00 \\ 0.15 & 0.80 & 0.05 \\ 0.54 & 0.06 & 0.40 \end{pmatrix}$$

En cambio, si se consideran los parámetros propios del operador selección proporcional a la calidad al cuadrado y se fijan las probabilidades $p_T(A,A) = 0.9$, $p_T(B,B) = 0.8$ y $p_T(C,C) = 0.4$, los parámetros apropiados de la matriz de transición son

$$\begin{pmatrix} p_T(A,A) & p_T(A,B) & p_T(A,C) \\ p_T(B,A) & p_T(B,B) & p_T(B,C) \\ p_T(C,A) & p_T(C,B) & p_T(C,C) \end{pmatrix} = \begin{pmatrix} 0.90 & 0.10 & 0.00 \\ 0.09 & 0.80 & 0.11 \\ 0.50 & 0.10 & 0.40 \end{pmatrix}$$

Si ahora se definen $p_T^A(i,j)$ y $p_T^B(i,j)$ como las probabilidades de transición de la estructura i a la estructura j con las dos matrices de transición originales, y se denota por $p_F(i)$ a la probabilidad de que aparezca la estructura i en cualquiera de las estructuras de la población final, entonces se tiene que

$$p_F(i) = \sum_{j=1}^{18} p_S^S(j) p_T^M(j,i),\ i = 1,...,18,\ S = A,B,\ M = A,B$$

Así, se obtienen las probabilidades finales por titulación que también se indican en el cuadro 6.

Las probabilidades anteriores $p_F(i)$ determinan las frecuencias relativas esperadas de la estructura i en la población predicha y éstas deberían aproximarse a las observadas en la población real. Para que esto ocurra, una posibilidad es que, si se utiliza el operador selección proporcional a la calidad y, por tanto

$$\begin{pmatrix} p_T(A,A) & p_T(A,B) & p_T(A,C) \\ p_T(B,A) & p_T(B,B) & p_T(B,C) \\ p_T(C,A) & p_T(C,B) & p_T(C,C) \end{pmatrix} = \begin{pmatrix} 0.90 & 0.10 & 0.00 \\ 0.15 & 0.80 & 0.05 \\ 0.54 & 0.06 & 0.40 \end{pmatrix},$$

se impongan las restricciones siguientes

$$\sum_{j=1}^{6} p_T(j,i) = 0.90, \sum_{j=7}^{12} p_T(j,i) = 0.10, \sum_{j=7}^{12} p_T(j,i) = 0.00, i = 1,...,6$$

$$\sum_{j=1}^{6} p_T(j,i) = 0.15, \sum_{j=7}^{12} p_T(j,i) = 0.80, \sum_{j=7}^{12} p_T(j,i) = 0.05, i = 7,...,12$$

$$\sum_{j=1}^{6} p_T(j,i) = 0.54, \sum_{j=7}^{12} p_T(j,i) = 0.06, \sum_{j=7}^{12} p_T(j,i) = 0.40, i = 13,...,18$$

Y una forma de que operen estas restricciones consiste en distribuir la suma correspondiente a cada restricción entre los elementos que intervienen en la suma proporcionalmente a los valores en las matrices originales.

En cambio, si se utiliza el operador selección proporcional a la calidad al cuadrado y, por tanto

$$\begin{pmatrix} p_T(A,A) & p_T(A,B) & p_T(A,C) \\ p_T(B,A) & p_T(B,B) & p_T(B,C) \\ p_T(C,A) & p_T(C,B) & p_T(C,C) \end{pmatrix} = \begin{pmatrix} 0.90 & 0.10 & 0.00 \\ 0.09 & 0.80 & 0.11 \\ 0.50 & 0.10 & 0.40 \end{pmatrix},$$

deben imponerse las restricciones siguientes

$$\sum_{j=1}^{6} p_T(j,i) = 0.90, \sum_{j=7}^{12} p_T(j,i) = 0.10, \sum_{j=7}^{12} p_T(j,i) = 0.00, i = 1,...,6$$

$$\sum_{j=1}^{6} p_T(j,i) = 0.09, \sum_{j=7}^{12} p_T(j,i) = 0.80, \sum_{j=7}^{12} p_T(j,i) = 0.11, i = 7,...,12$$

$$\sum_{j=1}^{6} p_T(j,i) = 0.50, \sum_{j=7}^{12} p_T(j,i) = 0.10, \sum_{j=7}^{12} p_T(j,i) = 0.40, i = 13,...,18$$

para que las frecuencias relativas esperadas de la estructura i en la población predicha se aproximen a las observadas en la población real. Y una forma de que operen estas restricciones consiste en distribuir la suma correspondiente a cada restricción entre los elementos que intervienen en la suma proporcionalmente a los valores en las matrices originales.

Este planteamiento implica que pueden definirse seis matrices de transición distintas: matriz A (probabilidades de transición inversamente proporcionales al número de características diferentes), matriz B (probabilidades de transición inversamente proporcionales a la diferencia de calidad), matriz C (matriz A corregida de acuerdo con la aplicación del operador selección proporcional a la calidad), matriz D (matriz A corregida de acuerdo con la aplicación del operador selección proporcional al cuadrado de la calidad), matriz E (matriz B corregida de acuerdo con la aplicación del operador selección proporcional a la calidad), matriz F (matriz B corregida de acuerdo con la aplicación del operador selección proporcional al cuadrado de la calidad). Y, por tanto, pueden construirse ocho algoritmos distintos, que, en función del

operador selección y la matriz de transición aplicados se denotarán por los indicadores AA, AB, AC, AD, BA, BB, BE, BF.

Dado que se dispone de los datos de matrícula de los cursos 02/03, 03/04 y 04/05, proporcionados por el Gabinete de Análisis y Planificación de la Universidad de La Laguna, estas poblaciones respectivas pueden usarse como poblaciones originales sobre las que se aplica el algoritmo genético para obtener las predicciones del año siguiente. En principio, cabría pensar en aplicar la población final derivada de la aplicación del algoritmo como población original para predecir la demanda del año siguiente; pero la demanda expresada en función de la preinscripción no coincide con la matrícula de ese año como consecuencia del efecto de los cupos por titulación. Estos cupos son: 350 en Empresariales, 150 en ADE (175 en 04/05 y 05/06) y 125 en Economía. Excepto en el caso de la licenciatura en Economía y en el caso de la licenciatura de ADE para el curso 2004/2005, la demanda es superior al cupo desde el curso 1998/1999 hasta 2005/2006.

Estas poblaciones originales se han construido seleccionando de la base de matrícula aquellos alumnos matriculados de primero como curso más alto que hayan accedido al centro en el año de la matrícula. Nótese que en la definición de la función de calidad a partir del logit estimado para el periodo 98/99-01/02, las variables determinantes de la probabilidad de encontrar empleo son la titulación, el género y la edad de egreso. Dado que la población original considerada para la predicción es la población matriculada en primero, las modalidades del atributo XEDAD:"edad en el curso de egreso", se han definido sumando a la diferencia entre el año de matrícula y el año de egreso 2 o 4 años en función de la duración teórica de que la titulación sea en la diplomatura de empresariales o en las licenciaturas de ADE o Economía (nótese que el nuevo plan de 5 años comenzó a impartirse en el curso 2002/2003). Por otra parte, el propio valor de la función de calidad debe modificarse teniendo en cuenta que algunos alumnos cursan titulaciones de 5 años. En este sentido, para efectuar estas predicciones se consideró inicialmente un factor de ponderación de 5/3 en la asignación de calidad a los alumnos que cursan la diplomatura. Ahora bien, el operador selección tendía a generar predicciones que sobrevaloraban claramente la frecuencia esperada de alumnos que demandan esta titulación. Por tanto, se optó por mantener el factor de ponderación 4/3 en la determinación del valor de la función de calidad. En los cuadros 7 a 9, se muestran las probabilidades de selección y las probabilidades finales para la predicción de los cursos 03/04, 04/05 y 05/06 usando las matrices de transición no corregidas. Los resultados de la aplicación de los otros algoritmos se espera que coincidan con las demandas registradas en el curso predicho. Estas demandas se recogen en el cuadro 10.

Cuadro 7. Probabilidades de selección y probabilidades finales predichas (2003/2004)

Titulación	$p_S^A(T_i)$	$p_S^B(T_i)$	$p_F^{AA}(T_i)$	$p_F^{AB}(T_i)$	$p_F^{BA}(T_i)$	$p_F^{BB}(T_i)$
Empresariales	0.6033	0.6593	0.5583	0.5329	0.6059	0.5812
ADE	0.2437	0.2308	0.2575	0.2885	0.2459	0.2809
Economía	0.1529	0.1100	0.1841	0.1786	0.1481	0.1379

Cuadro 8. Probabilidades de selección y probabilidades finales predichas (2004/2005)

Titulación	$p_S^A(T_i)$	$p_S^B(T_i)$	$p_F^{AA}(T_i)$	$p_F^{AB}(T_i)$	$p_F^{BA}(T_i)$	$p_F^{BB}(T_i)$
Empresariales	0.6108	0.6659	0.5647	0.5360	0.6117	0.5833
ADE	0.2367	0.2232	0.2517	0.2852	0.2397	0.2776
Economía	0.1526	0.1108	0.1836	0.1788	0.1486	0.1391

Cuadro 9. Probabilidades de selección y probabilidades finales predichas (2005/2006)

Titulación	$p_S^A(T_i)$	$p_S^B(T_i)$	$p_F^{AA}(T_i)$	$p_F^{AB}(T_i)$	$p_F^{BA}(T_i)$	$p_F^{BB}(T_i)$
Empresariales	0.5805	0.6281	0.5387	0.5230	0.5792	0.5642
ADE	0.2680	0.2635	0.2779	0.3049	0.2734	0.3036
Economía	0.1514	0.1084	0.1834	0.1721	0.1474	0.1323

Cuadro 10. Composición de la demanda de las titulaciones

		Titulación (%)	
Curso	Empresariales	ADE	Economía
03/04	0.6429	0.2955	0.0616
04/05	0.6805	0.2368	0.0827
05/06	0.6751	0.2680	0.0569

5. CONCLUSIONES

Este trabajo constituye una primera aproximación que, más que obtener unas predicciones concretas de la composición de la demanda de estudios universitarios, pretende mostrar una metodología útil para obtenerlas. Por tanto, siendo conscientes de las limitaciones en los datos y de las posibilidades de mejora de los elementos que configuran el algoritmo genético propuesto, no se han ofrecido siquiera predicciones de cursos recientes.

Ahora bien, el planteamiento metodológico adoptado ofrece la ventaja de su flexibilidad. Y, en particular, merece la pena destacar que la matriz de transición puede elegirse no sólo para introducir aleatoriamente cambios que no responden al principio de calidad, sino que también puede jugar un importante papel para introducir cambios en determinada dirección en función de la capacidad del operador selección para dirigir el cambio predicho en la misma dirección del cambio observado en la composición de la población. En este sentido, la matriz de transición puede ser un elemento activo para incorporar factores que intervienen en la decisión de elegir una licenciatura y que no están tan directamente relacionados con las expectativas laborales o que, aún estándolo, no sean correctamente capturados por la función de calidad empleada o por el operador selección definido.

BIBLIOGRAFÍA

GOLDBERG, DE (1989), Genetic Algorithms in Search Optimization, and Machine Learning, New York: Addison Wesley, Reading, M.A.

HERNÁNDEZ-LÓPEZ, M, Y CÁCERES-HERNÁNDEZ, JJ (2007), Forecasting tourists' characteristics by a genetic algorithm with a transition matrix, Tourism Management, 28, pp 290-297.

HOLLAND, JH (1975), Adaptation in Natural and Artificial Systems, Michigan: Ann Arbor. The University of Michigan Press.

DESEMPREGO, INFLAÇÃO, E POPULARIDADE DOS GOVERNOS EM PORTUGAL: QUE RELAÇÃO EXISTE?

ANTÓNIO CALEIRO
E-mail: caleiro@uevora.pt
Departamento de Economia & CEFAGE
Universidade de Évora

Resumo

Este trabalho apresenta uma abordagem de lógica difusa permitindo mostrar como pode ser determinado um indicador de popularidade à luz daquela abordagem. Esta pode ser vista como uma abordagem complementar à subjacente na determinação das 'tradicionais' funções de popularidade através do recurso a métodos de regressão. Através da consideração de dados (mensais, entre 1983 e 2007), para Portugal, para as taxas de inflação e de desemprego, o indicador de popularidade ajuda a compreender os resultados eleitorais que se verificaram em Portugal naquele período e, em particular, detecta claramente as vitórias e as derrotas eleitorais dos partidos no governo.

Palavras-chave: Desemprego, Eleições, Inflação, Lógica Difusa.
Área Temática: Métodos Quantitativos para Economia e Empresa.

Abstract

The paper presents a fuzzy logic approach showing how a popularity index can be determined. The approach here considered can be viewed as a complement to the 'traditional' popularity functions determined using regression methods. Considering monthly data, from 1983 until 2007, for Portugal, the fuzzy logic popularity index, based on inflation and unemployment rates, helps to understand the Portuguese election results and, in particular, clearly detects the electoral victories and defeats of the incumbents.

Keywords: Elections, Fuzzy Logic, Inflation, Unemployment.
Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUÇÃO

Durante muito tempo, a visão proeminente no campo da política económica assumiu que o Governo actuaria de acordo com uma função de bem-estar social e, assim, determinaria e, posteriormente, aplicaria medidas de política económica que, de acordo com aquela função objectivo, seriam as ideais. Como é sabido, esta visão do Governo, dita de ditador benevolente, foi colocada em causa, em primeiro lugar, pela escola da escolha pública [*public choice*] e, posteriormente, também pela abordagem dos ciclos políticos. De acordo com esta nova visão, o Governo, tal como qualquer outro agente, deveria ser considerado com um agente decisor egoísta, ou seja prosseguindo os seus próprios interesses, os quais, em algumas circunstâncias, poderiam não coincidir com a maximização de bem-estar social (Nordhaus, 1975). Por outras palavras, de acordo com esta abordagem alternativa, o Governo poderia ser visto como um agente essencialmente constituído por políticos, cujo interesse privado, ou seja, o desejo de serem reeleitos, poderia não ser (quase certamente) compatível com o interesse geral, ou seja, com a maximização do bem-estar social.

Na verdade, na base daquela visão alternativa está o facto de o Governo possuir duas

características: é um agente económico porque é um dos principais responsáveis pela política económica, embora seja também um agente político, uma vez que existe como resultado de um processo eleitoral nas democracias. Assim, a acção do Governo influencia o estado da economia, mas, por sua vez, também o Governo é influenciado pela situação económica, uma vez que o eleitorado normalmente o considera como sendo um dos principais responsáveis pelo estado da economia e, por este facto, o recompensa ou penaliza eleitoralmente. Neste facto consiste a chamada hipótese da responsabilidade ou da votação económica (Lewis-Beck & Paldam, 2000).

De acordo com a hipótese da votação económica, porque se admite que os eleitores consideram o Governo como sendo responsável pela evolução da economia, as funções de votação ou de popularidade, as quais explicam o apoio ao Governo como uma função de resultados económicos, podem facilmente tornar-se *as* funções objectivo da política económica governamental. Reflectindo esta abordagem, os primeiros estudos empíricos sobre o assunto começaram com a formulação e estimação de funções de popularidade ou de votação. De facto, este tipo de estudos empíricos foi, e ainda é, tão popular, que se podem facilmente contar centenas de referências (para um levantamento detalhado sobre as funções de popularidade ou de votação veja-se Nannestad & Paldam, 1994). Entre essas referências, apenas para referir algumas que ilustram a diversidade dos períodos de tempo e dos países sob investigação, podem referir-se Alvarez et al. (2000), Anderson (2000), Blount (2002), Byers et al. (2000), Carlsen (2000), Johnston (1999), Lewis-Beck & Stegmaier (2000), Sanders (2004), Tellier (2006), e Youde (2005).

A inclusão da popularidade na função objectivo do Governo, o que implica a existência de ciclos económicos como ciclos eleitoralmente induzidos, apresenta-se em contraste com as explicações tradicionais do ciclo económico, as quais incidiram sobre o papel dos choques aleatórios e da instabilidade estrutural. De facto, a teoria da estabilização baseava-se, obviamente, no pressuposto de que o Governo deveria estabilizar a economia. A capacidade de realmente o fazer foi questionada pelos monetaristas no final dos anos 1960 e 1970, mas a modelização político-económica veio apontar uma razão para acreditar que os governos não só não estão dispostos a estabilizar a economia, mas que intencionalmente criar alguma instabilidade cíclica, a fim de obter ganhos eleitorais. Esta desestabilização deliberada resulta do facto de que os eleitores tenderem a favorecer os governos que obtêm bons resultados económicos perto das eleições e, em consequência, esses mesmos governos, ao explorarem a memória (decrescente) dos eleitores, imporem, em algumas circunstâncias, um desempenho económico relativamente menos bom no início do mandato.

A perspectiva dos ciclos eleitorais, ao supor que a popularidade ou a votação depende da evolução das variáveis económicas, na verdade representa um problema para a estimação das funções de popularidade ou de votação, tendo em conta a endogeneidade das variáveis que lhe está implícita, como, por exemplo, ao supor-se também que as trajectórias das variáveis económicas eleitoralmente relevantes dependem da evolução da popularidade em todo o mandato.[1] Por exemplo, a popularidade pode ser encarada como uma função do desemprego, mas esta variável económica pode, também, depender da popularidade se o Governo a utilizar como forma de aumentar o seu nível de popularidade (veja-se, por exemplo, Blount, 2002, e Veiga & Chappell, 2002). Claramente, neste aspecto pode recordar-se também a abordagem complementar de Frey e Schneider, considerando uma função de reacção, a qual tem em conta aquele outro sentido na interacção entre aspectos económicos e aspectos eleitorais (Frey & Schneider, 1979).

Outro ponto que importa referir prende-se com a importância das características dos eleitores, enquanto elemento decisivo na gestão da popularidade [por exemplo, Sanders (2000,2003) e Glasgow & Weber (2005) apontam para a importância da percepção dos eleitores

[1] Este factor pode ser (parcialmente) explicativo de um aspecto muito reconhecido na literatura, ou seja a instabilidade das funções de popularidade estimadas econometricamente, quando esta endogeneidade não é tida em conta, i.e., por exemplo, aquando da estimação de modelos uni-equacionais.

na avaliação do desempenho dos governos acerca da economia]. Como tal, a memória dos eleitores é claramente importante no controle do ciclo eleitoral. Além disso, a (i-)racionalidade dos eleitores, ou mesmo o nível de conhecimento que caracteriza o eleitorado, são também questões essenciais no debate [veja-se, por exemplo, Aidt (2000) e Paldam & Nannestad (2000) para uma análise do nível de conhecimento que os eleitores dispõem sobre a evolução da economia e/ou sobre o desempenho do Governo].

De facto, tendo por base informações incompletas, os eleitores podem decidir votar (ou não) no actual Governo, nas eleições após o final do seu mandato, com base numa percepção vaga do desempenho governamental. Isto significa que, ao serem tomadas em conta, por exemplo, as trajectórias do desemprego e da inflação ao longo do mandato para decidir votar (ou não) no Governo actual – considerado responsável por essas trajectórias – os eleitores podem basear as suas decisões em conceitos bastante subjectivos como valores 'altos', 'normais' e 'baixos' para aquelas variáveis. Se este for o caso, a abordagem correcta é a da lógica difusa [*fuzzy logic*], como se clarificará de seguida.

Como é sabido, a lógica difusa tornou-se muito popular devido ao trabalho seminal de Zadeh (1965). A expectativa deste autor foi que a lógica difusa, pelas suas características, seria bem acolhida pelas ciências sociais, como a Economia e a Ciência Política, quando na verdade não foi, com algumas excepções (Chang & Stekler, 1977; Deboeck, 1994; Mansur, 1995; Lindström, 1998); ao contrário das suas próprias expectativas, foi, de facto, extremamente bem acolhida nas ciências 'puras' (ver o prefácio de Clark et al., 2008).

Efectivamente, desde Zadeh (1965), a lógica difusa passou por um tremendo crescimento, quer no campo teórico quer no campo das aplicações. Dado que a principal preocupação da lógica difusa é aproximar modelos de raciocínio, do tipo que caracteriza muitas situações económicas e políticas, é surpreendente que tão poucas aplicações da lógica difusa tenham sido consideradas em Economia e muito menos em Ciência Política. Para ser mais preciso, a lógica difusa permite valores 'intermédios' entre os valores a serem definidos por uma avaliação convencional como sim/não ou verdadeiro/falso. A imprecisão ou subjectividade de conceitos que caracterizam a forma de pensar humana são, assim, facilmente tidas em conta pela lógica difusa (Barrett & Pattanaik, 1989).

Se se partir do princípio de que agentes, ou seja, neste caso, os eleitores, não possuem as capacidades requeridas para adquirir, reter e processar todas as informações necessárias para tomar decisões rígidas – por outras palavras, se a sua racionalidade for limitada – como, por exemplo, para classificar 'rigidamente' os valores das taxas de desemprego ou de inflação como sendo 'altos', 'normais' ou 'baixos', a fim de calcular a popularidade do Governo, a lógica difusa é uma forma natural de lidar com este tipo de situações em que a fonte de incerteza é a ausência de critérios 'rígidos' resultantes de estados de conhecimentos incompletos ou imperfeitos.

Na realidade, a percepção como fonte de avaliação subjectiva dos eleitores é uma questão crucial nestas matérias (Sanders, 2000;2003), mas, tanto quanto sabemos, uma abordagem de lógica difusa, como uma forma de modelização da subjectividade na avaliação dos eleitores, ainda não foi utilizada na determinação de um indicador de popularidade (esta metodologia usada na determinação da disponibilidade para o investimento por Lindström, 1998).

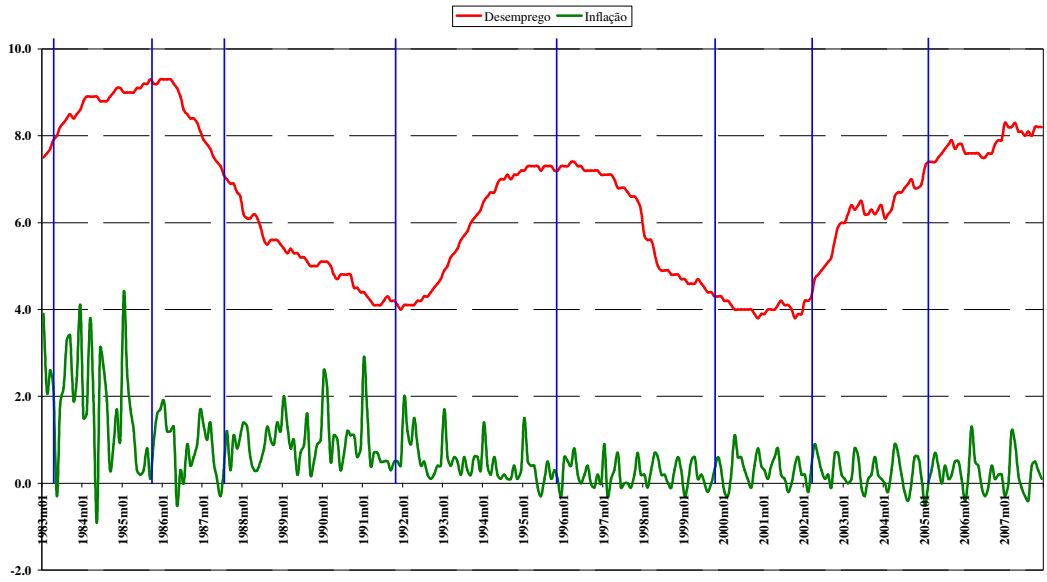
Partindo deste quadro, o objectivo neste trabalho é o cálculo de um indicador de popularidade em Portugal a partir de uma abordagem de lógica difusa. Tal será feito utilizando os dados mensais, a partir de 1983 até 2007, para o desemprego e inflação. Os resultados claramente explicam as vitórias e derrotas eleitorais do Governo durante as eleições que tiveram lugar no mesmo período. Além disso, o desemprego apresenta-se claramente mais importante que a inflação na explicação do nível de popularidade, o que está de acordo com o que a literatura, em geral, mostra.

Dito isso, o resto do documento está estruturado da seguinte forma. Na secção 2 serão apresentados os dados. A Secção 3 apresenta a metodologia da lógica difusa. Na secção 4, esta metodologia é aplicada aos dados, de forma a determinar o indicador de popularidade. A secção 5 termina este trabalho através da apresentação das principais conclusões assim como das vias de análise que poderão ser seguidas de forma a melhorar o presente trabalho.

2. DESCRIÇÃO DOS DADOS

Assumindo que a popularidade depende dos valores registados pelo desemprego e pela inflação, consideramos dados (mensais, desde 1983 até 2007) para Portugal relativos a estas duas variáveis, cujas fontes são o Eurostat (para a taxa de desemprego) e o Banco de Portugal (para a taxa de inflação). A sua representação gráfica apresenta-se na figura 1.

Figura 1: Os dados



Conforme a figura 1 mostra, a taxa de desemprego segue um claro padrão cíclico, apesar de ser igualmente perceptível que, no meio do período, o pico do ciclo não foi tão elevado. A taxa de inflação, por sua vez, foi mais elevada e mais volátil no início do período.[2]

No que diz respeito aos resultados eleitorais (as linhas verticais na figura 1 indicam a existência de uma eleição):

- Em Abril de 1983, teve lugar uma eleição em que o Partido Socialista [PS, doravante] obteve 36,11% dos votos e o Partido Social-Democrata [PSD, doravante] obteve 27,24% dos votos. Uma coligação entre estes dois partidos [o chamado Bloco Central] governou até as eleições de 1985.
- Em Outubro de 1985, realizou-se uma eleição, conduzindo a 29,87% dos votos para o PSD e 20,77% para o PS. Como o partido vencedor e o partido derrotado estiveram ambos no poder durante o mandato anterior, não consideramos esta eleição como representativa de uma vitória ou derrota eleitoral do Governo. Tal significa também que não iremos considerar esta eleição na determinação do indicador de popularidade do Governo, dado que ambos os partidos estavam no poder no mandato que precedeu esta eleição.
- Em Julho de 1987, teve lugar uma eleição que conduziu à reeleição do Governo, ou seja, o PSD com 50,22% dos votos, enquanto que o PS obteve 22,24%.
- Em Outubro de 1991, verificou-se outra reeleição do PSD, tendo este partido obtido 50,60% dos votos, enquanto que o PS obteve 29,13%.
- Em Outubro de 1995, o Governo foi derrotado pelo PS, o qual ganhou as eleições com 43,76% dos votos, enquanto que o PSD obteve 34,12%.
- Em Outubro de 1999, verificou-se uma reeleição do PS, tendo este partido obtido 44,06% dos votos, enquanto que o PSD obteve 32,32% dos votos.

[2] Em termos dos métodos tradicionais de regressão, esta característica dos dados poderia suscitar alguma preocupação. Não é o caso em questão dado que se irão utilizar médias móveis numa metodologia que comporta, de uma forma robusta, este género de dados.

- Em Março de 2002, o PS foi derrotado uma vez que obteve 37,79%, enquanto o que PSD obteve 40,21% dos votos. Este partido formou Governo com o Partido Popular [CDS/PP], o qual obteve 8,72%.
- Em Fevereiro de 2005, a coligação governante foi derrotada, tendo esta coligação obtido 36,11% dos votos (28,77% para o PSD e 7,24% para o CDS/PP), enquanto que o PS obteve 45,03%.

3. A METODOLOGIA DA LÓGICA DIFUSA

Considere-se um conjunto universal *U* e seja *A* um sub-conjunto de *U* em termos, digamos, clássicos, i.e. $A \subseteq U$. De acordo com a lógica clássica, digamos rígida [*crisp*] dos conjuntos, o grau com que um elemento de *U* pertence a *A* ou é 1 (o elemento de *U* está em *A*) ou é 0 (o elemento de *U* não está em *A*). Em termos formais, a *função característica* de *A* é tal que $\mu_A(x)=1$ para $x \in A$, e $\mu_A(x)=0$ para $x \notin A$. Assim, a função característica 'tradicional' discrimina, respectivamente, entre os membros e os não membros do conjunto 'rígido'. A generalização para um conjunto difuso faz-se através do relaxamento da separação estrita entre os elementos pertencentes ou não a *A*. Na lógica difusa a função característica pode, assim, assumir qualquer valor no intervalo [0,1].[3]

Os valores atribuídos pela função característica aos diversos elementos do conjunto indicam, assim, o grau de pertença dos diversos elementos no conjunto. Por exemplo, valores mais elevados indicam um maior grau de plausibilidade, ou seja uma maior correspondência com as características que descrevem o conjunto. Por conseguinte, utilizando a lógica difusa, pode modelizar-se o seguinte raciocínio: "o valor observado para a taxa de desemprego, digamos 6%, pode ser considerado 'alto', 'normal' ou 'baixo', com um certo grau de plausibilidade". Desta forma, a lógica difusa pode capturar o facto de os agentes económicos frequentemente avaliarem categorias como 'alto', 'normal' ou 'baixo' de maneira diferente, dada a sua inerente subjectividade.

Em termos da lógica difusa, taxas de inflação ou de desemprego 'altas', 'normais' ou 'baixas' são, assim, consideradas categorias subjectivas dado que os agentes geralmente avaliam esses conceitos de forma diversa. No que se segue, assumir-se-á, precisamente, que os eleitores usam este tipo de categorias subjectivas para determinar, de forma aproximada, o nível de popularidade do Governo. Em termos de lógica difusa, ter-se-á então um raciocínio aproximado ou qualitativo. Em geral, este facto corresponde à assunção de um mecanismo de inferência baseado em decisões do tipo '*se-então*' tal como se descreve de seguida:

P^1:	Se *X* for A^1, então *Y* é B^1, se não
…	
P^i:	Se *X* for A^i, então *Y* é B^i, se não
…	
P^n:	Se *X* for A^n, então *Y* é B^n.
Q: *X* é *A*'	*I*: *Y* é *B*'

O chamado *vector antecedente X* é constituído por variáveis linguísticas, possivelmente interligadas através de um operador E, no universo do discurso *U* enquanto o *vector consequente Y* é constituído por variáveis linguísticas no universo de discurso *V*. A partir do conjunto de regras *P* e da observação (difusa) *Q*: *X* é *A*' é então possível fazer a *inferência I*: *Y* é *B*'.

Em resumo, a estrutura de um sistema difuso, tal como será utilizado de seguida, pode ser apresentada da seguinte forma:

[3] No contexto de lógica difusa é habitual falar-se uma função de pertença [*membership function*] em vez de uma função característica embora estes conceitos sejam basicamente os mesmos.

1. *Fuzificação*:[4] Transformação de números (reais) em termos linguísticos, i.e. a transformação de *inputs* rígidos em difusos;
2. *Inferência*: Construção de regras difusas a partir das funções de pertença dos *inputs*/antecedentes e *outputs*/consequentes, seguida pela determinação de *outputs* termos linguísticos difusos;
3. *Desfuzificação*: Transformação dos termos linguísticos prévios em números reais, i.e. a transformação de *outputs* difusos em rígidos.

4. O INDICADOR DE POPULARIDADE

Usando a estrutura atrás indicada, esta secção estrutura-se da seguinte forma. Na subsecção 4.1 proceder-se-á ao processo de 'fuzificação' necessário para determinar o indicador de popularidade a partir das variáveis relevantes, i.e. desemprego e inflação, usando os dados atrás descritos. A subsecção 4.2 apresenta a inferência assumida pelos eleitores. A subsecção 4.3 apresenta o consequente processo de 'desfuzificação' conducente aos valores da popularidade.

4.1. A 'fuzificação'

De acordo com a, chamada, hipótese da responsabilidade, assuma-se que a popularidade do governo depende das taxas de desemprego e de inflação e que ambas as variáveis se podem caracterizar por três termos linguísticos de acordo com:

'taxa de desemprego' $\in$ {baixa, normal, alta}
'taxa de inflação' $\in$ {baixa, normal, alta}

A transposição das taxas de desemprego e de inflação, medidas em valores reais, naqueles três termos linguísticos será feita através do uso de funções de pertença associadas aos conjuntos difusos de valores 'baixos', 'normais' e 'altos'. No caso em questão, estas funções são determinadas através do uso das distribuições empíricas dos valores daquelas duas variáveis para Portugal durante o período 1983-2007.

No que diz respeito à forma das funções de pertença, serão consideradas funções de tipo Gaussiano por não excluírem, de uma forma *ad-hoc*, a possibilidade de as observações serem provenientes de todas as possíveis distribuições – valores 'baixos', 'normais', 'altos' – e por estarem de acordo com os dados observados. Para além disso, este tipo de funções de pertença está de acordo com o raciocínio que é assumido pelos eleitores aquando da determinação do indicador de popularidade. Assim é porque se assume que: (*i*) para a determinação dos graus de pertença de cada valor observado para as taxas de inflação e de desemprego àqueles três conjuntos difusos os eleitores consideram importantes quer a média quer a variância das observações que recordam – esta questão será clarificada de seguida – até ao momento *t* e (*ii*) em cada um daqueles conjuntos, quanto mais próximo for o valor observado para a inflação/desemprego em relação à média de cada conjunto em particular, maior deverá ser o grau de pertença a esse conjunto difuso, sendo a relação de natureza não linear.

Assim, cada função de pertença é do tipo:

$$f(x,\mu,\sigma)=\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \qquad (1)$$

[4] Embora se tenha traduzido o termo 'fuzzy' por 'difusa', tal como é comum na literatura, a tradução de 'fuzzification' ou 'defuzzification' recorrendo ao termo base 'difusa' não é de todo comum, sendo generalizado o uso dos termos 'fuzificação' e 'desfuzificação'. Por isso se empregam aqui também esses termos.

onde *x* representa o valor actual da taxa de inflação/desemprego, digamos π_t e u_t, e μ e σ são, respectivamente, medidas de localização e de dispersão nas observações da inflação/desemprego, digamos $\overline{\pi}_t, \overline{u}_t$ e $\sigma_t^{\pi}, \sigma_t^{u}$.

No que diz respeito às funções de pertença associadas às três categorias: 'baixa', 'normal', e 'alta' são determinadas, respectivamente, como:

$$\begin{matrix} f\left(\pi_t, \overline{\pi}_t - k\sigma_t^{\pi}, \sigma_t^{\pi}\right) \\ f\left(u_t, \overline{u}_t - k\sigma_t^{u}, \sigma_t^{u}\right) \end{matrix} \tag{2}$$

$$\begin{matrix} f\left(\pi_t, \overline{\pi}_t, \sigma_t^{\pi}\right) \\ f\left(u_t, \overline{u}_t, \sigma_t^{u}\right) \end{matrix} \tag{3}$$

$$\begin{matrix} f\left(\pi_t, \overline{\pi}_t + k\sigma_t^{\pi}, \sigma_t^{\pi}\right) \\ f\left(u_t, \overline{u}_t + k\sigma_t^{u}, \sigma_t^{u}\right) \end{matrix} \tag{4}$$

onde *f* é definido de acordo com a expressão (1). Os graus de pertença não normalizados para as observações correntes das taxas de inflação e desemprego, π_t e u_t, associados com as categorias 'baixa', 'normal' e 'alta', são, assim, dados por (2), (3), e (4), respectivamente. Note-se que $\overline{\pi}_t / \overline{u}_t$ representam a média actual do conjunto 'normal' da inflação/desemprego e que o parâmetro *k* serve para ajustar a separação entre os três conjuntos difusos, tornando-a proporcional à dispersão nos dados, medida por $\sigma_t^{\pi} / \sigma_t^{u}$. O valor de *k* será determinado como sendo o óptimo, i.e. aquele conduzindo ao melhor ajustamento entre o indicador de popularidade e os resultados eleitorais, enquanto os valores de $\overline{\pi}_t / \overline{u}_t$ e $\sigma_t^{\pi} / \sigma_t^{u}$ são determinados a partir das observações para a inflação e desemprego. Em particular, $\overline{\pi}_t / \overline{u}_t$ e $\sigma_t^{\pi} / \sigma_t^{u}$ são calculados, respectivamente, como médias e desvio-padrão móveis tal como se indica de seguida:

$$\begin{matrix} \overline{\pi}_t = \frac{\sum_{j=t-w+1}^{t} \pi_t}{w} \\ \overline{u}_t = \frac{\sum_{j=t-w+1}^{t} u_t}{w} \end{matrix} \tag{5}$$

$$\begin{matrix} \sigma_t^{\pi} = \sqrt{\frac{\sum_{j=t-w+1}^{t} \left(\pi_j - \overline{\pi}_t\right)^2}{w-1}} \\ \sigma_t^{u} = \sqrt{\frac{\sum_{j=t-w+1}^{t} \left(u_j - \overline{u}_t\right)^2}{w-1}} \end{matrix} \tag{6}$$

para $t \geq w$, onde *w* representa o número de observações usadas no cálculo das médias e desvios-padrão móveis. Dentro de uma grelha de valores plausíveis, o tamanho da janela móvel revelou-se quase inócuo em termos do resultado final. O valor de *w* escolhido foi então de 16 [veja-se Caleiro (2006) para uma confirmação desta escolha].

Seguindo a abordagem da ignorância racional dos eleitores tal como descrita em Downs (1957), a literatura tem vindo a apontar para o facto de os eleitores não serem totalmente racionais [para alguma análise desta questão veja-se, entre outros, Aidt (2000), Paldam & Nannestad (2000), e Sanders (2000)]. As expressões (5) e (6), de facto, capturam a assunção implícita da racionalidade limitada dos eleitores. Sargent (1993) argumenta que uma forma de substituir agentes totalmente racionais por outros possuindo racionalidade limitada é assumindo

que a memória dos agentes está limitada. Neste caso em particular, assume-se que a memória dos eleitores se associa ao tamanho da janela móvel, *w*, no cálculo das médias móveis (5) e dos desvios-padrão móveis (6), já que, por cada nova observação da inflação e do desemprego, os eleitores esquecem a observação mais remota no tempo.[5]

A tarefa seguinte consiste no cômputo dos graus de pertença para as categorias da inflação e do desemprego. Em termos mais claros, ir-se-á determinar um indicador de popularidade baseado nos valores da inflação e do desemprego, usando um processo de 'fuzificação', que o relaciona de forma óptima com os resultados eleitorais. Ao fazê-lo está a admitir-se que os eleitores formam as suas avaliações no momento *t*, num processo de 'fuzificação', através do mecanismo de inferência descrito acima, aplicado às *w* observações mais recentes paras taxas de inflação e de desemprego, i.e. para observações desde $t - w + 1$ até *t*, para os períodos $t \geq t_0 + w$ (sendo t_0, neste caso, Janeiro de 1983). Claramente, neste caso, tal significa um uso retrospectivo das taxas de inflação/desemprego por parte dos eleitores. Como é sabido, o comportamento retrospectivo dos eleitores é geralmente considerado como aquele que (mais) se verifica na realidade [veja-se, por exemplo, Lewis-Beck & Paldam (2000), para uma discussão de como este comportamento assumido é, de facto, o mais correcto.]

4.2. A inferência

Para estabelecer uma inferência assumiremos que o comportamento dos eleitores pode ser aproximado por regras simples de decisão do tipo '*se-então*' do género:
se ($u \in U$) e ($\pi \in \Pi$), então ($z \in Z$), onde *U*, Π, e *Z* são conjuntos difusos e *u*, π e *z* são números reais.

Já que existem três termos linguísticos quer para o desemprego quer para a inflação, existe um total de nove possíveis combinações e, assim, nove regras de votação serão consideradas. Assuma-se então que a predisposição para votar no (actual) Governo, *v*, pode ser descrita por três categorias de acordo com:

$v \in$ {forte, média, fraca}.

Dados estes três termos, as regras de decisão '*se-então*' podem ser definidas para cada par de categorias de desemprego e de inflação. Em termos julgados plausíveis, a predisposição para votar (no governo actual) decresce com o aumento do desemprego e/ou da inflação, de tal forma que as regras de votação são aquelas que constam na tabela 1.

Tabela 1: As regras de votação

Regra	Desemprego	Inflação	*v*	grau
1	baixo	baixo	forte	d_1
2	baixo	normal	forte	d_2
3	baixo	alto	normal	d_3
4	normal	baixo	forte	d_4
5	normal	normal	normal	d_5
6	normal	alto	fraco	d_6
7	alto	baixo	normal	d_7
8	alto	normal	fraco	d_8
9	alto	alto	fraco	d_9

[5] Isto naturalmente significa que ao longo do período amostral as médias e os desvios-padrão e, consequentemente, as funções de pertença, se vão alterando.

No que diz respeito aos graus, d_i, associados às categorias de predisposição para a votação, v, estes serão determinados de uma forma endógena, i.e. enquanto resultado de um programa de optimização. Neste caso, os valores de d_i serão aqueles conducentes ao um ajustamento máximo entre o indicador de popularidade e os resultados eleitorais (percentagem de votos obtida pelo partido no poder durante o mandato conducente àquela eleição).

O último passo consiste, então, na determinação dos graus de pertença associados aos termos linguísticos para a predisposição para a votação, v, tal como se apresentaram na tabela 1. De forma a clarificar este passo, julga-se esclarecedora uma pequena explicação teórica acerca de operadores difusos.[6] A presença do termo 'e' nas regras de decisão 'se-então' exige a definição de uma intersecção difusa. Um elemento de intersecção com um grau de pertença no conjunto A e um outro grau de pertença no conjunto B terá como grau de pertença o menor de entre aqueles valores. Por outras palavras a definição de E difuso (E-d) é a E-d $b = min\ (a,b)$.

A tabela 1 transforma os termos linguísticos para o desemprego e inflação em três termos caracterizando a predisposição para a votação. Como este termos resultaram da intersecção dos conjuntos difusos {baixo, normal, altao} temos que o grau de pertença de uma predisposição (linguística) para a votação será calculada como sendo o valor mínimo dos graus de pertença do desemprego e da inflação para cada regra de votação em particular ponderado pelos valores dos pesos d_i, os quais serão determinados de forma óptima. Em suma os valores de $d_{i\ =\ 1,\ldots,9}$ bem como k são aqueles que minimizam a soma de desvios (ao quadrado) entre os resultados eleitorais e os valores assumidos pelo indicador de popularidade no momento das eleições.

4.3. A 'desfuzificação'

O uso das regras de decisão difusas durante a fase de inferência produz *outputs* difusos, os quais devem ser transformados em *outputs* 'rígidos' de forma a serem úteis. No nosso caso, tal significa que as predisposições ponderadas para a votação, v, terão que ser 'desfuzificadas' em números reais para dar origem ao indicador de popularidade. Para tal existem vários métodos de 'desfuzificação'. Aqui será usado uma versão simples do método do máximo composto.[7] Tal como atrás, julga-se esclarecedora uma breve explicação sobre operadores difusos.

A partir das regras de votação atrás apresentadas podemos verificar que a mesma predisposição para votar, v, pode resultar – embora eventualmente associada a um grau diferente – de mais do que um par de categorias subjectivas para o desemprego e inflação. Tal significa que se deve utilizar o operador *OU difuso* (OU-d) definido como:

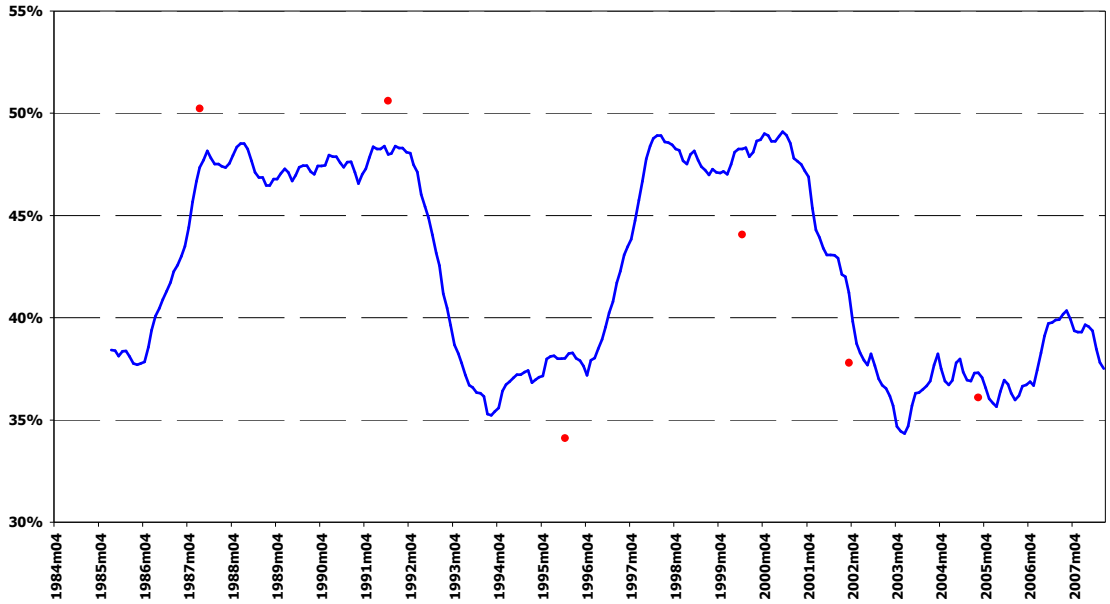
a OU-d $b = \max\ (a,b)$.

Assim, através do uso daquele operador difuso, o grau de pertença de um termo linguístico é o valor máximo dos graus de pertença associado àquele par. Estes graus de pertença (dos termos linguísticos 'forte', 'normal' e 'fraca') são então normalizados para que a sua soma seja igual a 1.

O passo final consiste na transposição dos termos linguísticos para a predisposição para votar, tal como se apresenta a tabela 1, no indicador de popularidade. Nesta subsecção a determinação do indicador de popularidade será feita assumindo que, após a normalização, aquele coincide com a predisposição 'normal' para votar. Considerando um valor de média móvel igual a $w = 16$, os resultados para todo o período podem ser visualizados na figura 2.[8]

[6] Veja-se, entre outros, Chen (1996) ou Zimmerman (1991).
[7] Outro método também muito utilizado é o método dos 'centróides', qual considera o centro de gravidade de todo o espaço difuso, sendo, por isso, sensível a todas as regras de decisão.
[8] Os resultados eleitorais do partido no poder estão representados na figura 2 por pontos.

Figura 2: O indicador de popularidade

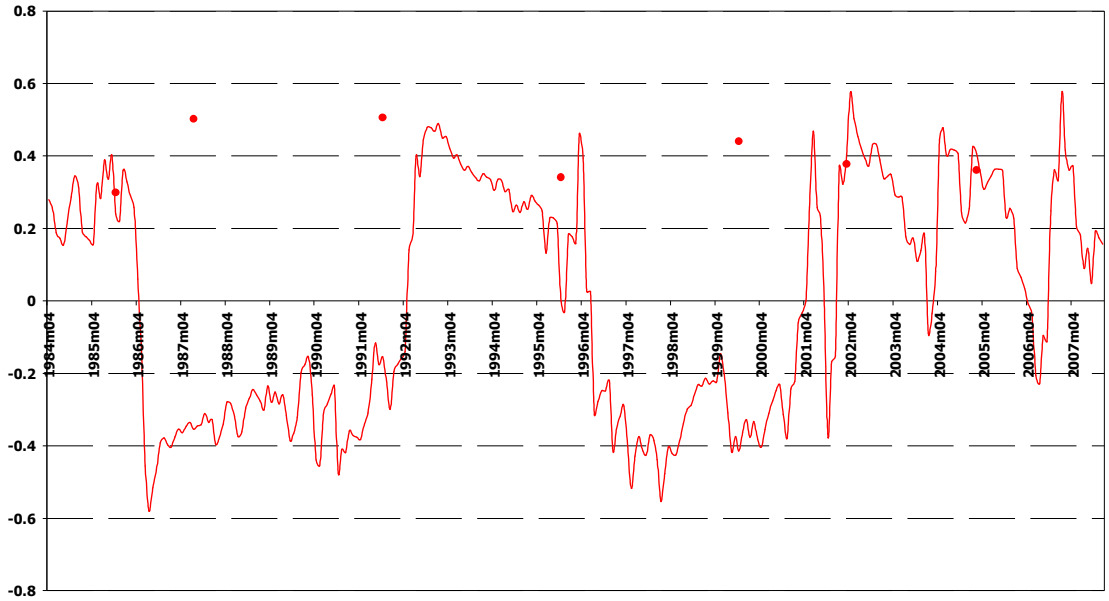


Claramente, o indicador de popularidade segue um padrão cíclico até Abril de 2003, data a partir da qual parece ter baixado para um nível sustentadamente mais baixo. Interessante é também verificar que as três reeleições (Julho de 1987, Outubro de 1991 e Outubro de 1999) se associaram a valores elevados do indicador de popularidade enquanto as derrotas eleitorais do partido no poder (Outubro de 1995, Março de 2002 e Fevereiro de 2005) se associaram a valores baixos do indicador de popularidade.

Para além de apontar o comportamento razoável do indicador de popularidade, gostaríamos de chamar a atenção para o seguinte facto. Verifica-se uma clara assimetria das preferências dos eleitores tal como está implícita nos graus d_i na tabela 1 em favor do desemprego. Os resultados foram $d_1 = 2{,}164$; $d_2 = 2{,}154$; $d_3 = 1{,}048$; $d_4 = 1{,}700$; $d_5 = 1{,}408$; $d_6 = d_8 = d_9 = 0$; $d_7 = 0{,}943$, o que significa também a categoria 'fraca' foi descartada. Se tivermos em conta a localização das eleições portuguesas no tempo parece ser evidente que as derrotas eleitorais tiveram lugar em momentos em que a trajectória do desemprego estaria a prejudicar a popularidade do partido no poder enquanto as vitórias eleitorais terão ocorrido quando a trajectória do desemprego estaria a beneficiar a popularidade do partido no poder. Na nossa opinião este resultado está bem de acordo com o facto estabelecido que os eleitores tendem a dar uma maior importância ao desemprego, tal como, aliás, a maior parte da literatura reconhece.

De forma a confirmar aquele facto, foi calculada a diferença entre os valores normalizados para os graus de pertença dos valores 'baixos' em relação aos 'altos', os quais se mostram na figura 3.

Figura 3: A importância do desemprego



Assumindo que os eleitores penalizam o Governo quando o desemprego é observado como pertencente a uma distribuição de valores 'altos' com um grau de pertença superior àquele respeitante a uma distribuição de valores 'elevados', ter-se-á que durante os períodos [1983m01,1986m05], [1992m07,1996m07], e [2001m04,…] a evolução da taxa de desemprego terá prejudicado a popularidade dos governos. De facto, todas as derrotas eleitorais dos governos tiveram lugar nestes períodos. De acordo com esta assunção, o caso oposto terá ocorrido durante os períodos [1986m05,1992m05] e [1996m07,2001m04]. De facto, as reeleições, i.e. as vitórias eleitorais dos governos tiveram lugar nestes períodos.[9]

5. CONCLUSÃO

Este trabalho mostrou como um indicador de popularidade pode ser determinado recorrendo a uma abordagem de lógica difusa. O uso desta abordagem foi justificado por ser esta a (mais) indicada para modelizar a situação – que é, de facto, a mais aceite na literatura – que as percepções económicas, provavelmente determinadas subjectivamente, dos eleitores influenciam, efectivamente, as intenções de voto ou a popularidade do Governo. Foi feita uma aplicação considerando o caso de Portugal (usando dados mensais, desde 1983 até 2007), para as taxas de inflação e de desemprego). Os resultados mostram que o indicador de popularidade ajuda a compreender os resultados eleitorais que se verificara em Portugal naquele período e, em particular, detecta claramente as vitórias e as derrotas eleitorais dos partidos no governo. Para além deste facto, foi realçada a muito maior importância do desemprego (relativamente à inflação) na determinação do nível de popularidade.

Claramente, a análise aqui considerada pode ser melhorada de diversas formas. Como, no que diz respeito aos resultados eleitorais do Governo, o desemprego se revelou substancialmente mais importante que a inflação, uma forma aparente de melhorar a análise será através da consideração de funções de pertença, tamanhos da janela, w, e parâmetro de separação, k, diferentes para aquelas duas variáveis. Além disso, sendo certo que estas duas variáveis são os 'dois-grandes' [*big*-2] nestas matérias, é possível conceber outras variáveis, mesmo económicas, que influenciem a popularidade e/ou outros mecanismos de transmissão dos efeitos das variáveis [por exemplo, Caleiro & Ramalho (2007) apontam para o facto de,

[9] No que diz respeito à inflação, a figura correspondente é tão pouco informativa que se torna inútil proceder à sua apresentação.

pelo menos em Portugal, a confiança dos consumidores – eventualmente outra variável importante para a popularidade – ser influenciada por circunstâncias políticas, nomeadamente, a ocorrência de eleições; Hardouvelis & Thomakos (2008) também exploram a ligação entre a confiança e as eleições na UE-15 durante 1985:1-2007:3, chamando a atenção para o aumento da confiança que se observa geralmente antes das eleições e o subsequente decréscimo após a data das eleições.].

REFERÊNCIAS BIBLIOGRÁFICAS

AIDT, TOKE S. (2000), "Economic voting and information", *Electoral Studies*, Vol. 19, pp. 349-362.

ALVAREZ, R. MICHAEL, JONATHAN NAGLER, & JENNIFER R. WILLETTE (2000), "Measuring the relative impact of issues and the economy in democratic elections", *Electoral Studies*, Vol. 19, pp. 237-253.

ANDERSON, CHRISTOPHER J. (2000), "Economic voting and political context: a comparative perspective", *Electoral Studies*, Vol. 19, pp. 151-170.

BARRETT, C. RICHARD, & PRASANTA K. PATTANAIK (1989), "Fuzzy Sets, Preference and Choice: Some Conceptual Issues", *Bulletin of Economic Research*, Vol. 41, nº 4, pp. 229-253.

BLOUNT, SIMON (2002), "Unemployment and economic voting", *Electoral Studies*, Vol. 21, pp. 91-100.

BYERS, DAVID, JAMES DAVIDSON, & DAVID PEEL (2000), "The dynamics of aggregate political popularity: evidence from eight countries", *Electoral Studies*, Vol. 19, pp. 49-62.

CALEIRO, ANTÓNIO (2006), "How is Confidence related to Unemployment in Portugal", *Applied Economics Letters*, Vol. 13, nº 3, pp. 887-890.

CALEIRO, ANTÓNIO, & ESMERALDA A. RAMALHO (2007), "Consumer Confidence in Portugal: what does it really matter?", *CEFAGE-UE Working Paper 2007/03* (disponível em http://www.cefage.uevora.pt/en/content/download/949/12218/version/1/file/200703.pdf).

CARLSEN, FREDRIK (2000), "Unemployment, inflation and government popularity: are there partisan effects?", *Electoral Studies*, Vol. 19, pp. 141-150.

CLARK, TERRY D., JENNIFER M. LARSON, JOHN N. MORDESON, JOSHUA D. POTTER, & MARK J. WIERMAN (2008), "Applying Fuzzy Mathematics to Formal Models in Comparative Politics", Springer-Verlag, Berlim.

CHANG, S.S.L., & H.O. STEKLER (1977), "Fuziness in Economic Systems, Its Modeling and Control", *Annals of Economic and Social Measurement*, Vol. 6, nº 2, pp. 165-174.

CHEN, C.H. (*ed.*) (1996), "Fuzzy Logic and Neural Network Handbook", New York: McGraw-Hill.

DEBOECK, GUIDO J. (*ed.*) (1994), "Trading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets", John Wiley & Sons, New York.

DOWNS, ANTHONY (1957), "An Economic Theory of Democracy", New York: Harper.

FREY, BRUNO S., & FRIEDRICH Schneider (1979), "An Econometric Model with an Endogenous Government Sector", *Public Choice*, Vol. 34, nº 1, pp. 29-43.

GLASGOW, GARRETT, & ROBERTO A. WEBER (2005), "Is there a relationship between election outcomes and perceptions of personal economic well-being? A test using post-election economic expectations", *Electoral Studies*, Vol. 24, pp. 581-601.

HARDOUVELIS, GIKAS A., & DIMITRIOS D. THOMAKOS (2008), "Consumer Confidence and Elections", *CEPR* Discussion Paper 6701, Fevereiro.

JOHNSTON, RICHARD (1999), "Business Cycles, Political Cycles and the Popularity of Canadian Governments, 1974-1998", *Canadian Journal of Political Science*, Vol. 32, pp. 499-520.

LEWIS-BECK, MICHAEL S., & MARTIN PALDAM (2000), "Economic voting: an introduction", *Electoral Studies*, Vol. 19, pp. 113-121.

LEWIS-BECK, MICHAEL S., & MARY STEGMAIER (2000), "Economic Determinants of Electoral Outcomes", *American Review of Political Science*, Vol. 3, pp. 183-219.

LINDSTRÖM, TOMAS (1998), "A Fuzzy Design of the Willingness to Invest in Sweden", *Journal of Economic Behavior and Organization*, Vol. 36, nº 1, pp. 1-17.

MANSUR, YUSUF M. (1995), "Fuzzy Sets and Economics: Applications of Fuzzy Mathematics to Non-cooperative Oligopoly", Edward Elgar, Aldershot.

NANNESTAD, PETER, & MARTIN PALDAM (1994), "The VP-Function: A survey of the literature on vote and popularity function after 25 years", *Public Choice*, Vol. 79, pp. 213-245.

NORDHAUS, WILLIAM (1975), "The political business cycle", *Review of Economic Studies*, Vol. 42, pp. 169-190.

PALDAM, MARTIN, & PETER NANNESTAD (2000), "What do voters know about the economy? A

study of Danish data, 1990-1993", *Electoral Studies*, Vol. 19, pp. 363-391.
SANDERS, DAVID (2000), "The real economy and the perceived economy in popularity functions: how much do voters need to know? A study of British data, 1974-97", *Electoral Studies*, Vol. 19, pp. 275-294.
SANDERS, DAVID (2003), "Party identification, economic perceptions, and voting in British General Elections, 1974-97", *Electoral Studies*, Vol. 22, pp. 239-263.
SANDERS, DAVID (2004), "Vote functions and popularity functions in British politics", *Electoral Studies*, Vol. 23, pp. 307-313.
SARGENT, THOMAS J. (1993), "Bounded Rationality in Macroeconomics", Clarendon Press, Oxford.
TELLIER, GENEVIÈVE (2006), "Effect of Economic Conditions on Government Popularity: The Canadian Provincial Case", *Canadian Journal of Political Science*, Vol. 39, nº 1, pp. 27-51.
VEIGA, LINDA GONÇALVES, & HENRY W. CHAPPELL, JR. (2002), "Politics and unemployment in industrialized democracies", *Public Choice*, Vol. 110, pp. 261-282.
YOUDE, JEREMY (2005), "Economics and government popularity in Ghana", *Electoral Studies*, Vol. 24, pp. 1-16.
ZADEH, L.A. (1965), "Fuzzy Sets", *Information and Control*, Vol. 8, pp. 338-353.
ZIMMERMANN, H.-J. (1991), "Fuzzy Set Theory - and Its Applications", Kluwer Academic Publishers, 2ª edição, Boston.

EXTERNALIDADES DE PRODUCCIÓN ENTRE LAS REGIONES ESPAÑOLAS

MIGUEL A. MÁRQUEZ PANIAGUA
E-mail: mmarquez@unex.es
JULIÁN RAMAJO HERNÁNDEZ
Department of Economics
University of Extremadura

Resumen

En este trabajo se detectan externalidades vinculadas al comercio interregional mediante un modelo dinámico espacio-temporal de corrección de error para datos de panel. De esta forma, se presenta prueba empírica acerca de la existencia de externalidades interregionales relacionadas con el comercio para las 17 regiones que componen el sistema económico regional español durante el período 1964-2003.

Área Temática: Métodos Cuantitativos para la Economía y Empresa.

Abstract

In this paper, trade-related externalities are measured be means a dynamic spatio-temporal panel error-correction model. This way, empirical evidence is presented about the existence of inter-regional externalities related to trade for the 17 regions integrating the Spanish regional economic system during the period 1964-2003.

Thematic Area: Quantitative Methods for Economics and Business.

1. INTRODUCCIÓN

Las interdependencias que existen entre las regiones deben ser consideradas a la hora de llevar a cabo un análisis de las políticas regionales. Para tener en cuenta dichas interdependencias, se deberían considerar los posibles canales diferentes de transmisión. Entre dichos canales, se encuentran los efectos espaciales, que han sido contemplados recientemente dentro del análisis empírico de los procesos de crecimiento regional (por ejemplo, ver Fingleton y López-Bazo, 2006 y las referencias ahí incluidas). Generalmente, la idea de interacción espacial se identifica con aspectos geográficos. No obstante, en un sentido amplio, "*interacción espacial se refiere a una amplia gama de flujos sujetos a diversos tipos de fricciones espaciales, tales como comercio de bienes, migraciones, movimientos de capital, convenios interregionales, remesas, y las transmisiones interregionales de conocimientos y efectos de los ciclos económicos*" (Behrens y Thisse, 2007, p. 460). Existe un amplio consenso en que las interconexiones entre las economías regionales provienen, probablemente, de la existencia de externalidades entre las regiones.

El reconocimiento de la importancia de las interacciones espaciales en los procesos de producción se ha visto estimulado por la dimensión espacial de las externalidades entre las economías. En este sentido, las teorías de crecimiento endógeno enfatizan los efectos de las externalidades (economías externas de escala) sobre la acumulación de los factores productivos, siendo las externalidades los motores del crecimiento endógeno. El papel importante jugado por las externalidades también ha sido destacado por la Nueva Economía Geográfica. Cuando las externalidades cruzan las barreras de las economías regionales (externalidades interregionales), estos spillovers podrían afectar al crecimiento económico regional. Sin embargo, las

externalidades no son directamente observables. Esto se debe a que no operan directamente sobre el crecimiento económico (Rigby y Essletzbichler (2002) y Duranton y Puga (2004)), siendo las expresiones de fuerzas como los enlaces input-output, efectos de innovación localizados o efecto de "labor pooling". El artículo seminal de Coe y Helpman (1995) puso de relieve que el comercio es un mecanismo de transmisión para tales spillovers. En este artículo se explorará en qué medida los patrones geográficos del comercio entre las economías regionales determinan la existencia de externalidades interregionales dentro de un sistema económico regional.

Desde una perspectiva metodológica, y con el objeto de permitir la identificación de forma indirecta de la posible existencia de externalidades interregionales, el trabajo de referencia del presente trabajo es Márquez et al (2009). Estos autores contrastaron la hipótesis de que los efectos espaciales derivados de la evolución a largo plazo de las regiones geográficamente adyacentes podría tener una influencia significativa en la respuesta a corto plazo del crecimiento regional. De esta forma, el enfoque geográfico puro adoptado por estos autores para las regiones peninsulares españolas es substituida por un enfoque de mercado, donde una medida ponderada de transmisión de externalidades basada en el comercio es considerada. La principal contribución de este artículo es posibilitar valorar si las externalidades de producción interregionales relacionadas con los patrones comerciales están operando dentro de un sistema económico regional.

Por otra parte, el análisis del crecimiento regional se ha basado, fundamentalmente, en datos de corte transversal o en datos de series temporales (Aghion y Durlauf, 2005). Hoy en día, los métodos de datos de panel permiten trabajar de manera más eficiente, puesto que combinan tanto información de series temporales como de corte transversal. Además, la falta de potencia de los contrastes de raíces unitarias tradicionales y de los contrastes de cointegración de las series temporales se ha visto mejorada por los contrastes de raíces unitarias para datos de panel y de los contrastes de cointegración del enfoque de datos de panel (Harris y Tzavalis, 1999). De esta forma, debido a la introducción de información de corte transversal en los datos, la existencia de relaciones de cointegración puede ser contrastada más apropiadamente.

En este trabajo, utilizando una función de producción agregada se estudia la distinción entre el crecimiento regional a largo y corto plazo dentro del enfoque proporcionado por un modelo de corrección de error basado en paneles de datos. El modelo es estimado para las 17 regiones españolas en el período 1964-2003. Para ello se utilizarán contrastes de raíces unitarias para datos de panel, contrastes de cointegración para paneles de datos y métodos de estimación recientemente desarrollados.

La hipótesis de este artículo es que los efectos espaciales derivados de la evolución de las regiones que mantienen flujos comerciales relevantes con una región podrían tener influencia significativa en la respuesta a corto plazo del crecimiento regional de dicha región. Contrastar esta hipótesis requeriría la reespecificación del modelo de correción de error tradicional en línea con la propuesta mostrada en Márquez et al (2009), introduciendo nuevos elementos espaciales determinados por las relaciones comerciales más relevantes. Así, la interacción espacial sería considerada a través de la incorporación de un retardo espacial-comercial de primer orden de las variables tradicionales del modelo de corrección de error. Como resultado, la introducción de un retardo espacial-comercial del término de error de equilibrio corregiría en el corto plazo desequilibrios en el largo plazo en la producción causados originados por desequilibrios en las regiones que mantienen flujos comerciales relevantes con la región analizada. La nueva especificación incorpora un retardo del término de corrección de error que se calcula en función del patrón geográfico comercial existente. Dicho nuevo término de corrección de error puede mejorar el conocimiento de los procesos de crecimientos regionales.

El resto del trabajo se organiza como se comenta a continuación. La sección segunda introduce la base teórica, describiendo cómo se especifican las externalidades. La sección tercera presenta la metodología empírica, estableciendo y discutiendo los métodos de estimación y los resultados obtenidos. Finalmente, en la sección cuarta se concluye.

2. MODELIZANDO EXTERNALIDADES: MARCO ANALÍTICO

El concepto de externalidades de producción entre las economías regionales se sitúa dentro del actual interés en el estudio acerca de cómo las externalidades interregionales afectan al crecimiento regional (por ejemplo, ver López-Bazo et al., 2004 y Egger y Pfaffermayr, 2006). Las externalidades han sido un tema muy tratado dentro de la reciente literatura sobre las teorías de crecimiento endógeno y la nueva economía geografica (ver Feser, 1998, para una revisión). La relevancia de las externalidades interregionales fue puesta de manifiesto por Kubo (1995), quién presentó un modelo de desarrollo regional que extendió el modelo desarrollado por Krugman (1981). El modelo de Kubo incorpora no sólo economías de escala dentro de cada región sino que, además, considera externalidades regionales entre las regiones. La principal contribución de este modelo es que las disparidades regionales pueden ser explicadas dentro de los argumentos teóricos que inciden en la existencia de externalidades interregionales. Consecuentemente, hay prueba teórica de que la literatura de crecimiento económico regional debería prestar atención al impacto de las externalidades interregionales.

Es habitual reconocer tres tipos de externalidades (Van Oort, 2007): externalidades de Marshall-Arrow-Romer -MAR- (especialización), externalidades de Jacobs (diversidad-competición) y externalidades de Porter (especialización-competición). Estas diferentes externalidades comparten el hecho de que son externalidades espaciales, en el sentido de que sus influencias sobre el proceso de creación y difusión del conocimiento tienen efectos geográficos. Sin embargo, en lo que se refiere a la prueba de la importancia de estas externalidades sobre el crecimiento regional, la diversidad y la competición ejercen efectos positivos sobre el crecimiento, mientras que no está clara cómo influye la especialización sobre el crecimiento (De Vor y De Groot, 2008). El nivel de agregación geográfica también influye en los resultados, dado que la magnitud de las externalidades incrementa cuando la agregación geográfica disminuye (Glaeser et al., 1992). En esta línea, con unidades geográficas mayores se suelen detectar más las externalidades MAR que las externalidades Jacobs (Beaudry y Schiffauerova, 2008). Cuando las externalidades entre las economías influyen en el proceso de producción regional, no es fácil adscribir los efectos a un tipo particular de externalidad. Según Duranton y Puga (2004), el crecimiento económico no recibe influencias directas de las externalidades, dado que las externalidades son el resultado de la combinación de fuerzas diferentes (como los enlaces input-output, efectos de *labor pooling* o efectos de innovación).

En cualquier caso, un objetivo fundamental sería contrastar la presencia de externalidades como uno de los determinantes del crecimiento a nivel agregado en un sistema regional, analizando cuál es el papel que juegan las externalidades interregionales en el crecimiento económico regional. En los trabajos empíricos de crecimiento, las externalidades de producción son conceptualizadas como modificadoras tecnológicas a través de los enfoques de la función de producción (Caballero y Lyons (1990, 1992)). Así, externalidades interregionales se especifican en la función de producción convencional (Feser, 2001) como:

$$Y_i = f(F_i)\, g(E_i)$$

donde Y_i representa el output in la region i, $f(F_i)$ symboliza los factores de producción internos para la región i, y $g(E_i)$ denota diferentes fuentes de efectos externos sobre la producción regional.

En el presente trabajo, la metodología también se basa en una función de producción agregada para un sistema económico regional. Por consiguiente, el punto de partida es asumir que el output regional puede ser descrito por una función de producción Cobb-Douglas. Sin embargo, en este caso, la función de producción sólo incorpora factores internos. Estos factores internos son el progreso tecnológico, el trabajo, el stock de capital humano, y el stock de capital físico (público y privado). El progreso tecnológico estimula mayores ratios de crecimiento regional. El crecimiento regional se vería favorecido mediante ratios de innovación elevados, transferencias internacionales de tecnología, y catching-up tecnológico internacional (Badinger y Tondl, 2003). El input del trabajo es el lado cuantitativo del capital humano. Los efectos derivados de cambios en la participación del trabajo influenciarían positivamente el crecimiento regional. De igual forma, mejoras en el capital humano de las personas empleadas también

afectarían positivamente al crecimiento económico regional. Aparte de los inputs del trabajo (cantidad y calidad), la acumulación de capital productivo es otra vía fundamental a la hora de favorecer el crecimiento regional. De este modo, el stock de capital productivo (público y privado) será incorporado en nuestra aplicación empírica.

En lo que concierne a la consideración de economías externas (externalidades interregionales) en el presente trabajo, dentro del modelo de crecimiento empírico que se propone se contemplan dos alternativas. En el largo plazo, externalidades interregionales serán detectadas a través de términos de corrección de error retardados y estadísticamente significativos, aunque esta forma de detección no proporcionará explicación de los orígenes de estas externalidades interregionales. En el corto plazo, elasticidades significativas asociadas con cada factor retardado mostrará prueba de la existencia de externalidades interregionales en el modelo de crecimiento. La significación del retardo de la variable dependiente indicaría la existencia de externalidades interregionales generadas, probablemente, desde el lado de la demanda.

3. METODOLOGÍA EMPÍRICA Y DATOS

Teniendo en cuenta los argumentos señalados en los apartados anteriores, en esta sección se lleva a cabo una aplicación empírica. Así, se analizará el crecimiento regional español para el período 1964-2003. Específicamente, el punto de partida es un modelo tipo Solow aumentado que produce la siguiente función de producción Cobb-Douglas para cada economía regional:

$$Y = AE^{\alpha} PK^{\beta} HK^{\gamma} \quad (1)$$

Donde Y es el nivel de output, E es el input del trabajo, PK es el stock de capital físico, HK es el stock de capital humano, y A es el progreso recnológico, que se puede considerar como el resultado de una actividad de innovación o de catching-up tecnológico.

3.1. Datos y propiedades estocásticas.

La base de datos utilizada consiste en series temporales anuales para las regiones españolas. España es un estado descentralizado compuesto de 17 regiones, además de Ceuta y Melilla (dos ciudades norteafricanas españolas),. Las regiones constituyen las denominadas Comunidades Autónomas. En este trabajo, el análisis utilizará sólo las 17 regiones españolas (ver figura 1), sin tener en cuenta las dos ciudades autónomas de Ceuta y Melilla.

Figura 1: Regiones Españolas



El panel de datos contiene información relativa a las 17 regiones españolas sobre el período 1964-2003 y para cada región, las variables utilizadas son las siguientes: Valor Añadido Bruto

real (VAB), *Y*, medido a precios básicos en miles de euros constantes del año 2000; empleo total (*E)*, en miles de personas empleadas; stock de capital total productivo (*PK*) (privado y público) en miles de euros constantes del año 2000; y stock de capital humano (*HK*), medido por el porcentaje de empleados totales con un nivel educativo alto (formación secundaria o grado universitario). Las series regionales para *Y* y *E* han sido tomdas de la base de datos BD.MORES (Bustos *et al*., 2008) y las series temporales para *PK* y *HK* han sido tomadas de la base de datos de la Fundación BBVA (Mas *et al*., 2007).

Como un paso previo al análisis de cointegración, las propiedades de integrabilidad de todas las variables fueron analizadas con el objetivo de determinar la existencia de raíces unitarias [I(1)] o estacionaridad [I(0)] en el comportamiento de las series temporales. En Breitung y Pesaran (2008) se puede encontrar una revisión reciente de la literatura sobre raíces unitarias y cointegración. Los resultados de los contrastes de raíces unitarias para datos de panel se presentan en la Tabla 2. Las filas de la Tabla proporcionan los resultados de dos contrastes utilizados de manera habitual en los contrastes utilizados en la literatura de raíces unitarias para datos de panel; concretamente, aquellos que permiten la presencia de heterogeneidad en el proceso de generación del panel de datos.

Tabla 1: Contrastes de raíces unitarias para $y = \log Y$, $e = \log E$, $pk = \log PK$ y $hk = \log HK$

	y	e	pk	hk
Im-Pesaran-Shin ADF	-0.28	5.72	-0.13	10.15
Maddala-Wu ADF-Fisher	42.95	6.41	27.24	1.93

NOTAS: Un * (**) [***] indica rechazo de la hipótesis nula (raíz unitaria) al 10% (5%) [1%] de nivel de significación.

Una visión general de los estadísticos presentados en la Tabla 1 muestra prueba que claramente favorece la hipótesis de que las cuatro variables básicas del modelo de producción (en logs) se comportan como variables no estacionarias con una raíz unitaria, al menos para una fracción muy importante de las 17 regiones que componen el panel. Por lo tanto, en lo que sigue, se considerará que los niveles de output, capital físico, capital humano e input del trabajo (todos en logaritmos) se comportan como procesos I(1).

3.2. Relación de producción a largo plazo

Para modelizar la dinámica de crecimiento en el corto plazo de las variables input-factores, primero se determina si existe una relación de producción en el largo plazo. Para contrastar la exitencia de cointegración en el panel de datos, se han llevado a cabo dos diferentes tipos de tests. Los contrastes de Pedroni (1999, 2004) se basan en los residuos estimados de la relación de cointegración analizada, mientras que el enfoque de Fisher propuesto por Maddala y Wu (1999) se aplica al análisis de múltiples ecuaciones de cointegración (CE). Ambos tipos de estadísticos permiten a los coeficientes de cada relación de cointegración variar líbremente para cada región, permitiendo heterogeneidad individual complea.

Las estimaciones de los diversos estadísticos de cointegración para paneles se presentan en la Tabla 2. Como valoración general de los valores presentados en la tabla, se puede deducir que existe una fuerte prueba indicando cointegración entre el nivel del output regional y los factores (inputs) del panel de 17 regiones. Así, en el caso de los estadísticos de Pedroni, ambos, los estadísticos Panel-ADF (versiones sin ponderar y ponderadas) y Group-ADF rechazan fuertemente la hipótesis de no cointegración. El estadístico Maddala-Wu también corrobora la existencia de una relación de cointegración estable. Por lo tanto, la prueba total es consistente con la hipótesis de existencia de una relación steady-state en el largo plazo entre las variables consideradas.

Tabla 2: Contrastes de cointegración de panel (Hipótesis Nula: No cointegración)

Test Cointegración	Estadístico
Kao ADF (Trend assumption: No deterministic trend)	-7.833***
Pedroni Panel-ADF (Trend assumption: Determ. intercept and trend)	-2.547*** (Unweighted) -2.961*** (Weighted)
Pedroni Group-ADF (Trend assumption: Determ. intercept and trend)	-2.192***
Maddala-Wu Fisher (Trend assumption: Determ. Trend. Number of CE: r=0)	221.9*** (Trace) 147.8*** (Max-eigen.)

Una vez detectada la función de producción en equilibrio en el largo plazo, el siguiente paso es estimar los parámetros de dicha función. Tomando en consideración la especificación de las funciones de producción regionales mostradas en (1), la relación de equilibrio en el largo plazo se tendrá en cuenta atendiendo a la expresión siguiente:

$$y_{it} = \beta_{0,i} + \delta_t + \beta_1 e_{it} + \beta_2 pk_{it} + \beta_3 hk_{it} + \mu_{it} \quad (2)$$

donde $y = \log Y$, $e = \log E$, $pk = \log PK$ y $hk = \log HK$.

La dinámica del output, y su ajuste al equilibrio de largo plazo entre las regiones vienen capturadas en los términso de error μ_{it}. Se asume homogeneidad en las pendientes, se permiten efectos de región fijos ($\beta_{0,i}$) con el objeto de controlar la heterogeneidad regional invariante, y se permite un efecto temporal común (δ_t) para tener en cuenta la posible presencia de dependencia de corte transversal en los términos de error de las diferentes ecuaciones (causados por la presencia de relaciones cruzadas entre las variables del modelo de producción regional o por la existencia de factores comunes inobservados dirigiendo el desarrollo económico en las diferentes regiones de España que podrían estar potencialmente correlacionadas con los regresores observados). Como resultado de la hipótesis de homogeneidad de pendientes asumida en la especificación anterior, la relación de cointegración estimada debe ser interpretada como una función de producción promedio en equilibrio para el panel de 17 regiones españolas.

Con respecto a la técnica elegida para estimar la relación de equilibrio, y dado que la estimación mediante Mínimos Cuadrados Ordinarios (OLS) del modelo de largo plazo sufriría de sesgo asintótico (Kao y Chiang, 2000), el método utilizado es el DSUR (Dynamic Seemingly Unrelated Regressions), propuesto por Mark *et al*. (2005). Este método permite la estimación simultánea eficiente de la relación de cointegración del panel con errores de desequilibrios correlacionados en las secciones cruzadas (cross-sections), trabajando con panel de datos en los que, como en este caso, la dimensión transversal (*N*) es mucho más pequeña que la dimensión temporal (*T*). Los resultados de la estimación "pooled" DSUR de la función de producción promedio en el largo plazo se presenta en la Tabla 4.

Tabla 4: DSUR estimación del modelo de largo plazo (2)

Variable	Coeficiente Estimado
Intercept	9.122*** (27.60)
e	0.461*** (16.11)
pk	0.241*** (9.31)
hk	0.078*** (4.85)
Número de observaciones = 629 (*N*=17, *T*=37)	

NOTA:1) Los estadísticos t de los parámetros están entre paréntesis; 2) Un * (**) [***] indica rechazo de la hipótesis nula de no significación al 10% (5%) [1%] de nivel de significación.

En términos de significación estadística, magnitud y plausibilidad teórica, las estimaciones DSUR obtenidas para las elasticidades-inputs de largo plazo son consistentes con las obtenidas en otros estudios con datos regionales españoles (ver, por ejemplo, De la Fuente y Vives, 1995; González-Páramo y Argimon, 1997; Dabán y Lamo, 1999; Gorostiaga, 1999; o Pablo-Romero y Gómez-Calero, 2008). Particularmente, nuestros resultados muestran un signo positivo y significativo para la variable de stock de capital humano *HK*, confirmando los resultados de la literatura previa sobre el papel positivo jugado por la fuerza de trabajo educada sobre el nivel de producción en el largo plazo.

3.3. Crecimiento regional en el corto plazo

Habiéndose establecido la cointegración del panel entre las variables y_{it}, e_{it}, pk_{it} y hk_{it}, la atención se dirige ahora a modelizar la dinámica de corto plazo del crecimiento regional español. Un modelo de corrección de error de panel standard de series temporales sería del tipo: [1]

$$\Delta y_{it} = \delta_{0,i} + \rho \Delta y_{i,t-1} + \delta_1 \Delta e_{it} + \delta_2 \Delta pk_{it} + \delta_3 \Delta hk_{it} + \phi \hat{\mu}_{i,t-1} + \nu_{it}$$

donde el crecimiento del VAB regional ($\Delta \log Y$) es explicado por su propio pasado ($(\Delta \log Y)_{-1}$), acumulación de factores ($\Delta \log E$, $\Delta \log PK$ y $\Delta \log HK$) y el *término de corrección de error temporal*, $\left(\log Y - \hat{\beta}_1 \log E - \hat{\beta}_2 \log PK - \hat{\beta}_3 \log HK\right)_{i,t-1}$. En lugar de formular este tipo de modelo, se especificó un modelo de retardo de primer orden de la primera diferencia del VAB, empleo, capital productivo, capital humano (todos en logs) y el témino de error de equilibrio, lo cual da la siguiente especificación:

$$\begin{aligned} \Delta y_{it} &= \delta_{0,i} + \rho_1 \Delta y_{i,t-1} + \rho_2 (W\Delta y)_{it} + \delta_1^1 \Delta e_{it} + \delta_1^2 (W\Delta e)_{it} + \delta_2^1 \Delta pk_{it} + \delta_2^2 (W\Delta pk)_{it} + \\ &+ \delta_3^1 \Delta hk_{it} + \delta_2^2 (W\Delta hk)_{it} + \phi_1 \hat{\mu}_{i,t-1} + \phi_2 (W\hat{\mu})_{i,t-1} + \nu_{it} \end{aligned} \quad (3)$$

Donde *W* es una matriz de pesos estandarizada por filas de dimensión $N \times N$ equivalente al número de regiones *N* consideradas (*N*=17 en nuestra aplicación) con elementos $w_{ij} = w_{ij}^* \Big/ \sum_j w_{ij}^*$, donde w_{ij}^* toma el valor de 1 si la i^{th} y la j^{th} regiones mantienen flujos comerciales relevantes en términos de share de mercado. El patrón geográfico de mercado fue aproximado mediante unos pesos de mercado fijos basados en los flujos comerciales promedios calculados para los años 1993, 1995 y 1997.

En este modelo espacio-temporal, el crecimiento del VAB de la región *i* es determinado por el crecimiento regional retardado temporalmente, el crecimiento promedio de las regiones vinculadas comercialmente a la región *i*, (efecto spillover de crecimiento), la acumulación de factores de la propia región y de las regiones vinculadas comercialmente con la región *i*, el término de corrección de error, y el retardo vinculado al comercio de este término de error de equilibrio.

La inclusión en la especificación dinámica local (3) del término de corrección de error retardado, $(W\hat{\mu})_{i,t-1}$ debe ser destacada. Este término fue introducido por primera vez en Márquez *et al*. (2006), quienes exploraron la posibilidad de que existieran dos efectos de externalidad espacial en el crecimiento regional español: un efecto nacional global y un efecto de vecindad local. dicho término representa el mecanismo de ajuste en la región *i* resultante de la existencia de un desequilibrio neto en el nivel de output de su vecindario (regiones vecinas). Este coeficiente de ajuste de corto plazo caracteriza la proporción por la que el desequilibrio en el largo plazo en la producción regional está siendo corregida en cada corto período por factores de vecindad. Mientras el término de correción de error ($\hat{\mu}_{i,t-1}$) corrige en cada corto período desequilibrios de largo plazo en el nivel de la producción regional causada por desequilibrios en

[1] Debido a restricciones de datos, se ha asumido que sólo un retardo de la variable dependiente y ningún retardo de las variables explicativas aparecen en esta especificación dinámica-temporal. A posteriori, esto fue una aproximación razonable del proceso de generación de datos.

su propia región, el término de correción de error espacial, $(W\hat{\mu})_{i,t-1}$, corrige en cada corto período desequilibrios en la producción en el largo plazo causados por desequilibrios en regiones que comparten patrones geográficos de comercio. Por lo tanto, un shock que ocurra en regiones vecinas (que produce un distanciamiento inesperado de la evolución estable de largo plazo de la producción en estos vecinos) podría transferir impulsos de crecimiento relevantes en el corto plazo a la región bajo análisis.

El modelo general anterior fue estimado y sobre dicha especificación se impusieron secuencialmente restricciones no significativas estadísticamente con el objeto de reducir las dimensiones de los parámetros vincualdos al espacio. Dado que la variable $W\Delta y$ está correlacionada con el término de error ν, el estimador OLS de (3) es inconsistente. Por lo tanto, se ha utilizado un estimador de variables instrumentales (IV), y $(W\Delta y)_{-1}$, X y WX [siendo X el vector $(\Delta e, \Delta pk, \Delta hk, \hat{\mu}_{-1})$] fueron utilizados como instrumentos. Para tener en cuenta la posible existencia de correlación espacial en el término de error, un procedimiento de Mínimos Cuadrados Generalizados espaciales factibles (FSGLS) fue usado para la estimación (Kelejian y Prucha, 1998). Los resultados para la especificación parsimoniosa final se muestran en la Tabla 5.

Tabla 5: FSGLS-IV estimación del modelo de corto plazo (3)

Variable	Coeficiente Estimado
Intercepto	-0.007*** (-3.36)
$(\Delta y)_{-1}$	0.105*** (3.03)
$W\Delta y$	0.369*** (6.34)
Δe	0.307*** (9.97)
Δpk	0.279*** (6.93)
Δhk	0.091*** (4.49)
$(\hat{\mu})_{-1}$	-0.111*** (-5.98)
$(W\hat{\mu})_{-1}$	-0.111** (-1.99)
$R^2 = 0.585$	Número de observaciones = 646 (N=17, T=38)

NOTE: 1) FSGLS-IV denota "Feasible Spatial Generalized Least Squares Instrumental Variable procedure"; 2) Los estadísticos t de los parámetros están entre paréntesis; 3) $\hat{\mu} = y - \hat{\beta}_1 e - \hat{\beta}_2 pk - \hat{\beta}_3 hk$ representa el error de desequilibrio de largo plazo estimado; 4) Un * (**) [***] indica el rechazo de la hipótesis nula al 10% (5%) [1%] de nivel de significación.

Explorando la perspectiva temporal, los resultados de la estimación indican que el crecimiento del VAB español está positiva y significativamente relacionado con el crecimiento de los tres input-factores considerados. El tamaño de los coeficientes indica una elasticidad output de corto plazo del 0.31 para el trabajo, 0.28 para el capital físico y 0.09 para el capital humano. Estos resultados son acordes con los valores que se encuentran en otros estudios de crecimiento regional.

El coeficiente estimado del retardo temporal del crecimiento del VAB también es positivo y significativo. También el coeficiente del término de corrección del error es estadísticamente significativo y tiene el signo negativo correcto, indicando que las desviaciones del VAB regional en torno a su tendencia a largo plazo están siendo corregidas en el tiempo.

Con respecto a los términos retardados en términos de los patrones geográficos comerciales, los resultados muestran la significación estadística del coeficiente del retardo de la variable dependiente ($W\Delta y$). Este resultado muestra la relevancia de las externalidades relacionadas con el comercio entre las regiones en el proceso de crecimiento. Este coeficiente mide la intensidad del efecto speillover interregional e indica que el ratio de crecimiento de una región está relacionado con las regiones con las que comparte un patrón geográfico de comercio. Desde otra perspectiva, este resultado significa que una realización no-cero de un shock en una región particular, teóricamente impactará no sólo esta región, sino que, a través de los patrones geográficos comerciales, a todas las regiones restantes (Anselin, 2003).

El crecimiento de las regiones españolas no es sensible al retardo espacio-comercial de los factores productivos, indicando que, en promedio, una región española no recibe efectos significativos de la evolución de los factores productivos de las regiones que comparten patrones geográficos comerciales con esta región.

La significación del retardo del VAB regional (variable dependiente) muestra prueba de la existencia de externalidades interregionales en el corto plazo.

Finalmente, el coeficiente negativo estimado para el término de corrección de error retardado también es estadísticamente significativo, implicando que desequilibrios en el largo plazo en el crecimiento de una región causado por desequilibrios en las regiones que comparten patrones geográficos comerciales con este región están siendo corregidos en cada corto período. La significación del término de corrección de error retardado muestra prueba empírica acerca de la existencia de externalidades interregionales relacionadas con el comercio operando en el largo plazo dentro del sistema económico regional español.

4. CONCLUSIONES FINALES

En este trabajo se muestra prueba empírica para el sistema regional económico español acerca de la existencia de externalidades interregionales relacionadas con el comercio. Bajo la hipótesis de que las externalidades entre regiones son una de las fuentes principales de dependencia espacial, el papel del comercio en el crecimiento regional es examinado mediante un nuevo enfoque. Tomando como referencia la metodología proporcionada por Márquez et al (2009), y partiendo de una función de producción agregada regional, la consideración de la teoría de la cointegración permite obtener una especificación empírica que destaca la dinámica (en el largo y corto plazo) del crecimiento económico de un sistema regional. El modelo es aplicado a las regiones españolas.

Las conclusiones obtenidas se pueden resumir como sigue. A través de la estimación del modelo empírico de crecimiento regional dinámico propuesto para las regiones españolas, se han detectado externalidades interregionales relacionadas con el comercio. En el largo plazo, estas externalidades interregionales se revelan mediante la significación estadística del término de corrección de error retardado atendiendo a los patrones geográficos comerciales más relevantes. Por otra parte, la obtención de elasticidades significativas asociadas con el retardo del VAB regional (variable dependiente) proporciona prueba de la existencia de externalidades interregionales relacionadas con el comercio en el corto plazo. Estas externalidades interregionales muestran que el ratio de crecimiento de una región está relacionado positivamente con la evolución de las regiones que comparten patrones geográficos comerciales con la región de referencia.

BIBLIOGRAFÍA

AGHION, P. y DURLAUF, S.N. (eds.) (2005): "*Handbook of Economic Growth*", North-Holland.

ANSELIN, L. (2003): "Spatial Externalities, Spatial Multipliers and Spatial Econometrics", *International Regional Science Review*, 26, 156-166.

BADINGER, H. y TONDL, G. (2003): "Trade, Human Capital and Innovation: The Engines of European Regional Growth in the 1990s", in Fingleton, B. (ed.), "*European Regional Growth*", Springer, Chapter 7, 215-239.

BEAUDRY, C. y SCHIFFAUEROVA, A. (2008): "Who's right, Marshall or Jacobs? The Localization versus Urbanization Debate", *Research Policy* (forthcoming).

BEHRENS, K. y THISSE, J.F. (2007): "Regional Economics: A New Economic Geography Perspective", Regional Science and Urban Economics, 37, pp. 457-465.

BREITUNG, J. y PESARAN, M.H. (2008): "Unit Roots and Cointegration in Panels", en Mátyás, L. y Sevestre, P. (eds.), "*The Econometrics of Panel Data*" (Third Edition), Springer, Chapter 9, 279-322.

BUSTOS, A., CUTANDA, A., DÍAZ A., ESCRIBÁ, F.J., MURGUI, M.J. y SANZ M.J. (2008): "La BD.MORES en base 2000: nuevas estimaciones y variables", *Documento de trabajo D-2008-02*, Dirección General de Presupuestos, Ministerio de economía y Hacienda.

CABALLERO, R. y LYONS, R. (1990): "Internal versus External Economies in European Industry", European Economic Review, 34, pp. 805-830.

CABALLERO, R. y LYONS, R. (1992): "External effects in U.S. Procyclical Productivity", Journal of Monetary Economics, 29, pp. 209-225.

COE, D.T. y HELPMAN, E. (1995): "International R&D Spillovers", European Economic Review, 39, pp. 859-887.

DABÁN, T. y LAMO, A. (1999): "Convergence and public investment allocation, Spain 1980-93", *Documento de Trabajo D-99001*, Dirección General de Análisis y Programación Presupuestaria, Ministerio de Economía y Hacienda.

DE VOR, F. y DE GROOT, H.L.F. (2008): "Agglomeration externalities and localized employment growth: the performance of industrial sites in Amsterdam", Annals of Regional Science (forthcoming).

DE LA FUENTE, A. y VIVES, X. (1995): "Infrastructure and Education as Instruments of Regional Policy: Evidence from Spain", *Economic Policy*, 20, 11-54.

DURANTON, G., y PUGA, D. (2004): "Micro-foundations of Urban Agglomeration Economies", in Henderson, J.V., Thisse, J.F. (eds.), "*Handbook of Regional and Urban Economics*", North-Holland, 2063-2117.

EGGER, P. y PFAFFERMAYR, M. (2006): "Spatial Convergence", *Papers in Regional Science*, 85, 199-215.

FESER, E. J. (1998): "Enterprise, External Economies and Economic Development", *Journal of Planning Literature*, 12, 283-302.

FESER, E.J. (2001): "Agglomeration, Enterprise Size, and Productivity", in Johansson, B., Karlsson, C.H. y STOUGH, R.R. (eds.), "*Theories of Endogenous Regional Growth*", Springer-Verlag, 231-251.

FINGLETON, B. y LÓPEZ-BAZO, E. (2006): "Empirical Growth Models with Spatial Effects", *Papers in Regional Science*, 85, 177-198.

GLAESER, E., KALLAL, H., SCHEINKMAN, J. y Shleifer, A. (1992): "Growth in Cities", *Journal of Political Economy*, 100, 1126-1152.

GONZÁLEZ-PÁRAMO, J. M. y ARGIMÓN, I. (1997): "Efectos de la inversión en infraestructuras sobre la productividad y la renta de las CC.AA", in Pérez E. (ed.), "*Infraestructuras y desarrollo regional: Efectos económicos de la Autopista del Atlántico"*, Editorial Civitas.

GOROSTIAGA, A. (1999): "¿Cómo afectan el capital público y el humano al crecimiento?: Un análisis para las regiones españolas en el marco neoclásico", *Investigaciones Económicas*, 23, 95-114.

HARRIS, R.D. y TZAVALIS, E. (1999): "Inference for Unit Roots in Dynamic Panels where the Time Dimension is Fixed", *Journal of Econometrics*, 91, 201-226.

KELEJIAN, H. y PRUCHA, I. (1998): "A generalized two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances", *Journal of Real State Finance and Economics*, 17, 99-121.

KRUGMAN, P (1981): "Trade, Accumulation, and Uneven Development", *Journal of Development Economics*, 8, pp.149-161.

KUBO, Y. (1995): "Scale Economies, Regional Externalities and the Possibility of Uneven Regional Development", *Journal of Regional Science*, 35, 318-328.

LÓPEZ-BAZO, E., VAYÁ, E. y ARTÍS, M. (2004): "Regional Externalities and Growth: Evidence from European Regions", Journal of Regional Science, 44, 1, pp. 43-73.

MADDALA, G.S. y WU, S. (1999): "A comparative study of unit root tests with panel data and a new simple test", *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

MARK, N.C., OGAKI, M. y SUL D. (2005): "Dynamic Seemingly Unrelated Cointegrating Regressions", *Review of Economic Studies*, 72, 797-820.

MÁRQUEZ, M.A., RAMAJO, J. y HEWINGS, G.J.D. (2006): "Dynamic effects within a regional system: an empirical approach", *Environment and Planning A*, 38, 711-732.

MÁRQUEZ, M.A., RAMAJO, J. y HEWINGS, G.J.D. (2009): "A Spatio-Temporal Econometric Model of Regional Growth in Spain", Mimeo.

MAS, M., PÉREZ, F. y URIEL, E. (2007): "*El stock y los servicios de capital en España y su distribución territorial (1955-1965). Nueva metodología*", Fundación BBVA.

PABLO-ROMERO, M.P., y GÓMEZ-CALERO, M.P. (2008): "Análisis por provincias de la contribución del capital humano a la producción en la década de los noventa", *Investigaciones Económicas*, 32, 27-52.

PEDRONI, P. (1999): "Critical values for cointegration tests in heterogeneous panels with multiple regressors", *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

PEDRONI, P. (2004): "Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis", *Econometric Theory*, 20, 597-625.

RIGBY,D.L. y ESSLETZBICHLER, J. (2002): "Agglomeration Economies and Productivity Differences in US Cities", Journal of Economic Geography, 2, pp. 407-432.

VAN OORT, F.G. (2007): "Spatial and sectoral composition effects of agglomeration economies in the Netherlands", *Papers in Regional Science*, 86, 5–30.

ESTACIONALIDAD DE PERIODO CAMBIANTE EN DATOS SEMANALES

GLORIA MARTÍN RODRÍGUEZ
E-mail: gmartinr@ull.es
JOSÉ JUAN CÁCERES HERNÁNDEZ
E-mail: jcaceres@ull.es
Departamento de Economía de las Instituciones, Estadística Económica y Econometría
Facultad de Ciencias Económicas y Empresariales
Universidad de La Laguna

Resumen

A seasonal model is proposed to deal with weekly seasonal patterns in which neither the length of the seasonal period nor the magnitude of the seasonal effects remain the same over time. To model this heterogeneous seasonal behaviour, the seasonal effect at a season is defined as a function of the proportion of the length of the seasonal period elapsed up to this season, and the seasonal pattern is modelled by means of evolving splines, defined in terms of the values of the seasonal effects at break points. The methodology is illustrated for a weekly series of tomato exports.

P*alabras clave:* weekly data, heterogeneous seasonal patterns, splines.
Área Temática: Métodos Cuantitativos para la Economía y Empresa.

1. INTRODUCCIÓN

En los últimos años ha crecido notablemente la disponibilidad de series semanales observadas durante largos periodos de tiempo. Este tipo de datos suele mostrar comportamientos estacionales caracterizados por su heterogeneidad, en tanto que la longitud del periodo de tiempo en que se completa la variación estacional no permanece constante a lo largo del tiempo. Puede pensarse, por ejemplo, en precios agrarios, que sólo se registran en las semanas en las que existe producción en el mercado, o también en producciones agrícolas, cuya campaña suele extenderse durante una parte del año que cambia en función de circunstancias climáticas o estrategias empresariales. En estos casos, la variación estacional que interesa modelar es aquélla que se observa durante un conjunto de semanas del año o, en general, de la campaña, y ese conjunto de semanas puede no estar formado siempre por las mismas semanas.

Los métodos convencionales no resultan apropiados si se desea considerar la complejidad de estos comportamientos estacionales. Por tanto, parece conveniente desarrollar modelos que conduzcan a formulaciones parsimoniosas, dado que el número de estaciones que define estos patrones periódicos es elevado, y, al propio tiempo, suficientemente flexibles para captar la heterogeneidad propia de estas variaciones. Desde este punto de vista, las funciones *splines* (Poirier 1976, Eubank 1988), en las que el efecto estacional en el instante del tiempo t se define como una función de la estación a la que corresponde el dato, constituyen una aproximación adecuada. En cualquier caso, la especificación apropiada de un modelo del componente estacional exige la extracción previa de la variación de largo plazo. Harvey y Koopman (1993) y Harvey *et al*. (1997) realizan esta descomposición a través de modelos estructurales de series temporales, y modelan la variación estacional recurriendo a funciones *splines* cúbicas periódicas estocásticas, a las que se exige que se ajusten perfectamente a los valores del componente estacional en estaciones concretas. Ferreira *et al*. (2000) también asumen que el componente estacional es una función alisada del tiempo y proponen especificar tendencia y variaciones estacionales a través de dos *splines* con diferentes parámetros de alisado.

Con ambas alternativas se pretende conseguir suficiente flexibilidad para modelar efectos estacionales cambiantes. Otros trabajos en la literatura especializada parecen inspirarse en la misma idea (Orbe *et al*. 2005, Pedregal y Young 2006). El inconveniente de tales procedimientos es que las restricciones que se imponen a los coeficientes introducidos en el modelo estacional no se derivan de los cambios observados en el patrón estacional.

Frente a aproximaciones que otorgan mayor flexibilidad en la variabilidad del proceso estocástico como mecanismo para ajustar el modelo a los datos, las funciones *splines* permiten que el modelo de las variaciones estacionales se adapte a los cambios observados en la magnitud y forma del patrón estacional a lo largo del periodo estacional. Esta aproximación trata de modelar explícitamente los cambios en la media del proceso estocástico causados por el carácter cambiante de las variaciones estacionales. De esta forma, la descripción y la predicción de este componente resultan más sencillas. La forma funcional polinómica definida a trozos y las restricciones de continuidad que suelen imponerse a las *splines* convencionales proporcionan a estos modelos un cierto grado de regularidad. Por otro lado, a través de funciones *splines* pueden definirse diferentes modelos estacionales para diferentes periodos de la muestra y, por lo tanto, se obtiene suficiente flexibilidad para recoger patrones estacionales heterogéneos.

La cuestión es que incluso estos modelos se encuentran limitados cuando, con relativa frecuencia, no existen observaciones correspondientes a algunas de las semanas del año, o bien, la longitud del periodo en el que se completa la variación estacional considerada no permanece constante a lo largo de los años. Además de un modelo capaz de recoger las variaciones estacionales, es necesario reflexionar sobre el concepto de estacionalidad. Por ejemplo, las semanas en las que comienza y termina el periodo del año en el que un producto agrícola está presente en los mercados internacionales suelen cambiar de un año a otro. Y podría asumirse que el comportamiento productivo es tal que la forma del patrón estacional esté asociada, no tanto a la semana del año en cuestión, sino más bien al número de semanas que han transcurrido desde que se inició la campaña de producción. En estas situaciones, el efecto estacional en un instante del tiempo depende de la proporción del periodo estacional completo que corresponde al subperiodo comprendido desde su inicio hasta la estación en cuestión. La formulación de los efectos estacionales como funciones de tal proporción implica que la periodicidad del patrón estacional se mide en una escala diferente a la habitual. Por supuesto, esta nueva aproximación sólo resulta útil cuando la longitud del periodo estacional no permanece constante a lo largo del tiempo. Si éste es el caso, la nueva formulación no significa una mayor complejidad matemática o estadística, pero proporciona mejoras significativas en flexibilidad y parsimonia. En este sentido, el modelo estacional propuesto se desarrolla en la sección siguiente. En la sección tercera, esta propuesta se aplica a una serie semanal de exportaciones de tomate con características estadísticas apropiadas para reflejar las ventajas del modelo. Por último, se exponen las conclusiones del trabajo.

2. FORMULACIÓN DEL COMPONENTE ESTACIONAL MEDIANTE UNA *SPLINE* CAMBIANTE

Sea $\{y_t\}_{t=1,...,T}$ una serie temporal registrada semanalmente tal que

$$y_t = \mu_t + \gamma_t + \varepsilon_t,\ t = 1,...,T, \tag{1}$$

donde μ_t y γ_t representan los componentes de nivel y estacional, mientras que ε_t es el componente irregular. Puede asumirse que el patrón estacional se caracteriza porque el número de semanas en las que se completa la variación estacional no permanece constante a lo largo del tiempo. Por supuesto, este supuesto implica admitir la posibilidad de que la variación estacional relevante se circunscriba a un conjunto de semanas que no necesariamente cubren un año completo. Aunque, especialmente en series agrarias, el periodo de la campaña agrícola puede no ser el año natural, el desarrollo metodológico que se expone a continuación parte del supuesto de que el conjunto de semanas que definen el periodo estacional completo está dentro del mismo año.

Suponga que la serie temporal se subdivide en tantos subperiodos de tiempo como periodos estacionales completos se registren. Es decir, cada subperiodo está formado por las semanas de un año en las que se observa una variación estacional completa. Suponga que existen m subperiodos de tiempo en cada uno de los cuales el periodo estacional se completa en s_c semanas, $c = 1,...,m$. En las observaciones del subperiodo c se define la variación estacional en el instante del tiempo t, γ_t, de forma que $\gamma_t = \gamma_{c,w}$, si la observación en ese instante del tiempo t corresponde a la semana j_c tal que $w = \frac{j_c}{s_c}$. Note que w es la fracción que indica la proporción del periodo estacional de s_c semanas que ha transcurrido en la semana j_c, $j_c = 1,...,s_c$, y, por tanto, $0 < w \leq 1$. Si se asume que la variación estacional cambia suavemente de una estación a la siguiente, una *spline* cúbica periódica puede ser un modelo adecuado de estas variaciones. Es decir,

$$\gamma_{c,w} = g_c(w) + \xi_{c,w}, \tag{2}$$

donde $\xi_{c,w}$ es un término residual y $g_c(w)$ es una función polinómica de grado tres definida a trozos,

$$g_c(w) = g_{i,c}(w) = g_{c,i,o} + g_{c,i,1}w + g_{c,i,2}w^2 + g_{c,i,3}w^3,\ w_{c,i-1} \leq w \leq w_{c,i},\ i = 1,...,k, \tag{3}$$

donde $w_{c,0} = 0$ y $w_{c,k} = 1$, a la que se imponen restricciones de continuidad de la propia función y de sus dos primeras derivadas. Estas restricciones se expresan a través de las condiciones siguientes:

$$g_{c,i,0} + g_{c,i,1}w_{c,i} + g_{c,i,2}w_{c,i}^2 + g_{c,i,3}w_{c,i}^3 = g_{c,i+1,0} + g_{c,i+1,1}w_{c,i} + g_{c,i+1,2}w_{c,i}^2 + g_{c,i+1,3}w_{c,i}^3, \tag{4a}$$

$$g_{c,i,1} + 2g_{c,i,2}w_{c,i} + 3g_{c,i,,3}w_{c,i}^2 = g_{c,i+1,1} + 2g_{c,i+1,2}w_{c,i} + 3g_{c,i+1,3}w_{c,i}^2 \tag{4b}$$

y

$$2g_{c,i,2} + 6g_{c,i,3}w_{c,i} = 2g_{c,i+1,2} + 6g_{c,i+1,3}w_{c,i}, \tag{4c}$$

para $i = 1,...,k\text{-}1$. Nótese que se está asumiendo que el número de tramos de la *spline* permanece constante para todos los subperiodos. De esta forma, la continuidad entre dos periodos estacionales completos consecutivos con el mismo patrón estacional, aunque posiblemente se completen en un número diferente de semanas, se garantiza a través de las siguientes condiciones

$$g_{c,k,0} + g_{c,k,1} + g_{c,k,2} + g_{c,k,3} = g_{c,1,0}, \tag{5a}$$

$$g_{c,k,1} + 2g_{c,k,2} + 3g_{c,k,3} = g_{c,1,1} \tag{5b}$$

y

$$2g_{c,k,2} + 6g_{c,k,3} = 2g_{c,1,2}. \tag{5c}$$

Debe notarse que la *spline* así definida es una función periódica de la proporción del periodo estacional, pero puede ser no periódica como función de la estación.

Dado que el número de parámetros desconocidos, $4k$, es mayor que el número de condiciones impuestas, $3k$, los parámetros del vector $G^c_{3kx1}=(\ g_{c,1,1},\ g_{c,1,2},\ g_{c,1,3},\ ...,\ g_{c,k,1},\ g_{c,k,2},\ g_{c,k,3}\ )'$ pueden expresarse como funciones de los parámetros del vector $G^{c,*}_{kx1} = (g_{c,1,0}, g_{c,2,0}, ..., g_{c,k,0})'$. En forma matricial, las restricciones de continuidad anteriores pueden expresarse como $T^c G^c = R^c$, donde

$$T^{c}_{3kx3k} = \begin{bmatrix} w_{c,1} & w_{c,1}^2 & w_{c,1}^3 & -w_{c,1} & -w_{c,1}^2 & -w_{c,1}^3 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & w_{c,k-1} & w_{c,k-1}^2 & w_{c,k-1}^3 & -w_{c,k-1} & -w_{c,k-1}^2 & -w_{c,k-1}^3 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 2w_{c,1} & 3w_{c,1}^2 & -1 & -2w_{c,1} & -3w_{c,1}^2 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 2w_{c,k-1} & 3w_{c,k-1}^2 & -1 & -2w_{c,k-1} & -3w_{c,k-1}^2 \\ -1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 2 & 6w_{c,1} & 0 & -2 & -6w_{c,1} & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 2 & 6w_{c,k-1} & 0 & -2 & -6w_{c,k-1} \\ 0 & -2 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 2 & 6 \end{bmatrix}$$

y

$$R^{c}_{3kx1} = \begin{bmatrix} -g_{c,1,0} + g_{c,2,0} \\ \vdots \\ -g_{c,k-1,0} + g_{c,k,0} \\ -g_{c,k,0} + g_{c,1,0} \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

Por lo tanto, $G^c = \left(T^c\right)^{-1} R^c$. Si la inversa de la matriz T^c se expresa como

$$\left(T^c\right)^{-1}{}_{3kx3k} = \begin{bmatrix} a_{c,1,1} & a_{c,1,2} & \cdots & a_{c,1,3k} \\ a_{c,2,1} & a_{c,2,2} & \cdots & a_{c,2,3k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{c,3k,1} & a_{c,3k,2} & \cdots & a_{c,3k,3k} \end{bmatrix}$$

y $g_{c,r}$, $r = 1, ..., 3k$, son parámetros tales que $g_{c,1} = g_{c,1,1}$, $g_{c,2} = g_{c,1,2}$, $g_{c,3} = g_{c,1,3}$, $g_{c,4} = g_{c,2,1}$, $g_{c,5} = g_{c,2,2}$, $g_{c,6} = g_{c,2,3}$, ..., $g_{c,3k\text{-}2} = g_{c,k,1}$, $g_{c,3k\text{-}1} = g_{c,k,2}$ y $g_{c,3k} = g_{c,k,3}$, entonces

$$g_{c,r} = \alpha_{c,r,1} g_{c,1,0} + \alpha_{c,r,2} g_{c,2,0} + ... + \alpha_{c,r,k-1} g_{c,k-1,0} + \alpha_{c,r,k} g_{c,k,0}, \quad (6)$$

donde $\alpha_{c,r,1} = -a_{c,r,1} + a_{c,r,k}$, $\alpha_{c,r,i} = a_{c,r,i-1} - a_{c,r,i}$, $i = 2,...,k$, para $r = 1, ..., 3k$. Ahora, la *spline* puede formularse como función de los parámetros $g_{c,i,0}$, $i = 1, ..., k$, como sigue,

$$g_c(w) = g_{c,1,0} X_{c,1,w} + g_{c,2,0} X_{c,2,w} + ... + g_{c,k-1,0} X_{c,k-1,w} + g_{c,k,0} X_{c,k,w}, \quad (7)$$

donde

$$X_{c,i,w} = D_{c,i,w} + \left(\alpha_{c,1,i} w + \alpha_{c,2,i} w^2 + \alpha_{c,3,i} w^3\right) D_{c,1,w} + \left(\alpha_{c,4,i} w + \alpha_{c,5,i} w^2 + \alpha_{c,6,i} w^3\right) D_{c,2,w} + ... + \left(\alpha_{c,3k-2,i} w + \alpha_{c,3k-1,i} w^2 + \alpha_{c,3k,i} w^3\right) D_{c,k,w} \quad (8)$$

para $i = 1,...,k$, $D_{c,i,w} = \begin{cases} 1, & w_{c,i-1} \le w < w_{c,i} \\ 0, & en\ otro\ caso \end{cases}$, $i = 1, ...,k\text{-}1$, y $D_{c,k,w} = \begin{cases} 1, & w_{c,k-1} \le w \le 1 \\ 0, & en\ otro\ caso \end{cases}$.

La especificación previa es suficientemente flexible para recoger una variación estacional γ_t que evolucione en el tiempo. El patrón estacional en los m subperiodos en los que se divide la serie puede modelarse conjuntamente como $\gamma_t = g(t) + \xi_t$, donde $g(t)$ es la *spline* cambiante

$$g(t) = \sum_{c=1}^{m} \left[g_{c,1,0} X_{1,t} + ... + g_{c,k,0} X_{k,t} \right] D_{c,t}^{sp}, \tag{9}$$

donde $D_{c,t}^{sp} = \begin{cases} 1, & t \in subperiodo\ c \\ 0, & en\ otro\ caso \end{cases}$, $c = 1,...,m$, y donde $X_{i,t} = X_{c,i,w}$, $i = 1,...,k$, si la observación en el instante del tiempo t corresponde a la semana j_c del subperiodo c de forma tal que $w = \frac{j_c}{s_c}$.

Una formulación alternativa permite expresar la variación estacional en cualquier estación como función de los valores del componente estacional en determinadas estaciones. Sean $\nabla^2 g_{c,i}(w_{c,i}) = a_{c,i}$, $i = 1,...,k$, los valores de la segunda derivada de la *spline* evaluados en los puntos de ruptura $w_{c,i}$, $i = 1,...,k$. Entonces, la continuidad de la segunda derivada de la *spline* se garantiza si se imponen las condiciones

$$\nabla^2 g_{c,i}(w_{c,i-1}) = a_{c,i-1}, \ i = 2,...,k, \tag{10.a}$$

y

$$\nabla^2 g_{c,k}(w_{c,k}) = a_{c,0}, \tag{10.b}$$

de forma que $a_{c,0} = a_{c,k}$. La segunda derivada de la *spline* es una función lineal

$$\nabla^2 g_{c,i}(w) = \left[(w_{c,i} - w)/(w_{c,i} - w_{c,i-1}) \right] a_{c,i-1} + \left[(w - w_{c,i-1})/(w_{c,i} - w_{c,i-1}) \right] a_{c,i}, \ i = 1,...,k. \tag{11}$$

Ahora, la *spline* puede obtenerse teniendo en cuenta que

$$\nabla g_{c,i}(w) = -\frac{1}{2}\left[(w_{c,i} - w)^2/(w_{c,i} - w_{c,i-1}) \right] a_{c,i-1} + \frac{1}{2}\left[(w - w_{c,i-1})^2/(w_{c,i} - w_{c,i-1}) \right] a_{c,i} + k_{c,1,i}, \tag{12}$$

para $i = 1,...,k$, de forma que

$$g_{c,i}(w) = \frac{1}{6}\left[(w_{c,i} - w)^3/(w_{c,i} - w_{c,i-1}) \right] a_{c,i-1} + \frac{1}{6}\left[(w - w_{c,i-1})^3/(w_{c,i} - w_{c,i-1}) \right] a_{c,i} + w k_{c,1,i} + k_{c,2,i} \tag{13}$$

para $i = 1,...,k$. Si se asume que la *spline* pasa por los puntos $\left(w_{c,i}, \gamma^+_{c,w_{c,i}} \right)$, $i = 0,...,k-1$, es decir, $g_{c,i}\left(w_{c,i-1} \right) = \gamma^+_{c,w_{c,i-1}}$ y $g_{c,i}\left(w_{c,i} \right) = \gamma^+_{c,w_{c,i}}$, entonces

$$k_{c,1,i} = \frac{(\gamma^+_{c,w_{c,i}} - \gamma^+_{c,w_{c,i-1}})}{(w_{c,i} - w_{c,i-1})} - \frac{1}{6}(w_{c,i} - w_{c,i-1})(a_{c,i} - a_{c,i-1}) \tag{14a}$$

y

$$k_{c,2,i} = \gamma^{+}_{c,w_{c,i}} - \frac{1}{6}(w_{c,i} - w_{c,i-1})^2 a_{c,i} - \frac{(\gamma^{+}_{c,w_{c,i}} - \gamma^{+}_{c,w_{c,i-1}})}{(w_{c,i} - w_{c,i-1})} w_{c,i} + \frac{1}{6}(w_{c,i} - w_{c,i-1})(a_{c,i} - a_{c,i-1}) w_{c,i} \quad (14b)$$

Por lo tanto,

$$g_{c,i}(w) = \frac{w_{c,i} - w}{6(w_{c,i} - w_{c,i-1})}\left[(w_{c,i} - w)^2 - (w_{c,i} - w_{c,i-1})^2\right] a_{c,i-1} + \frac{w - w_{c,i-1}}{6(w_{c,i} - w_{c,i-1})}\left[(w - w_{c,i-1})^2 - (w_{c,i} - w_{c,i-1})^2\right] a_{c,i} + \frac{w - w_{c,i-1}}{(w_{c,i} - w_{c,i-1})}\gamma^{+}_{c,w_{c,i}} + \frac{w_{c,i} - w}{(w_{c,i} - w_{c,i-1})}\gamma^{+}_{c,w_{c,i-1}} \quad (15)$$

para $i = 1,...,k$. En forma matricial, $G^c = W_A^c A^c + W_B^c \Gamma_c^+$, donde $G^c = (g_{c,1}(w), ..., g_{c,k}(w))'$, $A^c = (a_{c,0}, ..., a_{c,k-1})'$, $\Gamma_c^+ : \left(\gamma^{+}_{c,w_{c,0}}, ..., \gamma^{+}_{c,w_{c,k-1}}\right)$ y W_A^c y W_B^c son matrices definidas como

$$W_A^c = \begin{bmatrix} \frac{\delta_{c,1}}{6z_{c,1}}\left[\delta_{c,1}^2 - z_{c,1}^2\right] & \frac{-\delta_{c,0}}{6z_{c,1}}\left[\delta_{c,0}^2 - z_{c,1}^2\right] & 0 & 0 & \cdots & 0 \\ 0 & \frac{\delta_{c,2}}{6z_{c,2}}\left[\delta_{c,2}^2 - z_{c,2}^2\right] & \frac{-\delta_{c,1}}{6z_{c,2}}\left[\delta_{c,1}^2 - z_{c,2}^2\right] & 0 & \cdots & 0 \\ 0 & 0 & \frac{\delta_{c,3}}{6z_{c,3}}\left[\delta_{c,3}^2 - z_{c,3}^2\right] & \frac{-\delta_{c,2}}{6z_{c,3}}\left[\delta_{c,2}^2 - z_{c,3}^2\right] & \cdots & 0 \\ 0 & 0 & 0 & \frac{\delta_{c,4}}{6z_{c,4}}\left[\delta_{c,4}^2 - z_{c,4}^2\right] & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{-\delta_{c,k-1}}{6z_{c,k}}\left[\delta_{c,k-1}^2 - z_{c,k}^2\right] & 0 & 0 & 0 & \cdots & \frac{\delta_{c,k}}{6z_{c,k}}\left[\delta_{c,k}^2 - z_{c,k}^2\right] \end{bmatrix}$$

y

$$W_B^c = \begin{bmatrix} \frac{\delta_{c,1}}{z_{c,1}} & \frac{-\delta_{c,0}}{z_{c,1}} & 0 & 0 & \cdots & 0 \\ 0 & \frac{\delta_{c,2}}{z_{c,2}} & \frac{-\delta_{c,1}}{z_{c,2}} & 0 & \cdots & 0 \\ 0 & 0 & \frac{\delta_{c,3}}{z_{c,3}} & \frac{-\delta_{c,2}}{z_{c,3}} & \cdots & 0 \\ 0 & 0 & 0 & \frac{\delta_{c,4}}{z_{c,4}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{-\delta_{c,k-1}}{z_{c,k}} & 0 & 0 & 0 & \cdots & \frac{\delta_{c,k}}{z_{c,k}} \end{bmatrix}$$

con $z_{c,i} = w_{c,i} - w_{c,i-1}$, y $\delta_{c,i} = w_{c,i} - w$, $i = 0,...,k$.

Ahora, si se imponen condiciones de continuidad de la primera derivada de la *spline*, es decir,

$$\nabla g_{c,i}(w_{c,i}) = \nabla g_{c,i+1}(w_{c,i}), \; i = 1,...,k-1, \quad (16a)$$

$$\nabla g_{c,1}(w_{c,0}) = \nabla g_{c,k}(w_{c,k}), \quad (16b)$$

se obtienen k ecuaciones que permiten expresar los parámetros $a_{c,0},...,a_{c,k}$ como funciones de los parámetros $\gamma^{+}_{c,w_{c,o}},...,\gamma^{+}_{c,w_{c,k}}$. Teniendo en cuenta que $a_{c,0}=a_{c,k}$ y $\gamma^{+}_{c,w_{c,o}}=\gamma^{+}_{c,w_{c,k}}$, se obtiene que

$$\frac{w_{c,i}-w_{c,i-1}}{6}a_{c,i-1}+\frac{2\left[(w_{c,i}-w_{c,i-1})+(w_{c,i+1}-w_{c,i})\right]}{6}a_{c,i}+\frac{w_{c,i+1}-w_{c,i}}{6}a_{c,i+1}=$$
$$\frac{1}{w_{c,i}-w_{c,i-1}}\gamma^{+}_{c,w_{c,i-1}}-\left(\frac{1}{w_{c,i}-w_{c,i-1}}+\frac{1}{w_{c,i+1}-w_{c,i}}\right)\gamma^{+}_{c,w_{c,i}}+\frac{1}{w_{c,i+1}-w_{c,i-1}}\gamma^{+}_{w_{c,i+1}} \tag{17a}$$

para $i=1,...,k-1$, y

$$\frac{2\left[(w_{c,k}-w_{c,k-1})+(w_{c,1}-w_{c,0})\right]}{6}a_{c,0}+\frac{w_{c,1}-w_{c,0}}{6}a_{c,1}+\frac{w_{c,k}-w_{c,k-1}}{6}a_{c,k-1}=$$
$$-\left[\frac{1}{w_{c,k}-w_{c,k-1}}+\frac{1}{w_{c,1}-w_{c,0}}\right]\gamma^{+}_{c,w_{c,o}}+\frac{1}{w_{c,1}-w_{c,0}}\gamma^{+}_{c,w_{c,1}}+\frac{1}{w_{c,k}-w_{c,k-1}}\gamma^{+}_{c,w_{c,k-1}} \tag{17b}$$

En forma matricial, $W_1^c A^c = W_2^c \Gamma_c^+$, donde $A^c:\left(a_{c,0},...,a_{c,k-1}\right)$, $\Gamma_c^+:\left(\gamma^{+}_{c,w_{c,o}},...,\gamma^{+}_{c,w_{c,k-1}}\right)$,

$$W_1^c=\begin{bmatrix} \frac{z_{c,1}}{6} & \frac{z_{c,1}+z_{c,2}}{3} & \frac{z_{c,2}}{6} & 0 & \cdots & 0 \\ 0 & \frac{z_{c,2}}{6} & \frac{z_{c,2}+z_{c,3}}{3} & \frac{z_{c,3}}{6} & \cdots & 0 \\ 0 & 0 & \frac{z_{c,3}}{6} & \frac{z_{c,3}+z_{c,4}}{3} & \cdots & 0 \\ 0 & 0 & 0 & \frac{z_{c,4}}{6} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{z_{c,k}+z_{c,1}}{3} & \frac{z_{c,1}}{6} & 0 & 0 & \cdots & \frac{z_{c,k}}{6} \end{bmatrix}$$

y

$$W_2^c=\begin{bmatrix} \frac{1}{z_{c,1}} & -\left(\frac{1}{z_{c,1}}+\frac{1}{z_{c,2}}\right) & \frac{1}{z_{c,2}} & 0 & \cdots & 0 \\ 0 & \frac{1}{z_{c,2}} & -\left(\frac{1}{z_{c,2}}+\frac{1}{z_{c,3}}\right) & \frac{1}{z_{c,3}} & \cdots & 0 \\ 0 & 0 & \frac{1}{z_{c,3}} & -\left(\frac{1}{z_{c,3}}+\frac{1}{z_{c,4}}\right) & \cdots & 0 \\ 0 & 0 & 0 & \frac{1}{z_{c,4}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -\left(\frac{1}{z_{c,k}}+\frac{1}{z_{c,1}}\right) & \frac{1}{z_{c,1}} & 0 & 0 & \cdots & \frac{1}{z_{c,k}} \end{bmatrix}$$

con $z_{c,i}=w_{c,i}-w_{c,i-1}$, $i=1,...,k$. Por lo tanto, $A^c=\left(W_1^c\right)^{-1}W_2^c\Gamma_c^+$ y, dado que, $G^c=W_A^cA^c+W_B^c\Gamma_c^+$, se tiene que

$$G^c=W_A^cA^c+W_B^c\Gamma_c^+=W_A^c\left[\left(W_1^c\right)^{-1}W_2^c\right]\Gamma_c^++W_B^c\Gamma_c^+=W^c\Gamma_c^+. \tag{18}$$

Si se denota por $w_{c,p,q}$ al elemento de la matriz W^c en la fila p y la columna q, cada uno de estos elementos es una función de la fracción w, de forma que

$$\begin{aligned} g_c(w) = & \left[w_{c,1,1}D_{c,1,w} + w_{c,2,1}D_{c,2,w} + ... + w_{c,k,1}D_{c,k,w}\right]\gamma^+_{c,w_{c,0}} \\ & + \left[w_{c,1,2}D_{c,1,w} + w_{c,2,2}D_{c,2,w} + ... + w_{c,k,2}D_{c,k,w}\right]\gamma^+_{c,w_{c,1}} + ... \\ & + \left[w_{c,1,k}D_{c,1,w} + w_{c,2,k}D_{c,2,w} + ... + w_{c,k,k}D_{c,k,w}\right]\gamma^+_{c,w_{c,k-1}} \end{aligned} \quad (19)$$

donde $D_{c,i,w} = \begin{cases} 1, w_{c,i-1} \le w < w_{c,i} \\ 0, en\ otro\ caso \end{cases}$, $i = 1,...,k-1$, y $D_{c,k,w} = \begin{cases} 1, w_{c,k-1} \le w \le 1 \\ 0, en\ otro\ caso \end{cases}$. Así, la función *spline* puede expresarse como una función lineal

$$g_c(w) = \gamma^+_{c,w_0}X_{c,0,w} + ... + \gamma^+_{c,w_{c,k-1}}X_{c,k-1,w}. \quad (20)$$

Siguiendo a Koopman (1992) y Harvey *et al.* (1997), los valores de tal función pueden obtenerse usando los valores observados de alguna aproximación al patrón estacional, $\Gamma^+_c : \left(\gamma^+_{c,w_0}, ..., \gamma^+_{c,,w_{k-1}}\right)$. Sin embargo, en este trabajo se propone especificar el patrón estacional como $\gamma_{c,w} = g_c(w) + \xi_{c,w}$, donde la *spline* $g_c(w)$ se expresa como

$$g_c(w) = \gamma_{c,w_0}X_{c,0,w} + ... + \gamma_{c,w_{k-1}}X_{c,k-1,w}, \quad (21)$$

donde $X_{c,0,w},...,X_{c,k-1,w}$ son los regresores previamente definidos y $\gamma_{c,w_0},...,\gamma_{c,w_{k-1}}$ son parámetros libres que deben ser estimados.

La especificación previa es suficientemente flexible para capturar un patrón estacional en el que γ_t evoluciona a lo largo del tiempo. El patrón estacional en cada uno de los m subperiodos en los que se divide la serie temporal puede modelarse conjuntamente como $\gamma_t = g(t) + \xi_t$, donde $g(t)$ es la *spline* cambiante

$$g(t) = \sum_{c=1}^{m}\left[\gamma^c_{c,w_0}X_{c,0,t} + ... + \gamma^c_{c,w_{k-1}}X_{c,k-1,t}\right]D^{sp}_{c,t}, \quad (22)$$

donde $D^{sp}_{c,t} = \begin{cases} 1, & t \in subperiodo\ c \\ 0, & en\ otro\ caso \end{cases}$, $c = 1,...,m$, y $X_{i,t} = X_{c,i,w}$, $i = 0,...,k-1$, si la observación en el instante del tiempo t pertenece a la estación j_c del subperiodo c de forma tal que $w = \frac{j_c}{s_c}$. Entonces, si se asume que las localizaciones de los puntos de ruptura son constantes en cada uno de los m subperiodos y resulta que $\gamma_{c,w_{c,i}} = \gamma_{w_i}$, $i = 0,...,k-1$, $c = 1,...,m$, el modelo previo proporciona una importante ganancia en parsimonia con respecto a modelos convencionales en los que la variación estacional se define como una función de la estación. Y, por supuesto, también se produce una ganancia en parsimonia si la restricción anterior opera para un subconjunto de los m subperiodos en los que el periodo estacional no es siempre el mismo. Además, la especificación anterior permite la incorporación del componente estacional en un modelo conjunto de los diferentes componentes de la serie a través de la introducción de los regresores $X_{i,t}D^{sp}_{c,t}$.

Una cuestión que debe ser resuelta para obtener la especificación anterior es la selección del número y localizaciones de los puntos de ruptura. Para ello, es preciso obtener alguna aproximación previa del componente estacional. Cuando la forma del patrón estacional permanece constante, la ganancia en parsimonia es proporcional a la frecuencia con la que se producen cambios en la longitud del periodo estacional. E incluso cambios en las semanas concretas del año que definen el periodo estacional, aunque no se modifique el número de

semanas que lo componen, pueden constituir otro elemento favorable a la parsimonia de la especificación propuesta. Dado que los componentes tendencial y estacional son inobservables, la especificación de uno de ellos tiene consecuencias en la especificación del otro. Por lo tanto, el procedimiento para extraer una aproximación del componente estacional debería ser capaz de recoger sus características más relevantes desde el punto de vista de la formulación del modelo. En particular, la aproximación elegida debería mostrar los cambios en la forma del patrón estacional. En este sentido, un procedimiento adecuado puede basarse en el cálculo de medias móviles con un número de términos cambiante definido en función del número de semanas en las que se completa el periodo estacional al que pertenece el dato. Así, el número de puntos de ruptura puede elegirse a partir de la observación de las desviaciones $\hat{\gamma}_t^1$ entre la serie original y una serie de medias móviles de periodo s_c. Entonces, se ha tomado la decisión de elegir la combinación de localizaciones que minimiza la suma de cuadrados de los residuos en la estimación del modelo

$$\hat{\gamma}_t^1 = \gamma_{w_0} X_{0,t} + \gamma_{w_1} X_{1,t} + ... + \gamma_{w_{k-1}} X_{k-1,t} + \xi_t , \tag{23}$$

donde $X_{i,t} = X_{c,i,w}$, $i = 0,...,k-1$, si la observación en el instante del tiempo t corresponde a la estación j_c del subperiodo c en la que la proporción del periodo estacional es w. En resumen, el número de puntos de ruptura se elige de forma que la *spline* capture los principales cambios en la forma del patrón estacional observado, mientras que las localizaciones de estos puntos se seleccionan buscando el mejor ajuste posible.

3. UNA APLICACIÓN A UNA SERIE SEMANAL DE EXPORTACIONES DE TOMATE

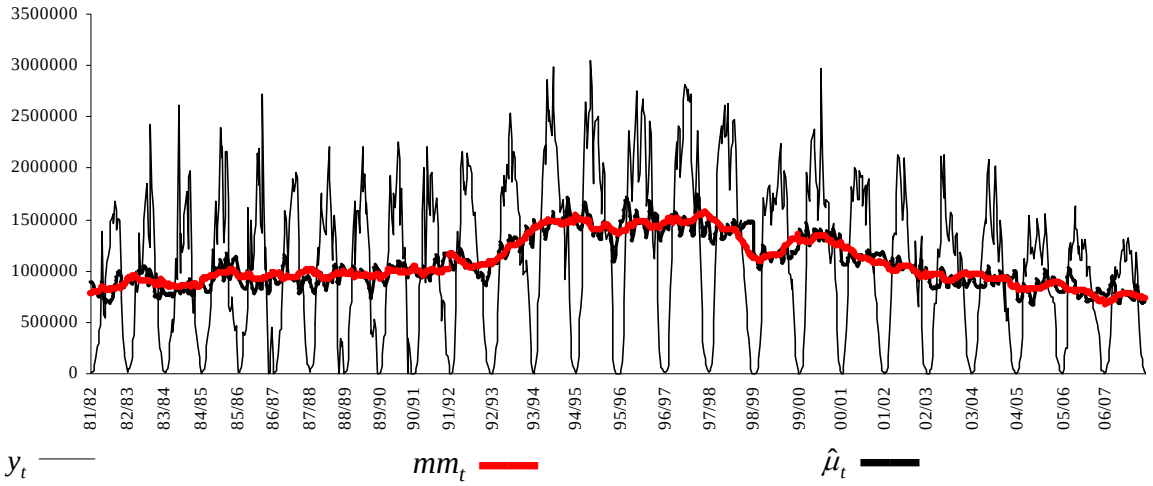
La propuesta metodológica planteada se aplica a continuación a una serie semanal de exportaciones de tomate de Canarias a los mercados europeos durante las campañas 1981/1982 a 2006/2007. Esta información, expresada en cestos de 6 kg, ha sido obtenida de los informes de campaña publicados por las asociaciones provinciales de cosecheros exportadores de tomate de Santa Cruz de Tenerife (ACETO) y Las Palmas de Gran Canaria (FEDEX).

La búsqueda de beneficios por parte de los exportadores canarios ha conducido a que la campaña exportadora se centre en el periodo de invierno, mientras que no suele existir exportación en los meses de verano, con mayor afluencia en los mercados europeos de la producción propia. Esta estrategia empresarial proporciona características estadísticas a la serie analizada de particular interés para mostrar la utilidad del método propuesto en la sección anterior. De hecho, la extensión de la campaña ha cambiado a lo largo del tiempo, en buena medida como resultado de los cambios en la normativa comercial de acceso a los mercados comunitarios. Durante la década de los 80, la zafra solía terminar a principios de mayo, cuando la confluencia con otras producciones reducía los precios del tomate canario hasta niveles que obligaban a pagar las tasas compensatorias resultantes de la aplicación del sistema de precios de referencia. La situación cambia cuando el sistema de precios de referencia se sustituye, primero, por un sistema de precios de oferta (Julio de 1991), y, finalmente, se elimina tras la incorporación plena de Canarias en la Política Agrícola Comunitaria (PAC) en enero de 1993. Así, desde la campaña 91/92 el periodo de exportación se prolonga hasta junio. Por otra parte, el comienzo de algunas campañas se ha adelantado con respecto a la semana habitual de inicio de zafra, que suele situarse en torno a la semana 40. Y, además de suprimir aquellas observaciones aisladas localizadas en semanas del verano en las que, de forma anómala, se ha registrado alguna exportación poco significativa, se ha optado por eliminar de la serie las exportaciones nulas durante dicho periodo estival.

Esta decisión implica que la variación estacional relevante es la que corresponde al periodo de exportación. De hecho, las exportaciones nulas en verano no proporcionan información más allá que la que se deriva de la extensión del periodo sin exportación, y esta información está implícitamente contemplada en la variación estacional que se modela. Así pues, se obtiene una

nueva serie que se denotará por $\{y_t\}_{t=1,...,942}$ (gráfico 1). Como se muestra en el cuadro 1, la longitud del periodo de exportación oscila entre 30 y 41 semanas. A pesar de este cambio en la longitud del periodo estacional, el patrón estacional se caracteriza por una cierta regularidad en su forma, que se corresponde con el propio ciclo productivo de la planta. Es decir, la exportación comienza con niveles bajos que van creciendo hasta alcanzar un máximo y a continuación se produce un descenso más o menos prolongado hasta que la exportación se interrumpe.

Gráfico 1. Exportaciones de tomate canario desde la campaña 1981/1982 a 2006/2007 (cestos de 6 kg)



Cuadro 1. Características de la serie de exportaciones

Campaña	Periodo(1)	Campaña	Periodo(1)
1981/1982	40-19 (32)	1994/1995	39-26 (40)
1982/1983	40-20 (33)	1995/1996	39-26 (40)
1983/1984	40-20 (33)	1996/1997	39-26 (40)
1984/1985	40-20 (33)	1997/1998	39-26 (40)
1985/1986	41-20 (32)	1998/1999	37-25 (41)
1986/1987	41-20 (32)	1999/2000	39-23 (37)
1987/1988	41-20 (32)	2000/2001	39-24 (38)
1988/1989	42-20 (31)	2001/2002	37-22 (38)
1989/1990	42-19 (30)	2002/2003	36-24 (41)
1990/1991	41-20 (32)	2003/2004	39-26 (40)
1991/1992	40-25 (38)	2004/2005	38-26 (41)
1992/1993	40-24 (37)	2005/2006	40-25 (38)
1993/1994	40-24 (37)	2006/2007	40-23 (36)

(1) El periodo estacional se identifica indicando la semana del año en la que se inicia y la semana del año siguiente en la que concluye. Entre paréntesis se señala el número de semanas en las que se completa la variación estacional.

Con respecto al movimiento de largo plazo, una aproximación de éste puede obtenerse a través del cálculo de medias móviles de periodo s_c en las observaciones que pertenecen a un periodo estacional de longitud s_c. Para el cálculo de estas medias, se han utilizado las observaciones necesarias de las campañas 1980/1981 y 2007/2008 de forma que pudiera obtenerse la aproximación de este movimiento de largo plazo para cada una de las observaciones de la serie de exportaciones considerada en el análisis. La serie de medias móviles así obtenida, $\{mm_t\}_{t=1,...,942}$, se representa también en el gráfico 1. La característica

más destacable de esta serie es que permite observar fases bastante claras de estancamiento, crecimiento y decrecimiento del volumen exportado. En concreto, la exportación experimentó un notable despegue durante la primera mitad de la década de 1990, que sin duda tuvo relación con las expectativas generadas tras la integración ya comentada en la PAC. Sin embargo, este crecimiento se interrumpe en 1996, coincidiendo con la aplicación de un importante acuerdo comercial entre la Unión Europea y Marruecos (Reg (CE) 747/2001; Reg (CE) 37/2004; Reg (CE) 503/2005), que facilitaba la entrada de la producción magrebí al mercado europeo. Estas inestabilidades en el componente de largo plazo aconsejan utilizar los modelos estructurales de series temporales (Harvey, 1989) como una herramienta estadística apropiada.

Como ya se comentó en la sección metodológica, la especificación propuesta para la variación estacional puede incorporarse en el modelo estructural. Pero la especificación de la *spline* estacional requiere obtener una aproximación previa de las variaciones estacionales que se pretende modelar a través de esta función. Así, el número y localización de los puntos de ruptura se decidió a partir de la observación de la serie de desviaciones, $\left\{\hat{\gamma}_t^1\right\}_{t=1,...,942}$, entre la serie original y la serie de medias móviles antes comentada. Y con objeto de obtener una formulación más parsimoniosa, se asumió que los puntos de ruptura son los mismos para todas las campañas. Se asume, además, que dichos puntos pertenecen al conjunto $\left\{\frac{l}{1000}\right\}_{l=1,...,1000}$, que divide el intervalo continuo $(0,1)$ en subintervalos suficientemente cortos. A partir de la estimación del modelo de regresión de la ecuación (23), se optó por una *spline* de 6 tramos, y los puntos de ruptura se localizan en las proporciones del periodo estacional $w_1 = 0.35$, $w_2 = 0.49$, $w_3 = 0.50$, $w_4 = 0.57$ y $w_5 = 0.9$.

De acuerdo a los resultados de la estimación del modelo paramétrico

$$\hat{\gamma}_t^1 = \sum_{c=1}^{26}\left[\gamma_{c,w_0} X_{0,t} + ... + \gamma_{c,w_5} X_{5,t}\right] D_{c,t}^{sp} + \xi_t, \tag{24}$$

se obtiene una nueva aproximación de la variación estacional definida como

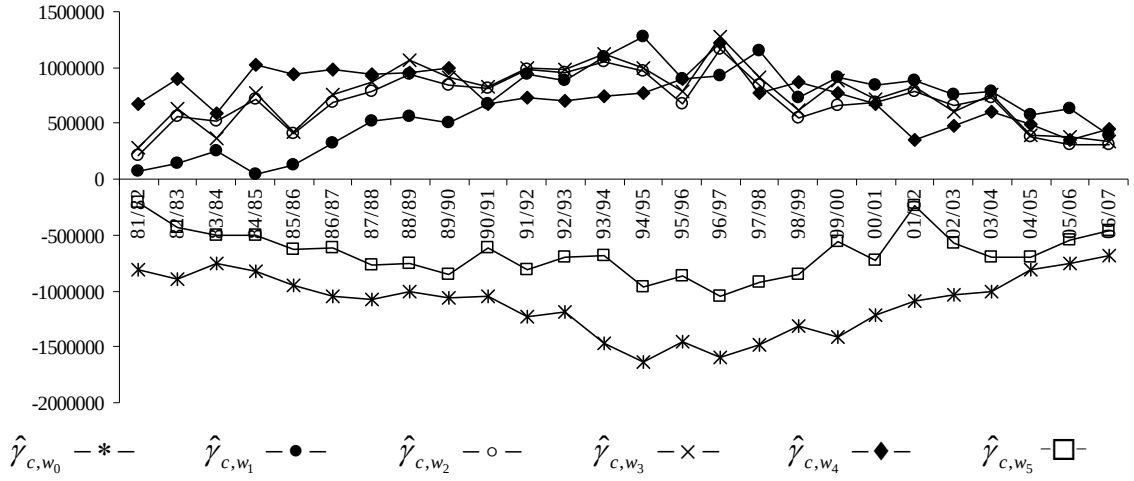
$$\sum_{c=1}^{26}\left[\hat{\gamma}_{c,w_0} X_{0,t} + ... + \hat{\gamma}_{c,w_5} X_{5,t}\right] D_{c,t}^{sp}. \tag{25}$$

Y dado que, como se explicó en la sección metodológica, la aproximación anterior admite una expresión equivalente en términos de los parámetros $g_{c,i,0}, g_{c,i,1}, g_{c,i,2}, g_{c,i,3}$, $i = 1,...,6$ (ecuación (3)), que definen la función polinómica de grado tres en cada tramo de la *spline* definida para cada campaña, es posible obtener una estimación corregida de las variaciones estacionales de modo que el área encerrada bajo la *spline* sea igual a cero para cada campaña exportadora. Es decir, imponiendo la condición

$$\sum_{i=1}^{6}\int_{w_{i-1}}^{w_i}\left(\hat{g}_{c,i,0} + \hat{g}_{c,i,1} w + \hat{g}_{c,i,2} w^2 + \hat{g}_{c,i,3} w^3\right) dw = 0, \tag{26}$$

se obtienen nuevas estimaciones $\left\{\hat{\gamma}_t^2\right\}_{t=1,...,942}$, que, a su vez, pueden expresarse en términos de adecuados valores de los parámetros $\gamma_{c,w_0},...,\gamma_{c,w_5}$ para las diferentes campañas. La evolución de estos parámetros se muestra en el gráfico 2 y es un reflejo de la evolución experimentada por el patrón estacional en su conjunto, puesto que las variaciones estacionales en los puntos de ruptura son los valores que explican las variaciones estacionales en cualquier otra fracción del periodo estacional.

Gráfico 2. Aproximación previa de las variaciones estacionales en puntos de ruptura, $\hat{\gamma}_t^2$.



Las estimaciones previas de la variación estacional no muestran por tanto un patrón común para todas las campañas de la muestra. De hecho, como parece lógico, las variaciones estacionales en los puntos de ruptura intermedios, que corresponden a las semanas de máxima exportación, adquieren valores mayores en las campañas de la primera década de los 90, mientras que los valores negativos de mayor magnitud se registran también durante estas campañas para las variaciones estacionales correspondientes a los extremos de la campaña, en la que se localizan las exportaciones menores. En definitiva, la amplitud de las variaciones estacionales es más acentuada en las campañas de mayor volumen de exportación. Ahora bien, el patrón estacional hasta principios de los 90 parece bastante regular. De hecho, una vez que se estima el modelo paramétrico

$$\hat{\gamma}_t^2 = \sum_{c=1}^{26} \left[\gamma_{c,w_0} X_{0,t} + ... + \gamma_{c,w_5} X_{5,t}\right] D_{c,t}^{sp} + \xi_t , \tag{27}$$

la aplicación de tests F convencionales conduce a la conclusión de que algunos patrones estacionales son comunes. En concreto, puede asumirse que el patrón estacional es homogéneo en las campañas 81/82 y 82/83, 84/85 a 86/87, 87/88 a 89/90, 90/91 a 92/93, 93/94 y 94/95, y, finalmente, 04/05 y 05/06.

Estas hipótesis sobre el componente estacional pueden introducirse en un modelo estructural formulado como

$$\begin{aligned} y_t = \mu_t &+ \left[\gamma_{w_0}^{1,2} X_{0,t} + ... + \gamma_{w_5}^{1,2} X_{5,t}\right]\left(D_{1,t}^{sp} + D_{2,t}^{sp}\right) \\ &+ \left[\gamma_{w_0}^{4,5,6} X_{0,t} + ... + \gamma_{w_5}^{4,5,6} X_{5,t}\right]\left(D_{4,t}^{sp} + D_{5,t}^{sp} + D_{6,t}^{sp}\right) \\ &+ \left[\gamma_{w_0}^{7,8,9} X_{0,t} + ... + \gamma_{w_5}^{7,8,9} X_{5,t}\right]\left(D_{7,t}^{sp} + D_{8,t}^{sp} + D_{9,t}^{sp}\right) \\ &+ \left[\gamma_{w_0}^{10,11,12} X_{0,t} + ... + \gamma_{w_5}^{10,11,12} X_{5,t}\right]\left(D_{10,t}^{sp} + D_{11,t}^{sp} + D_{12,t}^{sp}\right) \\ &+ \left[\gamma_{w_0}^{13,14} X_{0,t} + ... + \gamma_{w_5}^{13,14} X_{5,t}\right]\left(D_{13,t}^{sp} + D_{14,t}^{sp}\right) \\ &+ \left[\gamma_{w_0}^{24,25} X_{0,t} + ... + \gamma_{w_5}^{24,25} X_{5,t}\right]\left(D_{24,t}^{sp} + D_{25,t}^{sp}\right) \\ &+ \sum_{c=3,15,...,23} \left[\gamma_{w_0}^{c} X_{0,t} + ... + \gamma_{w_5}^{c} X_{5,t}\right] D_{c,t}^{sp} \\ &+ \left[\gamma_{w_0}^{26} X_{0,t} + ... + \gamma_{w_4}^{26} X_{4,t}\right] D_{26,t}^{sp} \\ &+ \lambda_{1,1984} I_{1,1984} + \lambda_{11,1986} I_{1,1986} + \lambda_{52,1990} I_{52,1990} + \varepsilon_t \end{aligned} \tag{28}$$

donde el componente de nivel se modela como

$$\mu_t = \mu_{t-1} + \eta_t , \tag{29}$$

y las variables $I_{\ell,p}$ son variables de intervención tipo impulso incluidas para recoger anomalías detectadas en la observación correspondiente a la semana ℓ del año p. Nótese también que para evitar problemas de multicolinealidad se ha eliminado el regresor $X_{5,t}D_{26,t}^{sp}$. Los resultados de la estimación del modelo se muestran en los gráficos 1, 3 y 4. Las estimaciones de las variaciones estacionales obtenidas a partir del modelo estructural se han corregido de forma que el área bajo la función *spline* a lo largo de cada campaña sea nula. Y, por supuesto la corrección aplicada al componente estacional estimado se ha tenido en cuenta para corregir oportunamente las estimaciones del componente de largo plazo.

Gráfico 3. Estimaciones finales de las variaciones estacionales en puntos de ruptura, $\hat{\gamma}_t$.

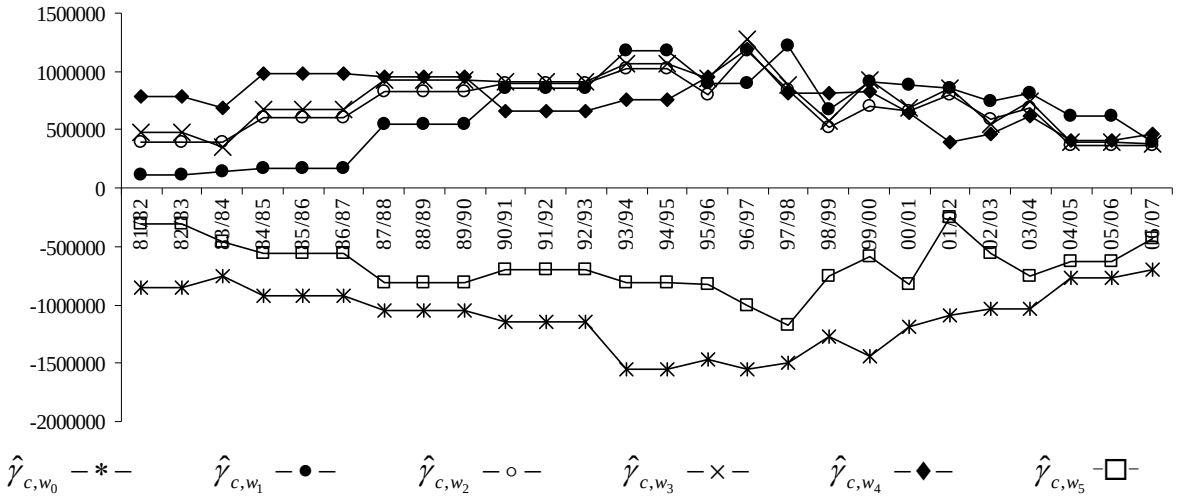
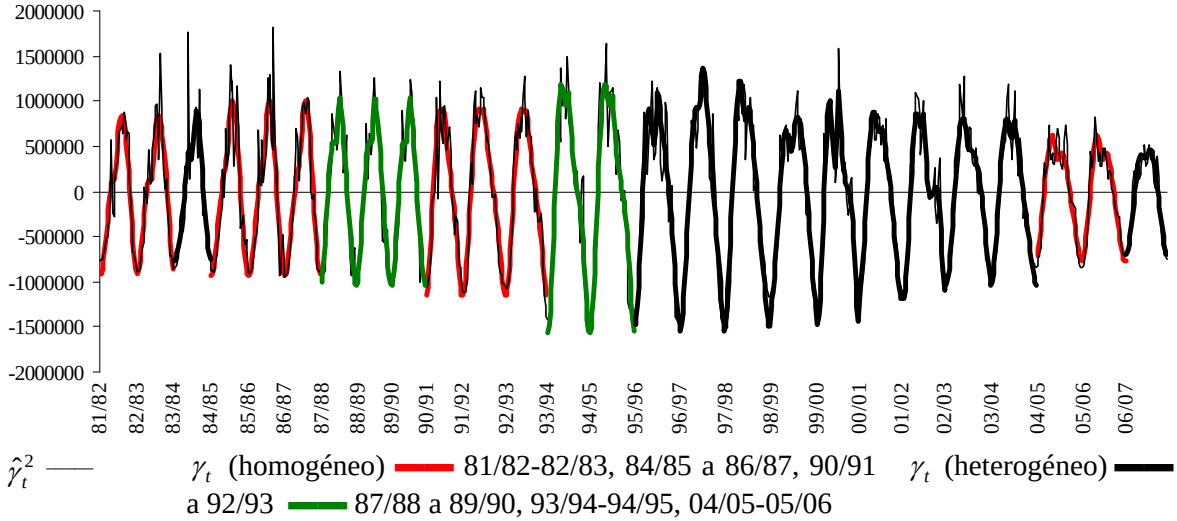


Gráfico 4. Variaciones estacionales (aproximación previa y estimaciones finales)



4. CONCLUSIONES

La metodología propuesta se ha mostrado como una herramienta estadística adecuada para recoger patrones estacionales heterogéneos en datos semanales, particularmente en series caracterizadas por periodos estacionales de longitud variable a lo largo de la muestra. La formulación del efecto estacional en una semana como una función de la proporción de la

longitud del periodo estacional que ha transcurrido desde su inicio hasta la semana en cuestión proporciona una ganancia en parsimonia y, sobre todo, hace posible modelar variaciones estacionales en las que la longitud del periodo estacional no es fija. Así, no es necesario eliminar observaciones relevantes o dividir la muestra artificialmente. Nótese que, en estas situaciones, el uso de formulaciones convencionales implica habitualmente que la serie tenga que considerarse como un conjunto de series correspondientes a diferentes submuestras y esta práctica puede distorsionar las estimaciones del movimiento de largo plazo.

Por otro lado, la metodología propuesta es suficientemente flexible para recoger inestabilidades en la magnitud de los efectos estacionales y el modelo puede adaptarse fácilmente a los cambios observados en la magnitud y forma del patrón estacional en el conjunto de semanas del periodo estacional. Es decir, la flexibilidad del modelo no depende de la varianza de los términos de perturbación que dirigen el comportamiento dinámico de los parámetros cambiantes en las formulaciones estocásticas convencionales. A partir de los cambios observados en la forma del patrón estacional, el modelo se formula con objeto de identificar los cambios en la media del proceso estocástico que recoge la variación estacional. Por lo tanto, los elementos de la *spline* no son siempre resultado de un proceso de optimización. Además, la *spline* ajustada a los datos proporciona estimaciones alisadas de los efectos estacionales, lo que facilita la descripción de los patrones estacionales y, por tanto, ayuda a los agentes que tienen que tomar decisiones teniendo en cuenta la naturaleza estacional de las magnitudes económicas de interés.

BIBLIOGRAFÍA

EUBANK, RL (1988), Spline smoothing and nonparametric regresión, M. Dekker: New York.

FERREIRA, E, NÚÑEZ-ANTÓN, V, y RODRÍGUEZ-PÓO, J (2000), Semiparametric approaches to signal extraction problem in economic time series, Computational Statistics & Data Analysis, 33, pp 315-333.

HARVEY, A C (1989), Forecasting, Structural Time Series and the Kalman Filter, Cambridge University Press: Cambridge.

HARVEY, AC, y KOOPMAN, SJ (1993), Forecasting hourly electricity demand using time varying splines, Journal of the American Statistical Association, 88, pp 1228-1236.

HARVEY, AC, KOOPMAN, SJ, y RIANI, M (1997), The modelling and seasonal adjustment of weekly observations, Journal of Business and Economic Statistics, 15, pp 354-368.

ORBE, S, FERREIRA, E, y RODRÍGUEZ-PÓO, J (2005), Nonparametric estimation of time varying parameters under shape restrictions, Journal of Econometrics, 126, pp 53-77.

PEDREGAL, DJ, y YOUNG, PC (2006), Modulated cycles, and approach to modelling periodic components from rapidly sampled data, International Journal of Forecasting, 22, pp 181-194.

POIRIER, DJ (1976), The Econometric of Structural Change with Special Emphasis on Spline Functions, North Holland Publishing Company: Amsterdam.

ÁREA XIV

DIDÁCTICA DE LA ECONOMÍA Y METODOLOGÍA

DIDÁCTICA DA ECONOMIA E METODOLOGIA

Evaluating the effect of public training programs. A survey from Spain.

JOSÉ MANUEL CANSINO MUÑOZ-REPISO
E-mail: jmcansino@us.es
ANTONIO SÁNCHEZ BRAZA
E-mail: asb@us.es
Departamento de Teoría Económica y Economía Política
Facultad de CC. Económicas y Empresariales
Universidad de Sevilla

Abstract
This paper estimes the average effect of a binary treatment on a scalar outcome (time needed to find a job). Such a treatment is the public training schools program implemented in the South of Spain (Seville). Speciffically, the paper estimes the average effect of this training program on the time needed to find a job. Difficulty in accesing microdata has obstructed in Spain this type of evaluation, largely developed in France, Germany, United Kingdom and USA.
Two methods are developed. The first uses an estimator which weights observations by the inverse of the propensity score. This estimator let us conclude that, for participants, the time needed to find a job is reduced in 471 days.
The second one is the differences estimator, that let us conclude that the time needed to find a job is reduced in 448 days.

Key words: Training programs evaluation, public policies, propensity score, differences estimator.
JEL Code: H30, C10
Thematic Area: Economic Didactics and Methodology.

1.- INTRODUCTION.

From Heckman, Clements and Smith (1997), it is generally accepted that social programs impact differently on individuals when they differ in characteristics like the previous educational level, professional experience, age, earnings, family background, etc. The individual's addecuate assignement to the set of programs disposable, begins a crucial issue in political decisions.

After controlling by covariates and by knowing the average effect of a program on an appropiatte outcome, the public decisor can decide which program would provoke the best impact on individuals by taking into account the program's average effect on subpopulations previously estimed. Social welfare can improve if public decisors follow an assignment rule which let them to determine which individuals must receive which treatment. Manski (2001) theorized this assuming the case of a finite set of rival treatments. Cansino and Roman (2007) explored this for the Spanish Accounting Court.

The aim of this paper is to estime the average effect of a binary treatment on a scalar outcome. Such a treatment is the training schools program implemented in the South of Spain (Seville) between 1997 and 1999. Speciffically, the paper estimes the average effect of this training program on the time needed to find a job.

We select the province of Seville as this is the zone with the most widely developed number of training policies until now. The evaluation is carried out by estimating the program's average effect over the individuals' ability of the sample to find a job, individuals being unemployed between 16 and 25 years old.

Following Hirano, Imbens and Ridder (2003), the paper estimes the average effect by using an estimator which weights observations by the inverse of nonparametric estimates of the propensity score. The differences estimator is also implemented to compare results.

Difficulty in accesing microdata has obstructed in Spain this type of evaluation, largely developed in France, Germany, United Kingdom and USA. This paper contributes to the literature in the sense that no evaluation based of non experimental methods has been applied to this training program in Spain before, apart from Cansino and Sanchez (2008a, 2008b).

The model and program characteristics are described in section 2. Section 3 is destined to explain the database used. In section 4, we define the estimator which weights observations by the inverse of the propensity -Hirano, Imbens and Ridder (2003)- and the results are obtained. The differences estimator is define in section 5 and the results are also obtained. Section 6 concludes.

2.- MODEL CHARACTERISTICS AND DATABASE.

2.1.- The average treatment effect.

The development of public policy evaluation has benefited from the use of causal inference[1]. One of the results is the Potencial Outcome Model –POM-, which allows us to compare participants and non participants in public programs[2]. A prolific development of the POM with regard to training programs evaluation comes thanks to Roy (1951) and Rubin[3] (1974, 1978). This paper support the Roy-Rubin Causal Model (RRM).

In the implementation of POM and RRM, the individual values of the main variables can be extracted from randomized experiments or from observational data. Both types of experiment will notably determine the evaluation and will promote different methodological developments.

In social sciences, randomized experiments face important problems related to cost, moral limitations, attrition and problems derived from the Hawthorne effect -Burtless (1995) and Cameron and Trivedi (2005)-. This can been solved by using observational data. In these cases, Rosenbaum (1999, p. 266) says that the researcher should design a treatment group and a control group from the individuals who have or have not been treated. The objective is to reproduce a scenario which is as similar as possible to a randomized experiment[4].

However, models which include counterfactual events (like an individual participating and not participating at the same time in a training program) are uneffective in individual causal effects estimation. Holland (1986) named this problem as the fundamental identification problem. The fundamental identification problem makes us look for second best solutions in which researchers leave the estimation of the individual causal effects, opting for an average effect estimation, which usually refers to the following specification.

To solve the identification problem, we mantein throughtout the paper the unconfounddedness assumption (Rubin, 1978; Rosenbaum and Rubin, 1983), which is also known as the selection on observables assumption (Barnow, Cain and Goldberger, 1980), which asserts that conditional on the observed covariates, the treatment indicator is independent of the potential outcomes.

[1] Refering to the theoretical approach of casuality and its use in randomized experiments, see Cox (1992). Other authors such as Dawid (1979, 2000), Holland (1986), Heckman (1990) and Pearl (2000) also discuss the meaning of casuality in such an enviroment. Finally, in the specific case of training programs, we have refered to the seminal papers of Rubin (1974) and Heckman and Hotz (1989).

[2] Cameron and Trivedi (2005, pp. 31 and onwards) expose the POM advantages compared to alternative models.

[3] The first references that Rubin considered were Neyman (1923, 1935) and Fisher (1928, 1935).

[4] The seminal papers in this environment were implemented in Medicine. The papers of Cameron and Pauling (1976), Billewicz (1965) and Cochran (1968), must be highlighted. An interesting comment about these seminals papers is contained in Rosenbaum (1995, 1996). Two of the best known papers in quasi-experimental methods are Kiefer (1979) and Bassi (1984).

The Average Treatment Effect of the program[5] (ATE) is addressed in a partial equilibrium environment and, by using the potential outcome notation popularized by Rubin (1974), it is obtained as the average expected value from the difference between the potential values of Y_1 (the case of an individual treated) and Y_0 (the case of a non treated individual)[6]. Implicit in this notation is the stability assumption or SUTVA (Rubin, 1978) that individuals are not affected by receipt of treatment by others, and there is only one version of the treatment. As a consecuence, no general equilibrium effects are considered (Cameron and Trivedi, 2005, p. 872).

The Average Treatment Effect on the Treated (ATET) is given as the average expected value from the difference between the potential values of Y_1 and Y_0 but only with respect to individuals who have received treatment.

Given the fact that the validity of average effects can be damaged if treated and non treated individuals have different characteristics apart from their participation or not in the program, it will be possible to control the differences if they can be observed and the treated and non treated only differ in these characteristics. This is the base of the selection on observables method in which the characteristics are noted as the covariate or vector *X*.

2.2.- The caracteristics of the program evaluated.

The training school program was designed as a nationwide experimental one implemented by the Spanish Department of Labor (more specifically, by the National Institute of Employment). Considering the first results of the program, the Spanish Department of Labor decided to convert it into a permanent program regulated by the Department Labor's order of march 29st - 1988. Finally, the Department Labor's order of august 3st - 1994 added this program into the set of the national policies of employment until now.

The training school organizes its activities into two steps; the first one gives a theoretical education to the unemplÿÿedÿÿÿÿ tÿÿ sÿÿonÿÿÿÿe offers a professional stage.ÿÿn order to judge the interest of this public training program, three parameters have been considered. The first is the number of participants. After the experimental period, the average number of participants had a range of between 45000 and 50000 young unemployed for every year. Compared to the whole of Spain, Seville is the zone with the largest number of implemented projects since 1985. This justifies the geographical focus of the paper.

Secondly, the size of the public funds absorbed as Table 1 shows.

Thirdly, the EU authorities supported this training program allowing the use of the European Social Fund to finance it.

Table 1 shows the total budget for joint employment actions between the Spanish Department of Labor and the European Social Fund for the period 2000-2006. This allows us to form an idea of the importance that the European Social Fund has as co-financer of the training programs, given that its participation in these joint budget actions is 60'32%.

[5] Although in this paper only the most well-known average effects are used, Imbens (2004, p. 4) has summarized all the possible types of average effects of treatment in literature. In this sense, he refers first to the PATE (Population Average Treatment Effect) as the average effect that treatment causes on population, and to the PATT (Population Average Treatment effect for the Treated) as the average effect when only treated are considered. Secondly, the SATE (Selected Average Treatment Effect) would show the average treatment effect when the evaluation is carried out only taking into account only a sample of the population, and the SATT (Selected Average Treatment effect for the Treated) when the sample is extracted only from the treated population. Finally, the CATE (Conditional Average Treatment Effect) would estimate the average effect of the treatment conditioned to the covariates' distribution and the CATT (Conditional Average Treatment effect for the Treated) would estimate the same but only considering the treated population.

[6] The *ATE* is addresed in a partial equilibrium enviorment diferent, for example, to the mega mentioned in Cansino, Cardenete and Roman (2007).

Table 1. Source: INEM (National Institute of Employment)

Joint budget actions between the Spanish Department of Labor and the European Social Fund (2000-2006)	
TOTAL ***9.540.537.406*** Euros	
ESF ***5.768.208.916*** Euros ***60'46*** %	***SDL*** ***3.772.328.490*** Euros ***39'54*** %

2.3.- Definition of *d* and *y* variables.

We define *D* as the binary variable which indicates the participation of the individuals in the sample, taking values 1 or 0 depending on if the individual considered participates or not in the program.

$D_i = 1$ will indicate that individual *i* has participated in one of the programs and $D_i = 0$ will indicate that individual *i* has not participated in any program.

The scalar Y, is the response variable from which the program's average effects will be evaluated. We define *Y* as the ability of the individual *i* to find a job, and shows how much time he has to spend searching for a job[7].

The choice of the response variable is justified because the individuals of the sample, both the participants and the non participants (control group), are initially unemployed and included in the oficial census of people who are searching for jobs. Most of them have not been working before or have a short labor experience because of their age and lack of experience.

For that reason, it is relevant for the program evaluation to consider a response variable which allows us to measure the abilities of these people to find jobs[8].

In this sense, Y_i will measure the "ability of individual *i* to find a job", and we define Y_i as

$$Y_i = 1 - \frac{number\ of\ consecutive\ days\ until\ the\ individual\ "i"\ find\ a\ job}{total\ duration\ of\ observation} \qquad (1)$$

The period of observation we have considered consists of three years[9] (1095 days). We started to measure this time from the moment the participants finished the training program (generally at the end of 1999) and january 1st 2000 for the individuals of the control group[10].

The value of *Y* varies between 0 and 1. If *Y* is equal to 0, it means that individual *i* has not found a job during the period considered. This is the worse scenario for the program's effectiveness. If *Y* is close to 1, the individual *i* has found a job in a short period of time and if *Y* is equal to 1, it implies that individual *i* has found a job the first day after finishing the training program.

[7] For a further investigation, we could define two outcomes in an alternative way. The first one let us to know the treatment average effect on the individual probability to find a job. The second would be the time needed to find a job conditioned to the unemployments subset (treated ond controls) who have find a job.

[8] The objective of this program is to act as an iniciative of young unemployed minors (less than 25 years old) in finding a job.

[9] A relatively broad period of time has been considered, three years (1095 days) due to specific problems of this collective in finding jobs.

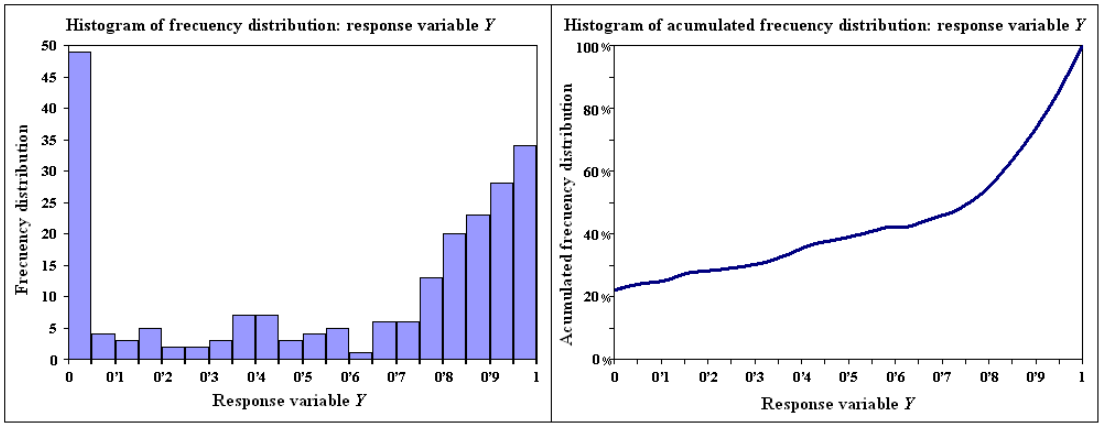
[10] The date fixed to start the test for the control group coincides with the starting date for many of the individuals in the participants group.

Table 2 summarizes the main descriptive data of Y, while *Figure 1* includes the frecuency distribution and the acumulated frecuency distribution of Y.

Table 2. Source: Own elaboration

Descriptive statistics related to response variable *Y* "ability to find a job"			
Mean	0,578147	Kurtosis	-1,392020
Median	0,773516	Coeffcient of Asymmetry	-0,533554
Mode	0	Minimum	0
Standard deviation	0,380657	Maximum	1
Variance	0,144900	Range	1

Figure 1. Source: Own elaboration



2.4.- The Model: methodology based on identification and selection on observables.

Because of the fact that the validity of average effects can be damaged if participants and controls show characteristics different from their participation or non participation in the training program, these charateristcs must be controlled because of their effect on the values of the response variable.

If the observed characteristics are the only individual (participants and controls) characteristics that differ, we can therefore control these differences. This is the base of the selection on observable model[11].

Selection on observables allows us to isolate the effect of a covariate (or a vector of covariates)[12] maintaining the independence between the treatment indicator variable *D* and the response variable, Y. This condition can be expressed as

$$(Y_1, Y_0) \perp D \mid X \qquad (2)$$

Selection on observables supports the independence assumption typical in randomized experiments, contributing to the comparison between participants and controls.

Following Heckman and Hotz (1989, p. 865), selection on observables is recommended when the independence between *D* and *Y* is because of the covariate *X* (or vector of covariates),

[11] When controlled and treated differ in unobserved characteristics like psicologycal ones, average effect can be estimed by the differences in differences estimator. For the context of this paper, see Cansino and Sanchez (2008).

[12] As an introduction to the framework of the observational methods we recommended the examples that are used by Rosenbaum (1995, pp. 2 and onwards) in his exposition about these methods; this also can by said of Cochran (1968) and Cameron and Pauling (1976). We also recommend the papers of Billewicz (1965) and Moertel et al. (1985), both of them refering to the two previous examples.

which has influence on the individual selection process, so by controlling *X* we give a solution to possible biased selection, making the dependency between *D* and *Y* disappear.

In the selection on observable context, when the independence assumption is guaranteed, we considered, according to Dehejia and Wahba (1999, p. 1057), that

$$\mathrm{E}\,[Y_1 - Y_0 \mid X] \;=\; \mathrm{E}\,[Y \mid X,\ D=1] - \mathrm{E}\,[Y \mid X,\ D=0] \tag{3}$$

Equation 3 lets us express the *ATE* as

$$\begin{aligned} ATE \;&=\; E\left[Y_1 - Y_0\right] \;=\; \int \left(E\left[Y_1 - Y_0 \mid X\right]\right)\, d\,P(X) \;= \\ &=\; \int \left(E\left[Y \mid X,\ D=1\right] - E\left[Y \mid X,\ D=0\right]\right)\, d\,P(X) \end{aligned} \tag{4}$$

In this way[13], it is possible to determine *ATE* from the difference between the average observed value of the response variable of the participants and the controls, calculating the difference for every possible value of *X*.

In a similar way, it is possible to calculate the average effect of the training program only for participants (*ATET*) as:

$$\begin{aligned} ATET \;&=\; E\left[Y_1 - Y_0 \mid D=1\right] \;= \\ &\int \left(E\left[Y \mid X,\ D=1\right] - E\left[Y \mid X,\ D=0\right]\right)\, d\,P(X \mid D=1) \end{aligned} \tag{5}$$

Therefore, *ATET* will be equal to the difference between the average observed values in the response variable of the participants and the average values of the controls for every different value of *X* when $D = 1$.

2.5.- The Covariate Vector Definition $X^3 = (X_1, X_2, X_3)$.

With *D* defined, *X* will be a covariate[14] with respect to *D* if, for each of the individuals observed, its values remain the same for each value of *D*. That is to say $X_{0i} = X_{1i}$, being X_{0i} the *X* value before the event *D* ($D = 0$) and X_{1i} the *X* value after happening *D* ($D = 1$).

X covariate is also named contaminant because of the fact that *X* can contamine *Y*, adding its own effects[15] to those provoked by D.

The fact that *X* is predetermined with respect to *D* does not imply that this independence is bidirectional, because it is possible that, as a characteristic of considered population, dependence in an opposite direction can appear, making the value of *D* be affected for *X*.

From the sample information included in the database, we consider three predetermined variables which form the vector of covariates $X^3 = (\boldsymbol{X_1}, \boldsymbol{X_2}, \boldsymbol{X_3})$. The database only allows us to include in the model a complete information about this three covariates. We define the covariates in the following way:

- $\boldsymbol{X_1}$: sex. This shows if the individual considered is male or female, taking $X_1 = 1$ in the case of a male and $\boldsymbol{X_1} = 2$ in the case of a female.
- $\boldsymbol{X_2}$: age. This shows the individual's age at the beginning of the observational period. In the case of participants, X_2 shows the individual's age when the training program is over.
  For controls, X_2 shows the individual's age as of January 1 st, 2000.
  Considering that age range for participants in one of the considered training programs is between 16 and 24 years old, and also considering that the program may extend for 1 or 2 years, X_2 covariate will have values of between 17 and 26 years old.

[13] To improve knowledge of selection on observables, we recomend Barnow, Cain and Goldberger (1980).

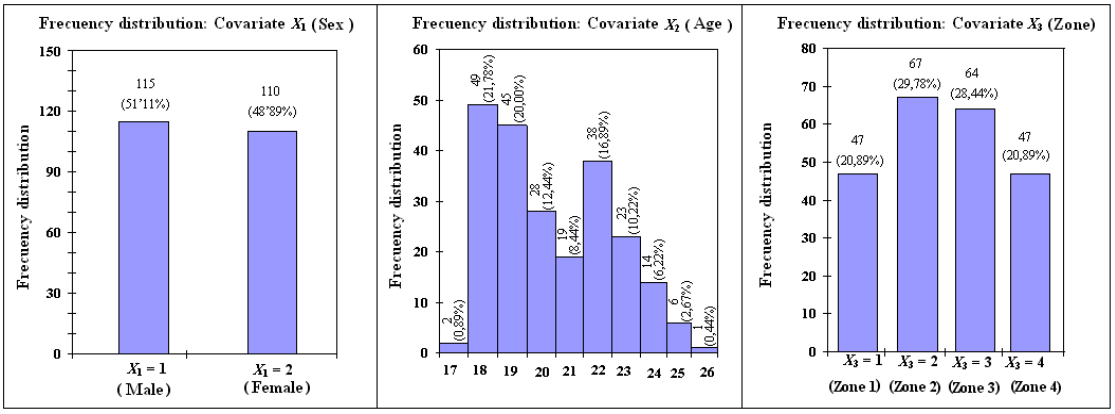
[14] We talk about one covariate but everything we state can be extrapolated for the case that *X* is a vector of n covariates, as $X^n = (X_1, X_2, \ldots, X_n)$.

[15] To read more, the comments of Rubin (1978) about covariates are very interesting.

- X_3: zone, showing the city where individuals took the training program or, in the case of controls, where individuals lived. For this last variable, we have divided the area of Seville (Spain) into four zones, named as zone 1 (Sevilla city), zone 2 (east and northeast of Seville), zone 3 (south and southwest) and zone 4 (west and northwest). The criterion of mapping is an operational one. So the X_3 covariate will take the following values: $X_3 = 1$ for zone 1, $X_3 = 2$ for zone 2, $X_3 = 3$ for zone 3 and $X_3 = 4$ for zone 4.

Figure 2 includes the frecuency distribution of the covariates X_1, X_2 and X_3.

Figure 2. Source: Own elaboration



The inclusion in the regressions of each predetermined variable, depending on each variable's characteristics, is done with the following procedure:

- X_1 is a qualitative variable and will be made up by dummy variables. To prevent perfect multicolineality, we introduce in the model as many dummies as categories less one.
  So in order to consider X_1, we will include the dummy variable X_{11} which can value 0 and 1.

$$X_{11} = \begin{cases} 1 & \textit{in the case of a male} \\ 0 & \textit{in the case of a female} \end{cases}$$

- X_2 is a quantitative variable taking values of between 17 and 26 years old. In this case, no dummies are necessary.
- Considering X_3 as a qualitative variable, we need to define three dummies noted as X_{31}, X_{32} and X_{33}, taking values 0 or 1.

$$X_{31} = \begin{cases} 1 & \textit{if the individual belongs to zone 1} \\ 0 & \textit{if the individual belongs to any of the other three zones} \end{cases}$$

$$X_{32} = \begin{cases} 1 & \textit{if the individual belongs to zone 2} \\ 0 & \textit{if the individual belongs to any of the other three zones} \end{cases}$$

$$X_{33} = \begin{cases} 1 & \textit{if the individual belongs to zone 3} \\ 0 & \textit{if the individual belongs to any of the other three zones} \end{cases}$$

3.- The "BASEVAFOR 96-03" database.

The "BASEVAFOR 96-03" database[16] has been constructed from individuals who have participated in the training programs carried out in the south of Spain (Seville) between 1996 and 2003[17]. We have selected those individuals who have finished the training program in 1999,

[16] The individual data came from the oficial employment agency (INEM).

[17] This time limit is due to the fact that it includes information related to the training programs that finished in 1999, meaning the program started in either 1997 or 1998. In some cases additional information of the individuals has been refered to from 1996. On other hand, as we refer to programs

the last year in which information was available when we started the evaluation[18]. Depending on the length of the program (1 or 2 years), programs finishing in 1999 started in either 1997 or 1998.

Only individuals who finished the training program have been considered, rejecting those who left the program before the end. The total of individuals was 1528, and from that figure we have selected a sample of 150. The selected individuals make up the participant group in our investigation. Similarly we have selected 75 individuals[19] to be the control group. The controls have similar characteristics to the participants[20].

The "BASEVAFOR 96-03" incudes two types of data related not only to participants but also with the controls. Firstly, "BASEVAFOR 96-03"gives us information about the periods of employment and unemployment of individuals, including data related to the number of times the individual has applied for a job. We will use this type of information to construct Y. Secondly, "BASEVAFOR 96-03" contains information related to the covariates considered; sex, age and resident zone.

4.- ESTIMATION OF THE AVERAGE EFFECT: EMPIRICAL RESULTS.

4.1. The propensity score.

To avoid the need to match individuals on the values of all covarities, Rosenbaum and Rubin (1983, 1984) develop an approach based on the propensity score[21], the probability of one individual to participate in a program (probability of $D = 1$), conditioned to the values of vector X. By making this probability $\varepsilon(X)$, we can express this as:

$$\varepsilon(X) = P(D = 1 \mid X) \tag{6}$$

which is assumed to be bounded away from zero and one.

This shows that this probability is a function of X, which is usually unknown, and therefore it should be estimated by using the database.

Rosenbaum and Rubin (1983) also define the assumption of propensity score independence as

$$(Y_1, Y_0) \perp D \mid \varepsilon(X) \tag{7}$$

$\varepsilon(X)$ being the probability of participating in a program conditioned on X. In this way, the independence assumption typical in randomized experiments is guaranteed. This assumption lets us argue that all of the observations with the same propensity score will have the same

finished during 1999, the information relative to the response variable extends until 2002, to which some information of the individuals of 2003 is added.

[18] The requirements of the individuals were to be between the ages of 16 and 25 years, be unemployed and have signed up to the unemployment office, and have the minimum requirements to begin a training contract.

[19] We maintain the ratio 2/1 which was used in the evaluation of the JTPA study. The National Job Training Partnership Act (JTPA) Study was commissioned by the U.S. Department of Labor in 1986 to measure the benefits and cost of selected employment and training programs for economically disadvantaged adults and out-of-school youths. See Heckman, Ichimura and Todd (1997, 1998) and Heckman, Ichimura, Smith and Todd (1998). The dimension of the database is comparable to similar evaluations included in Deheija and Wahba (1999, p. 1056-Table 2-). The control group was constructed by the regional labor authorities for our evaluation.

For data, we assume some of them are missing at random (Rubin, 1976; Little and Rubin, 1987).

[20] They are individuals who, during the period of time considered, show the same characteristics of participants and have the requirements to participate in the training program. Also, it is possible that some of them could have applied for a program and could not participate due to place limitation.

[21] Really, there is not exist consensus on the number of covariates which recommended the use of propensity score instead of the covariates vector (Imbens, 2004). Any case, this is recommended when the overlap assumption can not be guaranted for all the covariates.

distribution as the vector X, which means that we can compare the data observed for either participants or controls with the same propensity score.
Following Hahn (1998, p. 316), the calculation of the conditioned probability of participation in a program, given certain observable characteristics, plays a crucial role in controlling bias in order to obtain an estimator of the program's effects.
By using propensity score, we proceed as if it were the case of an unidimensional variable, improving evaluation efficiency by avoiding the management of a large number of covariates included in vector X.
The way to estimate the effects of a training program using selection on observables and by applying propensity score, is divided into "two-stage".

4.2. The calculation of propensity score on vector covariates $X^3 = (X_1, X_2, X_3)$ (sex, age and zone).

From (6), we can now express the probability of an individual's participation in a program conditioned on the value of vector X, as

$$\varepsilon(X) = P(D = 1 \mid X) = F(\beta X) \quad (8)$$

β is the parameter's vector associated with the covariates. The value of this probability will remain conditioned to the value of the distribution function at point βX_j ; X_j being every possible value that can adopt the vector of covariates X, with $j = 1, \ldots, k$.

Depending on the specific function of F, different selection models of binary response could be specified. From the possible non linear options, we have selected three: the Probit Model, the Logit Model and the Extreme Value Model Type I.

There is not a generally accepted selection criterion in choosing one of these three models for the estimation of the "propensity score", so the way in which the choice is made is due only to practical reasons. We will estimate the three models and after analysing the obtained results, we will choose the one with the best results according to the criteria specified later on.

Carrying out regressions on the vector of covariates $X^3 = (X_1, X_2, X_3)$, the results obtained are contained in *Tables 3*, *4 and 5*.

Table 3. Source: Own elaboration

The calculation of the "propensity score" by using the Probit Model					
Dependent Variable: *D* (Prob. *D* = 1)					
Method: ML - Binary Probit					
Variable	**Coefficient**	**Coefficient Value**	**Std. Error**	**t -Statistic***	**Prob.**
Fixed effect	μ	-2'366701	0'848762	-2'788414	0'0053
X_{11}	β_{11}	0'580077	0'185036	3'134942	0'0017
X_2	β_2	0'125535	0'040936	3'066652	0'0022
X_{31}	β_{31}	-0'254362	0'285963	-0'889491	0'3737
X_{32}	β_{32}	0'002826	0'254795	0'011091	0'9912
X_{33}	β_{33}	0'058707	0'257340	0'228131	0'8195
Mean dependent var		0'666667	S.D. dependent var		0'472456
S.E. of regression		0'460830	Akaike info criterion		1'249478
Sum squared resid		46'50776	Schwarz criterion		1'340574
Log likelihood		-134'5663	Hannan-Quinn criter.		1'286245
Restr. log likelihood		-143'2157	Avg. log likelihood		-0'598072
LR statistic (5 df)		17'29884	McFadden R-squared		0'060394
Probability (LR stat)		0'003967			
* t - Statistics adjusted by White's method					

Table 4. Source: Own elaboration

The calculation of the "propensity score" by using the Logic Model					
Dependent Variable: *D* (Prob. *D* = 1)					
Method: ML - Binary Logic					
Variable	**Coefficient**	**Coefficient Value**	**Std. Error**	**t -Statistic***	**Prob.**
Fixed effect	μ	-3'801614	1'384561	-2'745719	0'0060
X_{11}	β_{11}	0'959225	0'311150	3'082833	0'0021
X_2	β_2	0'201706	0'067164	3'003174	0'0027
X_{31}	β_{31}	-0'428509	0'477954	-0'896549	0'3700
X_{32}	β_{32}	0'007757	0'422600	0'018354	0'9854
X_{33}	β_{33}	0'090384	0'427989	0'211182	0'8327
Mean dependent var		0'666667	S.D. dependent var		0'472456
S.E. of regression		0'460893	Akaike info criterion		1'250555
Sum squared resid		46'52056	Schwarz criterion		1'341651
Log likelihood		-134'6875	Hannan-Quinn criter.		1'287322
Restr. log likelihood		-143'2157	Avg. log likelihood		-0'598611
LR statistic (5 df)		17'05642	McFadden R-squared		0'059548
Probability (LR stat)		0'004394			
* t - Statistics adjusted by White's method					

Table 5. Source: Own elaboration

The calculation of the "propensity score" by using the Extreme Value Model Type I					
Dependent Variable: *D* (Prob. *D* = 1)					
Method: ML - Binary Extreme Value Type I (Gompit function)					
Variable	**Coefficient**	**Coefficient Value**	**Std. Error**	**t -Statistic***	**Prob.**
Fixed effect	μ	-2'563204	1'079534	-2'374361	0'0176
X_{11}	β_{11}	0'773186	0'250257	3'089566	0'0020
X_2	β_2	0'156557	0'054409	2'877383	0'0040
X_{31}	β_{31}	-0'354458	0'376021	-0'942655	0'3459
X_{32}	β_{32}	-0'009549	0'331510	-0'028805	0'9770
X_{33}	β_{33}	0'048197	0'340991	0'141344	0'8876
Mean dependent var		0'666667	S.D. dependent var		0'472456
S.E. of regression		0'461152	Akaike info criterion		1'252412
Sum squared resid		46'57280	Schwarz criterion		1'343508
Log likelihood		-134'8964	Hannan-Quinn criter.		1'289179
Restr. log likelihood		-143'2157	Avg. log likelihood		-0'599539
LR statistic (5 df)		16'63867	McFadden R-squared		0'058090
Probability (LR stat)		0'005239			
* t - Statistics adjusted by White's method					

In every case, in order to avoid possible heteroskedasticity problems, the t-statistic values are adjusted by White's method.

According to the corresponding value estimations of the parameters and t-statistics contained in *Tables 3*, *4 and 5*, the variables X_{11} and X_2 appear to be significant when calculating the probability of participation in the three models considered, while the dummy variables X_{31}, X_{32} and X_{33}, defined to include the zone covariate, appear insignificant. In any case, we have decided to mantain them because they help to improve the significance of all the estimated parameters and to improve goodness of fits.

From the three methods, the most efficient will be the one that shows less values of information criterion of Akaike, Schwarz and Hannan-Quinn and a higher value of the log likelihood function for each of the three models. This information appears in *Table 6*.

Table 6. Source: Own elaboration

Comparison of the obtained results from the three binary response models applied			
	Probit Model	Logit Model	Extreme Value Model
Log likelihood function	-134'5663	-134'6875	-134'8964
Criterion Akaike	1'249478	1'250555	1'252412
Criterion Schwarz	1'340574	1'341651	1'343508
Criterion Hannan-Quinn	1'286245	1'287322	1'289179

The model that shows the lowest values of the criterion of Akaike, Schwarz and Hannan-Quinn and a the highest value of the verosimilitude function is the Probit Model, and for this reason, it has been selected.

Following the procedure of the Probit Model, the equation which reflects the participation probability of an individual of the sample in the evaluated program conditioned on the of vector values is as follows:

$$P = -2'366701 + 0'580077\, X_{11} + 0'125535\, X_2 - 0'254362\, X_{31} + \\ + 0'002826\, X_{32} + 0'058707\, X_{33}$$

Table 7 contains the main data of descriptive statistics related to the probability of participation. The probability has been estimated for every individual of the sample by using the Probit Model. *Table 8* contains the estimation of the participation probability for every possible individual depending on the different values that vector $X^3 = (X_1, X_2, X_3)$ can have.

Table 7. Source: Own elaboration

Descriptive statistics related to "Propensity score" obtained by using the Probit Model			
Mean	0'667158	Kurtosis	-0'795769
Median	0'681896	Coeffcient of Asymmetry	-0'145460
Mode	0'682903	Minimum	0'358889
Standard deviation	0'129550	Maximum	0'920798
Variance	0'016783	Range	0'561910

Table 8. Source: Own elaboration

"Propensity score" values depending on possible values that vector $X^3 = (X_1, X_2, X_3)$ can have, according to the Probit Model

	Male				Female			
	Zone 1	Zone 2	Zone 3	Zone 4	Zone 1	Zone 2	Zone 3	Zone 4
17 years old	0'537092	0'636943	0'657695	0'635882	0'313141	0'409132	0'430973	0'408034
18 years old	0'586537	0'682903	0'702538	0'681896	0'358889	0'458488	0'480714	0'457367
19 years old	0'634645	0'726203	0'74449	0'725261	0'406757	0'508494	0'530757	0'507366
20 years old	0'680721	0'766358	0'783123	0'765491	0'456062	0'558366	0'580319	0'55725
21 years old	0'724162	0'803015	0'818146	0'80223	0'506053	0'607328	0'628636	0'606241
22 years old	0'76448	0'835956	0'849398	0'835257	0'55595	0'654647	0'675006	0'653605
23 years old	0'801313	0'865097	0'876852	0'864483	0'604975	0'699663	0'718811	0'69868
24 years old	0'834439	0'890472	0'900591	0'889942	0'65239	0'74182	0'759548	0'740906
25 years old	0'863765	0'912224	0'920798	0'911773	0'697532	0'780683	0'796839	0'779847
26 years old	0'889321	0'930579	0'93773	0'930201	0'739839	0'81595	0'830443	0'815197

Finally, to every individual (participants and controls) the estimated value of his "propensity score" conditioned on vector X^3 has been assigned. After doing this we proceed to calculate *ATE* ($\hat{\alpha}_{ATE}$) and *ATET* ($\hat{\alpha}_{ATET}$).

4.3. Weighting observations by the inverse of the propensity score.

By weighting observations by the inverse of a nonparametric estime of the propensity score, we have an efficient estimator of the average effect. ATE and ATET's estimators are expressed as follows[22]:

$$\hat{\alpha}_{ATE \rightarrow WEIGHTING\ PS} = \frac{1}{n} \sum_{i=1}^{n} Y_i \cdot \frac{D_i - \hat{\varepsilon}(X_i)}{\hat{\varepsilon}(X_i) \cdot \left(1 - \hat{\varepsilon}(X_i)\right)} \quad (9)$$

$$\hat{\alpha}_{ATET \rightarrow WEIGHTING\ PS} = \frac{1}{n_1} \sum_{i=1}^{n} Y_i \cdot \frac{D_i - \hat{\varepsilon}(X_i)}{1 - \hat{\varepsilon}(X_i)} \quad (10)$$

where $\hat{\varepsilon}(X_i)$ is the estimed value of the propensity score for the i-individual on vector X.

The obtained results are $\hat{\alpha}_{ATE}$ = 0,421280 and $\hat{\alpha}_{ATET}$ =0,430409.

The estimed value of the ATE is positive. In average, the sample's individuals' ability to find a job increases by 0,421280. In the case of the ATET, the estimed value is also positive, meaning that there is a favourable causal effect from the program. This result indicates that participant's ability to find a job has increased, on average, by 0,430409.

[22] Hirano, Imbens and Ridder (2003) developed *ATE* ($\hat{\tau}$) and *ATET* ($\hat{\tau}_{treated}$) estimators obtaining,

$$\hat{\tau} = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{Y_i \cdot T_i}{\hat{p}(X_i)} - \frac{Y_i \cdot (1 - T_i)}{1 - \hat{p}(X_i)} \right) \qquad 0 = \sum_{i=1}^{N} \hat{p}(X_i) \cdot \left(\frac{Y_i \cdot T_i}{\hat{p}(X_i)} - \frac{Y_i \cdot (1 - T_i)}{1 - \hat{p}(X_i)} - \tau_{treated} \right)$$

where Y_i is the outcome, T_i the binary variable which indicates if individual it's treated or control and the X covariate's vector which let us to define $\hat{p}(X_i)$ as the probability to participate in the program, conditioned on X.

5.- ESTIMATION OF THE AVERAGE EFFECT BY REGRESSION.

It is possible to determine the average effect of a training program on participants (*ATET*) by regression by using Least Squares, given that independence assumption is also guaranteed. According to Stock and Watson (2003, pp. 375 and onwards), we can obtain an estimator of the average effects of the evaluated training program on partipants ($\hat{\alpha}_{ATET}$)sing a linear model.

It is also possible to introduce covariates in the model as additional regresors. By doing that, we can measure the effects that contaminant variables have on *Y*. The inclusion of covariates in the model is shown by the following expression:

$$Y = \mu + \alpha \cdot D + \beta \cdot X + \varepsilon \qquad (11)$$

Y is the dependent variable which shows the potential results of the individual and *D* is a binary explanatory variable. *X* is the covariates' vector and β is the vector of the associated parameters. The parameter μ collects the fixed effects in the model and the parameter ε collects the random error of the model, with an average value equal to 0, $E[\varepsilon \mid D, X] = 0$.

The parameter α will determine the average effect of the program on participants. This parameter will be the "differences estimator with additional regressors". As independence assumption is guaranteed, if the necessary assumptions[23] for multiple regression by Least Squares are guaranteed too, this estimator will be unbiased and consistent.

The inclusion of the additional regressors in the model lets us improve the estimator's efficiency, reducing the random error variance. On the other hand, the addition allows us to test the randomness in the individual assigning procedure between the participants group and the control group, in the case that the assigning procedure is related to the additional covariates. That is to say, by including the covariates in the model we can control the probability of individuals being assigned between the participants group and the control group by adding the characteristics in which participants and controls differ.

The inclusion of predetermined variables in the model will consist of inserting the variables included in the vector $X^3 = (X_1, X_2, X_3)$ as additional regressors.

Table 9 shows the correlation matrix between the explanatory variables included in the model, allowing us to analyze possible multicolineality in the model.

Table 9. Source: Own elaboration

Matrix correlation between the variable *D* and the variables included in the vector $X^3 = (X_1, X_2, X_3)$						
	D	X_{11}	X_2	X_{31}	X_{32}	X_{33}
D	1'000000	0'194895	0'168779	-0'007731	0'006873	0'027864
X_{11}	0'194895	1'000000	0'005540	0'218229	0'014691	-0'092845
X_2	0'168779	0'005540	1'000000	0'143666	-0'105606	0'030735
X_{31}	-0'007731	0'218229	0'143666	1'000000	-0'334617	-0'323978
X_{32}	0'006873	0'014691	-0'105606	-0'334617	1'000000	-0'410569
X_{33}	0'027864	-0'092845	0'030735	-0'323978	-0'410569	1'000000

Linear correlation between variables is not obvious because all the coefficients are very low and far from ±1. With regard to the relationship between X_{31}, X_{32} and X_{33}, these variables show slightly high correlation because of the fact that they have been constructed to include the zone variable in the model. However, the values are never over ±0'5. On the other hand, the value of the determinant of the correlation matrix is 0'4464, far from 0. In addition, the condition

[23] These are the four assumptions: a) the conditioned distribution of the random error, given the explanatory variates X_{1i}, X_{2i}, …, X_{ki}, is equal to 0 on average (in this case, the explanatory variables are D_i, that indicates program participation, and the covariates). b) all the observations $i = 1, 2, \ldots, n$ are distributed both independently and identically random. C) X_{1i}, X_{2i}, …, X_{ki} and ε_i have four moments. D) Perfect multicolineality does not exist.

number of the correlation matrix is C = 2'3240, far from the limits that determine multicolineality. Everything we have exposed lets us indicate that multicolineality problems are not relevant in the model.

With the previous specifications made, we implement the regression by Least Squares. Results are shown in *Table 10*.

Table 10. Source: Own elaboration

The calculation of the "differences estimator with additional regressors"					
Dependent Variable: *Y* (Ability to find a job)					
Method: MCO					
Variable	**Coefficient**	**Coefficient Value**	**Std. Error**	**t -Statistic***	**Prob.**
Fixed effect	μ	-0'206095	0'216058	-0'953891	0'3412
$\boldsymbol{D}$	α	0'409173	0'050124	8'163276	0'0000
$\boldsymbol{X_{11}}$	β_{11}	0'128764	0'045467	2'832055	0'0051
$\boldsymbol{X_2}$	β_2	0'023647	0'010274	2'301544	0'0223
$\boldsymbol{X_{31}}$	β_{31}	-0'040937	0'065065	-0'629174	0'5299
$\boldsymbol{X_{32}}$	β_{32}	-0'028182	0'059367	-0'474715	0'6355
$\boldsymbol{X_{33}}$	β_{33}	-0'075899	0'058980	-1'286872	0'1995
R-squared		0'360720	Mean dependent var		0'578147
Adjusted R-squared		0'343125	S.D. dependent var		0'380657
S.E. of regression		0'308514	Akaike info criterion		0'516520
Sum squared resid		20'74949	Schwarz criterion		0'622799
Log likelihood		-51'10855	F-statistic		20'50139
Durbin-Watson stat		2'023048	Prob(F-statistic)		0'000000
* t - Statistics adjusted by White's method					

After estimating the parameters of the regressors, the equation's model is:

$$Y = -0'206095 + 0'409173\,\boldsymbol{D} + 0'128764\,\boldsymbol{X_{11}} + 0'023647\,\boldsymbol{X_2} - 0'040937\,\boldsymbol{X_{31}} - 0'028182\,\boldsymbol{X_{32}} - 0'075899\,\boldsymbol{X_{33}}$$

The α parameter (0'409173) is the "differences estimator with additional regressors" from the program's effects on the participants ($\hat{\alpha}_{ATET}$).

The individual significance of the explanatory variables included in the model is demonstrated by the values obtained for the t-statistc and its associated probability. From these values D, $\boldsymbol{X_{11}}$ (sex) and $\boldsymbol{X_2}$ (age) appear to be significant variables. The $\boldsymbol{X_{31}}$, $\boldsymbol{X_{32}}$ and $\boldsymbol{X_{33}}$, defined to include the zone covariate, appear insignificant. In any case, we have decided to mantain them because they help to improve the significance of all the estimated parameters and to improve goodness of fit. In relation to the fixed effects, the adjustment shows non significance of the constant in the regression.

As in the previous estimation, the t-statistic values has been adjusted by White's method in order to correct possible problems of heteroskedasticity.

With respect to goodness of fit, the R-squared statistic equals 0'360720 and shows that the explanatory power of the considered variables is equal to 36'0720 percent, significantly improving the accuracy of the adjustment over the estimation without additional regressors. On the other hand, the value of R^2 adjusted is equal to 0'343125. The standard regression error is

very low, 0'308514, and the estimated residual variance is 0'09518. The low values of the information criteria of Akaike (0'516520) and Scharwz (0'622799) also support the accurracy of the model. In addition, the value of the Durbin-Watson statistic is close to 2, 2'023048, indicating that autocorrelation problems of the residuals are not relevant.

The joint significance of all the model estimated parameters can also be tested from the value of the probability of the F-Snedecor contrast. In this case the probability is equal to 0'00000, meaning the acceptance of the joint significance of all the parameters of the model. This implies that we can consider all the parameters significantly different from 0 whilst, at the same time, having a high probability.

With the estimation carried out, the meaning of all the parameters of the model is important. From the results obtained:

- The α coeficient, associated with the explanatory variable D, shows that when an individual has participated in the program ($D = 1$), the response variable increases by 0'409173. This is the effect on the response variable of the participants, and means that the ability to find a job increases by 0'409173 over the non participants' value.
- The β_{11} coefficient, associated with the explanatory variable $\boldsymbol{X_{11}}$, shows that in the case of a male participant ($\boldsymbol{X_{11}} = 1$), the response variable increases by 0'128794. In the case of a woman participant ($\boldsymbol{X_{11}} = 0$), this effect is not added. This means that males have a better situation than females in terms of Y, which is higher by 0'128794 than the registered value in the case of females.
- The β_2 coefficient, associated with the explanatory variable $\boldsymbol{X_2}$, collects the effect of age on Y. Due to variable $\boldsymbol{X_2}$ being a quantitive variable, this effect will be related to the possible values of this variable, adding 0'023647 to the value of the response variable *Y* for every unitary change registered by $\boldsymbol{X_2}$. Therefore, individuals belonging to the sample (individuals whose age is between 17 and 26) show a higher ability to find a job as the value of $\boldsymbol{X_2}$ increases.
- The β_{31}, β_{32} and β_{33} coefficients, associated with the explanatory variables $\boldsymbol{X_{31}}$, $\boldsymbol{X_{32}}$ and $\boldsymbol{X_{33}}$, show that:

  When the individual belongs to zone 1 ($\boldsymbol{X_{31}} = 1$), the effect on variable response *Y* is a reduction that equals 0'040937. When he belongs to zone 2 ($\boldsymbol{X_{32}} = 1$), the effect on variable response *Y* is a reduction that equals 0'028182. The reduction is equal to 0'075899 when the individual belongs to zone 3 ($\boldsymbol{X_{33}} = 1$). Finally, if the individual belongs to zone 4 ($\boldsymbol{X_{31}} = 0$, $\boldsymbol{X_{32}} = 0$ and $\boldsymbol{X_{33}} = 0$), no effect is added due to the fact that we take this category as the base.

  Therefore, this shows that individuals belonging to zone 4 have the best results with respect to the response variable Y, followed by individuals belonging to zone 2 and then individuals belonging to zone 1.
- The μ coefficient collects the effect on *Y* in the event that all of the explanatory variables become null. It shows the effect on *Y* in the event that the individual has not participated in the program ($D = 0$), is female ($\boldsymbol{X_{11}} = 0$) and belongs to zone 4 ($\boldsymbol{X_{31}} = 0$, $\boldsymbol{X_{32}} = 0$ and $\boldsymbol{X_{33}} = 0$). This value equals 0'206095. We have to add to this value the effect provoked by the covariate $\boldsymbol{X_2}$ (age).

To summarize, *Tables 11 and 12* contain the model estimated values for the response variable *Y* for every possible value of the explanatory variables. Table *11* contains the values for the case of participants while *Table 12* contains the values for the case of non participants.

Table 11. Source: Own Elaboration

Estimated values for the response variable Y for the participant individuals according to ″the differences estimator with addtional regressors″ model

	PARTICIPANT INDIVIDUALS							
	Male				Female			
	Zone 1	Zone 2	Zone 3	Zone 4	Zone 1	Zone 2	Zone 3	Zone 4
17 years old	0'692904	0'705659	0'657942	0'733841	0'56414	0'576895	0'529178	0'605077
18 years old	0'716551	0'729306	0'681589	0'757488	0'587787	0'600542	0'552825	0'628724
19 years old	0'740198	0'752953	0'705236	0'781135	0'611434	0'624189	0'576472	0'652371
20 years old	0'763845	0'7766	0'728883	0'804782	0'635081	0'647836	0'600119	0'676018
21 years old	0'787492	0'800247	0'75253	0'828429	0'658728	0'671483	0'623766	0'699665
22 years old	0'811139	0'823894	0'776177	0'852076	0'682375	0'69513	0'647413	0'723312
23 years old	0'834786	0'847541	0'799824	0'875723	0'706022	0'718777	0'67106	0'746959
24 years old	0'858433	0'871188	0'823471	0'89937	0'729669	0'742424	0'694707	0'770606
25 years old	0'88208	0'894835	0'847118	0'923017	0'753316	0'766071	0'718354	0'794253
26 years old	0'905727	0'918482	0'870765	0'946664	0'776963	0'789718	0'742001	0'8179

Table 12. Source: Own Elaboration

Estimated values for the response variable Y for the control individuals according to ″the differences estimator with addtional regressors″ model

	CONTROL INDIVIDUALS							
	Male				Female			
	Zone 1	Zone 2	Zone 3	Zone 4	Zone 1	Zone 2	Zone 3	Zone 4
17 years old	0'283731	0'296486	0'248769	0'324668	0'154967	0'167722	0'120005	0'195904
18 years old	0'307378	0'320133	0'272416	0'348315	0'178614	0'191369	0'143652	0'219551
19 years old	0'331025	0'34378	0'296063	0'371962	0'202261	0'215016	0'167299	0'243198
20 years old	0'354672	0'367427	0'31971	0'395609	0'225908	0'238663	0'190946	0'266845
21 years old	0'378319	0'391074	0'343357	0'419256	0'249555	0'26231	0'214593	0'290492
22 years old	0'401966	0'414721	0'367004	0'442903	0'273202	0'285957	0'23824	0'314139
23 years old	0'425613	0'438368	0'390651	0'46655	0'296849	0'309604	0'261887	0'337786
24 years old	0'44926	0'462015	0'414298	0'490197	0'320496	0'333251	0'285534	0'361433
25 years old	0'472907	0'485662	0'437945	0'513844	0'344143	0'356898	0'309181	0'38508
26 years old	0'496554	0'509309	0'461592	0'537491	0'36779	0'380545	0'332828	0'408727

5.- CONCLUSIONS.

The training schooll program's average effect estimed by weighting observations by the inverse of a nonparametric estime of the propensity score let us to conclude that, for treated, the time needed to find a job is reduced in 471 days. As the program was designed to improved the employ between youngers unemployed; this results supports the effectiveness of this public policy.

The training schooll program's average effect estimed by the differences estimator let us to conclude that, for treated, the time needed to find a job is reduced in 448 days. This result also supports the effectiveness of this public policy.

Another conclusion can be obtained from using the covariates information contained in the BASEVAFOR 96-03. More specifically, in the case of males, on average, the period needed for a treated for find a job, is reduced in 141 days. By considering age, the same period is reduced in 26 days per year from 16 to 25 years old.

According to the obtained results, the effectiveness of the training program from the "differences estimator with additional regressor" is positive too.

Both evaluations show evidence of that this program can reduce the time needed to find a job.

Further investigations might improve conclusions if public authorities let researchers to extend the database information with data related with others individual characteristics.

REFERENCES.

BARNOW, B., G. CAIN and A. GOLDBERGER (1980): "Selection on Observables", en Stromsdorfer, E. W. and Farkas, G. (eds.), *Evaluation Studies Review Annual*, vol. 5: 43-59, Beverly Hills, California, Ed.: Sage Publications.

BASSI, L. J. (1984): "Estimating the Effect of Training Programs with Non-Random Selection", *Review of Economics and Statistics*, 66 (1): 36-43.

BILLEWICZ, W. Z. (1965): "The Efficiency of Matched Samples: an Empirical Investigation", *Biometrics,* 21 (3): 623-644.

BURTLESS, G. (1995): "The Case for Randomized Field Trials in Economic and Policy Research," *Journal of Economic Perspectives*, 9 (2): 63-84.

CAMERON, E. and L. PAULING (1976): "Supplemental Ascorbate in the Supportive Treatment of Cancer: Prolongation of Survival Times in Terminal Human Cancer", *Proceedings of the National Academy of Sciences of the United States of America,* 73 (10): 3685-3689.

CAMERON, A. C. and P. K. TRIVEDI (2005): *Microeconometric. Methods and Applications*, New York, Ed.: Cambridge University Press.

CANSINO, J.M., M. A. CARDENETE and R. ROMAN: (2007): "Regional evaluation of a tax on retailer sales of some fuels through social accounting matrix", *Applied Economics Letters*, 14 (12): 877-880.

CANSINO, J.M. y A. SANCHEZ (2008a): "Evaluación del programa de Escuelas Taller y Casas de Oficios a partir de su efecto sobre el tiempo de búsqueda del primer empleo. El caso de Sevilla", *Estudios de Economía Aplicada*, *forthcoming*.

CANSINO, J.M. y A. SANCHEZ (2008b): "Average effect of training programs on the time needed to find a job. The case of the training programs schools in the south of Spain (Seville, 1997-1999)". *Funcas Workpapers*. Available in www.funcas.es.

COCHRAN, W. G. (1968): "The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies", *Biometrics,* 24 (2): 295-313.

COX, D. R., (1992): "Causality: Some Statistical Aspects," *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 155 (2): 291-301.

DAWID, A. P. (2000): "Causal Inference Without Counterfactuals", *Journal of the American Statistical Association*, 95 (2): 407-448.

DAWID, A. P. (1979): "Conditional Independence in Statistical Theory", *Journal of the Royal Statistical Society, Series B (Statistics Methodological)*, 41 (1): 1-31.

DEHEJIA, R. H. and S. Wahba (1999): "Causal Effects in Non-Experimental Studies: Reevaluating the Evaluation of Training Programs", *Journal of the American Statistical Association*, 94 (448): 1053-1062.

FISHER, R. A. (1935): *The design of experiments*, Edinburgh, Ed.: Oliver and Boyd.

FISHER, R. A. (1928): *Statistical Methods for Research Workers*, 2ª edición, London, Ed.: Oliver and Boyd.

HAHN, J. (1998): "On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects", *Ecometrica*, 66 (2): 315-331.

HECKMAN, J. J (1990): "Varities of Selection Bias", *American Economic Review*, Papers and Proceedings of the Hundred and Second Annual Meeting of the American Economic Association, 80 (2): 313-338.

HECKMAN, J. J. (1979): "Sample Selection Bias as a Specification Error", *Econometrica*, 47 (1): 153-162.

HECKMAN, J., N. CLEMENTS and J. SMITH (1997): "Making the Most out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts" *Review of Economic Studies*, 64 (4): 487-535.

HECKMAN, J. J. and V. J. HOTZ (1989): "Choosoing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training", *Journal of the American Statistical Association*, 84 (408): 862-874.

HECKMAN, J. J., H. ICHIMURA, J. SMITH and P. E. TODD (1998): "Characterizing Selection Bias Using Experimental Data", *Econometrica*, 66 (5): 1017-1098.

Heckman, J. J., H. ICHIMURA and P. E. TODD (1998): "Matching As an Econometric Evaluation Estimator", *Review of Economics Studies*, 65 (2): 261-294.

HECKMAN, J. J., H. ICHIMURA and P. E. TODD (1997): "Matching As an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme", *Review of Economics Studies*, 64 (4): 605-654.

HIRANO, K., G. Imbens and G. Ridder, (2003): "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score", *Econometrica*, 71(4): 1161-1189.

HOLLAND, P. W. (1986): "Statistics and Causal Inference" (with discussion), *Journal of the American Statistical Association*, 81 (396): 945-970.

IMBENS, G. W. (2004): "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review", *Review of Economics and Statistics*, 86 (1): 4-29.

KIEFER, N. (1979): *Economic Benefits from Four Manpower Training Programs,* Garland Series of Outstanding Dissertations in Economics, New York,. Ed.: Garland Press.

LITTLE, R. and D. RUBIN (1987): *Statistical Analysis with Missing Data*, Nueva York, Ed.: Wiley.

MANSKI, C. F. (2001): "Designing programs for heterogeneous populations: The value of covariate information, *American Economic Review*, 91 (2): 103-106.

MOERTEL, C., T. FLEMING, E. CREAGAN, J. RUBIN, M. O'CONNELL and M. AMES (1985): "High-dose Vitamin C Versus Placebo in the Treatment of Patients with Advanced Cancer Who Have Had no Prior Chemotherapy: a Randomized Double-blind Comparison", *New England Journal of Medicine,* 312: 137-141.

NEYMAN, J. (1935): "Statistical Problems in Agricultural Experimentation", *Supplement to the Journal of the Royal Statistical Society,* 2: 107-180.

NEYMAN, J. (1923): "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles.", re-edited en *Statistical Science* (con discusión), 1990, 5 (4): 465-472.

PEARL, J. (2000): *Causality: Models, Reasoning and Inference*, Cambridge, Ed.: Cambridge University Press.

ROSENBAUM, P. R. (1999): "Choice As an Alternative to Control in Observational Studies" (with discussion), *Statistical Science*, 14 (3): 259-304.

ROSENBAUM, P. R. (1996): "Observational Studies and Nonrandomized Experiments", en S. Ghosh and Rao, C. R. (eds.), *Handbook of Statistics*, vol. 13, capítulo 6: 1277-1366, New York, Ed.: Elsevier.

ROSENBAUM, P. R. (1995): *Observational Studies*, Springer Series in Statistics, Nueva York, Ed.: Springer-Verlag.

ROSENBAUM, P. R. and D. B. RUBIN (1984): "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score", *Journal of the American Statistical Association*, 79 (387): 516-524.

ROSENBAUM, P. R. and D. B. RUBIN (1983): "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrica*, 70 (1): 41-55.

ROY, A. (1951): "Some Thoughts on the Distribution of Earnings", *Oxford Economic Papers*, 3 (2): 135-146.

RUBIN, D. B., (1978): "Bayesian Inference for Causal Effects the Role of Randomization," *Annals of Statistics*, 6 (1): 34-58.

RUBIN, D. B. (1976): "Inference and Missing Data", *Biometrika*, 63 (3): 581-592.

RUBIN, D. B. (1974): "Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies", *Journal of Educational Psychology*, 66 (5): 688-701.

STOCK, J. M. and WATSON M. W (2003): *Introduction to Econometrics*, Boston, MA. Ed.: Addospm Wesley.

RETOS Y OPORTUNIDADES DE LOS ENTORNOS VIRTUALES DE APRENDIZAJE EN EL MARCO UNIVERSITARIO

LEIRE URCOLA CARRERA
E-mail: leire.urcola@ehu.es
AMAIA ALTUZARRA ARTOLA
E-mail: amaia.altuzarra@ehu.es
Departamento de Economía Aplicada
Escuela Universitaria de Estudios Empresariales
Universidad del País Vasco

Resumen

La presente comunicación pretende transmitir algunos elementos para la reflexión sobre la importancia creciente de adoptar nuevos enfoques de enseñanza-aprendizaje soportados en entornos virtuales en el contexto de las universidades tradicionales, y más concretamente, en el sistema universitario español.
Somos conscientes de que en la Sociedad de la Información y el Conocimiento, y en el ámbito universitario en particular, la adopción de paradigmas de aprendizaje desde una óptica abierta y constructivista del conocimiento, es un proceso complejo que debe hacer frente a importantes barreras tecnológicas, administrativas y humanas. Por ello, para poder plantear planes de intervención en nuestras universidades, resulta del todo necesario analizar en primer lugar, el estado del arte en el que se encuentran algunos indicadores relativos a los entornos virtuales de enseñanza-aprendizaje, de forma que nos permita identificar los principales obstáculos y plantear acciones para su plena integración en contextos educativos tradicionales.
Los apartados principales de la presente comunicación se estructuran de la siguiente manera. La primera parte expone las características de los entornos virtuales de enseñanza-aprendizaje y los principios de actuación a tener en cuenta en su implantación. La segunda parte presenta el panorama actual en el que se encuentran las universidades españolas, y se identifican las principales barreras y por lo tanto los retos a los que deben hacer frente las instituciones universitarias para la integración efectiva de los nuevos entornos de enseñanza-aprendizaje.
Por último, se ofrecen algunas tendencias e iniciativas en marcha en el ámbito nacional e internacional, así como las recomendaciones recogidas en los últimos informes publicados en esta materia.

Palabras Clave: e-learning, blended learning, entornos virtuales de enseñanza-aprendizaje, EVA.
Área Temática: Didáctica de la Economía y Metodología.

1. LAS PLATAFORMAS VIRTUALES UNA OPORTUNIDAD EN EL MARCO DEL EEES

La internacionalización y la globalización de la educación iniciadas tras el proceso de Bolonia ha originado en las propias universidades la necesidad de incrementar la calidad de sus procesos docentes, investigadores y administrativos. Los objetivos del nuevo Espacio Europeo de Educación Superior y las necesidades de la sociedad actual ponen especial atención en la capacidad de innovación y emprendizaje de los futuros graduados. El aprendizaje a lo largo de toda la vida, el "aprender haciendo", así como el aprendizaje colaborativo y activo constituyen aproximaciones fundamentales para alcanzar estos objetivos.

En torno al crédito europeo subyace la idea de que su puesta en práctica, debido a sus características organizativas (menor tiempo de clases teóricas y más de trabajo autónomo del alumno) impondrá un cambio metodológico en las formas de impartir la enseñanza y, en consecuencia, en las formas y procesos de aprendizaje. Esta innovación debiera consistir en

poner en práctica un modelo didáctico y de aprendizaje bien diferenciado del que en estos momentos está generalizado en las aulas universitarias tradicionales. Frente a la comunicación unidireccional que se estable entre el profesor y el alumno en la docencia tradicional, la metodología de enseñanza supuestamente impulsada por los ECTS asume muchos de los principios pedagógicos del llamado constructivismo (Carretero, 1993; Coll, 1996). La investigación psicopedagógica pone de relieve que el auténtico aprendizaje se produce cuando el alumno "construye" el conocimiento a través de la interacción con su entorno, a través de la búsqueda personal de información y mediante procesos de resolución constructiva de problemas (Rosales, 2001).

Este nuevo modelo abierto y constructivista de la enseñanza implica mantener los siguientes principios de actuación:

- Favorecer el trabajo en equipo y la colaboración.
- Favorecer el aprendizaje autónomo.
- Favorecer el desarrollo de destrezas y habilidades prácticas de aplicación del conocimiento conceptual (análisis de casos y situaciones reales del entorno).
- Utilizar diversos medios, tecnologías y formas de representación de la información con especial atención a los recursos informáticos.
- Favorecer y propiciar el análisis, debate y la crítica del conocimiento disponible.
- Atender el ritmo individual de aprendizaje de cada alumno mediante la tutorización y seguimiento de la marcha y la participación en las actividades desarrolladas en el marco de la asignatura.
- Utilizar distintas técnicas y procedimientos de evaluación continua del aprendizaje e informar de los resultados ofreciendo recomendaciones de mejora (feedback).
- Combinar distintos tipos de estrategias y técnicas de enseñanza como son los seminarios, las actividades prácticas, simulaciones o juegos de rol, análisis de casos, resolución de problemas, etc.
- Elaborar materiales didácticos de calidad propios o reutilizar otros ya creados, que además del contenido específico de la materia de estudio, incorporen propuestas de actividades prácticas.

Sin embargo, es obvio que algunos de estos principios se hacen difíciles de implementar debido a una serie de razones, como pueden ser: la existencia de un elevado número de alumnos, la configuración arquitectónica de las aulas, el ratio profesor/alumno. De ahí que aboguemos hacia un modelo de enseñanza semipresencial, o expresado en otros términos, una modalidad de enseñanza que combine la actividad docente presencial en las aulas, con el desarrollo de un aprendizaje autónomo y a distancia por parte del alumno a través de plataformas o aulas virtuales. Esta combinación de diferentes tareas de aprendizaje que el alumno tiene que realizar apoyándose en recursos diversos y utilizando espacios variados recibe el nombre de aprendizaje semipresencial, mixto o *Blended learning* (Garrison y Kanuka, 2004), el cual es una variante del concepto de e-Learning. Según diversos autores, esta modalidad educativa que facilita un aprendizaje más flexible, variado y adaptado a los propios ritmos del estudiante, será la que se generalice en los próximos años en la educación superior.

Con relación a los entornos virtuales de aprendizaje, *New Media Consortium* publicó en 2005 un informe sobre las tecnologías emergentes con mayor impacto en el ámbito educativo con un horizonte temporal de tres a cinco años. En dicho informe incluyó a las redes sociales unidas a las webs de conocimiento o espacios virtuales en los que se construye de forma colaborativa una base común de referencias y contenidos teóricos. Los entornos virtuales de enseñanza-aprendizaje presentan características específicas que permiten favorecer:

- el autoaprendizaje,
- el desarrollo de una enseñanza flexible,
- el desarrollo de competencias transversales y específicas

- la construcción del conocimiento por parte del estudiante,
- la colaboración entre grupos,
- el aprovechamiento de las ventajas de las redes telemáticas proporcionando acceso a grupos de personas y recursos compartidos,
- la búsqueda de datos colaborativa y la indexación de la información,
- la utilización de múltiples plataformas en función de las necesidades del creador, del receptor y del contexto.

Teniendo presente las anteriores características, los objetivos pedagógicos perseguidos a la hora de diseñar una asignatura apoyada en una plataforma virtual deberían orientarse a:

- Encontrar un equilibrio entre la autonomía del alumno y la flexibilidad de su aprendizaje.
- Desarrollar competencias transversales y específicas.
- Facilitar la interacción entre alumnos de diferentes universidades.
- Cimentar el aprendizaje en la resolución de problemas, análisis de casos reales y el espíritu crítico.
- Implicar al alumno en el proceso de evaluación mediante el desarrollo de una evaluación continua y una autoevaluación.
- Aprovechar las Tecnologías de la Información para facilitar el acceso a material de toda índole relacionados con su aprendizaje, estén o no dentro de su entorno universitario.

No cabe duda de que las metodologías que hemos venido utilizando hasta la fecha han de adaptarse a los cambios que se están produciendo en la sociedad, a los nuevos entornos, a las necesidades de las empresas, a las demandas de los estudiantes, más concretamente, a un tipo de estudiante que ya no sólo atiende una clase magistral, sino que debate, opina, critica. El perfil de los estudiantes de ahora poco tiene que ver con el de décadas pasadas. Las generaciones actuales de estudiantes son *nativos digitales*, han nacido con Internet, están experimentando la transición de la anterior generación web hacia la nueva filosofía de red (participativa, colaborativa y semántica) y, de cara al futuro no concebirán un espacio virtual que no les permita interactuar, crear y compartir conocimiento.

Por ello, es fundamental introducir y expandir una nueva cultura de conocimiento basado en usuarios activos que sean capaces de crear, buscar, analizar, comunicar y compartir información y conocimiento. En este sentido, consideramos que el *Blended learning* constituye un entorno estratégico que facilita la adopción del nuevo paradigma pedagógico basado en los aprendizajes.

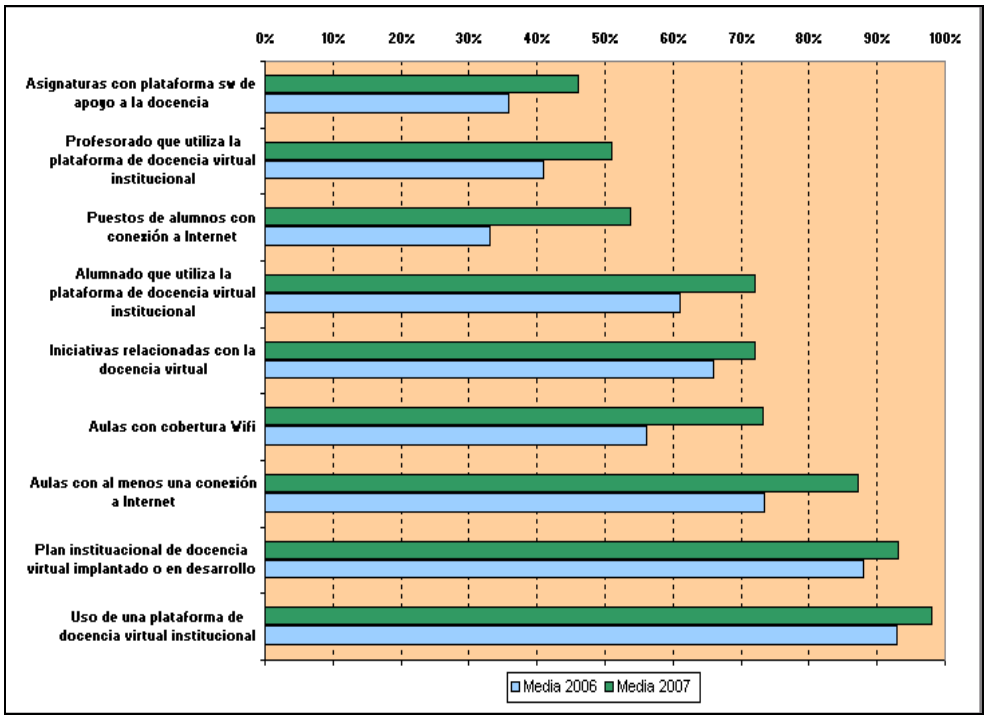
2. ESTADO DEL ARTE DEL E-LEARNING EN LAS UNIVERSIDADES ESPAÑOLAS

En relación con este aspecto, los resultados del último informe publicado por la Comisión Sectorial TIC de la Conferencia de Rectores de las Universidades Españolas[1] (UNIVERSITIC, 2008) constatan que cada vez se está extendiendo más el uso de tecnologías propias de docencia virtual, consolidándose además, las iniciativas de promoción de este tipo de docencia.

En el citado informe, tal y como queda reflejado en el Gráfico 1, se destacan importantes tendencias positivas respecto al año 2007 en determinados indicadores que se articulan en el Eje Estratégico de "Enseñanza-Aprendizaje".

[1] El Grupo de Trabajo de las TIC de la CRUE ha contado para realizar el estudio con la colaboración de 52 universidades, que supone casi el 90% de los alumnos del Sistema Universitario Español.

Gráfico 1: Indicadores del eje estratégico "Enseñanza-Aprendizaje"



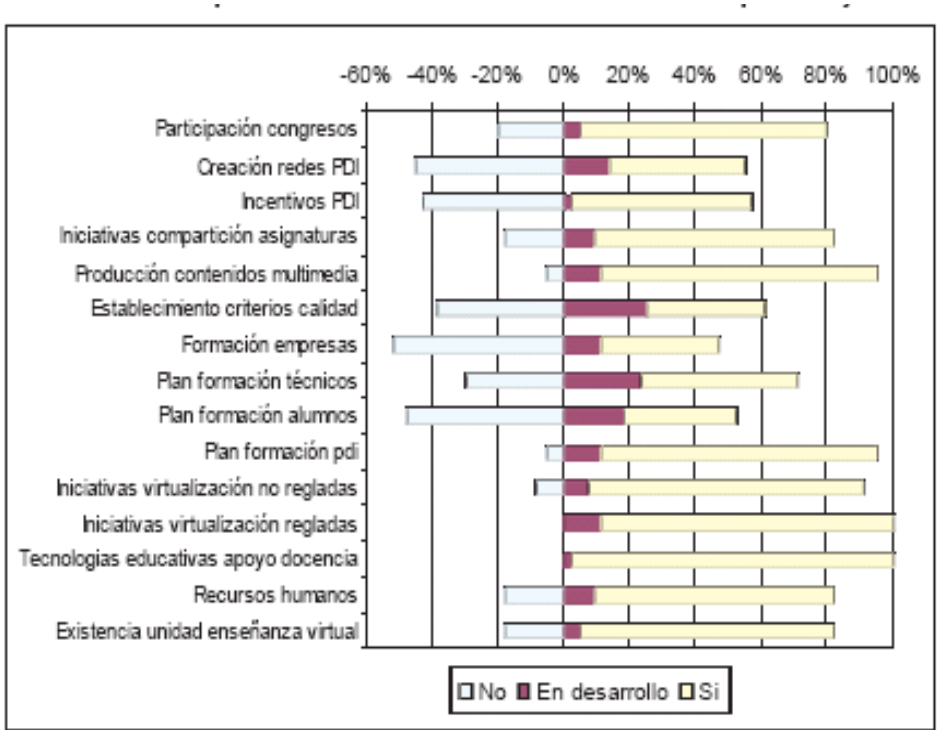
Fuente: Elaboración propia a partir de los datos del informe (UNIVERSITIC, 2008)

Con respecto a los indicadores más estrechamente vinculados con la docencia virtual, del Gráfico 1 podemos destacar las siguientes conclusiones. En primer lugar, tal y como queda reflejado, la práctica totalidad de las universidades españolas ya disponen de una plataforma institucional de docencia virtual. Asimismo, cabe señalar que un porcentaje muy elevado de las universidades (93%) posee un plan institucional de docencia virtual, bien ya implantado (79%) o bien en desarrollo (14%). En cuanto al uso de plataformas virtuales, puede afirmarse que algo más de la mitad del profesorado docente e investigador emplean la plataforma de docencia virtual, mientras que este porcentaje es muy superior en el caso de los estudiantes (82%). En términos generales, la mayoría de los indicadores correspondientes al Eje estratégico de Enseñanza-Aprendizaje presentan una importante mejora con respecto a la situación del año anterior.

Atendiendo a indicadores más específicos del estado actual del e-Learning en el sistema universitario español, el Gráfico 2 presenta los siguientes resultados.

Entre los aspectos positivos a destacar puede afirmarse que la práctica totalidad de las universidades españolas ya disponen de iniciativas para virtualizar asignaturas regladas y, además, realizan apoyo de docencia presencial mediante tecnologías propias de docencia virtual. Asimismo, las universidades españolas presentan un elevado porcentaje de iniciativas relacionadas con la docencia virtual (que se encuentran ya implantadas o en desarrollo) sobre el total de iniciativas consideradas como importantes por la Sectorial TIC de la CRUE.

Gráfico 2: Estado actual de indicadores de e-Learning



Fuente: "Las TIC en el Sistema Universitario Español" (UNIVERSITIC, 2008)

Sin embargo, tal y como queda reflejado en el gráfico 2, existen indicadores en los que es necesario avanzar más, como son en la asignación de incentivos y reconocimientos para el PDI que emplea estos entornos, el desarrollo de de planes de formación, el desarrollo de estándares y criterios de calidad para elaborar y publicar en la red materiales didácticos y científicos, así como la creación de redes interuniversitarias de colaboración e intercambio de experiencias, información y conocimiento.

Estos resultados vienen a corroborar las conclusiones del informe sobre el desarrollo del e-Learning en las universidades europeas, que pone de manifiesto que aunque todas las universidades europeas están acometiendo la incorporación de las TIC a sus servicios, en la actualidad sigue sin darse respuesta adecuada al reto de su integración pedagógica en la enseñanza superior (Pls Rambol Management, 2004). En este sentido, el último informe publicado por la Conferencia de Rectores de las Universidades Españolas refleja un claro desfase entre el nivel de implantación y el uso efectivo de los entornos virtuales de enseñanza-aprendizaje el sistema universitario español (UNIVERSITIC, 2008).

Tal y como se señala en el citado informe, si bien la integración de las TIC como herramienta de apoyo a la docencia se ha incrementado notablemente en los últimos años, su utilización aún se realiza bajo los clásicos entornos de docencia unilateral. Luego si comparamos nuestro sistema universitario tomando como referencia a universidades punteras en este terreno, como pueden ser algunas universidades norteamericanas y europeas (MIT, Stanford, Georgia Tech, Yale, Princeton, Oxford, Carnegie Mellon, por citar algunas) podría afirmarse que nos encontramos todavía en una etapa embrionaria en la utilización de los entornos virtuales de enseñanza-aprendizaje.

Como sabemos, el uso de la tecnología no garantiza por sí solo ningún cambio pedagógico. Muchas aulas virtuales de universidades e instituciones presenciales únicamente utilizan el medio de transmisión como un fin en sí mismo y no aprovechan las verdaderas posibilidades de estos nuevos entornos docentes. En la actualidad estas herramientas se utilizan mayoritariamente como soporte a las actividades presenciales, y tal como constatan diversos estudios realizados en esta línea, el uso que el profesorado hace de estas herramientas responde a una utilización más instrumental que didáctica y más relacionada con su dimensión investigadora que docente (Barro, 2004; Sangrá y González Sanmamed, 2004; Alba y Carballo,

2005). En este ámbito, algunos profesores comienzan a diseñar y desarrollar actividades *on line*, especialmente aquellas que por falta de tiempo y espacio no se pueden realizar en las aulas y laboratorios. Otro uso habitual consiste en emplearlas como apoyo en las tareas administrativas docentes: gestión de trabajos, evaluación de trabajos y distribución de calificaciones, calendarios, plazos, etc. Como regla general, puede afirmarse que el profesor comienza realizando lo mismo que hace presencialmente y, poco a poco, va explorando las nuevas posibilidades que le ofrece una plataforma virtual. Las herramientas menos utilizadas por el momento, son los recursos integrados en las plataformas virtuales que requieren una mayor elaboración didáctica previa, tales como: talleres, ejercicios de autoevaluación y otros módulos, como blogs, wikis, base de datos, portfolios, etc., pero precisamente es en estos recursos donde reside el verdadero potencial que ofrecen las plataformas virtuales y las nuevas posibilidades de su utilización.

En este escenario, puede afirmarse que las universidades no están realizando los esfuerzos suficientes para adaptarse a las nuevas demandas de una "sociedad en red". La falta de adaptación de nuestras instituciones puede explicarse por los obstáculos y temores con los que se enfrenta en la adopción del nuevo paradigma educativo basado en los aprendizajes. Una de las principales barreras a las que se debe hacer frente es la resistencia al cambio. Las universidades tradicionales son conservadoras y, en algunos casos, muestran una cultura de aversión hacia la innovación.

Es lógico pensar que como ocurre con cualquier innovación en general, existe un periodo inicial de resistencia; y en este caso, en particular, al tratarse de una innovación pedagógica con apoyo de medios tecnológicos y telemáticos, se produce una resistencia y un excepticismo hacia su empleo. Es evidente que la tecnología por sí misma no va a mejorar la productividad, o expresado en otros términos, el uso de plataformas virtuales de aprendizaje no implica necesariamente un incremento positivo en el rendimiento de los estudiantes, ni en la calidad de la docencia, ni en la investigación. No cabe duda de que los nuevos enfoques de aprendizaje requerirán de cambios metodológicos y organizacionales para mejorar la productividad.

Somos conscientes de que la adopción de un enfoque del aprendizaje desde una naturaleza abierta y constructiva en las universidades es un proceso complejo que debe hacer frente a importantes barreras tecnológicas, administrativas y humanas. Por eso, esta introducción debe realizarse a través de una estrategia adaptativa, que en su diseño pueda integrar experiencias previas de organizaciones educativas, de investigación y empresariales (Freire, 2008).

En este marco, podrían plantearse los siguientes retos en el progreso hacia la integración efectiva de los nuevos enfoques de aprendizaje soportados en entornos virtuales:

- Introducir de forma explícita las TIC dentro de los planes de estratégicos de las universidades, trasladando su utilización desde iniciativas individuales hasta llegar a ser un ingrediente docente básico (perspectiva *bottom-up*).
- Desarrollar estructuras profesionales de apoyo que sustenten los procesos de integración de las TIC en el rediseño de cursos. Más concretamente. Crear la figura de orientador de profesorado universitario, especialista en formación que entienda el uso de las herramientas tecnológicas como un uso didáctico, que integre el aprendizaje técnico-informático a través de su aplicación en el aprendizaje técnico-docente/didáctico. Incentivar desde el ámbito universitario la innovación docente e impulsar el empleo de procesos de evaluación y acreditación de la calidad en la docencia soportada con plataformas virtuales.
- Participar en las iniciativas europeas orientadas al desarrollo de estándares en el ámbito de los recursos digitales de enseñanza.
- Fomentar a nivel europeo, el desarrollo de estándares comunes de acreditación, evaluación y certificación de los alumnos bajo el nuevo entorno de formación y movilidad virtual.
- Desarrollar redes que faciliten el intercambio de experiencias y buenas prácticas entre universidades.
- Regular los derechos de propiedad intelectual y los sistemas de licencias para fomentar la reutilización y el intercambio de recursos digitales.

2.1 Desarrollo de campus virtuales en el sistema universitario español

Una vez analizado el estado actual de los indicadores de e-Learning en el sistema universitario español, en el presente apartado vamos a presentar el panorama actual de la oferta de e-Learning en nuestras instituciones universitarias.

En la actualidad, la práctica totalidad (90%) de las universidades españolas ofertan formación a través de algún Campus Virtual (distribuidos en 19 Comunidades Autónomas), existiendo un total de 67 Campus Virtuales, 60 pertenecientes a una sola universidad y 7 integrados por varias universidades. El panorama actual se caracteriza por una gran oferta de cursos, aunque muy dispersa, que principalmente corresponden a la formación de nivel de postgrado.

Los campus virtuales están concentrados principalmente en las Comunidades Autónomas de Madrid (16%), Cataluña (15%), Andalucía (13,6%) y Valencia (9,1%). Llama la atención que en la Comunidad Autónoma de Castilla y León (cuarta región con mayor número de universidades), tan sólo algo más de un tercio de ellas dispongan de un Campus Virtual. Consideramos que este dato requeriría un análisis diferenciado.

Por su parte, atendiendo al carácter público o privado de las instituciones, hay que señalar que existe un mayor desarrollo de los servicios virtuales en las universidades públicas respecto a las privadas.

Los Campus Virtuales existentes en el sistema universitario español se pueden clasificar en tres principales modelos organizativos:

- *Campus Virtual de una universidad presencial*: en este caso, la mayoría de los Campus ofrecen enseñanza virtual como apoyo a la presencial. Este sería el modelo predominante en el sistema universitario español.

- *Universidad Virtual*: este sería el caso de los Campus de las universidades originalmente constituidas como centros de formación a distancia. Como referentes en España hay que mencionar a la Universidad Nacional de Educación a Distancia (UNED) y a la Universitat Oberta de Catalunya (UOC). Además, hay que señalar la creación en 2007 de la Universidad a Distancia de Madrid (UDIMA) de carácter privado.

- *Campus Virtual Interuniversitario:* se trata de espacios virtuales compartidos por varias instituciones. Estos espacios se basan en disponer de un único punto de encuentro virtual entre varias universidades. Algunos ejemplos de referencia en España:
  - Grupo G9: asociación formada por nueve universidades públicas (Cantabria, Castilla La Mancha, Extremadura, Islas Baleares, La Rioja, Navarra, Oviedo, País Vasco y Zaragoza) que realizan una oferta conjunta de asignaturas de libre elección impartidas a través de medios telemáticos. Constituye el grupo de universidades españolas en red con mayor número de estudiantes.
  - "ADA Madrid" (Aula a Distancia y Abierta): proyecto formado por seis universidades de la Comunidad de Madrid: Alcalá, Autónoma de Madrid, Carlos III, Complutense, Politécnica, Rey Juan Carlos I.
  - Intercampus: red formada por universidades de la Comunidad de Cataluña, como la Universidad Autónoma de Barcelona, Girona, Lleida, Oberta, Politécnica, Pompeu Fabra, Rovira i Virgili.
  - Instituto Universitario de Postgrado (IUP): oferta fundamentalmente masters y está formado por la Universidad Carlos III, Autónoma de Barcelona y la Universidad de Alicante, en colaboración con el grupo Santillana Formación.
  - Instituto de Formación Continua (IL3)

- Instituto AEU (Altos Estudios Universitarios): ofertan estudios de postgrado a distancia. Participan las universidades de Alcalá, Barcelona, Granada, León y el Instituto de Formación Continua (IL3).
- Campus Virtual de las universidades andaluzas.
- Universia: portal universitario de habla hispano-portuguesa.

En esta línea, es importante destacar también los proyectos de colaboración Universidad-Empresa que tienen por objetivo ofrecer a profesionales del ámbito empresarial una variedad de cursos en algunas de las más prestigiosas universidades y escuelas de negocios españolas, como son la Universitat Oberta de Catalunya, la Universidad Europea de Madrid, el IL3-Universidad de Barcelona y la Escuela de Administración de Empresas, entre otras. Como ejemplo de esta colaboración se encuentran, entre otros, el desarrollado por la Universidad de Málaga, en colaboración con la empresa Ingenia, así como la iniciativa del Aula virtual de la Fundación Universidad-Empresa de la Universidad de Valencia (ADEIT), que nace con la filosofía de ayudar y asesorar al mundo empresarial en el desarrollo e integración de las TIC.

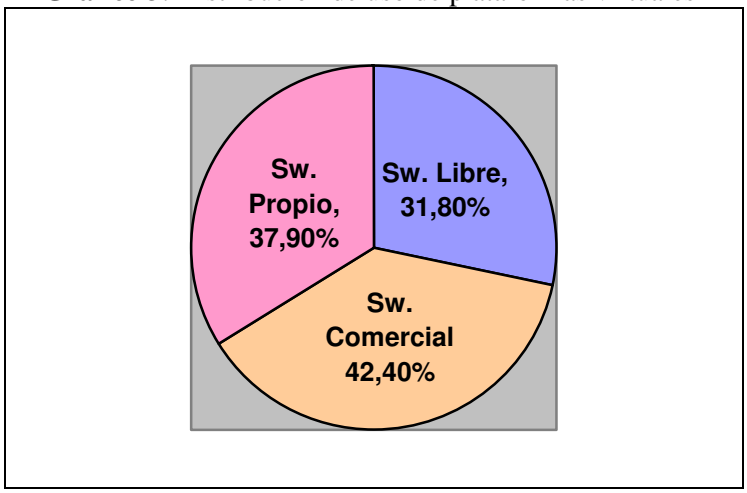
En otro orden de cosas, en cuanto al uso de plataformas virtuales, los sistemas de gestión de aprendizaje (LMS) que vienen utilizándose mayoritariamente en las universidades españolas desde hace una década aproximadamente, corresponden a software comercial o a desarrollos propios. Hay que señalar que las universidades han realizado, en muchos casos, grandes inversiones económicas para desarrollar software institucional o para comprar plataformas de software comercial. Sin embargo, de acuerdo con el profesor Juan Freire, esta infraestructura podría convertirse hoy en día, más que en un activo, en una barrera, ya que la mayoría de estas herramientas podrían empezar a atravesar en poco tiempo por una fase de obsolescencia acelerada, ya que necesitan configurarse, integrarse y combinarse con nuevas herramientas que están disponibles de forma gratuita o con un bajo coste en el mercado, para crear nuevas aplicaciones adaptadas a las necesidades personales de los usuarios (Freire, 2008). Con esto queremos señalar que en la sociedad "red 2.0" el concepto tradicional de TICs ya no supone una ventaja competitiva para las organizaciones (aunque de acuerdo con Nicholas Carr (2004) en su controvertido artículo "IT doesn't Matter" quien no disponga de ellas claramente se encontrará en desventaja). Y esto es aplicable a las plataformas virtuales de teleformación.

Hoy en día, casi todas las universidades disponen de un Campus Virtual y la opción de productos LMS de software libre está cogiendo fuerza, estableciéndose una positiva tendencia hacia este tipo de soluciones. Este es el caso de Moodle, plataforma de código libre y gratuita que cuenta con un fuerte soporte internacional y lleva camino de convertirse en un estándar de plataforma educativa virtual. Moodle ha sido diseñado para dar soporte a las nuevas metodologías de aprendizaje basadas en la teoría del constructivismo social, centrándose en ofrecer un entorno que facilite la colaboración, el intercambio de conocimiento y la generación de nuevas ideas.

En España, entre las universidades precursoras del uso de la plataforma Moodle se encuentran las universidades de Las Palmas de Gran Canaria y la Universidad Jaume I, quienes adoptaron en el 2004 Moodle institucionalmente. A partir de ahí, se han ido sumando varias universidades tanto públicas como privadas. Ahora bien, aunque existe una tendencia hacia la utilización de plataformas de código abierto o al menos a utilizarlas junto a otras de desarrollo propio o comercial, tal y como se ha señalado anteriormente, en la actualidad predomina el uso de plataformas comerciales y de desarrollo propio (ver Gráfico 3).

Existen algunas razones que están dificultando su plena adopción a nivel institucional. Quizás, una de las razones resida en que los bajos costes en numerosas ocasiones generan desconfianza en la toma de decisiones, y en la incomprensión que todavía hoy continúan produciendo conceptos como el de software libre o el código abierto.

Gráfico 3: Distribución de uso de plataformas virtuales[2]



Fuente: elaboración propia

Hay que señalar que entre los productos comerciales empleados, el 80% corresponde a la plataforma WebCT, mientras que entre el software de código abierto la plataforma dominante es Moodle, en el 60% de los casos.

No es el objetivo de esta ponencia entrar en el debate sobre la idoneidad en cuanto a la utilización de una plataforma u otra, existen varios organismos y proyectos que llevan a cabo comparaciones y evaluaciones de plataformas virtuales[3] (LMS o CMS). No obstante, aunque pueda parecer obvio, conviene señalar que la elección de un producto determinado dependerá de las circunstancias particulares de cada institución académica, por lo que dicha evaluación debería realizarse teniendo en cuenta los distintos escenarios o situaciones específicas.

3. TENDENCIAS E INICIATIVA EN MARCHA

El fenómeno e-Learning, como producto de los cambios que se avecinan en el panorama educativo, ha venido configurando gran variedad de proyectos, y también han surgido numerosas asociaciones y observatorios relacionados con esta modalidad educativa que sirven de punto de encuentro e intercambio de experiencias sobre el uso de las plataformas virtuales como apoyo a la formación, así como para investigar sobre las últimas tendencias en innovación educativa y proponer líneas de acción y estándares de calidad al objeto de regular la calidad educativa y técnica.

A nivel internacional, la UNESCO, a través de la Conferencia Mundial de Educación Superior, ya expresó en la declaración de 1998 la necesidad de extender las TIC "como herramienta para democratizar la educación", proponiendo para ello el e-Learning como herramienta para poner el conocimiento al alcance de todo el mundo. Prueba de ello ha resultado por ejemplo la transformación de la Universidad de las Naciones Unidas en una universidad virtual (www.unu.edu).

En esta misma línea se ha venido posicionando la Unión Europea adoptando un programa propio para el desarrollo de la educación virtual en los distintos países miembros[4], y la

[2] Hay que tener en cuenta que las universidades no emplean una determinada plataforma de forma exclusiva, pudiéndose encontrar universidades que utilicen plataformas de código abierto, plataformas de desarrollo propio y/o plataformas comerciales, en función de la tipología de cursos ofertados y el perfil de estudiante al que se dirigen.

[3] A nivel internacional, el grupo edutech (responsable del campus virtual suizo http://www.edutech.ch/lms/ev3) realiza evaluaciones de varios sistemas de gestión LMS. En España, también existen algunas iniciativas en este sentido, como la de la Universidad de Vigo, que lleva a cabo el análisis de los escenarios reales de uso de los sistemas LMS de software libre en las universidades españolas a través del proyecto JOIN! – *Guideline 'e-learning at Spanish universitities'*; proyecto financiado por la Unión Europea dentro de una iniciativa de divulgación sobre plataformas virtuales de software libre.

[4] www.elearningeuropa.info

Comisión Europea lo ha impulsado implantando las TIC en materia educativa[5]. La iniciativa e-Learning fue emprendida por la Comisión Europea en mayo de 2000, dentro del marco del Espacio Europeo de Educación Superior para concebir la "educación del futuro"[6]. Dentro de los objetivos propuestos por los programas Sócrates y Leonardo da Vinci ya se incluía el uso de las TIC y el desarrollo del aprendizaje abierto y a distancia en el ámbito de la educación superior.

Entre los numerosos proyectos financiados por la Comisión Europea cabe mencionar el proyecto EQIBELT. Este proyecto tiene por principal objetivo la integración efectiva del e-Learning en el sistema de educación superior. Es un proyecto financiado por el proyecto TEMP de la Comisión Europea, en el que colaboran universidades de varios países europeos, entre las cuales cabe mencionar las siguientes: la Universidad de Hull de Gran Bretaña, Universidad de Viena en Austria, Universidad de Tecnología de Helsinki, Universidad de Porto en Portugal, la Fundación de Tecnología de la Información en Estonia, Universidad de Leuven en Bélgica, y la Universidad de Valencia en España, entre otras.

En cuanto a las asociaciones internacionales que trabajan para la promoción de la educación virtual podemos destacar a la *Association for the Advancement of Computing in Education* (AACDE) que lleva más de 25 años dedicada al desarrollo de teorías y modelos de enseñanza-aprendizaje soportados con tecnologías de la información.

Otra tendencia que se viene desarrollando como producto del incipiente impulso de las TIC aplicadas a la educación, consiste en la creación de alianzas entre varias instituciones de educación superior con la finalidad de conformar nuevas organizaciones bajo la premisa de consolidar redes para compartir objetivos, conocimientos y estructuras tecnológicas. Como ejemplos de esta tendencia cabe mencionar a:

- El Campus Virtual Latino-Americano (CAVILA), proyecto que pretende contribuir a la democratización de la educación superior mediante el acceso de las TIC aplicadas a la educación. En esta iniciativa participan la Universidad de Santiago de Chile, las universidades argentinas (Nacional de Córdoba, Entre Ríos y Nacional de la Plata), la Universidad Santa María de Brasil, la Universidad de Guadalajara de México y la Universidad de Extremadura de España (http://blog.pucp.edu.pe/item/7905)
- La Finnish Virtual University (FVU) se trata de un consorcio conformado por veintiuna universidades finlandesas, con la finalidad de apoyar y desarrollar una red de colaboración que ofrece servicios educativos que impulsan el uso de las TIC (www.virtuaaliyliopisto.fi)
- La Universidad de Alicante, la Universidad Autónoma de Barcelona, la Universidad Carlos III de Madrid y el Grupo PRISA han creado el Instituto Universitario de Postgrado con la finalidad de aportar una amplia oferta educativa de postgrado (www.iup.es).
- Asimismo, es interesante destacar los proyectos de colaboración Universidad – Empresa. Como ejemplo de esta colaboración se encuentra, entre otros, el desarrollado por la Universidad de Málaga en colaboración con la empresa Ingenia, así como la iniciativa del Aula Virtual de la Fundación Universidad-Empresa de la Universidad de Valencia (ADEIT), que nace con la filosofía de ayudar y asesorar al mundo empresarial en el desarrollo e integración de las TIC.

4. REFLEXIÓN FINAL

A pesar de los buenos resultados que muestran los indicadores sobre penetración de las TIC en el sistema universitario español, según corroboran numerosos estudios en esta línea y, de las numerosas iniciativas que se vienen configurando en el ámbito de los entornos virtuales de enseñanza-aprendizaje, existe un claro desfase entre el nivel de implantación de estas tecnologías y su uso efectivo. Podemos afirmar que el desarrollo del *Blended learning* en el sistema universitario se encuentra en una etapa incipiente, en la que es necesario avanzar más allá de su aplicación en un entorno cerrado alumno-profesor hacia su integración efectiva como

[5] http://europa.eu.int/comm/education/elearning

[6] Sobre este particular existen numerosas referencias que se pueden consultar en el documento *'Education & Training 2010'. Main policy initiatives and outputs in education and training since the year 2000* (Comisión Europea, 2008).

herramienta interactiva, que permita la cooperación entre docentes y universidades tanto del ámbito estatal como internacional.

Es importante desarrollar un plan estratégico en el que se especifique de forma explícita cómo soportan los nuevos entornos de enseñanza-aprendizaje la visión, los valores y los principios de nuestras instituciones universitarias. Asimismo, es necesario identificar en dicho plan las necesidades, objetivos, costes, recursos disponibles y diversos parámetros de calidad para el desarrollo de un proceso de aseguramiento efectivo (Garnham y Kaleta, 2002). Esto, sin duda alguna, facilitaría el rediseño de los cursos, mejoraría la calidad de la docencia on-line, y en definitiva, ayudaría a utilizar la tecnología de forma efectiva en el proceso de enseñanza.

En el marco del EEES cada vez va a ser más necesaria la creación de redes o asociaciones universitarias, así como establecer unos estándares y criterios de calidad que permitan elaborar y publicar en abierto material de calidad tanto científica, como tecnológica como didáctica.

Tenemos por delante numerosos apartados en los que seguir avanzando hasta llegar a generalizar y extender la innovación pedagógica al conjunto del colectivo universitario, de forma que el nuevo modelo de enseñanza-aprendizaje deje de ser una novedad y se convierta en una práctica habitual en el sistema universitario español. Ahora bien, mientras la comunidad universitaria no detecte esta necesidad real es evidente que existirá una clara resistencia a dar este importante paso. Este es el gran reto.

BIBLIOGRAFÍA

ALBA, C. Y CARBALLO, R. (2005), "Viabilidad de las propuestas metodológicas para la aplicación del crédito europeo por parte del profesorado de las universidades españolas vinculadas a la utilización de las TIC en la docencia y la investigación". *Revista de Educación*, número 337.

BARRO, S. (2004), *Las tecnologías de la información y la comunicación en el sistema universitario español*. Conferencia de Rectores de las Universidades Españolas (CRUE).

CARR, N. (2003), "IT doesn't Matter". Harvard Business Review, Volumen 81, Nº 5.

CARRETERO, M. (1993), *Constructivismo y Educación*. Madrid. Edelvives.

COLL, C. (1996), *Aprendizaje escolar y construcción del conocimiento*. Barcelona: Paidós.

COMISIÓN EUROPEA (2008), "Education and Training 2010. Main policy initiatives and outputs in education and training since the year 2000". European Commission, Coordination of Lifelong Learning Policies.

CRUE (2006), *Las TIC en el Sistema Universitario Español*. Conferencia de Rectores de las Universidades Españolas.

FREIRE, J. (2008), "Universities and Web 2.0: Institutional challenges". Elearning Papers.

GARNHAM, C. Y KALETA, R. (2002), "Introduction to hybrid courses". *Teaching with technology today.* Volumen 8, No. 6.

GARRISON, D. Y KANUKA, H. (2004), "Blended learning: Uncovering its transformative potential in higher education". *The Internet and Higher Education*, Volumen 7, No. 2.

MCCRAY, G.E.(2000), "The hybrid course: Merging on-line instruction and the traditional classroom". *Information Technology and Management*, Volumen. 1.

PLS. RAMBOL MANAGEMENT (2004): *Studies in the Context of the E-learning Initiative: Virtual Models of European Universities. Draft Final Report to the EU Commission*. DG Education & Culture.

ROSALES, C. (2001): "Comunicación didáctica en la universidad". En A. García-Valcárcel (coord.): *Didáctica Universitaria*. Madrid, La Muralla.

SANGRÁ, A. Y GONZÁLEZ SANMAMED, M. (2004), *La transformación de las universidades a través de las TIC: discursos y prácticas*. Colección Educación y Sociedad Red. Editorial UOC.

UNIVERSITIC (2008), *Las TIC en el Sistema Universitario Español*. Informe publicado por la Conferencia de Rectores de las Universidades Españolas. Directores: Javier Uceda y Senén Barro. http://www.crue.org/UNIVERSITIC

VOOS, R. (2003), "Blended learning –What is it and where might it take us?". *Sloan-C View*, Vol. 2, No. 1.

ADE-DERECHO EN LA UNIVERSITAT DE VALÈNCIA: ESTUDIO DE UNA EXPERIENCIA PILOTO.

ANTONIA IVARS ESCORTELL
E-mail: Antonia.Ivars@uv.es
Mª ISABEL LÓPEZ RODRÍGUEZ
E-mail: Maria.I.Lopez@uv.es
FELIX RUIZ PONCE
E-mail: Felix.Ruiz@uv.es
Facultat d'Economia
Universitat de València

Resumen

Implementation of numerous educational innovation plans (EIP) in the Spanish universities, throughout these last years has produced a big amount of highly interesting experiences. Facultat d'Economia of Universitat de València is one of the biggest faculties in the whole Spanish university, having more than 9000 students. In this context, the global installation of innovation plans had serious difficulties and the pilot experience coincided with the start of the double degree "ADE-Derecho", that has been running for 6 years. This paper shows a descriptive study of the general characteristics in this group of students, as well as an inference analysis of their academic performance.

Palabras clave: doble titulación, innovación docente, rendimiento académico
Área Temática: Didáctica de la Economía y Metodología.

1. INTRODUCCIÓN

Desde hace unos años se está viviendo, en las universidades españolas, un proceso denominado, de forma general, Convergencia al Espacio Europeo de Educación Superior. La implantación de nuevos planes de estudio debía implicar un cambio radical en las metodologías docentes y en el peso de cada agente en la transmisión de conocimientos, exigiendo para los alumnos un papel mucho más activo del que actualmente desempeñan y el uso de los diferentes instrumentos y formas docentes por parte del profesorado.

Durante todo este periodo, cada universidad ha impulsado iniciativas tendentes a experimentar en sus titulaciones todos estos cambios, en la estructura de la docencia, que traía aparejado el denominado "Proceso de Bolonia". Se han desarrollado innumerables Proyectos de Innovación Educativa (PIE) y sus resultados han sido y siguen siendo debatidos en numerosos congresos, jornadas o reuniones, con el fin de intercambiar experiencias y formar al profesorado en esta ardua tarea que le espera.

La Facultat d'Economia de la Universitat de València fue una de las pioneras, recogió el testigo rápidamente y un conjunto de entusiastas profesores se involucró con distintos proyectos, tanto individuales como colectivos, iniciando un camino por el que todos hemos de transitar. En un primer paso, se aplicaron nuevas metodologías docentes en algunas asignaturas dentro de lo que podríamos considerar como iniciativas personales. Posteriormente, coincidiendo con la implantación de la doble titulación de Licenciado en Administración y Dirección de Empresas y Licenciado en Derecho (ADE-Derecho), desde los equipos decanales de ambos centros, se tomó la decisión de que esta titulación se desarrollase dentro de un PIE auspiciado por la Universitat

de València. Era la forma de añadir a los aspectos de las nuevas metodologías docentes, ya desarrollados en los proyectos individuales, todo aquello relacionado con la coordinación horizontal y vertical de los contenidos y la necesaria estructura de los estudios. El proyecto era ambicioso pues hay que añadir a todo lo anteriormente citado, el hecho de que se encuentren implicados los dos centros más grandes de nuestra universidad, con metodologías docentes tradicionalmente distintas.

Esta experiencia, que este curso cumple su sexto año, ha sido el detonante de otras experiencias interesantes como la implantación de un PIE en la Licenciatura de Economía, acogido a un proyecto del Ministerio de Educación, con el fin de analizar los resultados de las nuevas metodologías aplicadas a un grupo más numeroso.

La oferta de la doble titulación de ADE-Derecho supuso un éxito y la demanda, año tras año, supera con mucho la oferta de plazas, que es de 120. La nota de acceso suele superar el 7'5 y el bloque de estudiantes que cursa esta doble titulación presenta, por tanto, unas características peculiares dentro del alumnado matriculado en las distintas titulaciones que actualmente oferta la Facultat d'Economia. Hasta el curso 2007-08 el número de estudiantes que accedían por cada titulación era el mismo (60) y la nota de corte difería mínimamente. Sin embargo, en el curso 2007-08 la diferencia entre la nota de corte por cada centro difería sustancialmente lo que produjo agravios comparativos que desde los centros y el rectorado se intentó solucionar con el incremento, ese mismo curso, de 20 plazas que se distribuyeron por nota de acceso en un listado único de ambos centros. Ya en el 2008-09 se modificó el reparto de alumnos, accediendo 40 desde cada centro y 40 por nota según listado conjunto.

La Facultat d'Economia es, en estos momentos, el centro más grande de todas las universidades de la Comunidad Valenciana, con unos 9200 alumnos aproximadamente. Cada año acceden más de 1700 alumnos, se reciben más de 400 alumnos Erasmus y más de 350 de nuestros estudiantes se desplazan a alguna de las universidades extranjeras con las que se mantiene convenio.

La descripción, a grandes rasgos, de este escenario se considera conveniente para entender que los alumnos de la doble titulación ADE-Derecho constituyen un grupo claramente diferenciado con unas características particulares.

Algunos profesores, implicados desde el principio en este proyecto, publicaron en 2007 el libro "Innovación educativa en la Universidad: ADE-Derecho" donde se recogen, con gran profusión, los resultados académicos de los alumnos y se establece un estudio comparativo con los resultados de dos grupos convencionales de ambas titulaciones, es decir grupos no acogidos a ningún PIE.

Sin embargo, la peculiar situación ya comentada del acceso de los estudiantes durante el curso 2007-2008 creó un escenario distinto sobre el existente hasta entonces: las notas de acceso por la vía de Derecho habían descendido y el número de alumnos que accedieron por ADE se incrementó.

Este cambio motivó el planteamiento de una serie de cuestiones en las reuniones de coordinación que se realizan en los distintos cursos de la doble titulación, como por ejemplo la posible existencia de diferencias entre los rendimientos de unos y otros. Así, nuestro objetivo al desarrollar este trabajo, ha ido en la línea de dar respuesta a ese tipo de inquietudes planteadas y a otras cuestiones complementarias que han ido surgiendo y que parecen hacerse más patentes en los estudiantes del curso 2007-2008: ¿hay diferencias significativas en el rendimiento de los alumnos que acceden por Derecho y los que acceden por ADE? (Hay que considerar que la opción de Bachillerato de unos y otros puede ser sustancialmente distinta). ¿Son más acusadas estas diferencias cuando particularizamos a asignaturas de corte cuantitativo?. ¿Hay diferencias entre su procedencia (enseñanza privada o pública) y su rendimiento? etc.

El estudio se realiza, por tanto, utilizando los datos obtenidos a partir de las encuestas pasadas a dos grupos de alumnos de primero, durante el curso 2007-2008. Con esta información se ha realizado un estudio descriptivo de las características del alumnado de esta doble titulación y su relación con los resultados académicos obtenidos. Además, y siempre teniendo en cuenta las limitaciones de estas encuestas para ser consideradas una muestra representativa, se ha intentado en lo posible aplicar la inferencia para analizar la significatividad de las diferencias observadas desde el estudio descriptivo.

Debe tenerse en cuenta que este es solo un primer paso. En años sucesivos se pretende recabar más información para llegar a conclusiones que efectivamente puedan considerarse válidas para toda la población.

2. TRATAMIENTO DE LA INFORMACIÓN.

El interés suscitado por la experiencia que ha supuesto esta doble titulación en esta facultad y por analizar los resultados de la aplicación de las nuevas metodologías docentes, ha obligado a someter a los alumnos a una continuada batería de encuestas que, sin embargo, deja sin explorar distintos aspectos fundamentales para el estudio que se intentaba realizar.

Resultó necesario, por tanto, la elaboración de una nueva encuesta que se decidió pasar a los dos grupos de primer curso del curso 2007-08 ya que en este curso se comenzaron a apreciar diferencias entre distintos subgrupos de alumnos que parecían significativas.

La encuesta recogía la siguiente información:

- Nivel de estudios del padre y de la madre.
- Lugar de residencia durante el curso.
- Procedencia de centro privado o público.
- Opción de bachillerato.
- Notas de acceso a la universidad.
- Vía de acceso.
- Motivación en la elección de estudios.
- Notas obtenidas en las distintas asignaturas.
- Valoración sobre el interés del alumno.
- Metodología docente aplicada.
- Dedicación del alumno.
- Métodos de evaluación utilizados.

El tratamiento de las encuestas descubrió diversas deficiencias en la interpretación de las preguntas y en las contestaciones. Se procedió a una depuración detallada de las respuestas y, en algunos casos, se observó que los problemas eran de tal calibre que algunas variables debían quedar eliminadas del trabajo. Todas estas cuestiones sirven, por supuesto, para modificar las preguntas y las repuestas posibles en las futuras encuestas que se pretende realizar.

Por los motivos citados y porque los objetivos a analizar fueron acotándose según se desarrollaba este estudio, finalmente se decidió trabajar con las siguientes variables:

- CENTBACHI: tipo de centro donde el alumno ha cursado el bachillerato (público o privado)
- ESTUMADRE: nivel de estudios de la madre (1: Titulación superior, 2: Diplomatura universitaria o similar, 3: bachillerato, 4: estudios primarios completos y 5: sin estudios).
- ESTUPADRE: nivel de estudios del padre (1: Titulación superior, 2: Diplomatura universitaria o similar, 3: bachillerato, 4: estudios primarios completos y 5: sin estudios).
- VIAACC: vía de acceso seleccionada por el estudiante para realizar la titulación (ADE o Derecho).
- OPCBACHI: opción de bachillerato cursada por el estudiante (Biosanitario, Tecnológico, Humanidades, Artístico y Sociales).
- NOTAACC: nota de acceso a la Universidad.
- DRNOTA: nota obtenida en la asignatura Derecho Romano.
- MAT1NOTA: nota obtenida en la asignatura "Introducción a la Matemática Económico-Empresarial".
- TDCHONOTA: nota obtenida en la asignatura Teoría del Derecho.

En estas variables, no se detectaron prácticamente problemas de respuesta y en algunos casos se ha procedido a su agrupación para facilitar la aplicación de diversos métodos estadísticos de análisis.

3. ANÁLISIS DE LOS DATOS.

3.1. Criterios de agrupación.

Atendiendo a los objetivos mencionados en la introducción, se procede a analizar la correlación estadística entre pares de características y en la medida de lo posible la relación causa-efecto. Para ello, y fijando como objetivo prioritario estudiar el rendimiento académico de los estudiantes, se considerarán distintas formas de agrupar a los mismos para, posteriormente, ver si existen diferencias significativas entre los grupos formados. En concreto, los criterios de agrupación y las características sobre las que se analizará si tiene algún tipo de efecto son:

(I) Criterio de agrupación: procedencia de los alumnos, esto es si han cursado el bachillerato en un centro público o privado ("CENTBACHI").

Características sobre las que se estudia su posible efecto:

- "ESTUMADREPADRE": nivel de estudios de los padres (esta característica se obtiene considerando la suma del nivel de estudios de la madre y del padre, esto es ESTUMADREPADRE=ESTMADRE+ESTPADRE).
- VIAACC
- OPCBAHI
- NOTAACC
- DRNOTA
- MAT1NOTA
- TDCHONOTA
- NOTAMEDIA: nota media de las asignaturas mencionadas anteriormente.

(II) Criterio de agrupación: vía de acceso seleccionada por el estudiante para realizar la titulación ("VIAACC ").

Características sobre las que se estudia su posible efecto:

- DRNOTA.
- MAT1NOTA.
- TDCHONOTA.
- NOTAMEDIA.
- OPCBACHI.

(III) Criterio de agrupación: opción de bachillerato cursada por el estudiante ("OPCBACHI ").

Características sobre las que se estudia su posible efecto:

- NOTAACC.
- DRNOTA.
- MAT1NOTA.
- TDCHONOTA.
- NOTAMEDIA.

(IV) Criterio de agrupación: nivel de estudios de los padres ("ESTUMADREPADRE").

Característica sobre las que se estudia su posible efecto:

- NOTAMEDIA.

Por otra parte, teniendo en cuenta que algunas de las características consideradas son de carácter cualitativo y otras de tipo cuantitativo, se realizará el estudio comparativo entre los grupos confeccionados desde una perspectiva u otra, de manera que:

❖ Si la característica es cualitativa se procederá a comparar el porcentaje de alumnos que se clasifican en una u otra modalidad, utilizando también la obtención de las gráficas correspondientes a cada caso.

❖ Si por el contrario la característica es cuantitativa se hará uso, además del recurso gráfico, de un conjunto de herramientas de tipo descriptivo.

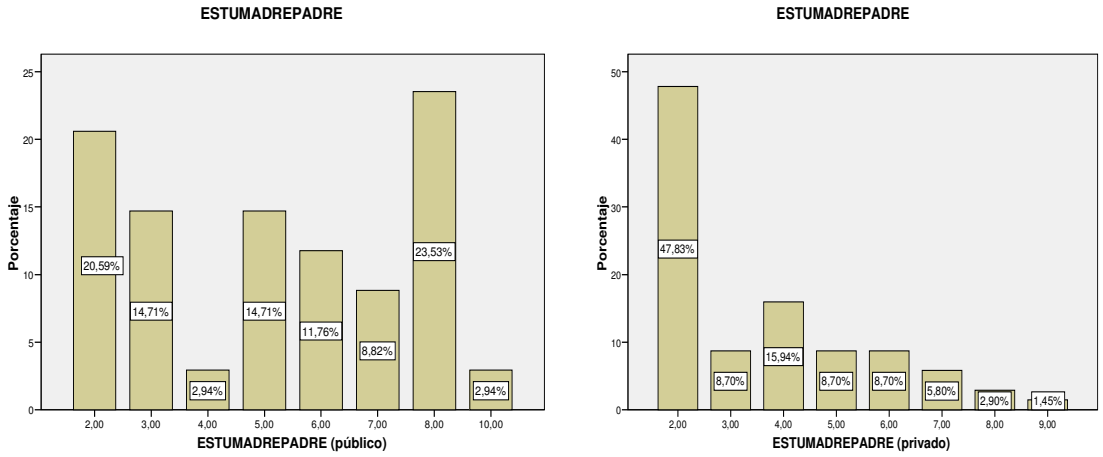
3.2. Análisis de los datos en los distintos grupos.

(i) efecto de la "procedencia" sobre el rendimiento académico y otras características.

En este apartado se analiza el efecto de la característica "procedencia" sobre otras.

a) "Procedencia" y nivel de estudios de los padres.

Para la característica ***"ESTUMADREPADRE"***, se puede comprobar como el reparto de porcentajes presenta claras diferencias entre los grupos considerados. Así caben destacar, por ejemplo, los casos en que ambos padres poseen "Titulación superior" y aquel en que ambos están en posesión de estudios "primarios completos", pues mientras que para el conjunto de estudiantes que proceden de un centro "PÚBLICO" los porcentajes son del 20,59% y 23,53% respectivamente, dichas proporciones son del 47,83% y del 2,90% para el grupo de alumnos que han cursado bachillerato en un centro "PRIVADO". Las siguientes gráficas muestran, con más detalle, lo expuesto:



Las diferencias que se observan parecen lo suficientemente grandes como para resultar estadísticamente significativas; para comprobarlo se realiza un contraste de independencia entre las características "procedencia" y "estumadrepadre", resultando que

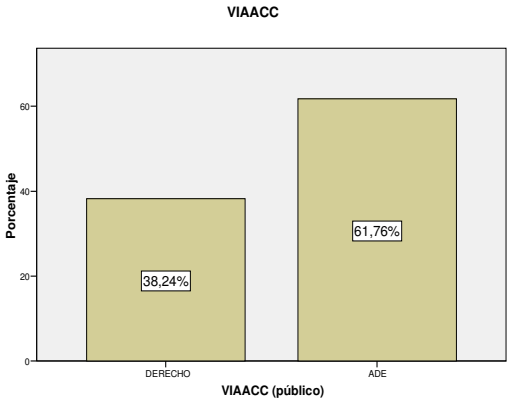
Pruebas de chi-cuadrado

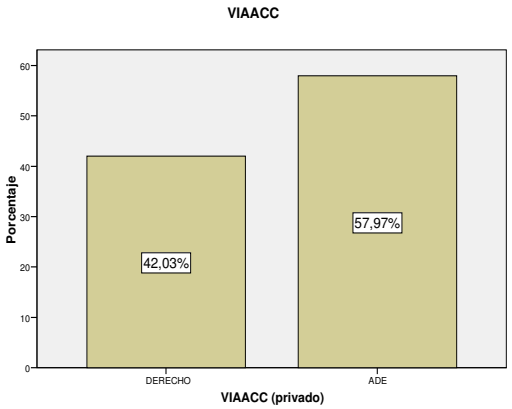
	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	13,580(a)	3	,004
Razón de verosimilitudes	13,464	3	,004
Asociación lineal por lineal	13,422	1	,000
N de casos válidos	103		

Luego se rechaza, con rotundidad, la independencia de las dos características analizadas. El porcentaje de alumnos cuyos padres tienen bajos niveles de estudios (7 a 10) es mucho mayor en los colegios públicos que en los privados: concretamente, el 35,29% frente al 10,15%.

b) "Procedencia" y vía de acceso.

No existen, sin embargo, grandes diferencias, en cuanto a la vía de acceso seleccionada (***"VIAACC"***). En efecto, en ambos grupos alrededor del 60% eligieron esta vía y entorno al 40% la de DERECHO.



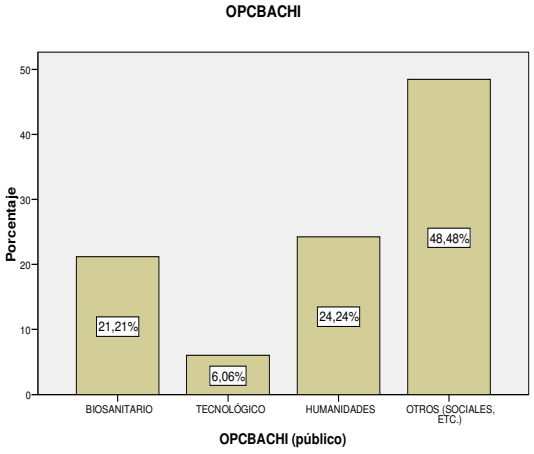


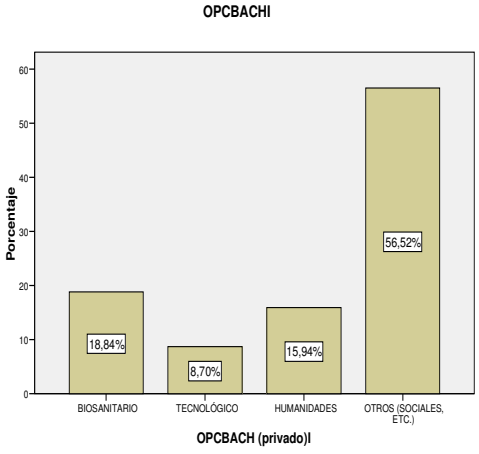
OPCBACHI (privado)

OPCBACHI (privado)

c) "Procedencia" y opción de bachillerato cursada.

Considerada la opción de bachillerato cursada (***"OPCBACHI"***) el perfil es similar entre ambos grupos, tal y como se muestra en los siguientes gráficos (que al igual que los anteriores contienen el porcentaje para cada modalidad u opción). Cabe destacar que en los dos casos la proporción mayoritaria corresponde a la de estudiantes que cursaron la opción de SOCIALES, proporción que se eleva al 48,48% para los procedentes de centro público y del 56,52% para los provenientes de privado.





SOCIALES

SOCIALES

Pero existen claras diferencias; ¿son estadísticamente significativas?. Pues realizando el contraste de independencia se obtienen los siguientes resultados:

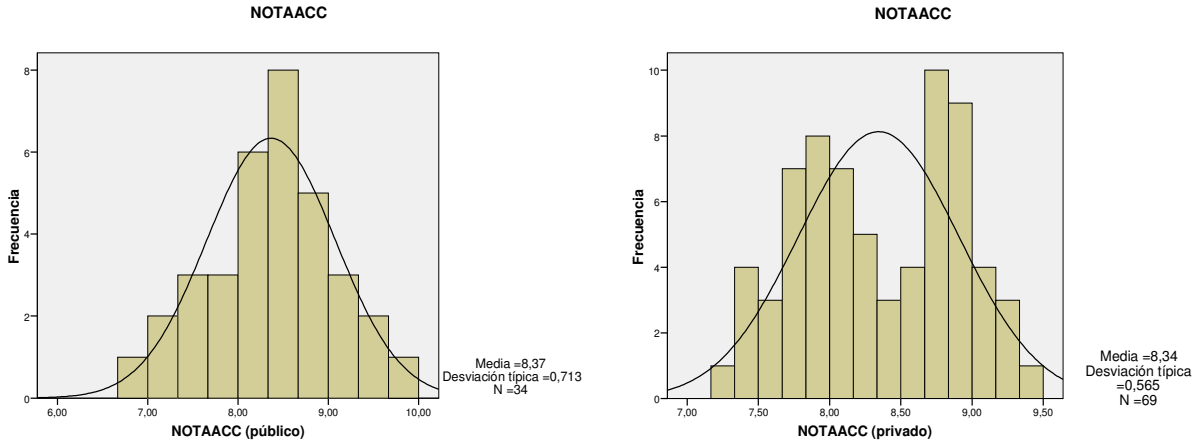
Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	1,355(a)	3	,716
Razón de verosimilitudes	1,333	3	,721
Asociación lineal por lineal	,276	1	,600
N de casos válidos	102		

Lo que lleva a no rechazar la hipótesis de independencia entre las características "procedencia" y "opción de bachillerato".

d) "Procedencia" y nota de acceso a la Universidad.

Considerada la nota de acceso a la Universidad (***"NOTAACC"***), se observa que en los dos conjuntos la nota media es similar, pues: $\overline{X}_{PÚBLICO} = 8{,}37$ y $\overline{X}_{PRIVADO} = 8{,}34$, tampoco hay gran diferencia en cuanto a las medidas de dispersión, ya que $S_{PÚBLICO} = 0{,}713$ y $S_{PRIVADO} = 0{,}565$ por lo que en términos relativos, haciendo uso del coeficiente de variación de Pearson se obtienen unos valores para dicho coeficiente de 0,085 y 0,068 respectivamente, no presentando tampoco una discrepancia importante. Sin embargo, si se representan gráficamente los datos se obtiene que:



Detectándose una diferencia en la forma, pareciendo que en el caso de "PRIVADO" se han mezclado dos poblaciones. Es más, también resulta curioso el recorrido de las variables en ambos casos. Concretamente, los valores mínimos y máximos, así como el rango de la variable para ambos grupos quedan reflejados en la siguiente tabla:

	Mínimo	Máximo	Rango
PÚBLICO	6,8	9,96	3,16
PRIVADO	7,3	9,37	2,07

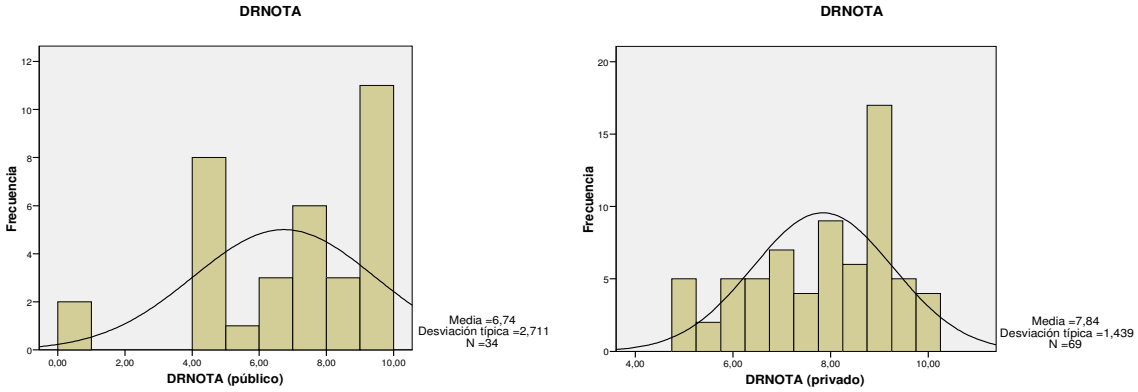
A pesar de la notable diferencia entre los gráficos, ésta no es suficiente como para rechazar la independencia entre las características "procedencia" y "NOTAACC".

e) "Procedencia" y nota en la asignatura Derecho Romano.

En cuanto a las notas obtenidas en las asignaturas del primer cuatrimestre cabe destacar una diferencia importante entre "PÚBLICO" y "PRIVADO" en la asignatura de Derecho Romano (***"DRNOTA"***). En efecto, las notas medias fueron: $\overline{X}_{PÚBLICO} = 6{,}74$ y $\overline{X}_{PRIVADO} = 7{,}84$, y las desviaciones típicas $S_{PÚBLICO} = 2{,}711$ y $S_{PRIVADO} = 1{,}439$

obteniéndose, por tanto, un valor para el coeficiente de variación de Pearson de 0,40 y 0,18 respectivamente. Es decir, por término medio los estudiantes provenientes de un centro privado obtuvieron mejores calificaciones, existiendo además una menor dispersión en sus notas.

Además, de la observación de las gráficas y de la tabla, se desprende que las calificaciones más bajas las obtuvieron los de "PÚBLICO", pero fueron más individuos de este grupo los que obtuvieron la máxima nota. Lo indicado pone de manifiesto las diferencias entre ambos grupos en "DRNOTA".



	Mínimo	Máximo	Rango	Moda
PÚBLICO	0	10	10	4
PRIVADO	5	10	5	9

En este caso las diferencias son suficientemente importantes, como se comprueba con el correspondiente contraste en el que se rechaza la igualdad de medias:

Prueba de muestras independientes

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
									95% Intervalo de confianza para la diferencia	
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	Superior	Inferior
DRNOTA	Se han asumido varianzas iguales	20,573	,000	-2,708	101	,008	-1,10529	,40823	-1,91509	-,29548
	No se han asumido varianzas iguales			-2,227	42,395	,031	-1,10529	,49622	-2,10642	-,10415

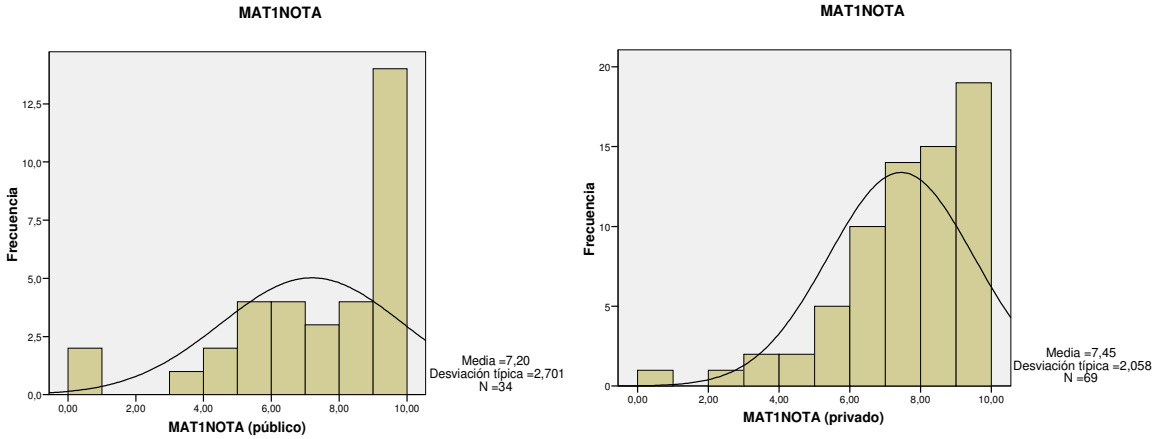
f) "Procedencia" y nota en la asignatura Introducción a la Matemática Económico-Empresarial.

Si se analizan las notas obtenidas en la asignatura "Introducción a la Matemática Económico-Empresarial" (***"MAT1NOTA***") las conclusiones son diferentes.

En efecto la nota media, así como las medidas de dispersión (desviación típica y coeficiente de variación de Pearson son parecidas, pues:

$\overline{X}_{PÚBLICO} = 7{,}2$, $\overline{X}_{PRIVADO} = 7{,}45$, $S_{PÚBLICO} = 2{,}701$ y $S_{PRIVADO} = 2{,}058$,

$CV_{PÚBLICO} = 0{,}375$, $CV_{PRIVADO} = 0{,}276$



Las representaciones gráficas no presentan una disparidad de forma tan acusada como en el caso anterior. Si acaso, una de las discrepancias que puede observarse es que las notas extremas se dan en mayor proporción en el grupo "PÚBLICO" (15% con notas inferiores al 4,5 y 45% superiores a 8,8 en el "público" frente 8% y al 28% respectivamente en el "privado"). En cualquier caso ambas distribuciones poseen el mismo valor mínimo y máximo, y modas y medianas similares:

	Mínimo	Máximo	Rango	Mediana	Moda
PÚBLICO	0	10	10	8,15	9
PRIVADO	0	10	10	7,9	10

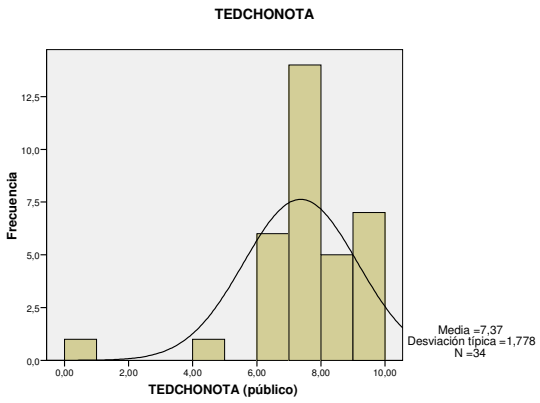
g) "Procedencia" y nota en la asignatura Teoría del Derecho.

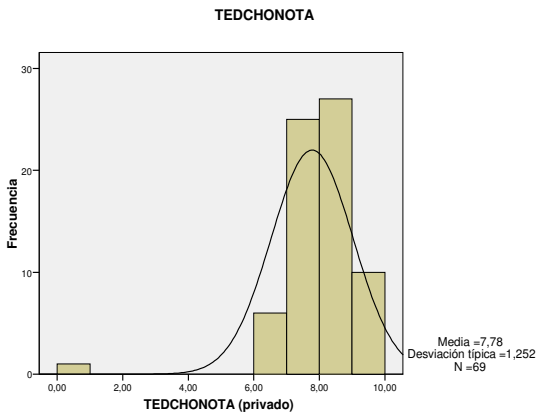
Consideradas las calificaciones obtenidas en la asignatura Teoría del Derecho (***TEDCHONOTA"***) las conclusiones son parecidas al caso anterior. Así, los valores de medias, desviaciones y coeficientes de variación de Pearson para este caso son:

$$\overline{X}_{\text{PÚBLICO}} = 7{,}37\,, \quad \overline{X}_{\text{PRIVADO}} = 7{,}78\,, \quad S_{\text{PÚBLICO}} = 1{,}778 \text{ y} \quad S_{\text{PRIVADO}} = 1{,}252\,,$$
$$CV_{\text{PÚBLICO}} = 0{,}24\,,\ CV_{\text{PRIVADO}} = 0{,}16$$

Es decir, la tendencia es muy similar al de las notas de la asignatura anterior, con una nota media ligeramente superior en el caso del grupo "PRIVADO" y además con menor dispersión.

Si se observan las gráficas en los dos grupos la gran mayoría de alumnos tienen una calificación por encima del 6, pudiéndose mencionar, si acaso, como una diferencia la aparición de más notas por debajo del 5 en el grupo "PÚBLICO". En cualquier caso ambas distribuciones poseen valor mínimo, máximo, y mediana similares:





	Mínimo	Máximo	Rango	Mediana
PÚBLICO	0	9,5	9,5	7,5
PRIVADO	0	9,6	9,6	8

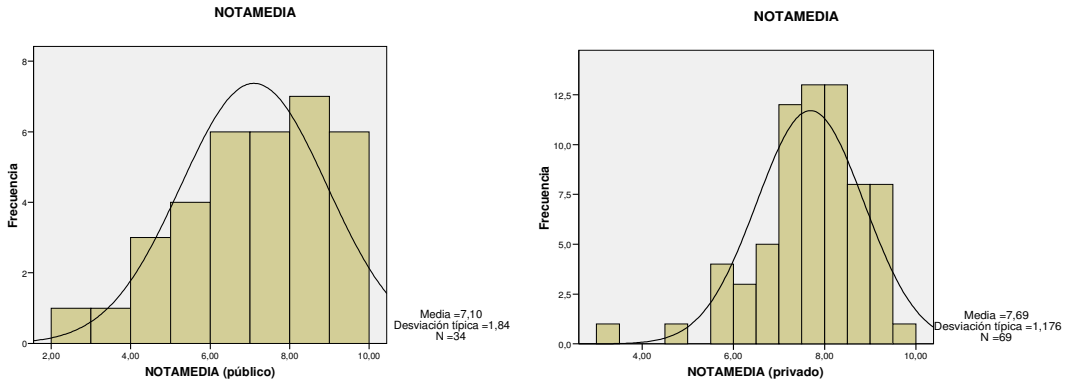
h) "Procedencia" y nota media.

Si se define una nueva variable, a partir de las tres últimas, correspondiente a la nota media de las tres asignaturas (***"NOTAMEDIA"***), y se comparan las medias y medidas de dispersión en los dos grupos de estudiantes no se detectan grandes discrepancias:

$$\overline{X}_{\text{PÚBLICO}} = 7{,}1\ ;\ \overline{X}_{\text{PRIVADO}} = 7{,}69\ ;\ S_{\text{PÚBLICO}} = 1{,}84\ ;\ S_{\text{PRIVADO}} = 1{,}176$$

$CV_{PÚBLICO}$=0,26; $CV_{PRIVADO}$=0,15

La observación de los gráficos así como valores mínimos corroboran lo indicado. Puede apuntarse que el hecho de ser multimodal en el grupo "PÚBLICO" confiere a la gráfica parte de su aspecto.
y máximos



	Mínimo	Máximo	Rango
PÚBLICO	2,67	9,83	7,17
PRIVADO	3,17	9,6	6,43

(ii) efecto de la "vía de acceso" sobre el rendimiento académico y otras características.

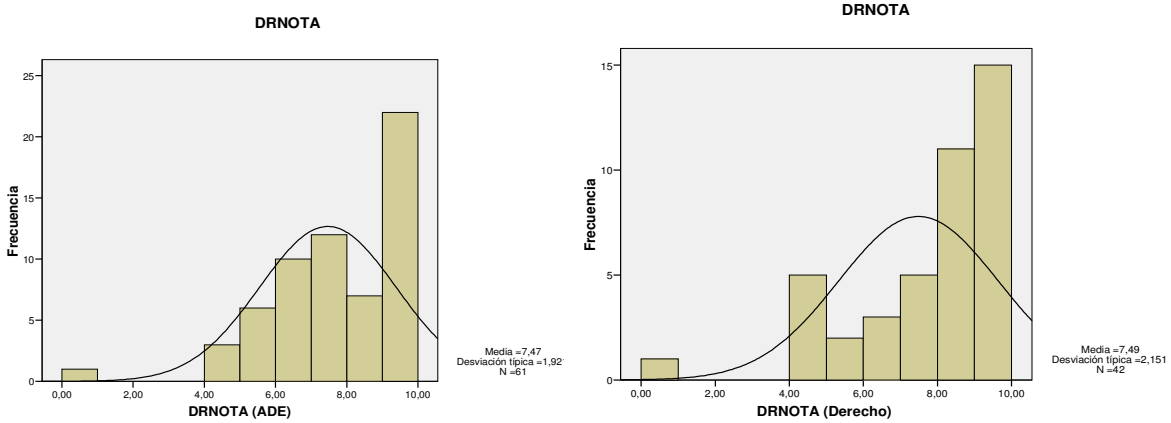
En el presente epígrafe se estudia el efecto de la característica "vía de acceso" sobre otras características, fundamentalmente aquellas relacionadas con el rendimiento académico.

a) "Vía de acceso" y nota en la asignatura Derecho Romano.

No existen grandes diferencias en ambos grupos (alumnos que accedieron a la doble titulación vía ADE o vía DERECHO) en cuanto a la calificación obtenida en Derecho Romano (***"DRNOTA"***). Efectivamente, de la comparación de las medias y otras medidas descriptivas, contenidas en la tabla siguiente, así como de la observación de las respectivas gráficas, se concluye una gran analogía entre las medidas de reducción.

Puede, en todo caso, constatarse un mayor porcentaje de notas elevadas en los alumnos que accedieron por Derecho, pero también hay una proporción ligeramente superior en dicho grupo de suspendidos (13% frente al 19%).

	Mínimo	Máximo	Rango	Media	D. Típica	C.V.	Mediana
ADE	0	10	10	7,467	1,921	0,257	7,5
DERECHO	0	10	10	7,488	2,151	0,287	8



Por supuesto, no se puede rechazar la igualdad de medias, como se comprueba con el correspondiente contraste:

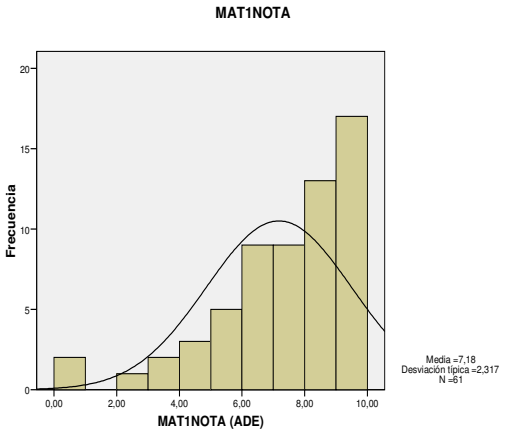
Prueba de muestras independientes

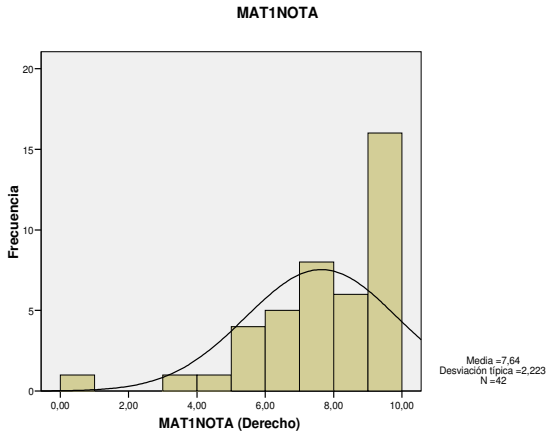
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
									95% Intervalo de confianza para la diferencia	
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	Inferior	Superior
DRNOTA	Se han asumido varianzas iguales	,349	,556	,052	101	,959	,02088	,40456	-,78166	,82342
	No se han asumido varianzas iguales			,051	81,578	,960	,02088	,41315	-,80106	,84282

b) "Vía de acceso" y nota en la asignatura Introducción a la Matemática Económico-Empresarial.

A la vista del contenido de la tabla y los gráficos siguientes, las conclusiones en el caso de esta variable son similares al anterior, esto es, medidas muy parecidas y apenas diferencia entre los dos grupos ("ADE" y "DERECHO") en las calificaciones obtenidas en la asignatura "Introducción a la Matemática Económica-Empresarial".:

	Mínimo	Máximo	Rango	Media	D. Típica	C.V.	Mediana
ADE	0	10	10	7,177	2,317	0,323	7,9
DERECHO	0	10	10	7,641	2,223	0,291	8





Puede resultar curioso para una asignatura como ésta que los alumnos procedentes de la vía Derecho obtengan calificaciones superiores, pero no tanto como para que resulte estadísticamente significativa, como se desprende del contraste de igualdad de medias:

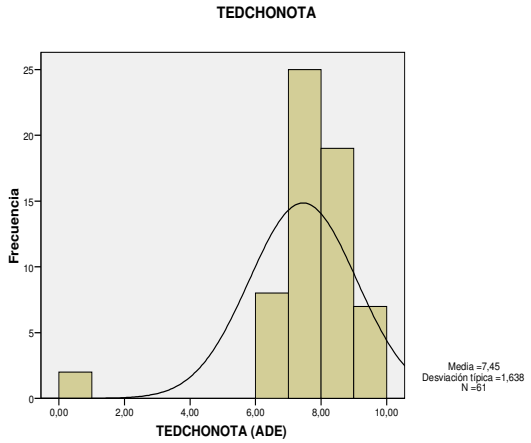
Prueba de muestras independientes

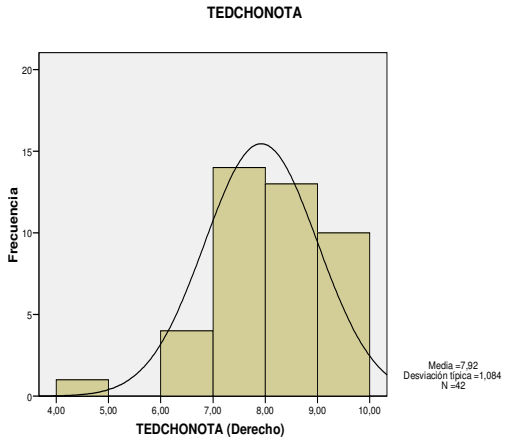
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
									95% Intervalo de confianza para la diferencia	
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	Inferior	Superior
MAT1NOTA	Se han asumido varianzas iguales	,102	,750	1,014	101	,313	,46343	,45703	-,44320	1,37005
	No se han asumido varianzas iguales			1,022	90,644	,310	,46343	,45349	-,43743	1,36428

c) "Vía de acceso" y nota en la asignatura Teoría del Derecho.

En este caso el análisis se vuelve a centrar en una asignatura de Derecho. Observando las gráficas se deduce que no hay grandes diferencias entre los dos grupos en cuanto a la nota obtenida. De los tres últimos casos es éste en el que se da una mayor semejanza en la forma de la representación gráfica.

Se puede, sin embargo, apuntar alguna diferencia, como que el valor mínimo para los que accedieron por ADE es inferior que para los alumnos del otro grupo.





Teniendo en cuenta ahora el valor de los parámetros recogidos en la siguiente tabla se puede concluir que la nota media y el porcentaje de notas elevadas, es algo superior entre los alumnos que accedieron por DERECHO.

	Mínimo	Máximo	Rango	Media	D. Típica	C.V.	Mediana
ADE	0	9,5	9,5	7,449	1,638	0,219	7,6
DERECHO	4	9,6	5,6	7,924	1,084	0,137	8

Como en los dos casos anteriores las diferencias no son significativas, como se desprende del contraste de igualdad de medias

Prueba de muestras independientes

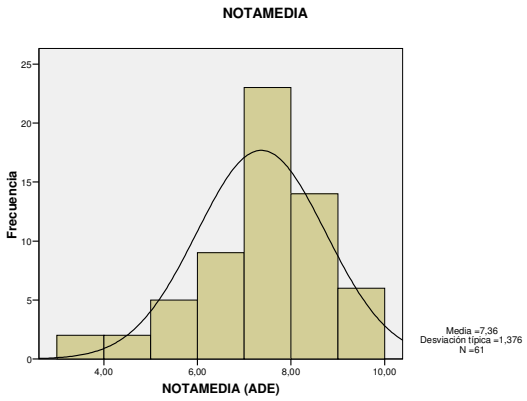
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
									95% Intervalo de confianza para la diferencia	
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	Inferior	Superior
TEDCHONOTA	Se han asumido varianzas iguales	,362	,549	1,644	101	,103	,47438	,28851	-,09794	1,04670
	No se han asumido varianzas iguales			1,769	100,874	,080	,47438	,26821	-,05769	1,00644

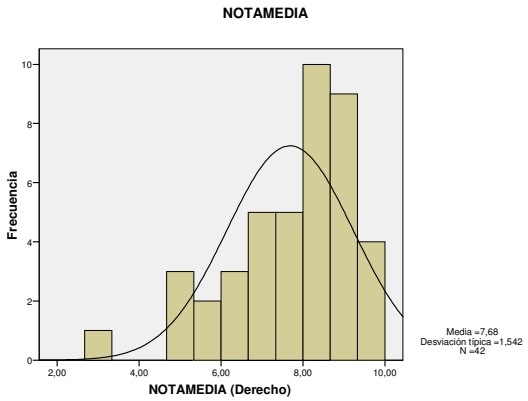
d) Vía de acceso y nota media.

Por último se pasa a analizar la relación entre la vía de acceso y la nota media de las tres asignaturas anteriores. La información que se posee se encuentra resumida en la siguiente tabla, a partir de la que se puede observar una nota media muy parecida en ambas vías de acceso, aunque un poco superior si se ha realizado el ingreso por Derecho, siendo las dispersiones en ambos casos prácticamente iguales.

	Mínimo	Máximo	Rango	Media	D. Típica	C.V.	Mediana
ADE	3,17	9,83	6,67	7,365	1,376	0,187	7,5
DERECHO	2,67	9,67	7	7,684	1,542	0,201	8

Respecto a las representaciones gráficas se puede comprobar una gran similitud entre ellas.





Por supuesto, y en congruencia con los tres resultados anteriores, no se puede rechazar la igualdad de medias

Prueba de muestras independientes

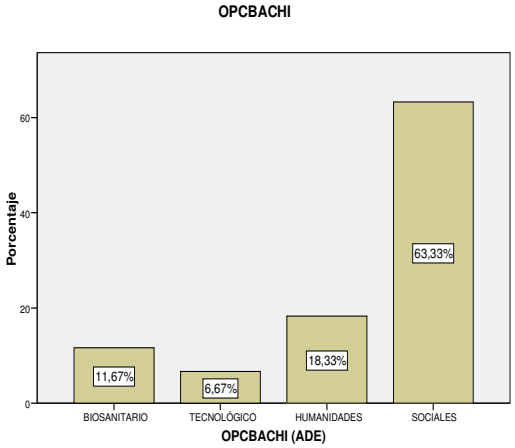
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
									95% Intervalo de confianza para la diferencia	
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	Inferior	Superior
NOTAMEDIA	Se han asumido varianzas iguales	,613	,435	1,102	101	,273	,31956	,28986	-,25544	,89457
	No se han asumido varianzas iguales			1,079	81,536	,284	,31956	,29605	-,26943	,90855

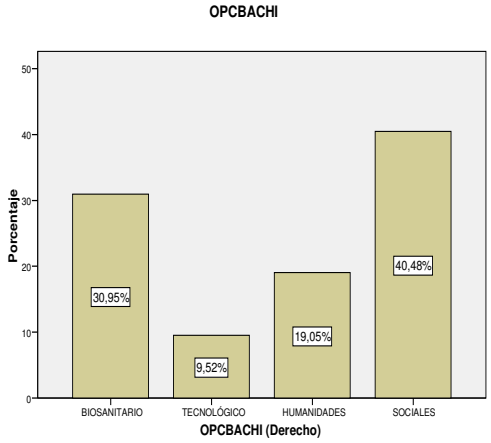
e) "Vía de acceso" y opción de bachillerato cursada.

Si se tiene en cuenta la opción de bachillerato cursada por el alumno, la distribución de los porcentajes presenta alguna discrepancia. Concretamente:

- Existe una diferencia aproximada de 20 puntos entre los que, habiendo cursado Biosanitario, decidieron entrar por ADE y los que eligieron Derecho (11,67% frente a 30,95%).
- La disparidad es análoga para los que cursaron Sociales, ya que un 40,48% seleccionaron Derecho y el 63,33% ADE.

En las gráficas de barras que siguen se tiene una información más amplia:





Dado que las diferencias observadas parecen lo suficientemente grandes como para resultar estadísticamente significativas, se procede a realizar un contraste de independencia entre las características "VIAACC" y "OPCBACHI", resultando que:

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	7,141(a)	3	,068
Razón de verosimilitudes	7,115	3	,068
Asociación lineal por lineal	6,548	1	,011
N de casos válidos	103		

Por lo que no se rechaza la hipótesis de independencia entre dichas características, al 5% de significación. Sin embargo, el resultado no es claro, pues a niveles superiores al 6,8% de significación se rechazaría la hipótesis.

(iii) efecto de opción de bachillerato sobre la nota de acceso y el rendimiento académico.

En este caso tendremos definidos 4 grupos, siendo el de más numeroso el formado por los alumnos que cursaron un bachillerato de Sociales y el menor el constituido por los que cursaron la opción Tecnológica, conteniendo respectivamente al 53,4% y 7,8% de los individuos.

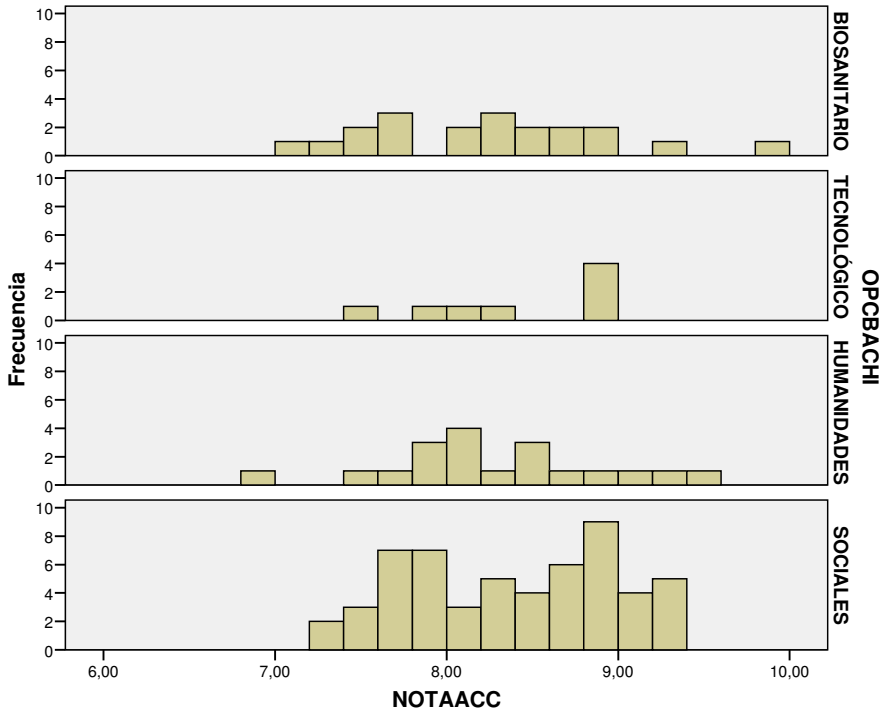
a) Opción de bachillerato y nota de acceso a la Universidad.

	Mínimo	Máximo	Rango	Media	D. Típica	C.V.	Mediana
BIOSANITARIO	7,14	9,96	2,82	8,262	0,718	0,0867	8,315
TECNOLÓGICO	7,57	8,90	1,33	8,385	0,521	0,062	8,515
HUMANIDADES	6,80	9,43	2,63	8,305	0,648	0,078	8,2
SOCIALES	7,3	9,37	2,07	8,400	0,589	0,070	8,46

En este punto se analiza la relación de la nota de acceso a la doble titulación con la opción de bachillerato cursada.

Se observa que las notas medias de acceso a la universidad son bastante homogéneas; la menor media se alcanza para la opción de Biosanitario con 8,26 y la máxima en la opción de Sociales con 8,40. Respecto a la dispersión de los valores, a pesar de estar bastante igualadas, la máxima se da en la opción de Biosanitario.

Si se observan los siguientes cuatro gráficos se ve como las opciones Tecnológico y Sociales son las que menor dispersión presentan.



Las diferencias de las medias no son significativas como se indica en el correspondiente contraste que ofrece los siguientes resultados:

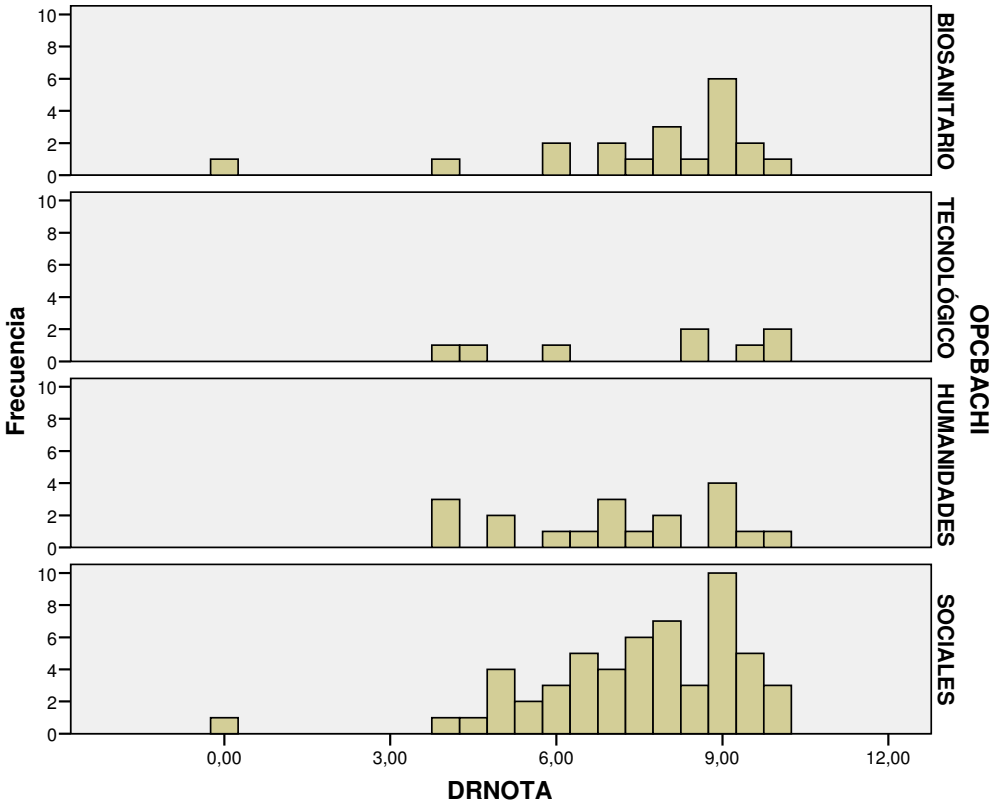
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	,343	3	,114	,295	,829
Intra-grupos	38,029	98	,388		
Total	38,372	101			

b) Opción de bachillerato y nota en la asignatura Derecho Romano.

En este apartado se analiza la relación de la nota en la asignatura de Derecho Romano con la opción de bachillerato cursada.

	Mínimo	Máximo	Rango	Media	D. Típica	C.V.	Mediana
BIOSANITARIO	0	10	10	7,65	2,323	0,304	8,25
TECNOLÓGICO	4	10	6	7,625	2,446	0,321	8,5
HUMANIDADES	4	10	6	7,079	1,967	0,278	7
SOCIALES	0	10	10	7,491	1,872	0,249	8

Las notas medias de Derecho Romano en las distintas opciones se encuentran en el intervalo [7,08, 7,65] siendo por tanto muy similares. Cabe comentar que la menor nota media y la menor mediana se alcanza en la opción de Humanidades, mientras las más altas se logran en la opción de Biosanitarios para la media y Tecnológico para la mediana. Respecto a la dispersión, esta es más alta en las opciones Biosanitario y Tecnológico. Si exceptuamos los casos en que la nota ha sido un cero, una en la opción de Biosanitario y otra en Sociales, la dispersión prácticamente es la misma en todos los casos como se puede intuir en los gráficos siguientes.



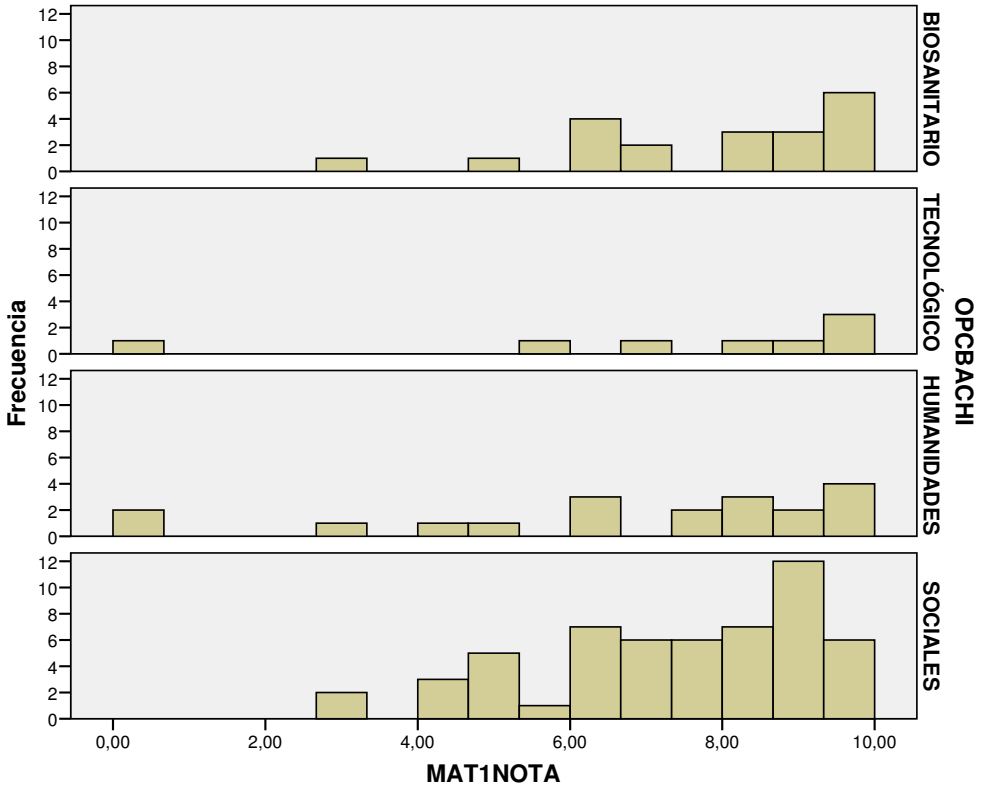
Tampoco son significativas las diferencias entre las notas medias, como se deduce del correspondiente contraste, cuyos resultados son:

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	3,749	3	1,250	,304	,823
Intra-grupos	403,302	98	4,115		
Total	407,051	101			

c) Opción de bachillerato y nota en la asignatura Introducción a la Matemática Económico- Empresarial.

	Mínimo	Máximo	Rango	Media	D. Típica	C.V.	Mediana
BIOSANITARIO	3	10	7	7,895	1,959	0,248	8,5
TECNOLÓGICO	0	10	10	7,438	3,417	0,459	8,5
HUMANIDADES	0	10	10	6,663	3,079	0,462	7,7
SOCIALES	2,7	10	7,3	7,367	1,852	0,251	7,7

En este apartado observamos cómo la nota media de Matemáticas alcanza el mínimo en la opción de Humanidades con un valor 6,66, siendo la nota para las demás opciones superior o igual a 7,36. La nota más alta, de 7,89, se logra en la opción de bachillerato de Biosanitario. Respecto a la dispersión, es en la opción de Sociales donde la desviación típica es menor, tal y como se desprende de la observación de los siguientes gráficos.



Como se observa en los cuadros siguientes, se rechaza (al 5% de significación) la igualdad de varianzas. Si a pesar de ello se realiza el análisis de la varianza, no se puede rechazar la igualdad de las medias de las notas.

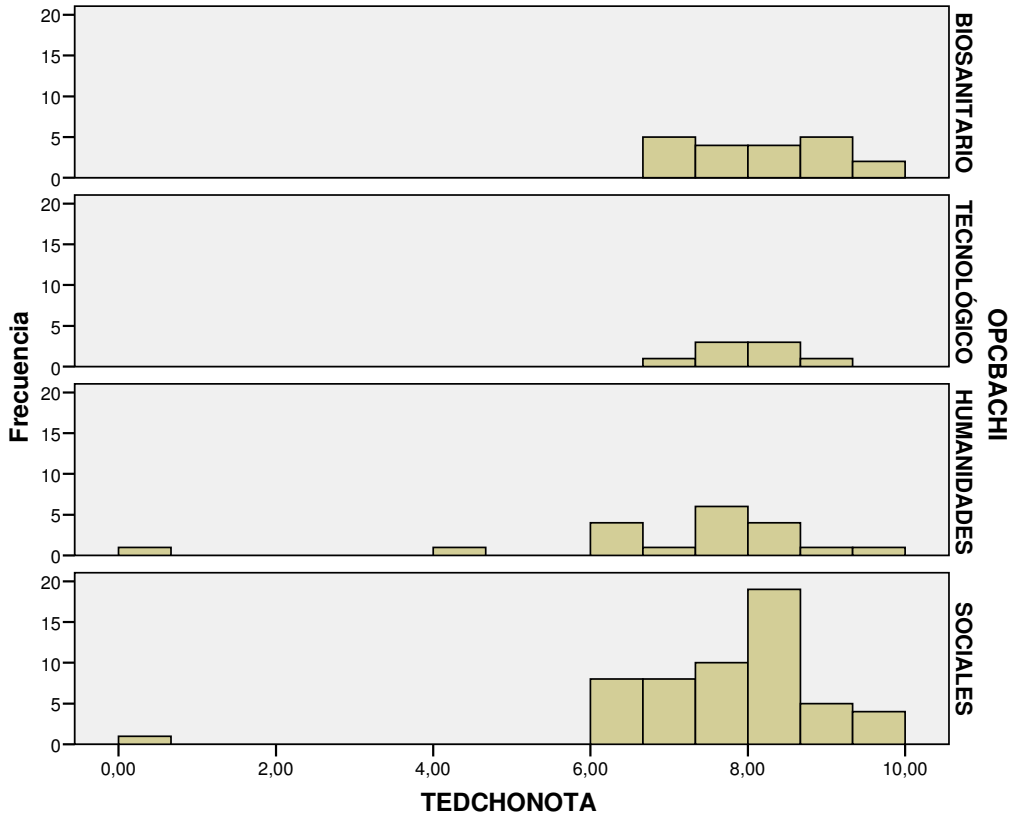
Estadístico de Levene	gl1	gl2	Sig.
3,191	3	98	,027

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	14,979	3	4,993	,958	,416
Intra-grupos	510,514	98	5,209		
Total	525,493	101			

d) Opción de bachillerato y nota en la asignatura Teoría del Derecho.

	Mínimo	Máximo	Rango	Media	D. Típica	C.V.	Mediana
BIOSANITARIO	7	9,5	2,5	8,11	0,888	0,109	8,15
TECNOLÓGICO	7,13	8,8	1,67	7,841	0,498	0,063	7,9
HUMANIDADES	0	9,5	9,5	7,001	2,092	0,299	7,5
SOCIALES	0	9,6	9,6	7,669	1,405	0,183	8

Como en los casos anteriores la nota media en Teoría del Derecho es superior para los alumnos que cursaron la opción de Biosanitario; en segundo lugar está la de Tecnológico, en tercero se sitúa la de Sociales y por último se encuentra la opción de Humanidades. La máxima dispersión también se alcanza para la opción de Humanidades, seguida por la de Sociales y con muy poca dispersión se sitúa la opción Biosanitario y Tecnológico, como se puede observar en los siguientes gráficos.



Una vez más no se puede rechazar la igualdad de las medias de las notas de las cuatro opciones como resulta del correspondiente contraste, cuyos resultados son:

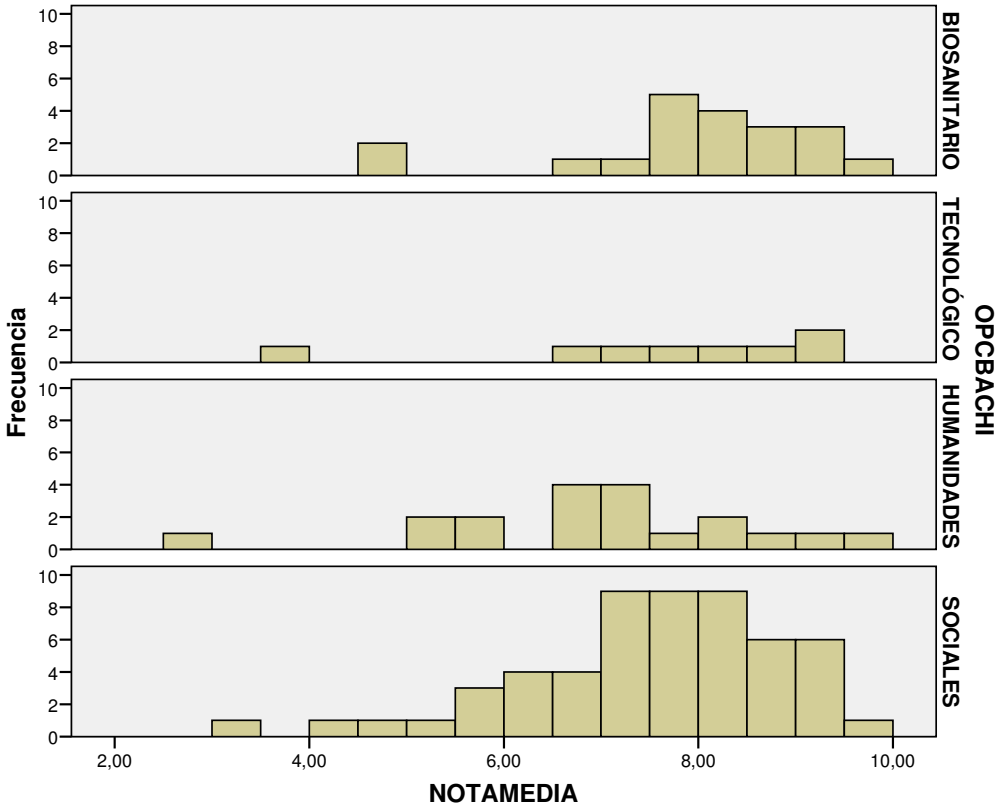
	Suma de cuadrados	Gl	Media cuadrática	F	Sig.
Inter-grupos	12,543	3	4,181	2,027	,115
Intra-grupos	202,145	98	2,063		
Total	214,687	101			

e) Opción de bachillerato y nota media.

Analizado el comportamiento de las notas de las tres asignaturas estudiadas por separado por último se repite el estudio entre la nota media global de las tres asignaturas con la opción de bachillerato realizada.

	Mínimo	Máximo	Rango	Media	D. Típica	C.V.	Mediana
BIOSANITARIO	4,67	9,6	4,93	7,885	1,308	0,166	8
TECNOLÓGICO	3,88	9,33	5,46	7,635	1,742	0,228	7,967
HUMANIDADES	2,67	9,67	7	6,914	1,657	0,239	7,167
SOCIALES	3,17	9,83	6,67	7,509	1,354	0,180	7,767

Como era de esperar, si ordenamos las opciones de bachillerato según la nota media, de mayor a menor, se tiene en primer lugar la opción de Biosanitario, en segundo lugar la Tecnológico, en tercer lugar la de Sociales y por último la opción de Humanidades, estando las notas comprendidas en el intervalo [6,91, 7,88].

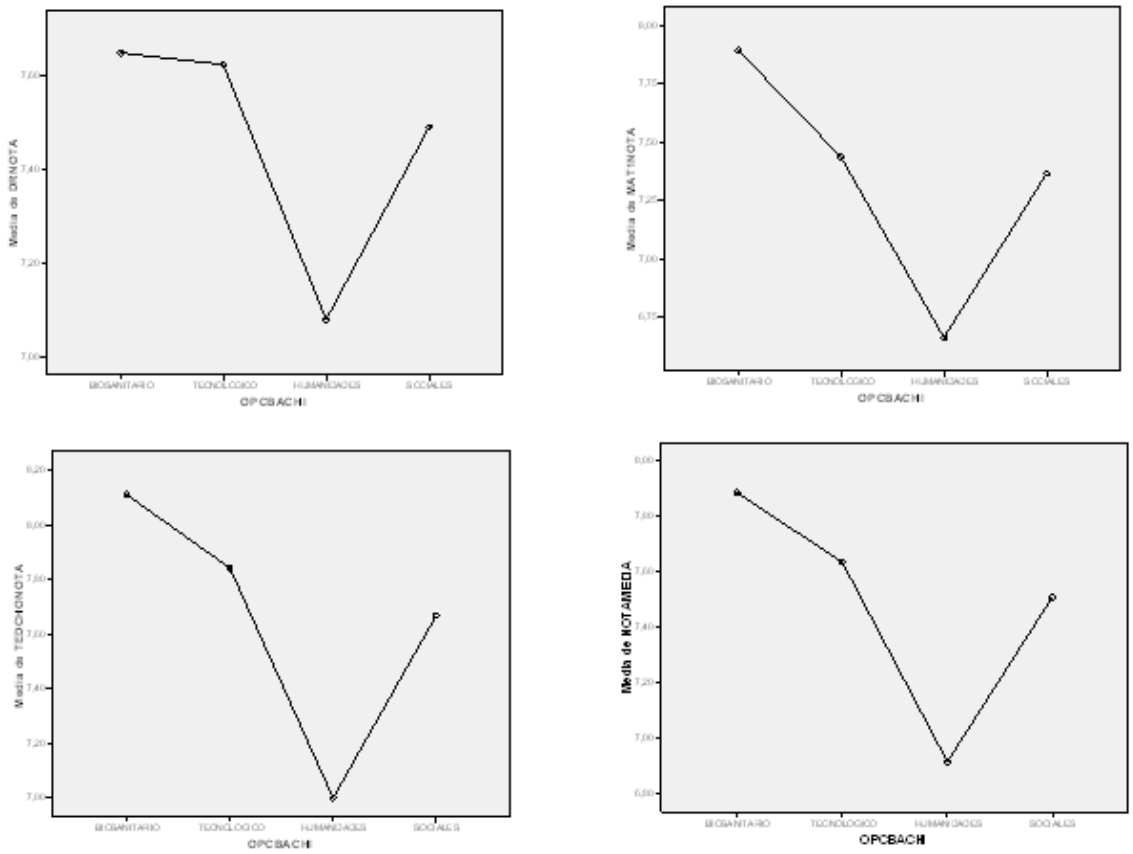


En este caso las dispersiones son muy similares.

Congruentemente con los tres casos anteriores no se puede rechazar que las medias de las notas por opción sean iguales, como se desprende del contraste realizado y cuyos resultados son:

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	9,596	3	3,199	1,550	,206
Intra-grupos	202,212	98	2,063		
Total	211,808	101			

Por último, antes de terminar este apartado debemos hacer una consideración. Si bien las notas no son significativamente distintas como para rechazar la igualdad de sus medias en cada opción de bachillerato, el orden de estas medias es el mismo en las tres asignaturas: la nota media más alta para los procedentes de Biosanitario, en segundo lugar los que eligieron la opción de Tecnológico, en tercero para la opción de Sociales y en último lugar los que proceden de Humanidades, como se puede comprobar en los siguientes gráficos



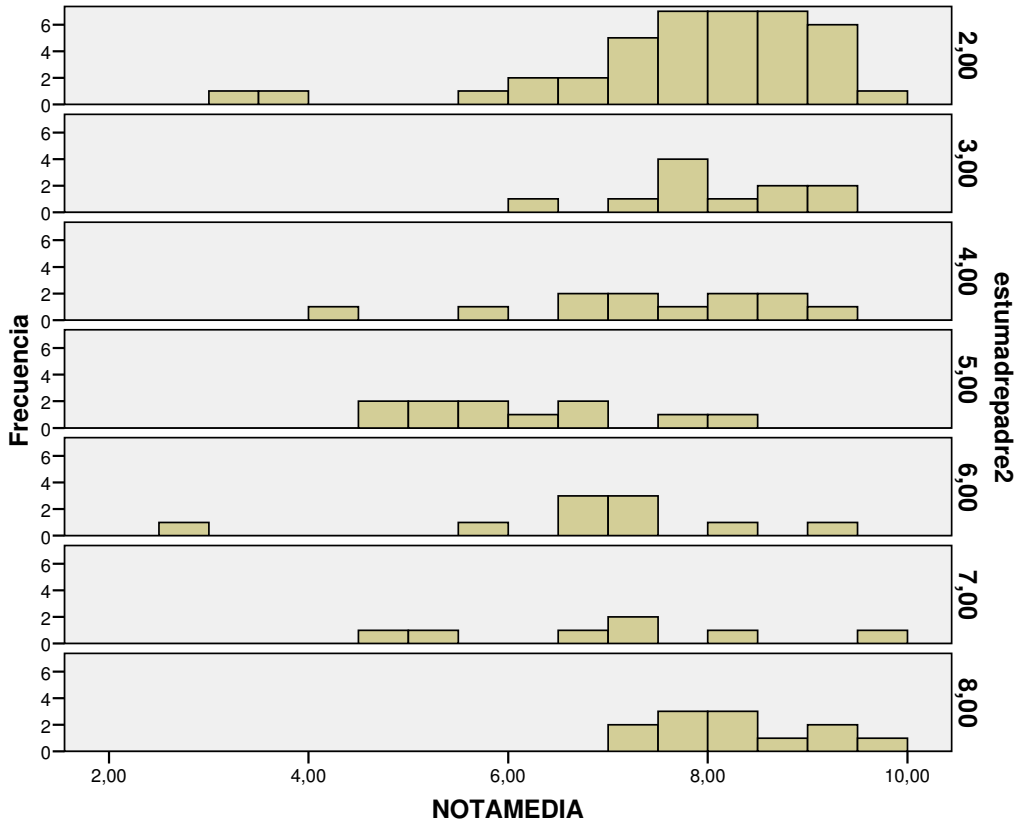
Es sumamente improbable que esta coincidencia en las cuatro ordenaciones sea puramente casual, pero dadas las características y limitaciones de estas encuestas, nos parece pronto para avanzar conclusiones al respecto. Anotamos el resultado y veremos si los próximos años los datos confirman la hipótesis de que los alumnos que llegan a estos estudios tienen ese orden de preparación para afrontarlos.

(iv) efecto del "nivel de estudios de los padres" y la nota media obtenida.

En este apartado se analiza el efecto de la característica "ESTUMADREPADRE"[1] sobre la nota media obtenida en las asignaturas del primer cuatrimestre.

De la observación de la representación gráfica para los 7 grupos a los que se da lugar, no se obtienen conclusiones claras. Se advierte que para el primer grupo (ambos padres con titulación superior) el rango es mayor que para el último (niveles de estudios bajos). Sin embargo, en este último la nota media mínima es la más alta de entre todos los grupos considerados.

[1] A efectos operativos se han agrupado aquellos casos en los que la característica "ESTUMADREPADRE" tomaba el valor 8, 9 y 10.



La tabla que sigue recoge algunos de los estadísticos para cada uno de los grupos, deduciéndose que los mayores valores medios se obtienen para aquellos grupos de alumnos cuyos padres tienen altos niveles de estudios (2, 3 y 4) o niveles de estudios bajos (7 y 8), con mayor dispersión en los grupos 6 y 7, correspondientes a situaciones intermedias.

	Mínimo	Máximo	Rango	Media	D. Típica	C.V.	Mediana
NIVEL PADRES=2	3,17	9,6	6,43	7,802	1,363	0,175	8
NIVEL PADRES=3	6	9,37	3,37	8,061	0,962	0,119	7,9
NIVEL PADRES=4	4,17	9,17	5	7,356	1,466	0,199	7,567
NIVEL PADRES=5	4,67	8,17	3,5	6,091	1,156	0,189	5,933
NIVEL PADRES=6	2,67	9,33	6,67	6,761	1,747	0,258	6,917
NIVEL PADRES=7	4,87	9,83	4,97	7,024	1,676	0,239	7
NIVEL PADRES=8 o más	7,13	9,67	2,53	8,267	0,815	0,175	8,167

La diferencia de las medias de las notas es significativa, como se deduce del contraste realizado, cuyos resultados son:

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	42,671	6	7,112	3,953	,001
Intra-grupos	169,115	94	1,799		
Total	211,786	100			

Si se realiza el análisis post-hoc Scheffé se tiene que:

ESTUMADRE-PADRE2	ESTUMADRE-PADRE2	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
		Límite inferior	Límite superior	Límite inferior	Límite superior	Límite inferior
2,00	3,00	-,25877	,45665	,999	-1,9166	1,3990
	4,00	,44628	,44148	,984	-1,1564	2,0490
	5,00	1,71092(*)	,45665	,038	,0531	3,3687
	6,00	1,04117	,47422	,570	-,6804	2,7628
	7,00	,77802	,54954	,917	-1,2170	2,7730
	8,00	-,50150	,47422	,980	-2,2231	1,2201
3,00	2,00	,25877	,45665	,999	-1,3990	1,9166
	4,00	,70505	,55989	,952	-1,3276	2,7377
	5,00	1,96970	,57193	,077	-,1066	4,0460
	6,00	1,29994	,58606	,557	-,8277	3,4275
	7,00	1,03680	,64851	,860	-1,3175	3,3911
	8,00	-,24273	,58606	1,000	-2,3703	1,8849
4,00	2,00	-,44628	,44148	,984	-2,0490	1,1564
	3,00	-,70505	,55989	,952	-2,7377	1,3276
	5,00	1,26465	,55989	,535	-,7680	3,2973
	6,00	,59489	,57431	,982	-1,4901	2,6799
	7,00	,33175	,63792	1,000	-1,9841	2,6476
	8,00	-,94778	,57431	,841	-3,0327	1,1372
ESTUMADRE-PADRE2	ESTUMADRE-PADRE2	Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95%	
		Límite inferior	Límite superior	Límite inferior	Límite superior	Límite inferior
5,00	2,00	-1,71092(*)	,45665	,038	-3,3687	-,0531
	3,00	-1,96970	,57193	,077	-4,0460	,1066
	4,00	-1,26465	,55989	,535	-3,2973	,7680
	6,00	-,66976	,58606	,970	-2,7974	1,4578
	7,00	-,93290	,64851	,911	-3,2872	1,4214
	8,00	-2,21242(*)	,58606	,035	-4,3400	-,0848
6,00	2,00	-1,04117	,47422	,570	-2,7628	,6804
	3,00	-1,29994	,58606	,557	-3,4275	,8277
	4,00	-,59489	,57431	,982	-2,6799	1,4901
	5,00	,66976	,58606	,970	-1,4578	2,7974
	7,00	-,26314	,66100	1,000	-2,6628	2,1365
	8,00	-1,54267	,59985	,367	-3,7203	,6350
7,00	2,00	-,77802	,54954	,917	-2,7730	1,2170
	3,00	-1,03680	,64851	,860	-3,3911	1,3175
	4,00	-,33175	,63792	1,000	-2,6476	1,9841
	5,00	,93290	,64851	,911	-1,4214	3,2872
	6,00	,26314	,66100	1,000	-2,1365	2,6628
	8,00	-1,27952	,66100	,710	-3,6792	1,1202
8,00	2,00	,50150	,47422	,980	-1,2201	2,2231
	3,00	,24273	,58606	1,000	-1,8849	2,3703
	4,00	,94778	,57431	,841	-1,1372	3,0327
	5,00	2,21242(*)	,58606	,035	,0848	4,3400
	6,00	1,54267	,59985	,367	-,6350	3,7203
	7,00	1,27952	,66100	,710	-1,1202	3,6792

Estos resultados indican que hay claras diferencias entre las notas según los niveles de estudios de los padres; de hecho son más acusadas entre niveles de estudios altos y medios y entre niveles bajos y medios.

El primer resultado (diferencias entre niveles de estudios de los padres altos y los medios, a favor de los niveles altos), se podrían deber al efecto beneficioso que supone para los hijos el que los padres tengan estudios superiores.

Pero las otras diferencias importantes (entre niveles de estudio de los padres bajos y los medios, a favor de los bajos), no pueden obedecer a las mismas causas. Podrían ser debidas a la mayor selección. Si los niveles de estudios de los padres son muy bajos, los alumnos solo acceden a estas titulaciones universitarias si son especialmente brillantes.

4. CONCLUSIONES

Éste ha sido el primer paso de un estudio que pretendemos continuar los próximos cursos. Como ya se ha dicho en apartados anteriores, es pronto para llegar a conclusiones.

De momento consideramos plenamente válido el nivel descriptivo pero no el inferencial, aunque éste último sí resulta útil para hacer conjeturas y establecer hipótesis que deberán ser contrastadas con información más amplia que se pueda recopilar los próximos años.

Un breve listado de hipótesis que se pueden deducir del estudio sería:

1) ¿Es cierto que el régimen jurídico público o privado del centro de estudios de bachillerato es independiente de la vía de acceso (ADE o Derecho), de la opción de bachillerato elegida y de la nota de acceso a la Universidad y de las notas obtenidas en la titulación?. Por el contrario ¿es cierto que hay una clara dependencia entre el régimen jurídico del colegio y el nivel de estudios de los padres?.

2) ¿Es cierto que la vía de acceso no es relevante para las notas obtenidas en la titulación y la opción de bachillerato cursada?.

3) ¿Es cierto que la opción de bachillerato cursada es independiente de la nota de acceso a la titulación?. ¿Es cierto que dicha opción es independiente de las notas obtenidas en la titulación?

4) ¿Es cierto que el nivel de estudios de los padres tiene relevancia en las notas obtenidas por los hijos y que esta dependencia lo es en el sentido que en su momento se apuntó?

Además, la contestación de estas preguntas no hace sino generar otras nuevas. Si entre dos características existe cierto grado de dependencia, cabe preguntarse siempre: ¿en qué grado?, ¿cuál es la relación causa-efecto?. Es decir, de cualquier conclusión siempre cabe preguntar cuanto y por qué.

Pero estas últimas respuestas ya son objetivo de un proyecto más ambicioso, que debe durar varios años. Esperamos que en los próximos congresos podamos presentar resultados más avanzados en esta investigación.

BIBLIOGRAFÍA.

CASAS SÁNCHEZ, J.M (1997).: "Inferencia Estadística", Editorial Centro de Estudios Ramón Areces, Madrid.

DASÍ, A.; GARCÍA, J.; HUGUET, A.; JUAN, R.; MONTAGUD, M.D.; GOLLNERT, G. (2007) "Innovación educativa en la Universidad: ADE-Derecho", PUV Universitat de València, Valencia.

ESTEBAN, J; BACHERO, J.M.; BLASCO, O.; COLL, V.; DÍEZ, R.; IVARS, A.; LÓPEZ, M.I.; ROJO, C.; RUIZ, F. (2005) "Estadística Descriptiva y nociones de Probabilidad", Thomson, Madrid.

ESTEBAN, J; BAHCERO, J.M.; IVARS, A.; LÓPEZ, M.I.; ROJO, C.; RUIZ, F.; (2008). "Curso Básico de Inferencia Estadística", Repro-Exprés, Valencia.

MEC (2006). "Propuestas para la renovación de las metodologías educativas en la Universidad", Ministerio de Educación y Ciencia, Madrid

PEREZ, C. (2005) "Métodos estadísticos avanzados con SPSS", Thomson, Madrid.

VISAUTA, B. (1997) "Análisis estadístico con SPSS para Windows", McGraw-Hill, Madrid.

LA DIDÁCTICA DE LA ECONOMÍA EN EL ESPACIO EUROPEO DE EDUCACIÓN SUPERIOR

MANUEL SALAS VELASCO
E-mail: msalas@ugr.es
MARÍA TERESA SÁNCHEZ MARTÍNEZ
NOELINA RODRÍGUEZ FERRERO
Departamento de Economía Aplicada
Universidad de Granada
MARÍA ISABEL SÁNCHEZ Y SÁNCHEZ-AMAYA
Departamento de Economía Aplicada I
Universidad del País Vasco

Resumen

La participación de España en el desarrollo del Espacio Europeo de Educación Superior (EEES) presenta un número importante de desafíos para los docentes. Movernos a una enseñanza basada en competencias es un cambio difícil de lograr pero fundamental en el contexto de las reformas actuales. En un sistema educativo en el que la clase magistral es la metodología docente predominante, la enseñanza centrada en el aprendizaje constituye en sí misma un cambio cultural de importantes consecuencias. El objetivo de este trabajo es proporcionarles a los profesores de *economía* técnicas prácticas para la enseñanza. En concreto, incluye sugerencias para la planificación, preparación e impartición de clases magistrales, seminarios, prácticas, estudio de casos, debates y enseñanza en el aula de Informática. También contiene consejos prácticos para evaluar a los estudiantes.

Spain's participation in the development of the European Higher Education Area (EHEA) presents a number of challenges for academics. The shift to competency-based teaching is a difficult change to achieve but fundamental in the context of the current reforms. In a system in which the lecture continues to be the dominant pedagogical practice, learning-centered teaching constitutes on its own a cultural change of great consequences. The purpose of this article is to provide lecturers in economics with practical techniques for teaching. Specifically, it includes suggestions on planning, preparation and delivery of lectures, seminars, tutorials, case studies, debates, and computers' lab. It contains some basic tips on assessment as well.

Palabras Clave: EEES, enseñanza basada en competencias, innovación docente
Área Temática: Didáctica de la Economía y Metodología.

1. INTRODUCCIÓN: QUÉ, CÓMO Y PARA QUIÉN[1]

«La economía es el estudio de cómo las sociedades utilizan recursos escasos para producir bienes valiosos y distribuirlos entre diferentes personas» (Samuelson y Nordhaus, 2006, p. 4). Los docentes de esta disciplina nos encontramos actualmente con tres problemas básicos para los que debemos encontrar respuestas: **qué** enseñar –es decir, qué contenidos deben formar parte de un curso de *economía*–; **cómo** hacerlo –esto es, qué estrategias docentes (métodos, recursos y materiales didácticos) debemos usar–; y **para quién** –es decir, cuál es el perfil de estudiante al que deben ir dirigidas nuestras clases–. En este trabajo nos centramos en responder al segundo interrogante –dejando para otros trabajos futuros las respuestas a los otros dos interrogantes–. Este aspecto es clave hoy día ya que el cambio de metodología en la enseñanza que se promueve en la convergencia hacia un Espacio Europeo de Educación Superior (EEES) es uno de los temas que suscitan más discusión dentro del colectivo de profesores universitarios.

[1] Este trabajo refleja gran parte de las ideas puestas en práctica en el proyecto de innovación docente *Enseñanza basada en competencias* coordinado por Manuel Salas Velasco y María Teresa Sánchez Martínez desde el curso académico 2008-2009. El proyecto nació como respuesta a las limitaciones de la enseñanza tradicional.

Hasta ahora, la enseñanza de la *economía* en las aulas universitarias ha puesto su énfasis en el aprendizaje de la materia, de su conocimiento y comprensión, lo que supone que el contenido mismo es lo realmente importante; el concepto de buen profesor se centra en indicadores como "explica con claridad" o "se le entiende bien". En el nuevo marco universitario, el papel que el estudiante ocupará será central, pues gran parte de su aprendizaje ocurrirá de manera autónoma. Caminamos hacia un modelo de docencia más centrado en los procesos de aprendizaje de los alumnos y que faciliten el desarrollo de competencias. La enseñanza basada en competencias, que se preocupa por desarrollar habilidades, destrezas y valores en los estudiantes integrando los contenidos en conocimiento personal y autónomo, requiere una metodología distinta que la tradicional de la que procedemos. Esto, sin duda, supone un reto para los docentes que deberán capacitarse para adaptarse a este enfoque cuyo papel requiere destrezas y habilidades distintas. No se trata de despreciar modelos tradicionales, ni siquiera el de la clase magistral, que deberá seguir en la universidad, pero sí que tenemos que hacer un ejercicio serio de reflexión para mejorar lo que estamos haciendo. Obviamente, los cambios en los métodos docentes requieren también de nuevos modos de evaluar, ya que habrá que evaluar las competencias y no sólo el conocimiento de los estudiantes. Evaluar competencias significa evaluar el desempeño de los estudiantes en la aplicación y resolución de situaciones en las que deben aplicar los conocimientos de modo adecuado, y demostrar las actitudes y valores adquiridos de forma integrada. En definitiva, el papel del profesorado es una pieza clave para el éxito de la convergencia europea. Requiere algunos hábitos nuevos y, sobre todo, un cambio de percepción del propio papel del docente y las funciones a desempeñar.

2. ESTRATEGIAS DOCENTES PARA LA ENSEÑANZA DE LA ECONOMÍA Y COMPETENCIAS QUE SE DESARROLLAN EN EL ALUMNADO A PARTIR DE LAS MISMAS[2]

El nuevo escenario educativo derivado del proceso de Bolonia exige cambios profundos en la forma en la que se han venido orientando los procesos de enseñanza-aprendizaje en la enseñanza de la *economía*. Permitirles a los estudiantes pensar como economistas incluye no solamente enseñarles teoría económica –o habilidades para la resolución de problemas de micro y macroeconomía–, sino también otra serie de competencias o habilidades requeridas por los empleadores a los recién licenciados: iniciativa, liderazgo, trabajo en equipo, creatividad, saber buscar información, comunicación oral y escrita, motivación, entre otras (Salas Velasco, 2000; Gómez García *et al*., 2008).

Esta parte del trabajo sugiere que una forma importante para desarrollar estas competencias es proporcionarles a los estudiantes oportunidades crecientes para llegar a estar más activamente "enganchados" en la aplicación de la *economía*. Clases magistrales, acompañadas de resolución de problemas en pequeños grupos, realización de ensayos sobre problemas económicos actuales, debates o prácticas con ordenador, constituye el paquete primario de técnicas de instrucción usadas para enseñar *economía*. Si queremos que nuestros estudiantes piensen como economistas necesitamos, necesariamente, considerar seriamente alternativas a los modos tradicionales de instrucción y evaluación.

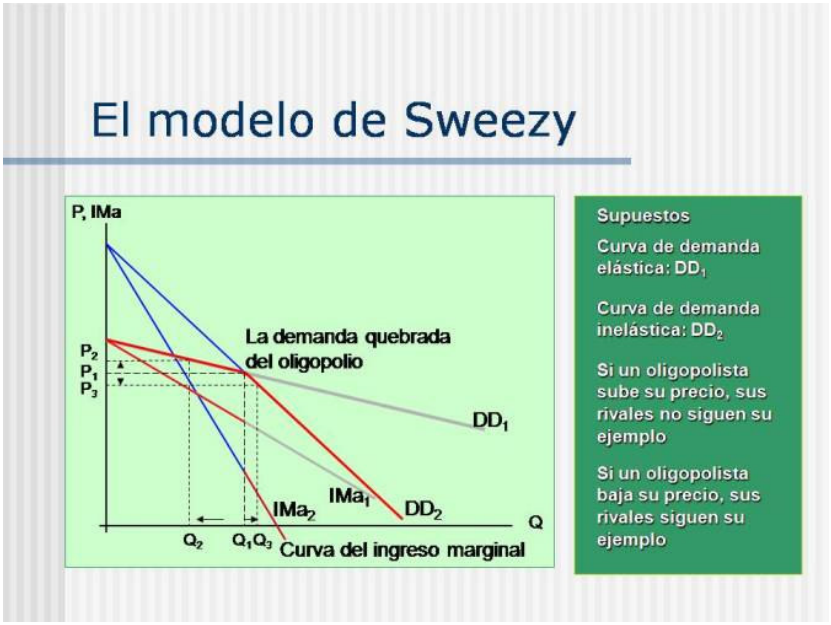
2.1. Clases magistrales de economía

«Sin duda, pasar de un enfoque centrado en la enseñanza a otro centrado en el aprendizaje, contrasta con la práctica habitual de muchos docentes, para quienes la clase magistral ha sido la forma de enseñar, la única y verdadera» (Benito y Cruz, 2006, p. 33). En la enseñanza de la *economía*, es la metodología didáctica más utilizada en la gran mayoría de las universidades españolas. Aunque este método docente resulta imprescindible en un principio para la transmisión de los conceptos fundamentales de micro y macroeconomía, por lo que puede ser una herramienta muy útil si se sabe utilizar en su justa medida, no obstante debe ir acompañado necesariamente de otros métodos de enseñanza para conseguir los mejores resultados en la adquisición de conocimientos del discente y en el desarrollo de competencias.

[2] Remitimos al lector al Capítulo III del libro de Salas Velasco (2007) en caso de querer tener más detalles de los expuestos en este apartado.

La idoneidad de esta estrategia es mayor cuanto más teóricos y abstractos sean los contenidos, o cuando se trata de planteamientos introductorios y panorámicos cuya finalidad sea establecer esquemas generales o marcos de referencia. En definitiva, las estrategias expositivas se recomiendan para trabajar contenidos de conceptos y principios, combinando la exposición propiamente dicha con apoyo de medios audiovisuales. Por ejemplo, una de las principales herramientas del análisis económico es el uso de gráficos. Frecuentemente se presenta el gráfico de alguna situación y luego se muestra qué ocurre si se presenta algún cambio, lo cual muchas veces implica el desplazamiento de los gráficos y al final se tiene la pizarra llena de curvas, y al estudiante le resulta a veces un tanto complicado seguir la exposición. Por ello, la opción de animaciones, permite hacer una exposición clara y amena de los conceptos teóricos que se enseñan en los cursos de *economía*. El uso de PowerPoint, flash, gifs animados y otros recursos permite animar los gráficos, facilitando así la comprensión de los análisis que en ellos se basan y que requieren trazar varias gráficas, desplazarlas, etc. (véase Figura 1)[3].

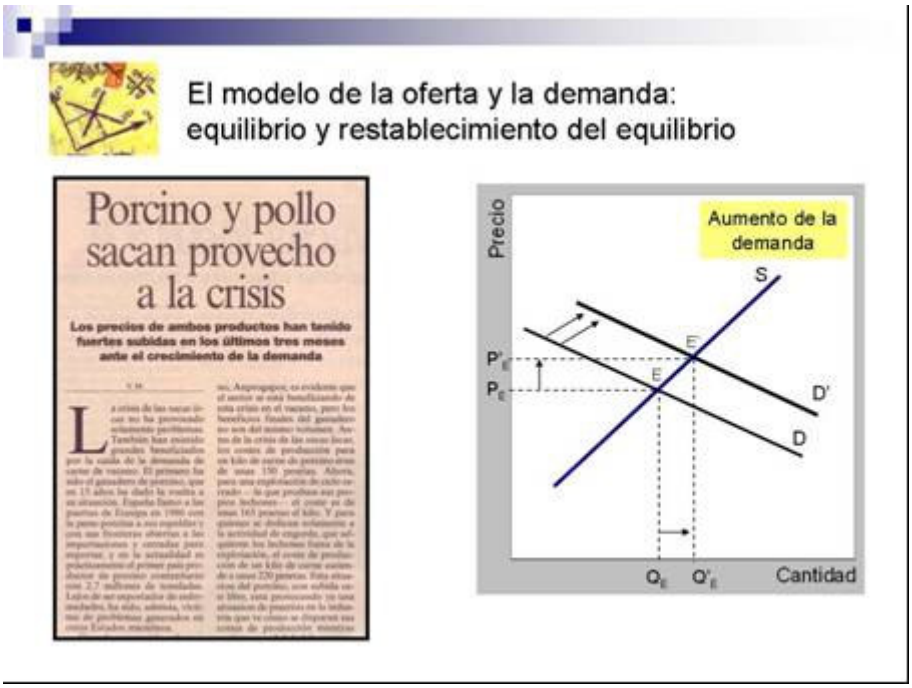
Figura 1. Alternativas al encerado y a la tiza: el uso de las nuevas tecnologías en la enseñanza de la economía



La enseñanza de la *economía* aún necesita desarrollar métodos, recursos y materiales didácticos que faciliten la tarea del profesor y la asimilación por parte de los alumnos de la disciplina. Una de las grandes dificultades desde el punto de vista del estudiante es la adecuada comprensión de los planteamientos gráficos, muy usuales en el conocimiento económico como hemos adelantado. Entre estos gráficos, destaca el modelo de la oferta y la demanda. Aplicar el instrumental de la oferta y la demanda a situaciones reales es relativamente fácil usando recortes de prensa. En la Figura 2 observamos que un aumento de la demanda, cuando las condiciones de oferta no cambian, lleva a un nuevo equilibrio de mercado en el que el nuevo precio de equilibrio es mayor y también lo es la cantidad de equilibrio. Si este análisis gráfico, cuando el profesor hace su exposición, va acompañado de noticias reales de prensa, el estudiante comprende mejor los aspectos teóricos y ve la utilidad de los modelos que aprende. En el caso de la Figura 2, el análisis gráfico explica la subida en el precio del porcino y pollo ante un crecimiento de su demanda (como indica el titular del recorte de prensa), debido a la crisis de las vacas locas que llevó a una disminución del consumo de carne de ternera (caída en la demanda de carne de vacuno como dice el recorte de prensa).

[3] El encerado y la tiza son dos recursos en vías de extinción en la enseñanza de la *economía*. Las Tecnologías de la Información y la Comunicación (TIC) se están convirtiendo en un recurso didáctico imprescindible en las exposiciones teóricas, destacando la pizarra digital (ordenador con un vídeo-proyector) y, un paso más allá, la pizarra digital interactiva (PDI), que permite subrayar, dibujar, tachar o cambiar de posición los elementos que se proyectan sobre el tablero, ya que éste es capaz de detectar la posición y presión de un puntero.

Figura 2. La economía es noticia: la prensa como recurso didáctico



Aunque el principal argumento a favor de la clase magistral es la denominada "autoridad científica del profesor", que con su exposición verbal de los contenidos de la materia facilita a los estudiantes una información organizada y adecuada a la finalidad, el uso de este método expositivo debe ir acompañado de estrategias docentes para la estimulación creativa. Por ejemplo, el dilema del prisionero, posiblemente el juego más conocido y estudiado en la teoría de juegos, tradicionalmente se enseña a partir de dos delincuentes, A y B, que son detenidos y encerrados en celdas de aislamiento de forma que no pueden comunicarse entre ellos. La policía le hace el siguiente ofrecimiento a cada uno de ellos: si ambos confesáis, cada uno queda en prisión 10 años; si ninguno de los dos confesáis, os caerá un año a cada uno; si uno de los dos confiesa pero el otro no, el que confiesa quedará libre pero el otro quedará entre rejas 20 años. Los docentes deberían mostrar una disposición favorable hacia la creatividad en la enseñanza del dilema del prisionero identificando, por ejemplo, a A y B con Ollie y Stand (el gordo y el flaco), como se muestra en la Figura 3.

Figura 3. Iconos cinematográficos en la enseñanza de la economía



En resumen, la clases magistrales deben utilizarse para dar da una visión general de cada tema, resaltando los aspectos nuevos o de especial complejidad. El problema con esta forma de enseñar no es su uso, sino el abuso que suele hacerse en los cursos introductorios de *economía*. Son principalmente tres los requisitos que deben darse para utilizar con eficacia esta metodología didáctica:

1.- Partir de una estructura clara y organizada de lo que se va a enseñar, lo cual se puede conseguir con la práctica docente y con una planificación de las clases que previamente habrá que preparar.

2.- Tener en cuenta los conocimientos y competencias que el alumno ya posee.

3.- Lograr una motivación positiva y actitud favorable en los alumnos.

Competencias que se desarrollan:
– Comprender las bases teóricas y epistemológicas de la disciplina.
– Conocer los métodos de investigación en la disciplina.
– Saber analizar e interpretar información.
– Ganas de aprender.

2.2. Seminarios

El seminario es un método de enseñanza tradicional en la universidad española, aunque poco usado en la enseñanza de la *economía*. Va dirigido normalmente a pequeños grupos con el objeto de lograr el conocimiento completo y específico sobre un tema. Su utilización es recomendada, bien para plantear los temas más polémicos (u opinables) del programa de la asignatura, bien para tratar de manera monográfica algún tema de actualidad relacionado con los contenidos de la materia. Por ejemplo, podría plantearse un seminario (de unas tres horas) sobre los modelos de crecimiento económico y el papel que actualmente juega el capital humano y la productividad del trabajo en el crecimiento.

Como método didáctico en la enseñanza de la *economía* presenta las siguientes ventajas:

- Permite la actualización, el perfeccionamiento y la profundización en un tema.
- Es útil para buscar información y discutir en grupo.
- Permite analizar a fondo datos e informaciones y confrontar puntos de vista.
- El aprendizaje logrado de esta forma es de mayor calidad.

Competencias que se desarrollan:
– Hábitos de trabajo intelectual.
– Obtener un enfoque crítico de la disciplina.
– Comunicación en una lengua extranjera.

Aunque muchas veces el seminario se le parece a una clase magistral, no obstante hay algunas diferencias. Por un lado, en lo relativo al número de asistentes; lo ideal es trabajar con unos 15 alumnos. Por otro lado, en lo relativo a la relación profesor-alumno; en el seminario hay normalmente una mayor participación de los estudiantes en la discusión e intercambio de ideas. A diferencia de una clase magistral, lo más habitual es que los alumnos interrumpan al profesor.

La aplicación de este método incrementa el esfuerzo del docente porque es él quien prepara normalmente los materiales: artículos de revistas de investigación, capítulos de libros, material de prensa, etc. Pero la efectividad de este método de enseñanza exige también una implicación importante de los alumnos, principalmente porque deben trabajar con antelación los materiales que el profesor pone a su disposición y, en ocasiones, porque deben complementar el material del seminario mediante una búsqueda en Internet, biblioteca, etc. Es fundamental que el profesor cuente con habilidades para estimular y dinamizar el grupo de manera efectiva hacia los objetivos propuestos, y cerciorarse de que el grupo es capaz de analizar críticamente y constructivamente.

2.3. Prácticas en clase (en grupos pequeños de trabajo)

La exposición teórica de la materia (clase magistral y/o seminario) debe ir acompañada de otras técnicas docentes, como las prácticas en clase, que incentiven la atención del estudiante,

siendo especialmente útil este modo de aprendizaje para evitar que la atención del discente decaiga.

La realización de prácticas de *economía* es imprescindible para que el estudiante aplique los conceptos teóricos a situaciones reales. Recomendamos que en las sesiones prácticas los estudiantes resuelvan en el aula los ejercicios y problemas propuestos con antelación por el profesor en pequeños grupos (trabajo cooperativo). Existen distintas formas de organizar el trabajo cooperativo, pudiendo hablar principalmente de grupos informales y grupos formales. El aprendizaje cooperativo informal es el que se organiza de forma temporal, no tiene solución de continuidad, y se forma habitualmente para llevar a cabo una determinada tarea académica en el transcurso de una clase. Es el recomendable para este tipo de actividades.

2.3.1. Ejercicios y resolución de problemas

Para que una sesión de prácticas sea ocasión para el aprendizaje tiene que estar bien planificada y bien fundada teóricamente –los alumnos deben haber recibido previamente una formación teórica sobre el objeto de la práctica–. La preparación de las clases prácticas comprende: (i) el establecimiento de objetivos; (ii) la elección de las actividades; (iii) la preparación de los materiales; y (iv) el diseño del procedimiento de trabajo. El profesor, para preparar bien la práctica, debe pensar qué necesitan los estudiantes para realizar la práctica, qué destrezas hay que emplear, qué recursos se necesitan, cuál es el producto esperado con la tarea y cuáles son los criterios de evaluación de los resultados. Asimismo, los estudiantes tienen que percibir las prácticas que se les proponen realizar como relevantes y significativas para sus intereses, si no la motivación por su realización decae; y en segundo lugar, para que la práctica sea ocasión de aprendizaje, debe ser evaluada en su proceso de realización.

Competencias que se desarrollan:
– Capacidad de análisis.
– Capacidad para la resolución de problemas.
– Capacidad de aplicar los conocimientos en la práctica.
– Capacidad intelectual.

2.3.2. Estudio de casos

El método del caso sumerge a los alumnos en problemas y escenarios reales en los que deben decidir una estrategia, plan o decisión que permita dar respuesta al dilema planteado: «El método del caso, como método activo de aprendizaje, parte de la descripción de una situación real, que normalmente tiene que ver con una decisión, un desafío, una oportunidad, un problema o cualquier otra cuestión, afrontada por un alumno o grupo de alumnos en un aula o entorno de aprendizaje concreto, en un momento determinado» (Benito y Cruz, 2005, p. 50).

En la enseñanza de la *economía*, para que sea eficaz como estrategia docente, debe tratarse de la descripción de un hecho real, que permita conocerlo, interpretarlo, crear diferentes hipótesis, realizar un diagnóstico, para llegar a la resolución del mismo y, en ocasiones, a encontrar posibles procedimientos alternativos de solución.

Permite desarrollar las siguientes **competencias**:
– Pensamiento crítico.
– Toma de decisiones.
– Confianza en sí mismo y en la expresión de ideas.

2.4. El uso del aula de informática en la docencia de la economía

Las aulas de Informática están concebidas como herramienta de apoyo a la docencia y formación del alumnado. El análisis de datos con ordenador, usando principalmente los programas informáticos SPSS (Figura 4) y Microsoft Excel (Figura 5), permite conocer e interpretar la realidad económica. La econometría –integración de la teoría económica, las matemáticas y las técnicas estadísticas– permite probar hipótesis sobre fenómenos económicos, al tiempo que potencia la capacidad analítica del estudiante para interpretar las relaciones cuantitativas entre variables económicas. Hoy día, la disponibilidad de estos programas

estadísticos universalmente extendidos hace bastante asequible la econometría y otras técnicas estadísticas, que acercan al estudiante a los enfoques, los métodos y las herramientas del análisis económico.

Competencias que se desarrollan:
– Manejo de un ordenador.
– Capacidad de aplicar los conocimientos en la práctica.
– Habilidad para analizar y buscar información proveniente de fuentes diversas.

Figura 4. El ordenador como herramienta docente: cálculo de elasticidades de demanda usando el programa SPSS

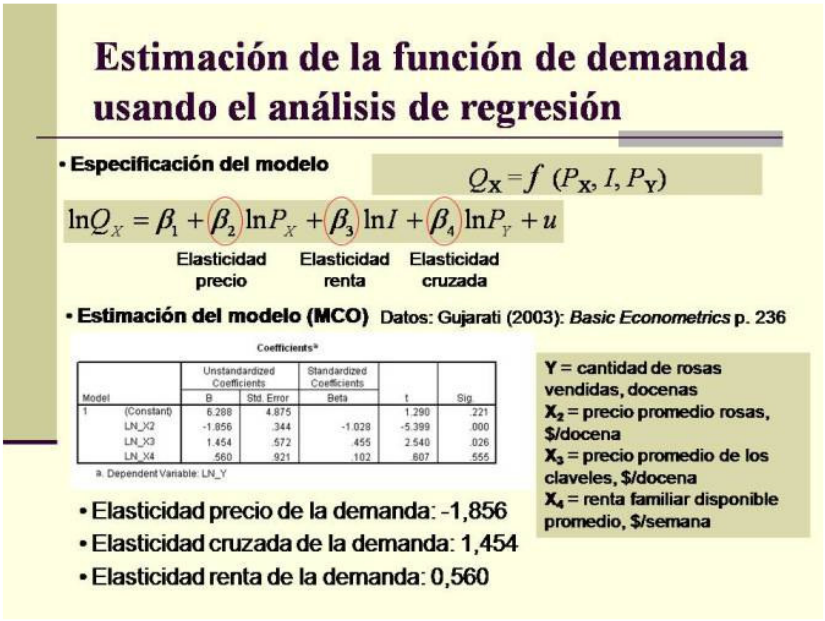
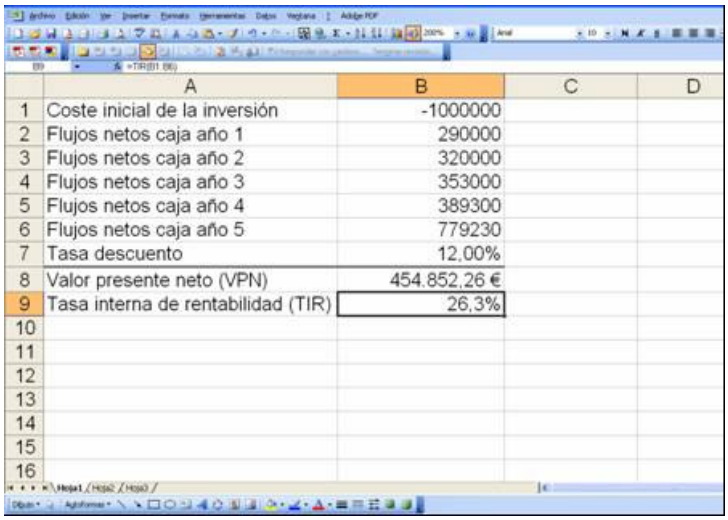


Figura 5. El cálculo del valor presente neto y de la tasa interna de rentabilidad con Microsoft Office Excel



	A	B	C	D
1	Coste inicial de la inversión	-1000000		
2	Flujos netos caja año 1	290000		
3	Flujos netos caja año 2	320000		
4	Flujos netos caja año 3	353000		
5	Flujos netos caja año 4	389300		
6	Flujos netos caja año 5	779230		
7	Tasa descuento	12,00%		
8	Valor presente neto (VPN)	454.852,26 €		
9	Tasa interna de rentabilidad (TIR)	26,3%		
10				
11				
12				
13				
14				
15				
16				

2.5. Debates

El debate es una situación en la que el profesor introduce a los alumnos en un proceso dialéctico a propósito de un tema adecuado para ello. El tema que se plantee ha de ser cuestionable para que surja el debate. También se exige que el tema sea conocido por todos para evitar improvisaciones; esta técnica implica el reparto previo de material especializado del tema

de que se trate (y/o búsqueda previa por parte del alumnado). Es aconsejable trabajar con grupos reducidos y su duración máxima debe ser de 60 minutos. El profesor actúa normalmente de moderador, por lo que debe manejar ciertas técnicas para dirigir grupos de discusión: controla la participación espontánea e imprevisible, da un tiempo limitado a cada expositor, evita que los participantes se aparten del tema, etc. El debate debe ser ante todo un ejercicio de desarrollo del pensamiento lógico y del rigor en la argumentación y no debe ser una situación donde cada uno deba adherirse a la tesis o antítesis que defienda, no se trata de valores personales sino de razonamiento lógico, de argumentación y rigor.

En la asignatura de *economía*, puede plantearse un debate, por ejemplo, sobre financiación universitaria en el EEES. Entre las preguntas a debatir tendríamos las siguientes: (i) ¿hasta qué punto está justificada la financiación pública de la enseñanza universitaria?; (ii) ¿qué fórmulas más racionales debería buscar el sector público para optimizar los recursos puestos a disposición de las universidades?; (iii) hoy por hoy, ninguna de nuestras universidades está entre las 50 mejores universidades europeas, ¿puede ser la financiación parte de la respuesta? El alumno debe preparar con suficiente antelación el debate, reflexionando sobre estas preguntas, buscando bibliografía que considere necesaria.

Competencias que se desarrollan:
– Comunicación oral.
– Capacidad crítica y autocrítica.
– Reconocimiento y respeto a la diversidad.
– Habilidades interpersonales.
– Liderazgo.

2.6. Talleres de trabajo

Los talleres de trabajo, muy frecuentes en el sistema educativo anglosajón, tienen como objetivo principal desarrollar habilidades (competencias, destrezas) en los estudiantes: habilidad para hablar en público, capacidad de negociación, dotes de mando y liderazgo, etc. A diferencia del seminario, los talleres suelen ser eminentemente prácticos. El profesor combina diferentes metodologías activas: debates, *role-playing* (juego de papeles), tormenta de ideas, simulaciones, etc.

Por ejemplo, en un taller para desarrollar la habilidad de hablar en público se proponen normalmente actividades –contar un cuento, autobiografía, etc.– que se graban con una cámara de vídeo y, posteriormente, el profesor analiza con este alumno y con el resto del grupo el vídeo. O, por ejemplo, en un taller para desarrollar la capacidad de negociación suele ser muy habitual usar el *role-playing*, donde el profesor y el alumno representan papeles (comprador/vendedor, entrevistador/entrevistado, etc.). Ni que decir tiene que para garantizar la efectividad de los talleres, éstos deben contar con un grupo reducido de asistentes.

Competencias que se desarrollan:
– Habilidades interpersonales.
– Comunicación oral.
– Toma de decisiones.
– Habilidad para hablar en público.
– Capacidad de negociación.

2.7. Las nuevas tecnologías en el proceso de enseñanza-aprendizaje

Las nuevas tecnologías de la información y la comunicación han tenido un gran impacto en la sociedad actual. Estas tecnologías auguran la progresiva desaparición de las restricciones de espacio y de tiempo en la enseñanza y la adopción de un modelo de aprendizaje más centrado en el estudiante. Ante nosotros hay unas herramientas impresionantes –Internet, correo electrónico, plataformas virtuales de aprendizaje, etc.– que permiten que los procesos de aprendizaje sean mucho más personalizados, mucho más flexibles. Al mismo tiempo, las nuevas tecnologías pretenden la capacitación del profesor universitario como usuario de recursos multimedia.

La introducción de las nuevas tecnologías no supone la desaparición del profesor, aunque obliga a redefinir sus funciones. En este contexto, el profesor ha de tender a reemplazar su

función de emisor de información por la función de tutor del proceso de aprendizaje, aspecto que se había ido perdiendo progresivamente en las aulas masificadas. Esto no hará más cómodo el trabajo de los profesores; más bien al contrario, ya que se les exigirá una mayor competencia pedagógica y un mayor grado de motivación.

Competencias que se desarrollan:
– Autonomía en el aprendizaje.
– Manejo de las nuevas tecnologías.
– Capacidad de trabajo.

En la enseñanza de la *economía* debe potenciarse el uso de las nuevas tecnologías de la información y la comunicación mediante el uso de plataformas virtuales de aprendizaje. Una plataforma como SWAD (Sistema Web de Apoyo a la Docencia; Figura 6), –o cualquier otra de las muchas que ya existen– permite, entre otras cosas:

- Una mayor interacción entre estudiantes y profesor mediante la realización de foros de debate, tutorías electrónicas, etc.
- Una más intensa comunicación entre estudiantes, mediante grupos de trabajo y de discusión.
- Apoyar a la docencia, permitiéndole al profesor "colgar" material didáctico de apoyo a la asignatura, tener la ficha con foto digital del alumno, saber cuándo y cuánto tiempo está un alumno en la plataforma, etc.
- Relaciones en tiempo real entre profesor y alumno, o entre alumnos, sin restricciones de distancia.
- Consulta de resultados (de pruebas objetivas, prácticas, ensayos, etc.) de forma individualizada.
- Trabajar en un entorno amigable, fácil de utilizar y bien estructurado.

Figura 6. La enseñanza de la economía en entornos virtuales: el caso de la plataforma SWAD <http://swad.ugr.es/>



3. EVALUACIÓN DEL APRENDIZAJE

La evaluación constituye una herramienta esencial para conseguir un aprendizaje efectivo; en función de cómo diseñemos la evaluación potenciaremos un tipo de aprendizaje u otro. La evaluación requiere el uso de instrumentos diversos, pues múltiples son los objetivos a cumplir. En la docencia de la *economía* proponemos cuatro instrumentos individuales y otro grupal para evaluar la asignatura.

3.1. Evaluación individual: potenciando el trabajo autónomo

3.1.1. Prueba objetiva

Esta modalidad de prueba consiste en formularle al alumno cuestiones (tipo test y/o preguntas de desarrollo y/o problemas) sobre algunos aspectos del programa de la asignatura. Los niveles que se evalúan son: información, comprensión, aplicación, capacidad de análisis y síntesis.

Competencias que se desarrollan al tener que preparar el alumno este tipo de prueba:
– Autonomía en el aprendizaje.
– Trabajar bajo presión.
– Gestión del tiempo.
– Capacidad de sacrificio.
– Habilidad para el aprendizaje.
– Planificación, coordinación, organización.

3.1.2. Ensayo

Es un instrumento indicado cuando se quiere evaluar la capacidad del alumno para organizar y estructurar información sobre un problema complejo y la búsqueda de soluciones. Concede más libertad al alumno para que exponga sus ideas con absoluta espontaneidad y creatividad. Ofrece un ambiente académico en el que se privilegia la labor crítica. Se incentiva el juicio crítico e independiente de los estudiantes.

Ejemplos de ensayos en *economía* pueden ser: a) «¿Qué política económica nos permitiría salir de la crisis?»; o b) «¿Deberían los estudiantes universitarios pagar el coste total de sus estudios?». Se plantea un problema en forma de pregunta o disyuntiva. Se espera que el estudiante asuma una postura, lo que implica desarrollar una línea de argumentación que respalde su propia postura frente al problema. En todo caso, el debido respaldo que requiere un ensayo académico está dado por pruebas que apoyen la posición de quien lo escribe, pruebas que pueden ser teóricas, empíricas o una combinación de ambas[4].

Competencias que se desarrollan:
– Capacidad de comunicación escrita.
– Organización.
– Pensamiento independiente.
– Superación.
– Gestión de la información.
– Ética.

3.1.3. Mapa conceptual

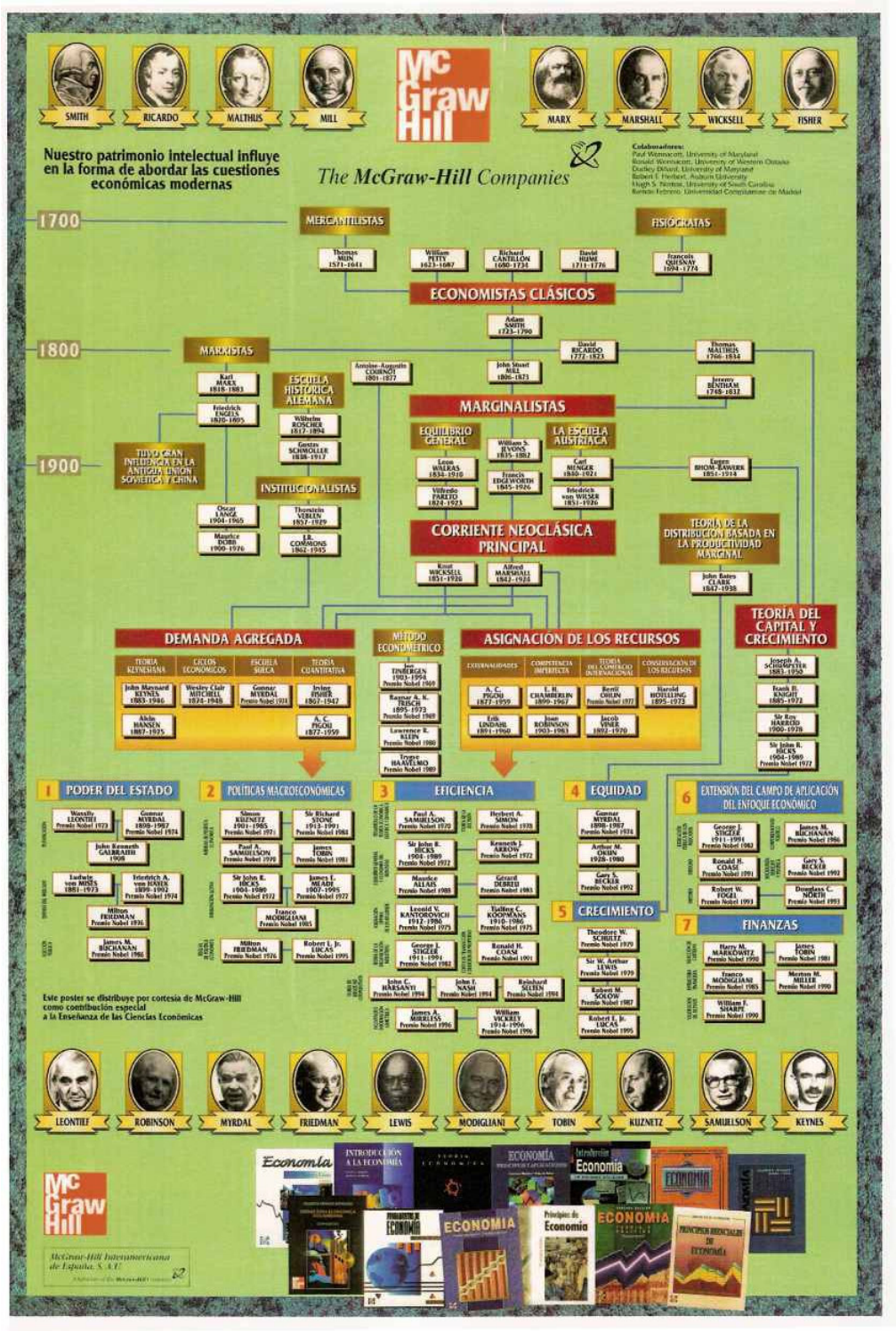
Es un procedimiento gráfico por medio del cual el alumno comunica sus conocimientos sobre conceptos (teorías, corrientes de pensamiento, etc.) y las relaciones entre éstos. En *economía* podemos evaluar mediante este instrumento temas introductorios relacionados con la

[4] A diferencia con las preguntas de desarrollo de un examen global (tradicional) de *economía*, donde el alumno que más recuerda es el que escribe más y mejor puntuación obtiene, el ensayo le permite al alumno pensar tranquilamente sobre un tema, redactando sus reflexiones.

historia del pensamiento económico (véase Figura 7), temas de macroeconomía relacionados con los modelos de crecimiento económico, etc.

Competencias que se desarrollan:
– Comprender conceptos.
– Desarrollar la capacidad de síntesis de la información.
– Ordenar ideas en esquemas lógicos.

Figura 7. Ejemplo de mapa conceptual: principales corrientes de pensamiento económico



3.1.4. Práctica con ordenador

Las sesiones prácticas realizadas en el aula de Informática deben evaluarse con alguna práctica que el alumno haga con el ordenador, usando los programas informáticos vistos en las diferentes sesiones. La evaluación puede ser presencial (en el aula de Informática) o no presencial –el alumno puede preparar la práctica de forma autónoma y entregársela posteriormente al profesor–. Los datos puede suministrarlos directamente el profesor, o bien es el estudiante quien debe descargarse esa información de Webs concretas (como INE o Banco de España), dependiendo del análisis que deba realizar.

Competencias que se desarrollan:
– Habilidades de búsqueda y análisis de información.
– Gestión del tiempo.
– Aplicación de las técnicas de análisis de la economía aplicada.
– Destrezas intelectuales relacionadas con la generación del conocimiento científico.
– Utilización de los diferentes recursos y aplicaciones tecnológicas.

3.2. Evaluación grupal: Potenciando el trabajo cooperativo

El aprendizaje cooperativo formal tiene, a diferencia del informal señalado anteriormente (apartado 2.3.), vocación de continuidad, en el sentido de formar grupos para una determinada tarea que puede durar desde varias semanas a todo el curso. En un grupo formal los estudiantes trabajan juntos para conseguir objetivos compartidos, intentando maximizar su aprendizaje y el de sus propios compañeros. El número de miembros que conforman un grupo formal es pequeño, 4 o 5 estudiantes, organizado por el profesor, y en ellos se espera que el estudiante interaccione con sus compañeros, compartiendo los conceptos y estrategias que aprendan, y que se consideren mutuamente responsables de la tarea asignada. Los estudiantes son responsables de su aprendizaje y del de sus compañeros de grupo, lo que provoca una corresponsabilidad para resolver las tareas académicas que les sean asignadas. Los objetivos de los participantes están vinculados, por lo que sólo pueden obtener el triunfo final si los demás miembros del grupo también lo consiguen.

En *economía*, la realización de trabajos en grupo (o en equipo) es también imprescindible. El trabajo en grupo, que se realiza principalmente de forma autónoma por sus componentes –aunque con la supervisión y asesoramiento del profesor–, persigue los siguientes objetivos:

- Iniciar al estudiante en la investigación aplicando las técnicas de análisis de la economía aplicada.
- Facilitar a los alumnos la construcción de sus aprendizajes.
- Proporcionar habilidades de búsqueda y análisis de información.
- Favorecer la capacidad crítica de los alumnos.

Ejemplos de trabajos podrían ser: a) «Situación actual del mercado de trabajo español»; o b) «El modelo de financiación de las comunidades autónomas españolas». Se evalúa la estructura del trabajo, la calidad de los razonamientos sobre los aspectos tratados, la riqueza de las fuentes utilizadas, la claridad en la redacción –lenguaje expositivo claro y científico– y la presentación de gráficos y tablas. Todos los trabajos hechos en grupo deben exponerse para toda la clase durante un tiempo limitado (unos 15 minutos).

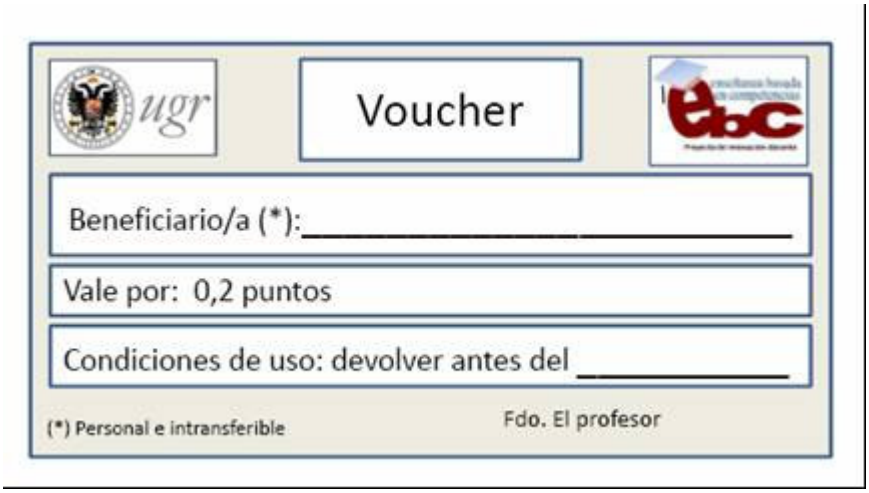
Competencias que se desarrollan:
– Capacidad coordinadora.
– Habilidades de búsqueda y análisis de información.
– Gestión del tiempo.
– Comunicación oral y escrita.
– Liderazgo.

4. AUMENTANDO LA EFICIENCIA EN LA ENSEÑANZA DE LA ECONOMÍA

El término eficiencia es uno de los conceptos más importantes en *economía*. Se utiliza principalmente a nivel microeconómico cuando analizamos la teoría de la producción de las

empresas. A nivel individual, una persona es eficiente si logra el mejor resultado posible con sus recursos y capacidades. Nuestro objetivo como docentes es que el estudiante trabaje a diario (o casi a diario) la asignatura. Por ello, además de la puntuación que puede obtener de la evaluación individual y grupal, puede obtener puntos adicionales que premien su esfuerzo extra. Por ejemplo, antes de empezar la clase magistral se pide que un alumno elegido al azar (o de forma voluntaria) haga un resumen de los contenidos que se explicaron en la sesión anterior. Se le entrega un vale (o *voucher*) para premiarle (véase Figura 8). Esto genera incentivos para que otros estén pendientes de repasarse la clase del día anterior por si el profesor pregunta. Se trata, en definitiva, de aumentar la eficiencia en la enseñanza de la asignatura.

Figura 8. Entrega de vales para aumentar la eficiencia de la asignatura



5. ADAPTACIÓN DE LOS DOCENTES AL ESPACIO EUROPEO DE EDUCACIÓN SUPERIOR

El reto que plantea el proceso de convergencia europea supone para el profesorado universitario español emprender una serie de actuaciones que implican asumir una dinámica de adaptación. La filosofía de Bolonia exige, entre otras tareas, que el profesor programe toda la asignatura, prepare con antelación materiales didácticos y tutorice a los estudiantes.

5.1. Planificación de actividades

El uso eficaz del tiempo de enseñanza exige que el profesor planifique en detalle cómo usar adecuadamente el tiempo que durará el curso y cada clase. De esta manera evitará uno de los errores más frecuentes observados en la actualidad en la universidad española: los profesores diseñan programas demasiado "ambiciosos" que son imposibles de desarrollar en el tiempo disponible. Cada contenido del programa y cada actividad de aprendizaje deben tener asignado su tiempo. Además, el profesor debe temporalizar cada parte y tarea programada, es decir, indicar qué día del calendario se va a acometer tal o cual actividad (elaboración del cronograma de la asignatura). Determinar qué hacer en el tiempo disponible y cuándo constituyen la esencia de la planificación[5].

5.2. Elaboración de material didáctico

Es importante precisar con antelación qué material (e instrumental) es el más adecuado para cada tema (y para apartados dentro del tema). El material cumple una función de apoyo a la

[5] En las actividades planificadas que tienen que realizar los alumnos, tanto presenciales como no presenciales, tendremos en cuenta los objetivos y contenidos que pretendemos desarrollar en cada tema, así como las características de los alumnos con los que trabajamos, el tamaño del grupo y la infraestructura y recursos de que disponemos.

actividad del alumno y le estimula a trabajar, pero hay que evitar que se convierta en protagonista principal del aprendizaje. El material debe ser variado: libros de prácticas, manuales para las clases magistrales, CDs con las diapositivas de las clases magistrales, material audiovisual (DVDs de *economía*), etc.

5.3. Tutorización (en el aula y virtual)

Para que el proceso de enseñanza-aprendizaje ocurra de manera efectiva es necesario el apoyo y orientación del profesor. El profesor debe desempeñar en todo momento el papel de tutor u orientador. La orientación puede dirigirse a todo el grupo en horas de clase, y estaríamos hablando de la tutoría presencial en el aula, o bien puede tener un carácter individualizado fuera de las horas de clase (en el despacho, por correo electrónico, etc.).

A diferencia de una clase magistral, en la tutoría presencial en el aula no se transmiten conocimientos teóricos a los alumnos sino que se les orienta sobre cómo usar un programa informático que será necesario para una sesión de prácticas en la sala de Informática, cómo hacer una búsqueda de bibliografía en Internet, cómo estructurar un trabajo de investigación, etc. La tutoría fuera de las horas de clase, por su parte, suele ser individualizada –o dirigida a un número reducido de alumnos; por ejemplo, a aquellos que están haciendo el trabajo en grupo–; el profesor debe adoptar un papel de facilitador, de apoyo, de seguimiento y de orientación, mientras que el alumno debe asumir siempre una responsabilidad dentro de la autonomía que genera este proceso de aprendizaje.

6. CONCLUSIONES

La *economía* proporciona a los estudiantes una base formativa que les permite construir conocimientos de carácter científico, apegados a la realidad, y que logra reestructurar sus esquemas cognoscitivos de forma significativa, incorporando toda la información que reciben diariamente sobre los problemas económicos, con los que la materia les va a proporcionar habilidades e instrumentos necesarios para emitir juicios críticos y constructivos de la realidad que les rodea a lo largo de toda su vida. No obstante, el modelo de enseñanza actual de la *economía* en las universidades españolas se basa fundamentalmente en el profesor –entendiendo al alumno esencialmente como receptor de la enseñanza–, fomenta la adquisición de conocimientos sobre todo a partir de la memorización, y plantea metodologías expositivas. Sin embargo, a partir de 2010, en el marco de un espacio común para la educación superior europea, pasaremos a un modelo basado en el aprendizaje.

Un modelo basado en el aprendizaje se basa en el alumno –el alumno es el responsable último de su aprendizaje–, favoreciendo su implicación, actividad y protagonismo. En este nuevo proceso de enseñanza-aprendizaje no sólo se definen los objetivos del aprendizaje, sino que se potencian las habilidades, actitudes y competencias necesarias para el futuro ejercicio profesional de los estudiantes. La labor del profesor no consiste en una mera transmisión del conocimiento, sino que está orientada a estimular el aprendizaje, ofrecer asesoramiento y tutorizar al alumnado. El trabajo dirigido o guiado se centra fundamentalmente en la adquisición de competencias. Mediante el trabajo dirigido, el profesor "monitoriza" el progreso del estudiante, ofreciéndole apoyo y *feedback* del trabajo realizado. El papel de las tutorías es clave en la actividad académica en el EEES.

La convergencia europea exige un replanteamiento de las técnicas de enseñanza y sistemas de evaluación usados hasta ahora en las clases de *economía*. Las lecciones magistrales deben combinarse con seminarios, sesiones de prácticas, estudio de casos, debates y enseñanza en el aula de Informática. No existe una estrategia docente ideal, sino que en cada momento habrá que hacer uso de los métodos de enseñanza que responden a las necesidades de los alumnos. Por eso, la metodología adoptada debe ser ecléctica, asumiendo diferentes métodos didácticos y variadas formas de organización y presentación de la materia. En relación con la evaluación, debe evaluarse no sólo el nivel de conocimientos, sino también la adquisición de habilidades, destrezas y actitudes de los alumnos. Los instrumentos de evaluación deben también diversificarse, combinando el examen tradicional con mapas conceptuales, ensayos, prácticas con ordenador y trabajos hechos en grupo.

REFERENCIAS

BENITO, A., y CRUZ, A. (coords.) (2005): *Nuevas claves para la docencia universitaria*, Narcea, Madrid.

BENITO, A., y CRUZ, A. (2006): «El nuevo papel del alumno bajo la perspectiva de Bolonia», *Revista Crítica*, núm. 934, pp. 32–35.

GÓMEZ GARCÍA, F., BORRA MARCOS, C., GIL GALVÁN, R., y SALAS VELASCO, M. (2008): *La inserción laboral de los economistas*, editorial Thomson, Madrid.

SALAS VELASCO, M. (2000): «Universidad y empleo: El proceso de inserción laboral de los graduados». En Salmerón, H., y López, V.L. (coords.): *Orientación educativa en las universidades*, Grupo Editorial Universitario, Granada, pp. 435–447.

SALAS VELASCO, M. (2007): *La elaboración del proyecto docente*, editorial UOC, Barcelona.

SAMUELSON, P.A., Y NORDHAUS, W.D. (2006): *Economía (decimoctava edición)*, McGraw-Hill, Madrid.

ÁREA XV

ECONOMÍA EMPRESARIAL

ECONOMIA EMPRESARIAL

HACIA UNA ORDENACIÓN DE LAS PEQUEÑAS EMPRESAS ATENDIENDO A SU POSIBLE SITUACIÓN DE FRACASO

ERNESTO J. VERES FERRER
E-mail: ernesto.veres@uv.es
Facultad de Economía
Dpto. Economía Aplicada
Universidad de Valencia (España)

Resumen

En la literatura contable son muchos los trabajos que han intentado explicar el posible fracaso empresarial a partir de la información contable y poner de manifiesto la relación entre dicha información y la situación de fracaso. Sin embargo, los resultados obtenidos hasta ahora no son definitivos, en cuanto no existe un mínimo común en todos ellos que permita modelizar el comportamiento de dicho fracaso.
Este trabajo pretende constatar, mediante la definición de un indicador en el que intervengan una gran batería de ratios contables, que efectivamente la información contable ofrece pistas sobre el posible fracaso de la empresa. Al canalizar toda la información de las ratios contables a través de un indicador global y único, puede clasificarse las empresas atendiendo a sus mayores o menores probabilidades de fracaso. La constatación de la no aleatoriedad en la clasificación resultante para una muestra de empresas confirmará la existencia de esa relación buscada.

Palabras Clave: Fracaso empresarial - información contable - ratio contable - PYME
Área Temática: Economía Empresarial.

Abstract

In the countable literature there are great the works that have tried to explain the possible managerial failure from the countable information and to reveal the relation between the above mentioned information and the situation of failure. Nevertheless, the results obtained till now are not definitive, in all that a common minimum does not exist in all of them that there allows modeling the behavior of the above mentioned failure.
This work tries to state, by means of the definition of an indicator with a great battery of countable ratios, that certainly the countable information offers tracks on the possible managerial failure. On having canalized all the information of the countable ratios across the global and unique indicator, one can classify the companies attending to her major or minor probabilities of failure. The verification of not randomness in the resultant classification for a sample of companies will confirm the existence of this looked relation.

Key words: Managerial failure - accounting information - accounting ratio - small bussines
Thematic Area: Entrepreneurial Economics.

1. INTRODUCCION

En la línea de investigación iniciada en la segunda mitad de los años sesenta por Beaver (1966) y Altman (1968), son muchos los trabajos que han intentado explicar el posible fracaso empresarial a partir de la información contable. Prácticamente en todos ellos, se pone de manifiesto la existencia de algún tipo de relación entre dicha información y la situación de fracaso. Sin embargo, los resultados obtenidos son parciales, no existe un mínimo común en todos ellos que permita modelizar el comportamiento del fracaso empresarial. De hecho, las ratios que anuncian o confirman la delicada situación de una empresa son distintas según cada uno de los autores y cada una de las aplicaciones. Dependiendo de la técnica estadística

utilizada, los resultados varían e, incluso, dependiendo del ámbito territorial y la muestra de empresas utilizada, también son diferentes las conclusiones finales y las descripciones obtenidas. La inexistencia de una teoría económica que la sustente ha sido determinante para que este planteamiento siga siendo un problema abierto, sin solución hasta ahora definitiva.

Al respecto y en relación a nuestro país, destacan los trabajos pioneros de Laffarga, Martín y Vázquez (1985), Pina (1989) y Rodríguez (1989) en el sector bancario y los de Rodríguez (1990) y Mora (1996) en el sector asegurador. Por otra parte, en las dos últimas décadas se han realizado interesantes contribuciones que abordan la realidad de las pymes en distintos ámbitos temporales, geográficos y sectoriales, como los trabajos realizados por Lizarraga (1997), López, Gandía y Molina (1998), Rodríguez López (2000, 2004), Somoza López (2002), De la Torre y Gómez (2005) y Mínguez (2006).

Este trabajo pretende dar respuesta a la pregunta inicial que ha motivado gran parte de los trabajos citados: ¿existe relación entre la información contable, canalizada a través de una gran batería de ratios, y el fracaso empresarial? No pretendemos diseñar un modelo de fracaso en el que sólo intervenga una selección de ratios, sino sólo constatar, mediante la definición de un indicador en el que intervengan todas ellas, que efectivamente la información contable ofrece pistas sobre el posible fracaso de la empresa.

La forma de abordar esta pregunta consiste en formular un indicador de fracaso, definido a partir de todas las ratios disponibles. Esto es, canalizar toda la información de las ratios contables a través de un indicador global y único, que permita una clasificación de las empresas atendiendo a sus mayores o menores probabilidades de fracaso. La constatación de la no aleatoriedad en la clasificación resultante para una muestra de empresas confirmará la existencia de esa relación buscada.

2. DEFINICION DEL INDICE DE CLASIFICACION "IC"

Partimos de 73 ratios R_i y de una doble muestra de empresas: las fracasadas Q_k (un censo en nuestra aplicación) y las sanas S_k. En este trabajo se asimila la condición de fracaso a las situaciones de quiebra o suspensión de pagos. El hecho de que a efectos de la definición se equipare la suspensión de pagos con la quiebra se justifica en que, a pesar de ser dos figuras concursales distintas, habiendo estado prevista la primera para superar estados transitorios de dificultad financiera y reservando la quiebra para situaciones de gravedad estructural, lo cierto es que en la práctica en nuestro país, la suspensión de pagos ha sido utilizada por el deudor como un mecanismo de defensa, en situaciones de deterioro patrimonial prácticamente irreversible, para impedir la solicitud de quiebra por parte de algún acreedor y evitar así posibles responsabilidades penales. Es decir, tras la mayoría de solicitudes de suspensión de pagos se esconden verdaderas situaciones de quiebra.

Una amplia discusión sobre el alcance y limitaciones de esta interpretación para el fracaso empresarial puede verse en Pozuelo (2007), y en Labatut, Pozuelo y Veres (2008a), en donde también pueden encontrarse la relación y definición de las 73 ratios contables utilizadas y que se recogen en el Anexo. Por otra parte, en Veres (2007) y en Labatut, Pozuelo y Veres (2008b) aparecen sendos intentos por formalizar la relación entre la información contable y el fracaso empresarial.

Calculamos para cada ratio y tipo de empresa Q ó S, las correspondientes medias y desviaciones standard: $\overline{R}_{iQ}$, $\overline{R}_{iS}$, σ_{iQ} y σ_{iS}. Denotamos por N_i el número total de empresas (activas y fracasadas) que proporcionan información sobre la ratio i-ésima.

La información inicial sobre las empresas sólo las clasifica como sanas (con actividad) o fracasadas (situaciones de quiebra o suspensión de pagos). Pero la realidad demuestra que existen empresas sanas con algunas ratios cuyo comportamiento es más propio de las fracasadas y viceversa. Esto demuestra que la realidad empresarial, necesariamente compleja, no presenta límites nítidos en relación a las situaciones de fracaso, y que muchas veces es la misma actitud del empresario –más o menos cansado o con mayores o menores expectativas de futuro-, la que determina su decisión sobre continuar o no con la actividad empresarial. De ahí la necesidad de

introducir una nueva categoría en las empresas y efectuar, con la información contable disponible, una nueva clasificación de las mismas en sanas, fracasadas y dudosas (D_k), contemplando de esta forma la existencia, para cada ratio, de una frontera difusa que no la identifique plenamente. El criterio es el básico en Estadística: todo lo que se "aleje" de la media es significativo de un comportamiento lejano al descrito por la variable –en nuestro caso, la ratio contable- estudiada.

Para clasificar una empresa fracasada como dudosa, según la ratio R_i, se considera el sencillo criterio siguiente:

$$\begin{cases} Si\ \overline{R}_{iQ} < \overline{R}_{iS} \rightarrow si\ R_{iQ} > \overline{R}_{iQ} + \sigma_{iQ}\ \ la\ empresa\ es\ dudosa \\ Si\ \overline{R}_{iQ} > \overline{R}_{iS} \rightarrow si\ R_{iQ} > \overline{R}_{iQ} - \sigma_{iQ}\ \ la\ empresa\ es\ dudosa \end{cases}$$

Análogamente para una empresa sana como dudosa:

$$\begin{cases} Si\ \overline{R}_{iS} < \overline{R}_{iQ} \rightarrow si\ R_{iS} > \overline{R}_{iS} + \sigma_{iS}\ \ la\ empresa\ es\ dudosa \\ Si\ \overline{R}_{iS} > \overline{R}_{iQ} \rightarrow si\ R_{iS} > \overline{R}_{iS} - \sigma_{iS}\ \ la\ empresa\ es\ dudosa \end{cases}$$

A cada empresa E_k se le asocia un vector de clasificación de i componentes, que recoge la situación de dicha empresa expresada en función de cada ratio:

$$V_k = \left(a_{k1} = S_1 \acute{o} D_1 \acute{o} Q_1; a_{k2} = S_2 \acute{o} D_2 \acute{o} Q_2; ...; a_{ki} = S_i \acute{o} D_i \acute{o} Q_i; ...; a_{k73} = S_{73} \acute{o} D_{73} \acute{o} Q_{73}\right)$$

Siempre que exista inicialmente información, se realiza, para cada ratio R_i, sendos análisis discriminantes utilizando como variable de clasificación la variable que define la situación de sana, dudosa o fracasada de la empresa. Del resultado del análisis obtenemos sendas funciones de clasificación de Fisher (una para cada posible situación de la empresa, *S* ó *D* ó *Q*) y una tabla que recoge el resultado de la clasificación:

Ratio R_i	Situación		
	sana	dudosa	fracasada
Coeficiente	c_{is}	c_{id}	c_{iq}
(Constante)	k_{is}	k_{id}	k_{iq}

Funciones discriminantes lineales de Fisher

Ratio R_i		Grupo de pertenencia pronosticado		
		sana	dudosa	fracasada
Grupo de pertenencia Original	SANA	S_{is}	S_{id}	S_{iq}
	DUDOSA	D_{is}	D_{id}	D_{iq}
	FRACASADA	Q_{is}	Q_{id}	Q_{iq}

Resultado de la clasificación por aplicación del análisis discriminante

La primera tabla permitirá clasificar una empresa, cuya situación respecto a un posible fracaso es desconocida, en activa, dudosa o fracasada. La segunda tabla permitirá definir las probabilidades a posteriori de activa, dudosa o fracasada para una empresa determinada, una vez conocido el resultado de aplicar sobre ella la función discriminante.

Sea un conjunto de empresas de las que desconocemos su situación como sana o fracasada. Dada una empresa E_k la aplicación de las 73 funciones discriminantes a su información contable da lugar a un vector de clasificación pronosticado V_k:

$$V_k = \left(a_{k1}; a_{k2}; ...; a_{ki}; ...; a_{k73}\right)$$

Para cada empresa E_k y vector correspondiente V_k se define el siguiente índice de clasificación:

$$IC_k = CTE_k \times \frac{\sum_{i=1}^{73} P(Q/a_{ki}) \times \frac{S_{is} + Q_{iq}}{N_i}}{\sum_{i=1}^{73} P(S/a_{ki}) \times \frac{S_{is} + Q_{iq}}{N_i}} = CTE_k \times \frac{\sum_{i=1}^{73} \frac{Q_{a_{ki}}}{S_{a_{ki}} + D_{a_{ki}} + Q_{a_{ki}}} \times \frac{S_{is} + Q_{iq}}{N_i}}{\sum_{i=1}^{73} \frac{S_{a_{ki}}}{S_{a_{ki}} + D_{a_{ki}} + Q_{a_{ki}}} \times \frac{S_{is} + Q_{iq}}{N_i}} = CTE_k \times [1]$$

donde

$$0 \le CTE_k = \frac{\sum_{i=1}^{73} \frac{S_{is} + Q_{iq}}{N_i}}{73 - n^{\circ}\, ratios\ con\ \text{inf}ormación} \le 1$$

Observemos que el *IC* así definido es una media agregativa ponderada definida sobre las probabilidades a posteriori de que la empresa sea fracasada (numerador) o sana (denominador), conocida la clasificación de la misma a través de las funciones discriminantes. La ponderación para cada ratio es el porcentaje de aciertos conjuntos en la clasificación como sana y fracasada deducida del análisis discriminante.

Existen empresas en las que la clasificación correcta como sana o fracasada deducida de los respectivos análisis discriminantes es mayor que otras. Pero también influye el número de ratios que facilitan información de cada empresa. Su influencia sobre el *IC* anterior debe ser superior en estos casos, pues no es lo mismo un cociente de probabilidades [1] alto cuando el porcentaje de aciertos en la clasificación deducido del análisis discriminante es pequeño que cuando es alto, o cuando el número de ratios sobre las que se dispone de información es superior a los disponibles en otra empresa. Por ello la media agregativa ponderada está multiplicada por el coeficiente CTE_k definido como la media de los porcentajes de acierto en las clasificaciones discriminantes. De esta forma se obtiene, finalmente, una medida de la mayor probabilidad de fracaso de una empresa respecto su posible situación de activa, una vez aplicada a su información contable las funciones discriminantes deducidas de la muestra original de empresas.

La ordenación de las empresas según su IC_k da lugar a un posible ranking de fracaso empresarial. Teniendo en cuenta esa ordenación, si existiera relación entre la información de las ratios contables y la situación de la empresa, deberán situarse en los primeros lugares de la ordenación aquellas empresas con mayor probabilidad de estar fracasadas, pues en el numerador que define el índice de clasificación aparece la probabilidad de estar fracasada. Esta ordenación, pues, da pie a que la entidad auditora, de control o crediticia observe con mayor cuidado las primeras empresas en la clasificación obtenida por el índice.

Nos planteamos, pues, hasta qué punto hay relación entre la información suministrada por las ratios contables, canalizada a través del índice de clasificación anterior, y la situación de fracaso de una empresa. La aplicación del apartado siguiente confirma esta apreciación.

3. APLICACIÓN

Las características de la muestra inicial de empresas utilizada en la definición del *IC* se recogen en la Tabla 1.

Tabla 1. Muestra de empresas para la definición del índice de clasificación

Tipo de empresa	Fracasadas	Sanas
Referencia temporal	Quebradas o en suspensión de pagos en 2003-2004	Activas en 2003-2004
Tamaño	191	500
Tamaño mínimo	120 para ROT8	201 para ROT8
Error muestral máximo (confianza 95,5 %)	0 (Censo)	7,05 % para ROT8
Tamaño máximo	190 para EF2	499 para EF7 y 10; SOLV5
Error muestral mínimo (confianza 95,5 %)	0 (Censo)	4,47% para EF7 y 10; SOLV5

La información contable no siempre es de gran calidad. Incluso, en ocasiones presenta problemas de homologación de unidades. De ahí que existan valores, para algunas ratios o empresas, que sospechosamente permanecen fuera de rango. Hacemos una selección de valores de R_i fuera de rango. Para cada ratio y para cada tipo de empresa (Q ó S) se excluyen aquellos valores que están fuera del intervalo $\overline{R_i} + 1{,}96 \times \sigma_i$. Suponiendo normalidad –no siempre satisfecha para todas las ratios-, la condición anterior implicaría excluir el 5% de los valores extremos de cada ratio, a los que se les considera como ausencia de información.

Para confirmar la bondad del *IC* trabajamos con una muestra de validación, sobre la que también se suprimen los valores fuera de rango con los mismos criterios anteriores. Las características de la muestra de validación son las de la Tabla 2:

Tabla 2. Muestra de validación

Tipo de empresa	Fracasadas	Sanas
Referencia temporal	Quebradas o en suspensión de pagos en 2005-2006	Activas en 2005-2006
Tamaño	76	200
Tamaño mínimo	34 para REN17	133 para ROT8
Error muestral máximo (confianza 95,5 %)	0 (Censo)	8,67 % para ROT8
Tamaño máximo	76 para REN1, 2, 3, 5,8, 10 11, 14, 15, 18 y 21; EF1, 2, 3, 6, 7, 10 y 11; ROT1, 4 y 7; SOL5; SOLVLP1; CF2, 4, 5, 6 y 7	200 para REN8, 10 y 18; EF1, 2, 3 y 11; SOLVLP1
Error muestral mínimo (confianza 95,5 %)	0 (Censo)	7,07 % para 200 para REN8, 10 y 18; EF1, 2, 3 y 11; SOLVLP1

Ambas muestras se obtienen de la última actualización –junio de 2007- de la base de datos financieros SABI (*Sistema de Análisis de Balances Ibéricos*). Sobre ambas muestras de empresas se aplica el *IC* definido en el apartado anterior obteniéndose una clasificación de las mismas atendiendo a su mayor o menor probabilidad de fracaso.

Conseguida la ordenación de empresas según el *IC* para ambas muestras de empresas, los resultados de la clasificación se recogen en la Tabla 3 siguiente, en donde aparecen los porcentajes de ubicación de las empresas fracasadas dentro de la ordenación:

Tabla 3. Ubicación de las empresas fracasadas en la ordenación según el *IC*

Cuartil	Muestra de validación	Muestra que define el *IC*
Antes del cuartil primero	Están el 52,63% de las fracasadas	Están el 62,30% de las fracasadas
Antes del número de empresas fracasadas de la muestra	Están el 56,58% de las fracasadas	Están el 65,97% de las fracasadas
Antes de la mediana	Están el 76,32% de las fracasadas	Están el 80,10% de las fracasadas
Antes del cuartil tercero	Están el 88,16% de las fracasadas	Están el 93,72% de las fracasadas

Se confirma que la mayor cantidad de empresas fracasadas se ubican en los primeros lugares de la ordenación. Este porcentaje es mayor para la muestra que define el *IC*, resultado totalmente esperado pues la definición del índice se ha realizado a partir de las empresas de esa muestra. Pero lo destacable es el buen porcentaje de empresas fracasadas ubicadas en los primeros lugares de la ordenación en el caso de la muestra de validación, si tenemos en cuenta que son empresas cuya información es posterior uno o dos años a la utilizada en la muestra de definición. Se confirma, pues, el carácter predictivo del *IC*.

Otras medidas estadísticas confirman las apreciaciones anteriores. En efecto, la Tabla 4 recoge la distribución de *unos* y *ceros* de ambas muestras, una vez se agrupan las empresas por cuartiles y se ordenan según su *IC*:

Tabla 4. Prueba χ^2 para tablas de contingencia

	Muestra de validación		Muestra que define el *IC*	
	0	1	0	1
Hasta 1º cuartil	29	40	54	119
Del 1º al 2º cuartil	51	18	139	34
Del 2º al 3º cuartil	60	9	146	26
Desde 3º cuartil	60	9	161	12
Total	200	76	500	191

Para la muestra de validación, el estadístico $\chi_3^2 = 46{,}629$ rechaza la hipótesis de independencia entre la ordenación por cuartiles y la situación de posible fracaso de la empresa para una significación del 0,000. Análogamente para la muestra que define el *IC*. En este caso, el valor del estadístico es $\chi_3^2 = 202{,}520$, nuevamente con significatividad nula del 0,000. Los respectivos *coeficientes de contingencia* para ambas muestras son 0,380 y 0,476, y los *coeficientes de contingencia cuadrática media o fi-cuadrado* y la *V de Cramer*, que coinciden en este caso, son 0,411 y 0,541 respectivamente. Todos ellos confirman la existencia de asociación.

Un resultado paralelo confirma la relación entre la ordenación obtenida por el *IC* y la situación de fracaso empresarial. Para ambas muestras de empresas –tanto para la muestra de empresas que da lugar a la definición del *IC*, como para las de la muestra de validación- es conocida la situación real de las mismas, por lo que la clasificación a partir del *IC* da lugar a una racha de *0* (situación *S*) y *1* (situación *Q*). Sendos tests de rachas aplicadas sobre ellas demuestra la significatividad nula de las mismas, confirmando la existencia de relación entre la clasificación por el *IC* y la situación de fracaso de una empresa.

Los resultados de los correspondientes tests de rachas son los de la Tabla 5:

Tabla 5. Prueba de rachas

	Muestra de validación	Muestra que define el *IC*
Valor de prueba[(a)]	1,000	1,000
Casos en total	276,000	691,000
Número de rachas	91,000	191,000
Z	-3,047	-8,227
Significatividad asintótica bilateral	0,002	0,000

[(a)] Especificado por el usuario

La significatividad nula del test de rachas para ambas muestras confirma la existencia de relación entre la clasificación por el *IC* y la situación de fracaso de una empresa. Como éstas se ubican con preferencia en los primeros lugares, la ordenación puede utilizarse para establecer más controles sobre éstas que sobre las que ocupan los últimos lugares. En efecto, para la muestra de validación la secuencia de rachas obtenida es la siguiente:

1	0	0	1	1	1	1	0	0	1	0	0	1	1	0	1	1	1	1	1
0	1	1	1	1	0	1	1	1	1	0	0	0	1	1	1	0	1	0	1
1	0	0	1	0	0	0	1	1	1	1	0	1	1	1	1	1	0	0	0
0	1	1	0	0	0	1	0	0	1	0	0	0	1	0	1	0	0	0	1
0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	1	1
0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0
0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	1	0
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	1				

Observando la secuencia deducimos que el mayor peso de *"unos"* (empresas fracasadas) se sitúa en su parte primera. No obstante, el hecho de que en la segunda parte de la misma sigan apareciendo empresas fracasadas –de hecho, la última empresa de la secuencia lo es-, o que en la primera parte existan empresas con actividad –la segunda y tercera empresas en la ordenación son activas-, da pie a pensar que los datos contables, que según se ha visto proporcionan información sobre el fracaso empresarial, son de compleja modelización para el objetivo propuesto. Es decir, en la decisión sobre cuándo finalizar la actividad de una empresa, por mala situación económica o por falta de expectativas de futuro, no sólo influyen aspectos meramente técnicos o de gestión, sino que hay una clara componente subjetiva del empresario que le lleva a adoptar una u otra decisión. Esta componente, cuya existencia es real como se aprecia en la secuencia de la tabla anterior, es de imposible cuantificación y, de resultas, de difícil concreción en una posible modelización. Así pues, la psicología empresarial se confirma como importante factor descriptivo del fracaso empresarial.

4. CONCLUSIONES

El índice de clasificación para las empresas se ha construido a partir de las probabilidades de clasificación como empresa sana o fracasada, condicionadas al conocimiento de su situación según funciones discriminantes deducidas para cada ratio contable. La importancia de cada ratio es respetada a través del porcentaje de aciertos conjuntos en la clasificación como sana y fracasada deducida del análisis discriminante. También se tiene en cuenta el número de ratios que facilitan información de cada empresa. El *IC* definido proporciona una medida de la posible probabilidad de fracaso para una empresa de forma que la ordenación de las empresas según su IC_k da lugar a un ranking de fracaso empresarial.

La aplicación confirma la existencia de relación estadística entre la información contable, canalizada por el índice de clasificación, y la situación de fracaso empresarial. El correspondiente test de rachas aplicado a la ordenación de las empresas según su *IC* determina la no aleatoriedad de la clasificación, ocupando las empresas fracasadas los primeros lugares de la ordenación.

Sin embargo, el hecho de que en ciertas posiciones primeras existan empresas sanas, coexistiendo con empresas fracasadas, mientras que en las últimas posiciones de la ordenación existan empresas fracasadas, coexistiendo con empresas sanas, sugiere que la modelización del fracaso empresarial es, necesariamente, compleja. En efecto, en la decisión de finalizar la actividad de una empresa, además de los aspectos técnicos o de gestión, influyen criterios subjetivos del empresario de muy difícil integración en un modelo de fracaso.

El anterior resultado apunta a una limitación de la modelización del fracaso empresarial a través de la información contable. En efecto, confirmándose la existencia de relación entre

información contable y fracaso empresarial, su concreción en un criterio de clasificación de empresas puede fallar cuando se aplica a una empresa concreta, restando valor a la aplicabilidad de un posible modelo *fracaso empresarial/información contable* cuando su uso se restringe a la búsqueda de decisiones concretas.

REFERENCIAS

ALTMAN, E.I. (1968): "Financial ratios, Discriminant Análisis and the Prediction of Corporate Bankruptcy". Journal of Finance, Vol. 23, nº 4, September. Págs. 589-609.

BEAVER, W.H. (1966): "Financial Ratios as Predictors of Failure". Journal of Accounting Research, Vol. 4, Empirical Research in Accounting: Selected Studies 1966. Págs. 71-111.

DE LA TORRE, J.M. y GÓMEZ, M. E. (2005): "Análisis de la sensibilidad temporal en los modelos de predicción de insolvencia: Una aplicación a las pymes industriales". XIII Congreso de A.E.C.A. (Armonización y Gobierno de la Diversidad).Comunicación. Universidad de Oviedo, Septiembre.

LABATUT, G., POZUELO, J. Y VERES, E. (2008a): "La insolvencia empresarial en las microempresas valencianas. Un estudio mediante metodología univariante". Partida Doble, nº 201, julio-agosto 2008. Págs. 68-77.

LABATUT, G., POZUELO, J. Y VERES, E. (2008b): "Las ratios contables como sistemas de control del fracaso empresarial en la PYME española". Revista Española de Financiación y Contabilidad. Artículo en segunda evaluación.

LAFFARGA, J., MARTÍN, J. Y VÁZQUEZ, J. (1985): "El Análisis de la Solvencia en las Instituciones Bancarias: Propuesta de una Metodología y Aplicaciones a la Banca Española", Esic Market, nº 48, Abril-Junio. Págs 51-83.

LIZARRAGA, F. (1997): "Utilidad de la información contable en el proceso de fracaso: Análisis del sector industrial de la mediana empresa española". Revista Española de Financiación y Contabilidad, Vol. 26, nº 93, Octubre-Diciembre. Págs 871-915.

LÓPEZ, J., GANDÍA, J.L., MOLINA, R. (1998): "La suspensión de pagos en las pymes: una aproximación empírica". Revista Española de Financiación y Contabilidad. Vol. XXVII, nº 94, Enero-Marzo. Págs 71-97.

MÍNGUEZ, J.L. (2006): "Factores explicativos de la insolvencia empresarial: una aplicación a la pequeña y mediana empresa constructora". VI Jornadas sobre Predicción de Insolvencia Empresarial. Comunicación. Carmona (Sevilla), Noviembre.

MORA, A. (1996): El contenido informativo de los datos contables para la toma de decisiones de inversión. Ed. ICAC. Madrid.

PINA, V. (1989). "La información contable en la predicción de la crisis bancaria". Revista Española de Financiación y Contabilidad, Vol. 18, nº 58. Págs 309-338.

POZUELO, J. (2007). La Predicción del Fracaso Empresarial: Un Estudio Empírico de la Pequeña Empresa de la Comunidad Valenciana. Tesis Doctoral, Mayo. Universidad de Valencia.

RODRÍGUEZ, J.M. (1989): "Análisis de las insolvencias bancarias en España: Un modelo empírico". Moneda y Crédito, nº189. Págs 187-227.

RODRÍGUEZ, M.C. (1990): La predicción de las crisis empresariales: Modelos para el sector de seguros. Departamento de publicaciones. Universidad de Valladolid.

RODRÍGUEZ, M. (2000): Métodos y Modelos de Pronostico del Fracaso Empresarial. Una Aproximación Empírica a la Realidad Empresarial de la Comunidad Autónoma de Galicia. Tesis Doctoral. Universidad de La Coruña.

RODRÍGUEZ, M. (2004): "Análisis de Sensibilidad del Modelo de Redes Neuronales Artificiales para la Predicción de la Insolvencia Empresarial. Una métrica para superar la solución de "Caja Negra" V Jornadas sobre la Predicción de la Insolvencia Empresarial: la gestión del riesgo financiero y la Nueva Ley Concursal. Comunicación. Noviembre Universidad de Oviedo-AECA.

SOMOZA, A. (2001): "La consideración de factores cualitativos, macroeconómicos y sectoriales en los modelos de predicción de la insolvencia empresarial. Su aplicación al sector textil y confección de Barcelona (1994-1997)". Papeles de Economía Española, nº 89/90. Págs 402-413.

VERES, E. (2007): "Clasificación de empresas según su solvencia a partir de información contable". XXI Reunión Anual de ASEPELT-España, Universidad de Valladolid. Comunicación. Junio Actas. Pág. 362 y CD.

ANEXO

Se consideran 73 ratios R_i agrupadas en siete categorías según la finalidad contable de las mismas: rentabilidad, estructura financiera, actividad, rotación, liquidez, solvencia a largo plazo y estructura económica. En su conjunto describen suficientemente la situación financiera y contable de la empresa. Reúnen tres condiciones esenciales: *identidad de componentes*, que permite que todo usuario del ámbito empresarial conozca el significado de cada uno de los componentes de los ratios; *existencias de normas contables* para homogeneizar el contenido; y *comparabilidad* para poder determinar la evolución de los valores a lo largo de los años estudiados. La tabla siguiente recoge su definición.

Ratios contables

RATIO	DEFINICION
Rentabilidad	
REN1	Resultado antes de intereses e impuestos / Activo total
REN2	Resultado del ejercicio / Activo total
REN3	Resultado del ejercicio / Fondos propios
REN4	Resultado del ejercicio / Pasivo total
REN5	Resultado de actividades ordinarias / Activo total
REN6	Resultado antes de intereses e impuestos / Ventas
REN7	Resultado de actividades ordinarias / Pasivo total
REN8	Resultado de actividades ordinarias / Fondos propios
REN9	Resultado del ejercicio / Ventas
REN10	Cash Flow recursos generados / Fondos propios
REN11	Cash Flow recursos generados / Activo total
REN12	Cash Flow recursos generados / Pasivo total
REN13	Cash Flow recursos generados / Pasivo circulante
REN14	Resultado antes de impuestos / Activo total
REN15	Resultado neto – Realizable – Existencias / Activo total
REN16	Cash Flow recursos generados / Ventas
REN17	Resultado antes de intereses e impuestos / Gastos financieros
REN18	Resultado antes de impuestos / Fondos propios
REN19	Resultado antes de impuestos / Ventas
REN20	Resultado antes de impuestos / Pasivo total
REN21	Resultado antes de int. e imp. / Gastos financieros + Pasivo circulante
REN22	Pasivo fijo / Cash Flow recursos generados
REN23	Resultado de actividades ordinarias / Ventas
Estructura financiera	
EF1	Pasivo total / Fondos propios
EF2	Pasivo fijo / Fondos propios
EF3	Pasivo circulante / Fondos propios
EF4	Gastos financieros / Pasivo total
EF5	Pasivo circulante / Pasivo total
EF6	Pasivo fijo / Activo total
EF7	Pasivo circulante / Activo total
EF8	Fondos propios / Pasivo total
EF9	Gastos financieros / Ventas
EF10	Pasivo total / Activo total
EF11	Pasivo fijo / Capital social
Actividad	
ACT1	Valor añadido / Ventas
ACT2	Gastos de personal / Valor añadido
ACT3	Gastos financieros / Valor añadido
ACT4	Ingresos de explotación / consumos de explotación
ACT5	Resultado del ejercicio / Valor añadido

ACT6	Pasivo fijo / Ventas
ACT7	Gastos de personal / Ventas
ACT8	Gastos de personal / Activo fijo
Rotación	
ROT1	Ventas / Activo total
ROT2	Ventas / Activo circulante
ROT3	Ventas / Activo fijo
ROT4	Ventas / Fondos propios + Pasivo fijo
ROT5	Ventas / Pasivo circulante
ROT6	Activo total / Ingresos de explotación
ROT7	Ventas / Capital circulante
ROT8	Ventas / Existencias
ROT9	Ventas / Realizable
Solvencia (liquidez)	
SOLV1	Activo circulante / Pasivo circulante
SOLV2	Activo circulante – Existencias / Pasivo circulante
SOLV3	Disponible / Pasivo circulante
SOLV4	Capital circulante / Pasivo circulante
SOLV5	Capital circulante / Activo total
SOLV6	Capital circulante / Ventas
SOLV7	Realizable / Capital circulante
SOLV8	Activo circulante – Existencias / Ventas
SOLV9	Recursos generados antes de impuestos / Pasivo circulante
SOLV10	(Activo circulante – Existencias) – Pasivo circulante / Gastos de explotación – Amortizaciones – Provisiones (Intervalo sin crédito)
SOLV11	Capital circulante / Ingresos de explotación
SOLV12	Realizable / Ingresos de explotación
Solvencia a largo plazo	
SOLVLP1	Activo fijo / Fondos propios
SOLVLP2	Recursos generados antes de impuestos / Pasivo total
SOLVLP3	Activo total neto / Pasivo total
Estructura económica	
CF1	Existencias / Activo circulante
CF2	Activo circulante / Activo total
CF3	Activo circulante / Activo fijo
CF4	Existencias / Capital circulante
CF5	Existencias / Activo total
CF6	Realizable / Activo total
CF7	Disponible / Activo total

DETERMINANTS OF ENTREPRENEURSHIP: ARE WOMEN DIFFERENT?

FERNANDA LLUSSÁ[1]
E-mail: fajl@fct.unl.pt
Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa

Abstract

In this paper we investigate, for the first time, how individual determinants of entrepreneurship - such as age, income, education, work status, skills, access to networks and fear of failure - differ between males and females. We conduct our exercise using individual data provided by the Global Entrepreneurship Monitor (GEM), available for 46 countries, between 2001 and 2004. The literature on entrepreneurship has uncovered differences in the rate of entrepreneurship between men and women, with women generally displaying lower entrepreneurial activity than men. This is important since, as we show, entrepreneurial activity is positively related across countries with the female to male entrepreneurial ratio. We examine total entrepreneurship rates, as well as entrepreneurship driven by opportunity and by need. We find that indeed entrepreneurial activity rates are lower for females across all but one of the countries in the sample. Looking at categorical groups – by age interval, education, work status, etc. – we find that female entrepreneurial rates are significantly lower than for males. For the first time we test for differences in the characteristics of female and male entrepreneurs and find that female entrepreneurs are slightly older, more frequently at home or not working, lower income and lower educated, and less access to business networks than their male counterparts. AS to the determinants of entrepreneurial rates themselves, the main differences across genders are the lower impact of secondary education and the larger impact of skills and fear of failure in female entrepreneurial rates relative to males. Results for entrepreneurship by opportunity and by necessity confirm the larger importance of specific skills for women creating new businesses,. Our results suggest that facilitating access to business networks and specific business skills are the most powerful instruments to increase the rates of female entrepreneurship.

Thematic Area: Entrepreneurial Economics.

1. INTRODUCTION

A new and growing literature has uncovered the importance of personal characteristics as determinants of entrepreneurial activity.[2] Characteristics such as education, personal income, work status and access to a network of entrepreneurs have been shown to affect the likelihood that any one person attempts to start a new business. A robust empirical fact receiving much less attention is the fact that, over time and across countries, entrepreneurship rates among women are about half those of males.[3] Though discussed and noted by several authors,[4] little has been done to explain what factors lie behind this important fact, including different personal

[1] I would like to thank Silvia Ardagna, José Mata, José Tavares, Stefan Mesiek, Rögnvaldur J Sæmundsson and the anonymous referees from the European Academy of Management for their valuable comments and the *Fundação para a Ciência e Tecnologia* for financial support. All remaining errors are my own. E-mail at **fajl@fct.unl.pt .**

[2] Notable examples are Ardagna and Lusardi (2008) and Djankov et al. (2006). There is no agreed to definition of entrepreneur. See Branco et al. (2008). Our task in the empirical exercise is facilitated by the use of a widely recognized data-set on intentions to start a business, the Global Entrepreneurship Monitor.

[3] Here we do not address the secular change in self-employment rates, as in Kuhn and Schuetze (2001), who finds that for women, most of the increase in self-employment is attributable to an increase in retention rates in self-employment, while for men, most is attributable to a decrease in the stability of paid employment.

[4] Ardagna and Lusardi (2008), for instance, use a male dummy indicator and find that males 1.1 percent higher probability of being an entrepreneur, for an average entrepreneurship arte of 5 percent.

characteristics, different returns to the characteristics and different goals when opening a business.[5] Figures 1 and 2 below plot the entrepreneurship rate by country against the female to male entrepreneurship rate ratio. There is a clear positive relation between the two variables, so that countries where women are relatively less entrepreneurial are also countries where total entrepreneurial activity is lower.[6] In other words, understanding the reasons why women are less frequently at the helm of new business ventures is a first step to devise policies that both help bridge the entrepreneurial gender gap as well as increase overall entrepreneurial activity. The latter is an important objective, as new firm creation is a major element in furthering economic growth and job creation.[7]

This paper undertakes a cross-country study of the determinants of entrepreneurship with a view to answer three related questions: first, do female entrepreneurs differ from their male counterparts?; second, do the personal determinants of entrepreneurial activity differ between females and males?

An important issue related to our empirical study is whether women suffer from discrimination when planning to start a business. Discrimination at the workplace is an important subject of study and recent studies have uncovered discrimination in other areas, including in the key area of access to credit.[8] Women receive less pay irrespective of their characteristics, as shown by the labor economics literature on gender discrimination.[9] Though we discuss the possible discrimination against women that want to start a business, our focus is instead on the differential determinants of entrepreneurial activity by females. However, we recognize that entrepreneurship may be an effective antidote to discrimination based on prejudice and on employers´ preferences, as it provides women with an autonomous avenue to circumvent social obstacles to employment, career progress and fair returns on effort.

Understanding the determinants of entrepreneurial activity by females is important also for policy reasons. First, increasing firm creation by females is a way to increase the productivity of the economy overcoming unnecessary barriers to women´s labor force participation, initiative and talent; second, females may be more able and more interested than males in undertaking activities in areas that that are particularly innovative and beneficial for the economy. Our discussion is interesting in the context of the allocation of talent model, which see the stock of talent – for instance, among women – as relatively constant but its allocation towards a range of activities possibly subject to major changes in response to institutions and policies.[10] More specifically, in the study of entrepreneurship, several authors have suggested that, while the stock of entrepreneurs is relatively constant, the nature and social impact of their activities can change dramatically with country institutions.[11]

2. GENDER AND ENTREPRENEURSHIP: DATA AND SUMMARY STATISTICS

[5] Cromie (1987) finds that both genders do have a variety of reasons for founding a business, primarily autonomy, achievement, a desire for job satisfaction and other non-economic rewards. The desire to make money, also present, is less important for women, who often choose entrepreneurship as a result of career dissatisfaction and as a means of meeting simultaneously their own career needs and the needs of their children.

[6] The simple correlation coefficient between the variables is 0.80 for total entrepreneurship and of 0.75 for entrepreneurship driven by opportunity, that is, intention to create a business that responds to a positive stimulus such as a new market opportunity, an innovation, etc.

[7] Entrepreneurship plays the vital role in the Schumpeterian process of "creative destruction" that fosters economic growth, innovation and employment.

[8] See Alesina et al. (2008) for an important study showing that, while women-led businesses are less prone to default, they have access to worse credit conditions than their male equivalent.

[9] See Cavalcanti and Tavares (2008) on the high aggregate output cost of gender discrimination.

[10] See Murphy et al. ??? (1991).

[11] See Baumol (1990). Niederle and Yestrumskas (2008) show that institutional design does affect the activity choices of women and men.

In this section we present the data on entrepreneurial activity across countries and over time, and relate it to individual characteristics of males and females.

2.1. The Data

In our empirical analysis we will draw on data from the Adult Population Surveys, collected by the Global Entrepreneurship Monitor (GEM). This data contain detailed information on individuals from 46 countries. W e can assess whether an individual is starting a new business, owns or manages a young firm, we can assess (at least partially) their motivation to start a firm, and take into account personal characteristics such as age, income, education, work status and skills. These micro survey data is collected annually and is made consistent across countries.[12] In this paper we use yearly data from 2001 to 2004.[13]

On the reliability of GEM data, Ardagna and Lusardi (2008) compare the GEM data with the Flash Eurobarometer Survey on Entrepreneurship collected by the European Commission for countries that are common to both data sets.[14] The percentage of individuals involved in entrepreneurial activity is very similar in both datasets. The same is true for individuals pursuing a business opportunity or for whom entrepreneurship is for necessity. Results are also very similar when the authors compare individual characteristics such as age, sex and work status.[15]

Acs, Desai and Klapper (2007) compare GEM data to the World Bank Group datasets (WBEGS) which reports formal entrepreneurial activity as the number of newly registered firms of limited liability corporations (LLCs). From GEM data the authors compute the "nascent entrepreneurship rate" – share of individuals actively involved in starting a new venture – and "baby entrepreneurship rate" – share of people that are owners or managers of a business less than 42 months old.[16] From the World Bank data, these authors compute the "corporate entrepreneurship rate" as the percentage of newly registered limited liability firms as a percentage of adult population. For the 41 countries examined between 2003 and 2005, Acs, Desai and Klapper (2007) find that GEM data tends to report significantly lower levels of early-stage that entrepreneurial activity in developed countries. Focusing on formal businesses, as WBEGS does, leads to the inclusion of initiatives that do not correspond to entrepreneurial activity, associated with legal incentives, in developed countries, to formally create new organizations.[17] As GEM data computes the number of individuals entrepreneurs, it may overlook individuals that are involved in multiple businesses. Acs, Desai and Kappler (2007)

[12] Each year a sample of at least 2,000 randomly selected individuals in each country are surveyed by phone or through face-to-face interviews. On average, a total of 35 national experts in each country are responsible for conducting the surveys. A coordination team at London Business School supervises and checks for inconsistencies.

[13] This is the set of surveys available to researchers who not directly involved in the GEM project, and also those for which the methodology is most consistent across time.

[14] Countries surveyed in both data sets are: Belgium, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, The Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Sweden, United Kingdom and United States. See Ardagna and Lusardi (2008).

[15] The only exception is the percentage of individuals who think that fear of failing could prevent them from starting a new business, which is higher in the Flash Eurobarometer Survey (47.5%) than in GEM data (33.3%).

[16] "Nascent" and "baby" entrepreneurship rates are two of the components of the Total Entrepreneurship Rate (TEA), the key variable in our study, as will become clear below.

[17] A couple of examples may illustrate this point. In Hong Kong all real estate sales, even those undertaken by individuals, are first converted to LLCs for tax reasons. In the United States firms may register several LLC´s to limit liability in different lines of businesses. Shell companies formally register new business ventures for tax reasons, (see Acs et al. (2007)). In Italy, labor laws restricting hiring and firing of employees apply to firms with more than 15 employees, which is an incentive for business owners to register multiple smaller firms to lower the regulatory burden (see Kappler et all 2006).

confirm that GEM data reports higher rates of entrepreneurship for developing countries, which the authors explain by the importance of the informal sector, captured by GEM data.

Reynolds et al. (2005) compare GEM estimates on new firm´s birth rate and national annual new firm´s estimates with the Official New Firm Census and data from the European Commission Report. They show that TEA index as well as other indexes calculated using GEM data are reliable and consistent with other datasets.

A study like ours, interested in assessing how personal characteristics affect entrepreneurial rates, particularly how they differ across gender, would like a data set that covers the widest possible number of individuals, independently of whether in the formal or informal sector, and give less salience to formal and legal aspects. We thus consider that GEM data is the appropriate choice.

2.2. Does Entrepreneurial Activity Differ Across Gender?

Table 1 and Figure 4 present total entrepreneurship rates (TEA) in the male and female population, for a cross-section of countries. These are individuals who are either starting a new business or are owners or managers of a young firm. In addition to total entrepreneurial activity, we also report rates of entrepreneurial activity driven by opportunity (TEA OPP) and by need (TEA NEC). Individuals who claim they are starting a new business to take advantage of a business opportunity are considered driven by opportunity; those that claim they could find no better job are classified as driven by necessity.[18] TEA OPP and TEA NEC sum up to total entrepreneurial activity, TEA. Please note that in virtually all countries in the sample TEA rates for females are smaller that TEA rates for males. This is evident in the last column of Table 1 and in Figure 4, where we compute the female to male TEA rates. The single exception to the rule is Thailand, where this ratio is equal to one. The lowest female to male TEA ratio is that of Croatia which compares, at 0.33, with the sample average of 0.53. On average females display half of the entrepreneurial activity than males. This gender imbalance is even more pronounced for entrepreneurial activity driven by opportunity - TEA OPP, in columns (6) and (7) -, and less pronounced for entrepreneurial activity driven by need – columns (9) and (10) -, as summarized in the Column "All". In general low income countries display higher entrepreneurial activity rates, and more balanced between females and males. In poor countries, entrepreneurs are relatively more driven by need than opportunity. In sum, entrepreneurs are more likely to arise in poor countries, where they are also more likely to be females driven by necessity.

Figure 3 shows that women are less likely to be entrepreneurs than their male counterparts irrespective of age. The entrepreneurship rate attains it maximum between the ages of 25 and 35 years old, for both men and women. The average age at which females and males become entrepreneurs is around 38 years.

Figures 5 and 6 display the age distribution of female and male entrepreneurs by motive. For both sexes and at all ages the opportunity motive is more prevalent than necessity. Women have lower rates but the rates of entrepreneurship by need, for females and males, are much closer than their equivalent for opportunity. Figures 7 and 8 show entrepreneurship rates by education and income levels, an important and clarifying exercise. Entrepreneurship rates increase with income and with education irrespective of gender. A closer look reveals that, while entrepreneurial activity driven by opportunity indeed rises with education and income, entrepreneurship by necessity behaves in the opposite fashion, decreasing with both income and education.

We now sharpen our question and try to answer whether female and male rates of entrepreneurial activity are different for a number of given characteristics. Our aim is to compute entrepreneurial rates by gender, for different ages groups, work status, education and income levels, social networks, etcetera. In Table 2 we present these different entrepreneurship rates and test whether the differences between females and males are statistically significant using the difference in means test. We reject the null hypothesis of equality of the female and male TEA and TEA OPP rates, at the 1% confidence levels, for most cases.

[18] Please consult Appendix I for more.

In the case of entrepreneurial activity by need - TEA NEC - we can not reject the null hypothesis of equality for female and males who work at home,[19] for individuals with same skills, or for countries that are classified as low or mid low income by the World Bank, or African countries in general.[20] Taking these results at face value, the only sensible policies that would raise female entrepreneurship rates to the levels of their male counterparts would be to generalize access to specific business skills.

2.3. Are Female and Male Entrepreneurs Different?

In Table 3 we try to answer a different but key question. Instead of comparing entrepreneurial rates across gender, we compare the characteristics of female and male entrepreneurs and ask: are they different? We test whether those differences are statistically significant using difference in means test and the 1% confidence level.

Table 3 shows that the average age of female and male entrepreneurs is very similar, at 38 years of age. Men that have a job have 8% higher probability of becoming entrepreneurs than women, 6% higher in the case of entrepreneurship by opportunity and 13% higher in the case of entrepreneurship by need. 5,2 % of women at home are entrepreneurs, compared to only 0,3% of men, and these numbers are similar for entrepreneurship by opportunity. Interestingly, 7.4% of women at home are entrepreneurs by need, compared with a paltry 0.2% for men.[21] Among the individuals not working at the time of the interview, 10% of the females are entrepreneurs, compared to 7.5% of males. 1.5% of male students are entrepreneurs, which compares to 2% of female students, and this is true for entrepreneurship by need as well as by necessity. 22.6% of female individuals who report their income in the lowest 33rd income percentile of the income distribution are entrepreneurs, as compared to 18% for males. The difference regarding entrepreneurs by need is even higher: at 31% for females, compared to 27% for males. The difference in means in the case of middle income is not statistically significant but it is in the case of upper income: 29.5% of males are entrepreneurs compared to 23% in the case of female. In terms of education there are almost no differences in gender probabilities of becoming an entrepreneur up to the college degree stage, when there are more women entrepreneurs.[22] Finally, it is more frequent that male entrepreneurs know someone who has started a business in the recent past.[23] The variable "skills" is also more relevant in the case of men: 84% of male entrepreneurs say they think they have the knowledge, skills and experience to start a new business, compared to 78% for females. This is consistent with the fear of failure results: 23% of female entrepreneurs compared to 19% of male entrepreneurs say that fear of failing can prevent them to start a new business.[24]

In sum, results in Table 3 show that female entrepreneurs are different as to whether they work – less do -, are at home – more do-, study – less do-, and whether they do not have a job – consistently, more female entrepreneurs are in this category. In addition, more female

[19] Edwards and Field-Hendrey (2002) find that home-based work is an attractive option for women for whom the fixed costs of work are high-women who have small children, are disabled, or live in rural areas-and that home-based workers are more likely to choose self-employment than are on-site workers.
[20] See Appendix I for variables definition.
[21] This result is consistent with Edwards and Field-Hendrey (2002) that using data from the 1990 Census, find that home-based work is an attractive option for women for whom the fixed costs of work are high. This is the case of women who have small children, are disabled, or live in rural areas-and that home-based workers are more likely to choose self-employment than are on-site workers.
[22] Vijverberg (1993) studies if women in the labor market enjoy the same returns to their human capital investments as men do and conclude that in Côte d'Ivoire, rates of return to education are high for both men and women, but men's wages exceed women's by a substantial margin for all but the most educated. Wong (1986) discusses the effect of spouse education on the productivity of entrepreneurial activity.
[23] Ibarra (1993) suggests that women have limited access to or are excluded from organizational networks and claims that the organizational context in which interaction networks are embedded produces unique constraints on women.
[24] Kalleberg and Leicht (1991) find that businesses headed by women were not more likely to go out of business, nor less successful, than those owned by men.

entrepreneurs are low income, when compared to males, and less are high income. No difference as to the gender incidence of entrepreneurship for middle income. Interestingly, using the income classification for countries, again women entrepreneurs are more frequent that their male counterparts in low to upper middle income countries and less so in high income countries. Also, female entrepreneurs are less connected to networks of entrepreneurs and are more fearful of being successful.

3. GENDER SPECIFIC DETERMINANTS OF ENTREPRENEURIAL ACTIVITY

We now formally analyze the individual determinants of gender activity and test how they differ across gender. Our dependent variables will be entrepreneurship rates- total, by opportunity and by need -, and our independent variables will include a host of individual characteristics, taken alone or interacted with a female indicator. For an individual *i*, in country j, at time t, we define the outcome of interest y_{ijt} :as one of the three measures of entrepreneurial activity: TEA, TEA OPP and TEA NEC. We estimate the equation for y_{ijt}:

$$y_{ijt} = \alpha + \beta_1 . X_{ijt} + \beta_2 . X_{ijt} . \text{Female}_i + \gamma_t + \varepsilon_{ijt} \quad (1)$$

where α is a constant and γ_t a vector of year dummies. X_{ijt} is our variable of interest, the vector of individual characteristics, including age, gender, employment status, education, income, the role of social networks, business skills, and fear of failure.[25] This same vector appears the interacted with the gender indicator, Female_i, which takes the value 1 when the individual is a woman. The coefficient on a variable such as "College", for instance, will give us the change in the probability of becoming an entrepreneur for an individual with the average characteristics in the sample. The variable "Female College" will give us the additional – positive or negative – effect on the probability of a female college graduate becoming and entrepreneur. Positive and significant coefficients on the variables interacted with the female indicator suggest females with that specific characteristic are more likely to be an entrepreneur that males with the same characteristic. The dependent variable is binary, and we use Probit estimation and correct the standard errors by clustering them at the country level. Our sample includes countries whose macroeconomic and institutional characteristics correlate both with the entrepreneurship indices. However, we control for countries' specific characteristics by including country fixed effects to our specification.

Tables 4 and 5 present our results. In Table 4, for a totals of 424 566 observations, and for total entrepreneurship rates – TEA -, we obtain that entrepreneurial activity for females starts earlier in life but then also drops at a faster rate at older ages – see the positive and negative signs of Female.Age and Female.Age Squared. Secondary education is less of an incentive for entrepreneurship for females, as well as a graduate education. More importantly, skills boost female entrepreneurship beyond that of males. Interestingly, while fear tends to discourage male entrepreneurship, the opposite seems true for females, as attested by the negative, significant, and large coefficient on Female.Fear.

Different pattern of results emerge for TEA OPP and TEA NEC, but the few differences are worth noting. Now age does not affect differently males and females, except that younger females are more likely than their male counterparts to engage in entrepreneurship driven by opportunity, TEA OPP. Also, females at home are less likely to engage in entrepreneurship by opportunity, the same is true for female with graduate education.[26] Female with entrepreneurial

[25] Ardagna and Lusardi (2008) estimate the probability of starting a business as a function of a set of demographic characteristics that are available in the data - age, gender, employment status, education, income, etc. - country by country for 2001 and 2002.

[26] Vijverberg (1993) finds that the return to education in self employment are high for both men and women, men´s returns exceeding that of women's by a substantial margin for all but the most educated.

skills and that state they are fearful they might not be successful are more likely to be entrepreneurs, relative to their male counterparts. As to entrepreneurship by need, we find women who hold a job are significantly less likely to be entrepreneurs, as well as those that are students, that do not work. Again, we confirm that the existence of skills is even more important for females than for male entrepreneurs.

Notice that in all specification we have included a gender indicator, which is negative and significant for total entrepreneurial activity and for the subset of that activity that is driven by opportunity. For entrepreneurship by need, the female dummy does not capture significant differences between males and females, once the individuals characteristics and the interacted terms are taken into account.

Table 5 reruns the basic specification for countries that are classified as high income by the World Bank and fro the other, middle and low income countries. The third column introduces country dummies so that the specification is now: given by

$$y_{ijt} = \alpha_j + \beta_1 . X_{ijt} + \beta_2 . X_{ijt} . \text{Female}_i + \gamma_t + \varepsilon_{ijt} \qquad (2)$$

as in equation (1), except that α_j is a set of country dummies. For High income countries the main differences are that an individual with a job now increases his or her chance of becoming an entrepreneur, females at home are significantly less likely to be entrepreneurs, low income individuals are *more* likely to be entrepreneurs but that effect is not present for females[27], and some graduate study increase the chance of being entrepreneurs for both male and female. For Low and Middle income countries age does differentiate entrepreneurial women and low income women indeed are les likely to be entrepreneurs. Fear of failure also does not distinguish female entrepreneurs. Interestingly, as was the case for entrepreneurship by need, the female dummy is not significantly different from zero.

5. CONCLUSION

This paper examines, for the first time, whether individual characteristics have a different impact on the likelihood of one becoming an entrepreneur, for males and females By examining total entrepreneurship rates, entrepreneurship driven by opportunity and by need, we find that indeed entrepreneurial activity rates are statistically and significantly lower for all categorical groups – by age interval, education, work status, network access, etc. –, except, in the case of entrepreneurship by need and the case of persons working at home, with specific entrepreneurial skills and that live in a middle or low income country.

We then estimate the differences in the personal characteristics of entrepreneurs across gender and find that female entrepreneurs are slightly older, more frequently at home or not working, lower income, lower educated, and with less access to specific skills than their male counterparts. As to the determinants of entrepreneurial rates themselves, the main differential impact across genders are the lower impact of secondary and graduate education, and the larger impact of skills and fear of failure in female entrepreneurial rates relative to males. Results for entrepreneurship by opportunity and by necessity confirm the larger importance of specific skills for women creating new businesses.

Combining the mean differences in male and female entrepreneurs´ and the how they impact entrepreneurship rates, the main policy implication of our paper is that creating business networks accessible to females and imparting the specific business skills associated with entrepreneurship may be the most potent levers to increase female and total entrepreneurship across countries.

[27] As shown by the almost identical size and opposite signs of the coefficients on Low Income and Female.Low Income.

REFERENCES

[1] ACS Z., S. DESAI, KLAPPER L., (2007), "What Does "Entrepreneurship" Data Really Show? A Comparison of the Global Entrepreneurship Monitor and World Bank Group Datasets."

[2] ALESINA, ALBERTO, LOTTI, FRANCESCA, AND PAOLO EMILIO MISTRULLI (2008), "Do Women Pay More For Credit? Evidence from Italy", NBER Working Paper 14202.

[3] ALFARO L., A. CHARLTON, (2007), "International Financial Integration and Entrepreneurial Firm Dynamics", Harvard Business School Working Paper, n. 07-012.

[4] ARDAGNA, SILVIA AND ANNAMARIA LUSARDI (2008). "Explaining International Differences in Entrepreneurship: The Role of Individual Characteristics and Regulatory Constraints", NBER Working Paper No. 14012.

[5] BAYOUMI T., D. LAXTON, P. PESENTI, (2004), "Benefits and Spillovers of Greater Competition in Europe: a Macroeconomic Assessment", Federal Reserve Bank of New York Staff Report 182.

[6] BARTON H. HAMILTON (2000). "Does Entrepreneurship Pay? An Empirical Analysis of the Returns of Self-Employment" The Journal of Political Economy, Vol. 108, No. 3, pp. 604-631.

[7] BERTRAND, M. AND F. KRAMARZ (2002). "Does Entry Regulation Hinder Job Creation? Evidence From The French Retail Industry." Quarterly Journal of Economics, November, 1369-1413.

[8] BLANCHFLOWER D., (2004), "Self-Employment: More May Not Be Better", Swedish Economic Policy Review, vol 11(2), Fall 2004, pp. 15-74.

[9] BLANCHFLOWER D., (2000), "Self-Employment in OECD Countries", Labour Economics, 7, September 2000, pp. 471-505.

[10] BLANCHFLOWER, DAVID G. AND ANDREW J. OSWALD (1998). "What Makes an Entrepreneur?". Journal of Labor Economics, Vol. 16, No. 1 (Jan., 1998), pp. 26-60.

[11] BLANCHFLOWER D., A. OSWALD, AND A. STUTZER, (2001), "Latent Entrepreneurship Across Nations", European Economic Review, 45(4-6) pp. 680-691, May 2001.

[12] BLANCHFLOWER D., A. OSWALD, (1998), "What Makes an Entrepreneur?", Journal of Labor Economics, January, 16(1) pp. 26-60, 1998.

[13] BLANCHARD, O. J. AND J. WOLFERS (2000). "The Role of Shocks and Institutions in the Rise of European Unemployment: The Aggregate Evidence." Economic Journal, 110, 1-33.

[14] BOWEN, DONALD D. AND ROBERT D. HISRICH (1986). "The Female Entrepreneur: A Career Development Perspective". The Academy of Management Review, Vol. 11, No. 2 pp. 393-407.

[15] BRANCO, RUBEN, LLUSSÁ, FERNANDA, AND JOSÉ TAVARES (2008), "Entrepreneurship: Concepts and Measurement", Mimeo, Universidade Nova de Lisboa.

[16] CICCONE A., E. PAPAIOANNOU, (2006), "Red tape and Delayed Entry", CEPR Discussion Paper n 5996.

[17] COLE, ARTHUR H. (1942). "Entrepreneurship as an Area of Research". The Journal of Economic History, Vol. 2, Supplement: The Tasks of Economic History, pp. 118-126.

[18] CROMIE, STANLEY (1987). "Motivations of Aspiring Male and Female Entrepreneurs". Journal of Occupational Behaviour, Vol. 8, No. 3, pp. 251-261.

[19] DESAI, M., P. GOMPERS AND J. LERNER, (2003), "Institutions, Capital Constraints and Entrepreneurial Firm Dynamics: Evidence from Europe", NBER Working Paper No. 10165.

[20] DJANKOV S., T. GANSER, C. MCLIESH, R. RAMALHO, A. SHLEIFER, (2007), The Effect of Corporate Taxes on Investment and Entrepreneurship", unpublished.

[21] DJANKOV S., Y. QIAN, G. ROLAND, AND E. ZHURAVSKAYA, (2006a), "Who are China's Entrepreneurs?", American Economic Review, Papers and Proceedings, vol. 96, May 2006, pp. 348-352.

[22] DJANKOV S., Y. QIAN, G. ROLAND, AND E. ZHURAVSKAYA, (2006b), "Entrepreneurship in China and Russia Compared", Journal of the European Economic Association, vol. 4, n. 2-3, April-May 2006, pp. 352-365.

[23] DJANKOV S., Y. QIAN, G. ROLAND, AND E. ZHURAVSKAYA, (2005), "Who are Russia's Entrepreneurs?", Journal of the European Economic Association, vol. 3, n. 2-3, April-May 2005, pp. 1-11.

[24] DJANKOV S., R. LA PORTA, LOPEZ-DE-SILANES F., SHLEIFER A., (2002), "The Regulation of Entry", Quarterly Journal of Economics, February, pp. 1-37.

[25] EDWARDS, LINDA N. AND ELIZABETH FIELD-HENDREY (2002). "Home-Based Work and Women's Labor Force Decisions". Journal of Labor Economics, Vol. 20, No. 1, pp. 170-200.

[26] EVANS D., B. JOVANOVIC, (1989), "An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints", Journal of Political Economy, vol. 97, no. 4, pp. 808-827.

[27] EVANS, DAVID S. AND LINDA S. LEIGHTON (1989). "Some Empirical Aspects of Entrepreneurship".The American Economic Review, Vol. 79, No. 3 , pp. 519-535.

[28] FAIRLIE R., A. ROBB, (2007), "Why Are Black-Owned Businesses Less Successful than White-Owned Businesses? The Role of Families, Inheritances, and Business Human Capital", Journal of Labor Economics, 25(2): 289-323.
[29] GARCIA, N. CHIN, (2005), "Gobal Entrepreneurship Monitor: Data Collection Design and Implementation 1998-2003," Small Business Economics, vol. 24, pp. 205-231.
[30] GENTRY W., R. G. HUBBARD, (2000), "Tax Policy and Entrepreneurial Entry," American Economic Review Papers and Proceedings, vol. 90, no. 2, pp. 283 - 287.
[31] GUISO L., F. SCHIVARDI, (2006),"What Determines Entrepreneurial Clusters?". Mimeo.
[32] GUISO L., P. SAPIENZA, L. ZINGALES, (2004), "Does Local Financial Development Matter?", Quarterly Journal of Economics, August, pp. 929-969.
[33] HARBURY, C. D. AND D. M. W. N. HITCHENS (1977). "Women, Wealth and Inheritance", The Economic Journal, Vol. 87, No. 345, pp. 124-131.
[34] HUNDLEY, GREG (2000). "Male/Female Earnings Differences in Self-Employment: The Effects of Marriage, Children, and the Household Division of Labor".Industrial and Labor Relations Review, Vol. 54, No. 1, pp. 95-114.
[35] HURST E., A. LUSARDI, (2004), "Liquidity Constraints, Household Wealth, and Entrepreneurship," Journal of Political Economy, April 2004, vol. 112(2), pp. 319-347.
[36] HURST E., A. LUSARDI, (2008), "Liquidity Constraints and Entrepreneurship. Household Wealth, Parental Wealth, and the Transition In and Out of Entrepreneurship", forthcoming in "Overcoming Barriers to Entrepreneurship," Rowman and Littlefield.
[37] IBARRA, HERMINIA (1993) "Personal Networks of Women and Minorities in Management: A Conceptual Framework." The Academy of Management Review, Vol. 18, No. 1, pp. 56-87.
[38] KALLEBERG, ARNE L. AND KEVIN T. LEICHT (1991). "Gender and Organizational Performance: Determinants of Small Business Survival and Success" .The Academy of Management Journal, Vol. 34, No. 1 pp. 136-161.
[39] KHAN, B. ZORINA (1996). "Married Women's Property Laws and Female Commercial Activity: Evidence from United States Patent Records, 1790-1895". The Journal of Economic History, Vol. 56, No. 2, Papers Presented at the Fifty-Fifth Annual Meeting of the Economic History Association, pp. 356-388
[40] KEHRER, BARBARA H. (1976). "Factors Affecting the Incomes of Men and Women Physicians: An Exploratory Analysis". The Journal of Human Resources, Vol. 11, No. 4, pp. 526-545
[41] KLAPPER L., L. LAEVEN, AND R. RAJAN, (2006) ,"Entry Regulation as a Barrier to Entrepreneurship," Journal of Financial Economics (forthcoming).
[42] KUHN, PETER J. AND HERB J. SCHUETZE (2001). "Self-Employment Dynamics and Self-Employment Trends: A Study of Canadian Men and Women, 1982-1998". The Canadian Journal of Economics / Revue canadienne d'Economique, Vol. 34, No. 3 pp. 760-784.
[43] LA PORTA R., F. LOPEZ-DE-SILANES, A. SHLEIFER, R.W. VISHNY, (2000), "Investor Protection and Corporate Governance," Journal of Financial Economics, vol. 58, pp. 1-25.
[44] LOUNSBURY, MICHAEL AND MARY ANN GLYNN.(2001). "Cultural Entrepreneurship: Stories, Legitimacy, and the Acquisition of Resources." Strategic Management Journal, Vol. 22, No. 6/7, Special Issue: Strategic Entrepreneurship: Entrepreneurial Strategies for Wealth Creation, pp. 545-564.
[45] MORRIS, MICHAEL H., DUANE L. DAVIS AND JEFFREY W. ALLEN.(1994) "Fostering Corporate Entrepreneurship: Cross-Cultural Comparisons of the Importance of Individualism versus Collectivism". Journal of International Business Studies, Vol. 25, No. 1, pp. 65-89.
[46] NANDA R., AND J. B. SORENSEN, (2007), "Peer Effects and Entrepreneurship", Harvard Business School Working Paper n. 08-051.
[47] NIEDERLE, MURIEL, AND ALEXANDRA H. YESTRUMSKAS (2008), "Gender Differences in Seeking Challenges: The Role of Institutions", NBER Working Paper 13922.
[48] PAPANEK, GUSTAV F. (1962). "The Development of Entrepreneurship". The American Economic Review, Vol. 52, No. 2, Papers and Proceedings of the Seventy-Fourth Annual Meeting of the American Economic Association, pp. 46-58.
[49] PARASURAMAN, SAROJ AND CLAIRE A. SIMMERS (2001). "Type of Employment, Work-Family Conflict and Well-Being: A Comparative Study". Journal of Organizational Behavior, Vol. 22, No. 5 (Aug., 2001), pp. 551-568.
[50] PURI M., D. T. ROBINSON, (2006), "Who Are Entrepreneurs and Why Do They Behave That Way", Working Paper Duke University. Reynonds P., N. Bosma, E. Autio, S. Hunt, N. De Bono, I. Servais, P. Lopez-
[51] SHANE, SCOTT.(2000) "Prior Knowledge and the Discovery of Entrepreneurial Opportunities". Organization Science, Vol. 11, No. 4, pp. 448-469.

[52] STEWART, ALEX.(1990) "The Bigman Metaphor for Entrepreneurship: A "Library Tale" with Morals on Alternatives for Further Research".Organization Science, Vol. 1, No. 2, pp. 143-159.
[53] VAN PRAAG, C. M. AND J. S. CRAMER (2001). "The Roots of Entrepreneurship and Labour Demand: Individual Ability and Low Risk Aversion".Economica, New Series, Vol. 68, No. 269, pp. 45-62.
[54] VIJVERBERG, WIM P. M. (1993). "Educational Investments and Returns for Women and Men in Côte d'Ivoire", The Journal of Human Resources, Vol. 28, No. 4, Special Issue: Symposium on Investments in Women's Human Capital and Development, pp. 933-974
[55] WONG, YUE-CHIM (1986). "Entrepreneurship, Marriage, and Earnings". The Review of Economics and Statistics, Vol. 68, No. 4, pp. 693-699.

Appendix I
VARIABLE DEFINITON

TEA = 1 if individuals are starting a new business or are owners and managers of a young firm, 0 otherwise. Source: Global Entrepreneurship Monitor (GEM).
TEAOPP = 1 if individuals are starting a new business or are owners and managers of a young firm to take advantage of a business opportunity, 0 otherwise. Entrepreneurship indices - source: Global Entrepreneurship Monitor (GEM).
TEANEC = 1 if individuals are starting a new business or are owners and managers of a young firm because they could find no better economic work, 0 otherwise. Entrepreneurship indices - source: Global Entrepreneurship Monitor (GEM).
AGE 14-20, AGE 21-25, AGE 26-35 etc.= age of individuals at the time of the interview categorized by range.
WORKING=individuals who work at the time of the interview.
RETIRED= individuals who are retired at the time of the interview.
AT HOME= individuals who work at home at the time of the interview. STUD= individuals who are students at the time of the interview.
NOT WORKING=individuals who do not work at the time of the interview (and are not students, not retired, and do not work at home).
LOW INCOME=individuals who report that their income is in the lowest 33rd income percentile of their country´s income distribution at the time of the interview.
MID INCOME = individuals who report that their income is in the middle 33rd income percentile of their country´s income distribution at the time of the interview.
UP INCOME = individuals who report that their income is in the upper 33rd income percentile of their country´s income distribution at the time of the interview.
HIGH SCHOOL (EDSEC)= only individuals with a high school degree.
COLLEGE (EDPSEC)= only individuals with a college degree.
GRADUATE (EDGRA) = only individuals with at least some graduate school education.
KNOWS ENTREPRENEUR (KNOW) = individuals who know someone who has started a business in the recent past.
HAS SKILLS= individuals who think they have the knowledge, skills and experience to start a new business.
FEAR OF FAILURE = individuals who answer that fear of failing can prevent them to start a new business.
LOW INCOME WB (LOWINWB)= individuals who are interviewed at India and Uganda.
MID LOW INCOME WB (MIDLOIWB) = individuals who are interviewed at China, Thailand, Chinese Shenzhen, Peru and Jordan.
UP MID INCOME WB (UPMIDIWB) = individuals who are interviewed at Argentina, Chile, Croatia, Hungary, Mexico, Poland, Russia, South Africa, Brazil, Venezuela.
HIGH INCOME WB (HIGHIWB) = individuals who are interviewed at Australia, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Iceland, Ireland, Israel, Italy, Japan, The Netherlands, Norway, New Zealand, Portugal, Scotland (UK), Singapore, Slovenia, Spain, South Korea, Sweden, Switzerland, Taiwan, United Kigdom and United States.
OECD = Australia,Belgium,Canada,Denmark, Finland, France, Germany, Greece, Iceland, Ireland,Italy, Japan, The Netherlands, Norway, New Zealand, Portugal, Scotland (UK), Spain, Swedeen, Switzerland, United Kingdom, United States.
EU = Belgium, Greece, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, The Netherlands, Portugal, Scotland (UK), Spain, Sweden, United Kingdom.
ECA = World bank classification-Europe & Central Asia - Croatia, (Hungary), Poland, Russia, (Slovenia)-in parenthesis are not according to world bank classification.
EAP = World bank classification-East Asia & Pacific - China, Hong Kong, Singapore, South Korea, (Taiwan), Thailand, (Chinese-Shenzhen)- in parenthesis are not according to world bank classification.
LATIN = World Bank classification Latin America - Argentina, Brazil, Chile, Mexico, Peru, Venezuela.
AFRICA = World Bank classification Sub-Saharan Africa - South Africa, Uganda.

Appendix II
GEM QUESTIONNAIRE QUESTIONS

The following are the questionnaires' questions that the GEM coordination team uses to generate the variables TEA, TEA OPP, TEA NEC. Questions are from the 2002 data documentation manual. Questions asked in 2001 were exactly the same, even though the numbering of the questions changes. The methodology followed to construct the indices is based on procedures previously used in the US Panel Study of Entrepreneurial Dynamics and it is described in detail in the 2001 and 2002 Adult Population Surveys' data documentation and in Reynolds et al. (2005).

1. Which of the following would apply to you? (Possible answers: Yes, No, Don't Know, Refused)
1a. You are, alone or with others, currently trying to start a new business, including any self-employment or selling any goods or services to others.
1b. You are, alone or with others, currently trying to start a new business or a new venture for your employer– an effort that is part of your normal work.
1c. You are, alone or with others, currently the owner of a company you help manage, self-employed, or selling any goods or services to others.
If "Yes", Or "Don't Know" To Qu. 1a or Qu. 1b, Ask Qu 2a. If "Yes", Or "Don't Know" To Qu. 1c, Ask Qu. 3a.

2a. Over the past twelve months have you done anything to help start a new business, such as looking for equipment or a location, organizing a start-up team, working on a business plan, beginning to save money, or any other activity that would help launch a business?
2b. Will you personally own all, part, or none of this business?
2d. Has the new business paid any salaries, wages, or payments in kind, including your own, for more than three months?
2d1. What was the first year the owners received wages, profits, or payments in kind?
2g. Are you involved in this start-up to take advantage of a business opportunity or because you have no better choices for work?

3a. Do you personally own all, part, or none of this business?
3c. What was the first year the owners received wages, profits, or payments in kind? Payments in kind refers to goods or services provided as payments for work rather than cash.
3g. Are you involved in this firm to take advantage of a business opportunity or because you have no better choices for work.

The following are the questionnaires' questions used to define the variables *Knowent*, *Skills*, and *Fear f ail* respectively. Questions are from the 2002 data documentation manual. Questions asked in 2001 were exactly the same even though the numbering of the questions changes.
1. Which of the following would apply to you? (Possible answers: Yes, No, Don't Know, Refused)
1g. You know someone personally who started a business in the past 2 years
1h. In the next six months there will be good opportunities for starting a business in the area where you live
1i. You have the knowledge, skill and experience required to start a new business 1j. Fear of failure would prevent you from starting a business

Table 1: Entrepreneurship Rate by Country

	N.Obs.	TEA (%)			TEA OPP (%)			TEA NEC (%)			TEA Fem/Male
		All	Fem	Male	All	Fem	Male	All	Fem	Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Argentina	7998	12.77	9.18	16.58	7.54	4.64	10.62	4.69	4.10	5.31	0.55
Australia	7661	7.68	6.00	10.10	6.13	4.81	8.05	1.16	0.82	1.66	0.59
Belgium	12158	2.37	1.31	3.63	1.94	1.03	3.04	0.21	0.18	0.25	0.36
Brazil	10000	12.05	10.01	14.01	6.18	4.83	7.47	5.52	4.92	6.10	0.71
Canada	5944	6.12	3.99	8.24	4.86	3.08	6.63	0.98	0.68	1.27	0.48
Chile	4008	13.37	10.75	16.05	7.58	5.53	9.69	5.09	4.69	5.50	0.67
China	3661	11.69	8.97	14.61	6.28	4.01	8.72	5.11	4.59	5.66	0.61
Chinese Shenzhen	2040	7.45	4.25	10.34	5.98	3.32	8.36	1.42	0.83	1.95	0.41
Croatia	6017	2.49	1.31	3.99	1.50	0.68	2.52	0.71	0.45	1.05	0.33
Denmark	8048	4.57	2.79	6.52	4.20	2.48	6.08	0.22	0.19	0.26	0.43
Finland	8011	3.37	2.45	4.30	2.83	2.13	3.54	0.24	0.17	0.30	0.57
France	7991	1.91	1.30	2.60	1.49	0.97	2.07	0.34	0.28	0.40	0.50
Germany	37156	4.37	2.91	6.05	3.20	2.01	4.58	0.98	0.77	1.23	0.48
Greece	4008	5.71	3.11	8.65	3.97	1.98	6.21	1.52	1.13	1.96	0.36
Hong Kong	6004	2.43	1.47	3.52	1.60	0.94	2.34	0.82	0.50	1.17	0.42
Hungary	6878	5.60	4.37	6.87	3.69	2.86	4.56	1.53	1.29	1.78	0.64
India	5058	13.56	10.02	16.91	7.55	5.09	9.88	5.16	4.36	5.92	0.59
Iceland	6013	9.00	6.27	11.82	7.47	5.19	9.82	0.63	0.39	0.88	0.53
Ireland	7920	6.59	4.02	9.49	5.38	3.21	7.82	1.00	0.62	1.42	0.42
Israel	5992	4.81	2.73	7.10	3.00	1.65	4.50	0.87	0.67	1.09	0.38
Italy	8887	3.39	2.50	4.34	2.50	1.76	3.29	0.35	0.37	0.33	0.58
Japan	7893	1.63	0.94	2.33	1.04	0.61	1.47	0.35	0.23	0.48	0.40
Jordania	2000	19.10	13.70	23.12	15.20	10.66	18.59	2.85	1.76	3.66	0.59
Korea (South)	4023	10.84	6.00	15.79	6.19	3.29	9.15	3.38	1.67	5.13	0.38
Mexico	3016	16.15	14.13	19.92	10.51	8.59	14.11	4.97	5.03	4.86	0.71
Netherlands	12535	3.18	1.95	4.79	2.77	1.74	4.12	0.26	0.11	0.44	0.41
New Zealand	7848	11.53	9.31	14.53	9.57	7.72	12.07	1.67	1.30	2.16	0.64
Norway	9833	5.55	3.06	8.14	4.75	2.48	7.10	0.41	0.24	0.58	0.38
Peru	2007	39.61	38.60	40.68	26.31	24.93	27.77	12.95	13.19	12.70	0.95
Poland	6001	5.48	3.58	7.44	3.25	1.94	4.60	2.13	1.54	2.74	0.48
Portugal	3000	4.47	2.81	6.28	3.40	1.72	5.24	1.00	1.02	0.98	0.45
Russia	2190	1.96	1.34	2.72	1.32	0.84	1.91	0.41	0.42	0.40	0.49
Scotland(UK)	2118	2.64	1.67	3.76	2.12	1.23	3.15	0.47	0.35	0.61	0.44
Singapore	9735	5.11	2.83	7.36	4.26	2.50	6.01	0.74	0.31	1.17	0.38
Slovenia	6045	2.58	1.31	3.88	2.03	0.95	3.14	0.51	0.33	0.70	0.34
South Africa	15519	5.39	4.23	6.54	3.36	2.44	4.28	1.61	1.52	1.70	0.65
Spain	27996	5.37	3.11	7.59	4.51	2.64	6.37	0.75	0.41	1.09	0.41
Sweden	32780	3.13	1.78	4.41	2.70	1.54	3.79	0.34	0.18	0.49	0.40
Switzerland	4004	5.09	3.39	7.14	4.32	2.93	5.99	0.60	0.37	0.88	0.47
Taiwan	2236	3.09	2.23	3.77	2.50	1.82	3.04	0.45	0.20	0.64	0.59
Thailand	1043	20.04	20.06	20.00	16.01	16.30	15.56	3.16	3.13	3.21	1.00
Uganda	3020	28.71	25.89	32.01	14.87	12.52	17.63	12.78	12.64	12.95	0.81
UK	66434	4.07	2.66	5.99	3.31	2.16	4.87	0.60	0.38	0.91	0.44
US	21056	8.57	6.17	11.07	6.98	4.95	9.10	1.03	0.81	1.26	0.56
Venezuela	2000	23.55	21.7	25.4	13.15	10.80	15.50	9.45	10.00	8.90	0.85
All	**425785**	**5.92**	**4.17**	**7.88**	**4.32**	**2.91**	**5.90**	**1.32**	**1.06**	**1.61**	**0.53**
Low Income WB	8078	19.23	16.35	22.17	10.29	8.05	12.58	8.01	7.66	8.37	0.74
Mid Low Inc. WB	10751	18.29	15.87	20.71	12.57	10.41	14.73	5.26	4.94	5.59	0.77
Up Mid Inc. WB	63627	8.59	6.79	10.50	5.02	3.65	6.49	3.15	2.84	3.48	0.65
High Income WB	343329	4.73	3.08	6.60	3.79	2.44	5.33	0.70	0.47	0.95	0.47
OECD	309294	4.74	3.12	6.58	3.85	2.50	5.39	0.66	0.46	0.89	0.47
EU	245055	4.13	2.61	5.87	3.35	2.08	4.81	0.59	0.40	0.81	0.44
ECA	27131	3.92	2.56	5.40	2.55	1.56	3.62	1.16	0.86	1.49	0.47
EAP	28742	6.74	4.58	8.95	4.64	3.08	6.25	1.80	1.25	2.35	0.51
Latin	29029	15.55	13.16	18.13	9.07	7.14	11.15	5.96	5.58	6.37	0.73
Africa	18539	9.19	8.00	10.40	5.24	4.19	6.31	3.43	3.46	3.41	0.77

Notes: See Appendix I for exact variable definition.

Table 2: Are Female Entrepreneurship Rate Different than Male?

	TEA			TEA OPP			TEA NEC		
	Mean Fem=1	Mean Male=1	St.Error of Diff	Mean Fem=1	Mean Male=1	St.Error of Diff	Mean Fem=1	Mean Male=1	St.Error of Diff
Age 14-20	0.0279	0.0512	0.0021***	0.0189	0.0355	0.0018***	0.0082	0.0135	0.0011***
Age 21-25	0.0520	0.0964	0.0027***	0.0357	0.0737	0.0023***	0.0139	0.0192	0.0013***
Age 26-35	0.0656	0.1216	0.0020***	0.0465	0.0933	0.0018***	0.0163	0.0225	0.0010***
Age 36-45	0.0553	0.1018	0.0018***	0.0390	0.0767	0.0016***	0.0136	0.0208	0.0009***
Age 46-55	0.0393	0.0734	0.0017***	0.0268	0.0542	0.0014***	0.0103	0.0155	0.0008***
Age 56-65	0.0196	0.0429	0.0014***	0.0138	0.0300	0.0012***	0.0046	0.0101	0.0007***
Age old 65	0.0062	0.0157	0.0010***	0.0041	0.0113	0.0008***	0.0017	0.0031	0.0004***
Working	0.0611	0.1009	0.0011***	0.0447	0.0774	0.0010***	0.0134	0.0190	0.0005***
Retired	0.0069	0.0155	0.0010***	0.0042	0.0108	0.0009***	0.0020	0.0044	0.0006***
At Home	0.0167	0.0370	0.0055***	0.0098	0.0295	0.0049***	0.0061	0.0051	0.0021
Student	0.0148	0.0307	0.0021***	0.0113	0.0235	0.0018***	0.0029	0.0055	0.0009***
Not Working	0.0258	0.0408	0.0014***	0.0152	0.0252	0.0011***	0.0095	0.0143	0.0009***
Low Income	0.0337	0.0622	0.0013***	0.0202	0.0403	0.0011***	0.0119	0.0192	0.0008***
Middle Income.	0.0415	0.0733	0.0014***	0.0298	0.0553	0.0012***	0.0097	0.0149	0.0007***
Up Income	0.0568	0.1022	0.0018***	0.0460	0.0853	0.0017***	0.0079	0.0123	0.0007***
High School	0.0341	0.0710	0.0012***	0.0242	0.0538	0.0011***	0.0081	0.0142	0.0006***
College	0.0499	0.0950	0.0017***	0.0396	0.0768	0.0015***	0.0076	0.0130	0.0007***
Graduate	0.0598	0.1049	0.0030***	0.0490	0.0892	0.0028***	0.0082	0.0115	0.0011***
Knows Entrepreneur	0.0993	0.1495	0.0019***	0.0721	0.1147	0.0016***	0.0226	0.0276	0.0009***
Has Skills	0.1206	0.1542	0.0018***	0.0862	0.1172	0.0016***	0.0289	0.0299	0.0009
Fear of failure	0.0334	0.0594	0.0013***	0.0207	0.0403	0.0010***	0.0112	0.0163	0.0007***
Low Income WB	0.1635	0.2217	0.0088***	0.0805	0.1258	0.0068***	0.0766	0.0837	0.0060
Mid Low Inc. WB	0.1587	0.2071	0.0074***	0.1041	0.1473	0.0064***	0.0494	0.0559	0.0043
Up Mid Income WB	0.0679	0.1050	0.0022***	0.0365	0.0649	0.0017***	0.0284	0.0348	0.0014***
High Income WB	0.0308	0.066	0.0007***	0.0244	0.0533	0.0007***	0.0047	0.0095	0.0003***
OECD	0.0312	0.0658	0.0008***	0.0250	0.0539	0.0007***	0.0046	0.0089	0.0003***
EU	0.0261	0.0587	0.0008***	0.0208	0.0481	0.0007***	0.0040	0.0081	0.0003***
ECA	0.0256	0.0540	0.0024***	0.0156	0.0362	0.0019***	0.0086	0.0149	0.0013***
EAP	0.0458	0.0895	0.0030***	0.0308	0.0625	0.0025***	0.0125	0.0235	0.0016***
Latin America	0.1316	0.1813	0.0043***	0.0714	0.1115	0.0034***	0.0558	0.0637	0.0028***
Africa	0.0800	0.1040	0.0042***	0.0419	0.0631	0.0033***	0.0346	0.0341	0.0027

Notes: Difference in means statististically different from zero at 1%(***). TEA= Total Entrepreneurship Rate, TEA OPP= Entrepreneurship Rate by Opportunity and TEA NEC = Entrepreneurship Rate by Necessity. ECA = Europe and Central Asia. EAP = East Asia and Pacific. See Appendix for exact definition of the variables.

Table 3 Are Women Entrepreneurs Different?

	TEA			TEA OPP			TEA NEC		
	Mean Fem=1	Mean Male=1	St.Error of Diff	Mean Fem=1	Mean Male=1	St.Error of Diff	Mean Fem=1	Mean Male=1	St.Error of Diff
Age	38.144	37.879	0.1594*	38.109	37.697	0.1869**	37.874	38.100	0.3349
Work	0.7220	0.8096	0.0056***	0.7587	0.8292	0.0063***	0.6248	0.7464	0.0125***
Retir	0.0166	0.0191	0.0017	0.0146	0.0177	0.0019	0.0185	0.0262	0.0039*
Home	0.0519	0.0028	0.0023***	0.0437	0.0029	0.0026***	0.0742	0.0019	0.0054***
Stud	0.0153	0.0196	0.0017***	0.0167	0.0200	0.0020	0.0118	0.0173	0.0032*
Nwork	0.1039	0.0751	0.0038***	0.0875	0.0618	0.0041***	0.1505	0.1287	0.0094**
Lowinc	0.2262	0.1819	0.0053***	0.1949	0.1573	0.0059***	0.3149	0.2749	0.0123***
Midinc	0.2521	0.2514	0.0057	0.2600	0.2536	0.0067	0.2336	0.2505	0.0116
Upinc	0.2319	0.2952	0.0057***	0.2700	0.3290	0.0070***	0.1269	0.1743	0.0095***
Lowinwb	0.0714	0.0558	0.0032***	0.0504	0.0423	0.0033**	0.1320	0.1031	0.0088***
Midloiwb	0.0912	0.0701	0.0036***	0.0858	0.0666	0.0042***	0.1121	0.0926	0.0082**
Upmidiwb	0.2384	0.2036	0.0054***	0.1836	0.1681	0.0059***	0.3938	0.3305	0.0130***
Highiwb	0.5990	0.6705	0.0063***	0.6802	0.7230	0.0071***	0.3621	0.4739	0.0132***
Edsec	0.2749	0.2860	0.0058*	0.2800	0.2896	0.0069	0.2589	0.2792	0.0120*
Edpsec	0.2523	0.2704	0.0057***	0.2873	0.2920	0.0070	0.1522	0.1805	0.0100***
Edgra	0.1050	0.1073	0.0040	0.1234	0.1218	0.0051	0.0565	0.0574	0.0063
Know	0.5789	0.6666	0.0063***	0.6030	0.6834	0.0074***	0.5198	0.6029	0.0134***
Skills	0.7873	0.8481	0.0051***	0.8069	0.8612	0.0058***	0.7441	0.8047	0.0113***
Fear	0.2337	0.1920	0.0054***	0.2076	0.1740	0.0061***	0.3082	0.2579	0.0122***

Notes: Difference in means statististically different from zero at 1%(***) , 5% (**) and 10% (*). TEA= Total Entrepreneurship Rate, TEA OPP= Entrepreneurship Rate by Opportunity and TEA NEC= Entrepreneurship Rate by Necessity. See Appendix for the exact definition of the variables.

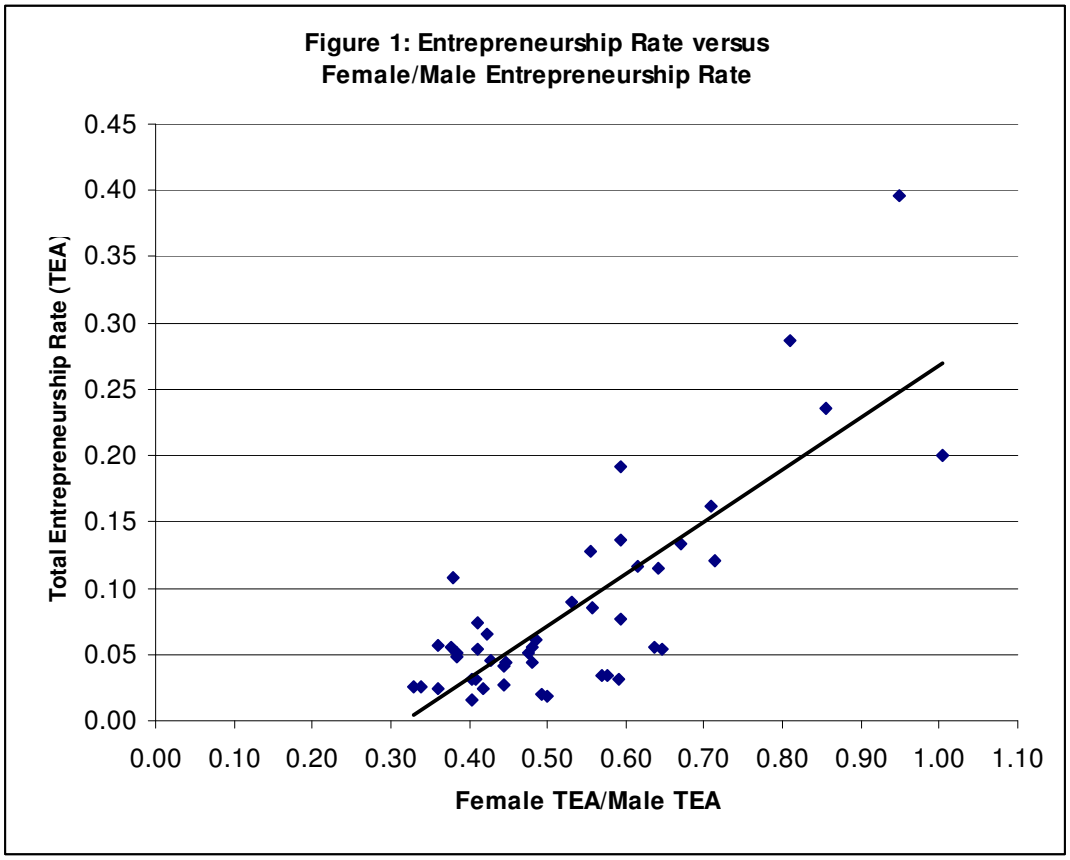
Table 4 . PROBIT RESULTS

PROBIT			
Dependent Variable:	**TEA**	**TEA OPP**	**TEA NEC**
	(1)	(2)	(3)
Age	0.0012811 (***)	0.0007071(**)	0.0004609(***)
	(0.000396)	(0.0002893)	(0.0001132)
Age Squared	-0.0000248 (***)	-0.000015(***)	-0.00000733(***)
	(0.000005)	(0.00000353)	(0.00000139)
Female.Age	0.0007036 (**)	0.0004696(*)	0.000111
	(0.0003281)	(0.0002692)	(0.0001226)
Female.Age Squared	-0.000007 (*)	-4.54e-06	-0.174000
	(0.000004)	(0.00000319)	(0.00000146)
Working	0.0076458	0.0057601	0.0011893
	(0.0060855)	(0.0038491)	(0.0018578)
Female.Working	-0.0022279	0.0003538	-0.0020516(**)
	(0.0021099)	(0.0015014)	(0.0009416)
At Home	-0.0042308	0.0011817	-0.0044772
	(0.0098782)	(0.0086039)	(0.0027565)
Female.At Home	-0.0134579	-0.0109482(*)	0.0010409
	(0.0073586)	(0.0046752)	(0.0051682)
Student	-0.0175774 (***)	-0.0108275(***)	-0.0044322(**)
	(0.004828)	(0.0029325)	(0.0017395)
Female.Student	-0.0056529	-0.0008879	-0.003522(**)
	(0.0040784)	(0.0031686)	(0.0011993)
Not Working	-0.0074277	-0.0070455(**)	0.0015583
	(0.0056668)	(0.0030436)	(0.0026615)
Female.Not Working	-0.0012184	0.0011949	-0.0024747(***)
	(0.0028817)	(0.0020342)	(0.0009146)
Low Income	0.0018601	-0.0016818	0.0032337(***)
	(0.0023676)	(0.0015527)	(0.0011427)
Female.Low Income	0.00051	0.0002493	-0.0001898
	(0.0015955)	(0.0012306)	(0.0004905)
Upper Income	0.0004073	0.0031265(**)	-0.0034215(***)
	(0.0020432)	(0.001402)	(0.0010048)
Female.Upper Income	-0.0019611	-0.0010211	-0.0004877
	(0.0014898)	(0.0010019)	(0.0008291)
Secondary Education	-0.0043625 (**)	-0.0016939	-0.001841(***)
	(0.0019644)	(0.0011623)	(0.0008013)
Female. Secondary Education	-0.002201 (*)	-0.0010864	-0.0009289
	(0.0011981)	(0.0009371)	(0.0006421)
Graduate Education	0.0042906	0.0058832(***)	-0.0030409(**)
	(0.0029192)	(0.001745)	(0.0012842)
Female.Graduate Education	-0.00491 (**)	-0.0029503(**)	-0.0007536
	(0.0019009)	(0.0011753)	(0.0008394)
Access to Network	0.0325205 (***)	0.0222001(***)	0.0051558(***)
	(0.0048195)	(0.0031912)	(0.0012469)
Female.Access to Network	-0.0004686	0.0001596	0.0002834
	(0.0011939)	(0.0007552)	(0.0005956)
Skills	0.0754828 (***)	0.0521001(***)	0.0153624(***)
	(0.006847)	(0.0043116)	(0.0027961)
Female.Skills	0.0045518 (***)	0.0028749(***)	0.0022702(**)
	(0.0018929)	(0.0011066)	(0.0011687)
Fear of Failure	-0.0124748 (***)	-0.0101741(***)	-0.000039
	(0.001688)	(0.001107)	(0.0006989)
Female.Fear of Failure	0.0037597 (***)	0.0019234(**)	0.0005197
	(0.0010835)	(0.0009526)	(0.0003831)
Female Dummy	-0.022924 (***)	-0.0187569(***)	-0.0013471
	(0.007935)	(0.006515)	(0.001992)
Dummy Year 2002	-0.0051625	-0.0033524	-0.0015987
	(0.0032687)	(0.0022346)	(0.0010325)
Dummy Year 2003	0.0024746	0.0026509	-0.0004325
	(0.004243)	(0.0024567)	(0.0016228)
Dummy Year 2004	-0.0020377	0.0001733	-0.0010156
	(0.0062819)	(0.0037226)	(0.0021955)
Number of Observations	424 566	424 566	424 566
Wald Statistics	10607.75	16451.57	8643.30
Pseudo R^2	0.1904	0.1920	0.1087
Log pseudo likelihood	-77333.3	-61098.6	-26589.2
Observed P	0.059	0.043	0.013
Predicted P	0.028	0.018	0.007

Note: Significant at 1% (***); 5% (**) and 10% (*). Standard errors adjusted for clustering on country.

Table 5 . ROBIT RESULTS - High, Middle and Low Income, Country Dummies
Note: Significant at 1% (***); 5% (**) and 10% (*). Standard errors adjusted for clustering on country. Country dummies not reported for parsimony reasons.

PROBIT			
Dependent Variable:	**TEA if** highiwb	**TEA if** Lowinwb,midloiwb,upmidiwb	**TEA With country dummy (a)**
	(1)	(2)	(3)
Age	0.0010058(***)	0.002179(**)	0.0011555(***)
	(0.0003375)	(0.0009658)	(0.0002931)
Age Squared	-0.0000178(***)	-0.0000437(***)	-0.0000206(***)
	(0.00000375)	(0.0000118)	(0.00000335)
Female.Age	0.0010298(***)	0.0004601	0.0007871(***)
	(0.0002942)	(0.0007182)	(0.000277)
Female.Age Squared	-0.0000103(***)	-0.00000385	-0.00000812(**)
	(0.00000344)	(0.00000849)	(0.00000332)
Working	0.0099294(***)	0.0278687	0.0138492(***)
	(0.0032134)	(0.0191623)	(0.0024626)
Female.Working	-0.0002121	0.0069066	-0.0005118
	(0.0012289)	(0.0067294)	(0.0014188)
At Home	0.0038664	-0.0108563	0.0063047
	(0.0099792)	(0.0259339)	(0.007768)
Female.At Home	-0.0126125(*)	-0.0244354	-0.0168033(***)
	(0.0047583)	(0.0193017)	(0.0033363)
Student	-0.0083534(**)	-0.0461429(***)	-0.011673(***)
	(0.0035513)	(0.016008)	(0.0032002)
Female.Student	-0.0033981	-0.0044515	-0.0046466
	(0.0039441)	(0.0127087)	(0.0032096)
Not Working	-0.002405	-0.0128382	-0.0007438
	(0.0039859)	(0.0180596)	(0.0033046)
Female.Not Working	0.0010477	-0.0048955	-0.0013213
	(0.0023035)	(0.006536)	(0.0018664)
Low Income	0.0024392(*)	-0.0134449	-0.0008159
	(0.001275)	(0.0090946)	(0.0016588)
Female.Low Income	-0.0018494(**)	0.0100232(**)	0.0006914
	(0.0008563)	(0.0062816)	(0.0013198)
Upper Income	0.0020413	-0.0081072	0.0024229(*)
	(0.0015904)	(0.0069403)	(0.0013375)
Female.Upper Income	-0.0019223	0.0029284	-0.0014872
	(0.0013442)	(0.0045783)	(0.0012768)
Secondary Education	-0.000431	-0.0170448(***)	0.0000294
	(0.0010276)	(0.0057334)	(0.0008396)
Female. Secondary Education	-0.0004096	-0.0021769	-0.0002917
	(0.0007992)	(0.0048668)	(0.0009337)
Graduate Education	0.0061239(***)	0.000362	0.00455(***)
	(0.0015423)	(0.012995)	(0.001507)
Female.Graduate Education	-0.0016214	-0.0146482(**)	-0.0030858(**)
	(0.0014744)	(0.0059412)	(0.0014739)
Access to Network	0.0266864(***)	0.0485451(***)	0.0294661(***)
	(0.0047317)	(0.0086378)	(0.0019402)
Female.Access to Network	-0.0006732	-0.000071	-0.0011524
	(0.0009054)	(0.0050795)	(0.0009255)
Skills	0.061087(***)	0.1182471(***)	0.0639777(***)
	(0.0064087)	(0.0115175)	(0.0026207)
Female.Skills	0.0019328(*)	0.0128501(**)	0.0023726(**)
	(0.0010596)	(0.0069365)	(0.001137)
Fear of Failure	-0.0093896(***)	-0.0220796(***)	-0.0113851(***)
	(0.0008314)	(0.0070753)	(0.0013204)
Female.Fear of Failure	0.0034975(***)	0.0040806	0.0039459(***)
	(0.0008805)	(0.0036144)	(0.0008916)
Female Dummy	-0.0335328(***)	-0.026891	-0.0256169(***)
	(0.007826)	(0.0181669)	(0.0063351)
Dummy Year 2002	-0.0045209(***)	-0.0062545	-0.0044844(*)
	(0.0016134)	(0.0139302)	(0.002344)
Dummy Year 2003	0.0010918	0.0203659	0.0035603
	(0.0025109)	(0.0210512)	(0.0030562)
Dummy Year 2004	-0.0038349	0.0266158	0.0052034(**)
	(0.0039439)	(0.020274)	(0.0027282)
N. obs	342129	82437	424566
Wald	na	na	na
Pseudo R^2	0.2026	0.1664	0.2294
Log pseudo likelihood	-51982.627	-23662.101	-73615.532
Obs. P	0.0473155	0.1089559	0.0592841
Pred P	0.0194202	0.0702947	0.0238631


Figure 1: Entrepreneurship Rate versus
Female/Male Entrepreneurship Rate
Total Entrepreneurship Rate (TEA)
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.00 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90 1.00 1.10
Female TEA/Male TEA

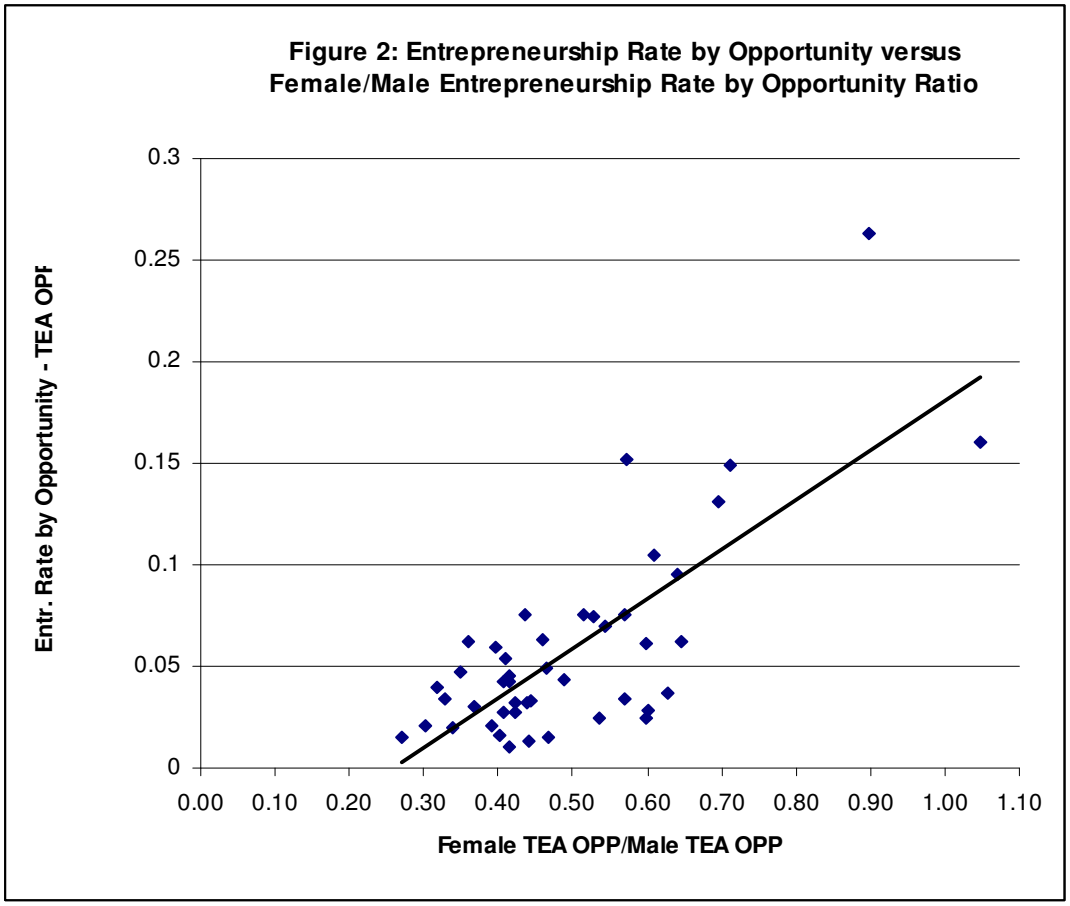
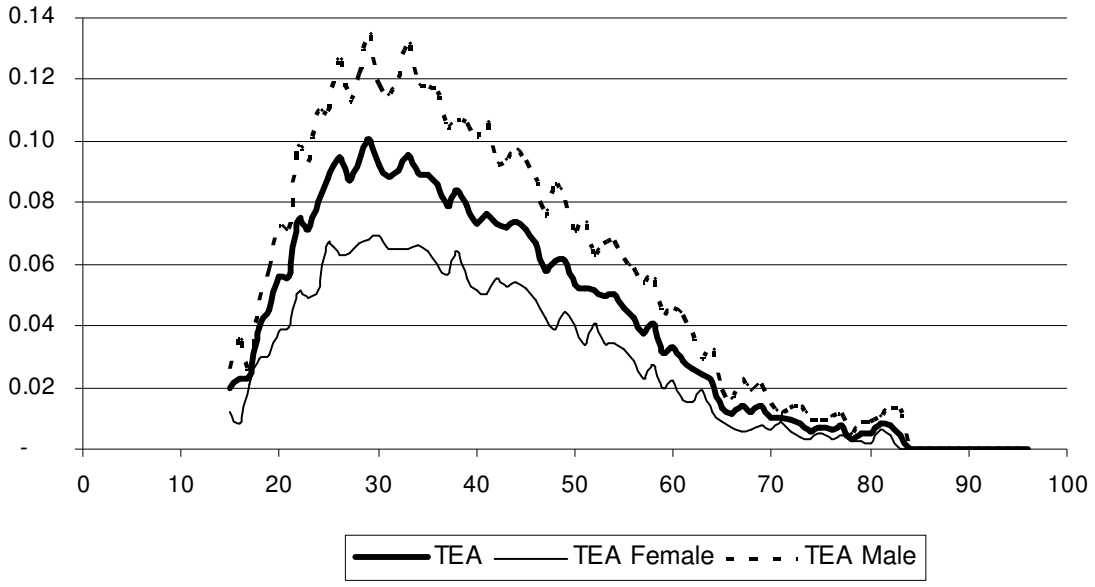
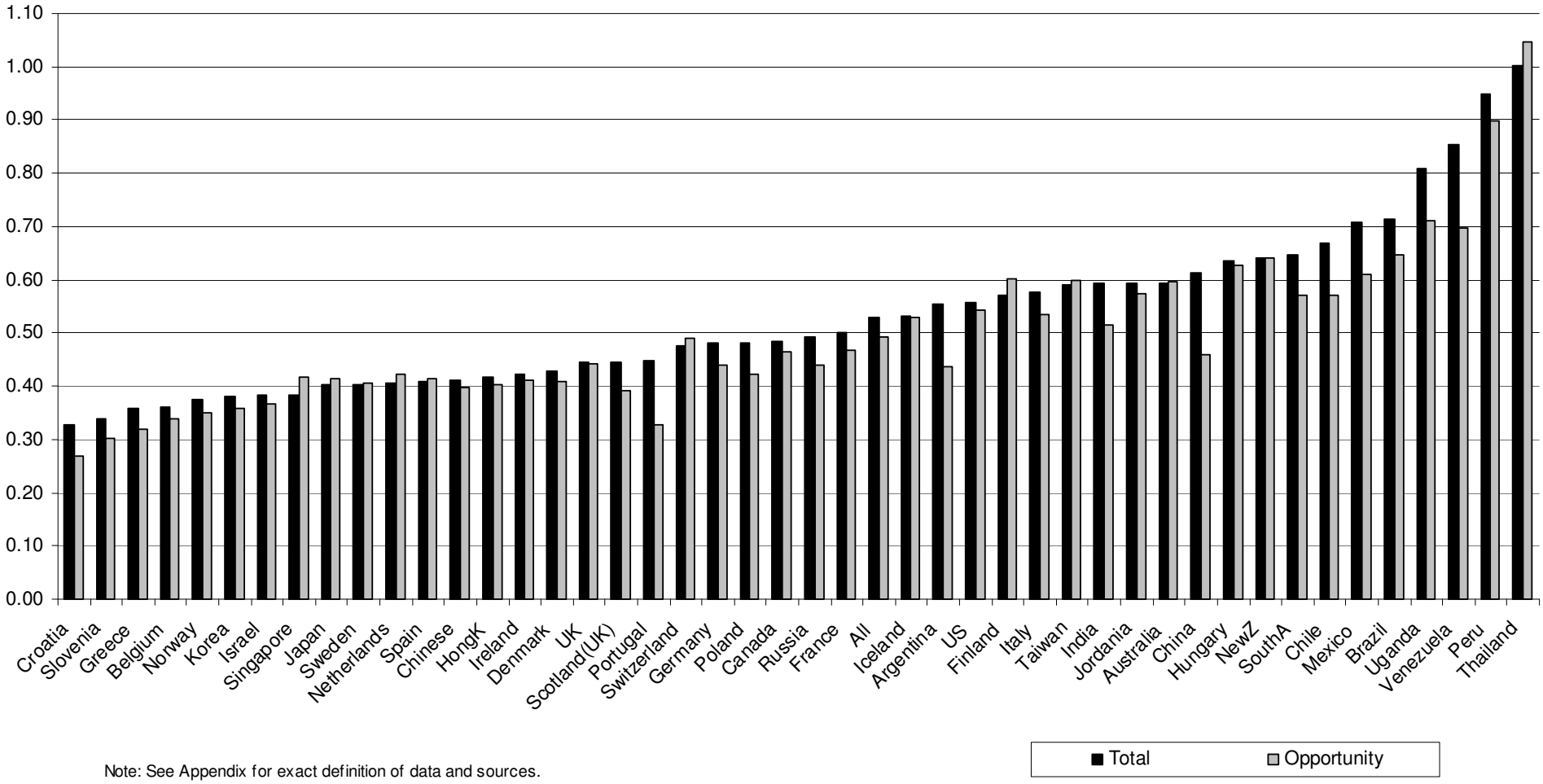

Figure 2: Entrepreneurship Rate by Opportunity versus
Female/Male Entrepreneurship Rate by Opportunity Ratio
Entr. Rate by Opportunity - TEA OPP
0.3
0.25
0.2
0.15
0.1
0.05
0
0.00 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90 1.00 1.10
Female TEA OPP/Male TEA OPP

Figure 3: Total Entrepreneurship Rate (TEA) by Sex and Age



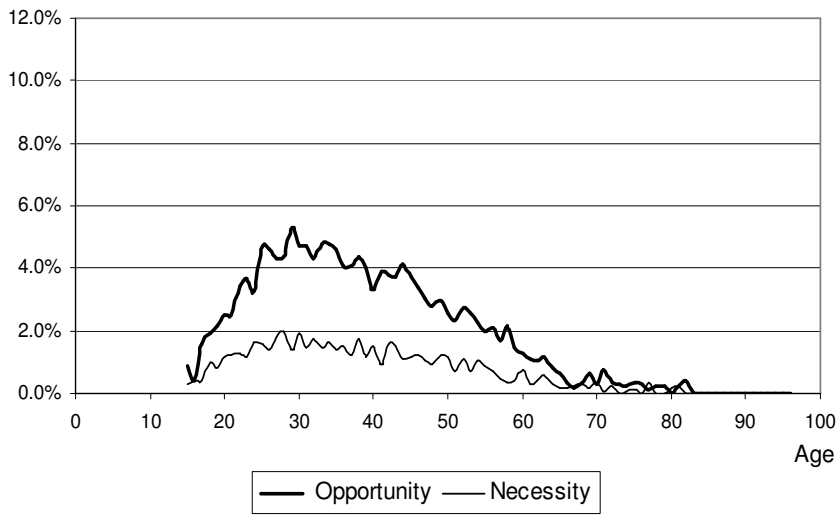
Note: Age from 15 to 96 years old. See Apendix for exact data description and source.

Figure 4: Entrepreneurship Rates - Female to Male Ratios



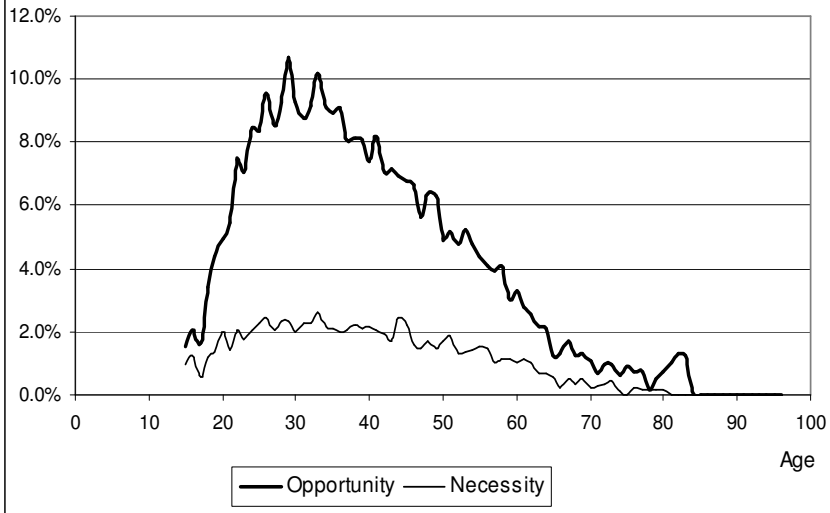
Note: See Appendix for exact definition of data and sources.

Figure 5: Entrepreneurship Rates by Age - Female



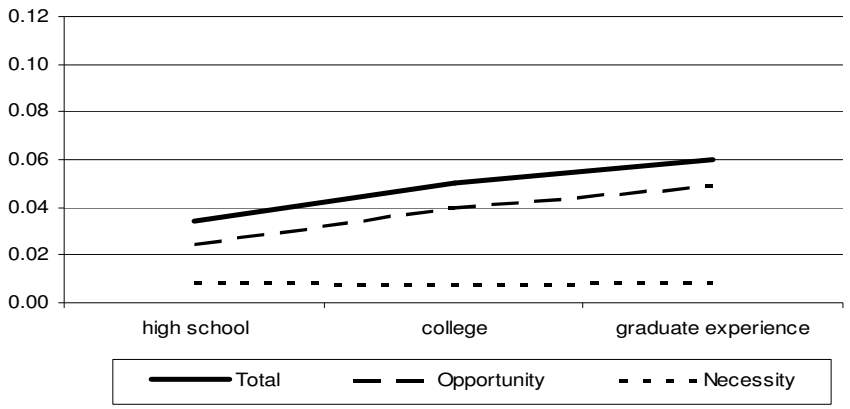
Note: Age from 15 to 96 years old. See Appendix for exact data description and source.

Figure 6: Entrepreneurship Rates by Age - Male



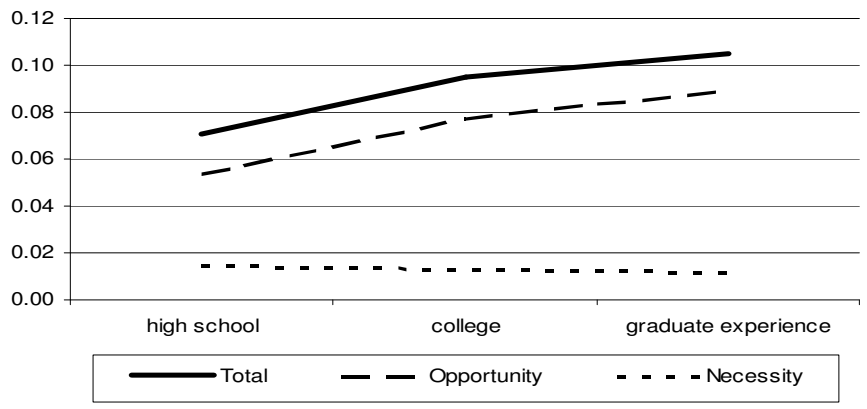
Note: Age from 15 to 96 years old. See Appendix for exact data description and source.

Figure 7: Entrepreneurship Rates by Education - Females



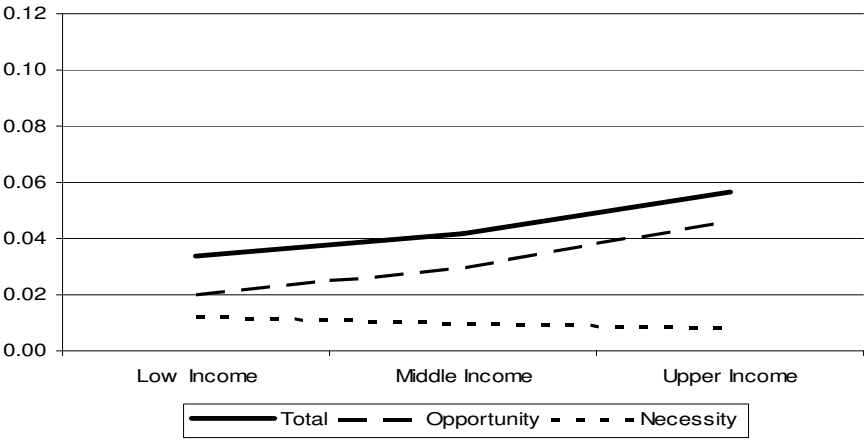
Note: See Appendix for exact data description and source.

Figure 8: Entrepreneurship Rates by Education - Males



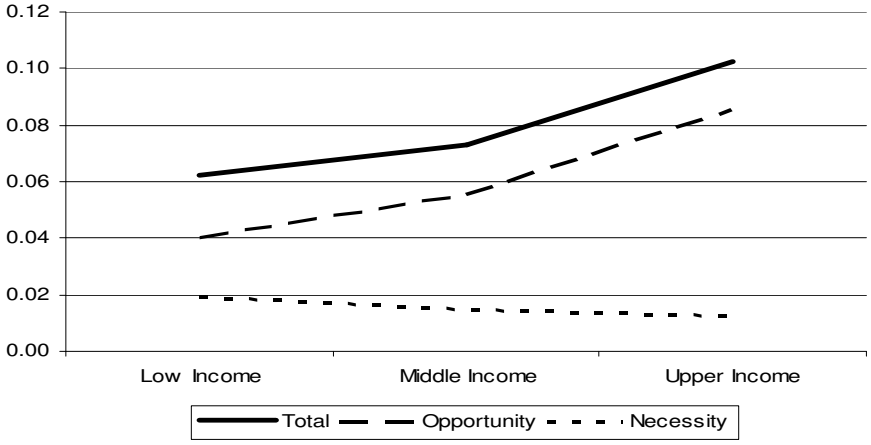
Note: See Appendix for exact data description and source.

Figure 9: Entrepreneurship Rates by Income - Female



Note: See Appendix for exact data description and source.

Figure 10: Entrepreneurship Rates by Income - Male



Note: See Appendix for exact data description and source.

AS MICRO E PEQUENAS EMPRESAS EM UM CONTEXTO DE DESENVOLVIMENTO ECONÔMICO - REALIDADE BRASILEIRA E CONFRONTO COM PORTUGAL

GÍLSON DE LIMA GARÓFALO
E-mail: songil@pucsp.br; songil@usp.br; songil1@uol.com.br
Faculdade de Economia, Administração e Contabilidade
Universidade de São Paulo

Resumo

O estudo está voltado ao papel estratégico desempenhado pelas micro e pequenas empresas, um universo heterogêneo nos setores primário, secundário e terciário de qualquer economia, além de importante fonte de geração de emprego / formação de mão de obra, embrião para grandes empreendimentos e responsável, nos diferentes países, por significativa parte dos negócios. Entretanto, essas firmas, relevantes no processo de desenvolvimento, por não terem um conceito único, acabam dificultando comparações entre as nações, inclusive no caso de Portugal e o Brasil, objetivo deste trabalho. O texto, após considerar este aspecto da definição, analisa a participação do segmento na estrutura econômica dos dois países e, na seqüência, os problemas enfrentados e, por vezes, impossibilitando que tais negócios prosperem ou tenham um período existencial maior. A crise atual, com a repercussão e os problemas que acarreta no conjunto da micro e da pequena empresa portuguesa e brasileira, precede a parte conclusiva.

Palabras Claves: Micro e Pequena Empresa, Desenvolvimento, Crise, Relações Internacionais.
Área temática: Economía Empresarial.

Abstract

The study is turned to the strategic role played by the micro and small enterprises, a heterogeneous universe in the primary, secondary and tertiary sectors of any economy and, in addition, an important source of job generation / formation of manpower, embryo for large companies and responsible, in different countries, by a significant part of business. However, these firms, relevant in the development process, for not presenting a single and only one concept, end up making difficult the comparisons between nations, even in the case of Portugal and Brazil, objective of this work. The text, after considering this aspect of the definition, analyses the contribution of the segment in the economic structure of the two countries and in sequence, the problems faced and, sometimes, making it impossible for such business to thrive or have a large existential period. The current crisis, with the repercussion and the issues that brings in the set of the Portuguese and Brazilian micro and small enterprise, precedes the conclusive part.

Key Words: Micro and Small Enterprise, Development, Crisis, International Relations.
Thematic Area: Entrepreneurial Economics.

1. INTRODUÇÃO

As micro e pequenas empresas, empreendimentos com flexibilidade, capacidade de adaptação rápida a alterações de cenário econômico e se constituindo em agentes de mudança, assumem importância ascendente no processo de desenvolvimento. Através delas, os países têm a possibilidade de atenuar o poderio abusivo geralmente exercido pelas grandes organizações no mercado. Por outro lado, em períodos conturbados, como o da crise atual, repercutindo, mais intensamente ou em menor magnitude, em todas as partes do mundo, com os desdobramentos negativos e preocupantes no mercado de trabalho, as atenções se voltam a esse segmento

credenciado para, dentro das características de que se reveste constituir alternativa para minorar o desemprego e contrabalançar a queda do nível de atividade.

Desta forma, o tema deste trabalho, ao pretender efetuar um diagnóstico das micro e pequenas empresas analisando a contribuição do segmento no processo de desenvolvimento econômico, particularmente do Brasil e de Portugal, certamente representa uma contribuição relevante não apenas à academia, mas, e esta intenção se sobressai, ao mundo dos negócios e à elaboração de políticas governamentais específicas de estímulo a criação, viabilização e preservação das mesmas. Outro propósito volta-se a possibilidade de agregar algo à, ainda hoje, limitada bibliografia sobre o assunto.

A dificuldade de uma definição uniforme e internacionalmente adotada para a micro e pequena empresa é a preocupação da primeira parte. Hoje, a questão está mais bem equacionada em Portugal que adotou a recomendação da Comissão das Comunidades Européias enquanto, na América Latina, e de forma mais particular, no âmbito dos países do Mercado Comum do Sul - MERCOSUL, ainda não se chegou a um consenso. Esta ressalva, embora possa colocar limitações, contudo não invalida o confronto da situação brasileira com a realidade lusitana, como acontecerá em partes subseqüentes do texto.

A incorporação das pequenas e médias empresas na economia é o objetivo do item seguinte. Certamente não se caracterizam por uma participação significativa na geração do produto nacional; entretanto, o setor constitui destaque, entre outras peculiaridades, pela quantidade de firmas, proporção delas relativamente à população, geração de postos de trabalho, além das condições que oferecem aos investidores com espírito empreendedor. Mas, entraves acontecem comprometendo a longevidade (ou o horizonte temporal de existência), das firmas em referência, preocupação a que está voltada a parte subseqüente do trabalho. Com efeito, se flutuações na conjuntura e transtornos no ambiente econômico podem contribuir à precoce mortalidade, obstáculos compreendidos na gestão informal, na baixa qualidade gerencial (de certa forma, conseqüência do anterior) e na escassez de recursos contribuem, e muito, na questão.

O estudo certamente não estaria completo, caso deixasse de inserir uma reflexão relativa às micro e pequenas empresas no contexto da crise atual a qual, com raiz no sistema financeiro, nascida em países avançados, atingiu todos os países de forma sincronizada. A comparação das realidades de Portugal e do Brasil é a preocupação maior nesta etapa. Os dois países, afetivamente e culturalmente irmanados, partilhando valores e interesses comuns e com grande responsabilidade frente aos demais componentes da comunidade das nações de língua portuguesa, não podem permitir que esta crise comprometa os laços existentes. A integração, com cooperação recíproca, é uma imposição do momento. Há oportunidades abertas no contexto econômico de ambos os lados que urge sejam aproveitadas. Este propósito é o destaque neste item desta contribuição.

Considerações finais enaltecendo a possibilidade de programas conjuntos de cooperação entre Portugal e o Brasil na busca do fortalecimento das micro e pequenas empresas encerram este trabalho direcionado a abordagem sobre o relevante tema da contextualização de tais empreendimentos em um processo de desenvolvimento econômico. Complementarmente, neste particular, ao comparar as realidades dos dois países lhe confere, com certeza, ineditismo.

2. ESTABELECENDO O CONCEITO

Não há consenso - internacional e nacional - quanto à conceituação de microempresas e pequenas empresas (MPEs). As contribuições neste assunto levam em conta os objetivos perseguidos (ou setores produtivos contemplados), precariedade de informações e o contexto considerado (blocos econômicos, países ou regiões destes). As definições geralmente utilizam critérios qualitativos e quantitativos recaindo em aspectos de pessoal ocupado ou de nível de faturamento.

No **Brasil**, fundamentando-se na Carta Magna promulgada em 05 de outubro de 1988, com base nos artigos 146, 170 (inciso IX) e 179 [os quais consideram que em legislação posterior (contemplada na Reforma Tributária aprovada em 2003), haveria definição de procedimento

diferenciado e favorecido às empresas nacionais de micro e pequeno porte e, também, se reportando a simplificação do enquadramento jurídico destas nas obrigações administrativas, tributárias, previdenciárias e creditícias], são encontradas definições na Lei Complementar 123, de 14/12/06 (Diário Oficial da União, 2009). Este diploma legal, instituindo o ***Estatuto da Microempresa e da Empresa de Pequeno Porte*** no País[1], estabelece a definição destas segundo o faturamento[2] de modo que todo empresário (pessoa que exerce profissionalmente atividade econômica organizada para a produção ou a circulação de bens ou de serviços) e pessoa jurídica (ou a ela equiparada) poderão ser enquadrados em uma dessas alternativas, conforme tabela adiante apresentada.

ESTATUTO DA MICROEMPRESA E EMPRESA DE PEQUENO PORTE	
MICROEMPRESA	**PEQUENA EMPRESA**
▪ Faturamento Bruto Anual até R$ 240.000,00	▪ Faturamento Bruto Anual de R$ 240.000,01 a R$ 2.400.000,00 inclusive

O diploma legal em referência também enumera as normas gerais relativas ao tratamento tributário nacional diferenciado e favorecido a serem dispensadas as MPEs no âmbito da União, dos Estados, do Distrito Federal e dos Municípios e, igualmente, especifica que o enquadramento como empreendimento de um desses portes **não** será possível quando:

- Houver a participação no capital de outra pessoa jurídica ou de empreendedor (necessariamente residente no país) que seja sócio de empresa com fins lucrativos onde detenha mais de 10% do capital social, ou na qual exerça cargo de administração ou equiparado.
- Constituir filial, sucursal, agência ou representação de pessoa jurídica com sede no exterior ou, que participe do capital de outra pessoa jurídica ou, ainda, seja constituída sob a forma de sociedades por ações.
- Estar organizada sob a forma de cooperativas (salvo as de consumo) ou que exerça atividade de banco comercial, de investimentos e de desenvolvimento, de caixa econômica, de sociedade de crédito, financiamento e investimento ou de crédito imobiliário, de corretora ou de distribuidora de títulos, valores mobiliários e câmbio, de empresa de arrendamento mercantil, de seguros privados e de capitalização ou de previdência complementar.
- Ser resultante ou remanescente de cisão ou qualquer outra forma de desmembramento de pessoa jurídica que tenha ocorrido em um dos 5 (cinco) anos-calendário anteriores.

O ***Banco Nacional de Desenvolvimento Econômico e Social - BNDES***, instituição governamental brasileira de fomento, na classificação para o porte da empresa aplicável à indústria, comércio e serviços, adota valores substancialmente elevados. Dessa forma:

BNDES - BANCO NACIONAL DE DESENVOLVIMENTO ECONÔMICO E SOCIAL	
MICROEMPRESA	**PEQUENA EMPRESA**
▪ Receita Operacional Bruta Anual ou Anualizada até R$ 1.200.000,00	▪ Receita Operacional Bruta Anual ou Anualizada superior a R$ 1.200.000,00 e inferior ou igual a R$ 10.500.000,00

A ***Fundação Instituto Brasileiro de Geografia e Estatística – FIBGE***, ao conceituar as MPEs e justificando como sendo o critério que melhor se adequa às pesquisas econômicas que realiza, recorre à definição do Statistical Office of the European Communities - EUROSTAT. Conseqüentemente adota faixas de número de trabalhadores e de proprietários das empresas como forma de se contar com informações acerca da quantidade de micro unidades empresariais

[1] O primeiro Estatuto da Microempresa e da Empresa de Pequeno Porte surgiu em 1984, tendo sido modificado em 1994 e, posteriormente, em 1999. A Lei Complementar 123, de 14/12/08 teve dispositivos alterados pela Lei Complementar 128, de 19/12/08, porém sem modificar o conceito de MPEs.

[2] Para efeito de referência, as taxas de conversão prevalecentes em 07/04/09 eram: EUR $ 1,00 = R$ 2,9818; US$ 1,00 = R$ 2,2252, ou, alternativamente, R$ 1,00 = EUR $ 0,3354 e R$ 1,00 = US$ 0,4494.

que não empregam trabalhadores, mas funcionam como relevante fator de geração de renda para os respectivos proprietários, ou:

FIBGE - FUNDAÇÃO INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA	
MICROEMPRESA	**PEQUENA EMPRESA**
▪ 0 a 09 Pessoas Ocupadas	▪ 10 a 49 Pessoas Ocupadas

No caso brasileiro é importante acrescentar a conceituação do ***Serviço Brasileiro de Apoio às Micro e Pequenas Empresas – SEBRAE*** (2007), entidade criada em 1972 objetivando promover a competitividade e o desenvolvimento sustentável dos empreendimentos de micro e pequeno porte. Esta fonte adapta os informes da FIBGE e quando os utiliza como fontes nos estudos e levantamentos que elabora, recorre ao conceito de pessoas ocupadas nas firmas, diferenciando entre indústria e comércio/serviços (tabela abaixo).

SEBRAE - SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS	
MICROEMPRESA	**PEQUENA EMPRESA**
Indústria	**Indústria**
▪ Até 19 Pessoas Ocupadas	▪ Entre 20 e 99 Pessoas Ocupadas
Comércio e Serviços	**Comércio e Serviços**
▪ Até 09 Pessoas Ocupadas	▪ Entre 10 e 49 Pessoas Ocupadas

O SEBRAE defende o uso do critério fundamentado no número de pessoas ocupadas pela vantagem dessa informação ser mais facilmente obtida e estar menos sujeita a restrições derivadas de sigilo comercial ou estatístico e, no Brasil, a única informação quantitativa disponibilizada para todo o universo de pessoas jurídicas ativas no país.

No âmbito do ***Mercado Comum do Sul - MERCOSUL*** (1998), para efeito de exportações, a Resolução Grupo Mercado Comum (GMC) MERCOSUL/GMC/RES 59, de 07/12/98, contempla como parâmetros de definição das MPEs o pessoal empregado e o nível de faturamento. Entretanto, para fins classificatórios prevalecerá o último, enquanto o primeiro critério servirá apenas como referencial. Ademais, essa Resolução distingue entre Indústria e Comércio/Serviços, estabelecendo uma restrição na conceituação, isto é, de que as MPEs não deverão estar controladas por outra empresa ou pertencer a grupo econômico que, no conjunto, supere os valores estabelecidos. Certamente os valores monetários expressos estão de certa forma consentâneos com o segmento em que essas firmas exercerão atividades, isto é, no intercâmbio comercial entre países.

MERCOSUL - MERCADO COMUM DO SUL	
MICROEMPRESA	**PEQUENA EMPRESA**
Indústria	**Indústria**
▪ Vendas Anuais até US$ 400.000,00	▪ Vendas Anuais até US$ 400.001,00 e US$ 3.500.000,00, inclusive
▪ Até 10 Pessoas Ocupadas	▪ Entre 11 e 40 Pessoas Ocupadas
Comércio e Serviços	**Comércio e Serviços**
▪ Vendas Anuais até US$ 200.000,00	▪ Vendas Anuais até US$ 200.001,00 e US$ 1.500.000,00, inclusive
▪ Até 05 Pessoas Ocupadas	▪ Entre 06 e 30 Pessoas Ocupadas

Por outro lado, em **Portugal**, de acordo com a Recomendação da ***Comissão das Comunidades Européias*** 2003/361/CE, de 06/05/03 (2003), publicada em 20/05/03 e o Decreto-lei 372/2007, de 06/11/07 (República Portuguesa, 2007), tem-se, desde 01/01/05 a classificação de MPE levando em conta o número de trabalhadores (efectivos) - considerado o

critério mais relevante - ou no faturamento, este baseado no volume de negócios (faturamento) ou no balanço total (patrimônio global).

PORTUGAL / COMISSÃO DAS COMUNIDADES EUROPÉIAS	
MICROEMPRESA	**PEQUENA EMPRESA**
▪ Entre 0 e 09 Trabalhadores (Efectivos) ▪ Faturamento ou Patrimônio Global até EUR $ 2.000.000,00	▪ Entre 10 e 49 Trabalhadores (Efectivos) ▪ Faturamento ou Patrimônio Global superior a EUR $ 2.000.000,00 até EUR $ 10.000.000,00

Cabe ressaltar que, caso o número de trabalhadores (efectivos) seja igual ou superior a 50, porém inferior a 250, o empreendimento será classificado como média empresa (volume de negócios no intervalo entre 10 milhões e 50 milhões de euros e balanço total no intervalo de 10 milhões e 43 milhões de euros).

Considerando ainda a situação em Portugal, a exemplo do que estabelece o Estatuto da MPE no Brasil, e seguindo diretrizes emanadas da Comissão das Comunidades Européias, há exigências complementares para o enquadramento na categoria desses negócios. Desta maneira, não podem ser propriedade em 25% ou mais do capital ou dos direitos de votos de outra MPE ou, conjuntamente, de várias empresas que não sejam definidas nessa categoria. Existem, entretanto, exceções a serem contempladas como no caso do capital se encontrar disperso de maneira que não seja possível determinar quem o detém.

Resumindo, se é admissível - nacional e internacionalmente - a inexistência de conceituação única, talvez fosse mais apropriado, prudente e oportuno que as legislações pertinentes e os órgãos envolvidos adotassem nesta tarefa critérios considerando as peculiaridades regionais dos territórios abrangidos ou o público-alvo na destinação dos recursos subjacentes a linhas de crédito / financiamentos concebidos por meio de instituições de fomento.

3. PARTICIPAÇÃO NA ECONOMIA

As médias e grandes firmas, premidas pela globalização, promovem automação crescente de atividades, contribuindo cada vez menos na geração de emprego. Puga (2002) coloca que essa mudança do paradigma tecnológico trouxe o reconhecimento da importância da MPEs uma vez que não replicam simplesmente o comportamento das médias e grandes empresas, mas atuam como agentes de desenvolvimento. O universo amplo e heterogêneo dessas organizações representa, em princípio, fator primordial na criação de postos de trabalho em qualquer país que seja analisado e não raro significativo volume de inovações, demonstrando flexibilidade e capacidade de rápida adaptação às mudanças que ocorrem.

No passado recente, em muitos países, o desempenho das MPEs tem superado o das grandes firmas. A falta de homogeneidade de critérios para conceituá-las certamente introduz distorções no dimensionamento da participação econômica que possam ter no país comprometendo, assim, as comparações que aconteçam entre nações. Tendo em consideração estes argumentos, serão apresentadas quantificações evidenciando o papel relevante que apresentam nas estruturas produtivas do Brasil e de Portugal.

No Brasil, em 2006 existiam 6.073.056 empresas formais em atividade no comércio (46,7%), serviços (40,6%), indústria (10%) e na construção (2,7%), das quais 36,9% com empregados e as restantes 63,1% sem empregados (Serviço Brasileiro de Apoio às Micro e Pequenas Empresas, 2008). O número das MPEs no conjunto dos estabelecimentos formais com empregados somava 2.184.934 (e considerando exclusivamente as micro, estas eram 1.873.241, ou 85,7% do agrupamento e 83,6% do total das empresas formais com empregados) enquanto em 2002 totalizavam 1.860.170, ou seja, um aumento de 17,5% no espaço de 6 anos. Segundo Ferro (2005)[3], na América Latina o país é líder em número de MPEs, estando estas concentradas

[3] Não obstante diferenças conceituais, em número de MPEs, depois do Brasil, estão: México, Argentina, Colômbia, Chile, Bolívia, El Salvador e Guatemala.

principalmente na região sudeste[4] do território nacional. Historicamente, de acordo com o Cadastro Central de Empresas publicado pelo IBGE (2004), o comércio é o segmento abrigando a maioria dessas firmas, com 49,1% delas, seguido do setor de serviços, distribuição e transportes (38,2%), indústria (10,1%), construção (2,4%) e agricultura/pesca (ínfima participação de 0,2%).

Em Portugal, igualmente em 2006, o Instituto Nacional de Estatística (2008) relata a presença de 1.085.435 empresas. Neste número, as MPEs representavam 99,4% do total (ou 1.078.922 firmas, sendo 95,4% micro e 4% pequenas), padrão próximo ao da União Européia onde ocupa a oitava posição no ranking das firmas desse porte (Eurostat, 2008)[5]. De acordo com o Instituto de Apoio às Pequenas e Médias Empresas e à Inovação - IAPMEI (2008) constituem a maioria no conjunto empresarial de todas as regiões do país, inclusive nas regiões autônomas da Madeira e dos Açores, porém, adquirindo especial relevância no Alentejo, Algarve e Centro[6]. A taxa média de crescimento entre 2000 a 2005 foi de 7,2% a.a. com o aparecimento em torno de 17,1 mil firmas desse porte a cada ano desse período. Quanto à distribuição por segmentos econômicos, esta, diferentemente do Brasil, privilegiava o setor de serviços, distribuição e transportes (51,1%), seguido pelo comércio (27,6%), construção (11,2%), indústria (9,4%) e agricultura/pesca (0,7%).

Contrasta sobremaneira a comparação do coeficiente no número de empresas por 1.000 habitantes entre os dois países: 232,9 em Portugal contra 11,4 no Brasil. Certamente isto se deve a diferença de dimensão territorial existente e as desigualdades abruptas na densidade demográfica (113 hab/km^2 e 22,5 hab./km^2, respectivamente)

O SEBRAE (2008), fundamentado em informes do Ministério do Trabalho e do Emprego através da Relação Anual de Informações Sociais (RAIS/MTE), estima que as PMEs sejam as responsáveis pela geração de 51% da totalidade dos trabalhadores formalmente empregados no Brasil, ou seja, cerca de 13.250 mil pessoas no ano de 2006 correspondendo a aproximadamente 19% da população economicamente ativa nos setores secundário e terciário da economia nessa data. Portugal apresenta uma situação diferenciada. De acordo com o Instituto Nacional de Estatística (2008), as PMEs tinham ocupadas 66,2% de pessoas em um universo 3.739 mil trabalhadores, ou 2.475 mil trabalhadores (47% da população economicamente ativa) e, portanto, bem mais da metade de todo o conjunto empresarial.

Seja como for, quando se leva em conta a redução do número de empregos causada pelas mudanças tecnológicas e dos processos de trabalho nas grandes empresas, fica explícita a tendência do segmento das MPEs tornar-se cada vez mais representativo para a geração de trabalho. No Brasil e em Portugal, fundamentando-se em dados de 2006, nos micro e pequenos empreendimentos, existiam 6,1 e 2,3 trabalhadores por firma respectivamente, ou seja, uma média compatível com o porte delas.

Outra evidência da importância das PMEs nos contextos brasileiro e lusitano é fornecida pelo volume de negócios que geram. Em Portugal, o INE (2008) relata para 2006, que no valor total deste abrangendo as empresas nacionais, ou 331.632 mil euros, as PMEs foram responsáveis por quase a metade, isto é 49,1% ou em torno de 21% do PIB nacional estimado nesse ano em 155.278 milhões de euros. No Brasil a informação neste particular é diversa e defasada com o SEBRAE dando conta de que no ano de 2000, elas respondiam por 28% do faturamento do total das empresas do país ou, então, que participavam com 20% do PIB nacional, este estimado em 2007 em US$ 1,310 trilhão. Teixeira e Pereira (2001) apontam que, em termos de faturamento e produção, enquanto as grandes organizações participam com 56,4%, os micros, pequenos e médios empreendimentos totalizam 43,6%. Seja como for, neste caso, nos dois países, as MPEs tem participações próximas, porém modestas, na totalização do produto nacional. Não obstante, compensam este fato por se caracterizarem e se revestirem, ao menos em princípio, pelo caráter inovador e pela flexibilidade em adaptação rápida às flutuações de demanda.

[4] A região sudeste abrange os estados do Espírito Santo, Minas Gerais, Rio de Janeiro e São Paulo.

[5] Estão à frente de Portugal, pela ordem, Itália, Espanha, França, Alemanha, Reino Unido, Polônia e Checoslováquia.

[6] A região Centro corresponde aproximadamente à antiga região da Beira, compreendendo integralmente, os distritos de Coimbra, Castelo Branco e Leiria, a maior parte dos distritos de Aveiro e Guarda, e cerca de um terço do Distrito de Santarém.

Algumas outras informações estão em ordem. A primeira diz respeito ao uso de equipamentos das chamadas tecnologias de informação e comunicação (TCIs) nas MPEs, com destaque para computadores, internet e celular/smartphone (celular de acesso a internet). Ao contrário do desejável, casos existem nesses empreendimentos (e são de natureza cultural), consubstanciados na resistência de utilizar essas ferramentas como instrumento de gestão. Independentemente disto, o SEBRAE (2008) e o INE/UMIC - Agência para a Sociedade do Conhecimento (2009) apresentam dados bastante interessantes sobre este aspecto. Assim, no Brasil, 75% dessas firmas (64% das micro e 95% das pequenas em Portugal) possuem o primeiro desses equipamentos, contra 23% ao início deste século XXI; 71% (12% no ano de 2000) recorrem à internet na viabilização e agilização dos negócios (55% das micro e 92% das pequenas em Portugal) e 91% contam com o recurso da telefonia móvel (em Portugal praticamente a totalidade das MPEs o faz). Pesquisa realizada pela Microsoft, em dez grandes mercados de atuação dessa multinacional, informa que, atualmente, as MPEs brasileiras são rápidas em adotar novidades tecnológicas e o ritmo de atualização no país supera, inclusive, o dos Estados Unidos, sendo que o percentual das que possuem página corporativa na web (54%) ultrapassa também o do território norte americano, Reino Unido, Japão e China (Aragão, 2009).

O descrito no parágrafo precedente certamente explica a participação das MPEs no comércio exterior brasileiro contribuindo, embora modestamente, para elevar a competitividade do país no cenário internacional e criando, no segmento, demanda por mão de obra qualificada. O SEBRAE (2007) reporta para 2006, que a participação dessas empresas no valor das vendas brasileiras ao exterior (as quais alcançaram nesse ano US$ 1376.470 milhões FOB), foi de 1,4%, porém no número total de empresas exportadoras representaram 65% (em um universo de 19.956 firmas), com o valor médio exportado por firma crescendo com o passar dos anos (US$ 275 mil em 2006). Porém, outras fontes estimam (O Estado de São Paulo, 2004) que tais empresas apresentam grande potencial e possam contribuir com até 20%, uma vez que nos países desenvolvidos as exportações por elas efetuadas atingiriam 70%. O Ministro do Desenvolvimento, Indústria e Comércio Exterior do Brasil Miguel Jorge (2009) alimenta o ideal de até o próximo ano, ou seja, 2010, elevar a base das empresas exportadoras e, nesse esforço, aumentar em 10% o número das MPEs locais inseridas no mercado internacional.

4. PROBLEMAS DEFRONTADOS PELAS MPEs

As MPEs apresentam características diferentes entre os países, com esses atributos seguramente condicionando-se aos estágios de desenvolvimento que as nações apresentam. Normalmente, estão inseridas em três grandes campos: gestão informal, baixa qualidade gerencial (um desdobramento da gestão informal) e escassez de recursos. Cezarino e Campomar (2009) detalham estes aspectos enfatizando, de modo particular, a situação brasileira.

Na questão "gestão informal", a não distinção entre o patrimônio pessoal e empresarial compromete a avaliação de desempenho, formulação de estratégias e análise detalhada da situação financeira das MPEs. Além disto, o alto grau de centralização do poder decisório na figura do empreendedor faz com que fiquem dependentes e impossibilitadas de uma gestão autônoma. A presença de um número relativamente significante de funcionários e sócios com laços familiares também dificulta a capacidade de racionalização de cargos, funções, salários e responsabilidades podendo até interferir nos comandos hierárquicos. A administração acaba acontecendo de forma precária com muitas delas, até por questão de sobrevivência, optando pela informalidade e, desta forma, prejudicando as legalmente constituídas.

A "reduzida qualidade gerencial" é refletida na ausência de informações sobre processos, controles, registros contábeis e financeiros pouco adequados no mais das vezes não se distinguindo a pessoa física da pessoa jurídica; desconhecimento do mercado, incapacidade de construção de uma estratégia competitiva proporcionando o surgimento de relação de complementaridade e subordinação com empresas médias e de grande porte. Há dificuldade de

identificação dos custos fixos, de tomada de decisões com avaliação de riscos e de formação de preços incorporando as margens de lucro desejadas e/ou possíveis e que sejam competitivos. Complementarmente, outros entraves também aqui se fazem presentes como a contratação simplista/direta de mão-de-obra e/ou utilização desta sem as qualificações desejáveis (no fundo, as MPEs são, no geral, até mesmo formadoras e capacitadoras de mão de obra para as congêneres maiores), processo de terceirização pobre, não administração de estoque conforme o poder de compra da clientela, elevadas perdas decorrentes de furtos internos e externos (os equipamentos eletrônicos para controle e fiscalização apresentam preços exorbitantes) e, eventualmente, o desenvolvimento inadequado (e, não raro problemático) de tecnologias de informações e comunicações. No caso daquelas voltadas ao comércio exterior deve ser incorporada no item a questão da complexidade dos procedimentos e os elevados custos inerentes.

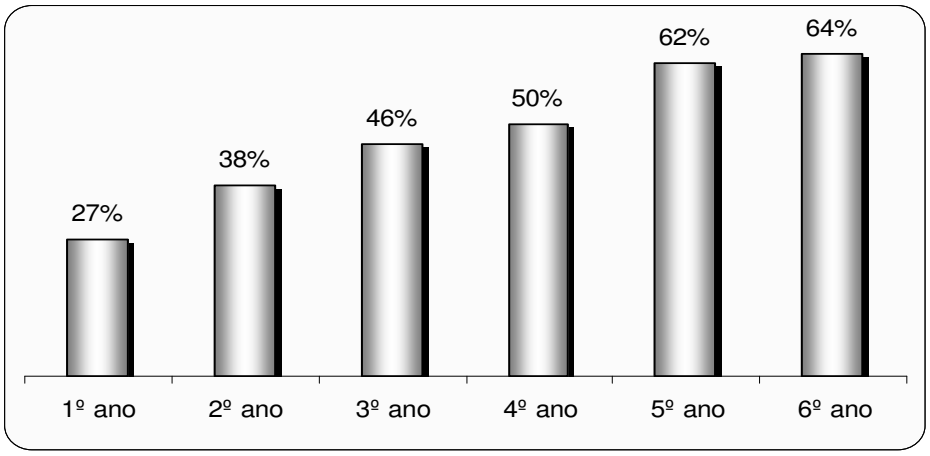
A terceira característica diz respeito à "escassez de recursos". As MPEs caracterizam-se pelo baixo volume de capital empregado (embora com elevado comprometimento de recursos próprios), por questões relativas à administração do fluxo de caixa, dificuldades de análise de viabilidade econômica-financeira para captação de recursos para investimentos. Defrontam, também, obstáculos no acesso a financiamentos, tanto públicos como privados, principalmente para capital de giro e às linhas de crédito que visem promover a atividade exportadora. Neste particular da exigüidade de recursos, há nítida diferenciação entre a maioria dos países emergentes e os países desenvolvidos nos quais, mesmo com certo grau de baixa qualidade gerencial e gestão informal, os micro e pequenos negócios normalmente contam com financiamentos e créditos despidos de restrições descabidas ao porte das mesmas, desburocratizado e com encargos (taxas de juros, comissões e outras cobranças) reduzidos, evitando que o empresário opte por contratar empréstimos como pessoa física em detrimento de fazê-lo na condição de pessoa jurídica, para a qual as taxas cobradas são usualmente menores.

Levando em conta exposição efetuada em aula magna na Universidade Clássica de Lisboa, Morais (2007)[7] enumera os principais problemas que condicionam a existência das MPEs em Portugal e que podem ser inseridos nos campos das características acima enumeradas e descritas. Compreendem a formação empresarial e a própria maneira como são constituídas essas firmas com excessiva centralização na tomada de decisões e estratégias de conduta, os elevados custos de produção e a falta de competitividade com atrelamento de atividades às médias e grandes organizações, a onerosa o obsoleta legislação trabalhista que deixa bastante a desejar, a elevada carga tributária e fiscal a que estão sujeitas obrigando-as a contratar profissional competente para que possam atender às exigências legais (sendo este um incentivo à informalidade), a falta de institutos e órgãos governamentais de apoio ao investimento e a inovação, a carência de linhas de financiamento específicas e de crédito mais acessível e, por último, mas não menos importante, a burocracia com os custos correlatos que incorpora.

Desta forma, no universo das MPEs o horizonte de sobrevivência é, lamentavelmente, uma estatística negativa. No Brasil, a exemplo de Portugal, apesar da importância do setor e de algumas conquistas no decorrer dos anos no sentido de preservá-las, estas benesses, em princípio, ainda são poucas, dada a fragilidade dessas empresas e a suscetibilidade que apresentam às oscilações de mercado e da economia acarretando, conseqüentemente, um ciclo de vida menor. Prova direta dessa realidade é a elevada taxa de mortalidade. Dados concernentes à realidade brasileira mostram que 27% das MPEs encerram as atividades no primeiro ano de constituição; 38% sobrevivem até o segundo ano; 46% fecham antes do terceiro ano; 50% não concluem o quarto ano; 62% não ultrapassam o quinto ano; e 64% desaparecem antes de completar seis anos. Na realidade, como o gráfico adiante permite inferir, apenas quatro empresas, entre dez, conseguem sucesso depois de decorridos cinco ou seis anos de atividade.

[7] Fernando Augusto Morais é atualmente o presidente da Associação Nacional de Pequenas e Médias Empresas de Portugal - ANPME e Vice Presidente da Confederação Européias de Pequenas e Médias Empresas.

Gráfico 1 - Brasil: Taxa de Mortalidade das Micro e Pequenas Empresas



Fonte dos dados brutos: SEBRAE

Certamente, além das flutuações na conjuntura e de transtornos no ambiente econômico e das características acima apontadas, fatores diversos, como a falta de planejamento prévio (verificação de mercado, isto é, da existência de clientela na área prevista para atuação e avaliação de potenciais concorrentes na região escolhida), o comportamento pouco desenvolvido do empreendedor[8], a gestão deficiente do negócio (inexistência de plano de atuação detalhando inclusive custos), a insuficiência de políticas de apoio, as imposições diversas como as impostas pelos órgãos encarregados de vigilância sanitária, a inexata idéia do porte / despesas de reformas requeridas para adequar o local de trabalho, as dificuldades ou carência de idealismo para inovar (empreendimentos inovadores provavelmente levam ao êxito, pois são normalmente bem recebidos pela população), além de problemas pessoais e/ou com a escolha de sócios que nada agregam em valor ao negócio, colaboram para o horizonte curto de sobrevivência das MPEs. Ademais, podem chegar ao ponto de terem um único e grande cliente (ou um conjunto de clientes atuando de forma conjunta) com todos os malefícios inerentes ao regime mercadológico do monopsônio.

Desta maneira, neste conjunto de problemas provocando elevada taxa de mortalidade para as MPEs, são requeridas ações pontuais, algumas imediatas e outras de prazo mais dilatado, da parte das autoridades governamentais e através de organismos próprios, objetivando melhorar as condições e/ou permitir que sobrevivam através dos anos. Com efeito, por terem personalidade própria, podem assim permanecer durante toda a existência sem deixar de serem eficientes e viáveis econômica e financeiramente. Conseqüentemente, a idéia de que sejam o embrião para uma média e grande empresa pode não se concretizar (Garófalo, 2009). Em outras palavras, considerando as peculiaridades dos diferentes países, especialmente as de Portugal e do Brasil que contariam com as intervenções do IAPMEI e do SEBRAE respectivamente, medidas de alivio fiscal e tributário (subsídios à abertura, períodos de carência relativamente a tributos que possam ser devidos, redução de alcunhas fiscais) e ações direcionadas ao crédito (criação de seguro de crédito, fundos de aval ou garantidor de empréstimos, e linhas de financiamento fundamentadas nos recebíveis de cartões e taxas de desconto nas operações nestes

[8] A Global Entrepreneurship Monitor - GEM analisa a relação entre o nível de empreendedorismo e o patamar de crescimento econômico e, simultaneamente, determina as condições que fomentam e/ou entravam as dinâmicas empreendedoras nos países componentes do universo do levantamento. O Brasil (GEM, 2008) está na 13ª. posição entre os 43 países mais empreendedores do mundo e no 3º. posto entre os integrantes do G-20 (este formado pelos ministros de Finanças e presidentes de Bancos Centrais da União Européia e de outros 19 países: África do Sul, Alemanha, Arábia Saudita, Argentina, Austrália, Brasil, Canadá, China, Coréia do Sul, França, Índia, Indonésia, Itália, Japão, México, Rússia, Turquia, Reino Unido e Estados Unidos) que representam cerca de 90% do Produto Interno Bruto mundial e 80% do comércio internacional (incluindo o comércio interno da UE), e 81,24% do total de pessoas empreendedoras no mundo. Porém convém lembrar que o empreendedorismo exige visão e conduta disciplinada e muitas das MPEs no Brasil são montadas amadoristicamente por desempregados ou por pessoas interessadas em aumentar a renda pessoal. Portugal embora não conste explicitamente da listagem em apreço sabe-se, por levantamento relativo ao ano anterior, ter sido o melhor classificado dentre os países da União Européia.

fundamentadas), deveriam ser contempladas e implantadas. Algo semelhante deveria acontecer, após devidamente analisados, quanto aos estímulos objetivando à modernização da gestão empresarial e simplificação da burocracia para abertura, manutenção e fechamento dos empreendimentos sob consideração.

5. AS MPEs E A CRISE ATUAL

Como anteriormente salientado, as MPEs dominam os empreendimentos e o universo dos negócios lusitanos e brasileiros. Foram igualmente apontados fatores que devem ser equacionados ou colocados em prática no sentido de evitar a elevada mortalidade dessas firmas e, inclusive, permitindo sejam administradas de forma mais eficiente. Entretanto, hoje o mundo capitalista enfrenta uma crise sem precedentes. Esta é entendida como a fase mais aflitiva e perigosa de passagem de uma época de prosperidade a outra de depressão. Com conotação mais grave do que as que a precederam, em especial a de 1929, teve igualmente origem no sistema financeiro (embora, naquela ocasião, o contexto fosse outro), nasceu nos países avançados e atingiu o mundo de forma sincronizada. O tamanho e a data do término constituem exercícios de adivinhação. Projeções otimistas dão conta de recuperação, se efetivamente acontecer, apenas ao final deste ano de 2009 ou no começo de 2010. Os desdobramentos em curso demonstram impactos sobre a demanda global e o nível de emprego, com os países experimentando forte retração da demanda e o protecionismo em alta.

No contra-ciclo, a perspectiva no Brasil, relativamente bem posicionado comparativamente aos demais países, é a de que venha superar rapidamente os entraves que afloram. As explicações para tanto estão alicerçadas na menor dependência e no baixo nível de exposição financeira que apresenta, complementadas pelo ajuste cambial que teve reduzido impacto sobre a inflação e a flexibilidade existente no mercado de trabalho. Entretanto, em Portugal (e na zona do euro) os sinais de recessão são inequívocos e ela encontraria mais dificuldades para ser contornada. Não obstante, foram oportunas as iniciativas governamentais lusitanas objetivando ao amparo da atividade econômica graças ao reforço da estabilidade financeira conjugado com a concessão de garantias e aumentos de capital propiciados pelo Estado. Também foram tomadas medidas de apoio à economia real visando estimular a procura, aumentar seletivamente o investimento público, ajudar às empresas em programas de incentivo a modernização, maior competitividade propiciando garantir o emprego, além de ações de natureza social. O sistema financeiro, como no caso brasileiro caracteristicamente sólido e competitivo, conta com supervisão organizada, ativa e profissional. Oxalá essas providências produzam rapidamente os efeitos desejados.

Mas o grande questionamento é como ficam as MPEs nessa crise global. Existem prós e contras para todos os empreendimentos, independentemente do porte e do setor em que atuem. Apostas sugerem que as firmas menores possam sofrer menos na medida em que priorizam o mercado interno dos países em que estão sediadas e demonstram maior flexibilidade de adaptação às mudanças e replanejamento contínuo do que as médias e grandes organizações. Seja como for, os desafios para esses negócios são a queda da demanda (por conta do aumento do desemprego) e, dependendo da essencialidade do produto que transacionem, podem ser mais ou menos afetadas. Setores ligados à produção de bens de consumo durável tendem a ser os mais atingidos uma vez que as pessoas podem adiar esse tipo de compra, no caso de produtos não tão requeridos ou necessários (roupas e calçados) os reflexos são menores e, no caso de itens básicos (alimentos) nos quais a questão de marca e fidelização costuma acontecer, o ideal é convergir para os produtos em que essa fidelidade seja menor podendo-se mesmo recorrer à divulgação/publicidade com promoções e campanhas pró-ativas para venda da marca. Em todos os casos, são fundamentais estratégias para manutenção e fidelização da clientela, diversificando-a e atraindo novos clientes oferecendo produtos e serviços com qualidade e preços competitivos, porém tomando cuidados na concessão indiscriminada de créditos de modo a evitar prejuízos futuros (Brasil, 2009).

Ao lado da queda defrontada na procura, tem-se a elevação do risco de inadimplência e a redução (ou falta) do crédito, uma vez que os bancos se retraem e priorizam as companhias

maiores. Seja como for, a crise deve ser vista pelas MPEs como geradora de oportunidades, tornando-se um exercício de criatividade e de gestão de resultados. É uma época ideal para redimensionar planos/investimentos, avaliar gastos cortando-os sempre que possível, identificando a urgência dos mesmos e, também, para rever a situação financeira equacionando o fluxo de caixa de modo a manter um nível estável de endividamento e efetivando ajustes nos recebimentos e pagamentos adequando-os a melhores condições e prazos, particularmente quando de desembolsos maiores. Outras providências (Brasil, 2009) são evitar excesso de estoques e, adicionalmente, motivar e envolver os funcionários para que, conjuntamente, busquem uma gestão eficaz de modo a sentir menos os efeitos da turbulência, evitando demiti-los uma vez que o desligamento e treinamento de novos custa caro.

Cabe acrescentar que em tempos de crise, muitas das MPEs exercendo a criatividade estão adotando a governança corporativa, ou seja, transparência do próprio negócio com o objetivo de aprimorar a comunicação com os fornecedores, clientes e instituições financeiras. Quando buscam o sistema bancário, recorrem a práticas mais transparentes e procuram evidenciar possuírem total controle sobre as contas, as despesas, as vendas (e correspondente faturamento), além do lucro da firma, ou seja, esperam potencializar a reciprocidade de modo a, em um segundo momento, terem acesso facilitado a outras formas de recursos que não as linhas de crédito de curto prazo (capital de giro) como, por exemplo, para viabilizar investimentos e/ou para poderem continuar contando com recursos para financiar as vendas ou a prestação de serviços.

Por outro lado, nesse conjunto de reflexões, chama a atenção o resultado de uma sondagem efetuada pelo SEBRAE e divulgada ao início do mês de Abril deste ano pelo jornal O Estado de São Paulo ("Micro e pequenos empresários estão confiantes", 2009). De acordo com a notícia, no Brasil os micro e pequenos empresários estariam otimistas com o desempenho da economia em 2009 e "prevêem um ano melhor do que o anterior, mesmo com a crise financeira". Fundamentando-se nessa pesquisa, 62% dos entrevistados estavam esperançosos de vender e faturar mais neste exercício sendo que 79% têm perspectiva de que terão bons ou muito bons resultados. Adicionalmente, mais da metade (56%) pretendem manter o quadro de funcionários e 35% teriam expressado o desejo de aumentar as contratações. De certa forma, esta enquete estaria em linha com outro indicador divulgado mensalmente e que revela as expectativas favoráveis dos economistas brasileiros relativamente ao futuro[9].

Uma última observação está em ordem, a de alternativas de negócios para as MPEs. Mesmo com toda a instabilidade decorrente da crise, no Brasil, e com certeza em Portugal, campos de atuação continuam abertos principalmente nas áreas de saúde, educação, serviços pessoais, qualidade de vida, inovação e solução tecnológica, processos e meio ambiente.

6. CONSIDERAÇÕES FINAIS

O segmento das MPEs é bastante representativo tanto em países emergentes, como o Brasil, e em economias consolidadas, caso de Portugal. Exigências à constituição desses negócios são, por vezes, desestimulantes, descabidas, complexas e onerosas fazendo com que ocorra, não raro, a implantação e o crescimento de negócios informais trazendo consigo a evasão tributária, a injustiça fiscal (contribuintes pagando muito por outros que nada pagam), subempregos e prejuízos à livre concorrência, pilar para o desenvolvimento de qualquer economia.

O trabalho inicialmente realçou a ausência de uma definição única no que se refere à MPE para, em seguida, evidenciar a importância que o segmento apresenta nas economias brasileira e portuguesa. Não obstante o estágio de desenvolvimento diverso em que estas se encontram e guardadas as proporções de ambos os países, eventuais dificuldades decorrentes da maneira de divulgação dos informes pelos órgãos governamentais de estatística e por entidades outras voltadas especificadamente à área (particularmente o SEBRAE e o IAPMEI) associada, em algumas circunstâncias, a não unicidade perfeita dos dados, os perfis são bastante próximos,

[9] Este indicador é o ISE - Índice de Sentimento dos Especialistas em Economia, elaborado e divulgado em parceria existente entre a Ordem dos Economistas do Brasil - OEB e a Federação do Comércio do Estado de São Paulo - Fecomercio.

com a diferenciação maior ocorrendo na maneira como as firmas se distribuem entre os diversos setores.

Em Portugal a maior concentração está no segmento de Serviços, Distribuição e Transportes, enquanto no Brasil isso acontece na área do comércio. Foi possível observar, também, em ambos os casos, que nos dois países e a exemplo do que ocorre em outros que integram a União Européia e o MERCOSUL, a importância do segmento para o desenvolvimento interno por meio da geração de empregos e divisas, ou até como regulador de mercado e, ao contrário do que poderia ser pensado, são empreendimentos abertos à inovação, geralmente adequando-se de forma positiva e rápida às alterações no mercado.

A reflexão sobre as dificuldades com que as MPEs se defrontam demonstrou estarem estas situadas em três grandes áreas: gestão informal, baixa qualidade gerencial (um desdobramento da gestão informal) e escassez de recursos. Cada uma destas mereceu detalhada análise e, na medida do possível, novamente as situações do Brasil e de Portugal foram comparadas. Certamente a Ordem dos Economistas Portugueses e a Ordem dos Economistas do Brasil, conjuntamente com o IAPMEI e o SEBRAE, pudessem promover eventos conjuntos em que os problemas existentes viessem a ser discutidos, encaminhados para um equacionamento uniforme e, posteriormente, com contribuições disseminadas pelas demais nações, especialmente para aquelas integrantes da Comunidade dos Países de Língua Portuguesa.

Por outro lado, no momento em que com maior intensidade a crise atual se irradiou entre todos os países trazendo preocupação sobre os efeitos que possa acarretar dentro do mundo empresarial, o trabalho estaria incompleto caso deixasse de correlacionar essa turbulência global com as MPEs. Não obstante não haja consenso sobre a duração da mesma e as medidas adequadas para localmente contorná-la, indicações quanto à maneira de agir desses negócios frente à mesma indicam que cada setor onde atuem demanda análise própria. Aos governos cabe a adoção de medidas para resguardá-las. Não é o caso de adotar medidas protecionistas, mas antes de eliminar obstáculos à constituição, tomar medidas na área fiscal e creditícia reduzindo carga tributária, taxas de juros e aumentando as linhas de crédito, além de promover programas de capacitação, orientação e indicar procedimentos adequados, propagando oportunidades de negócio, simplificando a burocracia, além de providências outras que possibilitem a vitalidade dessas MPEs. É relevante que prossigam existindo gerando negócios, criando empregos, e cooperando no processo de desenvolvimento econômico dos países em que se localizem. Aos empresários do setor, entre outras ações, fica reservado o exercício da criatividade e da gestão de resultados, privilegiarem a liquidez preocupando-se com o fluxo de caixa e controle do endividamento, fidelizar a clientela oferecendo serviços variados e mostrar pró-atividade e flexibilidade, com rapidez para converter oportunidades em negócios.

Finalmente, Portugal e Brasil poderão adotar programas conjuntos de engrandecimento das MPEs, a serem estendidos aos demais países lusófonos, favorecendo as respectivas economias e possivelmente contribuindo no incremento dos fluxos comerciais, financeiros e de serviços, com aumento dos postos de trabalho e, provavelmente, da qualidade de vida proporcionando inequívocas vantagens e não pensando exclusivamente em interesses locais. Há que se fortalecer economicamente o bloco das nações de língua portuguesa.

REFERÊNCIAS BIBLIOGRÁFICAS

ARAGÃO, MARIANNA (2009), "Pequena empresa é rápida em adotar novidades tecnológicas". *O Estado de São Paulo*, 17/03, p. B-12.

BRASIL, ANDRÉIA (2009), "Receitas para a Crise - Atenção, Cautela e Planejamento", *Comércio & Serviços*, Fevereiro e Março, pp. 40-43.

BRASIL. MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA E COMÉRCIO EXTERIOR (2000). "A Micro e a Pequena Empresa no Brasil: Dados Segmentados", Brasília (DF), Disponível em http://www2.desenvolvimento.gov.br/sitio/sdp/proAcao/micEmpPequeno/micEmpPequeno.php Acesso em 12/03/09.

CEZARINO, LUCIANA O. E CAMPOMAR, MARCOS C. (2006), "Micro e Pequenas Empresas: Características Estruturais e Gerenciais", *Revist@ Fafibe On Line* - Revista Eletrônica das Faculdades Integradas de Bebedouro - SP, Ano 2, nº. 2, Maio, Disponível em

http://www.fafibe.br/revistaonline/arquivos/lucianacezarino_microepequenasempresas.pdf Acesso em 04/03/09.

COMISSÃO DAS COMUNIDADES EUROPÉIAS (2003), "Recomendação 2003/361/CE, de 06/05/2003", *Jornal Oficial da União Européia*, 20 de maio, pp. 36-41.

"EMPRESAS TAMBÉM ENFRENTAM EXCLUSÃO DIGITAL" (2004), *O Estado de São Paulo - Economia,* 02/11, p. B12.

EUROSTAt (2008), "Enterprises by size class - Overview of SMEs in the EU", *Statistics in Focus 31/2008*, Gabinete de Estatísticas da União Européia, 25 de março, Disponível em http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-08-031/EN/KS-SF-08-031-EN.PDF Acesso em 08/03/09.

FERRO, RAÚL (2005), "Pequenas Maravilhas", *AméricaEconomia*, 27/09 a 10/10/05, p. 28.

GARÓFALO, GÍLSON DE LIMA (2005), "Crédito para as Micro e Pequenas Empresas", *Cadernos Fecomercio de Economia*, nº. 2, Federação do Comércio Estado de São Paulo, São Paulo, Junho.

GARÓFALO, GÍLSON DE LIMA E OUTROS (2009). "Entraves ao Desenvolvimento das Micro e Pequenas Empresas no Brasil - Diagnóstico e Perspectivas", Trabalho em fase de revisão para impressão, Fecomercio, São Paulo, Fevereiro.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (2004). "Estatísticas do Cadastro Central de Empresas 2002", Serviço Gráfico do IBGE, Rio de Janeiro.

GEM (2008), "Global Entrepreneurship Monitor 2008 Executive Report", Disponível em www.gemconsortium.org/download/1237755070404/GEM_Global_08.pdf Acesso em 10/03/09.

JORGE, MIGUEL (2009). "Desenvolvimento Produtivo para Evitar a Volta da Vulnerabilidade Externa", in Velloso, João P. R. e Albuquerque, Roberto C. (Org.), "Na Crise Global como ser o Melhor dos BRICS", Elsevier, Rio de Janeiro, pp. 35-47.

KOTESKI, MARCOS A.(2004), "As Micro e Pequenas Empresas no Contexto Econômico Brasileiro", *Revista FAE Business*, Vol. 8, nº. 1, Maio, pp. 16-18.

MERCADO COMUM DO SUL - MERCOSUL (1998), "Resolução Grupo Mercado Comum (GMC) 59, de 07/12/98", Disponível em http://www.mercosul.gov.br/normativas/lista.asp?cod_tipo_normativa=2&ano=1993 Acesso em 05/03/09.

"MICRO E PEQUENOS EMPRESÁRIOS ESTÃO CONFIANTES" (2009), *O Estado de São Paulo - Economia*, 03 de abril, p. B-2.

MORAIS, FERNANDO A.(2007), "Estrutura Empresarial - Portugal", *Aula Magna da Universidade Clássica de Lisboa, Lisboa*, 11 de dezembro, Disponível em http://www.ctoc.pt/fotos/editor2/EstruturaEmpresarial.pdf Acesso em 12/03/09.

PEREIRA, ANÍSIO C. E TEIXEIRA, SILVIO A.(2001), "Planejamento financeiro de curto prazo como ferramenta de gestão na PME", *Unopar científica: ciências jurídicas e empresariais*, Londrina, Vol. 2, nº. 2, Setembro, pp. 31-49. Disponível em www13.unopar.br/unopar/pesquisa/getArtigo.action?arquivo=00000392 Acesso em 12/03/09.

PORTUGAL. INSTITUTO NACIONAL DE ESTATÍSTICA (2008). "Empresas em Portugal 2006". Lisboa.

PORTUGAL, INE - INSTITUTO NACIONAL DE ESTATÍSTICA / UMIC - Agência para a Sociedade do Conhecimento (2009), "A Sociedade da Informação em Portugal 2008", Disponível em http://www.umic.pt/images/stories/osic/VI_Economia_2008.xls Acesso em 09/03/09.

PORTUGAL. MINISTÉRIO DA ECONOMIA E INOVAÇÃO, Instituto de Apoio às Pequenas e Médias Empresas e à Inovação (2008), "Sobre as PME em Portugal", Fevereiro, Disponível em www.iapmei.pt/resources/download/sobre_pme_2008.pdf Acesso em 09/03/09.

PUGA, FERNANDO P. (2002), "O Apoio Financeiro às Micro, Pequenas e Médias Empresas na Espanha, no Japão e no México", *BNDES – Textos para Discussão 96,* Agosto. Disponível em http://www.bndes.gov.br/conhecimento/td/Td-96.pdf Acesso em 09/03/09.

REPÚBLICA FEDERATIVA DO BRASIL (2006). "Constituição promulgada em 05/10/88", Senado Federal, Brasília (DF), Disponível em http://www.senado.gov.br/sf/legislacao/const/ Acesso em 28/02/09.

REPÚBLICA FEDERATIVA DO BRASIL (2009), "Lei Complementar 123, de 14/12/2006", *Diário Oficial da União*, Seção 1, 15/12/08, p. 01. Republicada no *Diário Oficial da União - Edição Extra,* 31/01/09.

REPÚBLICA FEDERATIVA DO BRASIL (2006), "Lei Complementar 128, de 19/12/2008", *Diário Oficial da União*, Seção 1, 22/12/2006, p. 01 e p. 65.

REPÚBLICA PORTUGUESA - MINISTÉRIO DA ECONOMIA E DA INOVAÇÃO (2007), "Decreto-lei 372/2007, de 06/11/07", *Diário da República - Electrónico,* 1ª. Série, nº. 213, 06 de novembro, pp. 8080-8084, Disponível em http://dre.pt Acesso em 02/03/09.

SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS DE SÃO PAULO - SEBRAE-SP (2008), "As Tecnologias da Informação e da Comunicação nas MPEs Brasileiras", *Observatório das MPEs,* SEBRAE-SP, São Paulo, Disponível em http://www.sebraesp.com.br/sites/default/files/informatizacao_br_2008.pdf Acesso em 12/03/09.

SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS - SEBRAE (Org.) (2008), "Anuário do Trabalho na Micro e Pequena Empresa: 2008", 2ª. edição, Dieese, Brasília (DF).

SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS (2007), "As Micro e Pequenas Empresas na Exportação Brasileira - Brasil e Estados 1998-2006", Brasília (DF).

WAHLMANN, GABRIELA C. (2004). Um Estudo Exploratório sobre a Atividade da Controladoria nas Microempresas na Cidade de Ubatuba", Faculdades Integradas Módulo, Caraguatatuba (SP). 2º. Prêmio Professor Eliseu Martins, Disponível em www.fucape.br/premio_profeliseu/arquivos/79.pdf Acesso em 08/03/09.

SÍTIOS

www.bndes.gov.br
www.ibge.gov.br
www.iapmei.pt
www.ine.pt
www.ipea.gov.br
www.presidencia.gov.br/casacivil/site/static/le.htm
www.sebrae.org.br
www.sebraesp.com.br
www.senado.gov.br

SP/ 08-04-09

Trabalho encaminhado ao Comité Organizador para ser submetido a avaliação pelo Comité Científico do XXIII Congresso/Encontro Anual da ASEPELT - Associação Europeia de Economia Aplicada - 17 a 20 de junho de 2009 - Faculdade de Ciências Sociais e Humanas – Departamento de Gestão e Economia - da Universidade da Beira Interior, Covilhã (Portugal).

ANALYSING RURAL TOURISM COMMERCIALIZATION IN EXTREMADURA, AN EMPIRICAL APPROACH IN ORDER TO OPTIMIZE BUSINESS STRATEGIES.

LIDIA ANDRADES CALDITO
E-mail: andrades@unex.es

Dpto. de Dirección de Empresas y Sociología
Facultad de Ciencias Económicas y Empresariales
Universidad de Extremadura

Resumen
El turismo rural supone una oportunidad para propiciar el desarrollo sostenible de áreas rurales, caracterizadas comúnmente por disponer de valiosos recursos naturales susceptibles de ser aprovechados con fines turísticos. En el caso de Extremadura, se ha procurado en los últimos años favorecer la creación de infraestructuras turísticas, con el fin de promover el desarrollo de un modelo turístico sostenible que impulse las por lo general frágiles economías locales. Actualmente, según fuentes oficiales, la región cuenta con 191 apartamentos rurales, 388 casas rurales y 61 hoteles rurales. El objetivo del presente trabajo es analizar y revisar la estrategia de comercialización de estos establecimientos para tratar de incrementar la ocupación media de los mismos. En esta comunicación se describirá la metodología que se está aplicando para lograr el objetivo planteado, así como un avance de los principales resultados obtenidos, puesto que la investigación que se presenta se haya en fase de ejecución.

Palabras Claves: "Turismo rural sostenible", "Comercialización", "Estrategia Empresarial".
Área temática: Economía Empresarial.

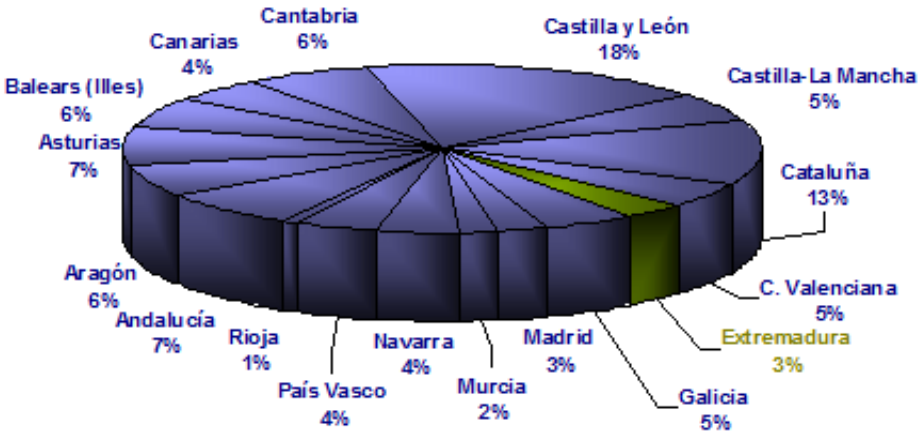
Abstract
Rural tourism represents a chance to promote sustainable development of rural areas, which are commonly characterized to have valuable natural resources susceptible to be used with tourist aims. In the case of Extremadura, it has been developed many tourist infrastructure during last years. The purpose of these investments is to promote the development of a sustainable tourist model that impels the commonly weak local economies. At the moment, according to official sources, the region counts with 191 rural apartments, 388 rural houses and 61 rural hotels. The objective of the present work is to analyze and to review the strategy of commercialization of this sector in order to increase their average occupation. This paper describes the methodology that is being applied to get our goal, as well as an advance of the main obtained results, since this research is in phase of execution.

Key Words: "Sustainable rural tourism, "Commercialization", "Touristic Business Strategy".
Thematic Area: Entrepreneurial Economics.

1. INTRODUCCIÓN Y ANTECEDENTES.

Actualmente una de las principales trabas que encuentran los pequeños y medianos empresarios turísticos extremeños, es lograr que sus establecimientos sean conocidos y elegidos por los turistas. Esta dificultad se agrava si tenemos en cuenta que como destino, gran parte de las comarcas extremeñas pueden calificarse todavía como destinos emergentes, y por lo tanto poco conocidos por los turistas. Las pernoctaciones en alojamientos rurales en Extremadura, representaron tan solo un 3% del mercado nacional en el último año, (Gráfico 1.1).

Gráfico 1.1. Participación de las diferentes regiones españolas en el Mercado de Turismo Rural Español, año 2008.



Fuente: elaboración propia a partir de los datos del IET.

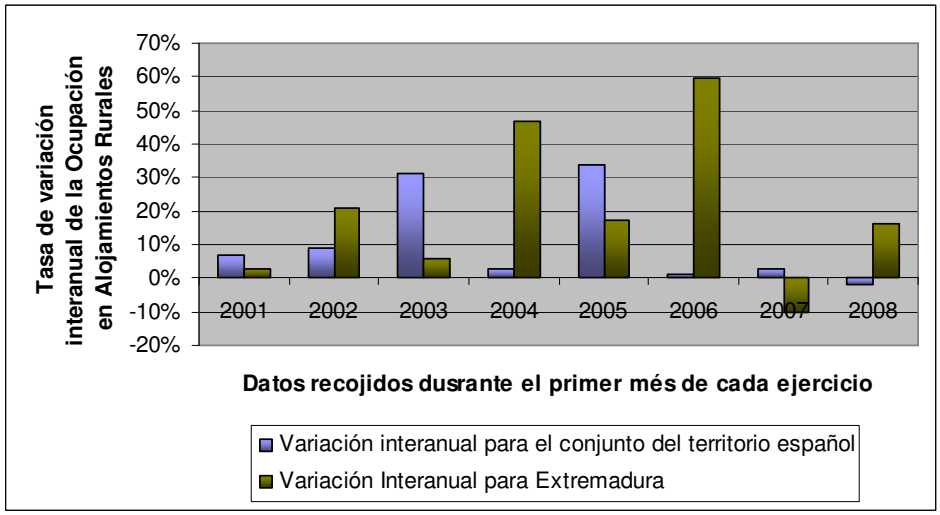
Observando los datos facilitados por el Instituto de Estudios Turísticos (IET), en la Encuesta de Ocupación en Alojamientos de Turismo Rural, (EOAT), las regiones con mayor cuota de mercado son Castilla y León, con un 18%, seguida de Cataluña, con un 13% y las de menor son La Rioja, con un 1%, Murcia con un 2%, seguidas de Extremadura y la Comunidad de Madrid, con un 3%. Por lo tanto Extremadura se sitúa a la cola de las regiones españolas en cuanto a participación en el mercado de turismo rural, entendido éste cómo un turismo alternativo, en oposición al turismo masivo, del que se contemplan una gran diversidad de vocablos que hacen referencia a distintas modalidades de turismo desarrollado en un entorno rural: ecoturismo, turismo ecológico, turismo natural, turismo verde, turismo de aventuras, turismo étnico, agroturismo, etc.

Sin embargo, a pesar de la aún incipiente participación de Extremadura en el mercado nacional, el análisis de la demanda turística, revela un interés creciente por el turismo de naturaleza en general y por el turismo sostenible en particular, entendido este último como el conjunto de modalidades turísticas que tienen por objeto ser consecuentes con los valores naturales, sociales y comunitarios, permitiendo además, tanto a los anfitriones como a los huéspedes, disfrutar de una interacción positiva que merece la pena, así como de las expectativas compartidas, *Wearing y Neil, 1999.*

Según datos oficiales disponibles, las pernoctaciones en alojamientos rurales han crecido para el conjunto del territorio español hasta el año 2008, observándose un ligero decremento en Enero de 2009, si bien estos datos son aún provisionales. Además en el caso particular de la región extremeña, la ocupación de estos establecimientos ha evolucionado favorablemente, experimentando un importante incremento hasta el año 2006. En el año 2007, la ocupación creció, si bien a un menor ritmo que en años anteriores, lo que revela una tasa de variación interanual negativa (Gráfico 1.2.).

Los cifras ponen de manifiesto que existe un interés en la demanda turística por el turismo de naturaleza, y que se trata de un segmento en alza que merece especial atención por parte de los empresarios turísticos extremeños. Si tenemos en cuenta que existe una demanda interesada en el disfrute de vacaciones en el medio natural, y a su vez que los establecimientos extremeños no son capaces de lograr atraer a una demanda suficiente, algo está fallando en el sistema.

Gráfico 1.2.: Evolución de la Ocupación en Alojamientos de Turismo Rural.

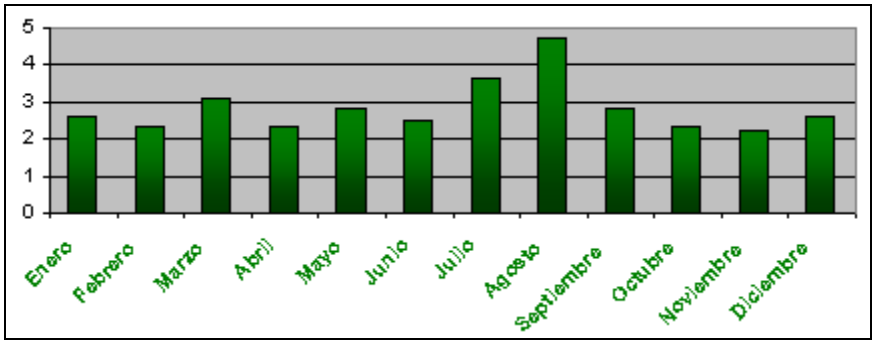


Fuente: elaboración propia a partir de los datos de IET.

Estudios realizados al respecto en el año 2007, para algunas comarcas extremeñas, *La Serena, La Siberia, y Olivenza*, revelan falta de presencia de la oferta alojativa existente en los principales canales de distribución turística regionales y nacionales[1], y especialmente en Internet. Por lo tanto, uno de los factores que podría explicar la aún baja cuota de mercado de Extremadura en el mercado de turismo rural, pueda ser la falta de presencia en los puntos de venta del producto "turismo rural en Extremadura". Ante esto, un hecho esperanzador es la reciente inauguración, en 2009, de una central de reserva online oficial para Extremadura[2].

Otro posible factor explicativo de la baja cuota de mercado podría ser la accesibilidad de la región para los principales mercados emisores de turistas. En este sentido, Extremadura se sitúa al Oeste, y teniendo en cuenta que la duración media de la estancia del turista rural se sitúa en dos días la mayor parte del año[3], (Gráfico 1.3.), es posible que su ubicación resulte excesivamente periférica para algunos turistas. Sin embargo, los turistas que vistan la región provienen fundamentalmente de Andalucía, Madrid, y Castilla León[4], comunidades altamente pobladas, lo que hace pensar que este factor no es la principal barrera a superar por la región de cara a lograr un mejor posicionamiento en el mercado turístico.

Gráfico 1.3.: Duración en días de la estancia media en establecimientos de turismo rural durante el año 2008.



Fuente: elaboración propia a partir de los datos de IET.

[1] Estudio Junta de Extremadura.
[2] http://www.extremaduraenunclick.com/

[3] Según los datos oficiales publicados por el IET, la estancia media oscila a lo largo del año, dependiendo de que se trate de meses con mayor o menor tendencia vacacional. En particular los meses de verano y Semana Santa, son los que registran estancias medias mas elevadas.
[4] Según datos facilitados por las oficinas de turismo de las principales localidades extremeñas.

En realidad podrían analizarse multitud de factores que en mayor medida pueden explicar la consolidación de un destino turístico. Otro factor importante es la competitividad turística del destino, que depende a su vez de muchas variables que van desde su atractivo intrínseco, su imagen de destino y notoriedad, su nivel de precios, la calidad proyectada al mercado, etc., hasta la capacidad de los empresarios turísticos para hacer llegar el destino al cliente potencial.

En este trabajo se aborda uno de los factores, que en nuestra opinión resulta clave a la hora de lograr una buena cuota de mercado: la capacidad empresarial, la destreza de los empresarios para comercializar sus productos. Desde nuestro punto de vista, una planificación errónea de la distribución comercial del servicio de alojamiento, conllevará menores tasas de afluencia de turistas, menores tasas de ocupación, y una menor rentabilidad que agravará el problema de la estacionalidad de la demanda propia del sector turístico. La planificación de una estrategia de comercialización eficaz, será crucial para mejorar el posicionamiento de Extremadura como destino de turismo rural.

En concreto, la presente comunicación muestra un avance de los resultados de una investigación que se está llevando a cabo en la Universidad de Extremadura desde el año pasado y que concluirá a finales del presente año. El primer objetivo del estudio es analizar cómo están comercializando actualmente los gestores de alojamientos rurales extremeños sus productos, qué herramientas de marketing emplean, y cuáles resultan a su juicio más efectivas. Tras este diagnóstico, se plantea el segundo objetivo de la investigación, consistente en proponer mejoras en las prácticas identificadas, potenciando aquellas qué se revelen cómo más efectivas y modificando o eliminando aquellas que no ofrezcan los resultados esperados.

En definitiva se pretende realizar un diagnóstico de la comercialización de turismo rural en Extremadura a fin de implementarla de cara a lograr un mejor posicionamiento en el mercado turístico que posibilite un mayor aprovechamiento, rentabilidad y rotación de las infraestructuras existentes. Además, alcanzar una mayor ocupación, y la consiguiente rotación de activos, facilitaría estrategias promocionales que fijaran precios más competitivos[5], sin ver mermada la rentabilidad de los establecimientos.

A continuación, pasamos a presentar la metodología que se definió para la consecución de objetivos fijados.

2. METODOLOGÍA.

Dado que se pretende identificar cuáles son los principales problemas que hayan las pymes turísticas extremeñas a la hora de comercializar sus servicios, a fin de formular propuestas dirigidas a paliar las deficiencias halladas, facilitando la labor comercial de los pequeños y medianos empresarios turísticos extremeños, se diseñó un cuestionario para recabar información sobre las prácticas más habituales en el sector[6].

Se definió la población objetivo del estudio cómo aquellos establecimientos (hoteles rurales, casas rurales y apartamentos turísticos) que ofrecen entre sus servicios alojamiento, y que estando ubicados en un entorno rural, no pertenecen a ninguna cadena nacional. La población censada inicialmente, según datos facilitados por la Dirección General de Turismo de la Junta

[5] Según un estudio elaborado por la Web especializada en turismo rural, Toprural.com., en el que analizaron los datos de más de 4.000 alojamientos rurales anunciados en su Site, las regiones españolas con precios más asequibles son Murcia, Navarra y Aragón, mientras que Baleares, Madrid **y** Extremadura**,** registran los precios más elevados. El precio medio en temporada alta es de 33,3€ por persona y noche**.** El hotel rural es el tipo de alojamiento más caro y la casa rural de alquiler íntegro la más económica. Sin embargo para el caso de Extremadura, los hoteles rurales resultan más económicos que la media nacional, costando por término medio 41.2€ por persona y noche, frente a los 44.8€ de media nacional.

[6] El cuestionario íntegro se presenta en el Anexo II.

de Extremadura, fue de 639 establecimientos[7]. Debido a la dispersión de los alojamientos sobre el territorio regional, y teniendo en cuenta que la casi totalidad de ellos dispone una dirección de contacto a través de Internet, se decidido administrar el cuestionario online.

Asumiendo un error máximo del 5%, se estableció como objetivo lograr un tamaño muestral mínimo de 240 entrevistas. Se diseño un cuestionario y se remitió por correo electrónico junto a una carta de presentación a la totalidad de la población. La primera oleada se realizó el 12 de Marzo de 2009 lográndose 64 respuestas. El 19 de Marzo se lanzó la segunda oleada, dirigida a quienes aún no habían respondido, lográndose 80 respuestas. Se descansó dos semanas antes de lanzar la tercera oleada, para no molestar a los establecimientos durante el periodo vacacional de semana santa. Actualmente, seguimos recabado las respuestas de la tercera oleada. Hasta la fecha disponemos de 160 encuestas válidas[8].

Pasamos a comentar los resultados más relevantes obtenidos hasta la fecha.

3. RESULTADOS PRELIMINARES.

De las 26 entidades territoriales delimitadas en la región se están representadas en la muestra 24. Únicamente no se dispone de respuestas de establecimientos ubicados en la comarca "Tierra de Barros" y en el área de los "Llanos de Cáceres".

La tasa de respuesta lograda, por ahora, es de un 22% de Casas Rurales, un 26% de Apartamentos Rurales y un 35% de Hoteles Rurales. Pensamos que la mayor profesionalizada de la gestión de los hoteles puede explicar los resultados obtenidos hasta la fecha. En base a estudios anteriores conocemos que por lo general los propietarios de casas y apartamentos rurales no tienen en esta actividad su principal fuente de ingresos[9], lo que les lleva a dedicar menos tiempo a la gestión de los establecimientos, y estar menor dedicación e interés podría verse reflejada en una menor tasa de respuesta.

Sólo un 10% de los establecimientos entrevistados lleva ejerciendo su actividad más de 10 años, mientras que un 24% inició su actividad en los últimos dos años. Por lo tanto, más del 75% de los establecimientos lleva en el sector al menos dos años, tratándose de una oferta turística experimentada.

El 86% de los establecimientos entrevistados tiene una capacidad de seis habitaciones o menos, son establecimientos de pequeña dimensión. Un 42% de los establecimientos ofrece exclusivamente el servicio de alojamiento, mientras que el resto ofrece restauración y/o servicios complementarios (actividades al aíre libre).

Al preguntarles por la ocupación que suelen registrar en las distintas temporadas del año, comprobamos que presenta un fuerte componente estacional, variando de forma significativa a lo largo del año. Más del 70% de los establecimientos declara una ocupación media en invierno inferior al 30%. En general es la peor época, la mejor es el verano, seguida de la primavera y el otoño. A pesar de ello, sólo un 15% de los establecimientos registra ocupaciones superiores al 70% en primavera, hecho que sorprende puesto que se trata de temporada alta. Según declaran

[7] El listado inicial contenía 640 establecimientos, pero comprobamos que uno de ellos había cesado su actividad en el momento de realizar la encuesta.

[8] La elaboración del la base de datos, se está llevando a cabo con la ayuda del alumno de la Licenciatura en Marketing e Investigación de Mercados de la UEX, D. Manuel Moreno López, becario de colaboración del Ministerio de Educación Política Social y Deporte para el curso académico 2008-2009.

[9] "Según los resultados obtenidos en el análisis de la oferta turística de las comarcas extremeñas de La Serena, La Siberia y Olivenza, la actividad turística no es la principal fuente de ingresos para una parte importante de los empresarios turísticos. Se constata que los propietarios de hoteles, hostales o pensiones suelen, en general, explotar el negocio a título principal, los empresarios que ofertan alojamientos en el medio rural o servicios de restauración suelen dedicarse a otras actividades económicas, lo que convierte al negocio turístico en un mero complemento de rentas. Prácticamente la mitad de los empresarios consultados obtienen sus rentas personales de forma exclusiva del negocio turístico, mientras que la otra mitad está integrada por empresarios para los que la actividad turística no es su principal fuente de ingresos" (Andrades y Sánchez, 2007).

la mayoría de los entrevistados, la ocupación en primavera oscila entre el 30 y el 70%. En verano es cuando registran las mayores tasas de ocupación del año, época en la que un 20% de los establecimientos superan el 70% de ocupación, y el 40% logra al menos un 30% de ocupación. En Otoño la ocupación cae respecto al verano aunque es más favorable que en Invierno y menos favorable que en Primavera. A priori no parecen unos resultados muy satisfactorios, y ante ellos se plantea la cuestión que constituye el primer objetivo de nuestra investigación: ¿Cómo comercializan los empresarios turísticos extremeños sus productos?

Los datos analizados revelan que la mayoría de los establecimientos coinciden en que su clientela acostumbra a reservar contactando directamente con el establecimiento 55%, o bien a través de Internet, 30%. Un menor porcentaje de clientes acude a los establecimientos sin reserva previa, 8%, y son aún menos los clientes los captados a través de agencias de viajes, 7%.

La baja captación de clientes a través de agencias de viajes quizás pueda ser explicada por la baja promoción que los establecimientos efectúan sobre este canal, ya que la gran mayoría considera una perdida de tiempo y dinero distribuir material promocional entre las agencias o diseñar viajes de familiarización para este colectivo.

En cuanto a los viajes de familiarización, sólo un 8% de los encuestados los califica cómo "efectivos" o "muy efectivos", siendo junto con la asistencia a ferias especializadas y la publicidad en medios de comunicación las actividades menos rentables en términos de captación de clientes de todas las que se le preguntaron.

Sin embargo, aunque tampoco destaque por la eficacia que le atribuyen, sí hay más establecimientos, un 20%, que estiman conveniente el diseño y distribución de folletos promocionales entre los potenciales clientes.

Si se establece un ranking de las actividades de marketing que consideran más acertadas, de cara a facilitar la comercialización de sus productos y servicios, la actividad mejor valorada es situar su oferta en una central de reservas online, seguida de la mejora de las instalaciones, la mejora de la calidad del servicio y la formación de los empleados que atienden al público.

Los últimos datos publicados por el IET[10] indican que los turistas extranjeros que llegan a España bajo la modalidad de turismo rural, representan únicamente el 7% de las plazas ocupadas. Al preguntar a los entrevistados, por este segmento de turistas, comprobamos que sólo un 20% de los establecimientos colabora con intermediarios internacionales para captar clientes extranjeros, lo que podría explicar la escasa representación de los turistas extranjeros en la ocupación actual de los establecimientos de la región. Sin embargo, un 60% manifiesta que aunque ha pensado que sería rentable acceder a éstos mercados, y les gustaría hacerlo, no sabe cómo. El 20% restante declara que o bien no se lo habían plantado o que habiéndolo hecho no lo juzgaron interesante.

El 63% de los establecimientos dispone de bases de datos de clientes, en la que consignan información útil afín de establecer contactos comerciales, sin embargo exclusivamente el 35% de los establecimientos las usan y además en muy pocas ocasiones. El 15% de los consultados considera que contactar con sus clientes por motivos promocionales resulta inefectivo. La práctica más común, es contactar con antiguos clientes en contadas ocasiones, y sólo con quienes se mantenga una relación especial. El escaso coste de hacer mailing periódicos con fines promocionales una o dos veces al año, puede justificar que sean realizados por un tercio de los establecimientos, auque opinen en su mayoría que resultan poco o nada efectivos.

En cuanto a la efectividad de ofrecer descuentos y ofertas a los clientes para aumentar la ocupación, resulta curioso que más de un tercio no se pronuncie al respecto. Sólo el 30% declaró que les resulta efectivo aplicar estrategias promocionales de precios, ofreciendo descuentos que generan una mayor ocupación de sus negocios.

Finalmente, tras haber evaluado estrategias de distribución, comunicación y precios, pasamos a valorar qué estrategias están desarrollando en materia de producto turístico. Además de apostar por la mejora de las instalaciones, la mejor capacitación profesional de sus empleados y la calidad del servicio, queríamos evaluar cómo diseñan sus productos turísticos. En particular preguntamos por el hecho de si mejoran el producto que ofrecen completando la oferta de alojamiento con otras ofertas que incrementen el atractivo conjunto del servicio. Según

[10] Instituto de Estudios Turísticos: Ficha de Resumen Mensual de Turismo Rural. Febrero 2009.

los datos recabados hasta la fecha, más de la mitad de los establecimientos diseñan paquetes turísticos para aumentar el interés de su oferta y lograr así mayores tasas de ocupación, además el 40% apunta esta opción como efectiva o muy efectiva de cara a alcanzar este objetivo.

4. CONCLUSIONES.

Hoy en día, se acepta la idea de que el turismo rural puede representar una vía de diversificación económica que vigorice las, por lo general, débiles economías rurales. En este sentido la actividad turística puede promover el desarrollo sostenible del medio rural. Extremadura es una comunidad predominantemente rural, que adolece de los problemas típicos de este hecho: baja densidad de población en el medio rural, población envejecida, población poco cualificada, elevadas tasas de desempleo, malas infraestructuras, etc. En este contexto, los gestores públicos están apostando decididamente por impulsar la actividad turística sostenible como apoyo a la actividad económica en el medio rural. Asimismo asociaciones como la Red Extremeña de Desarrollo Rural, CETEX[11] o FEXTUR[12], están destinando recursos para dinamizar este sector en nuestra región. Se pretende que sumando los esfuerzos de los muy diversos agentes implicados en el desarrollo regional, se logre mejorar los estándares de vida de la población local, satisfaciendo una demanda de turismo diferente[13] y conservando el entorno donde lleva a cabo su actividad. Hasta ahora, el fruto de esta apuesta se hace patente en el hecho de que Extremadura está registrando unas llamativas tasas de crecimiento del número de turistas que la visitan cada año, observándose en los últimos años un crecimiento tanto de su demanda como de su oferta turística.

A pesar de estos datos, y aunque es indudable el atractivo turístico de Extremadura, de su valioso patrimonio natural y cultural, observando los datos oficiales que publica el IET sobre participación de las distintas Comunidades Autónomas en el mercado de turismo rural, se comprueba que Extremadura sigue siendo un destino emergente y que aún quedan mucho por delante para lograr una participación más sustancial en el mercado. Por estas razones, se diseñó una investigación para identificar posibles causas de la aún baja presencia de los establecimientos extremeños en el mercado de turismo de naturaleza, así como las necesidades potenciales del sector, a fin de obtener información para la revisión de las estrategias que se están aplicando y la optimización de las mismas.

Según las opiniones reveladas en la encuesta que se está efectuando, las mayores tasas de actividad comienzan a registrarse en primavera, elevándose en verano y cayendo a partir del Otoño hasta quedar reducidas a menos del 30% de ocupación media durante los meses de invierno. En relación a estos datos, consideramos preocupante resaltar que ni en temporada alta los establecimientos logran ocupaciones superiores al 70% de media. Sí es cierto que en primavera, en concreto en semana santa registran ocupaciones superiores al 90%, pero se trata de una fecha puntual. Estas cifras sugieren que no se está llegando de forma adecuada a los clientes, y que urge perfeccionar las estrategias de cara a conseguir una mayor penetración en el mercado de turismo rural, tanto nacional como internacional.

Un aspecto muy satisfactorio a resaltar y que nos ofrece una ventaja competitiva importante, es que el perfil empresarial dedicado al turismo de naturaleza en Extremadura es de un empresario con años de experiencia en el sector y que además es consciente de la relevancia que tiene en su negocio ofrecer calidad en el servicio, disponer de buenas instalaciones y contar con profesional competente. Además la mayoría utiliza el diseño de paquetes turísticos como una herramienta de trabajo clave a la hora de incrementar el poder de atracción de su oferta. Estas serían las principales fortalezas detectadas en este análisis preliminar.

[11] Confederación de Empresarios Turísticos de Extremadura.
[12] Federación Extremeña de Turismo Rural.
[13] El turismo sostenible en un ámbito rural, irá dirigido a turistas, que previamente informados, motivados y "formados" busquen, principalmente, un conocimiento del medio visitado, para el mayor disfrute del mismo, en armonía con el entorno y con los habitantes locales, Mediano, 2004.

En cuanto a las debilidades, aparecen ligadas especialmente al ámbito de las políticas de distribución, precios y comunicación.

En la actualidad, los empresarios opinan que estar presentes en centrales de reservas online es la mejor vía para captar clientes, valorándola más positivamente que el resto de acciones comerciales por las que se les preguntó. Pero realmente esta es la única acción que estiman positiva, tachando de infructuosas el resto de acciones que las teorías de marketing para turismo rural vigentes aconsejan.

La escasa efectividad que atribuyen a la captación de clientes a través de agencias de viajes, y las acciones promocionales sobre las mismas (distribución de material promocional y diseño de viajes de familiarización) sorprende teniendo en cuenta que son estrategias altamente recomendables para el turismo de naturaleza (Mediano, 2004) y así lo confirma la experiencia en otras comunidades autónomas españolas (Aranda y Cantanero, 2004). Ante esto, la pregunta ineludible es ¿realmente no son efectivas las acciones de promoción sobre intermediarios, o no están siendo efectivas porque quizás no lo hacen de forma adecuada? Este será, sin duda, un aspecto a investigar con más profundidad en la investigación que estamos desarrollando.

En cuanto a la baja efectividad atribuida a la asistencia a ferias especializadas para la captación de clientes, posiblemente se deba a que si bien sí es interesante hacerlo de forma conjunta, amparados en asociaciones sectoriales o territoriales, ya que los costes de abordar estas actividades se diluyen, no es rentable hacerlo a título individual puesto que los costes son muy elevados. En este sentido, la recomendación sería llevarlos a cabo a través de asociaciones empresariales, como por ejemplo FEXTUR, o patronatos de turismo, que den a conocer el destino, generando una marca paraguas debajo de la cual puedan desarrollar los establecimientos sus estrategias de marcas individualizadas.

Por último, la escasa valoración estratégica de la publicidad en medios de comunicación seguramente se deba a los mismos motivos de coste aludidos en el punto anterior, ya que resulta muy gravosa para un establecimiento aislado. Difícilmente un único establecimiento podría amortizar una inversión en publicidad en medios de comunicación, máxime cuando la dimensión media de los establecimientos no supera las seis habitaciones. En este sentido, la responsabilidad de dar a conocer el destino, en nuestra opinión, recaería sobre las administraciones públicas competentes en sus diferentes niveles territoriales.

Finalmente, insistir en que los resultados presentados forman parte de una investigación que aún se haya en fase de desarrollo y por lo tanto los resultados deben ser interpretados con prudencia puesto que podrían verse alterados con la obtención de más información adicional en los próximos meses. En cualquier caso, parece que un aspecto clave para desarrollar el turismo rural en Extremadura es identificar nuevos canales de comercialización, especialmente agencias de viajes, nacionales e internacionales, especializadas en esta modalidad turística de cara a estimular la distribución indirecta de la oferta turística rural extremeña.

5. BIBLIOGRAFÍA.

ALTÉS, C. (1997): *Marketing y Turismo. Introducción al marketing de empresas y destinos turísticos.* Ed. Síntesis, S.A. Pág. 88, 121, 225.

ANDRADES, L. Y SÁNCHEZ, M. (2007): "Posibilidades de Aprovechamiento de los Embalses de Alqueva y La Serena como potencial fuente de diversificación económica en el medio rural". Estudio realizado para la Consejería de Desarrollo Rural de la Junta de Extremadura.

ARANDA DELGADO, M., Y CANTANERO QUESADA, J.M. (2004): Cuaderno de Promoción y Comercialización Turística. Junta de Andalucía.

ARAQUE, E. Y SÁNCHEZ-MARTINES, J.D. (2002): "Cazorla, Segura y Otras Villas: oportunidades para la reconstrucción sostenible de un espacio turístico de interior". *Cuadernos de Turismo*, Nº10; pp. 85-100.

ÁVILA, R. Y BARRADO, D. A. (2005): "Nuevas tendencias en el desarrollo de destinos turísticos: marcos conceptuales y operativos para su planificación y gestión". *Cuadernos de Turismo*, Nº15; pp. 27-43.

CÁNOVES, G. Y HERRERA, L. (2005): "Turismo Rural en España: Paisajes y Usuarios, Nuevos usos y nuevas visiones". *Cuadernos de Turismo*, Nº15; pp 63-76.

CASILLAS BUENO, J. C.; MORENO MENÉNDEZ, A. M. Y OVIEDO GARCÍA, M.A. (1995): "El turismo alternativo como un sistema integrado: consideraciones sobre el caso andaluz". *Estudios Turísticos*, Nº. 125, pp 53-74.

COCCOSSIS, H. (1996): Tourism and sustainability: perspectives and implications. *En Priestley, G. K.; Edwards, J.A. y Coccossis, H (eds.) Sustainable tourism? European Experiences.* CAB International, Wallingford.

COCCOSSIS H.(2001): "Tourism Development and Carrying Capacity". In *Tourism Sustainable Development and Natural Resource Management: experiences of Carribean, Pacific and Mediterranean islands chapter 8.* Apostolopoulos G. and D.Gayle, (eds). Greenwood, Connecticut.

CROSBY, A., DARIES, J. O. Y OTROS (1993): *El desarrollo turístico sostenible en el medio rural*. Centro Europeo de Formación Ambiental y Turística. Madrid.

CROSBY, A. Y MOREDA, A. (1996): *Desarrollo y gestión del turismo en áreas rurales-naturales.* Centro Europeo de Formación Ambiental y Turística. Madrid, p.182.

GARCÍA, B. (2003): *Marketing del Turismo Rural*. Ed. Esic. Pág, 213.

Instituto de Estudios Turísticos (2009): Ficha de Resumen Mensual de Turismo Rural. http://www.iet.tourspain.es/paginas/rsvisor.aspx?ruta=/Fichas%20Resumen/Estructura/Mensual/Ficha%20Resumen%20Mensual%20de%20Turismo%20Rural%20-%20Ref.709&par=1&idioma=es&mes=Febrero&anio=2009

KOTLER, P.; BOWEN, J.; MAKENS, J.; RUFÍN, R. Y REINA, M.D. (2004): *Marketing para turismo*. Ed. Pearson Educación-Prentice Hall. Madrid, pp.41-66.

MATHIESON, A. Y WALL, 0. (1986): *Turismo: Repercusiones económicas, físicas y sociales.* Trillas, México.

MEDIANO, L. (2004): *La Gestión del Marketing en el Turismo Rural.* Pearson Educación-Prentice Hall. Madrid, pp.186-202.

MORENO, S. Y OTROS (2002): "Análisis de la satisfacción del turista de paquetes turísticos respecto a las actividades de ocio en el destino: el caso de República Dominicana". *Cuadernos de Turismo*, Nº 9; pp. 67-84.

TURESPAÑA, EL MINISTERIO DE COMERCIO Y TURISMO. SECRETARÍA GENERAL DE TURISMO. (1994): *Manual del planificador de turismo rural*. Instituto de Estudios Turísticos. Madrid, p.33.

TOPRURAL (2009): Estudio comparative de precios de turismo rural en España. http://www.toprural.com/docs/comparativanacionalprecios.pdf

PAMADERO, M. Y OTROS (2002): "Turismo en espacios naturales: oportunidades en el corredor biológico Mesoamericano". *Cuadernos de Turismo*, Nº10; pp.69-83.

POPOVICH, M. R. Y TOSELLE, C. (2006): "Planificación de una estrategia participativa y comunidad local. Desarrollo de un destino turístico en Argentina". *Cuadernos de Turismo*, Nº17; pp. 167-188.

REIG, C. Y COENDERS, G. (2002): "Segmentación del mercado turístico según las preferencias ambientales". *Cuadernos de Turismo*, Nº 9; pp. 123-135.

REGUERO OXINALDE, M, (1994): Ecoturismo. Nuevas formas de turismo en el espacio rural. Bosch Turimo. Barcelona.

RUIZ, A.V., VÁZQUEZ, R. Y DÍAZ A.M. (1995): "La calidad percibida del servicio en establecimientos hoteleros de turismo rural". *Papers de Turisme*, Nº 19, pp. 17-33.

VÁZQUEZ, R. (1996): *"Estrategias de marketing para las empresas de turismo rural"* en Valdés Peláez, L. y Ruiz Vega, A. (coord.): op. Cit., pp.79-98.

VÁZQUEZ, R Y DÍAZ, A.M. (1997): "Calidad del servicio en turismo: gestión empresarial en función del conocimiento de las expectativas de los clientes". *Papers de Turisme*, Nº 20, pp. 87-114.

Anexo I: Encuesta.

Encuesta sobre comercialización de alojamientos rurales en Extremadura 2009. Universidad de Extremadura.

Aclaraciones con respecto a las encuestas:
La encuesta va dirigida a gestores de alojamientos de turismo rural ubicados en Extremadura.
Se preservará en cualquier caso el anonimato de los encuestados.
Por favor, no deje ninguna pregunta sin contestar, los resultados de la investigación se verían perjudicados.
Contestar al cuestionario le llevará unos seis minutos o menos.
Gracias por su colaboración,

1. ¿Podría indicar la comarca extremeña en la que está ubicado su negocio?

2. Indique por favor el tipo de establecimiento que regenta:

Apartamento Rural ○ Casa Rural ○ Hotel Rural ○

3. ¿Cuántos años lleva gestionando su actual negocio turístico?

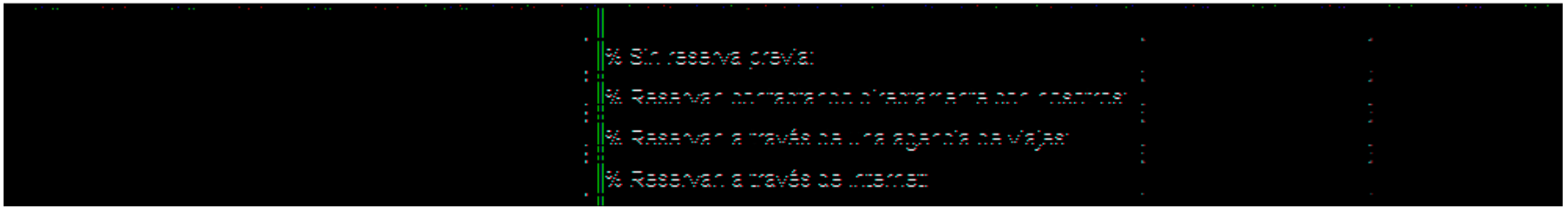
- ○ Menos de 2 años.
- ○ Entre 2 y 4 años.
- ○ Entre 4 y 10 años.
- ○ Más de 10 años.

4. ¿Cuántas habitaciones tiene disponible en su establecimiento?

5. ¿Además de alojamiento ofrece otros servicios en su establecimientos?

- ◯ No.
- ◯ Sí, restauración.
- ◯ Sí, otros (actividades al aire libre, etc).
- ◯ Sí, restauración y otros (actividades al aire libre, etc).

6. ¿Podría indicarnos aproximadamente el porcentaje de clientes que acceden a su establecimiento por los siguientes medios? (escriba el porcentaje aproximado en el espacio en blanco, por ejemplo: 30, 20, 10, 40).



7. ¿Dispone de una base de datos de clientes que incluya los datos de contacto de los mismos (dirección postal, correo electrónico, o teléfono?

No. ◯ Sí. ◯

8. ¿Alguna vez contacta con sus clientes para animarles a volver a visitarles con motivo, por ejemplo, de algún evento especial que se produzca en la zona?

(Si su respuesta es "no", marquela y pase a la pregunta número 10)

No. ◯ Sí. ◯

9. ¿Con qué frecuencia contacta con sus clientes para animarles a volver?

- ◯ Pocas veces, y sólo con quienes mantenemos una relación muy especial.
- ◯ Una o dos veces al año.
- ◯ Tres o cuatro veces al año.
- ◯ Con mucha frecuencia, más de cuatro veces al año.

10. Nos gustaría saber si cuando ha contactado con antiguos clientes para promover que vuelvan a su establecimiento, ha resultado efectivo, por favor marque la opción que mejor refleje su opinión:

Nada efectivo.	Poco efectivo.	Efectivo.	Muy efectivo.
○	○	○	○

11. ¿Podría indicarnos el grado medio de ocupación, aproximado, que tuvo su establecimiento el pasado año, en cada una de las estaciones?

	30% o menos	más del 30% y menos del	más del 70%
Invierno	○	○	○
Primavera	○	○	○
Verano	○	○	○
Otoño	○	○	○

12. ¿Desarrolla alguna estrategia específica para atraer a su establecimiento a turistas extranjeros?

- ○ No, no lo había pensado nunca.
- ○ No, porque no creo que sean interesantes para mi negocio.
- ○ No, pero me gustaría hacerlo aunque no sé cómo.
- ○ Sí, de hecho trabajo con mayoristas extranjeros (como Ornitholidays, Ramblersholidays, o Blakes-cottages).

13. A continuación enumeramos, algunas actividades de marketing que pueden desarrollarse para incrementar la ocupación de un establecimiento turístico.
¿Podría marcar aquellas que efectuó, indicando cómo de efectivas resultaron?
(siendo 5 "muy efectiva, atrajo a muchos clientes" y 1 "Nada efectiva, no logré clientes adiccionales")

Así por ejemplo, si realizó folletos y resultaron muy efectivos cuando los repartió entre clientes marcará un 5,
sin embargo si relutaron poco efectivo cuando los distribuyo entre agentes de viaje, marcará un 2 para esta segunda actividad.

	1	2	3	4	5
Promociones y ofertas en precio.	○	○	○	○	○
Promoción en ferias especializadas (Intur, FIO, etc.).	○	○	○	○	○
Publicidad en medios de comunicación (radio, prensa, etc.)	○	○	○	○	○
Reparto de folletos para dar a conocer mi establecimiento a clientes potenciales.	○	○	○	○	○
Reparto de folletos para dar a conocer mi establecimiento a agencias de viajes.	○	○	○	○	○
Diseño de paquetes turísticos.	○	○	○	○	○
Mejora de la calidad del servicio.	○	○	○	○	○
Mejora de las instalaciones.	○	○	○	○	○
Mejora de la formación de los empleados que atienden el público.	○	○	○	○	○
Introducir mi establecimiento en centrales de reserva de Internet (tales como toprural, ruralgest, crastur, TURIEX, etc.)	○	○	○	○	○
Organización o participación, en viajes de familiarización para dar a conocer mi establecimiento a los agentes de viaje y así lo recomienden.	○	○	○	○	○

14. Para agilizar sus tareas de captación de clientes y comercialización de su establecimiento, ¿utiliza el ordenador con alguno de los siguientes fines?

(marque todas las que realice, y si realiza alguna más puede indicarlo, si lo desea, en el espacio en blanco):

- ☐ Relaciones con mayoristas y otros intermediarios (agencias de viaje, centrales de reserva online, etc.)
- ☐ Gestión de reservas.
- ☐ Publicidad de mi establecimiento a través de mi página web.
- ☐ Envío de correos electrónicos a clientes ofreciéndoles ofertas especiales.
- ☐ Otros usos relacionados (por favor, especifique) _____

15. ¿Mantiene relación con otras empresas turísticas y entidades relacionadas con el turismo (tales como oficinas de información turística, museos, centros de interpretación,etc.) de su comarca, a fin de diseñar paquetes turísticos que contengan sus productos/servicios y facilitar así su comercialización?

(Si su respuesta es "no", marquela y pase a la pregunta 16)

No. Si.

16. ¿En que medida resulta efectivo el diseño de paquetes turísticos conjuntos para lograr una tasa de ocupación mayor en su establecimiento?

Nada efectivo. Poco efectivo. Efectivo. Muy efectivo.

17. Si desea recibir, despúes del verano, un informe con los resultados obtenidos globales y por comarcas, deje su email:

DESIGNING AND COMMUNICATING SUSTAINABLE TOURIST EVENTS IN RURAL ENVIRONMENTS, AS A DIFFERENTIATION MARKETING STRATEGY IN ORDER TO RE-INVIGORATE A TOURISTIC DESTINATION.

LIDIA ANDRADES CALDITO
E-mail: andrades@unex.es
MIGUEL ÁNGEL FAJARDO CALDERA
E-mail: fajardo@unex.es
Facultad de Ciencias Económicas y Empresariales
Universidad de Extremadura

Resumen

El diseño de eventos en espacios rurales se configura como una alternativa de diferenciación para destinos rurales. En los últimos años la demanda de turismo de naturaleza ha ido en aumento, lo que ha venido acompañado de una expansión de la oferta turística. Si a éste incremento de la oferta, le añadimos la estacionalidad propia del sector turístico, y la actual coyuntura de crisis económica generalizada, resulta que en los últimos años ha crecido notablemente la rivalidad turística entre destinos y municipios que se ubican en entornos muy próximos. Ante esto, la gestión de eventos se configura como una herramienta para potenciar el atractivo turístico de un destino. El objetivo de este trabajo es analizar las posibles estrategias de comunicación que podrían desarrollarse con respecto a eventos turísticos sostenibles que se programen con el fin de dinamizar turísticamente un territorio.

Palabras Claves: "Turismo"; "Eventos"; "Sostenibilidad"; "Marketing".
Área temática: Economía Empresarial.

Abstract

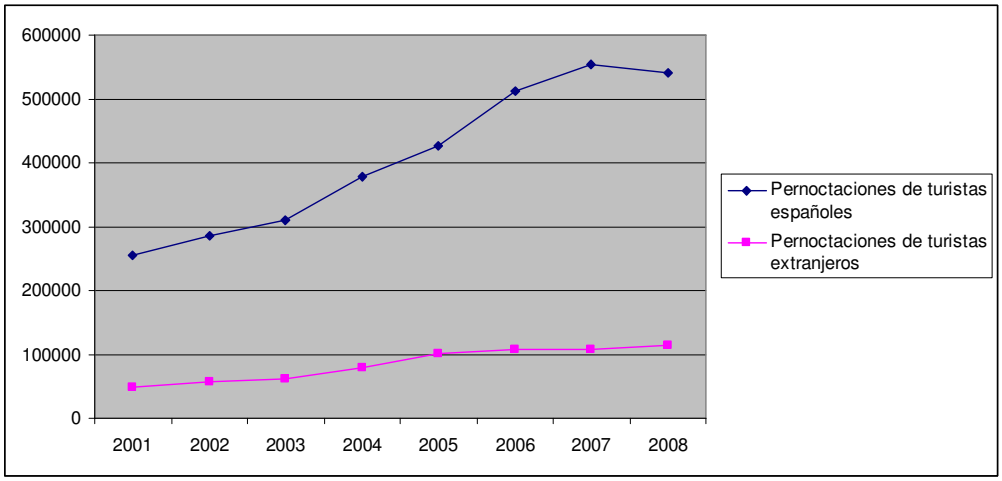
The design of events in rural spaces constitutes an alternative of differentiation for rural tourist destination. In the last years, the rural tourist demand has increased, which has come accompanied by an expansion of the rural tourist supply. If to this one increase of the supply, we added the own seasonality of the tourist sector, joint to the actual generalized economic crisis, it can be deducted that the tourist rivalry between destinations and municipalities, that are located in very next surroundings, has being growing remarkably. Due to this, the management of events is formed as a useful tool to harness the touristic attractiveness of a destination. The objective of this work is to analyze the possible strategies of communication that could be developed with respect to sustainable tourist events, which could be planned with the aim of to stimulate a tourist destination.

Key Words: "Tourism"; "Events", "Marketing", "Sustainability".
Thematic Area: Entrepreneurial Economics.

1.1. INTRODUCCIÓN.

Durante la última década, hemos asistido a un incremento paulatino de la demanda de turismo de naturaleza, hecho que comenzó a contabilizarse oficialmente en España desde el año 2001. En el gráfico 1.1.1. se aprecia el incremento de pernoctaciones medias anuales tanto de residentes en España como de turistas provenientes de otros.

Gráfico 1.1.1. Evolución del número medio de pernoctaciones anuales en alojamientos de turismo rural.



Fuente: elaboración propia a partir de los datos de la Encuesta de Ocupación de Turismo Rural, publicados por el I.E.T.[1]

El aumento del interés por esta modalidad turística, ha ido acompañado de una ampliación de la oferta alojativa rural, a fin de ajustarse al mercado y dar respuesta a la creciente demanda. Además, la expansión de la oferta también ha sido impulsada por numerosas políticas fomentadas por las distintas administraciones públicas, debido a que el desarrollo del turismo rural representa una vía de diversificación económica de este entorno, (Rico, 2005). En este sentido, el turismo rural es visto como una herramienta para vigorizar las economías rurales y una forma de que las familias de este medio puedan lograr rentas complementarias a las agropecuarias, que son las predominantes en este ámbito.

El modelo turístico promovido, desde la óptica del desarrollo rural, ha sido el de un desarrollo turístico ordenado, no masificado y respetuoso con el medio ambiente, en una palabra, sostenible. Entendiendo por "Sostenible" la actividad turística capaz de mejorar los estándares de vida de la población local, satisfaciendo una demanda de turismo diferente y conservando el entorno donde lleva a cabo su actividad, (Casillas, Moreno y Oviedo, 1995). En este sentido, cuando el desarrollo turístico respete los límites de la sostenibilidad, será capaz de satisfacer las necesidades actuales sin comprometer la capacidad de las generaciones futuras para satisfacer sus propias necesidades (Comisión Mundial sobre Medio Ambiente y Desarrollo, 1987).

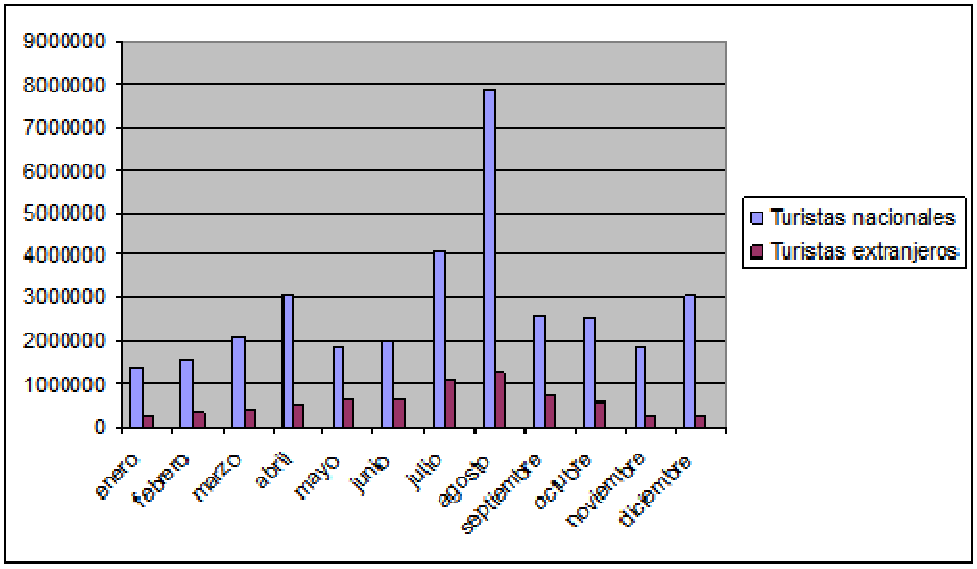
En definitiva, en los últimos años el sector turístico y el medio rural se han beneficiado mutuamente y gracias a su interacción se ha creado empleo, infraestructuras, se ha preservado y divulgado la cultura local, costumbres, etc. Sin embargo, en la actualidad nos hayamos ante un contexto socioeconómico inquietante, ya que por primera vez en muchos años, la renta nacional se haya en recesión. Además, si bien la actual crisis económica es un hecho constatado, aún existe mucha incertidumbre acerca de por cuanto tiempo se prolongará y cuáles serán sus efectos.

El sector turístico, no es ajeno a la crisis y está padeciendo los efectos de la coyuntura económica, reflejándose en los principales indicadores del sector (descenso del número de visitas, pernoctaciones, acortamiento de la duración media de los viajes, descenso del gasto medio por persona y día, etc.[2]). Junto a la coyuntura poco halagüeña, la estacionalidad propia del sector turístico se suma a los males del sector, gráfico 1.1.2. En este marco económico, desde un punto de vista optimista, podríamos pensar que el turismo rural podría resultar beneficiado, en el sentido de que pueda ser un producto turístico sustituto de vacaciones más largas a destinos más remotos, al ser un producto más accesible en cuanto al factor precio. Sin embargo, los últimos datos publicados por el I.E.T., para los meses de Enero y Febrero del

[1] Instituto de Estudios Turísticos.
[2] Datos disponibles en FAMILITUR; FRONTUR Y EGATUR, elaboradas por el I.E.T.

presente año[3], invalidan esta interpretación optimista, revelando que el turismo rural también se está viendo perjudicado por la crisis económica. Tras muchos años de crecimientos continuos, en 2009 se consolida la tendencia a la baja de finales de 2008, y por primera vez cae el número de pernoctaciones de turistas en alojamientos rurales, disminuyendo un 1.8% en Enero y un 11.2% en Febrero respecto a los mismos meses del pasado año.

Gráfico 1.1.2. Pernoctaciones medias mensuales (calculado en base a últimos 8 años)



Fuente: elaboración propia a partir de los datos de la EOTR, publicados por el I.E.T.

Asumiendo que esta tendencia se mantenga en los próximos meses y/o años, cabe esperar que aumente considerablemente la rivalidad turística entre destinos, municipios y establecimientos rurales ubicados en áreas próximas. Máxime cuando en muchas ocasiones los establecimientos rurales suelen ser entrañables casas acogedoras, ubicadas en entornos de gran calidad paisajística, próximos a pequeños municipios en los que ver alguna iglesia o/y otras construcciones significativas y en los que se puede degustar platos típicos en algún restaurante local. Y es que en mayor o menor medida, se trata en definitiva del "producto turístico rural típico". Ante esta estandarización, en el contexto actual de descenso de la demanda, es de esperar que aumente la competencia por captar clientes en el segmento de turismo de naturaleza, lo que se traducirá en una mayor rivalidad entre establecimientos y destinos, que por definición, son muy similares.

Ante la intensificación de la competencia, una posible reacción defensiva sería disminuir los precios, pero estratégicamente esta no suele ser la mejor opción. Desde una perspectiva de marketing, son bien conocidas las repercusiones negativas que se derivan de las "guerras de precio" para un sector en su conjunto. Por lo tanto, sería desaconsejable emplear esta estrategia como única alternativa empresarial ante un descenso de la demanda.

El objetivo de este trabajo es aportar otras alternativas plausibles, de cara a mantener y captar turistas en tiempos de recesión de la demanda. En concreto, en esta comunicación, se plantea la posibilidad de desarrollar una estrategia de diferenciación del destino, basada en el diseño de eventos singulares que aumenten el atractivo del emplazamiento, reforzando el interés de los turistas por visitarlo. Consideramos que la gestión de eventos puede resultar clave, como herramienta para potenciar el atractivo turístico de un destino y dinamizarlo. La cuestión es ¿Cómo diseñar eventos que siendo sostenibles, y consecuentes con la modalidad turística, sean interesantes para los turistas? Para dar respuesta a esta cuestión, desarrollamos un análisis cualitativo a fin de identificar eventos turísticos de éxito reconocido que se hayan celebrado en los últimos años. A partir de ahí, analizamos sus principales características con el fin de aislar los factores de éxito de cada uno de ellos. Tras el análisis realizado, presentamos a modo de

[3] Aún provisionales.

conclusión los aspectos que en nuestra opinión resultan claves a la hora de diseñar un evento atractivo para los turistas, destacando la relevancia de una comunicación eficaz.

1.2. EVENTOS TURÍSTICOS.

Si bien resulta complicado caracterizar los distintos tipos de eventos que se pueden diseñar para incrementar el atractivo de un destino, puesto que pueden ser muy diversos, tras la exploración efectuada, podrían clasificarse atendiendo a los recursos que ponen en valor. Por ejemplo eventos que pongan en valor recursos naturales, artísticos, culturales o combinaciones de éstos. Así, por ejemplo, la *"Fiesta del Cerezo en Flor"* que se celebra anualmente en el Valle del Jerte, Cáceres, utiliza un recurso natural para construir una imagen de destino diferenciado con respecto al resto de comarcas extremeñas, figura 1.2.1.

Figura 2.1.1. Imágenes de la Festa do Vinho, Madeira, y del cerezo en flor, Valle del Jerte.





Fuente: Internet[4].

Otro ejemplo parejo es la *"Festa do Vinho da Madeira"*, Portugal, que rememora la vendimia tradicional, invitando a los turistas a participar de la experiencia de recoger la uva y pisarla con la población de acogida.

Otra categoría de eventos, son los que aprovechan los recursos artísticos de un territorio, como por ejemplo la exposición de arte sacro *"Las Edades del Hombre"* desarrollada en repetidas ediciones en Castilla-León. Otro evento similar diseñado tras el éxito del anterior, fue la exposición artística *"Huellas"* celebrada en Murcia.

Otros recursos culturales susceptibles de ser utilizados para diseñar eventos, son los gastronómicos, como la *"Feria del Queso"* en Trujillo, Cáceres, o la *"Feria del Jamón"* en Jerez de los Caballeros, Badajoz, o artesanales como la *"Festa das Fores"* en Campomaior, Alentejo, Portugal.

Siguiendo con el análisis de eventos culturales, encontramos que también pueden plantearse a partir de tradiciones o hechos de un pasado común o una historia. Es el caso de la celebración del aniversario de la *"Batalla de la Albuera"* en en municipio de La Albuera, Badajoz, o la representación de la obra clásica de Calderón de la Barca, "El Alcalde de Zalamea[5]" en Zalamea de la Serena, Badajoz.

Por último, una categoría de eventos más sofisticados, la constituyen los "macro-eventos" que utilizan múltiples recursos, de muy diversa índole, para ofrecer a los turistas experiencias

[4]http://media.photobucket.com/image/cerezo%20en%20flor%20valle%20del%20jerte%202009/yolytesoro/Valle%20del%20Jerte/DSC06503.jpg
http://www.funchal500anos.com/galeria.asp?ano=2008&idAct=58&i=101#show

[5] Obra literaria del siglo de Oro de la literatura española, de Calderón de la Barca.

únicas, como por ejemplo la Exposición Universal del Agua, celebrada el pasado año en Zaragoza.

Ante la gran diversidad de experiencias, catalogar los eventos e identificar los factores generadores de su éxito, requiere la aplicación de un análisis sistemático, que chequee, al menos, los factores principales que definen el evento, así como el alcance del mismo (poder de atracción), y las repercusiones socioeconómicas inducidas en su entorno (impactos socioeconómicos, directos e indirectos), Cabrini, 2002.

En cualquier caso, si bien hacer un análisis exhaustivo de los eventos celebrados en España y Portugal escapa al ámbito de esta comunicación, las experiencias revisadas muestran cómo el diseño de eventos contribuye al afianzamiento de la imagen del destino, a una mayor notoriedad del mismo y a un mayor interés, por parte de los turistas, por conocerlo o por regresar tras haberlo visitado.

En un estudio efectuado con el fin de valorar los factores generadores de competitividad turística en la región italiana del Abruzzo, los autores aislaron cuatro factores que afectaban notablemente a la imagen del destino: Las atracciones (tangibles e intangibles); Los equipamientos; La eficacia de las acciones de comunicación; y el desarrollo de eventos. Distinguiendo entre eventos dirigidos a touroperadores (viajes de familiarización), o a los empresarios del sector (para la creación de redes entre los alojamientos), o a los turistas (eventos clásicos deportivos, de cultura y/o gastronomía). El estudio concluía que la oferta de eventos y la adecuada difusión de información sobre los mismos, mejoraba la imagen del destino, logrando un mayor número de visitas y alargando la duración de las mismas (Presenza, Moreno y Lobo, 2007).

Por lo tanto, si desde el punto de vista de la estrategia de marketing de un destino turístico, el diseño de eventos favorece el desarrollo de la imagen territorial y su consolidación, será interesante delimitar qué rasgos caracterizan a los eventos más aceptados por los turistas. Con este fin, en el siguiente epígrafe, pasamos a analizar algunos eventos de reconocida relevancia por el número de turistas que los han visitado, y por su consiguiente impacto sobre la imagen del destino. A través de su análisis podremos dar respuesta al objetivo de este trabajo: detectar factores de éxito a imitar al diseñar eventos en un espacio rural de forma que sirvan para diferenciar al destino y potencien su poder de atracción.

1.3. ANÁLISIS DE EVENTOS: DETECCIÓN DE FACTORES GENERADORES DE ÉXITO.

1.3.1. Eventos culturales, Las Edades del Hombre.

El turismo cultural comprende gran variedad de recursos y actividades capaces de lograr que el turista se desplace desde su lugar de residencia hasta otros lugares con el fin de conocer, participar, enriquecer su formación estético-cultural, disfrutar con su contemplación, degustación o participando en los mismos. Al promocionar y desarrollar el turismo cultural, se pretende poner en valor y hacer un uso sostenible de unos recursos poco aprovechados y muchas veces desconocidos, olvidados, expoliados, incluso mal conservados.

Hasta hace pocos años, el Turismo Cultural en España estaba restringido a unas cuantas ciudades monumentales, museos importantes y algunas fiestas como la Semana Santa, Feria de Sevilla o las Fallas. Tenía escasa aceptación porque sus recursos eran poco conocidos y además las deficiencias en promoción e infraestructuras eran grandes (García, 2002). En este contexto las repercusiones socioeconómicas del turismo cultural en España eran poco relevantes. Sin embargo, en la última década el turismo cultural en España ha cobrado fuerza.

Las Edades del Hombre es una exposición de arte sacro que se convirtió en un auténtico evento cultural en Castilla-León y que transcendió más allá de sus fronteras. El proyecto se desarrolló en ocho fases, convirtiéndose en una exposición itinerante que se presentó en distintas ciudades españolas, de Castilla y León, e internacionales, como Amberes y Nueva York. Su éxito quedó demostrado por la masiva afluencia de visitantes, más de 6 millones de personas.

Originalmente, Las Edades del Hombre, surge con el fin de dar a conocer parte de la riqueza histórico-artística de Castilla-León, a la vez que recuperar el mensaje religioso y cultural que, durante siglos, se habían transmitido a través de imágenes, esculturas y pinturas, mostradas a los fieles en las iglesias.

Figura 1.3.1. Las Edades del Hombre.



Fuente: Website oficial: http://www.lasedades.es/

Según un estudio realizado en la Universidad de Salamanca, García 2002, los factores determinantes del éxito de Las Edades del Hombre fueron los siguientes:

- Abundancia, diversidad y calidad de los recursos (Patrimonio Artístico y monumental existente en Castilla y León), y la adecuada selección de las piezas necesarias, eligiendo aquellas que mejor respondían a los objetivos generales de la muestra.
- Equipo humano multidisciplinar que trabajó coordinadamente para presentar la muestra para que fuera comprendida, y gustase a los visitantes.
- El tema central de la exposición, era religioso, pero se presentó como parte de la historia cultural de España.
- La ambientación de la muestra presentaba las obras con originalidad e imaginación, creando un ambiente adecuado a la iconografía expuesta y propicio para cumplir con los objetivos didáctico-religioso-culturales del proyecto inicial.
- Proyección de una imagen de marca de gran calidad.
- Apoyo institucional decidido (Episcopado regional, Caja Salamanca y Soria y Junta de Castilla y León), que viabilizó la entrada gratuita del público, reforzando la idea de que la muestra presentaba la historia y un patrimonio común, propiedad de todos los ciudadanos.

Las consecuencias derivadas de este evento, fueron muy numerosas, relevantes, trascendiendo a múltiples ámbitos: cultural, turístico, social, económico y urbanístico. Sus efectos se percibieron fundamentalmente en las ciudades-sede, pudiendo decirse que hubo un *antes y un después* de dicho acontecimiento (García, 2002). A grandes rasgos, los efectos de la exposición se sintetizan en cuatro grandes bloques: Notoriedad de la Imagen de Castilla León;

Posicionamiento dentro de destinos culturales; Mejora de la Autoimagen de la población endógena; e Incremento de la Receptividad de la población hacia turismo cultural.

Los resultados muestran la sostenibilidad del evento desde la perspectiva económica, social y ambiental, demostrando que el *Turismo Cultural* no constituye una oportunidad remota, sino una tangible, cuando se saben aprovechar los muchos recursos disponibles en los territorios, y se presentan de forma original, atrayente y accesible.

1.3.2. La Expo del Agua, Zaragoza, 2008.

En la pasada Exposición Universal de Zaragoza 2008, "Agua y Desarrollo Sostenible" fueron los temas centrales de este evento internacional. El binomio "agua-desarrollo sostenible" se abordó como un concepto indisoluble, que debe ser afrontado como un reto inteligente: la innovación en el uso del agua para la sostenibilidad del desarrollo humano. Este reto se convirtió en un compromiso de la Expo Zaragoza, plasmado en la Declaración de Impacto Ambiental de la propia Expo. El objetivo perseguido por esta exposición fue lograr ser un referente en la organización de eventos sostenibles y mostrar cómo el impacto ambiental de una exposición internacional puede tender a cero si se efectúa una buena gestión.

Para tal fin, se adoptaron multitud de medidas para minimizar el impacto medioambiental negativo del evento, partiendo desde la propia construcción del Recinto donde se ubicó la Expo. Así, se trabajó para reducir el impacto de las obras de urbanización. Además, en la construcción del recinto ferial se aplicaron criterios de racionalidad ambiental: usando materiales autóctonos, limos y zahorra provenientes del meandro de Ranillas, así como productos reciclados; también se propició el emplazamiento de plantas de hormigón en la misma obra para reducir el impacto del transporte de materiales; se aseguró el empleo de madera proveniente de bosques sostenibles; y la utilización de plásticos, polietileno o polipropileno, sustitutos del PVC. Aparte de esto, también se planificó que en la fase post-expo los pabellones del recinto se reutilizarían, convirtiéndose en un Parque Empresarial de 160.000 metros cuadrados, y otorgándoles así un uso alternativo.

Además se analizaron al detalle las necesidades energéticas de un evento como Expo 2008, considerando las altas temperaturas en la ciudad de Zaragoza durante la celebración de la Exposición, del 14 de junio al 14 de septiembre. En función de estas necesidades, se previó cubrirlas mediante la instalación de sistemas de generación de electricidad a partir de fuentes renovables, y una planta de trigeneración (frío/calor y electricidad) que permitieron ahorrar recursos energéticos y económicos.

Asimismo, conscientes de las emisiones de CO2 derivadas de la celebración de la Exposición Internacional, se propusieron disminuir el impacto ambiental repoblando terrenos aragoneses con la plantación de tres millones y medio de árboles.

En definitiva, el objetivo marcado por la Expo 2008 de ser un referente sostenible a nivel mundial abarcó todas las facetas del evento.

Otro aspecto que la Muestra Internacional cuidó especialmente en su Programa de Desarrollo Sostenible, fue la coordinación de la oferta turística a fin de que existiera un compromiso unánime por la sostenibilidad. Así, involucró a más de quinientos comercios, hosteleros y grandes empresas de Aragón. En el Programa se recogieron medidas muy específicas y sencillas, para infundir un cambio de conciencia ambiental y, en definitiva, promover un entorno urbano más sostenible. En ejemplo de estas medidas, fue el empleo de materiales promocionales y de uso diario sostenibles (vajillas biocompostables, bolígrafos hechos con algas o papel, lápices de CDs reciclados, entrega de bolsas en las tiendas de regalos descomponibles en un plazo máximo de 36 semanas, café y azúcar provenientes del comercio justo, etc.). Otra medida más ambiciosa fue la de procurar un cambio en los hábitos de los ciudadanos, usando, por ejemplo, más racionalmente el agua o favoreciendo el usos de los transportes públicos. Una de las acciones innovadoras previstas para involucrar a la ciudadanía y lograr un consumo responsable, fue la emisión en todo Aragón de la "tarjeta eco-fluvi". Esta tarjeta premiaba con regalos el ahorro en los consumos de gas, electricidad, agua y gasolina. Los organizadores distribuyeron más de medio millón y prolongaron su uso más allá de la clausura de Expo Zaragoza.

Si bien la planificación descrita de la exposición deja claro su apuesta por la sostenibilidad, desde el punto de vista de los visitantes, ¿qué se les ofreció para atraerlos? La clave fue una

amplia oferta de espectáculos[6], numerosas exposiciones[7], actividades, talleres, viajes en telecabina, zonas de descanso, zonas de visualizaciones, bailes, etc., ofrecidos todos en clave de sostenibilidad.

Figura 1.3.2.1. Tarjeta ECO-Fluvi y usos.



Fuente: http://www.ecofluvis.com/asp/catalogo.asp

1.3.3. Eventos en espacios naturales.

Entre los muchos eventos celebrados en torno a recursos naturales, hemos optado por comentar dos tipos de eventos. El primer evento es un concurso para diseñar una camiseta. Este evento, fue promovido por dos empresas privadas cuya oferta de productos y servicios es complementaria: una vende material deportivo de montaña y la otra es un alojamiento del Pirineo español. El segundo es "El cerezo en flor", promovido a nivel institucional para potenciar la imagen de marca del destino turístico "Valle del Jerte" en Extremadura.

La razón de seleccionar eventos tan dispares es mostrar que para diseñar un evento lo imprescindible realmente es la creatividad para definir un acontecimiento que genere interés entre los clientes potenciales de un destino.

El concurso de camisetas, denominado "Concurso de amigotes" y publicitado en la página web corporativa de la tienda de montaña Barrabes, figura 1.3.3.1., consistió en diseñar una camiseta inspirada en el refugio de montaña "Refigi d´ Amitges"[8], figura 1.3.3.2. El premio para los ganadores del concurso fue de 1500€.

Figura 1.3.3.1. Página web de Barrabes en Marzo de 2009.

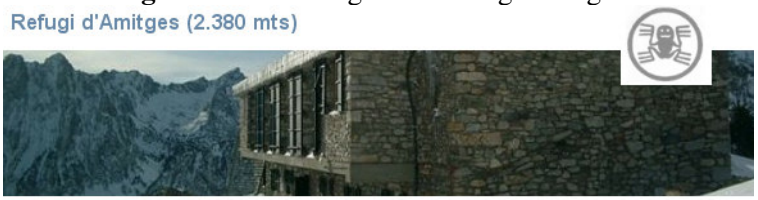


Fuente: Website corporativo http://www.barrabes.com/home.asp

[6] Se presentaron más de 5000 espectáculos, en trece espacios en los que mas de 300 compañías presentaron su arte.
[7] Agua para la vida, Paraísos del Agua, Agua Extrema, Sed, Agua compartida, etc.
[8] Propiedad del Centre Excursionista de Catalunya.

La singularidad de esta simbiosis, entre una tienda y un alojamiento de un destino turístico, es lo más relevante de este ejemplo. Describe perfectamente cómo dos empresas complementarias pueden asociarse para incrementar el interés de sus potenciales clientes (que en este caso constituyen el mismo segmento objetivo: gente aventurera a quienes les gusta disfrutar de la montaña) por sus respectivas ofertas empresariales. La empresa Barrabes ha hecho de la planificación de eventos un factor clave de su éxito en el segmento de montaña. En la actualidad animan su tienda, a través de su espacio virtual, poniendo al alcance de sus clientes, previo registro, una densa agenda de eventos, figura 1.3.3.3., que podrían resultarles interesantes y para los que es posible que necesiten adquirir algunos de los materiales que comercializa.

Figura 1.3.3.2. Imagen del Refugi dámigets.



Fuente: http://www.amitges.com

Así, una empresa que opera en un sector muy técnico, ofreciendo unos productos en los que el asesoramiento es fundamental, crea a través de los eventos una comunidad en torno a su tienda, yendo más allá de la simple interacción con el cliente, buscando, además de satisfacerlo, implicarlo, ofreciéndole un trato personalizado, adaptado a sus gustos, necesidades y expectativas. Tenemos pues, una prueba más de cómo un evento potencia la imagen, en este caso de su promotor, satisfaciendo mejor las inquietudes de sus turistas/clientes, creando vínculos entre ellos, incitando la fidelidelización del cliente, y favoreciendo la comunicación boca-oído del destino (en la caso de Barrabes, de su tienda).

Figura 1.3.3.3. Información de eventos facilitada por Barrabes en su Web corporativa.



Fuente: Agenda de Montaña: http://www.barrabes.com/revista/agenda/enviar.asp

El segundo evento natural que comentaremos, es promovido a nivel institucional en Extremadura cada año. Consiste en la visita del Valle del Jerte durante los escasos días en los que el árbol del cerezo está florecido. Basándose en un recurso paisajístico de innegable belleza, y coordinando tanto a la oferta turística local como a la población de la zona, se prepara la recepción del turista en esos días, para lograr que la experiencia de su visita sea inolvidable. Las claves del éxito de este evento son por una parte la alta calidad del recurso base: el paisaje del valle florecido, y por otra parte el esfuerzo coordinado de los agentes involucrados. Un aspecto

importante, a destacar en este evento es la difusión de información con respecto al mismo, la estrategia de comunicación.

Para el éxito del evento, es preciso poner a disposición de los potenciales visitantes información con valor añadido, explicando la experiencia que se brinda, detallando cómo, cuándo y dónde se ofrece.

Estudios realizados sobre la calidad percibida en la prestación de servicios turísticos, demuestran que uno de los errores más comunes que se cometen a la hora de tratar de mejorar la calidad de servicio, es identificar incorrectamente las mejoras potenciales del servicio, destinando recursos económicos a variar aspectos de la prestación que realmente no mejoran el servicio desde el punto de vista de los clientes (Berry, L., 1995). Sin embargo, la práctica demuestra que existen servicios que se pueden realizar sin apenas coste, y que a la vez proporcionan una satisfacción importante a los turistas (Ortega y Loy Puddu, 2003). Uno de estos "servicios ideales" es ofrecer información útil sobre el destino para que la experiencia del turista sea más rica. Estudios al respecto indican el valor que los turistas otorgan a la información que los establecimientos del destino le suministren sobre el mismo a través de folletos, explicando este factor, un importante porcentaje de la satisfacción/insatisfacción del turista una vez que se hayan en los destinos elegidos (Ortega, E. y Rodríguez, B., 2005). Paradójicamente, proporcionar información útil sobre el evento es relativamente barato si lo comparamos con otras posibles acciones que podrían emprenderse de cara a mejorar la calidad del evento. En consecuencia, será vital establecer con los turistas una comunicación estratégica, que con un coste reducido, les permita disfrutar plenamente del evento diseñado para diferenciar y animar dicho destino.

En el caso del evento "cerezo en flor", la Oficina de Turismo del Valle ofrece en su espacio web información sobre el evento y actividades complementarias, medios para que el turista pueda planificar su viaje, imágenes para hacer tangible el valor que se ofrece, etc., figura 1.3.3.4.

Figura 1.3.3.4. Información oficial, Oficina de Turismo del Valle del Jerte, Cáceres.



Fuente: http://www.turismovalledeljerte.com/

1.3.4. Factores determinantes del éxito de un evento.

Interpretando las experiencias revisadas a la luz de las directrices básicas que los manuales de planificación de marketing para turismo rural proponen, pueden establecerse una serie de pautas a tener en cuenta a la hora de diseñar un evento que dinamice un destino rural:

1. Orientación al turista: importancia de reflexionar sobre el/los segmentos de turistas objetivos del destino, determinando qué tipo de eventos resultarían más sugestivos

para este público. Así por ejemplo, el Refugi d´Amigets propone una actividad divertida para el público que pretende atraer. Selecciona un concurso frente a otras posibilidades porque sus clientes potenciales son deportistas que disfrutan con los retos. Asimismo, fija un premio de una cuantía relevante para su perfil de clientes.

2. Es posible crear un evento en base a recursos naturales, artísticos o históricos, si bien será recomendable optar por utilizar aquellos recursos que posean una singularidad y autenticidad especialmente relevante para potenciar la imagen del destino, como por ejemplo la Feria del Jamón, en Jerez de los Caballeros, robustece la imagen de la comarca[] como productora de jamón. Pero en cualquier caso, el ingrediente más importante será la creatividad para poner en valor los recursos y transformarlos en una experiencia interesante para el turista. Así por ejemplo, el evento "El cerezo en Flor" es algo más que la visita a un paisaje, sino que es también una puesta en escena y un programa de actividades complementarias al eje central del evento que son los cerezos florecidos.
3. En la programación de actividades del evento, será necesario aunar esfuerzos y coordinar la participación de todos los agentes implicados (oferta turística del destino, población de acogida y autoridades locales). En el caso de eventos en entornos rurales, el rol desempeñado por la población local en la recepción del turista cobra especial relevancia, contribuyendo a incrementar la autenticidad de la experiencia que se le ofrece al turista (Ambientación de los comercios, Oferta de productos típicos, Potenciación gastronomía tradicional, Muestras de artesanía local, etc.).
4. Cuando sea posible, se procurará tangibilizar la calidad del evento, con sellos de calidad, como por ejemplo, la fiesta del cerezo en flor, utilizando la denominación "*fiestas de interés turístico regional*" como argumento al comunicar el evento. De este modo se consigue enfatizar la identidad territorial marco del evento (imagen del destino).
5. Otros elementos identificativos útiles para comunicar el evento y potenciar su imagen son la utilización de logotipo destino y/o evento, como por ejemplo el del Refigi d´Amigets que aparece en la figura 1.3.3.2., o mascotas, como "*Fluvi*" en el caso de la Expo de Zaragoza 2008, figura 1.3.4.1.

Figura 1.3.4.1. Mascota y logotipo de Expo Zaragoza 2008.



Fuente: Página Web Oficial de la Expo.

6. La comunicación del evento no concluye con la captación del turista, sino que abarca todo el proceso, debiendo informarse antes, durante y después del evento (Kotler, Bowen & Makens, 1996; Reich, 1997). Así, por ejemplo el evento del cerezo en flor se anuncia en Internet y en los principales medios de comunicación antes del evento, durante el evento se reparten programas del mismo en los establecimientos, y tras el evento pueden consultarse fotos y otros materiales en la página Web de la oficina de turismo del Valle del Jerte.
7. Al definir los contenidos de la información que se difundirá, es importante seleccionar aquellos mensajes que reporten verdadera utilidad al turista, por ejemplo mapas, alojamientos, precios, etc. Además, deberá tenerse presente que la comunicación debe contribuir a la sostenibilidad ambiental del evento, formando al

[] Sierra Suroeste.

turista a la vez que se le informa y persuade para que participe en el evento (Mediano, 2004).

8. Debido a que el éxito del evento pasa por la coordinación y colaboración de todos los agentes implicados, la estrategia de comunicación del evento debe comenzar por informar a la población anfitriona, a los empresarios turísticos del territorio, a los gestores locales etc., para que cada uno contribuya en la medida de sus posibilidades al desarrollo del evento. Así por ejemplo, gracias a una labor previa a la exposición las Edades del Hombre, ésta se pudo ofrecer gratuitamente, lo que indudablemente contribuyó a un éxito que benefició a todos.
9. En la aceptación del evento por parte de los turistas, un aspecto importante es el precio. Generalmente, los eventos desarrollados en el contexto del turismo rural son gratuitos, puesto que su objetivo fundamental es dinamizar el territorio. En cualquier caso, el precio a fijar para el evento dependerá de los objetivos perseguidos y de la valoración de las rentabilidades directas e indirectas esperadas con la ejecución del mismo.
10. Por último, las experiencias revisadas muestran la utilidad de Internet como canal de comunicación y comercialización. Por lo tanto, una buena manera de dar a conocer un evento es anunciarlo en centrales de reservas online especializadas en turismo rural.

1.5. CONCLUSIONES.

Mediante el diseño de eventos turísticos, los destinos generan interés atrayendo a nuevos turistas, o incitando a volver a turistas que ya lo visitaron en el pasado. Los eventos turísticos ayudan a reforzar imagen del destino, poniendo en valor los recursos disponibles.

Partiendo del contexto socioeconómico actual, de la incipiente contracción de la demanda turística, en el segmento de turismo rural se plantea un previsible exceso de oferta que presumiblemente aumentará la rivalidad competitiva del sector.

Una alternativa para mantener la participación en el mercado de un destino de naturaleza, es desarrollar una estrategia de diferenciación, ofreciendo a los turistas una experiencia distinta (Harsha, 2009). En este trabajo se plantea la posibilidad de diferenciar un destino o una empresa ubicada en el destino a través del diseño de eventos.

Tras repasar eventos de diferente naturaleza, promovidos por distintos agentes con diversos objetivos, **se extrajeron una serie de pautas para diseñar eventos de calidad** que aumenten el atractivo del destino para los turistas.

La calidad del evento es un aspecto intangible que se configura a través de aspectos tangibles: calidad de los recursos (obras artísticas, paisajes, autenticidad de la historia revivida, etc), originalidad, creatividad de la "puesta en escena", cooperación de la población de acogida para incrementar la autenticidad de la experiencia, utilidad de la información difundida, etc.

Junto con la calidad, la **sostenibilidad del evento es esencial** cuando hablamos de destinos de naturaleza, elegidos por turistas sensibilizados con el medio y deseosos de tener experiencias auténticas de contacto con una realidad distinta a la de los destinos masivos (motivaciones principales: contacto con la naturaleza, relajación, etc.). En este sentido, se ha visto como un macro-evento esboza un programa de sostenibilidad, procurando la sostenibilidad ambiental, social y económica del mismo. La cuestión a plantearse es si realmente lo consigue. Con respecto a la sostenibilidad de la Expo de Zaragoza hay opiniones muy dispares, habiendo denuncias de grupos de interés, como ecologistas en acción, que alertan sobre el impacto negativo de las obras de la Expo sobre el cauce del Ebro. En cualquier caso, la descripción de cómo la Expo trató de abordar su sostenibilidad resulta útil, puesto que aporta ideas que pueden incorporarse al diseño de un evento sostenible en el medio rural: uso de materiales biodegradables, racionalización de los consumos, concienciación de la población local, etc. Además, pone de manifiesto que **diseñar un evento en clave de sostenibilidad implica incorporar criterios sociales y ambientales a los criterios económicos que se apliquen para tomar decisiones referentes al evento.**

Otro **aspecto clave para el éxito del evento es la definición de una estrategia de comunicación acertada.** Gestionando el flujo de información sobre el evento, se logrará visitas de los turistas a un mayor número de lugares o atracciones; prolongará la duración de la visita del turista; incrementará el gasto del turista; proporcionará un mayor conocimiento de los lugares y atracciones existentes, y por tanto, existirán más posibilidades de que el turista lo dé a conocer entre sus familiares y amigos; incrementará la satisfacción del turista propiciando su regreso en el futuro; y mejorará la notoriedad de la imagen del destino.

Cuando se diseñe el evento hay que **valorar la rentabilidad inmediata, directa, y la rentabilidad indirecta, más a largo plazo,** que se logra al robustecer la imagen del destino y posicionarlo de forma diferenciada respecto a sus competidores.

Finalmente, tras el análisis exploratorio realizado, pensamos que una **futura línea de trabajo** sería elaborar una base de datos o **catálogo,** que recopile información de manera sistemática, sobre **eventos** que se desarrollen en un entorno natural, a fin de recabar experiencias útiles para los gestores de destinos que deseen dinamizarlo a través del diseño de eventos. Otra línea complementaria a la anterior, sería **cuantificar el impacto de los eventos**, que se vayan catalogando, sobre las economías locales.

1.6. BIBLIOGRAFÍA.

ALBACETE, C.A. (2004): "Calidad de servicio en alojamientos rurales". Cáp. 3, Págs. 27-45. Ed. Universidad de Granada.

CABRINI, L. (2002): "Turismo, desarrollo rural y sostenibilidad". Presentación en el VII Congreso Asociación Expertos Científicos en Turismo.

CÁNOVES, G.; HERRERA, L. y VILLARINO, M. (2005): "Turismo Rural en España: Paisajes y Usuarios, nuevos usos y nuevas visiones". Cuadernos de Turismo, Nº15, Págs. 63-76.

CASILLAS J.C.; MORENO, A.M.; OVIEDO M.A., (1995) : "El turismo alternativo como un sistema integrado: consideraciones sobre el caso andaluz" Estudios turísticos nª 125 pp. 53-74.

CHACKO, H. (2009): "Positioning a Tourism Destination to Gain Competitive Edge". Documento digital, disponible en:
http://www.hotel-online.com/Trends/AsiaPacificJournal/PositionDestination.html

COMISIÓN MUNDIAL SOBRE EL MEDIO Y EL DESARROLLO. Declaración de Tokio (1987): documento digital disponible en:
http://www.cma.gva.es/comunes_asp/documentos/legislacion/val/006000225_1.htm

FUNDACIÓN LAS EDADES DEL HOMBRE: http://www.lasedades.es/index.htm

GALÍ, N., MAJÓ, J. y VIDAL, D. (2000): "Patrimonio Cultural y Turismo: Nuevos Modelos de Promoción vía Internet". Cuadernos de Turismo, Nº6, Págs. 73-87.

GARCÍA, B. (2005): "Características diferenciales del producto turismo rural". Cuadernos de Turismo, Nº15, Págs. 113-133.

GARCÍA, E. (2002): "El Turismo Cultural en Castilla León. El caso singular de las edades del hombre." Cuadernos de Turismo, Nº10, Págs. 23-67.

INSTITUTO DE ESTUDIOS TURÍSTICOS: Encuesta de Ocupación Turismo Rural. Documento digital disponible en:
http://www.iet.tourspain.es/paginas/EOTR.aspx?seleccion=tablas&periodo=mensual&idioma=es-ES&option=otr&option2=EOTR

KOTLER, P.; BOWEN, J.; MAKENS, J. (1996): Marketing for hospitality and tourism. Ed. Prentice Hall, Upper Saddle River, N.J.

LOPES, D. (2003): "El desarrollo turístico integrado en los espacios rurales de interior: su aplicación al producto Salud". Cuadernos de Turismo, Nº11, Págs. 107-126.

MEDIANO, L. (2004): La gestión de Marketing en el Turismo Rural. Ed. Pearson Prentice Hall. Págs. 149-178.

MIDLETON, V..(1994): Marketing in Travel and Tourism. Oxford: Butter Worth, Heinemann.

MORA, A.; SUCH, M.J.; CÓRDOBA, J.L. (2001): "El desarrollo del turismo en el Parque Natural del Alto Tajo". Cuadernos de Turismo, Nº 7, Págs. 111-130.

ORTEGA, E. y RODRÍGUEZ, B. (2005): "Importancia de la comunicación de los destinos turísticos internacionales: la percepción de los turistas extranjeros en los destinos españoles". Cuadernos de Turismo, Nº15, Págs. 169-188.

PERLES, J.F. (2006): "Análisis del impacto económico de eventos: una aplicación a fiestas populares de proyección turística". Cuadernos de Turismo, Nº17, Págs. 147-166.

PRESENZA, A.; MORENO, S.; y LOBO, A. (2007): "El funcionamiento de las organizaciones de marketing de destinos regionales. El caso de la organización regional de turismo Arruzzo Promozione Turismo (Italia)". Cuadernos de Turismo, Nº19, Págs. 133-146.

RICO, M. (2005): "El turismo como nueva fuente de ingresos para el medio rural de Castilla y León". Cuadernos de Turismo, Nº16, Págs. 175-195.

REICH, A. Z. (1997): Marketing Management for the Hospitality Industry. A Strategic Approach. Ed. John Wiley & Sons.

ROYO, M. y Serarols, C. (2005): "el turismo rural-cultural: un modelo de gestión del marketing turístico a nivel local basado en la medida de la imagen del destino". Cuadernos de Turismo, Nº16, Págs. 197-222.

Turismo Sostenible Zaragoza: http://www.turismososteniblezaragoza.com/Noticia%202.html

VÁZQUEZ, R. y DÍAZ, A.M. (2001): "El conocimiento de las expectativas de los clientes: una pieza clave en la calidad del servicio en turismo". Documento de Trabajo 114/96. Facultad de Ciencias Económicas y Empresariales. Universidad de Oviedo. Disponible en: http://euturismogijon.en.eresmas.com/euturismo/estudio_uo.pdf

EMPRENDIMIENTOS EN MICROEMPRESAS MEXICANAS Y SUS PERSPECTIVAS DE CONSOLIDACIÓN

MICHELLE TEXIS FLORES
mtexis@uabc.mx
MARTÍN A. RAMÍREZ URQUIDY
martinramirez@uabc.mx
Facultad de Economía y Relaciones Internacionales
Universidad Autónoma de Baja California (México)

Resumen

En México la importancia de la microempresa, considerada como aquella que cuentan con diez empleados como máximo, deriva del gran número de establecimientos y personal que labora en ellas, representan más del 90% de la estructura empresarial del país. Sin embargo, la relativa facilidad con la que se insertan en el mercado tiene su contraparte en la alta tasa de cierre. En este contexto el trabajo aborda el tema de la empresarialidad en microempresas orientado a conocer las posibilidades de consolidación de sus emprendimientos. Para ello, se estima la probabilidad condicional de superar la fase de desarrollo inicial de la microempresa a través de un modelo probit, considerando una muestra de 750 microempresas de bajo valor agregado. Los resultados muestran evidencia de que variables como ventas y educación del microempresario están relacionadas con la posibilidad de que las microempresa permanezcan en el mercado por más de tres años.

Palabras clave: empresarialidad, microempresas, etapa de desarrollo inicial.
Área Temática: Economía Empresarial.

Abstract

The importance of microenterprises in Mexico, considered as those with ten employees as maximum, is consequence of the big quantity of establishments and employment generated, represent more than 90% of the business structure of the country. However, the relative ease with which they enter in the market has its counterpart in the high rate of closure. This research addresses the topic of entrepreneurship at the microenterprises to identify possibilities for consolidation of its businesses. The research proposes estimating the conditional probability of exceeding the initial stage of development of microenterprises through a probit model, considering a sample of 750 low-value added microenterprises. Preliminary results show that education of microentrepreneurs and incomes are associated to the possibility of staying in the market for over three years.

Keywords: entrepreneurship, microenterprises, initial stage of development.
Thematic Area: Entrepreneurial Economics.

1. INTRODUCCIÓN

La actividad empresarial tiene impactos significativos en el desarrollo de una región. En la medida en que las empresas consolidan su participación en el mercado, su actividad se torna eficiente y contribuyen en la dinámica social económica y cultural. En México la estructura empresarial ha estado dominada por la presencia de pequeñas unidades económicas, en su mayoría generadoras de bajo valor agregado. La creciente presencia de estas empresas ha orientado los estudios empresariales y esfuerzos de política económica a entender y fortalecer la creación de nuevos emprendimientos, vigorizar su proceso de desarrollo e integrar sus actividades de forma dinámica al proceso de globalización, ya que la configuración de la

estructura empresarial se ha caracterizado por comportamientos muy disímiles que acentúan su composición heterogénea.

En este contexto resulta pertinente establecer un diagnóstico sobre la empresarialidad en microempresas de bajo valor agregado, orientado a conocer las posibilidades de consolidación de sus emprendimientos. El enfoque en las microempresas se fundamenta en los esfuerzos de emprendimiento que inician en pequeño y constituyen una base importante de los aprendizajes empresariales e industriales que a lo largo de los años forjan las grandes industrias que articulan pequeñas y grandes empresas.

Por otro lado, estudiar experiencias de emprendimientos en microempresas permite identificar patrones de comportamiento que se orienten a reflexionar y constituir elementos de soporte y desarrollo inicial de microempresas, para fortalecer la estructura empresarial de la región y el nivel de bienestar de las personas dependientes de estas actividades empresariales.

El objetivo del documento es determinar los factores que impulsan a las microempresas de bajo valor agregado a consolidarse en el mercado, a partir del análisis probabilístico. Así como analizar las experiencias de emprendimiento y desarrollo de las microempresas. El trabajo se compone de cinco partes que inician con una breve descripción de las características de las microempresas en México. La segunda parte aborda el tema de la empresarialidad enfocando el análisis a los emprendimientos de microempresas. La tercera parte muestra la metodología seguida para estimar el modelo de probabilidad planteado, seguida de los resultados y finalmente las conclusiones.

2. EMPRESARIALIDAD Y MICROEMPRESAS EN MÉXICO

Estudios como los de Audretsch y Thurik (2001) demuestran la importancia que la creación de nuevas empresas tiene en términos del crecimiento económico de una región, al hallar una relación directa entre la actividad empresarial y crecimiento del PIB, acompañado de una reducción en el nivel de desempleo. La incorporación de nuevas empresas al mercado incentiva procesos de competencia que benefician al consumidor.

La literatura sobre empresarialidad considera tres etapas del proceso emprendedor de la empresa: formación del proyecto; puesta en marcha de la empresa y; desarrollo inicial de la empresa. Esta última, considera los primeros tres años de vida de la empresa, estimados como fundamentales para su supervivencia. Durante este periodo el empresario enfrenta los retos de posicionarse en el mercado (Kantis et al, 2002). Con respecto a la formación de empresas algunos trabajos tratan de explicarla con base en la dinámica que muestran las economías. Las dos hipótesis básicas en estas perspectivas son que el desempleo incentiva la formación de empresas, y la otra que el incremento de empresas se asocia con economías en crecimiento en las que la propagación de empresas reduce el desempleo. Estas dos hipótesis son conocidas en la literatura como los efectos "Refugee" y "Shumpeter" (Audretsch, 2002). Sin embargo, percibir oportunidades de beneficio se relaciona con la capacidad de identificar elementos que impulsen o retrasen el proceso de emprendimiento y el desarrollo económico regional, de tal forma que las instituciones son un determinante substancial de la empresarialidad (Viego, 2004).

Durante la fase de gestación, las nuevas empresas operan con costos fijos y se enfrentan a procesos de aprendizaje que las coloca en puntos subóptimos debido a que no han maximizado la escala de operación. El éxito de las nuevas empresas se relaciona con superar la fase inicial, fortalecer su participación en el mercado y mantener un desempeño eficiente. En este sentido, la estrategia de desarrollo emprendedor implica no sólo identificar los estímulos para la puesta en marcha de nuevas empresas sino del diseño e implementación de mecanismos de fortalecimiento de los emprendimientos (Kantis, 2003).

Otros estudios sobre empresarialidad se basan en la teoría de la elección de ingreso, que se enfoca en las decisiones confrontadas por los individuos para iniciar un negocio y volverse empresarios. Al respecto destacan los trabajos de Evans y Leighton (1990) y Blanchflower y Meyer (1994), quienes proponen que altos niveles de desempleo se asocian con altos niveles de apertura de negocios bajo el argumento de que el costo de oportunidad de no empezar un

negocio es reducido. Por otro lado, existe también el argumento opuesto de que el desempleo alto va acompañado de un nivel de empresarialidad bajo dado que los desempleados tienden a poseer bajo nivel de capital humano y talento empresarial requerido para abrir y mantener los negocios (Lucas, 1978; Jovanovic, 1982).

Los estudios sobre creación de empresas se han concentrado en los efectos de la presencia de pequeñas, medianas y grandes empresas dinámicas en la economía debido al peso que representan en el desarrollo de la región, al ser generadoras de valor agregado, actuar bajo lógicas de exportación y desarrollar cadenas de proveedores. Sin embargo, en la configuración de la estructura empresarial existen estratos poco dinámicos caracterizados por procesos de subsistencia de valor social, económico y cultural como las microempresas.

En México el criterio comúnmente utilizado para definir a la microempresa corresponde al número de empleados, que la define como aquella unidad económica que cuentan con diez empleados como máximo. Su importancia deriva del gran número de establecimientos y personal que labora en ellas. La propagación de microempresas fue considerada como un factor atenuante de la situación vivida durante los años posteriores a la crisis de México de 1994, debido principalmente a que contribuyeron a disminuir los niveles de desempleo y apoyar la estabilidad social de la región (Carrasco, 2005). En el aspecto social destacan como estrategia de sobrevivencia y punto de partida a la empresarialidad. En el aspecto económico, su importancia radica en que generan ingreso y empleo. Aunado a ello se colocan como elemento de apoyo al desarrollo de las empresas más grandes y de procesos de encadenamiento bajo figuras de subcontratación (Morales et al, 2003).

A pesar de la relativa facilidad con que entran al mercado, derivada de las nulas barreras de entrada al mercado, su tasa de mortandad es alta. Sin embargo, existen microempresas que han demostrado tener la habilidad de permanecer en el mercado, ya que ventajas de las unidades de escala menor, como adaptación, flexibilidad, motivación y talento empresarial (North, 1993), deben aprovecharse para disminuir sus restricciones de desarrollo (Deneffe y Masson, 1994). Por las bondades mencionadas y la importancia que ha cobrado en últimos años, se ha convertido en un factores estratégicos para el desarrollo de una región (Ruiz, 1995).

En México, según datos del Censo Económico 2004 del Instituto Nacional de Estadística y Geografía (INEGI), en 2003 había 3 millones 5 mil 157 unidades económicas que generaban 16 millones 239 mil 536 empleos que corresponden al 97.3% de las unidades económicas y el 88.7% del empleo de la industria manufacturera, comercio y servicios. De ellas las microempresas absorben el 95.5%. Destaca el hecho de que tan sólo las que cuentan hasta con dos empleados concentran el 71.1%. Por su parte, las pequeñas empresas representan tan sólo el 3.5%, las medianas 0.8% y las grandes el 0.2% del total. Las micro y pequeñas empresas (MYPES) concentran el 99%.

Participación MYPES sobre el total de establecimientos (%)

	Manufactura	Comercio	Servicios	Total
MYPES	97	99	99	99
Micro	91	97	95	96
Micro (1-5)	83	93	88	88
Pequeñas	6	2	4	3

Fuente: Elaboración propia con datos de INEGI, Censos Económicos, 2004

La economía bajacaliforniana, contaba con 57,404 empresas en 1999, mientras que para 2003 éstas llegaron a 61, 812, lo cual reportó un incremento de la participación de las MYPES, de 97.5% a 97.81%, respectivamente, y el de las medianas y grandes, cayó de 2.5% a 2.19%, entre los periodos censales 1999 y 2003. Lo anterior fue paralelo a la caída en la tasa de crecimiento promedio anual del PIB estatal de 6.5% en el periodo 1993-2000 a 2.34% durante el periodo 2001-2006, y a la consecuente disminución de la participación del PIB estatal en el nacional de 3.37% en el 2000 a 3% en 2005 y 2006, nivel similar al de 1993.

Ello muestra la importancia de la microempresa en el aspecto cuantitativo, no así en el cualitativo al generar bajos niveles de valor agregado y, por tanto, bajos niveles de bienestar por

las remuneraciones que perciben los trabajadores. Si bien es claro que este tipo de empresas han sostenido un bajo nivel en la generación de valor agregado, el creciente número de establecimientos ha tenido efectos positivos en la producción, el empleo y la equidad. De acuerdo con Ruiz (1995) ello ha situado a la microempresa como parte integral del desarrollo local y nacional, con un alto valor social en el orden económico.

Dentro del estrato microempresarial existen empresas de subsistencia que en la mayoría de los casos se encuentran marginadas de la dinámica regional debido a que son emprendimientos originados fuera de la cultura de la legalidad, al no participar totalmente de la vida económica y social del territorio. La puesta en marcha de la microempresa, en ocasiones, se relaciona con la búsqueda de estabilidad económica, como fuente principal de ingreso o generación de fuentes complementarias al ingreso familiar. La pertinencia de estos emprendimientos es notable en contextos de pobreza y desempleo, ya que son un medio a través del cual los microempresarios obtienen una fuente de ingreso y se incorporan al mercado, es decir, configura una estrategia de desarrollo local integradora (Drucaroff y Padin, 2004).

La relativa facilidad que las microempresas poseen para iniciar actividades, relacionada con la baja dotación de activos fijos requerida para la puesta en marcha de la empresa, tiene su contraparte en la también alta tasa de cierre de las mismas. Lo anterior ha llevado a realizar estudios sobre las capacidades empresariales, en cuanto a la creación de empresas, establecimiento de redes de desarrollo del negocio, programas de capacitación y financiamiento de las nuevas empresas (Angelelli y Llisterri, 2003). Destaca la relevancia de estas empresas producto de su capacidad de generación de empleo, configuración de redes de proveedores y establecimiento de procesos de subcontratación (Pratten, 1991).

3. METODOLOGÍA

Se utilizó información generada por el Centro de Investigación Asistencia y Docencia a la Micro y Pequeña Empresa (CIADMYPE) de la Facultad de Economía y Relaciones Internacionales de la Universidad Autónoma de Baja California, sobre 750 emprendimientos microempresariales en los municipios de Mexicali, Tijuana, Ensenada y Tecate, Baja California. La información se ha recopilado en distintos periodos a lo largo de los últimos ocho años (Mungaray et. al, 2007). La sistematización de la información permitió estructurar una base de datos combinada de corte transversal a lo largo del tiempo, que deriva en diferentes muestras de una población objetivo.

Se aplicaron cuestionarios a las microempresas con el objetivo de conocer los factores básicos de sus emprendimientos y se les brindo asistencia por un periodo de tres meses para determinar niveles de venta. La asistencia que ofrece el CIADMYPE a través de un programa de asistencia tiene el objetivo de fortalecer el desarrollo de las microempresas de subsistencia a través de servicios no financieros.

Con la finalidad de determinar las posibilidades de permanencia en el mercado, se utiliza un modelo probit (Greene, 1998). La variable dependiente permanencia en el mercado, considera el tiempo de operación de la microempresa, toma valor de uno si ha operado por tres años o más, lo que implica que la empresa ha superado la etapa de desarrollo inicial, cero en caso contrario.

$$P_i = 1/(1+ e^{-(\beta 1+\beta 2)}) \quad (1)$$
$$Pr\,(Y_i=1) = P_i$$
$$Pr\,(Y_i=0) = (1-P_i)$$

La probabilidad de respuesta para *Y*:

$$P\,(Y=1|X_i) = 1-G\,[-(\beta_0+X\beta)] \quad (2)$$

donde:

G = es la función de distribución acumulada (FDA) de la normal estándar.

Las variables explicativas X_i son: tiempo de operación ajustado -cociente de tiempo de operación de la microempresa, y edad del microempresario, medidos en años-; ventas promedio de la microempresa, pesos deflactados a precios de 2003; sexo; nivel de estudios, primaria,

secundaria, preparatoria y licenciatura y; estado civil del microempresario, son datos a nivel individual.

La ecuación estimada es:

$$Doper=1-G(-(\beta_0+\beta_1 Toperajus+\beta_2 Ventaspr+\beta_3 Dse+\beta_4 Dprim+\beta_5 Dsec+\beta_6 Dprep+\beta_7 Dlic)) \quad (3)$$

donde:

Doper es el tiempo de operación de la empresa, variable dicotómica donde Y_i=1 si la microempresa tienen operando el mercado tres años o más; Y_i=0 si el tiempo de operación de la microempresa es de menos de tres años, *Toperajus* es el tiempo de operación ajustado es el cociente de tiempo de operación de la microempresa y edad del microempresario, variable proxy de aprendizaje. Es la proporción de tiempo que la microempresa ha operado con respecto a la edad del microempresario, *Ventaspr* corresponde al nivel de ventas mensuales promedio de la microempresa, deflactadas a precios de 2003, *Dse* es una variable dicótoma de sexo, *Dprim* es variable dicótoma, nivel de estudios del microempresario correspondiente a primaria, *Dsec* corresponde a secundaria, *Dprep* es nivel preparatoria y *Dlic* correspondiente a licenciatura, finalmente *Dcas* es el estado civil casado del microempresario.

Las variables explicativas utilizadas en el modelo corresponden al perfil del microempresario, no obstante, no se ignora el hecho de que factores como las instituciones influyen en la consolidación de emprendimientos, mostrar su relación empírica queda para estudios posteriores.

4. RESULTADOS

Los emprendimientos analizados son el resultado de iniciativas personales inducidas por obtener un complemento al ingreso familiar o generar una fuente de empleo estable, ya que los microempresarios, en promedio, tienen dos dependientes económicos. Las fuentes de financiamiento utilizadas para inicial el negocio han sido ahorros y préstamos de familiares y amigos del microempresario. Los emprendimientos se ubican en los subsector de alimentos y bebidas 50%, comercio 27%, servicios 17% y metalmecánica 6%, con destino de su producción o servicios al mercado interno. En promedio, cuentan con tres empleados, generar ventas mensuales de $11,454 y bajos niveles de valor agregado. El rango de edad de los microempresarios comprende de los 16 a los 78 años.

El 10% de los emprendimientos hay evidencia de experiencia previa, por lo que, no hay un proceso de incubación que permita adquirir conocimiento técnico. A pesar de no ser emprendedores experimentados, se han interesado por el ámbito empresarial con expectativas de permanecer en el mercado y obtener mejores niveles de bienestar. Los hombres incursionan a menor edad en la empresarialidad que las mujeres, las frecuencias son mayores en los estratos de 36 a 50 años. Resalta el hecho de que la actividad empresarial de adultos es mayor que la de los jóvenes, convirtiéndose la microempresa en oportunidad laboral para personas de edad mayor.

En 52% de los casos, los ingresos generados por la microempresa representa la principal fuente de ingreso de la familia al oscilar dichas aportaciones entre el 50% y 100%. Para el 30% de los casos, las aportaciones al ingreso familiar son de 91 a 100%. El 70% de las microempresas estudiadas operan en la informalidad, situación que les impide el acceso al financiamiento. Por lo que, la mayoría de los emprendimientos estudiados ha enfrentado el reto de insertarse y operar en el mercado con recursos propios. En este sentido apoyos no financieros como asistencia y provisión de información son relevantes para fortalecer una cultura de la legalidad que ofrezca beneficios a las microempresas que operan en la esfera formal para integrarse a la dinámica de la región.

Los resultados de la ecuación estimada (3), calculados mediante Stata se muestran en la tabla 1, exhiben el signo esperado, en el caso de tiempo de operación ajustado, ventas promedio,

nivel de estudios de primaria, sexo y estado civil del microempresario la relación es positivamente con la probabilidad de que la microempresa se mantenga operaciones en el mercado por más de tres años.

Tabla 1. Resultados econométricos

	Coeficiente	Std. Err.	Z
Toperajus	66.70234	8.278595	8.06 (0.000)[a]
Ventaspr	6.48e-06	3.14e-06	2.06 (0.039)[a]
Dse	1.232799	.6135842	2.01 (0.045)[a]
Dprim	1.148682	.4608281	2.49 (0.013)[a]
Dsec	-.2653864	.437508	-0.61 (0.544)
Dprep	-.8235295	.4276698	-1.93 (0.054)[b]
Dlic	-1.575112	.7714984	-2.04 (0.041)[a]
Dcas	.540768	.3013011	1.79 (0.073)[b]
Cons	-5.382522	.7317687	-7.36 (0.000)[a]

LR chi2(8)= 488.21 Prob > chi2= 0.0000 Log likelihood= -56.499056 Pseudo R2=0.8120
note: 9 failures and 83 successes completely determined.
(P>|z|). a. Significativa al 5%, b. Significativa al 10%.

A mayor proporción de tiempo que el microempresario se dedica a la esfera empresarial se adquiere mayor experiencia que estimula procesos de aprendizajes, lo cual explica la relación positiva entre la proporción de tiempo como microempresario y la posibilidad de que la microempresa supere la fase inicial. El costo de oportunidad de cerrar un negocio es alto cuando se ha invertido una proporción de tiempo relativamente alta, con respecto a la edad del microempresario, en el desarrollo de la empresa.

El coeficiente de la variable sexo, *Dse,* revela una relación directa con la permanencia en el mercado, por lo que, la probabilidad de exceder tres años de operación no es independiente del sexo. Por lo tanto, el sexo es una variable que determina la probabilidad con que la microempresa puede superar la fase inicial. En términos de probabilidad[1] los resultados indican que los microempresarios de sexo masculino son 7% más propensos a superar la fase inicial de la microempresa. Del total de la muestra el 55% de los microempresarios son mujeres y 45% hombres. Por su parte, las ventas de los microempresarios hombres son 57% superiores a las de mujeres empresarias. Considerando que la variable *Ventaspr* está relacionada directamente con el tiempo de operación de la empresa, implica que los hombres tienen incentivos para permanecer en la actividad empresarial. No obstante, la participación de las mujeres en la actividad microempresarial ha sido importante como complemento al ingreso familiar.

La posibilidad de superar la fase inicial de la microempresa disminuye a aumentar el nivel de estudios del microempresario, sólo en el caso de microempresario con nivel de estudios de primaria existe probabilidad de superar la fase inicial de la microempresa, la relación cambia cuando el microempresario cuenta con estudios de secundaria, preparatoria y licenciatura. Un microempresario con mayor nivel de estudios incrementa sus posibilidades de reinserción a la esfera laboral, sobretodo en el caso de que el emprendimiento sea una válvula de escape al desempleo o las oportunidades laborales se incrementan.

El estado civil del microempresario exhibe una relación positiva con microempresas que superan los tres años de operación. Los microempresarios casados son 2 veces más propensos a mantener a la microempresa operando por más de tres años.

5. CONCLUSIONES

La gran proporción que las microempresas representan en la estructura empresarial de México muestra una actividad emprendedora caracterizada por un espíritu de sobrevivencia con necesidades insatisfechas. Tales emprendimientos reducen la tensión de bajo crecimiento económico y problemas de desempleo en la región, preocupaciones latentes en el estímulo a la economía. También representan nichos de oportunidad para fortalecer el tejido empresarial, debido a que son empresas que ya se incorporaron al mercado, generan ingresos y tienen expectativas de crecimiento. En este sentido, las microempresas son factores estratégicos del desarrollo económico y mecanismo de equilibrio social. Sin embargo, la empresarialidad en microempresas se ha reducido a la creación de empresa y no al desarrollo de dicha iniciativas empresariales.

La puesta en marcha de la microempresa requiere de información sobre el mercado, adaptar servicios y/o producto, mejorando la operación de la microempresa, para conseguir clientes que genere un nicho de mercado que favorezca condiciones para que los emprendimientos prosperen en una economía de mercado. Algunas estrategias se relacionan con la provisión de servicios gratuitos para el desarrollo de negocios.

Dada la relación, establecida a partir de los resultados econométricos, entre el nivel de estudios con permanencia de la microempresa en el mercado, mejorar la actividad empresarial en la fase inicial requiere fortalecer el capital humano no escolarizado para perfeccionar la formación empresarial de los microempresarios, estimular procesos de aprendizaje empresarial que mejoren el control de sus operaciones y facilite el acceso al crédito de las microempresas con potencial de desarrollo o experiencias innovadoras.

Por su parte, la trayectoria de la microempresa en cuanto al número de establecimientos, indica que en el tema de creación de empresas y desarrollo empresarial continuará en el centro el análisis de la microempresa con temas referentes al diseño de estrategias que atenúen restricciones a su desarrollo, como capacitación, canales de información y acceso a servicios financieros, de lo contrario se configurará un estrato empresarial con pocas posibilidades de éxito.

BIBLIOGRAFÍA

ANGELELLI, P.J y LLISTERRI, J.J. (2003), El BID y la promoción de la empresarialidad: Lecciones aprendidas y recomendaciones para nuevos programas. *Informe de trabajo,* Marzo, Washington, D.C., pp. 1-19.

AUDRETSCH, D., CArree, M. A. and THURIK A. R. (2002), Does Entrepreneurship reduces unemployment? *Tinbergen Institute Discussion Paper*, Tinbergen Institute, pp. 1-12.

AUDRETSCH, D. B. and R. THURIK (2001), "Linking Entrepreneurship to Growth", OECD Science, Technology and Industry Working Papers, 2001, OECD Publishing, pp. 1-34.

BLANCHFLOWER, D. and BRUCE, M. (1994) A Longitudinal Analysis of Young Entrepreneurs in Australia and the United States, *Small Business Economics*, Vol. 6(1), pp. 1-20.

DRUCAROFF, S. y PADIN, M. (2004), Del Desempleo a la Creación de Empresas: ¿Realidad o Utopía?, pp. 1-19.

EVANS, D. S. and LEIGHTON, L. (1990), "Small Business Formation by Unemployed and Employed Workers," Small Business Economics, 2(4), pp. 319-330.

GREENE, W. (1998) Análisis Econométrico, Prentice Hall. pp. 750-755.

INEGI, censos económicos (1999) Micro, Pequeña, Mediana y Gran Empresa, México Ariel.

_____, 1995, Economía de la pequeña empresa, México, Ariel.

JOVANOVIC, B. (1982) Selection and Evolution of Industry, *Econométrica*, Vol. 50, pp. 649-670.

KANTIS, H. ISHIDA, M. y KOMORI, M. (2002), Empresarialidad en economías emergentes: Creación y desarrollo de nuevas empresas en América Latina y el Este de Asia. Banco Interamericano de Desarrollo, Departamento de Desarrollo Sostenible, División de Micro, Pequeñas y Medianas Empresas, pp. 1-115.

KANTIS, H. (2003). Creación y fortalecimiento de nuevas empresas; Estudio 1. Componente E. Coordinación del Estudio: Oficina de la CEPAL-ONU en Bs As, a solicitud de la Secretaría de Política Económica, Ministerio de Economía de la Nación, pp. 3-53.

LUCAS, R. E., (1978) On the Size Distribution of Business Firms, *Bell Journal of Economics*, Vol. 9, pp. 508-523.

MORALES, A. RENDÓN, A. y POMAR, S. (2003) Las micro y pequeñas empresas ante la apertura comercial: El caso del subsector alimentos, bebidas y tabaco en México, *Administración y Organizaciones,* julio, pp.93-112.

MUNGARAY, A. RAMÍREZ-URQUIDY, M., TEXIS, M., LEDEZMA, D.and RAMÍREZ, N. (2007), Promoting Learning in Small Entrepreneurs and Higher Education Students Through Service Learning Programs, *International Journal of Business Research*, Vol. VII, 3, pp. 10-28.

NORTH, D. (1993) Instituciones, Cambio Institucional y Desarrollo Económico, Fondo de Cultura Económica.

PRATTEN, C., (1991) The Competitiveness of Small Firms, Cambridge, Cambridge University Press.

RUIZ, C. (1995) Economía de la pequeña empresa, México, Ariel.

VIEGO, V. (2004) Empresarialidad e instituciones: dos nuevas perspectivas del análisis regional contemporáneo, Revista EURE, Vol. XXX, (90), septiembre, Santiago de Chile, pp. 41-63.

A RELEVÂNCIA DO *RATING* E DE OUTROS FATORES NA DETERMINAÇÃO DO RENDIMENTO DAS DEBÊNTURES EMITIDAS NO MERCADO BRASILEIRO

PAULO BELTRÃO FRALETTI, PROFESSOR
E-mail: paulo.fraletti@fgv.br
WILLIAM EID JÚNIOR, PROFESSOR
E-mail: william.eid@fgv.br
Escola de Administração de Empresas de São Paulo – FGV

Resumo

Compreender os fatores relevantes na precificação de debêntures é primordial para investidores e empresas interessadas em captar recursos dessa forma. A inexistência de um mercado secundário líquido de títulos privados de renda fixa restringe a análise empírica das informações. Utilizaram-se dados referentes a 67 debêntures indexadas ao CDI-over e 30 indexadas ao IGP-M, emitidas entre abril/2000 e setembro/2004. Utilizou-se o método dos mínimos quadrados ordinários para testar a influência do rating e de um conjunto de variáveis de controle na formação da taxa de remuneração dos títulos. Estimaram-se relações lineares e não-lineares entre as variáveis independentes e o spread expresso em nominal e over Treasury. Dentre as evidências encontradas, destacam-se: o rating é variável significante para os dois grupos de debêntures; algumas outras variáveis (como prazo e volume da emissão) são relevantes, especialmente para títulos corrigidos pelo IGP-M; a precificação de debêntures é imperfeita e aparentemente sujeita a fatores não-econômicos.

Palavras-Chave: mercado financeiro, debêntures, precificação, rating.
Área Temática: Economia Empresarial.

Abstract

Understanding relevant factors for debenture pricing is essential for investors and companies interested in this way of attracting funds. The inexistence of a secondary liquid market of private fixed income bonds restricts the empirical analysis of information. It was used data referring to 67 debentures index-linked to CDI-over and 30 index-linked to IGP-M, issued from April, 2000 to September, 2004. The ordinary least squares method was used in order to test the influence of rating and of a set of control variables in the formation of the bonds remunerative tax. Linear and non-linear relations were estimated between the independent variables and the spread expressed as normal and over Treasury. Among the found evidence, what can be highlighted is: the rating i a significant variable for both debenture groups; some other variables (such as time and volume of the issuing) are relevant, especially for bonds adjusted with IGP-M; debenture pricing is not perfect and is subject to non-economic factors.

Key Words: financial market, debentures, pricing, rating.
Thematic Area: Entrepreneurial Economics.

1. INTRODUÇÃO

Debêntures são instrumentos financeiros de renda fixa emitidos por empresas para captação de recursos por prazos tipicamente mais longos que os de empréstimos bancários. Dentre suas vantagens destaca-se a flexibilidade para estabelecimento de prazos, garantias e condições de pagamento, o que permite adequação dos desembolsos à capacidade de geração de caixa da companhia. As linhas de financiamento do Banco Nacional de Desenvolvimento Econômico e

Social (BNDES) são alternativas vantajosas em termos de custo e prazo, mas a dificuldade e demora para aprovação do crédito, as restrições à utilização dos recursos e a imposição de monitoramento intensivo representam barreiras insuperáveis para muitos potenciais tomadores.

Ao definir as características de determinada emissão as empresas podem ainda incluir cláusulas como as de participação nos lucros, conversibilidade e repactuação. Tais facilidades tornaram a debênture o instrumento mais utilizado no mercado de capitais brasileiro para captação de recursos de médio e longo prazos, tendo sido utilizada por centenas de empresas de diferentes setores da economia, seja para financiamento de novos projetos, securitização de ativos, reestruturação de passivos financeiros ou obtenção de capital de giro, dentre outros objetivos.

As debêntures brasileiras são usualmente emitidas com valor nominal unitário de R$ 10.000 e prazo entre 3 e 8 anos. Quanto às garantias oferecidas, são classificadas segundo as seguintes espécies:

- Real - Garantidas por bens integrantes do ativo da emissora, ou de terceiros, que ficam indisponíveis para negociação por força de hipoteca, penhor ou anticrese;
- Flutuante - Asseguram privilégios sobre o ativo da emissora, não impedindo, entretanto, a negociação dos bens que compõem esse ativo;
- Quirografária - Não oferecem nenhum tipo de garantia, concorrendo em igualdade de condições com os demais credores quirografários no caso de liquidação;
- Subordinada - Asseguram privilégio aos debenturistas, em caso de liquidação, somente em relação aos acionistas da emissora no ativo remanescente.

Ao contrário do mercado internacional, onde comumente são remuneradas a uma taxa fixa, o histórico de instabilidade econômica consolidou no mercado brasileiro a prática de remuneração das debêntures com base em um indexador mais um *spread* nominal. No final de setembro de 2004 o estoque de títulos registrados no Sistema Nacional de Debêntures (SND) apresentava o seguinte perfil: 45% indexados ao CDI-over, 24% ao IGP-M, 18% prefixados, 6% indexados à TJLP, 3% ao INPC e 4% aos demais indexadores (variação cambial, taxa Anbid, participação nos resultados, IGP-DI, TR e outros).

Em tese o rendimento de um título de dívida emitido por empresa privada incorpora: i) o retorno assegurado por um título livre de riscos de mesmo prazo; ii) um adicional de crédito (*credit spread*) correspondente à inadimplência esperada (prêmio de inadimplência) e à incerteza quanto à efetiva inadimplência (prêmio de risco). A determinação objetiva do *spread* de crédito apropriado exigiria um esforço analítico considerável e não justificável para a maioria dos agentes econômicos. Por esse motivo é usual que emissores e emissões sejam avaliados por agências especializadas independentes e classificados segundo sistemas simplificados de categorização (*rating*) que em tese indicam a probabilidade de inadimplência (*default*) de cada um deles.

O presente estudo visa acrescentar informações que aprofundem o entendimento e ajudem a promover o desenvolvimento do mercado nacional de debêntures. O trabalho segue na linha trilhada por Mellone, Eid Júnior e Rochman (2002), e sua contribuição está centrada no cuidadoso tratamento dado à base de dados expandida, descrita no item 2., e na introdução de duas inovações metodológicas: i) análise da rentabilidade diferencial entre títulos privados e títulos públicos de mesmo prazo (*spread over Treasury*) em acréscimo à análise baseada no *spread* nominal; ii) verificação da hipótese de não linearidade na relação *spread-rating*, mais consistente com o modelo de Merton, em adição ao modelo linear empregado em estudos precedentes.

2. REVISÃO BIBLIOGRÁFICA

A relação *rating* versus diferencial de rendimento entre *corporate* e *government bonds* foi amplamente estudada na literatura internacional. Estudos mais recentes como os de Duffee (1998), Elton et alii (2001), Sironi e Gabbi (2002) e Kose et alii (2003) analisaram essa relação, controlando para outras variáveis, e concluíram que a precificação de debêntures é

influenciada, dentre outros, pelos seguintes fatores: *rating* (inclusive as diferenças entre classificações de agências distintas e as diferenças entre *ratings* atribuídos à emissora e à sua dívida), prazo, liquidez, garantias, setor da economia ao qual a emissora do título pertence, estágio do ciclo econômico, prêmio de risco (correlacionado ao risco sistemático do mercado de ações) e regime de taxação dos títulos.

No Brasil poucos trabalhos foram realizados para investigar a relação entre a variável *rating* e a remuneração das debêntures. Mellone, Eid Júnior e Rochman (2002) utilizaram o modelo de regressão linear (considerando o *rating* uma variável escalar) e empregaram dados relativos a 36 emissões indexadas ao CDI-over e 43 indexadas ao IGP-M lançadas entre 01/jan/2000 e 31/mar/2002. As regressões múltiplas com variáveis de controle prazo, tipo de garantia, classe (conversível ou não), forma (escritural ou nominal) e ano de emissão apresentaram baixo poder explicativo e reduzida significância estatística em praticamente todas as variáveis. Apenas a regressão simples do *spread* nominal de títulos corrigidos pelo CDI em função do *rating* mostrou evidência de que esta variável contribui para a formação do preço.

Um estudo mais amplo foi desenvolvido por Sheng e Saito (2005) que também controlaram a influência das principais características escriturais da emissão e introduziram o indicador EMBI-Brazil como *proxy* para a expectativa do mercado internacional em relação ao ambiente econômico brasileiro. Foram aplicados testes paramétricos e não paramétricos, bem como os métodos dos mínimos quadrados ordinários e dos momentos generalizados, aos dados relativos a 138 debêntures emitidas no período janeiro/1999 a dezembro/2002 (89 indexadas ao CDI-over e 49 indexadas ao IGP-M). Dentre as evidências encontradas pelos autores estacam-se: i) o *rating* afeta o *spread* nominal independentemente do indexador; ii) o ambiente econômico, o volume e o setor do emissor são importantes variáreis de controle.

3. COMPOSIÇÃO DA AMOSTRA E DESCRIÇÃO DOS DADOS

O Sistema Nacional de Debêntures (SND) foi criado em 1988 pela Associação Nacional das Instituições do Mercado Financeiro (ANDIMA) e pela CETIP - Câmara de Custódia e Liquidação para processar eletronicamente o registro, a custódia, a negociação e a liquidação financeira das operações com debêntures. Uma ampla gama de informações referentes a esses instrumentos financeiros pode ser extraída de seu portal internet www.debentures.com.br. Não existe, entretanto, um registro integrado com todas as informações referentes à época da emissão de cada título. A base geral de dados históricos disponível a qualquer instante contém somente as informações mais recentes de cada variável. Caso tenha havido, no decorrer da vida do título, revisão de *rating* pelas agências de crédito ou mudança no *spread* contratual em razão de repactuação, por exemplo, consta na base o dado mais recente. Foi necessário, portanto, um extenso trabalho de recomposição, com consultas aos prospectos e a diversas outras fontes do próprio SND.

O processo de *bookbuilding* consiste na verificação, junto aos investidores, da demanda pelas debêntures em diferentes níveis de taxa de juros. Esta consulta permite a identificação de uma taxa de remuneração de mercado consistente com as características do ativo ofertado. A informação sobre *spread* contida no prospecto inicial de cada debênture foi substituída pelo resultado de *bookbuilding* nas emissões em que esse mecanismo foi utilizado.

Para o investidor em instrumento de renda fixa o prazo relevante para determinação da remuneração adequada é o tempo em que os recursos financeiros ficam indisponíveis e sujeitos ao risco de inadimplência. A repactuação é um mecanismo tipicamente brasileiro mediante o qual a remuneração da debênture pode ser adequada, periodicamente, às condições vigentes no mercado. A adoção dessa cláusula transforma, efetivamente, um título nominalmente de longo prazo em um instrumento de menor duração, pois é conferido ao investidor o direito de ter seu papel resgatado caso não concorde com as novas condições oferecidas pela emissora.

Adicionalmente, muitas debêntures não prometem reembolso do valor nominal em uma única parcela na data de vencimento contratual mas prevêem um plano de amortização. Nesses

casos foi calculada a data média financeira de vencimento do título (ponderada pelos valores reembolsados a cada data de amortização).

Foi definida então a "data de vencimento virtual" do título respeitando-se a seguinte ordem de precedência: i) a data da primeira repactuação dos papéis com esta cláusula; ii) a data média financeira de vencimento dos títulos com plano de amortização; iii) a data de vencimento contratual. Para efeito do presente estudo a variável Prazo, em anos, foi calculada pela divisão por 365 do número de dias corridos entre a data do início da distribuição e a "data de vencimento virtual" do título. Dentre as debêntures analisadas cerca de 23% tiveram prazo estabelecido pela primeira repactuação e 42% pela data média de vencimento.

A base de dados coletada junto ao SND incluiu todos os títulos registrados junto à Comissão de Valores Mobiliários (CVM) entre o início do segundo trimestre de 2000 e o final do terceiro trimestre de 2004. A data inicial do período foi condicionada pela existência de séries de rendimentos dos títulos públicos federais no mercado secundário (necessárias para cálculo dos *spreads over Treasury*), divulgadas diariamente pela ANDIMA desde abril de 2000.

Essa base inicial, compreendendo 233 códigos de registro, foi inicialmente reduzida a 212 pela eliminação de papéis não indexados ao depósito interfinanceiro de curtíssimo prazo CDI-over ou ao índice geral de preços de mercado IGP-M. A exclusão dos títulos desprovidos de *rating* inicial resumiu a amostra a 191 códigos.

O tratamento dos dados prosseguiu com a eliminação das debêntures enquadradas em pelo menos uma das seguintes categorias:

- Com prazo (conforme a definição acima) inferior a um ano;
- Sem indicação de colocação ou oferta pública, ou com registro de um único subscritor;
- Objeto de renegociação de dívida ou com evidência de preço de emissão fora do mercado;
- Com valor total emitido inferior a R$ 30 milhões;
- Contendo alguma cláusula de opcionalidade (debêntures resgatáveis e conversíveis);
- Emitida por empresa controlada por instituição financeira. Sarkar (2003) e outros relataram peculiaridades nos títulos emitidos pelo setor financeiro. No Brasil esses papéis tendem a não ser precificados para o mercado, sendo tipicamente adquiridos por empresas coligadas, clientela de varejo ou fundos administrados do próprio grupo financeiro.

Adicionalmente, foram consolidados sob um único código dois conjuntos de operações atípicas de CESP (duas séries, cada uma contendo 18 títulos *zero coupon* com vencimentos mensais sucessivos e demais características idênticas, registradas sob 18 diferentes códigos mas equivalentes a um único instrumento com amortização mensal), bem como dois pares de emissões da CPFL Energia e um da Nova Dutra com códigos diferentes mas em tudo mais idênticas. Restaram, finalmente, 67 debêntures indexadas ao CDI e 30 indexadas ao IGP-M.

Para construção das séries de *spread over Treasury* (ver item 4.1.), foram obtidos junto à ANDIMA os rendimentos diários das séries de títulos públicos federais indexados à taxa over Selic (Letras Financeiras do Tesouro ou LFT) e ao IGP-M (Notas do Tesouro Nacional - Série C ou NTN-C).

4. VARIÁVEIS ESTUDADAS

As variáveis selecionadas para estudo atendem a dois requisitos: i) evidência na literatura internacional de relevância para a precificação de títulos privados; ii) existência de séries de dados confiáveis.

a. Variável dependente: *spread* nominal ou *spread over Treasury*

Os sessenta e sete (67) títulos que compõem a amostra de debêntures indexadas ao CDI-over Cetip podem ser classificados, segundo o tipo de juro, em duas categorias: i) taxa fixa

(CDI+) - acrescida multiplicativamente ao fator de capitalização correspondente a 100% do CDI-over acumulado no período; ii) percentual do CDI (%CDI) - calculado pela acumulação dos fatores de capitalização diários correspondentes a um percentual da taxa efetiva de cada CDI-over do período.

Os rendimentos das vinte e seis (26) emissões com *spread* na forma de percentual da taxa over (%CDI) foram ajustados para que essas pudessem ser agrupadas às quarenta e uma (41) com *spread* aplicado sobre o indexador (CDI+). A conversão consiste na estimação da taxa fixa (Sprd) que iguale os dois fatores de rentabilização, operação equivalente à solução da seguinte equação: (1+Sprd) * □(1+CDI) = □(1+%CDI). O processo depende, entretanto, do patamar das taxas de mercado e exige uma projeção das taxas over futuras. Como estimativa de CDI acumulado foi utilizada a taxa prefixada de prazo igual ao do título, extraída da curva livre de riscos (ver Fraletti (2004)) na data em que teve início a distribuição da debênture, metodologia largamente utilizada no mercado financeiro.

A preponderância na amostra de títulos com atualização tipo CDI+ (41) sobre os corrigidos na modalidade %CDI (26) deve-se à forte preferência pela primeira forma, na proporção de mais de quatro para um, nos anos 2001 e 2002. O tradicional equilíbrio entre eles foi restaurado a partir de 2003. No trabalho de Mellone, Eid Júnior e Rochman (2002) os títulos indexados a percentual do CDI-over (%CDI) foram simplesmente excluídos da amostra analisada.

Foram definidos então dois tipos de variável dependente: i) *spread* nominal sobre o indexador, tal como estabelecido nos termos contratuais da debênture (também utilizada por Mellone et alii (2002) e Sheng e Saito (2005)); ii) *spread* acima do *spread* nominal de títulos do governo de características similares, denominado *spread over Treasury* na literatura internacional. Esta última é economicamente relevante porque investidores podem escolher, a qualquer instante, entre um título teoricamente livre do risco de inadimplência e um título corporativo com risco de crédito. A opção pelo segundo dependerá da percepção de risco e do prêmio oferecido.

Para cálculo de cada *spread over Treasury* foi estimada, por interpolação, a rentabilidade teórica de um título público federal de prazo igual ao da debênture na data de início da distribuição desta. Para os papéis indexados ao CDI-over foram utilizados os rendimentos das Letras Financeiras do Tesouro (LFT), e para os indexados ao IGP-M os das Notas do Tesouro Nacional - Série C (NTN-C). A apuração da variável baseou-se na divisão do fator de capitalização do *spread* nominal da debênture pelo fator de capitalização do *spread* nominal do título público.

b. Variável teste: *rating* da emissão

Como o *rating* atribuído pela agência reflete, em tese, a probabilidade de inadimplência da debênture, espera-se que o *spread* acima do indexador aumente com a queda de categoria na escala que classifica os investimentos segundo a qualidade de crédito.

Apesar da baixa relevância do mercado brasileiro de dívida privada, os títulos que compõem a amostra foram avaliados por cinco diferentes agências nacionais e internacionais: Standard & Poor's (63 emissões), Fitch Atlantic Ratings (60), Moody's Investors (25), SR Rating (17) e Austin Rating (7). Dentre as noventa e sete (97) debêntures estudadas, cinqüenta e nove (59) foram analisadas por dois serviços de crédito e oito (8) por três deles.

A divergência de opinião entre duas ou mais agências na classificação de uma mesma emissão introduz o problema de seleção para o pesquisador, assim como para o investidor. Partindo-se da atribuição de uma escala quantitativa ordinal (entre 0,5 e 10,0) aos códigos da agência internacional Standard & Poor's, foi criado um quadro de equivalência entre os *ratings* das cinco agências de crédito representadas na amostra (Tabela 1 a seguir). No caso das agências Standard & Poor's e Moody's foi respeitada a relação sugerida por Fabozzi (2000).

TABELA 1
Valores atribuídos aos *ratings* das cinco agências de crédito

S&P = Standard & Poor's ; FAR = Fitch Atlantic Ratings
MOD = Moody's Investors ; SRR = SR Rating ; AUS = Austin Rating

S&P		FAR		MOD		SRR		AUS	
brAAA	10,0	AAA	10,0	Aaa.br	10,0	brAAA	10,0	AAA	10,0
brAA+	9,5	AA+	9,5	Aa1.br	9,5	brAA+	9,5	AA+	9,5
brAA	9,0	AA	9,0	Aa2.br	9,0	brAA	9,0	AA	9,0
brAA-	8,5	AA-	8,5	Aa3.br	8,5	brAA-	8,5	AA-	8,5
brA+	8,0	A+	8,0	A1.br	8,0	brA+	8,0	A+	8,0
brA	7,5	A	7,5	A2.br	7,5	brA	7,5	A	7,5
brA-	7,0	A-	7,0	A3.br	7,0	brA-	7,0	A-	7,0
brBBB+	6,5	BBB+	6,5	Baa1.br	6,5	brBBB+	6,5	BBB+	6,5
brBBB	6,0	BBB	6,0	Baa2.br	6,0	brBBB	6,0	BBB	6,0
brBBB-	5,5	BBB-	5,5	Baa3.br	5,5	brBBB-	5,5	BBB-	5,5
brBB+	5,0	BB+	5,0	Ba1.br	5,0	brBB+	5,0	BB+	5,0
brBB	4,5	BB	4,5	Ba2.br	4,5	brBB	4,5	BB	4,5
brBB-	4,0	BB-	4,0	Ba3.br	4,0	brBB-	4,0	BB-	4,0
brB+	3,5	B+	3,5	B1.br	3,5	brCCC	2,0	B+	3,5
brB	3,0	B	3,0	B2.br	3,0	brCC	1,5	B	3,0
brB-	2,5	B-	2,5	B3.br	2,5	brC	1,0	B-	2,5
brCCC	2,0	CCC+	2,0	Caa1.br	2,0	brD	0,5	CCC+	2,0
brCC	1,5	CCC	2,0	Caa2.br	2,0			CCC	2,0
brC	1,0	CCC-	2,0	Caa3.br	2,0			CCC-	2,0
brD	0,5	CC+	1,5	Ca.br	1,5			CC+	1,5
		CC	1,5	C.br	1,0			CC	1,5
		CC-	1,5					CC-	1,5
		C+	1,0					C+	1,0
		C	1,0					C	1,0
		C-	1,0					C-	1,0
		DDD+	1,0						
		DDD	1,0						
		DDD-	1,0						
		DD+	0,5						
		DD	0,5						
		DD-	0,5						
		D+	0,5						
		D	0,5						
		D-	0,5						

Devido à limitação das séries históricas de dados praticamente não existem no Brasil estudos empíricos conclusivos a respeito da qualidade relativa das agências de crédito, e nem mesmo sobre a consistência entre *rating* e inadimplência efetiva. O trabalho de Sheng e Saito (2005) indica, entretanto, que para determinação do *spread* nominal não há diferença significativa entre *ratings* de agências nacionais e internacionais ou entre o maior e o menor *rating* atribuídos à mesma emissão.

A solução adotada no presente estudo consistiu na criação de uma série baseada na importância relativa de cada agência, de acordo com o número de avaliações realizadas por ela na amostra em estudo. O *rating* numérico atribuído a cada emissão corresponde ao da Standard & Poor's quando este existe, pois tal agência avaliou 45 das debêntures indexadas ao CDI-over (67% do total) e 18 das indexadas ao IGP-M (60% do total). Na sua ausência utilizou-se

primeiramente o *rating* da Fitch Atlantic Ratings (em 21 casos), e sucessivamente o da Moody's Investors (8), o da SR Rating (1) e o da Austin Rating (4).

A adequação dessa escolha foi averiguada mediante a repetição de algumas das regressões apresentadas no item 6 (resultados empíricos). Nas análises adicionais foi empregada uma série alternativa de *ratings*, construída com base na média aritmética dos valores numéricos atribuídos a cada emissão. Apesar de as conclusões gerais não se alterarem, tais regressões apresentaram estatísticas marginalmente mais favoráveis, com impacto mais evidente no nível de significância da variável teste *rating*. As médias, ao capturarem todas as informações disponíveis, parecem melhor refletir o processo decisório dos investidores.

c. Variáveis de controle

É possível que duas debêntures de mesmo *rating*, porém com características contratuais heterogêneas, ofereçam rendimentos diferentes. A rentabilidade de títulos de renda fixa, assim como a de qualquer investimento, depende dos riscos envolvidos. Conseqüentemente, outras variáveis devem ser controladas para possibilitar o estudo dos efeitos provocados pela variável teste sobre a variável dependente.

Investidores demonstram preferência pela liquidez, e ativos com mercado secundário limitado ou inexistente tendem a oferecer prêmios de rendimento mais elevados. O alongamento do prazo de um instrumento privado aumenta seus riscos de mercado (maior sensibilidade do valor presente do *spread* fixo a mudanças no custo de oportunidade do dinheiro) e de crédito (maior incerteza quanto à capacidade futura de pagamento do devedor e problemas de *agency*). A literatura internacional mostra que a estrutura temporal dos *spreads* de crédito tende a ser ascendente.

No caso de um evento de inadimplência (*default*), as garantias disponíveis podem ser determinantes na velocidade e na taxa de recuperação do valor em haver. Existe também evidência empírica internacional a respeito da influência que o estado da economia exerce sobre a capacidade de pagamento dos agentes privados, e conseqüentemente sobre os *spreads* no mercado de crédito.

Utilizamos as seguintes variáveis de controle:

i. Volume: Na ausência de um mercado secundário desenvolvido, foi utilizado como *proxy* para liquidez o valor das debêntures ofertadas para distribuição (produto da quantidade emitida pelo valor nominal unitário).

j. Prazo

k. Garantias

Estado da economia: Segundo o modelo de Merton, discutido no item 4.1., o *spread* de crédito para determinada categoria de emissores é função do valor de mercado dos ativos dessas empresas, e este valor depende das condições vigentes no ambiente econômico em que operam. Sheng e Saito (2005) utilizaram o indicador *Emerging Markets Bond Index - Brazil* (EMBI-Brazil) para capturar a expectativa do mercado internacional com relação ao ambiente econômico brasileiro, seu efeito sobre o risco-Brasil e conseqüentemente sobre o rendimento de títulos brasileiros no momento da emissão.Pode-se argumentar, entretanto, que indicadores domésticos mais abrangentes estão sujeitos tanto aos choques externos quanto aos de origem interna. Como *proxy* para o estado da economia foi adotado o índice de ações da Bolsa de Valores de São Paulo (Ibovespa). É de se esperar que durante as crises econômicas, com o índice Bovespa em baixa, a taxa de captação aumente.

5. METODOLOGIA

Para estabelecer a relação entre as variáveis de estudo foram estimados os parâmetros de modelos lineares múltiplos empregando-se a metodologia de regressão por mínimos quadrados ordinários, também utilizada nos estudos de Mellone et alii (2002) e Sheng e Saito (2005). Segundo o modelo da firma de Merton, entretanto, a relação entre *rating* e *spread* de crédito de títulos privados não é linear. Esta hipótese também foi testada no presente trabalho.

a. Modelo de Merton

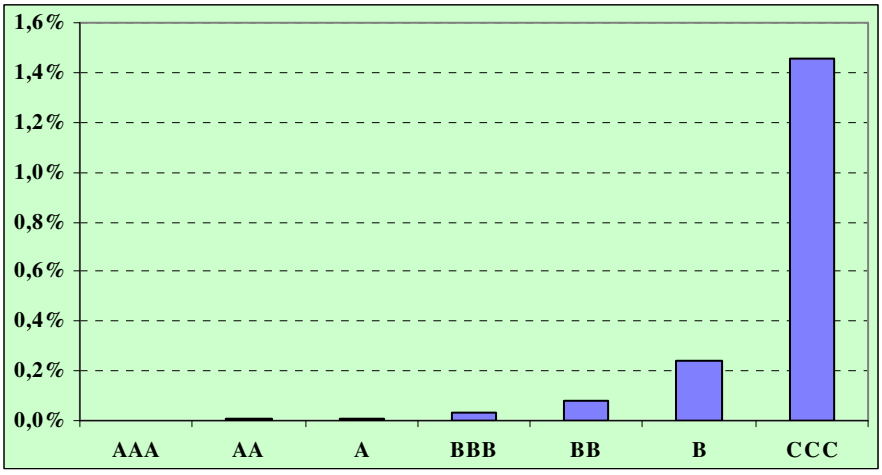
O modelo de avaliação de ativos proposto originalmente por Merton (1974) fundamenta-se no fato de que no regime legal de responsabilidade limitada os acionistas têm o direito de quitar o passivo mediante a entrega da empresa aos seus credores, mesmo que o valor dos ativos seja inferior ao da dívida. O direito de falência corresponde, portanto, a uma opção de venda (*put*) concedida pelos financiadores aos acionistas, e o valor de mercado de um título corporativo equivale ao valor de um título com características similares mas livre do risco de inadimplência menos o prêmio dessa opção.

Partindo da estrutura simplificada de Merton, na qual o passivo da empresa é constituído pelo capital acionário e por um único *zero coupon bond* com prazo residual de um ano, foram simulados os *spreads* de crédito teóricos das dívidas de sete empresas que diferem apenas quanto ao grau de alavancagem adotado, e por decorrência quanto ao *rating* atribuído ao título de sua emissão.

Tal como no modelo de Black & Scholes assume-se que o valor futuro dos ativos da empresa é log-normalmente distribuído e segue um movimento browniano geométrico. Para um título de determinado *rating*, com inadimplência esperada dada pelo Quadro 1, obtém-se a distância estatística padronizada entre os valores correntes dos ativos e da dívida (denominada *distance to default* no modelo KMV). Para implementação das simulações foram assumidos ativos da empresa com rendimento esperado de 23,8%a.a. e volatilidade de 30%a.a. no cálculo do valor de face da dívida de cada uma das sete empresas.

A taxa livre de riscos (16%a.a.) adotada no cálculo do valor das opções de venda por Black & Scholes foi também utilizada na precificação do *zero coupon bond* livre de riscos de um ano de prazo. O Gráfico 2 a seguir apresenta os *spreads* de crédito (diferencial de rendimento entre títulos corporativos e título livre de riscos) estimados para cada *rating*.

GRÁFICO 2
Spread* de crédito teórico de título *zero coupon* corporativo em função do *rating



Os valores do gráfico, consistentes com o modelo de Merton, mostram uma relação claramente não linear. Os parâmetros da função SPREAD = EXP[□ + □*RATING], proposta como modelo para as regressões não lineares do trabalho, foram estimados com base nos *spreads* simulados e o R-quadrado ajustado da regressão foi 0,98. Já a função linear SPREAD = □ + □*RATING apresentou aderência bem menor, com R-quadrado ajustado de 0,40. Convém salientar, entretanto, que a análise de uma amostra contendo apenas títulos *investment grade* (isto é, com *rating* BBB ou superior) deve apresentar diferença bem menos marcante entre o poder explicativo dos modelos linear e exponencial.

Nas debêntures indexadas ao CDI-over o *spread* nominal representa essencialmente o adicional pelo risco de crédito, pois o indicador utilizado é uma taxa de mercado que incorpora tanto a expectativa quanto à inflação futura quanto a taxa real de juros da economia. Isto não ocorre nos títulos corrigidos pelo IGP-M, um índice de inflação. Nestes o *spread* acima do indexador deve incorporar o adicional de crédito e a taxa real de juros. As estatísticas descritivas no início dos itens 5.1 e 5.2 ilustram tal diferença. O *spread* nominal médio das debêntures do segundo grupo é 11,60%, mas apenas 1,67% no primeiro grupo.

O *spread over Treasury* captura, em tese, apenas o risco de crédito das emissões privadas. Na amostra estudada, entretanto, observa-se adicional médio mais elevado (1,61%) nas debêntures indexadas ao IGP-M que nas corrigidas pelo CDI-over (1,34%). Fato notável se considerada a qualidade relativa das duas amostras: títulos IGP-M com *rating* médio 6,90 e títulos CDI-over com *rating* médio 6,62. Cançado e Falcão (2002) observaram o mesmo e atribuíram tal resultado à adoção do CDI-over como *benchmark* pela maioria dos investidores e ao baixo interesse por títulos corrigidos por índices de preços face à expectativa de inflação futura declinante que reinava no período estudado.

Para atender à demanda dos investidores não é raro que uma empresa emita simultaneamente duas séries de debêntures semelhantes com indexadores distintos. Por estarem sujeitas ao mesmo risco de crédito deveriam exprimir *spread over Treasury* idêntico, caso contrário seria possível realizar arbitragem utilizando contrato de swap CDI versus IGP-M. A última coluna da Tabela 2 abaixo é uma ilustração adicional da imprecisão no processo de precificação das debêntures brasileiras evidenciada nas análises apresentadas adiante.

TABELA 2
Emissões de Debêntures com Séries CDI-over e IGP-M – Abr/00 a Set/04

Empresa	Código	Indexador	Registro na CVM	Prazo (anos)	*Spread* Nominal	*Spread over Treas.*	Diferença IGPM-CDI
CPFL Energia	DRFT21/41	CDI-over	18-mai-00	2,45	0,73%	0,72%	-0,05%
	DRFT11/31	IGP-M		2,45	13,00%	0,67%	
AES Sul	AESL11	CDI-over	5-mar-01	5,07	2,00%	1,97%	0,16%
	AESL21	IGP-M		5,07	12,90%	2,13%	
Marlim	MRLM11	CDI-over	7-mar-01	4,73	0,34%	0,31%	-0,16%
	MRLM21	IGP-M		4,73	10,70%	0,15%	
Braskem	COPE10	CDI-over	26-dez-01	2,76	3,31%	3,05%	-0,43%
	COPE20	IGP-M		2,76	13,25%	2,61%	
Sabesp	SBSP15	CDI-over	14-mai-02	1,38	1,85%	1,30%	0,79%
	SBSP25	IGP-M		1,38	13,25%	2,09%	
CPFL Paulista	PALF12	CDI-over	30-jul-04	4,91	1,60%	1,14%	0,67%
	PALF22	IGP-M		4,91	9,80%	1,81%	
Baesa	BESA11	CDI-over	29-set-04	6,84	0,85%	0,30%	0,91%
	BESA21	IGP-M		7,34	9,55%	1,22%	

A especificação da variável dependente *spread over Treasury* possibilitou a consolidação dos dados relativos às 97 debêntures em uma única amostra, e a realização de algumas regressões adicionais não apresentadas nos itens 5.1 e 5.2. Os resultados mostraram elevação generalizada no grau de significância estatística dos parâmetros estimados, porém baixo impacto no poder explicativo (R-quadrado ajustado) dos modelos.

Análise baseada na segmentação da amostra de títulos indexados ao CDI-over segundo o tipo de juro remuneratório, taxa fixa (CDI+) e percentual do CDI (%CDI), não revelou informação relevante para o estudo.

As variáveis utilizadas no estudo são definidas a seguir:

SPNM – *Spread* Nominal: taxa de juros expressa em percentual ao ano na convenção do Banco Central (BaCen), isto é, capitalização composta por dias úteis na base de um ano de 252 dias úteis. O fator baseado neste *spread* é acrescido, de forma multiplicativa, ao fator de atualização especificado na escritura da debênture (CDI-over ou IGP-M).

SPOT – *Spread over Treasury*: taxa de juros expressa em percentual ao ano na convenção do BaCen. Representa a remuneração adicional do título privado, na forma multiplicativa dos fatores, em relação a um hipotético título com características idênticas às da debênture porém emitido pelo Tesouro Nacional.

RTNG – *Rating*: Valor entre 0,5 e 10, com pontuação máxima atribuída às emissões com risco de crédito mínimo.

VOLM – Volume: Valor total ofertado em bilhões de reais (com três casas decimais).

PRAZ – Prazo: Calculado segundo critérios descritos no item 2. e expresso em anos (com duas casas decimais).

DSUB – *Dummy* que assume valor 1 quando a garantia da debênture é subordinada (característica encontrada em 19% das indexadas ao CDI e em 20% das indexadas ao IGP-M).

DFLU – *Dummy* que assume valor 1 quando a garantia da debênture é flutuante (18% das indexadas ao CDI e 7% das indexadas ao IGP-M).

DREA – *Dummy* que assume valor 1 quando a garantia da debênture é real (26% das indexadas ao CDI e 17% das indexadas ao IGP-M). As três variáveis *dummies* assumem valor zero quando a garantia é quirografária, isto é, não há preferência (37% das indexadas ao CDI e 56% das indexadas ao IGP-M).

IBVS – Índice Bovespa, expresso em milhares de pontos (com três casas decimais), no fechamento do mercado na data em que teve início a distribuição da debênture.

6. RESULTADOS EMPÍRICOS

a. Debêntures indexadas ao CDI-Over

O Anexo 1 no final do artigo apresenta as principais informações a respeito de cada debênture indexada ao CDI utilizada no estudo. Uma análise estatística descritiva das variáveis encontra-se resumida nas Tabelas 3 e 4 a seguir. Dentre as 67 emissões há uma considerável dispersão do *spread* nominal (SPNM), com valores entre 0,34%a.a. e 5,00%a.a., embora 53 delas tenham remuneração inferior a 2,00%a.a.

TABELA 3
Estatística Descritiva de Variáveis Selecionadas – Debêntures indexadas ao CDI

Variável	SPNM	SPOT	RTNG	VOLM	PRAZ	IBVS
Média	**1,67**	**1,34**	**6,62**	**0,288**	**2,95**	**16,523**
Mediana	**1,60**	**1,05**	**7,50**	**0,150**	**2,81**	**15,259**
Desvio padrão	**1,02**	**1,04**	**2,16**	**0,291**	**1,47**	**4,010**
Mínimo	**0,34**	**-0,65**	**0,50**	**0,030**	**0,90**	**9,884**
Máximo	**5,00**	**4,88**	**9,50**	**1,300**	**8,21**	**23,385**

TABELA 4
Distribuição de Freqüência de Variáveis Selecionadas – Total 67 Emissões

SPNM	*Freqüência*
0 a 1	24
>1 a 2	29
>2 a 3	8
>3 a 4	3
>4 a 5	3

RTNG	*Freqüência*
0 a 2	8
>2 a 4	1
>4 a 6	10
>6 a 8	37
>8 a 10	11

PRAZ	*Freqüência*
0 a 2	21
>2 a 4	31
>4 a 6	12
>6 a 8	2
>8 a 10	1

SPOT	*Freqüência*
-1 a 0	2
>0 a 1	30
>1 a 2	23
>2 a 3	6
>3 a 4	3
>4 a 5	3

VOLM	*Freqüência*
0,0 a 0,3	46
>0,3 a 0,6	12
>0,6 a 0,9	5
>0,9 a 1,2	3
>1,2 a 1,5	1

IBVS	*Freqüência*
9 a 12	6
>12 a 15	25
>15 a 18	14
>18 a 21	6
>21 a 24	16

Como o rendimento acima da taxa over normalmente aproxima-se de zero nos títulos públicos federais, os valores da variável *spread over Treasury* (SPOT) são apenas um pouco inferiores aos da variável SPNM. Destaca-se ainda a ocorrência de dois valores negativos nessa variável, um da empresa Duratex (DURA12, a -0,32%, com início da distribuição em 13/12/2002 e seis subscritores registrados) e um da empresa Pão de Açúcar (PCAR15, a -0,65%, com início da distribuição em 14/11/2002 e sem indicação do número de subscritores). Além da possibilidade de serem emissões "privadas" (não precificadas no mercado), foram emitidas no auge da crise que expôs a percepção de risco de inadimplência nos títulos público federais.

Embora a amostra contenha uma ampla dispersão de *ratings* (RTNG), com valores entre 0,5 e 9,5, pouco mais de 50% das emissões foram classificadas na faixa entre 6,5 e 8,0 (ou seja, entre BBB+ e A+ na escala da agência Standard & Poor's) e 84% na categoria *investment grade* (com *rating* maior ou igual a 5,5). Quanto ao volume (VOLM), embora existam emissões na casa de um bilhão de reais mais de dois terços delas não superam o valor nominal total de 300 milhões e a mediana é igual a 150 milhões.

Apesar de haver um caso de prazo (PRAZ) superior a 8 anos, há uma forte concentração de títulos (aproximadamente 40%) na faixa entre dois e três anos e na categoria inferior a dois anos (aproximadamente 30% da amostra). A variável índice Bovespa (IBVS) mostra que o período analisado inclui momentos de crise (com o índice abaixo dos 10.000 pontos no final de 2002) e de otimismo (particularmente 2004, com o índice superando os 23.000 pontos).

i. Variável dependente *Spread* Nominal (SPNM)

A Tabela 5 mostra que o modelo proposto apresenta um poder explicativo muito baixo, com R-quadrado ajustado de 0,13. Apesar disto, a constante (Interseção) e a variável teste *rating* (RTNG) são estatisticamente significantes ao nível de 1%, e as variáveis de controle garantia subordinada (DSUB) e índice Bovespa (IBVS) aos níveis de 5% e 10% respectivamente. A variável teste apresenta o sinal esperado, isto é, quanto mais elevado o *rating* (menor a probabilidade de inadimplência) menor o *spread*. As outras duas, porém, apresentam sinais que conflitam com as expectativas: segundo o modelo de Merton o *spread* deveria cair com a elevação dos valores dos ativos (assumindo-se constante a alavancagem das empresas, o que poderia não ser o caso na amostra); títulos subordinados oferecem menor garantia ao investidor e deveriam render mais (o parâmetro da variável DSUB deveria ter sinal positivo).

TABELA 5
Resultado da Regressão Múltipla Completa – SPNM de debêntures indexadas ao CDI

Variável	*Coeficiente*	*Erro padrão*	*Estat. t*	*Probab.*
Interseção	2,173	0,713	3,046	0,3%
RTNG	-0,181	0,059	-3,067	0,3%
VOLM	0,504	0,474	1,065	29,1%
PRAZ	-0,087	0,093	-0,941	35,1%
DSUB	-0,692	0,335	-2,064	4,3%
DFLU	0,104	0,373	0,279	78,1%
DREA	0,017	0,320	0,052	95,9%
IBVS	0,055	0,033	1,694	9,6%
Estatísticas da Regressão				
R-quadrado		0,222		
R-quadrado ajustado		0,129		
Erro padrão		0,948		
Observações		67		

A insignificância estatística e/ou o sinal contraditório dos parâmetros das variáveis *dummies* (DSUB, DFLU e DREA) pode ser conseqüência do limitado número de observações em cada uma das categorias. Como observado no item 3.3.3., entretanto, as cláusulas contratuais de garantia eram pouco relevantes nos processos falimentares sob o regime legal que vigorou no Brasil durante o período do estudo, e ainda que assim não fosse tal característica já teria sido considerada na atribuição do *rating* da debênture. Decidiu-se, portanto, pela re-estimação da regressão sem essas variáveis.

A Tabela 6 mostra que de fato a decisão tem impacto marginal nos resultados da regressão, com destaque para a redução do R-quadrado ajustado de 0,13 para 0,09 e para a elevação da significância estatística da variável de controle Ibovespa (IBVS) ao nível de 5%.

TABELA 6
Resultado da Regressão Múltipla sem *Dummies* – SPNM debêntures indexadas ao CDI

Variável	*Coeficiente*	*Erro padrão*	*Estat. t*	*Probab.*
Interseção	1,848	0,638	2,896	0,5%
RTNG	-0,170	0,060	-2,838	0,6%
VOLM	0,614	0,443	1,388	17,0%
PRAZ	-0,108	0,088	-1,234	22,2%
IBVS	0,066	0,032	2,034	4,6%
Estatísticas da Regressão				
R-quadrado		0,144		
R-quadrado ajustado		0,089		
Erro padrão		0,970		
Observações		67		

Eliminando-se as variáveis não significativas deste último modelo, volume (VOLM) e prazo (PRAZ), observa-se que o modelo reduzido apresenta R-quadrado ajustado ainda mais baixo e que a variável Ibovespa torna-se insignificante ao nível de 10% (Tabela 7 abaixo).

TABELA 7
Resultado da Regressão Múltipla Reduzida – SPNM de debêntures indexadas ao CDI

Variável	*Coeficiente*	*Erro padrão*	*Estat. t*	*Probab.*
Interseção	1,797	0,613	2,932	0,5%
RTNG	-0,130	0,056	-2,325	2,3%
IBVS	0,044	0,030	1,470	14,6%
Estatísticas da Regressão				
R-quadrado		0,101		
R-quadrado ajustado		0,073		
Erro padrão		0,979		
Observações		67		

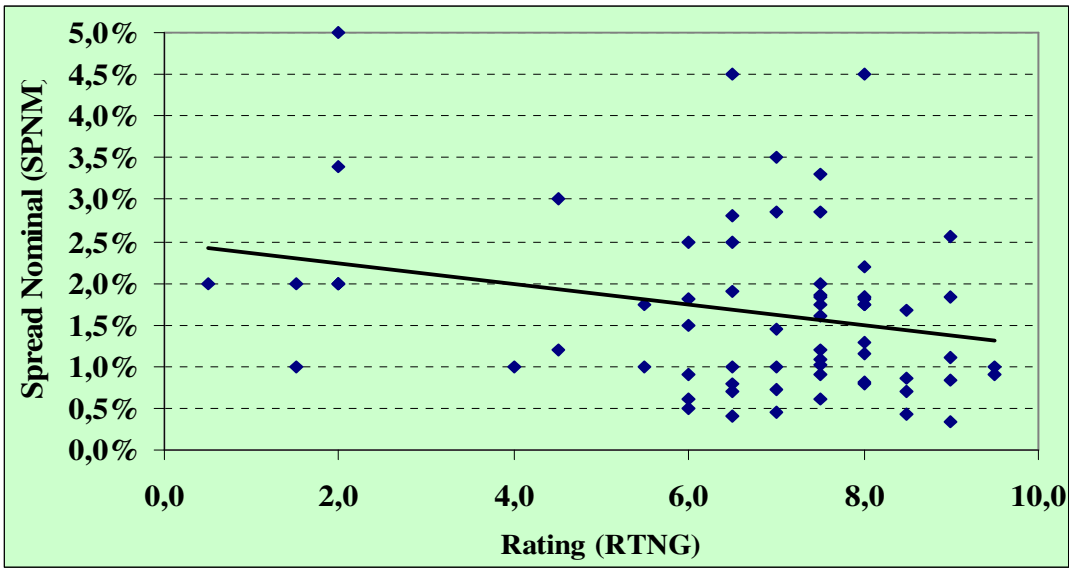
O mais simples dos modelos, baseado apenas na variável teste *rating* (RTNG), confirma que ela é sempre estatisticamente significante e apresenta o sinal esperado (Tabela 8).

TABELA 8
Resultado da Regressão Simples – SPNM de debêntures indexadas ao CDI

Variável	*Coeficiente*	*Erro padrão*	*Estat. t*	*Probab.*
Interseção	2,493	0,392	6,365	0,0%
RTNG	-0,125	0,056	-2,217	3,0%
Estatísticas da Regressão				
R-quadrado		0,070		
R-quadrado ajustado		0,056		
Erro padrão		0,987		
Observações		67		

O Gráfico 3 mostra a relação inversa entre *spread* nominal e *rating* na amostra estudada. Mais importante, porém, ele salienta a elevada dispersão na remuneração de debêntures que receberam um mesmo *rating*. O título BRKM11 de janeiro de 2004 da empresa Braskem, por exemplo, foi emitido com *rating* A+ (equivalente ao valor 8,0) e remuneração de CDI-over mais 4,50%. De acordo com a regressão simples apresentada na Tabela 8 o título deveria pagar apenas 1,49% acima do CDI, ou 2,42% se considerado o modelo completo da Tabela 5.

GRÁFICO 3
Relação entre *Spread* Nominal e *Rating* – Debêntures indexadas ao CDI



O processo *stepwise* apresentado não confirma qualquer dos modelos de estimação para o *spread* nominal de debêntures indexadas ao CDI, significando que as variáveis de controle sugeridas na literatura e aqui utilizadas não são capazes de explicar a dispersão observada no Gráfico 3. Ainda que os investidores reconheçam a relevância do *rating* e de algumas dessas outras variáveis, é possível que o processo de precificação no ainda pouco evoluído mercado brasileiro de debêntures esteja sujeito a fatores não econômicos e até mesmo espúrios.

A hipótese de não linearidade, mais consistente com o modelo de Merton descrito no item 4.1., foi verificada através do ajuste das mesmas séries de dados a um modelo de regressão do tipo SPNM = EXP [□+ □1*RTNG + □□*VOLM + □□*PRAZ + □□*IBVS]. Os resultados da análise, resumidos na Tabela 9 abaixo, mostram estatísticas em todos os casos desfavoráveis quando comparadas com as do modelo linear equivalente apresentado na Tabela 6.

TABELA 9
Resultado da Regressão Exponencial Múltipla – SPNM de debêntures indexadas ao CDI

Variável	*Coeficiente*	*Erro padrão*	*Estat. t*	*Probab.*
Interseção	0,654	0,392	1,668	10,0%
RTNG	-0,093	0,037	-2,521	1,4%
VOLM	0,154	0,272	0,566	57,4%
PRAZ	-0,062	0,054	-1,151	25,4%
IBVS	0,026	0,020	1,321	19,1%
Estatísticas da Regressão				
R-quadrado		0,107		
R-quadrado ajustado		0,049		
Erro padrão		0,596		
Observações		67		

O modelo exponencial simples, correspondente ao linear apresentado na Tabela 8, apresentou R-quadrado ajustado de 0,058 e *rating* com significância estatística ao nível de 2,8%.

ii. Variável dependente *Spread over Treasury* (SPOT)

A análise mediante a substituição da variável dependente *spread* nominal (SPNM) pela variável alternativa *spread over Treasury* (SPOT), mantidas as demais séries da base de dados, produz regressões com estatísticas similares às das anteriormente apresentadas mas com R-quadrado ajustado e significância da maioria dos parâmetros marginalmente mais elevados.

TABELA 10
Resultado da Regressão Linear Múltipla – SPOT de debêntures indexadas ao CDI

Variável	*Coeficiente*	*Erro padrão*	*Estat. t*	*Probab.*
Interseção	1,168	0,609	1,918	6,0%
RTNG	-0,210	0,057	-3,664	0,1%
VOLM	0,630	0,422	1,493	14,1%
PRAZ	-0,133	0,084	-1,596	11,6%
IBVS	0,108	0,031	3,483	0,1%
Estatísticas da Regressão				
R-quadrado		0,258		
R-quadrado ajustado		0,210		
Erro padrão		0,926		
Observações		67		

A Tabela 10 acima apresenta os resultados de uma regressão que pode ser comparada ao modelo representado na Tabela 6. A variável teste *rating* e a variável de controle Ibovespa tornam-se estatisticamente significantes ao nível de 0,1% mas a significância da constante cai para o nível de 10%, patamar este ainda não alcançado pelas variáveis volume e prazo. Apesar da elevação do R-quadrado ajustado (de 0,09 para 0,21), o procedimento não reduz de forma significativa a aleatoriedade na precificação de debêntures indexadas ao CDI-over identificada no item 5.1.1..

b. Debêntures indexadas ao IGP-M

O Anexo 2 no final do artigo contém as principais informações a respeito de cada debênture indexada ao IGP-M utilizada no trabalho. As Tabelas 11 e 12 a seguir apresentam uma análise estatística descritiva das variáveis estudadas. Há uma considerável dispersão de *spread* nominal (SPNM) dentre as 30 emissões da amostra, com valores entre 8,00%a.a. e 14,50%a.a., mas a remuneração de 26 delas está contida no intervalo 9,00% - 13,50%.

TABELA 11
Estatística Descritiva de Variáveis Selecionadas – Debêntures indexadas ao IGP-M

Variável	SPNM	SPOT	RTNG	VOLM	PRAZ	IBVS
Média	**11,60**	**1,61**	**6,90**	**0,299**	**5,20**	**16,295**
Mediana	**11,25**	**1,95**	**7,50**	**0,275**	**4,91**	**14,435**
Desvio padrão	**1,71**	**1,16**	**2,21**	**0,213**	**2,30**	**4,452**
Mínimo	**8,00**	**-0,72**	**1,50**	**0,075**	**1,38**	**10,140**
Máximo	**14,50**	**3,61**	**10,00**	**0,800**	**9,92**	**23,245**

Como a variável *spread over Treasury* (SPOT) reflete apenas o risco de crédito privado, apresenta desvio-padrão (1,16%) inferior ao do *spread* nominal (1,71%) que captura também alterações no estado da economia ao longo do período de estudo. Salienta-se aqui a ocorrência de três valores negativos na série, dois referentes à Itá Energética (ITAE11 e ITAE21, a -0,72%, com início da distribuição em 23/03/2001 e vinte cinco subscritores em cada série) e o terceiro à Suzano Bahia Sul (SUZB13, a -0,15%, com início da distribuição em 20/08/2004 e cinqüenta e dois subscritores). Ao contrário dos títulos indexados ao CDI-over com SPOT negativo, entretanto, essas debêntures tiveram um número considerável de subscritores e não foram emitidas em período de crise.

TABELA 12
Distribuição de Freqüência de Variáveis Selecionadas – Total 30 Emissões

SPNM	*Freqüência*
7,5 a 9,0	**1**
>9,0 a 10,5	**9**
>10,5 a 12,0	**7**
>12,0 a 13,5	**10**
>13,5 a 15,0	**3**

RTNG	*Freqüência*
0 a 2	**3**
>2 a 4	**0**
>4 a 6	**6**
>6 a 8	**14**
>8 a 10	**7**

PRAZ	*Freqüência*
0 a 2	**2**
>2 a 4	**8**
>4 a 6	**12**
>6 a 8	**4**
>8 a 10	**4**

SPOT	*Freqüência*
-1 a 0	**3**
>0 a 1	**7**
>1 a 2	**5**
>2 a 3	**12**
>3 a 4	**3**

VOLM	*Freqüência*
0,05 a 0,20	**12**
>0,20 a 0,35	**11**
>0,35 a 0,50	**3**
>0,50 a 0,65	**0**
>0,65 a 0,80	**4**

IBVS	*Freqüência*
9 a 12	**3**
>12 a 15	**14**
>15 a 18	**4**
>18 a 21	**0**
>21 a 24	**9**

Embora diferente da amostra de debêntures indexadas ao CDI, nota-se também uma ampla dispersão da variável *rating* (RTNG). Apesar de oscilar entre 1,5 e 10,0, cerca de 87% dos títulos enquadram-se na categoria *investment grade* (*rating* maior ou igual a 5,5). Nesta amostra o volume (VOLM) médio e a mediana estão mais próximos (cerca de 300 milhões de reais), é menor a freqüência de emissões de pequeno porte e a maior delas não supera o valor nominal de 800 milhões. É menor, portanto, o desvio-padrão da variável.

O prazo (PRAZ) médio das debêntures IGP-M (5,20 anos) é claramente superior ao das indexadas ao CDI (2,95), atingindo o máximo de quase 10 anos. A maior concentração (40%) encontra-se na faixa entre quatro e seis anos. Durante os períodos em que a variável índice Bovespa (IBVS) registrava menos de 15.000 pontos foram emitidos proporcionalmente mais títulos corrigidos pelo IGP-M (57% da amostra) que indexados ao CDI-over (apenas 46% colocados nesse tipo de ambiente).

i. Variável dependente *Spread* Nominal (SPNM)

Também no caso da amostra aqui estudada, com um número de observações ainda mais limitado em cada uma das categorias de garantia, a eliminação das variáveis *dummies* tem pequeno impacto na análise. A Tabela 13 apresenta os resultados da regressão linear múltipla que descreve a rentabilidade nominal (SPNM) em função das variáveis *rating* (RTNG), volume (VOLM), prazo (PRAZ) e índice Bovespa (IBVS).

TABELA 13
Resultado da Regressão Linear Múltipla – SPNM de debêntures indexadas ao IGP-M

Variável	*Coeficiente*	*Erro padrão*	*Estat. t*	*Probab.*
Interseção	17,317	1,035	16,738	0,0%
RTNG	-0,247	0,094	-2,645	1,4%
VOLM	0,621	1,075	0,578	56,8%
PRAZ	-0,364	0,092	-3,971	0,1%
IBVS	-0,142	0,051	-2,770	1,0%
Estatísticas da Regressão				
R-quadrado		0,693		
R-quadrado ajustado		0,643		
Erro padrão		1,018		
Observações		30		

O poder explicativo desse modelo, com R-quadrado ajustado igual a 0,64, é muito superior ao de seu equivalente para títulos indexados ao CDI-over (ver Tabela 6). Com exceção da variável volume todos os parâmetros da regressão são altamente significantes. A variável teste *rating* apresenta o sinal esperado, pois à medida que aumenta a probabilidade de pagamento da emissão (maior RTNG) reduz-se o *spread* de crédito: 0,25% para cada ponto de elevação no *rating* (equivalente, por exemplo, a um *upgrade* de A- para A+).

O fato de o *spread* nominal diminuir com o aumento do prazo (coeficiente PRAZ negativo) é consistente com a inclinação invertida freqüentemente observada na estrutura temporal do rendimento de títulos públicos indexados ao IGP-M (função da expectativa de redução futura na taxa real de juros da economia). Segundo o modelo o alongamento do prazo em um ano reduz o *spread* em 0,36%. Ao contrário do verificado na amostra de títulos indexados ao CDI, a variável IBVS tem coeficiente negativo e aderente ao modelo de Merton que associa *spreads* menores a valores maiores nos ativos das empresas (ver item 4.1.). O modelo sugere que a elevação de 1.000 pontos no índice Bovespa provoca redução de 0,14% no custo de captação das empresas.

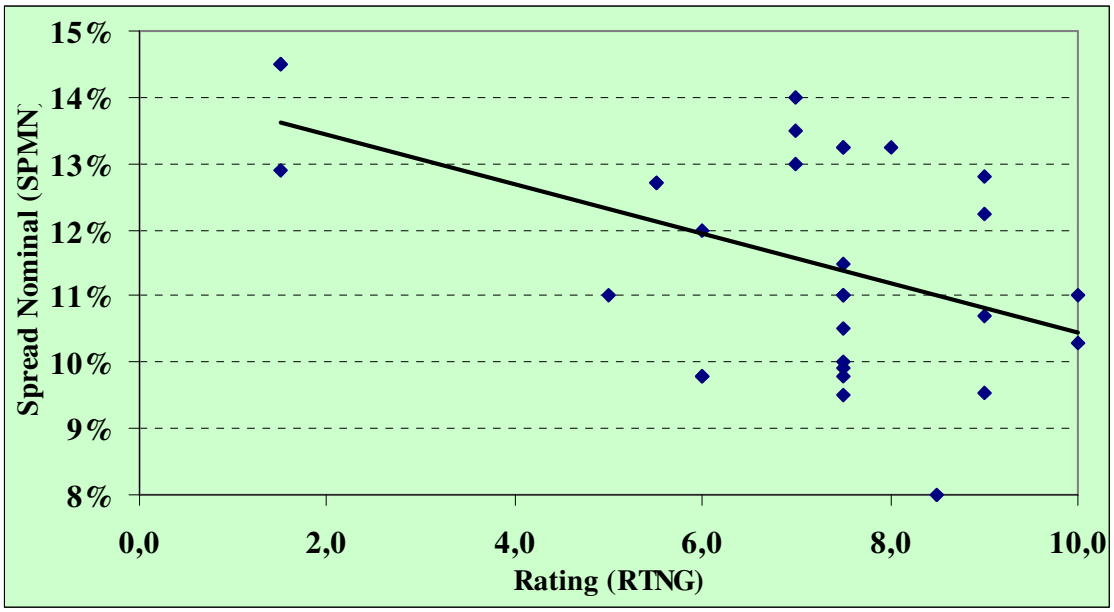
O baixo R-quadrado ajustado do modelo baseado apenas na variável teste *rating* (Tabela 14) ressalta a importância das variáveis de controle prazo e índice Bovespa na explicação do *spread* nominal de emissão das debêntures indexadas ao IGP-M.

TABELA 14
Resultado da Regressão Linear Simples – SPNM de debêntures indexadas ao IGP-M

Variável	*Coeficiente*	*Erro padrão*	*Estat. t*	*Probab.*
Interseção	14,167	0,924	15,334	0,0%
RTNG	-0,372	0,128	-2,915	0,7%
Estatísticas da Regressão				
R-quadrado		0,233		
R-quadrado ajustado		0,205		
Erro padrão		1,520		
Observações		30		

O Gráfico 4 destaca, além da relação inversa entre *spread* nominal e *rating*, a alta dispersão na remuneração de debêntures de mesmo *rating* já observada no estudo dos títulos indexados ao CDI. São focados, a título ilustrativo, os papéis NDUT11 e NDUT21 da concessionária Nova Dutra, emitidos em agosto de 2004 com *rating* A (equivalente ao valor 7,5) e remuneração de 9,50%. De acordo com o modelo retratado na Tabela 13 essas debêntures deveriam render 11,04% acima do índice de inflação.

GRÁFICO 4
Relação entre *Spread* Nominal e *Rating* – Debêntures indexadas ao IGP-M



A interferência de fatores não econômicos ou espúrios na precificação das debêntures talvez mascare a eventual diferença entre o modelo linear retratado na Tabela 13 e o modelo não linear apresentado na Tabela 15 abaixo, este mais consistente com o modelo de Merton e do tipo SPNM = EXP [□+ □1*RTNG + □□*VOLM + □□*PRAZ + □□*IBVS].

TABELA 15
Resultado da Regressão Exponencial Múltipla – SPNM debêntures indexadas ao IGP-M

Variável	*Coeficiente*	*Erro padrão*	*Estat. t*	*Probab.*
Interseção	2,944	0,091	32,530	0,0%
RTNG	-0,020	0,008	-2,455	2,1%
VOLM	0,053	0,094	0,564	57,8%
PRAZ	-0,032	0,008	-4,018	0,0%
IBVS	-0,013	0,004	-2,936	0,7%
Estatísticas da Regressão				
R-quadrado		0,696		
R-quadrado ajustado		0,647		
Erro padrão		0,089		
Observações		30		

ii. Variável dependente *Spread over Treasury* (SPOT)

A Tabela 16 a seguir apresenta os resultados da regressão linear baseada nas mesmas séries de dados utilizadas na geração da Tabela 13, substituindo-se apenas a variável dependente original *spread* nominal (SPNM) pela variável alternativa *spread over Treasury* (SPOT).

TABELA 16
Resultado da Regressão Linear Múltipla – SPOT de debêntures indexadas ao IGP-M

Variável	*Coeficiente*	*Erro padrão*	*Estat. t*	*Probab.*
Interseção	2,093	0,880	2,380	2,5%
RTNG	-0,152	0,079	-1,913	6,7%
VOLM	2,321	0,914	2,540	1,8%
PRAZ	-0,321	0,078	-4,119	0,0%
IBVS	0,094	0,043	2,170	4,0%
Estatísticas da Regressão				
R-quadrado		0,520		
R-quadrado ajustado		0,443		
Erro padrão		0,866		
Observações		30		

De modo geral este modelo é inferior ao baseado no *spread* nominal, embora neste caso todas as variáveis sejam estatisticamente significativas ao nível de 10%. Este resultado contrasta com o verificado nas debêntures indexadas ao CDI-over, onde o procedimento causou elevação do R-quadrado ajustado (de 0,09 para 0,21).

Além da leve deterioração na estatística t da variável teste *rating*, destaca-se a significância da variável volume (VOLM) ao nível de 2%. Emissões de maior volume caracterizam-se, em tese, por maior liquidez e menor rendimento. Contrariamente ao esperado, entretanto, o coeficiente da variável volume é positivo neste e em todos os modelos apresentados no presente trabalho. Uma possível interpretação é que emissões volumosas devem ser ofertadas com remuneração "de mercado" a fim de conquistar uma ampla base de investidores.

O sinal negativo da variável prazo (PRAZ) no modelo descritivo do *spread* nominal (Tabela 13) foi justificado com base na forma invertida da curva de rendimento dos títulos públicos indexados ao IGP-M. Essa linha de argumentação não pode ser invocada, entretanto, para explicar o sinal do parâmetro no modelo de *spread over Treasury*. A redução no *spread* de crédito de debêntures mais longas (0,32% para cada ano adicional na regressão da Tabela 16) conflita com a evidência internacional.

Neste tipo de abordagem o coeficiente do índice bovespa (IBVS) continua significativo, mas assume o paradoxal sinal positivo já observado no estudo das debêntures indexadas ao

CDI: a desaceleração da economia parece reduzir o *spread* de crédito. A aversão do mercado a títulos públicos federais na crise de 2002, quando o Ibovespa encontrava-se deprimido, estimulou a demanda por emissões de empresas sólidas a taxas relativamente baixas. O fato pode ter enviesado a amostra e causado a inesperada relação.

7. CONCLUSÕES

Em conformidade com as conclusões do trabalho de Sheng e Saito (2005), mas ao contrário de parte dos resultados de Mellone, Eid Júnior e Rochman (2002), a variável teste *rating* revelou-se estatisticamente significante na determinação da rentabilidade de debêntures indexadas aos CDI-over e ao IGP-M. Uma análise complementar mostrou a superioridade informacional do valor médio dos *ratings* de títulos avaliados por mais de uma agência, dado não utilizado na estimação das regressões descritas no trabalho. Resta avaliar qual a percepção atual dos investidores em relação à qualidade dos *ratings* em face do *default* de debêntures com nota elevada na data do evento (Cemar e Banco Santos, por exemplo).

A insignificância das variáveis *dummies* relativas às garantias escriturais pode ser decorrência do número limitado de observações nas amostras estudadas, mas pode ser atribuída à ineficácia dessas cláusulas em processos de falência brasileiros ou ao fato de a característica ter sido indiretamente capturada pelo *rating* atribuído à emissão.

A relevância das variáveis de controle volume e prazo foi constatada apenas na precificação de debêntures corrigidas pelo IGP-M, embora nem sempre apresentando o sinal indicado na literatura internacional. O índice Bovespa mostrou-se significativo na maioria dos modelos, mas o sinal do parâmetro estimado foi em alguns casos contraditório. Interessante notar que no período analisado o *spread* de crédito médio foi mais elevado nas emissões indexadas ao IGP-M, mas o tamanho da amostra e a qualidade dos dados não permitem generalizações.

[REVER] O efeito das duas variantes de análise propostas no presente estudo não pôde ser comprovado de forma inequívoca. O modelo exponencial revelou-se equivalente ao linear para debêntures IGP-M, mas inferior para as debêntures CDI. Essa semelhança de resultados poderia ser atribuída ao fato de a amostra analisada ser 85% composta por títulos *investment grade*, e como antecipado no item 4.1. nessa faixa de *rating* é menos marcante a diferença entre modelos. Já a análise com base no *spread* de crédito puro (*over Treasury*) apresentou ganhos quando aplicada às debêntures corrigidas pelo CDI-over, mas mostrou-se inferior na amostra das indexadas ao IGP-M.

O cuidadoso tratamento dispensado à construção da base de dados e a detalhada descrição da amostra representam importantes contribuições deste artigo. Análise efetuada *a posteriori* referendou, por exemplo, a decisão de se eliminarem as vinte e duas debêntures de empresas ligadas a instituições financeiras (todas indexadas ao CDI-over). Elas foram emitidas, na quase totalidade dos casos, com *spread* de crédito nitidamente inferior ao sugerido por qualquer dos modelos aqui propostos.

A conclusão geral é que a determinação da remuneração de debêntures no mercado primário brasileiro é no mínimo imperfeita, e não baseada apenas em indicadores econômicos objetivos. A tabela 2, por exemplo, sustenta essa interpretação ao revelar diferença no *spread* de crédito de pares de títulos quase idênticos emitidos simultaneamente pela mesma empresa, um indexado ao CDI-over e outro ao IGP-M. A parcela do rendimento não explicada pelos modelos pode ser atribuída, ao menos em parte, à falta de sofisticação dos participantes e ao baixo grau de transparência e concorrência na distribuição desses papéis. É possível também que a amostra depurada e analisada ainda contenha emissões adquiridas apenas por acionistas e antigos credores em operação de reestruturação de dívida.

REFERÊNCIAS BIBLIOGRÁFICAS

CANÇADO, P.; FALCÃO, R. (2002), *Spread de Crédito*. Unibanco Research – Fixed Income. São Paulo.

CHAIA, A. J. (2003), *Modelos de Gestão do Risco de Crédito e sua Aplicabilidade ao Mercado Brasileiro*. Dissertação (Mestrado) – FEA Universidade de São Paulo.

CROUHY, M.; GALAI, D.; MARK, R. (2000), A Comparative Analysis of Current Credit Risk Models. *Journal of Banking & Finance* v. 24, p. 59-117, 2000.

DUFFEE, G. R. (1998), The Relation Between Treasury Yields and Corporate Bond Yield Spreads. *Journal of Finance* v. 53, p. 2225-2241.

ELTON, E. J.; GRUBER, M. J.; AGRAWAL D.; MANN, C. (2001), Explaining the Rate Spread on Corporate Bonds. *Journal of Finance* v. 56, p. 247-277.

FABOZZI, F. J., (2000), *Bond Markets, Analysis and Strategies*. 4th Edition. New Jersey: Prentice Hall.

FRALETTI, P. B. (2004), *Ensaios sobre Taxas de Juros em Reais e sua Aplicação na Análise Financeira*. Tese (Doutorado) – FEA Universidade de São Paulo.

KOSE, J.; LYNCH, A. W.; PURI, M. (2003), Credit Rating, Collateral and Loan Characteristics: Implication for Yield. *Journal of Business* v. 76, p. 371-410.

MELLONE, G.; EID JÚNIOR, W.; ROCHMAN, R. (2002), Determinação das Taxas de Juros das Debêntures no Mercado Brasileiro. In: Encontro Brasileiro de Finanças 2, Rio de Janeiro. *Anais Eletrônicos*.

MERTON, R. (1974), On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance* v. 28, p. 449-470.

SARKAR, S. (2003), *Risk-free Interest Rates, the Call Feature, and Corporate Bond Yield Spreads*. Department of Finance, McMaster University.

SHENG, H. H.; SAITO, R. (2005), Determinantes de Spread das Debêntures no Mercado Brasileiro. *Aceito para publicação na RAUSP* (provavelmente no.3 ou 4).

SIRONI, A.; GABBI, G. (2002), *Which Factors Affect Corporate Bonds Pricing? Empirical Evidence from Eurobonds Primary Market Spreads*. Newfin Research Center, Bocconi University.

EFECTOS SOBRE LA RENTABILIDAD DE LA DIVERSIDAD DE GÉNERO EN LOS CONSEJOS DE ADMINISTRACIÓN. COMPARATIVA EUROPEA

FRANCISCO JAVIER JIMENO DE LA MAZA
E-mail: jjimeno@efc.uva.es
MERCEDES REDONDO CRISTÓBAL
E-mail: mredondo@efc.uva.es
Departamento de Economía Financiera y Contabilidad
Universidad de Valladolid

Resumen

En los últimos tiempos son distintas las contribuciones empíricas que contrastan para diferentes ámbitos geográficos la hipótesis de que la diversidad de género en los consejos de administración ejerce una influencia positiva sobre la rentabilidad empresarial. Los argumentos teóricos para suponer que tal asociación es significativa remiten principalmente al impacto económico del buen gobierno corporativo y al reflejo en los resultados de la responsabilidad social empresarial. La evidencia previa señala que la falta de diversidad existente en las cúpulas de poder corporativo se debe a una generalizadamente escasa representación femenina. En este marco, el presente trabajo utiliza una muestra de sociedades anónimas pertenecientes a diez países europeos, incluyendo empresas españolas, para analizar si la composición heterogénea en términos de género de los consejos de administración tiene efectos relevantes sobre distintas medidas financieras de rentabilidad de las firmas, calculadas a partir de la información contable que presentan las compañías.

Palabras clave: diversidad de género ~ consejos rentabilidad.
Área temática: Economía Empresarial.

Abstract

The subject of gender diversity on corporate boards has received considerable attention in the last years. Particularly, several contributions have focused on the link between a higher degree of gender diversity on director's board and a better firm performance, mostly in single country studies. There are multiple theoretical perspectives offering a way to understand why diversity affects positively to performance. Theoretical cases provide a framework in which diversity as a part of good corporate governance can lead to improve the firm performance. In the same way, the diversity issue has been broadened to include a concern to corporate social responsibility. Additionally, the lack of diversity can be associated primarily to the global phenomenon of low levels of female representation on corporate boardrooms. On this background, this paper analyses whether greater gender diversity on board of directors has significant effects on financial measures of performance, using a sample of European firms belonging to ten countries. These financial data are obtained from their annual accounts.

Thematic Area: Entrepreneurial Economics.

1. INTRODUCCIÓN.

La literatura académica relacionada con el estudio de las organizaciones ha manifestado un creciente interés durante la última década por analizar la diversidad en los consejos de administración de las sociedades mercantiles. Particularmente, una de las cuestiones tratadas con frecuencia se refiere a la influencia de la diversidad de género sobre ciertas variables económicas, fundamentalmente de rendimiento empresarial. Dentro de las características

demográficas, la diversidad de género atiende al grado de heterogeneidad que presenta la composición por sexos de los órganos de gobierno societario. En esta corriente investigadora son numerosos los planteamientos que se encuadran en el marco teórico del gobierno corporativo, considerando que una mayor diversidad constituye uno de los aspectos del buen gobierno de las empresas, y, por ello, el impacto sobre la actuación de la organización se manifestaría en tales términos, en el sentido de que las prácticas de buen gobierno conducen a una mejora en el desempeño de la firma.

Los desarrollos asociados al gobierno corporativo no suelen establecer relaciones causales directas entre diversidad y rendimiento, sino que, en general, la diversidad suele resultar indicativa de la presencia de otros factores reconocidos como conductores de la *performance* de la firma. Así, la diversidad en un consejo estaría poniendo de manifiesto, al menos, dos características. (1) En primer lugar, una base de independencia de los integrantes del órgano de gobierno de la compañía que, de acuerdo a las clásicas teorías de la agencia, ayudaría a que el consejo de administración cumpla mejor su función de control, alineando los intereses de directivos y accionistas. La composición diversa de un grupo de decisión parece ir asociada al incremento de integrantes a los que se les puede atribuir la cualidad de independencia, justificándose así la influencia sobre el desempeño de la función de monitorización y control de los consejos de administración. (2) En segundo término, una ausencia de segregación por razón de género en los consejos de administración, puesto que para todos los contextos geográficos la evidencia estadística muestra que la generalizadamente escasa diversidad que se manifiesta se debe siempre a la reducida representación de mujeres. Como la segregación vertical se correlaciona inversamente con las posibilidades que tiene una organización de disponer de una base de talento más amplia, la selección sin perjuicio y la consiguiente apertura a una mayor diversidad incrementaría el capital humano de un consejo de administración. En el campo del gobierno corporativo, estas proposiciones pueden enlazarse con las teorías de los *stakeholders* y de dependencia de recursos, lo que permite establecer hipótesis entre una mayor diversidad y un superior rendimiento de la firma.

A este respecto, conviene señalar que la confluencia de esa escasa diversidad de género que muestran las cúpulas corporativas con el hecho de la sistemática segregación vertical existente en el tejido empresarial ha ocasionado que los consejos de administración de las sociedades mercantiles se hayan convertido también en centro de atención de las políticas públicas de igualdad de oportunidades. Conforme a los estándares de las actuales sociedades del bienestar europeas, no parece aceptable que haya en la esfera de los negocios un considerable sesgo de género en posiciones de poder, cuando la paridad en los niveles educativos de grado superior, convencionalmente asociados a esos cargos, se ha conseguido ya desde hace años, con lo cual existe un número suficiente de candidatas adecuadas. Así, la acción política se dirige a que las compañías tengan unos consejos de administración más equilibrados, con el objeto de conseguir así un óptimo social, garantizando ciertos niveles de consumo de un bien preferente (la igualdad de género), que debiera satisfacerse al haber revelado sus preferencias la ciudadanía mediante los mecanismos de elección colectiva contemplados en el sistema.

Sin embargo, el punto de vista de la diversidad de género en el contexto empresarial se aparta de la visión institucional de las políticas de igualdad, pues lógicamente responde a una perspectiva orientada a la consecución de objetivos de éxito de la firma, tratando de obtener un aprovechamiento eficiente de todos los recursos, incluyendo a aquellos recursos humanos asignados a los niveles superiores del área de administración. Esta vertiente económica de la diversidad ha suscitado una progresiva atención por parte de las disciplinas financieras y contables. Así, y en lo referente a la línea de investigación que estudia los efectos económicos de la diversidad de género en puestos de poder de las organizaciones, se distinguen dos tendencias principales: (1) aportaciones que se centran en la escasa representación de las mujeres y su relación con factores corporativos, tales como las dimensiones o el sector de actividad (Fryxell y Lerner, 1989; Oakley, 2000; Singh et al, 2002); y (2) trabajos que analizan la posible influencia de una mayor diversidad sobre un rendimiento superior de la organización (Shrader et al, 1997; Erhardt et al, 2003; Carter et al, 2003) y cuyos resultados conducen a conclusiones contrapuestas.

Una última cuestión a contemplar es que el planteamiento de la diversidad de género en el marco del gobierno corporativo obliga a prestar atención al contexto cultural de cada nación, por las distintas características estructurales que presentan los regímenes de gobierno corporativo. Los estudios comparativos (LaPorta et al, 2002) muestran las importantes diferencias en cuanto a sistemas de gobierno empresarial que imperan en cada país, lo cual deriva de unas regulaciones heterogéneas que responden a una variedad de raíces, tradiciones y desarrollos históricos. Esta dispersión se manifiesta, por ejemplo, en las recomendaciones que incluyen los Códigos de Buen Gobierno de las compañías cotizadas que han venido emitiendo las instituciones nacionales reguladoras de los mercados de capitales.

De acuerdo a los razonamientos expuestos, en el presente trabajo se contrasta empíricamente si la diversidad de género existente en los consejos de administración de una muestra de sociedades mercantiles europeas incide en las medidas contables de rentabilidad de esas compañías, contemplando asimismo la significación de la variable representativa del país en que radica la sede social de la compañía. En función de tales objetivos, el trabajo se ha organizado de la siguiente forma: en el siguiente apartado se explican los desarrollos teóricos utilizados para el establecimiento de hipótesis, efectuando asimismo una revisión de los resultados de las aportaciones empíricas que nos preceden. A continuación, se describe la muestra de datos así como la metodología empleada, presentando posteriormente los resultados y procediendo a su discusión. Finalmente, se exponen las conclusiones más relevantes obtenidas.

2. MARCO TEÓRICO Y LITERATURA PREVIA.

2.1 El concepto de diversidad aplicado a los consejos de administración de las sociedades mercantiles.

El incremento del volumen de contribuciones a lo largo de la última década resulta indicativo del auge que ha experimentado la investigación enfocada a estudiar los efectos de la composición demográfica de los consejos de administración. Un precedente significativo de ese interés por analizar la diversidad de los recursos humanos en el contexto de las organizaciones se recoge en McDougall (1996), que remite a las postrimerías de los años ochenta del pasado siglo, período en el que comienzan a difundirse estadísticas laborales en Estados Unidos que incluían datos desagregados por sexos que ya ponían de manifiesto la segregación vertical existente.

El término diversidad suele emplearse en el ámbito de la literatura especializada con diferentes acepciones, muchas veces en función de factores contingentes al contexto de análisis, y que atienden principalmente al marco institucional de referencia. Desde la perspectiva de la demografía de las organizaciones, la diversidad generalmente alude al grado en el que se encuentran heterogéneamente distribuidos los atributos demográficos de los individuos que componen una unidad, grupo u organización (Pelled et al, 1999; Wagner et al, 1984). Conforme a esta concepción, cuando el foco de interés sea el consejo de administración de las sociedades por acciones, la diversidad representará la combinación variada de atributos, características y habilidades que aportan los individuos que integran el consejo (Van der Walt e Ingley, 2003). En los estudios de naturaleza empírica suele distinguirse entre dos categorías de diversidad (Pelled, 1996; Milliken y Martins, 1996): (1) la diversidad demográfica, que es la observable, basada en características fácilmente detectables, tales como sexo, raza o nivel educativo; y (2) la diversidad no observable o cognitiva, referida a atributos como el conocimiento, las habilidades, los perfiles psicológicos o las capacidades individuales. Obviamente, la difusión de bases de datos de carácter mercantil conteniendo información sobre los miembros que componen los órganos de administración en un determinado momento facilita el recuento de atributos demográficos como el género, siendo con frecuencia la variable más analizada (Rosenzweig, 1998). Por el contrario, la diversidad no observable directamente suele exigir el recurso a instrumentos de recogida de información específico, como pueden ser las encuestas. En todo caso, y como señalan, Milliken y Martins (1996), las variables demográficas suponen una

representación consistente y válida de atributos relativos a la diversidad no observable, motivo por el cual se asume frecuentemente que ambas categorías se correlacionan sistemáticamente (Peterson y Philpot, 2007; Smith, 2007; Rose, 2007).

2.2 Desarrollos teóricos relativos al impacto de la diversidad de género efectuados desde los planteamientos del buen gobierno corporativo.

El trabajo seminal de Kanter (1977) suele señalarse como una de las primeras aproximaciones en la literatura organizacional a la conjetura de que la composición en términos de género de las cúpulas de decisión empresariales puede tener un impacto significativo sobre el desempeño de la firma. Distintas proposiciones explican esta relación apoyándose en los efectos positivos que ejercería la diversidad sobre la estructura y sobre los procesos de la organización (Kesner, 1988; Bilimoria y Piderit, 1994). Aplicado específicamente al órgano de gobierno, una idea muy extendida es que la diversidad resulta importante para la dinámica de los consejos, lo cual conduce a un mejor desempeño de sus funciones. En última instancia, un buen gobierno corporativo se transmitiría al éxito en el cumplimiento de objetivos de la empresa (Hillman et al, 2002).

De acuerdo a los argumentos expuestos, las explicaciones más desarrolladas que asumen como hipótesis que el grado de diversidad de género que presenta un consejo de administración puede tener algún efecto relevante sobre los niveles de *performance* de la firma se basan en los planteamientos teóricos de mayor difusión en el estudio del gobierno corporativo. Así lo demuestra la inclusión de referencias en varias de las reformas de los Códigos de Buen Gobierno de los últimos años (Higgs Review, 2003 y Tyson Report, 2003 en Reino Unido, Código Unificado de Buen Gobierno, 2006 en España), con recomendaciones basadas en la necesidad de una apertura a la diversidad de los consejos, identificando la escasa presencia de mujeres como un síntoma de pobre gobierno corporativo. Por otro lado, la conexión entre buen gobierno corporativo y rendimiento ha supuesto una línea argumental muy destacable dentro de la literatura financiera, donde precisamente uno de los aspectos más extensamente analizados ha sido la influencia de la composición de los consejos de administración. Sin embargo, los resultados de la investigación empírica han venido ofreciendo una evidencia contradictoria respecto al impacto que una mayor diversidad ha de suponer sobre variables de rendimiento empresarial. Este hecho puede interpretarse en el sentido de que no exista una relación directa, sino que, en todo caso, el vínculo esté mediatizado por otros factores, lo cual lleva a Carter et al (2003) a sugerir que no existe un único enfoque teórico en el estudio del gobierno corporativo que proporcione una explicación completa para todos los elementos de incidencia. Por el contrario, si los factores que vinculan la diversidad con indicadores de *performance* de la empresa son contextuales, en función de las circunstancias concretas podrán ser de aplicación supuestos de diferentes teorías (Kiel y Nicholson, 2003). De acuerdo a esta idea, un enfoque que integre múltiples teorías sería el más adecuado para sustentar hipótesis acerca de la relación entre diversidad y desempeño. Así, y de entre las propuestas que la literatura sobre diversidad asume de forma recurrente, destacamos tres corrientes de amplia difusión, como son la teoría de la agencia, la teoría de los *stakeholderes* y la teoría de la dependencia de recursos.

La teoría de la agencia, una de las áreas más prolíficas de la literatura empresarial de las últimas décadas, ha sido la aproximación teórica más utilizada para justificar que una composición diversa del consejo pueda impulsar el rendimiento de la firma. En su versión más extendida se fundamenta en que la existencia de información asimétrica y de contratos incompletos conduce a conflictos de agencia entre los intereses de propietarios y directivos, conflictos que llevan asociado un coste. Determinados factores internos, como pueden ser las estructuras de gobierno corporativo, permiten la reducción de tales costes y en consecuencia un incremento del rendimiento (Core et al, 2006). Los consejos son un mecanismo clave de gobierno de cara a alinear los intereses de accionistas y directivos, ya que tienen asignada una función de control de la dirección, que será interpretada con mayor eficacia cuando los consejos presenten ciertas características entre las cuales destacaría la independencia (Forbes y Milliken, 1999).

En consecuencia, la diversidad promoverá que la función de control se ejecute más eficientemente si se acepta la hipótesis de que un consejo más heterogéneo y abierto a las minorías opera con mayor independencia que un consejo más monolítico y cerrado al debate, que actúa como si fuera una sola voz y por ello contempla menos alternativas de solución en el proceso de adopción de decisiones. De esta forma, la diversidad de género en el consejo se comporta como uno de los mecanismos de reducción de los costes asociados a los problemas de agencia, lo cual supone en última instancia incrementar el valor de la firma (Hillman y Daziel, 2003). En todo caso, la validez de las suposiciones relativas a los efectos de la diversidad que se basan en esquemas de agencia dependerá de la intensidad con la que se manifieste la hipótesis de partida de separación entre la propiedad y el control de la compañía. En este sentido, conviene mencionar una cuestión ampliamente documentada, y es que determinados países, sobre todo los de tradición basada en la ley civil, se caracterizan por un predominio de estructuras de propiedad accionarial concentrada (LaPorta et al, 2002), como pueden ser las que derivan de la presencia de redes familiares, lo cual obviamente influirá en el paradigma de gobierno corporativo aplicable a este tipo de empresas. Cuando la propiedad y la gestión de una firma recae en un mismo grupo de personas, es de esperar que los problema de agencia entre el accionariado y la dirección no sean particularmente relevantes, y, por lo tanto, el impacto de mecanismos como la diversidad en el consejo resulte escasamente significativo en la reducción de costes asociados a esos conflictos.

Hay otros desarrollos en el marco del gobierno corporativo que se enfocan hacia premisas teóricas que permiten explicar el impacto de la diversidad contemplando otros intereses además de los que provienen de la tenencia de derechos de propiedad de la compañía. La idea básica de este tipo de explicaciones parte de que la actuación de las firmas ha de tener en cuenta también a otros terceros involucrados en el entorno en el que radica la empresa. La relación con la *performance* de la compañía se establece por el hecho de la interrelación de la empresa con su entorno, el cual proporciona los recursos y del que también se obtienen los retornos. Un posible argumento que sugieren algunas aportaciones (e. g. Robinson y Dechant, 1997) es que la diversidad de género en el consejo puede ser uno de los factores que facilite ese flujo de intercambios con el entorno de una forma más eficaz. En esta línea, puede efectuarse un análisis del impacto de la diversidad en el consejo con base en la teoría de la dependencia de recursos, encuadrada en la investigación sobre comportamiento organizacional (Hillman y Dalziel, 2003). Esta teoría, utilizada cada vez con más frecuencia en el análisis de las funciones y la actuación de los consejos (Gabrielsson y Huse, 2004) desplaza el foco de la relación entre propiedad y dirección (que enfatizaba el enfoque de agencia) hacia los vínculos de la empresa con su entorno, y por ello se concentra en la función llamada de dependencia de recursos. De acuerdo a la centralidad de este rol del consejo, uno de los factores claves a considerar sería la habilidad de los miembros del órgano de gobierno para establecer vínculos externos, con el objeto de allegar recursos que se consideren estratégicos para la firma (Siciliano, 1996). La diversidad en este contexto supondría ampliar los perfiles de los componentes del consejo, y ello podría permitir un incremento de relaciones con clientes y competidores, así como ampliar el conocimiento sobre el sector de actividad en que opera la empresa o las posibilidades de acceso a fuentes de financiación; en suma, aumentar la obtención de recursos críticos que conduzcan a un mayor rendimiento de la compañía.

Otro de los enfoques utilizados desde el gobierno corporativo para justificar el impacto económico de la diversidad en el consejo es la teoría de los *stakeholders*, el cual, al igual que las ideas basadas en la dependencia de recursos, incide en la importancia de la vinculación de la empresa con su entorno, de una forma tal que a menudo esta perspectiva ha sido relacionada también con las corrientes de estudio de la responsabilidad social corporativa. Así, Fryxell y Lerner (1989) ya proponían el desarrollo de una teoría de los interesados orientada hacia la presencia de grupos demográficos minoritarios en los consejos, fundamentada en la necesidad de reconocer los intereses de otros actores vinculados a la firma (empleados, clientes, proveedores, financiadores, etc.). Atendiendo a este contexto, la diversidad de género y la incorporación de mujeres a los consejos podrían contemplarse como importantes indicadores de la responsabilidad social de una compañía, de manera tal que puedan ser interpretados como una señal dirigida a determinados terceros o a su entorno en general (Oakley, 2000; Ibrahim y

Angelidis, 1994; Webb, 2004). Además, como observan Hillman et al (2002), introducir un grado mayor de diversidad de género en el consejo permite unos procesos de gobierno más abiertos, que incrementan la reputación corporativa, considerada cada vez con mayor atención desde la perspectiva de la dirección estratégica como uno de los recursos intangibles más importantes para que las organizaciones obtengan una ventaja competitiva. De hecho, Wartick (1992) define a la reputación corporativa como la agregación de las percepciones de los *stakeholders* respecto a cómo la organización responde a sus expectativas y demandas. Así, distintas contribuciones resaltan el efecto de cuestiones sociales, éticas o de responsabilidad social en la reputación empresarial, apareciendo la diversidad como una de las dimensiones más relevantes a contemplar en este contexto (Berman et al, 1999; Kang et al, 2007; Tacheva y Huse, 2006), ya que suele obtener una elevada valoración por parte de analistas profesionales e inversores institucionales (Carter et al, 2003).

En conclusión, y recapitulando las consideraciones mencionadas, la revisión de aportaciones teóricas que se han efectuado desde el marco del gobierno corporativo permite suponer que el incremento de diversidad en el consejo de administración puede repercutir favorablemente en la rentabilidad de la compañía. Veamos seguidamente los resultados de los trabajos precedentes al respecto.

2.3 Revisión de la literatura.

En general, la mayoría de estudios empíricos difundidos que relacionan diversidad de género en los consejos y *performance* empresarial se centran en analizar la situación de un ámbito territorial concreto, predominando los trabajos referidos a grandes firmas estadounidenses, por la simple razón de la disponibilidad de datos existente (Fields y Keys, 2003). Las variables representativas de la diversidad se construyen a partir de la composición por sexos de los consejos de administración (Milliken y Martins, 1996), y en cuanto a indicadores de *performance* se utilizan tanto medidas de mercado (lógicamente, cuando se trata de compañías cotizadas) como medidas contables, o bien combinaciones de ambas. Cabe señalar que los diferentes estudios arrojan resultados opuestos, si bien algunos de ellos encuentran una evidencia consistente de correlaciones positivas entre unos consejos más heterogéneos y el rendimiento de las firmas (Carter et al, 2003; Erhardt et al, 2003).

Entre los trabajos que no obtienen evidencia se encuentra el de Kochan et al (2003), o el de Zahra y Stanton (1988), quienes consideran la diversidad en términos de minorías étnicas y de mujeres en el consejo, mientras que la *performance* se valora a través de diversas magnitudes contables. Por su parte, Farrell y Hersch (2004) estudian la reacción en términos de mercado a nombramientos femeninos para puestos en el consejo, sin obtener efectos significativos en los precios. Tampoco Wolfers (2006) encuentra diferencias en la relación entre valor de mercado y valor en libros en un conjunto de empresas cotizadas que se segmentan en dos grupos, atendiendo al hecho de que el consejero delegado sea un hombre o una mujer. Inclusive otros estudios encuentran una relación negativa entre diversidad y desempeño, como es el caso de Fryxell y Lerner (1989), respecto de algunas medidas de rentabilidad y de liquidez empresarial. En esta misma línea, Shrader et al (1997) indica una débil relación negativa entre la proporción de mujeres en el consejo y diversas medidas contables de compañías incluidas en el listado de "*Fortune 500*". Desde la perspectiva de las cotizaciones bursátiles, Lee y James (2003) observan reacciones negativas del mercado financiero a los nombramientos de mujeres como consejeras delegadas. También Ryan y Haslam (2005) encuentran cierta correlación negativa entre el porcentaje de mujeres en el consejo y los precios de mercado, diferenciando a las compañías en función de sus prácticas de diversidad y su estabilidad financiera.

En contraposición a los estudios citados, los resultados de otros trabajos permiten concluir que la diversidad de género ejerce efectos positivos sobre la valoración de la empresa o sobre sus indicadores de rendimiento. Así, Carleton et al (1998) observan que compañías marcadas por prácticas segregacionistas por parte de un inversor institucional relevante en Estados Unido como es TIAA-CREF (fondo de pensiones que establece criterios de selección de inversiones entre los cuales se incluyen directrices de diversidad) experimentaban retornos de mercado anormales en las fechas en que eran calificados por su falta de diversidad. Keys et al (2003)

concluyen que la promoción femenina a puestos de responsabilidad ejerce efectos significativos sobre la riqueza de los accionistas, mientras que en el estudio de Erhardt et al (2003) se obtiene evidencia del impacto positivo de la diversidad en medidas contables de rentabilidad. Ellis y Keys (2003) encuentran valoraciones superiores de mercado en firmas que son incluidas en índices institucionales de buenas prácticas por sus altos niveles de diversidad en la fuerza laboral. También sugieren efectos positivos Krishnan y Park (2005), en este caso respecto de la diversidad en cargos directivos, o Adams y Ferreira (2004), quienes estudian el impacto de la diversidad en el consejo, tanto en medidas contables como de mercado. Cabe citar asimismo el influyente trabajo de Carter et al (2003), el cual concluye que existen efectos significativos de una mayor diversidad sobre el valor de la firma. Asimismo, también Sudarsanam y Hiang (2007) observan una reacción favorable del mercado estadounidense en el período 1993-2004 cuando se produce una fusión o adquisición en la cual la empresa que lleva a efecto la operación tiene una mujer en el puesto de consejero delegado.

Si se efectúa un recorrido por estudios basados en datos que corresponden a otros ámbitos geográficos distintos a Estados Unidos, también los estudios existentes muestran resultados dispares. Así, Bonn et al (2004) comparan el impacto financiero para una muestra de empresas australianas y otra de japonesas, sin encontrar evidencia significativa para las compañías japonesas (que presentan unas tasas muy elevadas de desigualdad), pero obteniendo cierta correlación positiva de la diversidad con ratios contables y de mercado para el caso de Australia. Para Dinamarca, Smith et al (2006) encuentran un impacto positivo de la proporción de mujeres en los consejos sobre la rentabilidad para el período 1993-2001, al contrario que Rose (2006; 2007), quien obtiene resultados opuestos cuando considera a las empresas cotizadas en el mercado danés. Randøy (2006) analiza conjuntamente las 500 mayores compañías de Dinamarca, Suecia y Noruega y no concluye que la diversidad de género, edad o nacionalidad tenga efectos relevantes en los precios de mercado de las compañías. El trabajo de Mitchell (2004), para empresas sudafricanas cotizadas, estudia tanto la diversidad de género como la diversidad étnica, concluyendo que en este contexto geográfico existe una fuerte correlación entre mayor diversidad y mejor desempeño. El estudio de Martin et al (2007) se desarrolla utilizando una muestra de firmas del Reino Unido que contiene una amplia representación de medianas y pequeñas empresas, apreciándose para este tipo de compañías unos resultados que indican una relación no lineal entre una alta diversidad de género y ciertas variables representativas de la evolución del negocio, como es el incremento en la cifra de facturación.

3. DATOS Y METODOLOGÍA.

3.1 Datos.

La muestra utilizada en los análisis realizados consiste en una selección de compañías de diez países europeos,[1] obtenida de la base de datos *Amadeus*, correspondiente al cierre del ejercicio 2007. Como primer criterio para seleccionar las firmas que integrarían la muestra, y en función del objetivo de analizar el impacto de la diversidad de género existente en los órganos de gobierno societario, se ha optado por escoger aquellas empresas para las que estuviese disponible información relativa a la composición por sexos del consejo de administración para el período de referencia, así como dos indicadores contables calculados a partir de las cuentas anuales, la rentabilidad financiera y la rentabilidad económica[2]. Como en la fuente de datos original no se disponía de datos completos sobre la rentabilidad económica para Lituania y

[1] Los países considerados para seleccionar las compañías que integran la muestra han sido Austria, Alemania, Bélgica, España, Francia, Holanda, Lituania, Luxemburgo, Rusia y Suiza. Exceptuando a las empresas españolas, para las que se ha efectuado un recuento de los hombres y mujeres que ocupan puesto en el consejo de administración deduciendo la composición por sexos a partir de la consignación nominal que figuraba en la base, en el resto de casos la información ya se encontraba directamente disponible en la fuente de datos.

[2] La rentabilidad económica relaciona el resultado del ejercicio con el total de activos, mientras que la rentabilidad financiera relaciona el resultado con los fondos propios de la compañía. Las prácticas contables para cada país en el ejercicio económico de referencia (2007) se guían por criterios distintos, derivados de la diferente regulación nacional aplicable a la presentación de información financiera, aunque en ciertas empresas haya una convergencia en las normas contables con las que elaboran sus estados financieros, debido a su acceso a mercados de capitales.

Luxemburgo, las empresas correspondientes a estos dos países han tenido que excluirse de la muestra cuando dicho indicador intervenía en los análisis.

De la selección resultante de ese primer filtro, y tras efectuar algunos contrastes de coherencia destinados a eliminar observaciones que presentasen valores anormalmente extremos en las medidas contables, para cada país se escogieron las 300 mayores empresas de acuerdo a una ordenación basada en el importe de los ingresos de explotación durante el ejercicio. Por lo tanto, la muestra definitiva empleada se compone de 3.000 sociedades, para las cuales se consignan un total de 20.420 personas físicas identificadas nominalmente en su calidad de miembros de un consejo de administración, de las cuales 1.758 son mujeres, lo cual representa un 8,60% del total de cargos de consejero. Esa baja proporción femenina permite ya apreciar en términos globales el hecho de que en la muestra elegida se manifiesta una fuerte desigualdad de género, como cabía esperar a priori.

3.2 Medidas de diversidad.

En función de los objetivos del análisis a efectuar y del diseño que se plantea, se han empleado dos medidas de diversidad. En primer término un indicador inmediato y habitualmente utilizado en la literatura previa (Erhardt et al, 2003; Adams y Ferreira, 2004; entre otros), como es el porcentaje de mujeres en cada consejo, el cual representa realmente una medida de la desigualdad, más que de diversidad en sentido estricto. Por este motivo, y complementariamente, se ha calculado también como variable de diversidad una segunda medida, el Índice de Blau, que ha sido ampliamente utilizado para medir la heterogeneidad demográfica (Ruigrok et al, 2007), según la formulación que se expone en la ecuación [1], donde B es el índice de Blau y p es el porcentaje de miembros en cada i-ésima categoría de las k existentes (en este caso una para cada sexo, luego k=2). Cuanto más elevado el valor de B, mayor será el grado de diversidad, y dado que los valores varían entre 0 y $(k-1)/k$, la diversidad máxima será cuando B alcance el valor de 0,5.

$$B = \left[1 - \sum (p_i)^2 \right] \quad [1]$$

3.4 Metodología.

Tomando como base cada una de las dos medidas contables de *performance* de las compañías (rentabilidad económica y rentabilidad financiera), se ha contemplado de forma separada la realización de análisis estadísticos que comprueben si existe alguna influencia significativa de dos factores, uno de ellos representativo del país en el que radica la sede social y otro representativo de la diversidad de género en el consejo. De esta manera se establecen hipótesis tanto en el sentido de que los niveles de las medidas de diversidad como el sistema de gobierno corporativo, o incluso la interacción cruzada de ambas variables, conduzcan a diferencias significativas en los valores de las medidas financieras. Como la variable país representa un factor categórico, se ha procedido a la aplicación de técnicas de Análisis de Varianza para ver el efecto de ese factor sobre cada variable de rentabilidad, incluyendo en los modelos además otra covariable que sería la variable de diversidad. En los modelos se han contemplado las dos medidas de diversidad señaladas, por una parte la proporción de mujeres y por otro lado el índice de Blau. Lógicamente, por el evidente desequilibrio en la distribución de cargos de administración de las empresas de la muestra, ambas variables están muy relacionadas, ya que la generalizadamente escasa proporción de mujeres lleva a unos niveles bajos en el Índice de Blau. Como no es posible asumir la normalidad de las distribuciones en las rentabilidades, se ha optado por trabajar con sus funciones transformadas para combatir el principal problema de la asimetría de acuerdo al método de Box-Cox (García Ayuso, 1995), cuya formulación se expone en la ecuación [2], donde Y es la variable a transformar en cada caso (añadiendo el mínimo valor de la distribución cuando existan valores negativos) para cada firma $i = 1,\ldots,N$; y λ es el parámetro de transformación que optimiza la adecuación a la normalidad.

$$y_i^{(\lambda)} = \begin{cases} \dfrac{\left(y_i^{\lambda} - 1\right)}{\lambda} \text{ si } \lambda \neq 0 \\ \log\left(y_i\right) \text{si } \lambda = 0 \end{cases} \qquad [2]$$

Aunque no se obtendrá la normalidad en ninguna de las transformaciones (cuyos resultados, por razones de espacio, no se expondrán), tanto las pruebas estadísticas como su reflejo gráfico evidencian una mejora notable de la simetría en todos los casos, que sería realmente la característica que permite validar los resultados del análisis de la varianza. Además, se sortean los problemas de heteroscedasticidad de las funciones transformadas, como muestra la aplicación de la prueba de Levene.

4. RESULTADOS.

En la tabla 1 se presentan desagregados por países los estadísticos descriptivos para las medidas de diversidad y para las medidas contables de rentabilidad. Del análisis de los mismos podemos resaltar varios aspectos de interés. En primer lugar, respecto de la participación de mujeres en consejos de administración se observan diferencias en el promedio para cada país, donde la proporción más elevada corresponde a Lituania, con un 14,70%, confirmando así los resultados de trabajos como el de Terjesen y Singh (2008), que ya observa esta característica en economías post-comunistas, citando los casos en Europa de Eslovenia, de Lituania o de Rumanía. En el extremo opuesto, los países que presentan menor representación femenina serían los que corresponden a las empresas de la muestra que radican en Alemania (5,24%) y en Holanda (4,15%). El promedio para el total de la muestra difiere ligeramente de la proporción femenina calculada para el total de puestos de administrador, por el hecho de que cada empresa puede tener un distinto número de miembros del consejo, siendo por tanto distinta su ponderación.

Si se observa el índice de diversidad de Blau, llama la atención que la elevada participación femenina de las empresas de Lituania no se corresponde con unos niveles destacados en esta otra medida. Ello obedece a que las compañías de dicho país integradas en la muestra suelen tener un consejo con muy pocos miembros (derivado también del hecho de que corresponden a firmas de relativamente menor tamaño que las de otros países), y adicionalmente con una composición escasamente diversa; hasta 23 casos de firmas lituanas que tienen una representación del 100% de mujeres figuran como entidades de administración unipersonal. En todo caso, la divergencia de resultados que se observa al comparar las dos medidas de diversidad justifica la aplicación en los contrastes estadísticos de ambos indicadores, uno orientado específicamente hacia la desigualdad que experimentan las mujeres y otro que realmente se dirija a estimar el grado de equilibrio en la composición del consejo.

Tabla 1. Estadísticos descriptivos

	Media	Desviación típica	Asimetría tipificada	Curtosis tipificada
Proporción de mujeres (%)				
Austria	7,20	17,52	2.119,94	3.585,75
Bélgica	7,83	15,19	1.561,62	1.605,68
Francia	8,75	15,42	2.316,46	4.974,24
Alemania	5,24	12,97	2.876,85	7.541,02
Lituania	14,70	30,31	1.386,79	888,79
Luxemburgo	10,55	17,53	1.157,92	653,92
Holanda	4,15	8,49	1.673,15	2.268,49
Rusia	7,67	22,21	2.368,54	3.772,50
España	8,95	13,30	1.190,75	902,74
Suiza	8,30	9,25	959,70	770,07
Total	8,33	17,50	6.900,34	12.375,70

Índice de Blau				
Austria	0,0724	0,1587	13,5496	7,0306
Bélgica	0,0983	0,1657	9,5608	0,7437
Francia	0,1123	0,1470	7,1217	-0,9034
Alemania	0,0657	0,1291	13,9320	10,1881
Lituania	0,0675	0,1630	14,5486	8,3558
Luxemburgo	0,1275	0,1884	6,7050	-3,1594
Holanda	0,0653	0,1242	11,8365	5,4031
Rusia	0,0433	0,1246	19,5224	21,5849
España	0,1277	0,1656	6,2446	-2,3304
Suiza	0,1351	0,1338	4,5347	-1,8575
Total	0,0915	0,1543	32,3047	7,2695
Rentabilidad económica (%)				
Austria	11,42	30,90	6,06	9,01
Bélgica	19,96	23,27	15,94	26,98
Francia	20,02	19,47	18,14	42,03
Alemania	17,69	17,33	19,84	37,03
Holanda	25,32	26,82	15,25	22,56
Rusia	32,77	33,96	11,74	10,20
España	27,77	26,66	9,02	11,87
Suiza	12,20	11,64	5,65	12,47
Total	20,89	25,66	36,59	62,59
Rentabilidad financiera (%)				
Austria	4,28	49,53	-0,32	7,05
Bélgica	21,07	33,84	5,00	46,89
Francia	26,80	31,59	15,83	30,79
Alemania	34,86	39,62	20,50	43,19
Lituania	29,28	36,22	8,49	34,12
Luxemburgo	37,46	43,60	9,83	22,16
Holanda	30,34	38,34	-0,63	49,05
Rusia	36,98	48,97	7,70	23,44
España	28,18	23,63	11,04	13,50
Suiza	15,74	18,59	25,98	132,27
Total	26,50	38,82	23,42	103,15

Las diferencias entre países que se manifiestan en las variables de rentabilidad indican también la importancia de considerar la incidencia del factor nacional. Atendiendo a los promedios, destacan las elevadas rentabilidades que obtienen las empresas rusas, mientras que para el período considerado hay otros contextos geográficos cuyas firmas ofrecen unos niveles muy inferiores, como es el caso de Austria (sobre todo) y Suiza, tanto en los valores promedio como para muchas de las observaciones individuales de la submuestra nacional, como se deduce de la observación de la distribución.

En la tabla 2 se muestran los p-valores obtenidos para los análisis que se han realizado para contrastar los efectos de la medida de diversidad y del país de pertenencia de la empresa sobre las transformaciones Box-Cox de las variables de rentabilidad. El modelo 1 contempla como medida de diversidad la proporción de mujeres, mientras que el modelo 2 considera como medida de la diversidad al Índice de Blau. Como puede observarse de los resultados, en el primer modelo el factor país resulta significativo para explicar la rentabilidad en todos los casos. Ello resulta lógico no solamente por las distintas tradiciones de gobierno corporativo, ya que la variable diversidad se plantea como indicador de un mecanismo de gobierno corporativo que se presupone puede incidir sobre la *performance*, sino también porque las condiciones del entorno nacional recogerán el impacto de otras muchas variables que obviamente influyen en la rentabilidad empresarial y que, sin embargo, no se han contemplado en los modelos planteados.

Tabla 2: p-valores obtenidos para los contrastes de los modelos

	Transformada Rentabilidad económica	Transformada Rentabilidad financiera
Modelo 1	N=2.400	N=3.000
País	0,0000*	0,0000*
Proporción_ mujeres	0,1916	0,2956
País*Proporción_ mujeres	0,7118	0,9333
Modelo 2	N=2ÿÿ00	N=3ÿÿÿÿ
País	0,0000*	0,0000*
Índice_Blau	0,0017*	0,0451**
País*Índice_Blau	0,3282	0,3850

* denota significación al 1%, ** denota significación al 5%

Otro de los aspectos que merecen destacarse es que la proporción de mujeres no resulta significativa en ninguna de las dos rentabilidades, ni tampoco la interacción de primer orden de la proporción de mujeres con la variable representativa del país. Ello puede ser interpretado de la siguiente manera: aunque la razón principal de la desigualdad de género en los consejos sea la escasa representación femenina, no puede suponerse que la presencia de mujeres por si sola tenga efectos económicos directos sobre la actuación de la empresa. En cambio, cuando en el modelo se considera como variable de diversidad el Índice de Blau, se obtiene evidencia de un impacto significativo sobre las dos rentabilidades, más claramente en el caso de la rentabilidad económica (aunque la interacción con el país no se revela significativa). De acuerdo a estos resultados, podría afirmarse que las hipótesis derivadas de las teorías del gobierno corporativo que aluden a la importancia de la diversidad que puede conducir a un mejor desempeño de la firma hallan una confirmación empírica.

5. CONCLUSIONES

La diversidad de género en los consejos de administración de las sociedades mercantiles ha sido un tema que ha suscitado un creciente interés en la literatura empresarial de los últimos años. Particularmente, una línea de análisis destacable se enfoca hacia los potenciales efectos que puede ejercer una mayor diversidad de género sobre la *performance* de las organizaciones, puesto que, aunque los desarrollos teóricos relacionados con el buen gobierno corporativo avalan esa relación, las investigaciones empíricas realizadas ofrecen resultados no concluyentes. En el presente trabajo se analiza si la composición heterogénea en términos de género de los consejos de administración de compañías europeas tiene efectos significativos sobre distintos indicadores de rentabilidad. Para ello se ha empleado una muestra que contiene 3.000 empresas pertenecientes a diez países (300 empresas por país) con más de 20.000 cargos de administración en los consejos identificados para el período correspondiente al cierre del ejercicio 2007.

Los resultados indican que la participación femenina en los consejos es escasa (un 8,60% para el conjunto de cargos de administración de la muestra), pero la proporción es desigual para distintos países, lo cual, obviamente se transmite a otros indicadores de diversidad, que muestran también una tasa muy baja. Respecto al impacto sobre la rentabilidad empresarial medida a partir de datos contables, los indicadores de diversidad parecen tener algún efecto significativo, lo cual resulta explicable desde los desarrollos de la teoría de la agencia, de la dependencia de recursos o derivado de una orientación *stakeholder*. También hay una evidencia consistente acerca de que el factor país explica las diferencias de rentabilidad entre empresas. Sin embargo, la participación de mujeres considerada como variable de incidencia no se muestra significativa a la hora de determinar diferentes niveles en los indicadores de rentabilidad de las firmas.

REFERENCIAS BIBLIOGRÁFICAS:

ADAMS, R.B. Y FERREIRA, D. (2004), "Gender diversity in the boardroom", ECGI Working Paper Series in Finance 58.

BERMAN, S.L.; WICKS, A.C.; KOTHA, S. Y JONES, T.M. (1999), "Does stakeholder orientation matter? The relationship between stakeholder management models and firm financial performance", Academy of Management Journal, Volumen 42, nº 5, pp. 488-506.

BILIMORIA, D. Y PIDERIT, S. (1994), "Board committee membership: effects of sex-based bias", Academy of Management Journal, Volumen 37, nº 6, pp. 1.453-1.477.

BONN, I.; YOSHIKAWA, T. Y PHAN, P.H. (2004), "Effects of board structure on firm performance: A comparison between Japan and Australia", Asian Business & Management, Volumen 3, pp. 105–125.

CARLETON, W.T.; NELSON, J.M. Y WEISBACH, M.S. (1998), "The influence of institutions on corporate governance through private negotiations: evidence from TIAA-CREF", Journal of Finance, Volumen 53, nº 4, pp. 1335-1362.

CARTER, D.A.; SIMKINS, B.J. Y SIMPSON, W.G. (2003), "Corporate governance, board diversity, and firm value", The Financial Review, Volumen 38, nº 1, pp. 33-53.

COMISIÓN NACIONAL DEL MERCADO DE VALORES (CNMV) (2006), "Código Unificado de Buen Gobierno. Informe del grupo especial de trabajo sobre Buen Gobierno de Sociedades Cotizadas", Comisión Nacional del Mercado de Valores, Madrid.

CORE, J.E.; GUAY, W.R. Y RUSTICUS, T.O. (2006), "Does weak governance cause weak stock returns? An examination of firm operating performance and investors' expectations", The Journal of Finance, Volumen 61, nº 2, pp. 655-687.

ELLIS, K.M. Y KEYS, P.Y. (2003), "Stock returns and the promotion of workforce Diversity", Working paper, University of Delaware.

ERHARDT, N.L.; WERBEL, J.D. Y SHRADER, C.B. (2003), "Board of director diversity and firm financial performance", Corporate Governance: An International Review, Volumen 11, nº 2, pp. 102-111.

FARRELL, K.A. Y HERSCH, P.L. (2005), "Additions to corporate boards: the effect of gender", Journal of Corporate Finance, Volumen 11, nº 1-2, pp. 85-106.

FIELDS, M.A. Y KEYS, P.Y. (2003), "The emergence of corporate governance from Wall St. to Main St.: Outside directors, board diversity, earnings management, and managerial incentives to bear risk", The Financial Review, Volumen 38, nº 1, pp. 1-24.

FORBES, D.P. Y MILLIKEN, F.J. (1999), "Cognition and corporate governance: Understanding boards of directors as strategic decision-making groups", Academy of Management Review, Volumen 24, nº 3, pp. 489-505.

FRYXELL, G.E. Y LERNER, L.D. (1989), "Contrasting corporate profiles: women and minority representation in top management positions", Journal of Business Ethics, Volumen 8, nº 5, pp. 341-352.

GABRIELSSON, J. Y HUSE, M. (2004), "Context, behavior and evolution: Challenges in research on boards and governance", International Studies in Management and Organization, Volumen 34, pp. 11-36.

GARCÍA-AYUSO, M. (1995), "La necesidad de llevar a cabo un replanteamiento de la investigación en materia de análisis de la información financiera", Análisis Financiero, nº 66, pp. 36-61.

HIGGS, D. (2003), "The Higss Report: Review of role and effectiveness of non-executive directors", The Department of Trade and Industry.

HILLMAN, A.J.; CANNELLA, A.A. Y HARRIS, I.C. (2002),"Women and racial minorities in the boardroom: how do directors differ?", Journal of Management, Volumen 28, pp. 747-763.

IBRAHIM, N.A. Y ANGELIDIS, J.P. (1994), "Effect of board members' gender on corporate social responsiveness orientation", Journal of Applied Business Research, Volumen 10, nº 1, pp. 35-40.

KANG, H.; CHENG, M. Y GRAY, S.J. (2007), "Corporate governance and board composition: diversity and independence of Australian boards", Corporate Governance: An International Review, Volumen 15, nº 2, pp. 194-207.

KANTER, R.M. (1977), "Men and women of the corporation", Basic Books, New York.

KESNER, I.F. (1988), "Director's characteristics and committee membership: an investigation of type, occupation, tenure, and gender", Academy of Management Journal, Volumen 31, nº 1, pp. 66-84.

KEYS, P.Y.; ELLIS, K.M.; NEWSOME, P.T. Y FRIDAY, S.S. (2003), "Shareholders benefits of diversity", Working Paper. http://fisher.osu.edu/fin/dice/seminars

KIEL, G.C. Y NICHOLSON, G.J. (2003), "Board composition and corporate performance: how the Australian experience informs contrasting theories of corporate governance", Corporate Governance: An International Review, Volumen, 11, nº 3, pp. 189-205.

KOCHAN, T.; BEZRUKOVA, K.; ELY, R.; JACKSON, S.; JOSHI, A.; JEHN, K.; LEONARD, J.; LEVINE, D.; THOMAS, D. (2003), "The effects of diversity on business performance: report of the diversity network", Human Resource Management, Volumen 42, nº 1, pp. 3-21.

KRISHNAN, G.V. Y PARK, D. (2005), "A few good women - on top management teams", Journal of Business Research, Volumen 58, nº 12, pp. 1712-1720.

LAPORTA, R.; LÓPEZ DE SILANES, F.; SHLEILER, A. Y VISHNY, R. (2002), "Investor protection and corporate valuation", Journal of Finance, Volumen 57, pp. 1147-1170.

LEE, P.M. Y JAMES, E.H. (2003), "She'-e-os: Gender effects and stock price reactions to the announcements of top executive appointments", Working Paper, University of Virginia.

MCDOUGAL, M. (1996), "Equal opportunities versus managing diversity. Another challenge for public sector management?", The International Journal of Public Sector Management, Volumen 9, nº 5, pp. 62.

MARTIN, L.; WARREN-SMITH, I.; ROPER, S. Y SCOTT, J. (2007), "Board gender diversity in Uk firms-recent trends and performance effects", Aston Business School Research Papers, Aston University.

MILLIKEN, F.J. Y MARTINS I.L. (1996), "Searching for common treads: Understanding the multiple effects of diversity in organizational groups", Academy of Management Journal, Volumen 21, nº 2, pp. 402-433.

MITCHELL VAN DER ZAHN, J-L.W. (2004), "Association between gender and ethnic diversity on the boards of directors of publicly listed companies in South Africa and intellectual capital performance", Financial Reporting, Regulation & Governance, Volumen 3, nº 1, pp. 1-35.

OAKLEY, J.G. (2000), "Gender-based barriers to senior management positions: Understanding the scarcity of female CEOs", Journal of Business Ethics, Volumen 27, nº 4, pp. 321-334.

PELLED, L.H. (1996), "Demographic diversity, conflict, and work group outcomes: An intervening process theory", Organization Science, Volumen 7, nº 6, pp. 615–631.

PELLED, L.H.; EISENHARDT, K.M. Y XIN, K.R. (1999), "Exploring the black box: an analysis of work group diversity, conflict, and performance", Administrative Science Quarterly, Volumen 44, nº 1, pp. 1-28.

PETERSON, C.A. Y PHILPOT, J. (2007), "Women's roles on U.S. Fortune 500 boards: Director expertise and committee memberships", Journal of Business Ethics, Volumen 72, pp. 177-196.

RANDØY, T.; THOMSEN, S. Y OXELHEIM, L. (2006), "A Nordic perspective or corporate board diversity", Nordic Innovation Centre.

ROBINSON, G. Y DECHANT, K. (1997), "Building a business case for diversity", The Academy of Management Executive, Volumen 11, nº 3, pp. 21-31.

ROSE, C. (2006), "Board composition and corporate governance-A multivariate analysis of listed Danish firms", European Journal of Law and Economics, Volumen 21, pp. 113-127.

--- (2007), "Does female board representation influence firm performance? The Danish evidence", Corporate Governance: An International Review, Volumen 15, nº 2, pp. 404-413.

ROSENZWEIG, P. (1998), "Managing the new global workforce: Fostering diversity, forging consistency", European Management Journal, Volumen 16, nº 6, pp. 644-652.

RUIGROK, W.; PECK, S. Y TACHEVA, S. (2007), "Nationality and gender diversity on Swiss corporate boards", Corporate Governance: An International Review, Volumen 15, nº 4, pp. 546-557.

RYAN, M.; HASLAM, S.A. (2005), "The glass cliff: evidence that women are over represented in precarious leadership positions", British Journal of Management, Volumen 16, nº 1, pp. 81-90.

SHRADER, C.B.; BLACKBURN, V.B. Y ILES, P. (1997), "Women in management and firm financial performance: An explorative study", Journal of Managerial Issues, Volumen 9, nº 3, pp. 355-372.

SICILIANO, J.I. (1996), "The relationship of board member diversity to organizational performance", Journal of Business Ethics, Volumen 15, nº 12, pp. 1.313-1.320.

SINGH, V.; KUMRA, S. Y VINNICOMBE, S. (2002), "Gender and impression management: playing the promotion game", Journal of Business Ethics, Volumen 37, nº 1, pp. 77-89.

SMITH, E. (2007), "Gender influence on firm-level entrepreneurship through the power structure of boards", Women in Management Review, Volumen 22, nº 3, pp. 168-186.

SMITH, N.; SMITH, V. Y VERNER, M. (2006), "Do women in top management affect firm performance? A panel study of 2,500 Danish firms", International Journal of Productivity and Performance Management, Volumen 55, nº 7, pp. 569-593.

SUDARSANAM, S. Y HIANG, J. (2007), "Gender Diversity in US Top Management: Impact on Risk-taking and Acquirer Performance", Working Paper, Cranfield University.

TACHEVA, S. Y HUSE, M. (2006), "Women directors and board task performance: Mediating and moderating effects of board working style", European Academy of Management Conference, February, Oslo.

TERJESEN, S. Y SINGH, V. (2008), "Female presence on corporate board: A multi-country study of environmental context", Journal of Business Ethics, Volumen 83, pp. 55-63.

TYSON, L. (2003), "The Tyson Report on the recruitment and development of non-executive directors", Department of Trade & Industry.

VAN DER WALT, N. Y INGLEY, C. (2003), "Board dynamics and the influence of professional background, gender and ethnic diversity of directors", Corporate Governance: An International Review, Volumen 11, nº 3, pp. 218-234.

WAGNER, W.G.; PFEFFER, J.; O'REILLY III, C.A. (1984), "Organizational demography and turnover in top-management groups", Administrative Science Quarterly, Volumen 29, nº 1, pp. 74-92.

WARTICK, S. L. (1992), "The relationship between intense media exposure and change in corporate reputation", Business and Society, Volumen 31, pp. 33-42

WEBB, E. (2004), "An examination of socially responsible firms' board structure", Journal of Management and Governance, Volumen 8, pp. 255-277.

WOLFERS, J. (2006), "Diagnosing discrimination: Stock returns and CEO gender", Journal of the European Economic Association, Volumen 4, nº 2-3, pp. 531-541.

ZAHRA, S. Y STANTON, W. (1988), "The implications of board of directors' composition for corporate strategy and performance", International Journal of Management, Volumen 5, nº 2, pp. 229-236.

ENTREPRENEURSHIP: CONCEPT AND MEASUREMENT[1]

RÚBEN TIAGO BRANCO
E-mail: rubentmb@gmail.com
FERNANDA LLUSSÁ
E-mail: fajl@fct.unl.pt
JOSÉ ALBUQUERQUE TAVARES
E-mail: jtavares@fe.unl.pt
Faculdade de Economia
Universidade Nova de Lisboa

Resumo

É feita uma pesquisa acerca do conceito de empreendedorismo e das diferentes formas através das quais ele poderá ser medido e abordado. São apresentadas visões bastante diversas sobre quem é o empreendedor, é discutida a negligência com que a ciência económica tem tratado este tema e são exploradas diferentes abordagens sobre como medir a actividade empreendedora.

Palavras-chave: empreendedorismo, inovação, medição.
Área Temática: Economia Empresarial.

Abstract

We survey the concept of entrepreneurship and the different ways in which entrepreneurial activity might be measured and assessed. In the process we present the very diverse views of who the entrepreneur is, discuss how economic science has neglected the issue and explore different approaches to the measurement of entrepreneurial activity.

Key words: entrepreneurship, innovation, measurement.
Thematic Area: Entrepreneurial Economics.

1. INTRODUCTION

Entrepreneurship is at the heart of output and productivity growth, exports and job creation. At a personal level, it is associated with personal autonomy and success. However, to date, economic theory has been at best mildly interested in the concept of entrepreneurship.[2] In the traditional neo-classical model, individuals have perfect information, economic objectives are clear and can be stated in rational terms, markets exist and are assumed to clear in equilibrium at one set of prices.[3] A person with special ability to deal with change and uncertainty, as detailed below, is not a necessary character; an able manager that deals with quantities and prices is all that is needed.[4] The development of economic theory in the last few decades, with the emphasis

[1] Fernanda Llussá is a member of DCSA at FCT-UNL, and is thankful to the Fundação para a Ciência e Tecnologia for financial support. José Tavares thanks both Nova Forum and Fundação para a Ciência e Tecnologia for financial support.
[2] Economic History and the History of Economic Thought are, to an extent, exceptions. This may result from the salient role of personalities in the history of development and of economic theory. Entrepreneurs, including scientists, innovators and business owners and managers have an eminent role and deserved attention in Economic History; major economic thinkers are, to a considerable extent, successful intellectual entrepreneurs that easily become aware ÿÿ the role of individuals.
[3] Aÿÿexception to the absence of the entrepreneur in economics is the work of Schumpeter who, though preoccupied with the aggregate economy, gives emphasis to the role of entrepreneurs and their personal traits in economic development at large. See, for example, Schumpeter (1934). The Schumpeterian approach however, is distant from mainstream economic theory.
[4] Baumol (1993b) states that "[i]t seems to be taken for granted in the literature that, even if entrepreneurs are not in complete control of our economic destiny, they influence its direction as few, if any others, are able to do. But having acknowledged this,

on imperfect markets, information theory, and game theory, has still not suggested the importance of the entrepreneur.[5] Baumol (1968) states that this model "is essentially an instrument of optimality analysis of well defined problems which need no entrepreneur for their solution". Real world problems are not, however, always well defined, especially if they are the important problems of starting a firm *ab initio*. According to Wennekers and Thurik (1999), as neo-classical economics became more formalized, references to the entrepreneur disappeared from the textbooks. The sad result is that almost all social sciences have a theory of the entrepreneur, with the possible lone and notorious exception of economics.[6]

Economists are increasingly uncomfortable with the glaring absence of the entrepreneur from economic theory. The entrepreneur is more difficult to ignore as small and new firms survive and compete proficiently with the large corporations and organizations that, at one point, were predicted to ultimately gain control of the whole economy.[7] Since the 1980s it became clear that firms of different size would continue to coexist and the flexibility of small firms was identified as a competitive advantage, some arguing that large mass-producers, working according to "Fordist" principles actually lost importance as the share of small firms expanded in several sectors ion many industrial countries .[8] Recognizing this, large firms have started to create room for entrepreneurial activity within its ranks.

In this paper we discuss the concepts of entrepreneurship and entrepreneur as they have been used in economics. We then propose a typology of entrepreneurial activity using as criteria the main motivation of entrepreneurs and suggest different indicators that can be understood as measures of entrepreneurial activity. Our contribution is thus to bring into light the core characteristics of entrepreneurship and their manifestation in economic activity.

2. WHO IS THE ENTREPRENEUR?

The figure of the entrepreneur apparently made its first appearance in the texts of Richard Cantillon.[9] This eighteenth century thinker suggested the division of economic agents into three categories: landowners, entrepreneurs and employees. Cantillon´s division - updated by substituting property-owners or capitalists by landowners - is magnificently prescient as it recognizes the three main economic activities in relation to the characteristics and attitude of those exercising them. Cantillon's entrepreneur is that who exercises business engagements and combines production factors in the face of uncertainty. Without uncertainty, the role the entrepreneur would be redundant.[10] According to Wennekers and Thurik (1999) and Hébert and Link (1989), also Von Thünen and Menger distinguish clearly between the entrepreneur and the supplier of financial capital, similar to Cantillon's "landowner". Marshall, on the other hand,

implicitly or explicitly, normally no more is done to incorporate the entrepreneur's role into the mainstream models of value theory or the theory of the firm."

[5] Two theoretical developments brought entrepreneurship some attention: Leibenstein´s X-efficiency theory and institutional economics stemming from the work of Coase. X-efficiency refers to the degree of inefficiency within the firm. According to Leibenstein firms don´t necessarily realize their "theoretical " productive potential, there is a conflict between employer and employee over how hard the latter works and effort and alertness are required to change old routines and production techniques. Leibenstein proposes entrepreneurship as the creative response to X-efficiency, as in Leibenstein (1968; 1979). Coase (1937), on the other hand, who first suggested that firms exist precisely within the confines and the extent of "contracts" where the market does not work well, saw the entrepreneur as the coordinator of production within the firm, needed because the price mechanism is for the most part absent from the allocation of resources within the firm itself. Williamson (1975) furthered this idea suggesting that the firm´s task is economizing on transaction costs. An important and related question, according to Wennekers and Thurik (1999), is where entrepreneurs better appropriate the fruits of their important abilities: within the firm or or by starting a new firm?

[6] See Casson (1982).

[7] See Galbraith (1967) and Schumpeter (1942).

[8] See Piore and Sabel (1984).

[9] See Cantillon (1931). A completely different perspective is that of Stauss (1944), who considers that *the firm is the entrepreneur*, existing "apart from the individuals who compose its decision-making organization, but it does not function apart from them."

[10] The work of Frank Knight has proposed a distinction between risk and uncertainty that has been the norm in economic theory, where insurance is unique and uninsurable. One can view the entrepreneur as that who deals with uncertainty – both in the market or within firms -, while controlling risk – where insurance and other standard techniques are available – is a managerial task.

contrasts the entrepreneur, the "pioneer of new paths", and the "superintendent", who merely organizes the production.[11] See Marshall (1961). Baumol (1993b) point to two different meanings for entrepreneur: the organizer who culminates the process of organizing resources into a new productive venture, the firm; and the innovator who transforms ideas and inventions into "economically viable entities, whether or not these take the form of a business firm". Schumpeter (1934), on the other hand, focused his attention on the entrepreneur as innovator, though the degree of innovation varied, from business creation and expansion to the move to new markets and products. These entrepreneurial qualities would be present in only a small fraction of the population.

According to Hébert and Link (1989) and Wennekers, Uhlaner and Thurik (2000), three main traditions have tried to grapple with the meaning of entrepreneurship, all ultimately traced back to Cantillon (1931). In the German or Schumpeterian tradition, as in Schumpeter (1934) the entrepreneur is a creator of instability that changes the "rules of competition" for the industry. The Chicago tradition of Frank Knight, as in Knight (1921), where entrepreneurs lead markets to equilibrium. The role of entrepreneurs as agents of change or individuals with special ability to respond to change is not emphasized. The Austrian tradition of Ludwig von Mises and Israel Kirzner, as in Kirzner (1979,1997), emphasizes the role of the entrepreneur in identifying profit opportunities in the wake of aggregate changes, satisfying new needs and overcoming market inefficiencies. Wennekers and Thurik (1999) define entrepreneurship as "the perception and creation of new economic opportunities" combined with "decision-making on the (…) use of resources". The latter seems mostly related to the neoclassic perspective and the firm owner-manager, while the relation of the entrepreneurs to new opportunities connects back with the Schumpeterian and Austrian traditions.

Wennekers and Thurik (1999) enumerate several roles that have been associated with the entrepreneurs in the literature, which can be summarized as: person who starts a new business; person who takes the risk, innovator; supplier of financial capital; owner of the enterprise; decision-maker; industry leader; manager – or organizer, coordinator, contractor, allocator of economic resources among alternative uses; employer of factors of production; and, finally, even arbitrageur. Some of these definitions refer to tasks that the entrepreneur may undertake but that do not characterize her or him –the case of manager, owner, supplier of financial capital -, others point to characteristics that are possible among entrepreneurs but do not sum up his role – the case of innovator -, and still others are vague and may or may not be appropriate – industry leader and arbitrageur. We consider the entrepreneur a self-motivated high-achiever, a risk-taker and a non-specialist that intermediates between different functions to innovate within an organization or by creating a new organization.[12] The entrepreneur is an innovator that assumes the risk of new endeavours, often by starting a new firm or organization, and then making it grow and change to pursue objectives such as profit and, yes, but also personal or general human happiness, by combining resources such as ideas, techniques and inventions, workers and managers, and capital.[13] Entrepreneurial activity most often involves the creation of new organizations, like business firms, though this is not necessarily the case as there is entrepreneurial activity within firms.[14]

[11] Entrepreneurs themselves, according to Tuttle (1927), "have long distinguished, though in a loose and general manner, between employer and workman, between employer and capitalist, and between employer and landowner, on the one hand, and on the other, between profit and wages, between profit and interest on a money loan and between profit and the rent of instruments."

[12] McClelland (1971) identified the need for achievement as the most distinguishable quality of entrepreneurs. not motivated by money per se; they are motivated by the desire to succeed; profitability in money terms is useful only insofar as it provide concrete knowledge of how well one is doing one's job. There is here a clear connection with the protestant notion of "predestination" as highlighted in the classic "The Protestant Ethic and the Spirit of Capitalism", by Max Weber. See Weber (…).

[13] Hébert and Link (1989, p. 47) proposed a definition of entrepreneur as : "someone who specializes in taking responsibility for and making judgmental decisions that affect the location, form, and the use of goods, resources, or institutions".

[14] Contrast Gartner and Carter (2004), who argue that "entrepreneurial behaviour involves the activities of individuals who are associated with creating new organizations rather than the activities of individuals who are involved with maintaining or changing

3. THE MOTIVATIONS OF THE ENTRPRENEUR

In this section we propose a taxonomy of the entrepreneur´s motivations, as made explicit in Figure 1 below, and then motivate it with recourse to the literature. Individuals become entrepreneurs for different reasons: because it corresponds to innate preference, and they yearn for the personal autonomy, risk taking, innovativeness and social prestige associated with creating new ventures; because economic need or other oblige it to start anew and take the risk of attempting a new activity; or because opportunity arises, either in the form of economic rents that are available to be taken by those who make the effort, or because an important innovation, technical or otherwise, to which the individual has easier access is present.[15] Some authors differentiate between "push" and "pull" factors as determinants of entrepreneurship, where pull factors are associated with the prospect (in expectation) of being better off as an entrepreneur (in material or other terms), and "push" factors are associated with some dissatisfaction with the current job or position. In our typology need would be the case of a "push" factor.

Rent-seeking cannot be neglected as a source of motivation for entrepreneurial activity. It can be associated to many activities, such as corrupting public officials to build in otherwise environmentally-protected areas to suppliers of high-demand illegal drugs where market power is enforced by continuous – but ultimately fallible – police persecution.[16] Naturally, rent-seeking can be associated with completely legal activities such as creating mechanisms to differentiate products and create market power privately or through the enactment of government laws that biases the productive process in favour of some, or inducing government intervention that favours own private interests at the expense of the interests of others. However, in general, rent-seeking activities are not productive for society as a whole and correspond to second-best situations, where a more efficient allocation would yield a larger amount of total resources.[17]

Different societies encourage different types of initiative. Baumol (1990a) suggests that weak and unstable formal institutions encourage unproductive entrepreneurship or destructive entrepreneurship, rendering unproductive activities more profitable than productive innovation-driven activities. Baumol (1990) goes as far as suggesting that while the total "supply of entrepreneurs" does vary among societies, more variation is associated with the mix of productive - such as innovation - and unproductive activities - such as rent seeking. Substitutability among types of entrepreneurship may thus more easily harm total output than the mere absence of entrepreneurial activity.[18]

the operations of ongoing established organizations." The frequent neglect of entrepreneurial activity that occurs within business firms is largely the consequence of the difficulty of measuring entrepreneurial-like activities that take place within existing organizations.

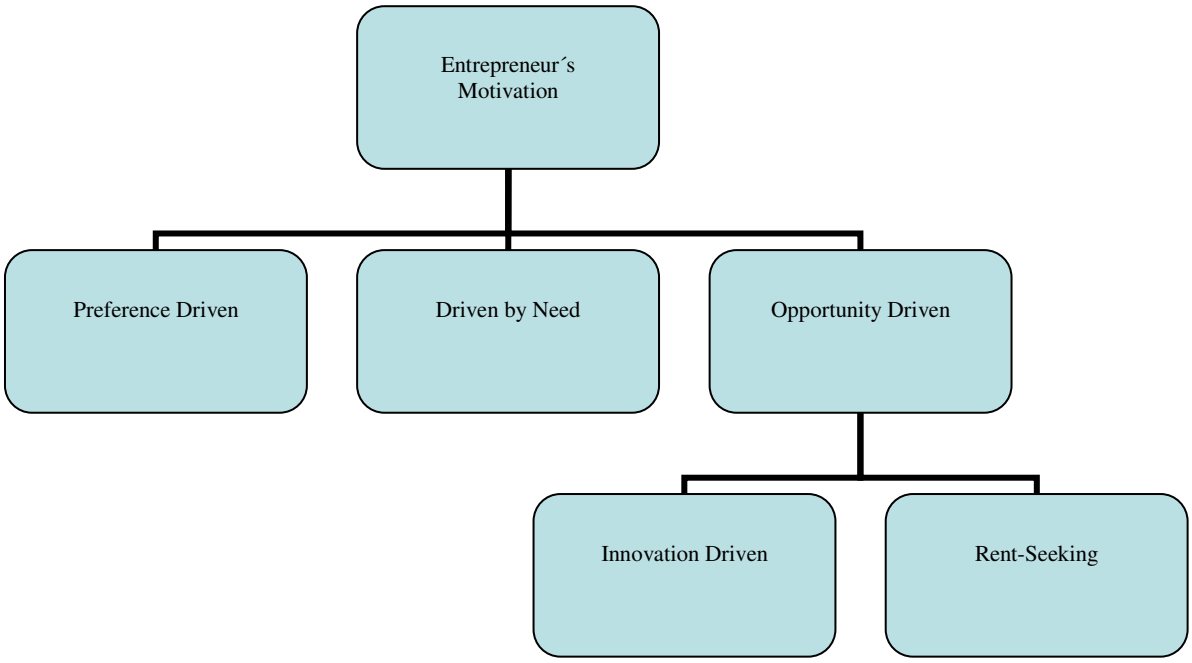
[15] An economic rent is an amount of resources that may be accessed in an environment with less than perfect competition, such as a competitor to a product theretofore supplied only by a monopolist. In a specific sense, privileged access to an innovation is a special case of economic rent.

[16] Other possible forms are litigation, lobbying, takeovers, tax evasion and tax avoidance, illegal and shadow activities, as well as corruption.

[17] In contrast, what Baumol (1993a) calls "productive entrepreneurship" is any activity contributing directly or indirectly to an increase in the net output of the economy."

[18] In some cases the same firms engage in both in legal and illegal rent-seeking, as well as more generally productive activities, as has been shown by Smallbone and Welter (2001) for new and small firms in the context of transition from a socialist to a market economy. When the rules are unclear, incomplete and changing very fast, this "portfolio" of Entrepreneurial activity may be key to ensure firm survival and growth, as suggested in Smallbone and Welter (2006).

Figure 1. The Motivations of Entrepreneurs



4. MEASURING ENTREPRENEURIAL ACTIVITY

Individuals who we labelled above as entrepreneurs can innovate and organize in different ways: by creating a new firm or organization,[19] a new product, perfecting or creating a new production or sales method, explore a new market, a new source of inputs or devise a new form of organization for a whole industry. Some of these manifestations of entrepreneurial activity are visible, explicit, and facilitate measurement, such as the creation of business firms. Others take place within existing organizations or have a qualitative element that makes them hard to measure. Here we will try to be inclusive, thus highlighting key dimensions of entrepreneurship that are tend to be neglected in empirical and theoretical studies.

In order to guide the description of the dimensions of entrepreneurial activity, we will build on Grilo and Thurik (2007) and their framework describing the typical life cycle of an enterprise, composed by the following stages:

(1) *outsider* - agent with preference for self-employment but "who never thought about it".
(2)*conception* – when an idea has appeared or the opportunity identified and the agent starts "thinking about it".
(3) *gestation* - the agent starts "taking steps for starting up"
(4) *infancy* - a young firm, product, method, market, source or industry organization is born.
(5) *adolescence* - the enterprise grows at a fast pace
6) *maturity* - the growth rate of the enterprise slows down and its size stabilizes
(7) *decline* - negative growth
(8) *exit* - the enterprise exits the market
(9) *step out* - the agent gives up on entrepreneurial activity[20]

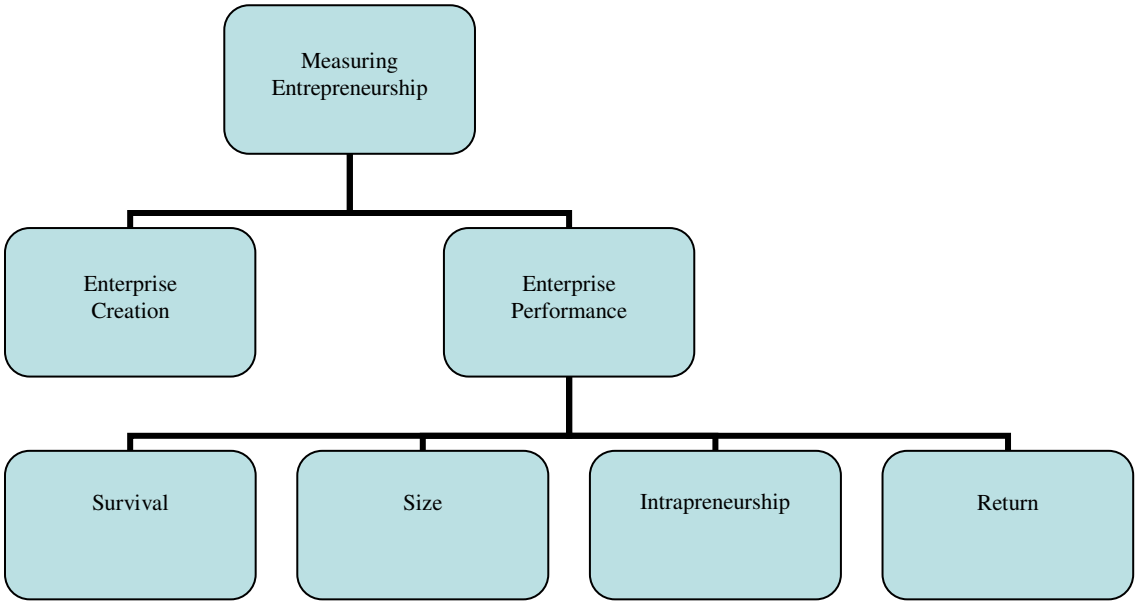
As in Thurik and Grilo (2005), we consider stages (1) to (3) as *Latent entrepreneurship*, and stages (4) to (7) as *Actual entrepreneurship.*[21] Moreover, stages (2)-(4) can be also labelled as *Nascent entrepreneurship.*[22]

[19] Below we will use enterprise to mean any for of organization, business firm or otherwise, which is affected by the object of entrepreneurial activity.
[20] See also Reynolds and White (1997) and Reynolds (1997)

One can see entrepreneurship in a static perspective, where the main indicator is the percentage of self-employed in the economy, and in a dynamic perspective, where entrepreneurs are agents of change who start new businesses, make them grow, change nature - new products or markets, new techniques or new organization – or increase their profitability.[23] In the dynamic perspective one can measure nascent entrepreneurs – who have made the decision to launch a firm -, gross and new entry of business start-ups, as well as the turbulence rate – the sum of entry and closures. Entrepreneurial activity that takes place within existing corporations - corporate entrepreneurship or "intrapreneurship" - are very extremely difficult to measure

In Figure 2 we present the key dimensions of entrepreneurial activity as we see them. The first stage, or *Enterprise Creation* considers all the activities between latent desire for entrepreneurial activity, idea selection and conception up to the formal creation of a new organization. In the framework above, steps (1) to (3) with the explicit addition of the moment of firm creation. The second possible moment of entrepreneurial activity is reflected in *Enterprise Performance*, either in the form of *Survival*, the very basic measure of performance, *Growth*, *Return* and internal change – qualitative or otherwise – which we group under the term *Intrapreneurship*.

Figure 2. Measuring Entrepreneuship



4.1. Enterprise Creation

As *Enterprise Creation*, we would ideally aim to measure the quantity and quality of ideas created in an economy. The complex nature of the creation of new enterprises must be considered, including the different stages through which an idea goes through until it materializes in a new organization. [24] As an example, a given country may have a substantial level of entrepreneurial activity that does not materialize in new firms. Through surveys, one can in principle break down the probability that a given individual is at any of the different stages of enterprise creation and break down those probabilities by different population

[21] These authors consider stages (8) and (9) as *Past entrepreneurship* and stages (7)-(9) as *Decadent entrepreneurship*. Here we will consider as outside the realm of entrepreneurship.
[22] See Reynolds et al., 2005
[23] See Wennekers, Uhlaner and Thurik (2000).

[24] Grilo, Thurik (2007) characterize it as "the result of a long series of complex choices. It is a process rather than the result of a single binary choice (...)". The same idea is also recognized by: Low and MacMillan (1988); Bull and Willard (1993)

subclasses (e.g. men *vs* women) or by countries. [25] Another possibility to measure activity before firm creation is the use of indicators that proxy *conception*, such as the number of patents.

One can distinguish, in a first pass, gross and net entry. A high intensity of *exits* may be the result of a natural selection of good ideas, and a high number of *step outs* a sign of a generational cycle or a faulty materialization entrepreneurs ideas. Entry and performance are, necessarily, related. For instance, several studies have found a positive correlation between initial firm size and market share with subsequent survival and growth.[26] Size at death may also be analysed over time, as well as the pattern of size-growth-return-productivity and its evolution before *exit*.[27]

4.2. Enterprise Performance

After the emergence of a new enterprise in the market, the important issue to which entrepreneurs dedicate themselves is to maintain a high performance, first by ensuring survival, next by devoting their energies to increase its size, modify its scope as needed and, ultimately, ensure high returns.

In the case of firms, when measuring size, it is important to consider two dimensions: *absolute size*, usually measured through the level of inputs used - number of employees or level of capital assets - or through the output level – sales -, and *relative size*, assessed through the market share of each firm. The latter may also suggest market concentration analysis. In addition to the analysis of the evolution of size and growth over the life-cycle, it may be important to investigate the presence of a minimum scale of operation in each industry.

Intrapreneurship, as mentioned above, is hard to measure but may be approached through the persistence of a high market value in the case of business firms, possibly complemented by direct measurement of observables, such as the number of patents produced or resources devoted to R&D. Both, however, have the important limitation of measuring technology-based intrapreneurship only, as highlighted inGeroski, Machin and Walters (1997).

As to return, we can consider it at two different levels: *venture level*, concerning the return to shareholders or even stakeholders, and at *societal level*, when the impact of the enterprise in the economy or society as a whole are considered.[28] An obvious example of divergence between the two is the case of an enterprise that explores rents in the form of illegal and criminal activities. At the *venture level*, one can use objective variables - net sales profits, sales turnover, employment, investment, export share, return on equity or assets[29]-, as well as subjective variables - assessed through interviews and surveys that assess personal fulfilment.[30] As to the societal level, entrepreneurship not only is often associated with beneficent or social activities, but it may also involve substantial spillovers that should be taken into consideration. This is key in the case of knowledge creation that enlarges a general pool that is accessible to many companies without significant additional costs. Still, assessing the societal-level contribution of entrepreneurial activity is one of the most challenging tasks in the entrepreneurship literature.[31]). Indicators such as employment generation and innovativeness are regarded as the most 19 appropriate for this purpose (e.g. Johnson et. al., 2000) as these are considered to be the main contribution of SMEs to economic development. Scholars have explored the empirical link between firm performance and the growth of an economy – see Thurik (1999), Audretsch et al. (2002).

[25] Grilo and Thurik (2007), for instance, use multinomial choice model and data from two Entrepreneurship Flash Eurobarometer surveys (2002 and 2003), covering the 15 earlier European Union member states plus Norway, Iceland, Liechtenstein and US. They consider the stages of outsider, conception and gestation.
[26] See Caves (1998)
[27] Caves (1998) presents some patterns.
[28] Sauka (2008).
[29] Geroski, Machin and Walters (1997), for instance, use return on sales
[30] See Sauka (2008)
[31] Sauka and Weler (2007).

5. CONCLUDING REMARKS

This paper will provide a systematic discussion of the characteristics of entrepreneurs in terms of motivation for entrepreneurial activity, as well as different measures of the latter. Our aim is to make progress in clarifying a very complex concept, whose many facets have often been neglected in the literature.

REFERENCES

ACS, Z.; AUDRETSCH, D. (2003): (Eds.), Handbook of entrepreneurship research: an interdisciplinary survey and introduction, Boston, Mass., London: Kluwer Academic.

BAUMOL, WILLIAM (1993b), "Formal Entrepreneurship Theory in Economics: Existence and Bounds," Journal of Business Venturing 8, 197210.

BAUMOL, WILLIAM (1990) "Entrepreneurship, Productive, Unproductive and Destructive," Journal of Political Economy, 98(5): 893921.

BAUMOL, WILLIAM J. (1968), "Entrepreneurship in Economic Theory," American Economic Review, May, 6471.

CANTILLON, RICHARD (1931), « Essai Sur la Nature du Commerce en Général », Edited and translated by H. Higgs, London: MacMillan.

CASSON, MARK (1982), The Entrepreneur: An Economic Theory, Barnes & Noble Books, Totowa, New Jersey.

CAVES, RICHARD E. (1998), "Industrial Organization and New Findings on the Turnover and Mobility of Firms", Journal of Economic Literature, 36(4), December, pp. 19471982.

COLE, ARTHUR H. (1968), "The Entrepreneur: Introduction Remarks," American Economic Review, May, 6063.

Chandler

GALBRAITH, JOHN K. (1967), "The New Industrial State", London: Routledge.

GARTNER, WILLIAM, and NANCY CARTER (2004), "Entrepreneurial Behavior and Firm Organizing Processes", in Acs and Audretsch (eds.) "Handbook of Entrepreneurship Research", Boston, Kluwer Academic Press.

GRILO, I. and J.M. IRIGOYEN, 2006, Entrepreneurship in the EU: to wish and not to be, Small Business Economics 26(4), 305318.

GRILO, I. and THURIK, R., 2007, "Determinants of entrepreneurial engagement levels in Europe and the US", Working Paper, Max Planck Institute of Economics

HÉBERT, R. F., and A. N. Link, 1989, 'In Search of the Meaning of Entrepreneurship', Small Business Economics, 1, 39–49.

KIRZNER, ISRAEL M., 1979, Perception, Opportunity and Profit: Studies in the Theory of Entrepreneurship, Chicago: University of Chicago Press.

KIRZNER, ISRAEL M., 1997, Entrepreneurial discovery and the competitive market process: an Austrian approach, Journal of Economic Literature 35, 6085.

KNIGHT, FRANK H., 1921, Risk, Uncertainty and Profit, New York: Houghton Mifflin.

KOOLMAN, G. (1971), "Say's Conception of the Role of Entrepreneur," Economica, New Series, Vol. 38, No. 151, August, 269286.

LAZEAR, EDWARD (2005), "Entrepreneurship, "Journal of Labor Economics, vol. 23, no. 4, pp. 649680.

LEIBENSTEIN, HARVEY (1968), "Entrepreneurship and Development", American Economic Review 58, 72–83.

LEIBENSTEIN, HARVEY (1979), 'The General XEfficiency Paradigm and the Role of the Entrepreneur', in Mario Rizzio (ed.), Time, Uncertainty and Disequilibrium, Lexington: Heath, 127–139.

MARSHALL, ALFRED (1961), Principles of Economics, 9th edition, London: Macmillan.

MCCLELLAND, DAVID C. (1971) "The Achievement Motive in Economic Growth," in Kilby, Peter ed. Entrepreneurship and Economic Development, The Free Press, New York.

PIORE, MICHAEL and C. SABLE (1984), "The Second Industrial Divide:Possibilities for Prosperity", New York: Basic Books.

SAUKA, A. (2008), "Productive, Unproductive and Destructive Entrepreneurship: A Theoretical and Empirical Exploration", William Davidson Institute Working Papr Number 917

SAUKA, A.; WELTER, F. (2007): Productive, unproductive and destructive entrepreneurship in an advanced transition setting: the example of Latvian small enterprises, in Dowling, M.; Schmude, J. (Eds.), Empirical entrepreneurship in Europe, Cheltenham: Edward Elgar, pp. 87111.

SCHUMPETER, JOSEPH A. (1934), "The Theory of Economic Development", Cambridge, Mass: Harvard University Press.

SCHUMPETER, JOSEPH A. (1942), "Capitalism, Socialism and Democracy", New York: Harper & Brothers

SMALLBONE, D.; WELTER, F. (2001): The distinctiveness of entrepreneurship in transition economies, in Small Business Economics, 16:4, pp. 249262.

SMALLBONE, D.; WELTER, F. (2006): Conceptualising entrepreneurship in a transition context, in International Journal of Entrepreneurship and Small Business, 3:2, pp. 190206.

STAUSS, JAMES H. (1944), "The Entrepreneur: The Firm," The Journal of Political Economy, Vol. 52, Issue 2, June, P 112127.

THURIK, ROY, and ISABEL GRILO (2005), "Latent and Actual Entrepreneurship in Europe and the U.S.: Some Recent Developments", Working Paper, MaxPlanck Institute of Economics.

WEBER, MAX (…), "The Protestant Ethic and the Spirit of Capitalism".

WENNEKERS, SANDER, and ROY THURIK (1999), "Linking Entrepreneurship and Economic Growth", Small Business Economics, 13, pp. 27 – 55.

WENNEKERS, SANDER, UHLANER, M., and ROY THURIK (2002) "Entrepreneurship and its Conditions: a Macro Perspective", International Journal of Entrepreneurship Education 1(1): 2565.

PRÁTICAS DE CONTROLO DE GESTÃO. O CASO DAS PME'S.

TIAGO ANDRÉ SERRA GONÇALVES
E-mail: tiago.andre25@hotmail.com
JOANA FIDALGO VAZ REIS SILVA
E-mail: fidalgo.jo@gmail.com
Departamento de Gestão e Economia
Universidade da Beira Interior

Resumo

A crescente internacionalização da gestão empresarial promoveu a introdução de sistemas de controlo, essencialmente financeiros, que passaram a ser institucionalizados e considerados como sistemas de controlo por excelência. Para muitos autores os sistemas de controlo de gestão são processos através dos quais os gestores asseguram que os recursos são obtidos e usados de forma eficaz e eficiente no acompanhamento da obtenção dos objectivos propostos pela empresa. Este trabalho procura incidir sobre o impacto das medidas não financeiras no controlo de gestão, partindo dos instrumentos tradicionais e fazendo a ponte com novas ferramentas de controlo de gestão, procurando entender o comportamento das PME's que adoptam um sistema de controlo de gestão para a obtenção da desejada afirmação no mercado global, bem como os factores que influenciam esse mesmo sistema.

Palavras-chave: Controlo de Gestão, PME's, Sistemas de Controlo, Instrumentos Não Financeiros.
Área Temática: Economia Empresarial.

Abstract

The increasing internationalization of business management promoted the introduction of control systems, mainly financial ones that have become institutionalized and considered as control systems for excellence. For many authors the systems of management control are processes through which managers ensure that resources are obtained and used effectively and efficiently to monitor the achievement of the objectives proposed by the company. This work is focused on the impact of non-financial measures in management control, from traditional instruments to new tools of management control, trying to understand not only the behaviour of SME's which adopt systems of management control as a way to obtain the desired assertion in the global market, but also the factors which influence those systems.

Key words: Management Control, SME, Control Systems, Non-financial Measures.
Thematic Area: Entrepreneurial Economics.

1. INTRODUÇÃO

No decorrer deste estudo propomo-nos tentar entender a importância e a pertinência da aplicação do controlo de gestão como ferramenta auxiliar do gestor nos dias de hoje, estudando a sua aplicação concreta nas empresas de pequena e média dimensão. O estudo será feito de acordo com as seguintes etapas: numa primeira fase iremos definir os conceitos de "gestão" e "controlo", fazendo a ponte com o conceito "controlo de gestão"; numa fase seguinte iremos analisar sucintamente a evolução deste último conceito ao longo do tempo, desde os seus fundamentos na própria contabilidade de gestão até ao conceito de controlo de gestão nos dias de hoje; numa terceira fase o objectivo será definir o conceito e as características das PME's, assim como entender o porquê deste estudo incidir no caso concreto deste tipo de empresas; posteriormente, tendo em conta os tipos de ferramentas existentes no controlo de gestão,

financeiras e não financeiras, daremos uma breve definição de ambas e centrar-nos-emos nas ferramentas não financeiras; por último, faremos uma análise aos factores contingênciais que influenciam directamente a escolha dessas mesmas ferramentas nas PME's.

2. CONCEITOS-CHAVE

2.1 Controlo

A importância do controlo na organização é explicada por Antony e Dearden através da metáfora *When the brake pedal is pressed, an automobile slows down or stops. When the accelerator is pressed, the automobile goes faster. When the steering wheel is rotated, the automobile turns. Whit these devices, the driver controls the speed and direction of the vehicle. Without them, the automobile would not do what the driver wanted it to do; that is, it would be out of control*[1] (Antony e Dearden, 1975:3). Numa organização o paralelismo adequa-se, embora na opinião dos autores o processo seja bem mais complicado.

Diversos autores (Williams *et al* (1982)) descreveram os passos necessários para que a função controlo seja uma realidade nos planos da empresa: o primeiro consiste em estabelecer as expectativas para que haja um caminho orientador; de seguida devem-se prever avaliações de *performance* ao longo do processo de modo a medir o alcance das expectativas previamente definidas; por fim devem-se tomar as medidas adequadas quando o caminho escolhido não está a corresponder, ou como medida de prevenção de possíveis adversidades. Numa vertente mais social, o controlo organizacional foi definido como *todas as acções ou actividades feitas em avaliação para influenciar a probabilidade do pessoal se comportar nos diversos cenários que conduzem à realização dos objectivos organizacionais* (Flamholtz (1983) citado por Fisher, 2005). Para Fisher (1995) o controlo trata-se de uma ferramenta capaz de criar as circunstâncias que motivam a organização a atingir os resultados previamente desejados. Já Herath (2007) dá o seu contributo considerando o controlo como uma importante função dos gestores, e cita ainda Merchant (1985) ao referir que se trata ainda de uma forma de *manter as coisas no bom caminho* e como *a última função no processo de gestão*.

O controlo foi caracterizado tendo em conta diversos cenários, formais e informais, assumindo a forma de controlo comportamental, de mercado, burocrático, administrativo, de resultados e de acções (Smith, 1997) em suma um elemento chave para minimizar a tensão nas empresas, especialmente nas mais novas (Davila, 2005).

Ao nível da gestão contabilística, o controlo é um dos conceitos mais importantes e complexos (Scheytt *et al*, 2003) visto que não é totalmente compreendido em termos da sua importância no contexto internacional. Este conceito é interpretado de forma diferente nas diversas culturas europeias, o que deriva por um lado da diferente percepção das práticas de gestão contabilística e de controlo, e por outro da maneira como essas práticas podem ser transferidas e percebidas.

Ao longo da presente exposição e seguindo Herath (2007), os termos "controlo" e "sistema de controlo" serão utilizados para dar sentido ao controlo de gestão e à própria gestão dos sistemas de controlo.

2.2 Gestão

No decorrer do estudo do conceito é possível verificar que "gestão" não tem uma definição universalmente aceite. No entanto, existe algum consenso no que diz respeito a um determinado conjunto de tarefas que lhe devem estar inerentes, procurando garantir a afectação eficaz de todos os recursos disponibilizados pela organização com o intuito de se atingirem os objectivos.

No início do século XX (1906), através da teoria da Administração Científica, Taylor diz-nos que o conceito se realiza partindo de incentivos financeiros e da separação das tarefas de planeamento, entre quem executa e o responsável, demonstrando já uma preocupação na

[1] Quando se carrega no travão, um automóvel fica mais lento ou parado. Quando se carrega no acelerador, o automóvel anda mais rápido. Quando se gira o volante, o automóvel muda de direcção. Com estes dispositivos, o condutor controla a velocidade e a direcção do veículo. Sem eles, o automóvel não iria fazer o que o motorista desejaria fazer, isto é, estaria fora de controlo.

uniformização metodológica no controlo de todas as organizações, considerando de certa forma o Homem como uma máquina. Em 1916 Fayol dá-nos uma nova visão baseada na teoria da Administração Industrial, considerando que para o bom funcionamento de uma organização é necessário prever, organizar, liderar, coordenar e controlar.

Mais tarde, em 1922, Max Weber vem referir a importância de uma autoridade carismática que contribua para existir um poder baseado na tradição, exercendo o dito controlo burocrático.

Decorria o ano de 1928 quando Urwich, citado por Amat e Gomes (2004), indicou cinco princípios indispensáveis na concepção de gestão: responsabilidade, evidência, uniformidade, comparação e utilidade, em suma, a mecânica e a dinâmica da administração.

Uma das evoluções mais significativas ao longo da história foi o Movimento Humanista associado a Elton Mayo que, em 1933, defendeu que os trabalhadores se comportam da melhor forma se forem regidos por valores e factores que os motivem. Na mesma década Glover e Maze concluíram que era necessário existir uma avaliação de *performance* e relacionar o controlo de gestão com a análise de custos. Esta perspectiva é o ponto de partida para a óptica financeira, nomeadamente no âmbito do controlo de gestão. Na década de 40, Holgen, Fish e Smith enunciaram que a gestão é uma responsabilidade primária da alta direcção das empresas, partindo de objectivos, procedimentos e avaliações. Na década seguinte Amat e Gomes (2004) referem que se introduziram novas inovações no sentido de um programa-orçamento, orçamento base-zero, divisionalização, análise custo-benefício e contabilidade de gestão por projectos.

Uma organização tem de ser dirigida e comandada em concordância com os seus objectivos. Neste sentido cabe à gestão fazer com que a estrutura organizacional e os seus objectivos estejam em sintonia (Cleland e King, 1972). O conceito de gestão que reúne maior consenso é partilhado por DuBrin (2000) e Herath (2007), que olham para a gestão como o processo que uma organização tem ao seu dispor (um recurso), que permite alcançar objectivos tendo em conta as fases de planeamento, organização, liderança e controlo. A importância da gestão numa organização deve-se essencialmente às ideias de negócio e à capacidade intrínseca do gestor para gerir o sucesso, alcançando os objectivos. Constatamos então que a gestão não é mais do que a optimização do funcionamento de uma organização, que através da tomada de decisões baseadas num processo de recolha e tratamento prévio de dados/informação relevante, se espera um contributo positivo para o desenvolvimento da organização, assim como a garantia de satisfação da sociedade em geral ou de um grupo em particular.

3. PORQUÊ AS PEQUENAS E MÉDIAS EMPRESAS?

Devido à diminuição das barreiras comerciais, ao maior acesso à informação e aos avanços tecnológicos, verificaram-se mudanças importantes na forma de gestão das empresas, pelo que estas, para sobreviverem no actual contexto económico, são obrigadas a formalizar os seus processos de planeamento e o controlo das suas operações (Lavarda, 2007).

Para Nobre (2001) uma PME é uma empresa independente de empresas do grupo ou de grandes empresas, de carácter familiar, com menos de 500 trabalhadores. Neste sentido, o estudo centra-se nas PME's uma vez que em Portugal, em 2005, existiam 296 928 mil empresas deste género representando 99,6% do tecido empresarial, o que corresponde a 75,2% do emprego gerado nesse ano, realizando 56,4% do volume de negócios nacional (170,3 mil milhões de euros), de acordo com os dados de 2005 do Instituto de Apoio às Pequenas e Médias Empresas e à Inovação – IAPMEI.

De acordo com a definição da União Europeia (Recomendação da Comissão 96/280/CE, de 3 de Abril), considera-se PME a empresa que tiver menos de 250 trabalhadores, volume de negócios anual que não exceda 40 milhões de euros ou balanço total anual que não exceda 27 milhões de euros e que não sejam propriedade, em 25% ou mais, do capital ou dos direitos de voto de uma empresa ou, conjuntamente, de várias empresas que não se enquadrem na definição de PME ou de pequena empresa, conforme seja o caso (critério de independência).

Russo (2006) reconhece como características específicas das PME's a estrutura organizacional simples; pouca delegação de poder de decisão por parte do dirigente aos subordinados; geralmente os sócios acumulam o cargo de gestores; recursos materiais,

financeiros e humanos escassos; poder negocial reduzido em relação a fornecedores, clientes, banca, entre outros; falta de informação pública e auditada; a sua continuidade depende de uma pessoa; operacionalidade predominantemente regional ou nacional e a sua localização são determinadas por critérios afectivos ou de residência. Jones (2003) refere a flexibilidade como a grande virtude das PME's e que as demais qualidades advêm desta. Segundo Lavanda (2007) as PME's possuem características que justificam uma linha de investigação concreta, devido essencialmente à sua estrutura organizacional e à sua flexibilidade, e também por terem uma gestão pouco burocrática que proporciona rapidez de resposta, capacidade de inovação e adaptação.

No entanto, a variável preponderante que identifica uma PME é o seu dirigente, que vai influenciar decisivamente não só a estrutura, como também a própria cultura organizacional (Abi Azar, 2007). A sua estrutura é tradicionalmente simples o que leva a que haja pouca especialização de tarefas, à existência de menos processos formalizados e a que as decisões sejam centralizadas no dirigente. Relativamente à cultura organizacional, Garengo e Bititci (2007) consideram que advém directamente do dirigente, do seu perfil e maneira de ser.

Entre os autores que se têm dedicado ao estudo desta temática, destacam-se Nobre (2001), Lavigne (2002), Abi Azar (2007), Condor e Rebut (2007), Garengo *et al* (2007), Meyssonnier e Zawadzki (2007), Garengo e Bititci (2007) e Lavarda (2008).

4. EVOLUÇÃO DO CONTROLO DE GESTÃO

A base do controlo de gestão numa PME deriva necessariamente da contabilidade de gestão, como tal, torna-se pertinente analisar os seus fundamentos, a evolução e a sua importância para o controlo de gestão.

Segundo Fournier (1992), citado por Nobre (2001), a função de controlo de gestão nas PME's é muitas vezes confundida com a função financeira, uma vez que neste tipo de empresas não existe uma diferenciação clara entre as funções de recursos humanos, comercial, produção e área financeira, todas elas asseguradas pelo controlo de gestão.

Dent (1953), citado por Amat e Gomes (2004), refere que o controlo de gestão está relacionado com a análise de orçamentos, tendo por base o planeamento de actividades, a delegação de autoridade, a delegação de autoridade associada à responsabilidade e a coordenação. Anthony (1965), citado pelos mesmos autores, refere que o controlo de gestão se trata do processo através do qual os administradores asseguram que os recursos são obtidos e usados eficaz e eficientemente. Na década de 60 começam a emergir alguns modelos quantitativos de controlo de gestão, influenciados por Taylor. Na década seguinte muitos foram os autores que se dedicaram ao estudo desta temática, entre eles, Lorange e Mortur, citados por Amat e Gomes (2004), consideram que o controlo de gestão tem como finalidade auxiliar a administração na concretização dos objectivos, através de uma estrutura formal. Por sua vez, Lowe, citado por Otley e Berry (1980), define controlo de gestão como um sistema que assegura a adaptação da empresa às possíveis mudanças ambientais, e que garante simultaneamente que o comportamento dos trabalhadores segue de encontro aos objectivos propostos pela empresa, demonstrando a integração necessária entre controlo e a própria estrutura da empresa. Já Famholtz, citado por Amat e Gomes (2004), complementa referindo que a temática abordada se trata de um processo que influencia o comportamento dos membros de uma organização formal para atingir os próprios objectivos da organização.

Em 2005 Jordan *et al* consideravam que faz parte das competências do controlo de gestão desenvolver instrumentos de gestão, concebidos pelos próprios gestores que os utilizavam no âmbito da estratégia definida pela empresa. Por se encontrar directamente dependente do gestor, nesta fase o controlo de gestão assume uma perspectiva contingêncial.

Apesar de existirem diversas definições para o conceito de controlo de gestão, prevalece a ideia de que se trata de *estimular, responsabilizar, gerir com rigor e êxito* (Jordan *et al*, 2005:17). Para estes autores existem alguns princípios que se associam ao conceito de controlo de gestão. Os objectivos da empresa são de natureza diversa, ou seja, os instrumentos de controlo não se referem apenas à dimensão financeira; procura-se descentralizar as decisões e

delegar autoridade; existe convergência de interesses em toda a empresa; os instrumentos são concebidos tendo em conta o futuro e não apenas o passado; o controlo actua essencialmente mais sobre os Homens do que sobre os números; existem sanções e recompensas; num primeiro impacto são os responsáveis operacionais que estão afectos ao controlo de gestão e posteriormente os próprios controladores de gestão. Os mesmos autores defendem ainda que o desenvolvimento ou a estagnação do sistema de controlo de gestão depende essencialmente da cultura da empresa.

Jordan *et al* (2005) consideram existir três etapas fundamentais para o sucesso do controlo de gestão. Numa primeira fase estabelecem-se os objectivos a realizar através de planos de acções, devidamente quantificados, escolhendo os meios mais adequados e compatíveis para o fazer; de seguida é necessário interpretar resultados e averiguar se existem desvios de acordo com o que foi previamente estabelecido e se necessário tomar decisões correctivas; por último, ao alcançar os objectivos previamente estabelecidos realiza-se uma avaliação de *performance* aos diversos intervenientes do processo.

Em suma, o controlo de gestão trata-se de uma poderosa ferramenta que os gestores têm ao seu dispôr e cuja utilização revela ser bastante pertinente no âmbito da organização, não só pelo facto de se basear fortemente nos principios incutidos pela própria contabilidade de gestão, mas também porque se encontra associado intrinsecamente a algumas etapas fundamentais, como referimos anteriormente, pelo que faz todo o sentido aprofundar o seu estudo e a sua contribuição para as empresas. Neste sentido, as ferramentas de controlo são conceitos construtivos, como é o caso das taxas de actualização e do custo completo, que se materializam após a formalização contabilística, financeira, quantitativa e qualitativa (Nobre, 2001).

Para Jordan *et al* (2005) o gestor tem à sua disposição diversos instrumentos tais como o plano operacional, o orçamento, o controlo orçamental e o *tableaux de bord*, para além da análise estratégica e do plano estratégico. O plano operacional consiste em definir as políticas de comercialização, de investimento e gestão da produção, de organização e gestão dos recursos humanos e de gestão financeira, diferindo do plano estratégico por ser de médio prazo e ter um carácter quantitativo (através de balanços e contas de resultados previsionais), enquanto que o plano estratégico tem uma visão de longo prazo e é sustentado pelo próprio plano operacional. A fase de orçamentação tem de ser devidamente articulada com o plano operacional, e caracteriza-se pela fixação de objectivos a curto prazo em concordância com os objectivos previamente estabelecidos pela empresa para que a definição, a coordenação e a aprovação dos planos de acção, em todas as áreas da empresa, não suscitem confusão (Jordan *et al*, 2005). A análise estratégica consiste em analisar o ambiente da empresa, da concorrência, dos clientes, dos fornecedores, ameaças a novas entradas, produtos substitutos, dos elementos da cadeia de valor, para além das forças e fraquezas da empresa – análise *SWOT*.

Jordan *et al* (2005) referem que o plano operacional e o orçamento são dois instrumentos previsionais que necessitam de ferramentas de acompanhamento dos resultados, tais como os *tableaux de bord* ou o *balanced scorecard* (ferramentas de substituição que permitem acompanhar os resultados com maior eficiência e eficácia) e o controlo orçamental, cuja principal função é comparar os resultados num determinado período com as previsões inicialmente elaboradas, determinando possíveis desvios e analisando-os para definir a natureza e a intensidade das acções correctivas necessárias (Williams *et al*, 1982). Esta ferramenta apresenta algumas lacunas, por analisar apenas a dimensão financeira e ser um instrumento lento, uma vez que os dados nele contidos são oriundos da contabilidade de gestão.

Simon (2007) define contabilidade de gestão através de duas perspectivas: considerando-a como uma simples vertente da contabilidade, ou vendo-a como um elemento do processo estratégico de tomada de decisões. Scapens (2006) começa por fazer uma retrospectiva sobre a contabilidade de gestão apartir da década de 70. Nesta altura verificou-se que era feita com base em análises económicas marginais e modelos quantitativos, fazendo uso de modelos simples, complexos e matemáticos, aproximando a contabilidade de gestão à Economia (Scapens, 1984). Faz parte das competências da gestão por em prática estes modelos, depois de devidamente explicados e demonstrados. Scapens (2006) centrou o seu estudo no modelo do lucro económico e na sua importância para a contabilidade de gestão, concluindo posteriormente que não tinha conteúdo prático.

Em suma, na década de 70 a contabilidade de gestão assentava em modelos matemáticos orientados para a economia, nos quais se enfatizava a teoria relativamente à prática.

Johnson e Kaplan (1987), citados por Burns e Vavio (2001), afirmaram que o controlo de gestão estaria em crise, visto que a informação que chegava ao responsável não era suficiente para este realizar a sua tarefa. Sugeriram ainda que o caminho a seguir passaria pela aposta nas novas tecnologias. Sulaiman e Mitchell (2005) revelaram que a contabilidade de gestão entrara numa fase de mudança devido à sua falta de prática. Aliás, para estes autores a mudança existe sempre e é plausível enquanto houver crescimento económico, pelo que no estudo elaborado constataram a importância da pesquisa na contabilidade de gestão principalmente, em termos de sistemas de planeamento. Introduziram-se medidas, tais como a imposição da contabilidade de gestão nas empresas, realização de entrevistas à gerência e novos estudos. Os métodos de controlo de despesas das empresas, nesta altura, baseavam-se no retorno do investimento e/ou nos orçamentos. Enquanto isso Scapens continuava a usar o lucro económico. Devido ao interesse na matéria e após estudos exaustivos, passou a haver uma diversidade de práticas sem qualquer paralelismo entre si (Scapens, 2006), o que proporcionou que o autor centrasse o seu estudo nesta matéria, procurando entender o porquê de existir tanta diversidade de práticas de contabilidade de gestão. A essência da contabilidade de gestão surge nos anos 90 (Scapens, 2006). Foram realizados novos estudos para investigar o papel da contabilidade de gestão no controlo de uma organização, no entanto detectou-se que a relação entre a prática e as teorias económicas convencionais eram úteis mas inconclusivas. Deste estudo retiram-se duas conclusões essenciais: cada empresa tem a sua maneira de ser e adequa a contabilidade de gestão à sua medida sem seguir modelos teóricos; por outro lado, os modelos teóricos económicos e neoclássicos servem apenas para estudar a prática da contabilidade de gestão.

A pesquisa realizada por Scapens permitiu constatar que algumas ferramentas ainda são estudadas nos dias de hoje, tais como o método ABC (*Activity-based Cost*), o *Balanced Scorecard* entre outros (Sulaiman e Mitchell, 2005). Considerando as diversas metodologias, teorias e dimensões práticas relevantes até ao início do novo milénio, o quadro seguinte sintetiza a evolução da contabilidade de gestão do ponto de vista de Scapens (2006).

Quadro n.1 – Resumo histórico da Contabilidade de Gestão

	Metodologia	Teoria	Dimensão Prática
1970s	Modelos Matemáticos	Económica	O que devem os gestores fazer?
1980s	Positivista	Contingêncial	O que faz o gestor?
1990s	Interpretativa	Estruturada	Pôr em prática
2000s	Pluralista/Pragmática	Institucional	Ajudar os investigadores

Fonte: Adaptado de Scapens (2006)

Burns e Vavio (2001), apoiados por Sulaiman e Mitchell (2005), consideram a mudança da contabilidade de gestão como um interesse actual devido aos avanços nas tecnologias de informação em direcção à inovação. Os sistemas de planeamento têm que ter uma base em tecnologia avançada como a *internet* e o *e-commerce*, e só assim se entende que o controlo de gestão tem de acompanhar esta mudança. O objectivo será munir a contabilidade com elementos de estratégia, tecnologias de informação e conceber a gestão como algo mutável. Burns e Scapens (2000) consideram que as rotinas organizacionais e os processos que as mesmas impõem à contabilidade de gestão são motivos preponderantes na mudança. Os autores, citando Gralund e Lukka (1998), consideram que em muitas organizações a contabilidade de gestão se está a descentralizar e a ser eliminada das áreas de negócio, alterando a perspectiva do controlo de gestão como uma ferramenta de controlo de negócios para se tornar na base de sustentação do mesmo. Para Burns e Vavio (2001) a contabilidade de gestão necessita de incorporar elementos como a estratégia, tecnologias de informação, capacidades de gestão para lidar com a mudança e ainda alguns aspectos de negociação. Estas medidas devem ser incorporadas tendo em conta três perspectivas na mudança relativamente a esta temática: a natureza epistemológica da mudança; a lógica da mudança, que varia consoante a estratégia adoptada, e por fim a gestão

da mudança como forma de acompanhar a tecnologia. A contabilidade analítica pretende atenuar as insuficiências da contabilidade geral, também denominada contabilidade financeira enquanto utensílio de gestão. Embora a contabilidade geral seja essencial e imprescindível para se levar por diante uma gestão racional, carece de operacionalidade na planificação, no controlo e na tomada de decisões pelos gestores e empresários.

5. FERRAMENTAS FINANCEIRAS

Os sistemas de controlo de gestão estão num processo de transição (Oatley, 2003). Actualmente a globalização pode ser associada a um vasto conjunto de transformações, que introduzem um novo paradigma tecno-económico, marcado pela ampliação da concorrência, pela flexibilização dos mercados e pela utilização de novas tecnologias de informação. A compilação destes factores origina a que diversas empresas aumentem o uso de ferramentas e sistemas de controlo de gestão.

O **orçamento** é o sistema de controlo de gestão mais tradicional das organizações como relatam Hope e Frase (2003) sustentados por Bajan-Banaszak (1993) e Nobre (2001). É importante para o planeamento (Azar, 2006), para o controlo e para o desempenho (Hansen *et al*, 2003). Os orçamentos não podem ser eficazes isoladamente e são apenas uma característica da estrutura organizacional. Podem fornecer os meios para traduzir os objectivos das organizações com estruturas detalhadas (Thornbory, 2007). *O orçamento é um instrumento de gestão de apoio ao gestor no processo de alcançar os objectivos definidos pela empresa, ou seja, um instrumento de decisão e de acção* (Jordan *et al*, 2005:77). Para os autores, é necessário distinguir o controlo de gestão do processo orçamental pois estes conceitos facilmente se confundem. Um orçamento é um instrumento técnico de controlo de gestão, que auxilia os responsáveis para dominar a gestão, desde os responsáveis da produção até ao próprio controlador de gestão. Schick (2007:110) considera o orçamento como *um processo que transforma informação em decisão*. O orçamento como ferramenta auxiliar da gestão possui diversas perspectivas de aplicação no seio de uma organização. Pode ser encarado como um instrumento de descentralização, porque a elaboração de um orçamento é feita através da colaboração dos diversos departamentos da empresa, fazendo com que a organização se encontre em perfeita concordância com a estratégia delineada. Para tal, o controlador de gestão tem um papel complexo a desempenhar. Incentivar e auxiliar os operacionais a realizar planos de acção economicamente viáveis, ajudar os operacionais e os seus superiores hierárquicos a avaliar esses planos de acção, assim como promover o diálogo entre o gestor operacional e o seu superior hierárquico, em suma, a elaboração do orçamento compete não ao controlador de gestão mas sim à linha hierárquica implicada. O orçamento pode ser visto como um instrumento para o planeamento, visto que ao longo do processo orçamental é necessário seleccionar o número de oportunidades tendo em conta os objectivos e o plano estratégico adoptado. Pode também ser encarado como um instrumento de motivação, por exemplo, incentivando o gestor a realizar melhores resultados e a estimular a satisfação pelas tarefas que desempenha. Outra vertente possível do orçamento é a coordenação, porque é necessário articular e manter a coerência entre os objectivos ao longo da linha hierárquica vertical e dos diversos departamentos. Por último, outra abordagem possível do orçamento é a capacidade de avaliação, uma vez que é necessário existir um cuidado permanente em acompanhar os resultados a realizar relativamente aos objectivos que previamente se propuseram. Os orçamentos tradicionais (Thornbory, 2007) são anuais, directos e baseados geralmente em dados de anos anteriores. São relativamente simples de preparar e permitem a coordenação através dos departamentos. Entretanto, discute-se que estes tipos de orçamentos não se ajustam para um negócio com um tempo algo indeterminado ou que sejam voláteis. Para tal urge a necessidade de criar alternativas como os "orçamentos rolantes" onde são revistos continuamente, os orçamentos flexíveis em que os rendimentos e os custos são ajustados durante o período de orçamento, o orçamento base-zero em que o processo incluído no orçamento inteiro começa como se fosse preparado pela primeira vez.

Nos últimos tempos vem-se verificando um acérrimo interesse no orçamento, nomeadamente na sua aplicação nos dias de hoje. Autores como Wallender (1999), Jensen (2001), Hope e Frase (2003), Neely, Bourne e Adams (2003) e Scarlett (2007) põem em causa a utilidade do orçamento. *Beyond Budgeting* é o nome genérico dado a um conjunto de práticas que se destinam a substituir a orçamentação.

Wallender (1999) apoiado por Samuelson (1999) tem uma visão mais radical da temática, considerando que os orçamentos convencionais nunca poderão ser válidos porque não são capazes de capturar a incerteza envolvida em ambientes voláteis, relativamente à mudança, e impedem frequentemente a perseguição de objectivos estratégicos suportando práticas mecânicas. Jensen (2001) afirma que o orçamento é uma piada e que todos sabem disso, visto que consome muito tempo, encoraja os seus elaboradores à mentira e à batota, inflaciona resultados, ou seja, impossibilita o relato da veracidade duma empresa. Num plano comportamental, o orçamento para o autor conduz a motivação dos trabalhadores a favor dos interesses da empresa em que o melhor é aquele que detiver menor custos e este cenário repete-se todos os anos visto que se torna uma rotina. Esta situação torna-se mais gravosa para empresas cotadas, manipula-se informação para que a rentabilidade da empresa seja maior. O orçamento é uma prática habitual das empresas mas devia ser abolido para Hope e Frase (2003). Neely, Bourne e Adams (2003) referem que 80% das empresas estão insatisfeitas com os seus processos de planeamento e orçamentação. Estes arrastam um montante considerável de juros em "reengenharia". O método tradicional de orçamentação é demasiado moroso, dispendioso, não acompanha a competitividade e a turbulência do mercado (barreira à mudança) e é contra producente, na medida em que é geralmente afectado por jogo, política e tácticas empresariais. Para além destas lacunas, os orçamentos possuem outras deficiências para os autores, raramente têm uma orientação estratégica, são muitas vezes contraditórios, concentram-se na redução dos custos e não sobre a criação de valor, tendem a ser burocráticos e a desencorajar o pensamento criativo, são baseados em hipóteses, incentivam o jogo perverso, reforçam o comando e o controlo vertical, não reflectem a rede emergente estruturas que as organizações estão a adoptar, reforçam as barreiras departamentais, em vez de incentivar a partilha de conhecimentos e fazem com que se sintam subvalorizadas. Scarlett (2007) explica como o orçamento difere das abordagens tradicionais de controlo financeiro. A orçamentação é conhecida como uma dependência modelo, em que o recurso é orientado por planos fixos que são baseados em informações que rapidamente se tornam arcaicas. Para o autor o orçamento é o exemplo crasso para a imposição da autoridade vertical.

Esta visão pessimista não é compartilhada pela maioria dos críticos do orçamento. Existem outros autores que consideram o orçamento uma das ferramentas mais viáveis do controlo de gestão como Rankin *et al* (2003), Combes (2007), Juozapavičiūtė e Stončiuvienė (2007). Rankin *et al* (2003) considera a orçamentação uma prova inequívoca da honestidade onde inibe o comportamento indesejado numa empresa. Para os autores os orçamentos são resultados de simples negociações, em que a motivação dos subordinados é dominada por questões estratégicas. As negociações quanto mais simples forem as previsões económicas são mais sustentáveis. Combes (2007) afirma que o orçamento ainda é "rei", um dos principais trunfos do mundo financeiro. Para a autora o orçamento prevê um foco para a organização, auxilia à coordenação das actividades e facilita o controlo visto ser uma declaração quantitativa de um período de tempo que pode incluir o planeamento de receitas, de activos, de passivos e de fluxos de caixa. Juozapavičiūtė e Stončiuvienė (2007) realçam que existem alguns métodos de planeamento e avaliação de empresas, mas o mais eficaz é a orçamentação. Libby (2007) veio pôr um pouco de "água na fervura" e realizou um estudo com o intuito de verificar se os orçamentos são "uma coisa do passado" ou "um mal desnecessário" num contexto norte-americano. A contraposição de ideologias é patente nas referências por parte da autora. A orçamentação oferece aos gestores uma maravilhosa oportunidade de rejuvenescer as suas organizações. Não existe nenhum outro processo gerencial. Umapathym (1987), citado por Libby (2007), diz-se consciente de que o orçamento traduz qualitativamente a missão, declarações e estratégias corporativas em planos de acção, a curto prazo, com ligações a longo prazo, reúne gestores de diferentes níveis hierárquicos e de diversas áreas funcionais, e ao mesmo tempo proporciona continuidade com a enorme regularidade do processo. Mas de

seguida a autora contrapõe citando Welch (1995) que considera o orçamento como algo que nunca devia ter existido, é um exercício de minimização, porque o melhor é aquele que tiver um valor mais baixo. O estudo realizado por Libby (2007) realça que os orçamentos continuam a ser utilizados para fins de controlo e que existe uma pequena minoria das empresas que estão receptivas ao *Beyond Budgeting*.

Outras ferramentas foram divulgadas mas sem o mesmo destaque/uso que as anteriores como são o ***Economic Value Added*** (EVA) desenvolvido por Stern Stewart Corporation (Otley, 1999), o **método ABC** Nobre (2001), a **contabilidade analítica** Bajan-Banaszak (1993) e Azar (2006) e a **análise de desvios** Nobre (2001).

6. ENFOQUE NAS FERRAMENTAS NÃO FINANCEIRAS

A realidade empresarial demonstra que as empresas não podem centrar-se apenas no contexto financeiro interno e/ou externo. Uma empresa que pretenda ser bem sucedida, para além de definir um rumo tem de o avaliar a fim de serem tomadas as medidas correctivas necessárias. Para o conseguir não basta avaliar o desempenho financeiro. A orientação para a quantificação dos objectivos, meios e factores críticos quer internos quer externos com a finalidade de obter uma informação em síntese, a personalização de responsabilidades, o interesse em promover a comunicação e o diálogo interno e a criação de bases para a comparação com as metas previamente fixadas são princípios base para o aparecimento dos novos instrumentos de controlo de gestão anteriormente referidos que têm como horizonte a avaliação de desempenho (Jordan *et al*, 2007). Russo (2006) refere que são os factores intangíveis e não financeiros, no contexto económico actual, que garantem a sustentabilidade da capacidade competitiva das organizações. O autor afirma que os indicadores financeiros têm como medição o desempenho financeiro passado, enquanto os indicadores não financeiros servem para medir o desempenho financeiro futuro.

Nesta base surgem novos métodos, o ***Tableaux de Bord*** e o ***Balanced Scorecard*** (Bourguignom *et al*, 2001). Para os autores e Jordan *et al* (2007), estas técnicas estão no seguimento da necessidade de obter resultados complementando a parte financeira com a avaliação de desempenho de pessoas, departamentos ou organizações. Anthony, Dearden & Vancil (1976), Parker (1979), Merchant (1985), Schoenfeld (1986), Eccles (1991) citados por Bourguignom *et al* (2001) demonstram que o uso de medições não-financeiras, combinada com medidas financeiras mais tradicionais, tornou-se uma recomendação geral na qual advém de alguns anos a esta parte onde foram plenamente desenvolvidos nos anos 90. Por sua vez, Otley (1999) vê um abandono crescente de indicadores puramente financeiros, existindo uma mudança (década de 90) no modelo da gestão de desempenho para outras formas de controlo. O autor mencionou como um aspecto negativo a utilização exclusiva do controlo orçamental, onde o risco é maior se a empresa se basear apenas num indicador de desempenho. Azar (2006) refere que a função controlo de gestão tem uma vertente para além da financeira, é também um instrumento contabilístico (e não só) e dá como exemplo o *Tableaux de Bord* que segundo o autor é uma ferramenta de acompanhamento.

No ano 1932 surge em França o *Tableaux de Bord (TDB)* como reacção à inadequação dos dados contabilísticos para a tomada de decisões, situação que se tornou mais evidente com a crise de 1929. Contudo, no início, o *TDB* consistiu mais num conjunto de medidas físicas de desempenho, do foro da linguagem das engenharias, não propriamente da contabilidade (Russo, 2003). Com maiores avanços a partir dos anos 50, em contrapartida da difusão dos métodos de gestão americanos (Bourguignom *et al,* 2001), dos centros de responsabilidade nas empresas, o *Tableaux de Bord* evoluiu, passando a incluir dados orçamentais e a desagregar-se por centros de custo/responsabilidade. Até aos anos 80, o *TDB* comparava e controlava os dados realizados com os previamente fixados (em orçamentos), com vista à tomada de medidas correctivas. Como refere Bugalho (2004) citado por Russo (2003), apesar da utilidade evidente do *Tableaux de Bord* para as empresas esta ferramenta revelava várias insuficiências, tais como: dados eminentemente financeiros; predominância de critérios financeiros que não garantia o alinhamento dos interesses da direcção e dos gestores dos diversos centros de responsabilidade;

não estabelecia relações de causa e efeito, apenas apurava o resultado final; não mostrava qual o impacto das decisões dos gestores sobre as variáveis não financeiras, como a qualidade, a satisfação dos clientes e a inovação; não fazia a ligação entre a estratégia e as decisões e acções ao nível operacional; analisava dados históricos, do passado, mas não considerava de forma sistemática; variáveis com impacto nos resultados futuros.

Só nos anos 90 surgem preocupações, de modo genérico, de coerência entre os objectivos estratégicos e as acções. No entanto, o *Tableaux de Bord* aborda esta ligação, deixando aos gestores a tarefa de definirem o que consideram mais relevante, o que acabou por manter o principal objectivo que consistia na perspectiva financeira para a análise do desempenho (Russo, 2003).

À medida que se foi focalizando em factores como a boa gestão das organizações (factores intangíveis e não financeiros), o *TDB* foi sendo aperfeiçoado, passando a considerar novos aspectos, conforme indica Jordan *et al* (2005). Segundo este autor, o *Tableaux de Bord* é desenvolvido a partir da missão e da definição dos objectivos de cada centro de responsabilidade que, por sua vez, dão origem à identificação dos factores críticos de sucesso, que serão quantificados pelos Indicadores Chave de Desempenho. Jordan *et al* (2005) identificam-no ainda como instrumento de informação rápida, para apoio ao processo de tomada de decisão, considerando que o *Tableaux de Bord* deve ter as seguintes características:

- ✓ Ser personalizado por gestor ou centro de responsabilidade, em função das necessidades e objectivos específicos de cada um.
- ✓ Conter indicadores financeiros e não financeiros que permitam avaliar melhor, em conjunto, o impacto das decisões operacionais.
- ✓ Ser sintético, conter o que é relevante para a tomada de decisões, mas com possibilidade de desagregação da informação.
- ✓ Conter informação diversificada, por exemplo, dados reais e previsionais (por estimativas), dados físicos, financeiros, indicadores de objectivos e de meios, entre outros.
- ✓ Ser convergente, i.e., os objectivos dos centros de responsabilidade mais elementares devem traduzir os meios de acção dos centros de responsabilidade hierarquicamente superiores.
- ✓ Estar disponível rapidamente, em tempo útil, de forma frequente e regular, pelo menos com periodicidade mensal.

Em suma, o *Tableaux de Bord* é, essencialmente, um instrumento de informação rápida, de comparação da realidade com as previsões, permitindo o apuramento de desvios, a elaboração de diagnósticos e a tomada de decisões a curto prazo (Russo, 2003). Jordan *et al* (2007:229) definem o *Tableaux de Bord* como *um conjunto de dados sobre os resultados de um centro de responsabilidade para a informação do gestor* a actuar para o curto prazo, que se torna muito rápido de consultar, contém dados históricos e previsionais (estimativas), contém indicadores diversificados e é muito sintético mas com possibilidade de desagregação. Como tal, esta ferramenta pode-se considerar um instrumento de comparação, favorece o diálogo, ajuda e estimula a tomada de decisão.

A pouca adesão e divulgação do *Tableaux de Bord* entre os gestores e académicos americanos por motivos culturais (Bourguignom *et al,* 2001), acabou por permitir uma boa oportunidade para o desenvolvimento do *Balanced Scorecard* (Otley, 1999), como alternativa para a resolução das insuficiências verificadas na avaliação do desempenho empresarial com base noutras metodologias.

De facto, o conceito e sistema de gestão *Balanced Scorecard* apresentado por Robert Kaplan e David Norton (1992) nos anos 90 nos Estados Unidos na *Harvard Business School*, surgiu na sequência de estudos realizados em várias multinacionais americanas, insatisfeitas com o facto de verificarem que os seus métodos de avaliação do desempenho empresarial, baseados em indicadores contabilísticos e financeiros, eram inadequados no actual contexto competitivo mundial. Os «cortes radicais» nas despesas das empresas, com o intuito de se conseguirem lucros imediatos, originavam a deterioração da qualidade dos serviços prestados, resultavam na maximização das remunerações dos gestores, a saída de quadros e trabalhadores indispensáveis,

aumentando a insatisfação dos clientes e a desmotivação dos empregados, ou seja, sacrificando o potencial de desenvolvimento (valor) a médio longo prazo das suas organizações. Nos anos 80 torna-se indispensável a gestão e o controlo de factores não financeiros e intangíveis, como forma de se alcançarem vantagens competitivas duradouras, o que não estava a ser assegurado pelas ferramentas de controlo de gestão anteriores ao *BSC* (Russo, 2003). A novidade do *Balanced Scorecard (BSC)* resulta essencialmente das seguintes características (Russo, 2003): o *BSC* focaliza a gestão na estratégia da empresa/organização, traduzindo-a em objectivos, iniciativas e indicadores, segundo quatro perspectivas (uma financeira e as restantes não financeiras); são então incluídas neste modelo três perspectivas adicionais – Clientes, Processos Internos, Aprendizagem e Desenvolvimento Organizacional; o *BSC* implica comunicação, envolvimento e alinhamento das pessoas, assim como o ajustamento dos recursos da organização, de acordo com a estratégia definida; esta metodologia considera a existência de relações de causa e efeito entre as diferentes perspectivas referidas e entre as medidas e as acções de desempenho, de curto e longo prazo; o *BSC* pretende assegurar o *feedback* e a aprendizagem resultante das acções e dos resultados alcançados (a curto e a longo prazo), com o objectivo de facilitar a revisão da própria estratégia.

O *Balanced Scorecard* ao estimular novas e diferentes estratégias, proporcionou às empresas/organizações a oportunidade de se afastarem da perspectiva tradicional de curto prazo, sustentada na redução de custos e na concorrência a preços baixos, reorientando-as para a criação de oportunidades de crescimento, para a oferta de produtos e serviços de elevado valor para o cliente. Permitiu também ajudar a resolver outra fragilidade detectada – a falta de ligação entre a estratégia de longo prazo e as acções de curto prazo, aspecto que Kaplan & Norton (1996) consideram estar na origem do insucesso na implementação de muitas estratégias empresariais. Segundo os criadores do *BSC*, o conjunto de indicadores de desempenho tem de ser escolhido de modo coerente com a visão, missão e estratégia organizacional, num processo de desdobramento hierárquico (top-down). A construção e definição dos indicadores deve ser realizada na fase do planeamento, quando se determina a missão e a estratégia da empresa, os factores críticos de sucesso da mesma, as metas de desempenho a alcançar, os indutores desses resultados e os indicadores de causa-efeito, para posterior controlo. Define-se assim um cenário pretendido, com os meios para o atingir, consciencializando as pessoas para o seu alcance, através de medidas de compensação e incentivo (Ramos e Gonçalves, 2002). Como exemplo de empresas que foram bem sucedidas na implementação no *BSC*, são de referir a Mobil, a Oracle, Cigna, SAP, Bain & Co., Rockwater, Apple Computer, entre outras, que em poucos anos superaram situações de prejuízos significativos para, actualmente, serem empresas com elevada rentabilidade (Rolph, 1999, Sim & Koh, 2001).

O actual ambiente competitivo exige das empresas a definição de uma orientação estratégica e o domínio de factores não financeiros e intangíveis, necessários à sustentação futura das suas vantagens competitivas. Neste contexto, o *Balanced Scorecard* apresenta-se como uma solução apelativa, nomeadamente para o tecido empresarial português (constituído maioritariamente por PME´s), confrontado hoje com graves problemas de competitividade internacional (Russo, 2003). Contudo, deve-se ter alguma precaução e ponderar alguns aspectos mais críticos: o *BSC* foi desenvolvido num contexto diferente daquele que é o da maioria das empresas portuguesas, que são de reduzidíssima dimensão. Nessa perspectiva, o grau de formalismo e complexidade inerente à implementação desta ferramenta pode não ser compatível com a gestão fortemente informal e simplificada das empresas mais pequenas, numa perspectiva de custo-benefício (Russo, 2003). No entanto, este problema também se coloca na implementação de um modelo formal de *Tableaux de Bord*.

Outras ferramentas foram divulgadas mas sem o mesmo destaque/uso que as anteriores, tais como a definição de objectivos individuais e/ou colectivos, a produção de resultados intercalares, como referem Nobre (2001) e Meyssonnier e Zawadzki (2007) e a utilização de previsões, como indicam Bajan-Banaszak (1993) e Meyssonnier e Zawadzki (2007).

7. FACTORES DE CONTINGÊNCIA QUE INFLUENCIAM O SISTEMA DE CONTROLO

A abordagem contingêncial enfatiza a importância das características da organização e do seu ambiente. Nos últimos anos, muitas organizações têm experimentado numerosas mudanças nas suas estruturas como resultado da mudança empresarial. Em resposta, os sistemas de controlo de gestão nas organizações também estão a mudar e o seu interesse é intensificado pelos teóricos. Estas mudanças abriram novas perspectivas para os investigadores e criaram uma oportunidade para estudar a natureza da evolução da gestão, e para tal a investigação empírica desempenha um papel importante no estudo das mudanças no desenvolvimento da gestão (Herath, 2007). O estudo desta temática pelos diversos autores tem como finalidade absoluta o desempenho organizacional, que para Simon (2007) deriva do ajuste de factores contingênciais identificados e por conseguinte dos sistemas de controlo de gestão.

Os principais factores de contingência que influenciam a adopção de um sistema de controlo de gestão resultam do Tamanho da Empresa (Lavigne, 2002), do Perfil do Dirigente (Nobre, 2001), da Cultura Organizacional (Abi Azar, 2007), da Orientação Estratégica (Porter, 2004), da Dinâmica do Mercado (Condor e Rebut, 2007), e dos Sistemas de Informação (Condor e Rebut, 2007). Outros factores poderiam ser analisados mas devido ao seu menor impacto estas não se tornam tão relevantes, tais como o tipo de produção (Nobre, 2001). De realçar que estes factores estão constantemente ligados entre si como poderemos observar devido à própria caracterização de uma PME.

O **tamanho da empresa** é um dos factores contingênciais que influencia a escolha de um sistema de controlo de gestão (Lavigne, 2002), nomeadamente a insuficiência de recursos humanos que pode mesmo ser uma barreira para a adopção e uso (Garengo *et al*, 2007). Os autores referem que os sistemas são vistos como causadores de burocratização e como um obstáculo à flexibilização das PME's, que pelas características que apresentam (pequena dimensão), em termos de actividade e de pessoal, torna limitados o volume de recursos técnicos e financeiros à disposição do funcionário numa PME (Abi Azar, 2007). *Uma característica fundamental das pequenas empresas é o papel desempenhado pelo seu líder* (Fallery, 1983 citado por Abi Azar 2007). Condor e Rebut (2007) constataram que o **líder**, os seus objectivos e o facto das PME's serem maioritariamente familiares, são consideradas como relíquias de família, controladas por um membro da família para determinar a forma de controlo de gestão a ser implementado. Portanto, o objectivo principal do líder será manter a empresa, para transmiti-la a um membro da família, por vezes sem experiência e garantir a sua independência. Neste caso, a gestão só irá satisfazer as necessidades do líder para tomar uma decisão estratégica para a sobrevivência da sua empresa no longo prazo. Os autores referem que esta situação advém igualmente pela formação do dirigente e na vontade de impor uma ferramenta mais formalizada para que o seu negócio se torne mais sustentável. Abi Azar (2007) e Nobre (2001) afirmam que existe uma forte influência do dirigente no sistema de controlo de gestão duma PME para adequar as suas prioridades como a autonomia, o crescimento ou a sustentabilidade, pelo que se torna importante estudar a sua formação devido a esta personalização da ferramenta de controlo de gestão. Um sistema de avaliação de desempenho por excelência é influenciado pela cultura organizacional (Otley, 1999). Abi Azar (2007) caracteriza uma PME pela especialização baixa de tarefas dentro da empresa, concluindo que essa característica é directamente proporcional ao tamanho da empresa, sendo que nas empresas mais pequenas têm uma estrutura com menos processos formalizados e as decisões centram-se no dirigente pelo que influenciará a sua estrutura organizacional. Esta ideia vem de encontro ao que Davila (2005) sustenta, ou seja, que os recursos humanos são vistos como parte integrante do sistema de controlo de gestão, são informais. O estilo de gestão (perfil do dirigente) é um factor a ter conta, considerado mesmo indissociável nas PME's, visto ser considerado como um dos aspectos-chave para a compreensão da cultura organizacional (Garengo e Bititci, 2007) e por conseguinte da **aprendizagem** na empresa, onde a mudança na cultura é frequentemente descrita como parte de um complexo e longo processo que deverá ser dissecado. *Cada empresa que compete em uma indústria possui uma estratégia competitiva, seja ela explícita ou implícita* (Porter, 2004), neste sentido independentemente da **estratégia** (liderança no custo total, diferenciação e enfoque) que

uma empresa detém esta é um factor que condiciona a ferramenta de controlo de gestão. O modelo das Cinco Forças de Porter foi concebido pelo mesmo em 1979 e destina-se à análise competitiva entre empresas para que as suas estratégias sejam eficientes. Rivalidade entre os concorrentes, Poder Negocial dos clientes, Poder Negocial dos fornecedores, Ameaça de Entrada de Novos Concorrentes e a Ameaça de produtos substitutos são os factores-chave (segundo Porter) para a obtenção da eficiência estratégica. Herath (2006) diz que a estratégia é a linha condutora para qualquer esforço organizacional e, assim, um factor primordial para a concepção e manutenção de um sistema de controlo de gestão eficaz. Brochand (2000) apresenta dois modelos para a avaliação da **dinâmica do mercado**, a matriz *BCG* (*Boston Consulting Group*) e o modelo *McKinsey* (ou *General Electric*). A *BCG* avalia a actividade de uma empresa com base na taxa de crescimento do mercado e na quota relativa de mercado tendo em conta a rendibilidade e as necessidades financeiras. O modelo *McKinsey* apoia-se, também, em dois critérios de avaliação, a atractividade do mercado e na posição competitiva da empresa. Tendo em conta os modelos apresentados, Condor e Rebut (2007) afirmam que o crescimento de mercado é um factor de contingência mesmo que estes se apresentem estáveis ou fortes, com a nuance de que quanto mais fortes se apresentam os mercados maior será sofisticação do sistema de controlo. Este pensamento é partilhado por Fisher (1995) que refere que a adequação de diferentes sistemas de controlo depende da definição do negócio. Condor e Rebut (2007) afirmaram que as características da oferta referem-se à forma como a empresa oferece o produto/serviço ao cliente. Os autores suscitam algumas questões para apurar esta abordagem como a venda de um produto padronizado ou como a venda de uma gama de produtos. Nobre (2001) no estudo que realizou verificou que o processo de produção é direccionado para o cliente e é feito à sua imagem, ou seja, o cliente está na origem do produto e do processo de produção. Por via dos dados apresentados, Condor e Rebut (2007) consideram a hipótese do mercado como um factor de contingência. O sistema de gestão das PME's é caracterizada por um lado pelo seu **sistema de informação** interno bastante complexo e muitas vezes mal organizado, onde na maior parte dos casos é feito pelo contacto e pelo diálogo (informal), e por outro lado um simples sistema de informação externo como o contacto com outros actores do mesmo meio tais como clientes, fornecedores e bancários (Abi Azar, 2007). Herath (2007) diz que o sistema de informação facilita o controlo de gestão, fornecendo informações sobre o funcionamento da organização de cada um e de cada departamento. Condor e Rebut (2007) referem que a origem dos instrumentos de controlo de gestão está nos sistemas de informação (factor de contingência), sendo que quanto mais agressivo é for o mercado mais sofisticado é o sistema de informação. No entanto os sistemas mais utilizados para o controlo segundo os autores são o Excel, base de dados, software de gestão, sistemas integrados de gestão e programas informáticos de contabilidade.

8. CONCLUSÃO

Os inúmeros autores que se debruçam sobre o tema demonstram que a temática tem tanto de pertinência como de proficuidade, apesar de as PME's continuarem a ser o "parente pobre" face à sua preponderância no tecido empresarial. As ferramentas financeiras são aquelas que detêm um maior peso nos sistemas de controlo, particularmente o orçamento (Abi Azar, 2007) (Nobre, 2001). A sua "discussão" relativamente à sua utilidade dá ênfase a esta ideia, embora a sua unanimidade depara-se em declínio com o aparecimento das ferramentas não financeiras como o *Balanced Scorecard* e o *Tableaux de Bord* (Jordan *et al*, 2007), onde não se encontram ainda totalmente harmonizáveis com as PME's e daí o seu menor uso.

Os factores contingênciais são fundamentais para a selecção da ferramenta de gestão (Herath, 2007). Neste aspecto constatou-se duas ideologias, a francesa que menciona a dimensão da empresa como o factor crucial (Condor e Rebut, 2007) (Nobre, 2001) e a anglo-saxónica onde elege o ciclo de vida da empresa (Condor e Rebut, 2007). Devido ao controlo de gestão nas PME's ser assegurado pelo dirigente (Abi Azar, 2007) este ainda é bastante informal. As PME's têm características peculiares como a sua flexibilidade (Lavanda, 2007) o que dificulta a distinção dos factores de contingência e o seu impacto nas escolhas das ferramentas

(Condor e Rebut, 2007), visto que na maioria deles estão relacionados entre si (Garengo e Bititci, 2007).

REFERÊNCIA BIBLIOGRÁFICA

ABI AZAR, J. (2006), «Les outils de contrôle de gestion dans le contexte des PME : cas de PMI au Liban», *Congrès de l'Association Francophone de Comptabilité*, Tunis.

AMAT, J. E GOMES, J. (2004), "Controle de Gestão – Uma perspectiva Global"

ANTONY, R. E DEARDEN, J. (1975), "Management Control Systems – Text and Cases.", Instituto Politécnico da Covilhã.

ANTHONY, R. N. (1965), "Planning and Control Systems: A Framework for Analysis", *Harvard Business School Press*, Boston, MA.

BAJAN-BANASZAK (1993), « L'expert-comptable et le conseil en gestion», *Revue Française de Comptabilité*, n° 249, p. 95-101.

BOURGUIGNON, A., MALLERET, M. E NØRREKLIT, H. (2001), "Balanced scorecard versus French tableau de bord: beyond dispute, a cultural and ideological perspective."

BURNS, J. E SCAPENS, J. (2000), "Conceptualising management accounting change: an institutional framework, forthcoming."

BURNS, J. E VAVIO, J. (2001), "Management accounting change", *Management Accounting Research*, 12**,** 389 – 402.

CLELAND, D. E KING, W. (1972), "Management: A Systems Approach", *McGraw-hill Series in Management.*

COMBES, H. (2007), "Budget is still king", *Chartered Accountants Journal*, Jul, Vol. 86 Issue 6, p44-44.

CONDOR, R. E REBUT, K. (2007), "Déterminants et modes opératoires du contrôle de gestion dans les PME: une approche qualitative comparative."

DAVILA, T. (2005), "An exploratory study on the emergence of management control systems: formalizing human resources in small growing firms", *Accounting, Organizations and Society*, 30, pp. 223-248.

BROCHAND, BERNARD (2000), "Mercator 2000: teoria e prática do marketing", Publicações Dom Quixote, 9ª edição, Lisboa.

DUBRIN, A. J. (2000), "Essentials of Management, South-Western College Publishing, New York, NY."

FISHER, J. (1995), "Contingency-based research on management control systems: categorization by level of complexity", *Journal of Accounting Literature*, 14, 24–53.

FISHER, J.G, MAINES, L.A., PEFFER, S.A. E SPRINKLE, G.B. (2005), "An experimental investigation of employer discretion in employee performance evaluation and compensation", *The Accounting Review*, 80(2), 563-583.

GARENGO, P. E BITITCI, U. (2007) "Towards a contingency approach to performance measurement: an empirical study in Scottish SMEs", *International Journal of Operations & Production Management*. Bradford. Vol. 27, nº 8, pg. 802.

HANSEN, S., OTLEY, D. E VAN DER STEDE, W. (2003), "Practice Developments in Budgeting: An Overview and Research Perspective", *Journal of Management Accounting Research,* 15, pg. 95.

HERATH, S. K. (2007), "A framework for management control research.", *Journal of Management Development*, Vol. 26, nº9, pg. 895-915.

HOPE, J. E FRASER, R. (2003), "Who needs budgets?", *Harvard Business Review* 81(2), 108-15,126.

JENSEN, M. (2001), "Corporate Budgeting Is Broken - Let's Fix It", *Harvard Business Review, Vol*. 79. Issue 10, p94-101.

JORDAN, H., NEVES, J. E RODRIGUES, J. (2005), "O Controlo de Gestão – Ao Serviço da Estratégia e dos Gestores", Áreas Editora, SA, 6ª edição, Lisboa.

JORDAN, H., NEVES, J. E RODRIGUES, J. (2007), "O Controlo de Gestão – Ao Serviço da Estratégia e dos Gestores", Áreas Editora, SA, 7ª edição, Lisboa.

JONES, O. (2003), "Competitive Advantage in SMEs: Towards a Conceptual Framework. Competitive Advantage in SMEs – Organising For Innovation and Change", Wiley, Southern Gate – England.

JUOZAPAVIČIŪT, A. E STONČIUVIEN, N. (2007), "Flexible budgets and the problems of their preparation"

KAPLAN, R.S. E NORTON, D.P. (1992), "The balanced scorecard - measures that drive performance", *Harvard Business Review, 70*, *1,* 61-66.

KAPLAN, R. S. E NORTON, D. P. (1996), "Using the balanced scorecard as a strategic management system", *Harvard Business Review, 74, 1*, 75-85.

LAVARDA, C. (2008), 2El Alcance de los Sistemas Contables de Gestión en las PYME: Su Impacto en la Eficiencia Empresarial. Estudio Empírico en el Sector de la Madera y del Mueble de la Comunidad Valenciana", Universitat de València – Facultat d'Economia.

LAVIGNE, B. (2002), « Contribution à l'étude de la genèse des systèmes d'information comptable des PME : une approche empirique », *Actes du XXIIIème Congrès de l'Association Française de Comptabilité*, Toulouse.

LIBBY, T. (2007), 2Beyond Budgeting or Budgeting Reconsidered? A Comprehensive Survey of North-American Managers' Views"

MERCHANT, K.A. (1985), "Control in Business Organizations", Pitman, Boston, MA.

Meyssonnier, F. e Zawadzki, C. (2007), « L'introduction du contrôle de gestion en PME: étude d'un cas de structuration tardive de la gestion d'une entreprise familiale en forte croissance », *Congrès de l'Association Francophone de Comptabilité*, Poitiers

NEELY, A., BOURNE, M. E ADAMS, M. (2003), "Better budgeting or beyond budgeting?" *Measuring Business Excellence*, Vol. 7, nº3; pg. 22.

NOBRE, T. (2001), « Méthodes et outils du contrôle de gestion dans les PME », *Finance Contrôle Stratégie.* Volume 4, nº2, juin, p.119-148.

OTLEY, D., (1999), **"**Performance management: A framework for management control systems research", *Management Accounting Research*, Kidlington, Dec., Vol. 10, nº4, pág. 363.Oatley, D., (2003), "Management control and performance management: whence and whither?" *The British Accounting Review*, 35 (4), 309-326.

OTLEY, T. E BERRY, J. (1980), "Control, Organisation and Accounting*" Accounting, Organizations and Society,* Oxford, Vol. 5, nº2, pg. 231.

PORTER, MICHAEL (2004), "Estratégia Competitiva – Técnicas para análise de indústrias e da concorrência", Editora Campus, 2ª edição, São Paulo.

RAMOS, C. E GONÇALVES, D. (2002), "Balanced Scorecard: Uma ferramenta de gestão estratégica" *Jornal de Contabilidade, nº* 298, 4-8.

RANKIN, F., SCHWARTZ, S. E YOUNG, R. (2003), "Management Control Using Nonbinding Budgetary Announcements", *Journal of Management Accounting Research,* 15, pg. 75.

ROLPH, P. (1999), "The balanced scorecard: Get smart and get control**",** *Chief Executive,* Jul/Aug, 146, pg. 52.

russo, j. (2003), « Balanced Scorecard *versus* Tableau de Bord", *Artigo – CTOC.*

RUSSO, JOÃO (2006), « Balanced Scorecard para PME », Lidel – edições técnicas, lda, 3ª edição, Lisboa-Porto.

SAMUELSON, L. (1999), The effects of increasing turbulence on organizational control: Some reflections. In The Current State of Business Disciplines, Management Volume 5, edited by S. B. Dahiya, 2017-2090. Rohtak, India: Spellbound.

SCAPENS, R.W. (1984), Management Accounting — a survey paper. In: Scapens, R.W., Otley, D.T., Lister R.J., Management Accounting, Organisation Theory and Capital Budgeting—Three Surveys. Macmillan/ESRC, London, 15-95.

SCAPENS, R.W. (2006), "Understanding management accounting practices: A personal journey", *The British Accounting Review,* 38(1), pp.1-30.

SCARLETT, B. (2007), "Management Accounting – Performance Evaluation", *Financial Management*, London, Sep., pg. 54.

SCHICK, A. (2007), "Performance Budgeting and Accrual Budgeting: Decision Rules or Analytic Tools?" *OECD Journal on Budgeting,* 7, 2, pg. 109.

SCHEYTT, T., SOIN, K. E METZ, T. (2003), "Exploring notions of control across cultures: a narrative approach", *European Accounting Review*, Sep., Vol. 12, Issue 3, p515-547.

SIM, K. E KOH, H. (2001), "Balanced scorecard: A rising trend in strategic performance measurement", *Measuring Business Excellence,* 5, 2, pg. 18.

SIMON, C. (2007), "A Configuration Form of fit in Management Accounting Contingency Theory: An Empirical Investigation", *The Business Review, Cambridge,* Summer, 7, 2, pg. 220.

SMITH, K. (1997), "Management Control Systems and Strategy: A Critical Review", *Accounting, Organizations and Society*, Vol. 22, nº2, pp. 207-232.

SULAIMAN, S. E MITCHELL, F. (2005), "Utilising a typology of management accounting change: An empirical analysis", *Financial Management*, London, pág. 36.

THORNBORY, G. E FARLEY, L. (2007), "Budget basics", *Occupational Health,* Aug., 8, pg. 20.

WALLANDER, J. (1999), "Budgeting: An unnecessary evil", Scandinavian. *Journal of Management* ,15, 405421.

WILLIAMS, J., DUBRIN, A. E SISK, H. (1982), "Management & Organization", South-Western Publishing Co, Ohio.

EVOLUTION PATTERNS OF FOREIGN FIRMS IN PORTUGAL: A COMPARATIVE ANALYSIS WITH DOMESTIC FIRMS

CELESTE AMORIM VARUM
E-mail: camorim@ua.pt
VERA CATARINA ROCHA
E-mail: verarocha@ua.pt
Department of Economics, Management and Industrial Engineering
University of Aveiro

Abstract

In this paper we assess the potential contribute of MNEs to the Portuguese economy, by conducting a dynamic and comparative analysis showing evolutional differences between foreign and domestic firms, a topic particularly important when FDI has become more important to economic growth. The study is based on microecomic panel data from Quadros de Pessoal database, for the period 1985-2005. FDI and foreign firms in Portugal reveal a positive evolution at technological complexity, performance and human capital levels. These trends may have beneficial effects on industrial diversification and regeneration of Portuguese productive structure. Given the exporting nature of foreign firms, FDI also contributes for Portuguese growth and structural change on its exports. Our results support the need for a FDI policy focused on the quality of investments and on the reduction of gaps between foreign and domestic firms, in order to Portuguese firms benefit from potential positive FDI spillovers.

Key Words: Foreign Direct Investment (FDI), Multinational Enterprises (MNEs), Performance, Human Capital.
Thematic Area: Entrepreneurial Economics.

1. INTRODUCTION

Since the EEC membership in 1986, the Portuguese economy has been exposed to higher international openness and registered an outstanding growth of inward Foreign Direct Investment (FDI). From the early literature on FDI, it has been suggested that multinational firms have a positive contribution to host country development by supplying capital, technology and management resources that would not otherwise be available. However, the potential impact of MNEs for host economies depends on a variety of factors and, in particular, on the differences between domestic and foreign firms (Bellak, 2004a). Hence, in this paper we assess the potential contribute of MNEs to the Portuguese economy using a systematic analysis of the specificities and quality of the FDI.

Multinational companies, owing to their ownership advantages and firm-specific assets as better access to financial markets and superior managerial practices, are systematically expected to have better performance than purely domestic firms. There are two main theoretical approaches that are important to explain performance differences between firms – Industrial Organization (IO), a neoclassical perspective based on industry characteristics and firm specificities, and International Business (IB), a more evolutionary theory based on arguments as ownership-advantages, location factors and incentives for internalization of some activities by foreign firms. Both viewpoints are crucial for understanding firm performance and why foreign firms may differ from domestic ones.

The international business (IB) literature has well established that a reason why firms invest abroad is that they possess firm-specific advantages, not available to domestic firms in the host country (Vernon, 1966; Hymer, 1976; Dunning, 1981; Caves, 1974). These advantages assist MNEs to achieve a differentiated behaviour and superior performance from domestic owned establishments. Moreover, some forms of FDI are clearly better than others. A good investment should create linkages with the local economy and contribute to a qualitative improvement of the host environment, for example, by hiring skilled labour and investing in high-value added knowledge based activities. Ben Hamida (2007) demonstrated that MNEs are an important agent in host countries, being a provider of knowledge, capital, capabilities and markets, as well as a creator of jobs and a competition stimulator, inducing intra-industry spillover effects – demonstration, competition and worker mobility effects. So, subsidiaries evolve with the country and must contribute to the development of the host economy.

The international literature points that the main areas where foreign firms widely differ from domestic ones are wages, work relations, workforce, human capital, productivity, growth, profitability, survival patterns and innovation activities (Blomström and Kokko, 1998; Ramstetter, 1999; Hanson, 2001; Sleuwaegen and DeBacker, 2003a; Bellak, 2004b). However, Bellak (2004b) also concludes that the findings of the empirical studies are not so supportive of the superior performance hypothesis. In most cases, structural variables of the firm (in particular, size) or industry characteristics soften the effects of foreign ownership. Though, the literature tends to suggest on balance that foreign ownership impacts positively on firms' performance.

This paper focus on Portuguese case, a small economy with, not only internal structural challenges, but also strong external ones, imposed by the emergence of new competitors and, more recently, by the international crisis. Portuguese attractiveness for foreign investors has been threatened by new EU members, in particular by CEECs (Central and Eastern European Countries), a fact that is now largely discussed in the newly literature (Buckley and Castro, 1998; Galego et al., 2004; Ferragina and Pastore, 2007). The effect of the current international crisis is not yet measurable.

The literature on the Portuguese case supports that the future of Portuguese economy depends on the evolution of production structures. Portugal must become more productive, hence the development of medium-high/high technology industries and the strengthening of human capital competencies are imperative. Some recent empirical studies confirm the potential contribute of FDI on international dynamics (Cabral, 1996) and on the regeneration of Portuguese productive structures (Farinha and Mata, 1996; Gonçalves and Guimarães, 1997; Tavares, 2001; Crespo, Proença and Fontoura, 2007; Flores, Fontoura and Santos, 2007; Teixeira and Tavares-Lehmann, 2007; Freitas and Paes Mamede, 2008).

The difficulty in attracting new investments, worsened in a crisis environment, appeals to the need of policies focused on the quality of investments (not only past investments, but also future ones). An analysis based on capital flows is misleading and does not reveal the potential of investments. In this paper, we study the evolution of foreign firms located in Portugal, according to a set of relevant indicators for their competitiveness. We conduct a comparative and dynamic analysis between foreign and domestic firms, in order to assess if foreign firms really make the difference. If they systematically reveal a superior behaviour at some critical levels (as performance, exporting strategies, technological intensity, formation and remuneration of human capital), thus FDI could have a positive impact on the economic development (Bellak, 2004a, 2004b).

Over the 80s and 90s, the attitude towards FDI has changed considerably. Many governments around the world have liberalized their FDI regulations since then and are now actively providing generous investment incentives to attract inward FDI, expecting FDI spillovers. Economies in general devote great resources to attract investments, in particular foreign investments (Oxelheim e Ghauri, 2004; Marchick e Slaughter, 2008). Peripheral economies, in convergence or in transition, look for a significant bonus with the entry of foreign firms. The literature suggests that this benefit may occur through a "presence effect" (Bellak, 2004a), due to an increasing share of more productive firms (Sleuwaegen and De Backer, 2003a), or through a "demonstration effect" (Ben Hamida, 2007), allowing the imitation of

foreign technologies by domestic firms, or indirectly through a "transmission effect" (Bellak, 2004a), through spillovers from foreign firms to domestic ones (Crespo and Fontoura, 2007; Blomström and Kokko, 1998).

Portugal has also implemented, over the last decades, an active policy to attract FDI (Buckley and Castro, 1998; Tavares, 2001; Tavares, 2002; Teixeira and Tavares-Lehmann, 2007). FDI in Portugal grew from about 1% of GDP in 80s and 90s to 3,65% of GDP in the period 2000-2006 (UNCTAD, 2007). Low labour costs were one of the main motivations of MNEs' location decisions (Ribeiro and Santos, 2001; Galego et al., 2004) as low labour costs are viewed as a location advantage by many MNEs (Brenton et al., 1998). This study is not a mere analysis of FDI flows, it assesses the quality of foreign investments and how do they differ from domestic investments.

For the Portuguese case, there is some empirical evidence on a differentiated behavior between foreign and domestic firms. In table 1, we present the broadest comparative studies conducted for Portugal. These studies cover differences at different levels, as exporting behaviour, size and scale, survival trends, specialization patterns, productivity and profitability, wage structures and human capital endowments. In our study, we contribute for a deepening of knowledge on these issues, by analyzing the dimensions of proximity and differentiation between foreign and domestic firms. Conclusions of previous studies are considered and support our own results. We investigate this question and how it has evolved over the period 1985-2005. From our knowledge, this is the first study with such a long time span dimension. Thus, more than allowing for a comparative and static analysis, with this time horizon it is possible to investigate how the differences evolved over time. We pay special attention to issues of scale, efficiency, human capital, technological complexity and dynamics.

Table 1. Summary of comparative studies: DF versus FF in Portugal

Variable	Databases	Period	References
Exporting Behaviour	QP; Central Balance of BP; INE	1986-1992	Cabral (1996)
	OECD, IMF, PWT, International Financial Statistics	1995-2000	Magalhães and Africano (2007)
	Questionnaire	1999	Tavares, 2001; Tavares and Young (2002)
Size	QP	1983-1992	Mata and Portugal (2004)
Survival Rate	QP	1983-1992	Mata and Portugal (2004)
Specialization Patterns	QP	1982-1992	Gonçalves and Guimarães (1997)
	QP,UN-COMTRADE,IMF-World Economic Outlook Database	1995-2005	Freitas and Paes Mamede (2008)
Ability to overcame entry barriers – economies of scale and concentration	QP	1983-1989	Mata and Portugal (1999, 2001)
	QP	1983-1989	Geroski, Mata and Portugal (2003)
Ability to overcame entry barriers – differentiation and industry concentration	QP	1982-1990	Barbosa, Guimarães and Woodward (2004)
Productivity and Profitability	SABI	2006	Cardoso (2008)
	QP; Central Balance of BP	1986-1992	Farinha and Mata (1996)

Table 1. (…)

Variable	Databases	Period	References
Productivity	QP,UN-COMTRADE,IMF-World Economic Outlook Database	1995-2005	Freitas and Paes Mamede (2008)
Profits Expectations	QP	1992	Barbosa and Louri (2005)
	QP	1982-1990	Barbosa, Guimarães and Woodward (2004)
Human Capital: wages	QP	1983-1992	Mata and Portugal (2004)
	QP	1991-2000	Martins (2008)
Human Capital: formation	QP	1983-1992	Mata and Portugal (2004)
	QP	1981-2000	Barbosa (2007)
FDI, Human Capital and R&D	Questionnaire	2004	Teixeira and Tavares-Lehmann (2007)

We look at data at subsidiary level. Although theories pertaining to Multinational Enterprises have traditionally focused upon the parent group, recent works have increasingly recognized the importance of the MNE subsidiary as the unit of analysis (Birkinshaw, 2001). However, little has been written about the assessment of subsidiary performance and dynamics.

The research is based on a panel data at firm level from Quadros de Pessoal database for the period 1985-2005. This database belongs to the Portuguese Ministry of Employment and it is a longitudinal data file matching firms and workers. The data is based on an annual questionnaire that every firm with wage-earners is legally obliged to fill out.

Aggregate data on FDI flows and stocks is important and relevant in many analyses. However, it does neither reflect real decisions and motivations of investors nor qualitative characteristics of subsidiaries and their investments. With data at firm level, we can assess with more accuracy the quality and the type of foreign investments, in order to judge the potential impact of MNEs in the host economies, on a dynamic context (Griffith, 1999; Birkinshaw, 2001; Markusen, 2002).

For the classification of firms, the share of equity owned by foreign investors is a common criterion used in the literature for firms' discrimination. Some studies consider as FDI a foreign participation of, at least, 10%. In our study, we classify a firm as foreign if its share of equity owned is, at least, 50%. This participation level is also commonly used on studies at microeconomic level and ensures a significant level of foreign influence in domestic firms' operations.

Our study complements the existing comparative analyses applied to the Portuguese case with empirical information at firm level (Table 1). These empirical evidences are needed to evaluate the efforts put in action to attract FDI, as well as open new lines of discussion and policy measures.

Why do MNEs pay more to their employees? Why do they retain a more qualified workforce? Should governments give priority to foreign investments? Is foreign ownership in itself an advantage or does it loses significance against structural aspects as size, home country or industry? This paper raises the discussion at this level and is a first step for further econometric analysis.

The paper is organized as follows: in section 2 we present a brief analysis of foreign investments in Portugal during the period 1985-2005, focusing on Portuguese Manufacturing Industry (PMI), where some variables of scale, efficiency, technological complexity and human capital are described. In section 3, we evaluate the asymmetries between and within foreign and domestic firms, based on an analysis by percentiles and the construction of transition probability

matrices. Section 4 briefly describes the demography of both groups of firms, highlighting the differences on their survival trends. Section 5 concludes.

2. FOREIGN SUBSIDIARIES IN PORTUGAL

After the membership to EEC in 1986, FDI inflows increased substantially. Many studies have been demonstrating a positive association between economic integration processes and FDI inflows (Brenton et al., 1998; Buckley and Castro, 1998; Jimeno et al., 2000; Tavares, 2001; Galego et al., 2004). This evidence was supported by Portuguese experience, as Portugal was understood by foreign investors as a platform to easily reach other EEC countries. Between 1985 and 2005, the number of foreign and domestic firms (henceforth 'FF' and 'DF', respectively) located in Portugal grew in a considerable way, although the growth had been greater in foreign case. During this period, the average annual growth rate was about 7,72% in FF and 6,24% in DF. However, 1992 was an adverse year, since both firms had declined in number. In the beginning of 2000s, a little stagnation was, again, evident, but in the recent years the growth of firms has recovered, as table 2 reveals.

Table 2. Evolution of the number of firms in Portugal

	1985		1990		1995		2000		2005	
	Total	Manuf.	Total	Manuf.	Total	Manuf.	Total	Manuf.	Total	Manuf.
No. FF	872	344	1274	434	1964	551	2314	582	3530	830
No. DF	101.329	24.459	139.032	32.337	190.306	37.718	266.387	45.132	337.252	46.059

FF: stands for foreign firms, DF stands for domestic firms

At sector level, the economic sector with greater number of FF during the period studied was Wholesale and Retail, followed by Manufacturing Industry, probably to export activities (Cabral 1996). The data reveals that, despite the growth in number of firms, in relative terms manufacturing reduced its importance as destiny of foreign investments. By 1985 nearly 40% of foreign firms were concentrated in manufacturing, but the value declined to 23,5% in 2005[1].

Table 3. Concentration of investments in PMI

	1985	1990	1995	2000	2005
No. FF	39,4	34,1	28,1	25,2	23,5
No. DF	24,1	23,3	19,8	16,9	13,7

The relative importance of manufacturing declined far more for the group of DF. In reality, foreign investments reveal a higher concentration on manufacturing than domestic firms. In fact, as we show below, FF gained outstanding relevance as players on Portuguese industry over the time period under study.

Regarding the employment and the turnover, the evolution of FF at aggregate level and for manufacturing in particular was positive and greater than that for DF. The average rate of growth of turnover was similar for both groups. Overall, turnover growth rate for domestic firms reached 19% and 17% for foreign ones. Turnover in manufacturing increased slightly less, 14% for DF and 15% for FF. Conversely, the growth of employment was relatively low. Nevertheless, foreign firms registered considerable better results. FF presented a greater growth

[1] The number of FF in PMI increased over the years under study. Although, the average rate of growth was just 5%, lower than the growth rate at aggregate level (7,72%).

rate of employment (4,87%) against the average of DF (2,33%).The average rate of growth of employment in manufacturing registered even lower values - for foreign firms it was about 2,33%, while in domestic firms the growth rate was negative or almost null.

Table 4. Evolution of Turnover and Employment in PMI

	1985		1995		2005	
Turnover (1.000 €)		Total	Manuf	Total	Manuf	Total	Manuf
Foreign Firms	3.672.208	1.812.359	18.756.836	8.571.768	48.931.190	17.419.227
Domestic Firms	32.917.523	12.593.584	128.435.816	33.986.575	243.084.394	47.311.437
Employment (1.000 workers)		Total	Manuf	Total	Manuf	Total	Manuf
Foreign Firms	116	78	170	103	269	109
Domestic Firms	1782	793	2066	744	2801	663

Values in thousands

For FF, the beginnings of 1992 and 2000 were recorded by a slowing in their growth or even by a negative growth. Jimeno et al. (2000), in a study conducted for Portugal and Spain for the period 1960-1996, have also detected a setback in foreign investments received by both countries in the early 90s. A possible explanation for this is the increase of labour costs, due to currency depreciations, and the resulting decrease on productivity levels. Additionally, the early 90s and 2000s were marked by economic instability environments in Portugal (Bank of Portugal, 2003). Galego et al. (2004) and Ferragina and Pastore (2007) also suggest a possible deviation of FDI to the new EU-members (CEECs), mainly to Poland, Hungary and Czech Republic. The reduction of multinational enterprises (MNEs) activities, mainly those ones focused to export, is considered by some authors as a threat for countries like Portugal, where FDI has been crucial, not only to economic growth, but also to industrial diversity and access to a new markets (Ribeiro and Santos, 2001; Castro, 2004; Pereira, 2007). Consequently, the Portuguese position, not yet well-built, may be affected by these setbacks in FF growth.

Despite those fluctuations, next table shows that foreign firms have been gaining increasing expressivity in Portuguese economy. Although their reduced weight in the whole set of firms (just over 1%), the proportion of employment guaranteed by FF was almost 9% in 2005. Additionally, they were responsible by almost 17% of the total turnover reached in Portugal in the same year, representing a clear evolution against 1985, where their influence was just 10%. On average, FF assured 7% of total employment and 13% of total turnover, between 1985 and 2005.

Table 5. Relative Importance of Foreign Firms

	1985		1995		2005	
	Total	Manuf	Total	Manuf	Total	Manuf
Share of FF in number of firms	0,85	0,3	1,02	0,3	1,04	0,2
Share of FF in total turnover	10,04	12,6	12,74	20,1	16,76	26,9
Share of FF in total employment	6,10	8,9	7,61	12,1	8,76	14,1

FF: stands for foreign firms, DF stands for domestic firms.
Values in percentage

Moreover, foreign firms' contribution to total turnover of Portuguese manufacturing more than doubled in two decades, accounting for 27% of turnover in 2005. Multinationals contribution to employment in manufacturing also increased substantially, from 9% in 1985 to 14% in 2005.

2.1. Size and Scale

The following data represents the average dimension of FF and DF, not only at aggregate level, but also in PMI. From tables 6 and 7, we can conclude that FF are larger in dimension, both in turnover and employment. In addition, we can realize that both firms have been reducing their average dimension, since the level of employment has been decreasing. In foreign case, FF were, in the first years, large enterprises, becoming small-medium enterprises (SMEs) in the last years under study. This validates the empirical evidence about the increasing involvement of SMEs in internationalization activities, through foreign investments. For the Portuguese case, Mata and Portugal (2004) had already confirmed, for the period 1983-1992, that foreign firms have, in average, higher dimensions than domestic-owned firms. Following the studies of Mata and Portugal (1999, 2001) and Geroski, Mata and Portugal (2003), focused on entry patterns for the period 1983-1989, our study reveals that foreign firms tend to come into sectors where the foreign presence is already evident and that these firms present stronger abilities to overcome entry barriers in terms of scale. The importance of dimension variables is also highlighted by Barbosa and Louri (2005) for the year 1992, in a comparative study between foreign and Portuguese firms.

Table 6. Evolution of Average Turnover

	1985	1995	2005
Aggregate Level			
Foreign Firms	4.211	9.550	13.862
Domestic Firms	325	675	721
PMI			
Foreign Firms	5.268	15.557	20.987
Domestic Firms	123	177	139

Values in 1.000 €

Table 7. Evolution of Average Employment

	1985	1995	2005
Aggregate Level			
Foreign Firms	133	87	76
Domestic Firms	18	11	8
PMI			
Foreign Firms	225	186	131
Domestic Firms	8	4	2

No. Workers

2.2. Technological Complexity

An interesting point that differentiates FF from DF is their relevance in more advanced industries. According to Markusen (1995), multinational firms have assumed greater importance in industries with higher intensity on R&D activities, higher proportion of qualified professionals and in industries where new and technologically advanced products are produced, as well as with higher levels of advertisement. The literature supports the magnitude of FF in industries with intense R&D activities, so their presence could have positive effects on

industrial diversification of host economies. Moreover, technology embeddedness seems to have a positive impact, not only on subsidiary market performance, but also on subsidiary organizational performance (Andersson, Forsgren and Pederson, 2001).

We analyzed foreign technological complexity based on OECD classification. We present in table 8 the evolution of foreign investments in Low Technology Industries (LTI), Medium-Low Technology Industries (MLTI) and Medium-High/High Technology Industries (MH/HTI)[2].

Table 8. Technological Complexity of Foreign Investments

	1985	1995	2005
Concentration of FF by Industry			
LTI	48,0	48,1	40,2
MLTI	25,6	17,6	21,7
MH/HTI	26,5	34,3	38,1
Concentration of FF in total turnover by Industry			
LTI	30,6	30,3	25,1
MLTI	27,1	9,1	13,1
MH/HTI	42,3	60,6	61,8
Concentration of FF in total employment by Industry			
LTI	38,7	40,2	34,0
MLTI	19,3	9,4	16,5
MH/HTI	41,9	50,5	49,5

In the first years under analysis, FF were more concentrated in Low Technology Industries, as textiles and food products. However, over the years and especially after 1990, we observed a positive evolution in foreign concentration in Medium-High/High Technology Industries, as chemicals, machinery and equipment, electrical machinery and motor vehicles. In 2005, FF in MH/HTI represented about 40% of total FF located in Portugal, while, in 1985, their share was less than 30%.

The change in the level of technological complexity of FF located in Portugal was more evident in what concerns to turnover and employment. In recent years we found that more than 60% of foreign turnover comes from MH/HTI and these are the industries where an half of the employment of FF is collected.

Regarding the share assumed by FF in each stage of technological complexity, we observe that, in spite of representing only 7% of total firms in MH/HTI in 2005, FF contributed with more than 50% at turnover level and with almost 40% of total employment. The reasons for that concentration may be associated to the results of Mata and Portugal (1999, 2001) and Geroski, Mata and Portugal (2003), whose studies revealed that foreign firms prefer sectors with foreign agglomeration and where entry barriers are important. Barbosa, Guimarães and Woodward (2004) have also used Quadros de Pessoal database for the period 1982-1990, showing that foreign firms are able to avoid entry barriers as industrial concentration and differentiation.

[2] Due to the high level of aggregation in economic activities present in 'Quadros de Pessoal' Database, Medium-High and High Technology Industries were joined in our analysis. Additionally, although this classification is a simple and consistent instrument for international comparisons, it presents some limitations, mainly the insufficient sectoral disaggregation of data. Moreover, this classification weakly reflect the real technological development of firms located in certain sectors (Hatzichronoglou, 1997).

Table 9. Relative Importance of FF in each stage of technological complexity

	1985	1995	2005
Share of FF in number of firms			
LTI	1,0	1,0	1,1
MLTI	2,5	1,0	1,5
MH/HTI	1,7	5,2	6,7
Share of FF in total turnover			
LTI	7,2	11,6	14,1
MLTI	11,2	7,7	15,9
MH/HTI	32,5	51,5	55,6
Share of FF in total employment			
LTI	5,7	7,7	8,0
MLTI	9,5	5,9	10,6
MH/HTI	17,5	36,0	38,8

The Portuguese example corroborated the argument of Markusen (1995): MNEs tend to be more important in industries with higher R&D intensity, with higher proportion of qualified professionals, with new products and/or more technologically advanced products and with higher levels of product differentiation and advertisement (see also Kuemmerle, 1999; Andersson, Forsgren and Pederson, 2001).

The concentration of foreign firms in more technologically complex sectors and with higher entry barriers may result in positive effects, as an industrial diversification and a regeneration of productive structure of Portugal (Freitas and Paes Mamede, 2008). Previously, Gonçalves and Guimarães (1997) had revealed that foreign firms presented different specialization patterns from their domestic counterparts. Our study on Portuguese case shows that foreign firms have a determinant role in MH/HTI. Thus, given their exporting behaviour, foreign investments received by Portuguese economy contributed for economic growth and for the structural change of Portuguese exports during the period 1985-2005.

2.3. Operational Performance[3]

Some authors have already found a positive effect of foreign presence in domestic productivity. This can occur directly, if FF are themselves more productive, or indirectly, with FF encouraging competitiveness or transferring technology and knowledge to domestic firms (Crespo and Fontoura, 2007). In this study, we evaluate the level of productivity of firms with the ratio turnover per employee. Table 10 presents the results, showing that foreign productivity was always greater than domestic one. Similar evidences were found for other developed countries like UK (Haskel, Pereira and Slaughter, 2004) and Canada (Globerman, Ries and Vertinsky, 1994; Boardman, Shapiro and Vining, 1998). The supremacy from FF was also evidenced in some emerging economies from Eastern Europe (Sabirianova, Svejnar and Terrel, 2005; Yasar and Paul, 2007), in Japan (Kimura and Kiyota, 2007) and other Asian countries (Ramstetter, 1999). As for the Portuguese case, the differences in productivity levels are obvious, not only in PMI, but also at aggregate levels, with a statistical significance at 1% level[4]. During all the period, subsidiaries' performance largely overcame domestic one. Performance differences are even more evident when we disaggregate the industries by technological complexity levels, as table 11 confirms.

[3] This is a measure of employees' productivity and may be applied in comparative analyses between firms within the same industry. However, low values may represent a labour-intensive industry and high values may be associated to capital-intensive industries.

[4] However, average rates of growth had been greater in domestic firms, 16% against 11,4% for all sectors, and 14,35% against 11,5% in manufacturing.

Table 10. Evolution of Productivity

	1985	1995	2005
Aggregate Level			
Foreign Firms	32	110	182
Domestic Firms	18	62	87
PMI			
Foreign Firms	23	83	160
Domestic Firms	16	46	71

€1.000/Worker

Table 11. Evolution of Productivity in each stage of technological complexity

	1985	1995	2005
Low-Technology Industries			
Foreign Firms	19	63	118
Domestic Firms	14	40	63
Medium-Low Technology Industries			
Foreign Firms	33	81	126
Domestic Firms	27	61	79
Medium-High/High Technology Industries			
Foreign Firms	24	100	199
Domestic Firms	10	53	101

€1.000/Worker

Some studies applied on other economies show that productivity differences are not explained by foreign ownership itself. In most cases, when dimension, capital intensity and human capital are controlled, productivity differences are mitigated or even vanished (Globerman, Ries and Vertinsky, 1994). On the other hand, some authors suggest that foreign ownership is an important source of firm heterogeneity affecting productivity dynamics (Sleuwaegen and De Backer, 2003a, 2003b). So, productivity differences must be carefully evaluated, taking into account that FF are typically more specialized, since they have access to a more productive technologies and to a more qualified workers. Besides this, FF can eventually use public infra-structures in a more efficient way, or even can enter the domestic markets by acquiring more productive firms.

Studies on the Portuguese case, as the study of Barbosa, Guimarães and Woodward (2004) for the period 1982-1990, show that foreign firms not only have better abilities to get through entry barriers, but also have different profit expectations[5]. Cardoso (2008), in a study based on SABI database for the year 2006, notices that there are considerable and significant differences at performance levels, with multinationals presenting higher profitability and productivity, even after controlling for other relevant variables. Recently, Freitas and Paes Mamede (2008) evidence an increasing specialization of Portuguese economy in more productive products and sectors during 1995-2005, which reflect a higher concentration of FDI, thus contributing to economic growth and structural change in Portugal.

[5] However, in particular for 1992, Barbosa and Louri (2005), after controlling for industry and firm specificities, reveal that foreign ownership loses significance.

These performance divergences may be explained in different ways. Foreign subsidiaries have many ownership advantages, including better access to capital and modern management practices, which help them to perform better than domestic firms. The differences may also be associated to different activities performed by both groups of firms (Aitken and Harrison, 1994). Frequently, foreign firms operate at larger scale and carry out more capital-intensive activities (Blomström and Wolff, 1994; Chuang and Lin, 1999; Harris and Robinson, 2002). On the other hand, foreign firms are often more specialized and detain more qualified workers (Andersson, Forsgren e Pederson, 2001; Griffith e Simpson, 2001).

Studies applied in other countries expose that higher performance levels do not result from foreign ownership, but from many other factors. In most cases, when dimension, capital intensity and human capital are controlled, performance differences are soften or even vanished (Globerman, Ries and Vertinsky, 1994; Pessoa, 2007). As a result, the reasons for productivity differences are not exhausted and, in Portuguese case, they are not clear, so that a future research could explain this fact.

2.4. Human Capital

Today, human capital is probably one of the most important competitive factor at firm level, with reflects in the evolution and growth of firms. Some evoke a clear superiority of foreign firms regarding human capital indicators (Teixeira, 2002; Almeida, 2003; Bellak, 2004a, 2004b; Barbosa, 2007; Teixeira and Tavares Lehmann, 2007). The difference may well reflect in the competitiveness for firms. In this sense, for Teixeira (2002), education and skills may have particular effects at *top levels* of the firms, as human capital is viewed as a determinant of productivity and education constitutes a source of information which tends to be highly relevant to "decode" new technical information and to incorporate it into manufacturing process. Many foreign subsidiaries, having pools of skilled labor, have higher survival chances and lower likelihood to exit the market (Mata and Portugal, 2001; Andersson and Vejsiu, 2001; Mata and Portugal, 2004; Mamede, Mota and Mira Godinho, 2007). Bates (1990) also argues that college education improves access to debt capital, so the probability of business discontinuance falls sharply with higher endowment of human capital. Similarly, Hamermesh (1998) proved that additional years of schooling by workers, ceteris paribus, reduce the probability of plant closure. Next we explore our data in an attempt to disclosure potential differences between foreign and domestic firms regarding human capital.

2.4.1. Wages

The dominant perception in the literature is that export oriented FDI received by Portugal is essentially attracted by Portugal's cheaper labour, a fact even more visible in PMI (Ribeiro and Santos, 2001; Tavares, 2001; Barbosa, Guimarães and Woodward, 2004). However, this evidence does not prevent FF from paying higher wages. Moreover, over all the period under study, average earnings paid by FF largely overcame the corresponding earnings in DF (€675 and €282, respectively), and a study conducted by Martins (2008), for the period 1991-2000 and with Quadros de Pessoal database, revealed that FF have greater "generosity" in the remuneration practices, comparing to Portuguese firms, allowing them to attract some of the "best" workers of DF, a fact also corroborated by Crespo, Proença and Fontoura (2007). The study of Mata and Portugal (2004) reveal that domestic entrants are smaller and pay lower wages. So, FDI seems to have a significant effect upon wage inequality (Taylor and Driffield, 2004). The asymmetry in wages remained all over the time (1985-2005), as table 12 shows. The differences in average wage growth rates were irrelevant (12% for FF and 10,6% for DF).

Table 12. Evolution of Average Wages

	1985	1995	2005
Foreign Firms	180	697	1.322
Domestic Firms	82	295	476

Values in €

As for productivity, the differences in wages are not absolutely explained in the literature. Other studies justify those differences with performance gaps (better performed firms may pay higher wages) or with management policies adopted by firms. Other ones explain those divergences evoking the differences in qualifications of the workforce (more productive and qualified workers usually earn higher wages) or even the high technology used by firms (Globerman, Ries and Vertinsky, 1994; Bellak, 2004a, 2004b). For Jimeno et al. (2000), the explanation is the firm size – their study release that larger firms tend to pay higher wages. Higher wages can also be a strategy to motivate the workers, to make them more productive. Furthermore, some authors show that inward FDI result in higher wages to more qualified workers (Ernst, 2005). In Portuguese case, this persistent asymmetry must be carefully observed in future research.

2.4.2. Qualifications of the workforce

In recent decades, the studies applied to the effects of human capital on economic growth have proliferated, mainly in conjunction with FDI effects. So, it is also important to study the relationship between foreign ownership and the workers' qualifications. The literature suggests that FF generally employ a higher fraction of qualified workers (Narula and Marin, 2003; Barbosa, 2007; Teixeira and Tavares-Lehmann, 2007), but many empirical studies conclude that this relationship is complex and nonlinear (e.g.: Blomström and Kokko, 2003). To understand the relation between foreign ownership and human capital, we need to consider, in addition, the relationship between FDI and technology. Since FF are commonly concentrated in sectors with higher technological complexity, it is expected that respective workers are more qualified, so better compensated. On the other hand, the workers of FF can be motivated to invest in their trainings and capabilities, in order to receive higher wages.

Table 13 shows a positive evolution in qualifications level of PMI's workers. However, the differences between FF and DF persist. In 1985, the disparity in average schooling was slightly larger than a year of schooling. Two decades later, this disparity overcame two years of schooling and the FF's workers presented, on average, 9 years of schooling concluded.

In 1985, only 4,5% of DF's workers had concluded, at least, secondary education, while 9% of FF's workers had that qualifications level. Two decades later, the proportion of workers with secondary education was 16,6% and 31,7% in DF and FF, respectively. From these results, we see a positive trend in human capital in Portugal, but a large distance between DF and FF. The proportion of workers with higher education has also evolved in a positive way, but foreign superiority has been always evident, as well as on other variables. Over the time, the differences have been augmenting and the disparity found at human capital level was also statistical significant at 1% level. These conclusions corroborate those ones of Almeida (2003), whose study, applied to PMI for the period of 1991-1998, found that FF had a proportion of low educated workers 7 percentage points lower than DF and pay 15 percent higher wages. Those differences, revealed by Almeida (2003), remained even after controlling for region and sector composition, as well as size and age of firms.

Table 13. Average qualifications of workers

	1985	1995	2005
Average years of schooling			
Foreign Firms	6,1	7,3	9,6
Domestic Firms	5,0	6,2	7,5
Proportion of workers with, at least, 12 years of schooling			
Foreign Firms	9,0	15,2	31,7
Domestic Firms	4,5	7,7	16,6
Proportion of workers with higher education			
Foreign Firms	3,8	2,1	11,5
Domestic Firms	1,8	1,0	5,2

Next table reflects a positive association between technological complexity, human capital quality and FDI, a fact also demonstrated for Portugal by Teixeira and Tavares-Lehmann (2007). Similarly, Barbosa (2007) states that there is a positive impact of foreign presence on human capital formation, in both groups of firms.

Table 14. Qualifications of workers by Industry

Foreign Firms				Domestic Firms			
	1985	1995	2005		1985	1995	2005
Average years of schooling				Average years of schooling			
LTI	5,9	6,8	8,6	LTI	4,9	5,6	6,9
MLTI	6,6	7,1	10,5	MLTI	5,3	6,7	7,6
MH/HTI	6,3	8,2	10,2	MH/HTI	5,1	6,7	8,4
Proportion of workers with, at least, 12 years of schooling				Proportion of workers with, at least, 12 years of schooling			
LTI	6,0	9,7	23,0	LTI	3,1	6,3	13,5
MLTI	13,2	16,0	27,9	MLTI	8,2	8,2	17,6
MH/HTI	9,9	19,3	38,9	MH/HTI	5,3	14,7	30,1
Proportion of workers with higher education				Proportion of workers with higher education			
LTI	1,8	1,0	8,8	LTI	1,0	0,6	3,9
MLTI	5,5	2,2	8,5	MLTI	4,0	1,3	5,7
MH/HTI	4,8	2,9	14,3	MH/HTI	2,4	2,5	10,9

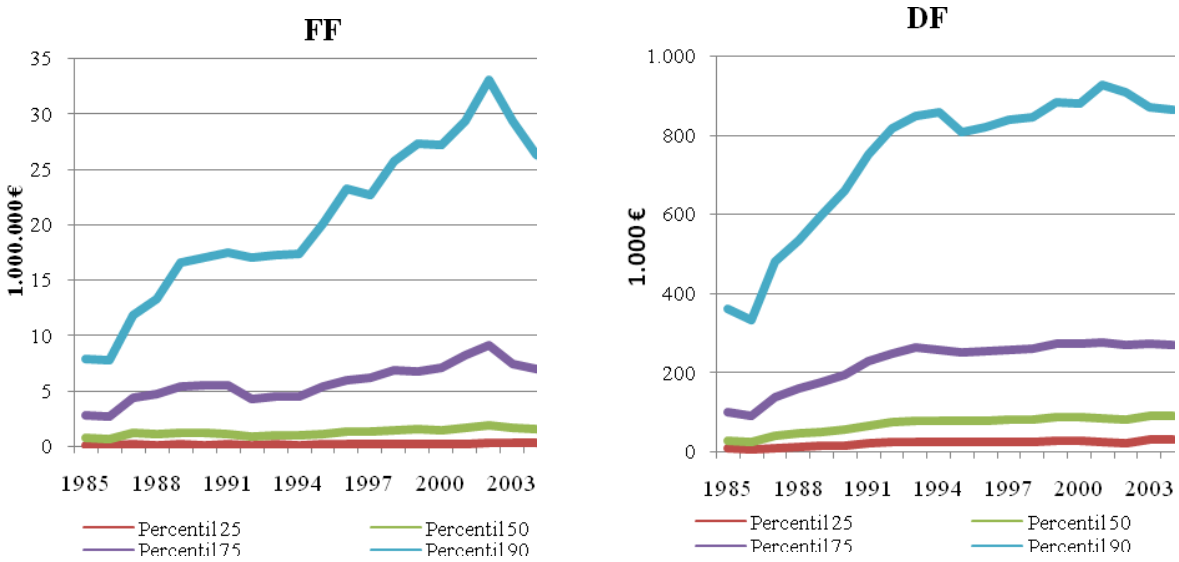
The results reveal that MH/HT industries present higher levels of human capital, both in domestic and foreign firms. However, comparing both groups, foreign firms always present higher endowments of human capital than domestic ones, corroborating the reviewed literature.

3. ASYMMETRIES IN THE EVOLUTION OF FOREIGN AND DOMESTIC FIRMS

In order to evaluate the asymmetries between FF and DF, as well as within each one of these groups, we made a succinct observation by percentiles, using the data on turnover, employment, productivity and human capital, the main variables used to observe the comparative evolution of both groups of firms. Next figures illustrate the comparative distribution by percentiles of each variable. From the distribution of turnover, visible on figure 1, we conclude that the distribution of median firm's turnover is relatively steady all over the years. Nonetheless, it is observable the increasing trend of turnover, particularly for firms located above the median one. However,

these ones also present a higher dispersion between them. Figure 1 shows, not only the huge differences in the values over which the turnover of DF and FF are distributed (highest values of domestic turnover are close to 1.000.000, while the highest values for FF are close to 35.000.000), but also a greater evolution of FF. On the other hand, the dispersal trend between upper percentiles is more apparent in DF.

Figure 1: Evolution of turnover distribution – analysis by percentiles



From Figure 2 we can confirm the negative trend of employment all over the years under observation, already noted in previous section. The median domestic and foreign firms seem relatively steady, while firms located above those ones present a downward trend over the time, not only in employment, but also in the magnitude of dispersion between upper percentiles. With those two preceding figures, we can forecast a growth in productivity of both groups of firms, due to the upward trend of turnover and the downward one in employment.

Figure 2: Evolution of employment distribution – analysis by percentiles

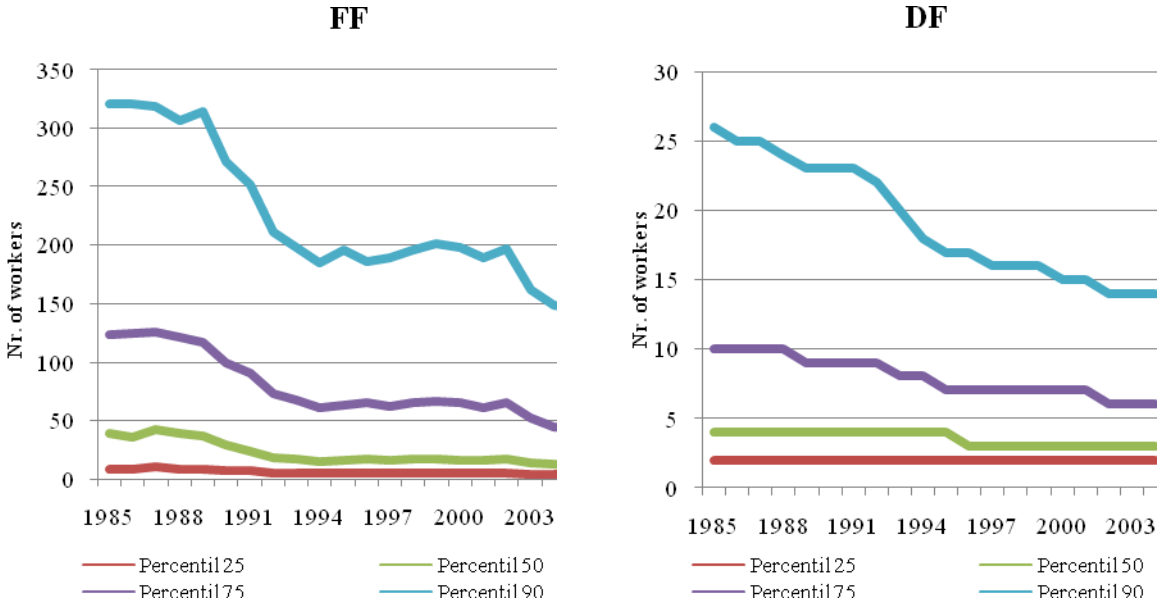
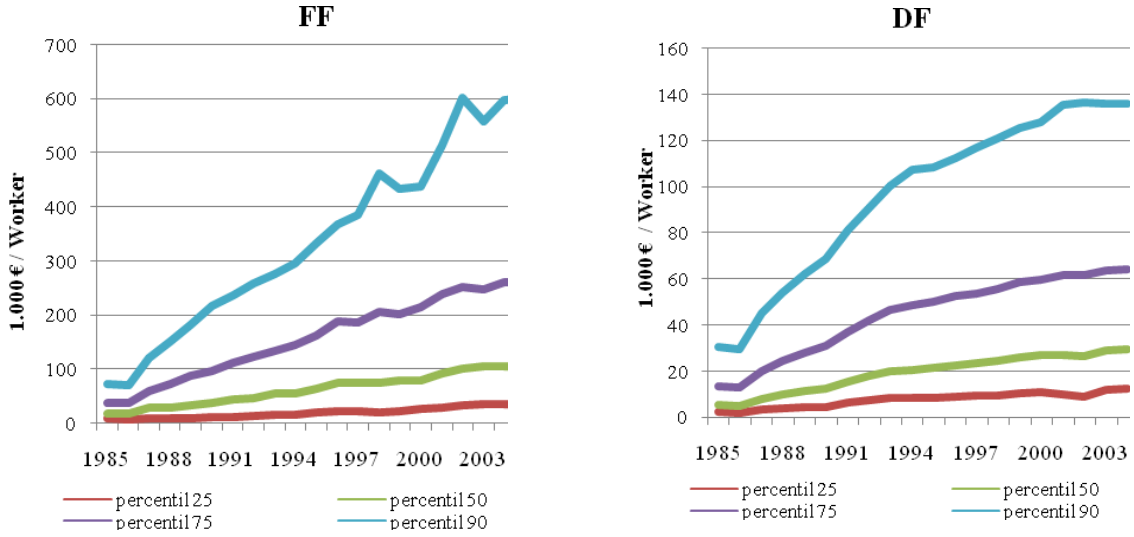


Figure 3 confirms it, illustrating an accelerated expansion of productivity in DF and FF, even in their median firms. The dispersion is persistently larger between upper percentiles and the values of foreign productivity are more than four times higher than domestic ones.

Lastly, figure 4 represents the distribution of human capital (proportion of workers with, at least, secondary education) with a very particular evolution. As a result of very low values present by DF, only three percentiles could be created – the percentiles of orders 75, 90 and 95. So, even the median domestic firm had no worker with, at least, 12 scholar years, all over the two decades.

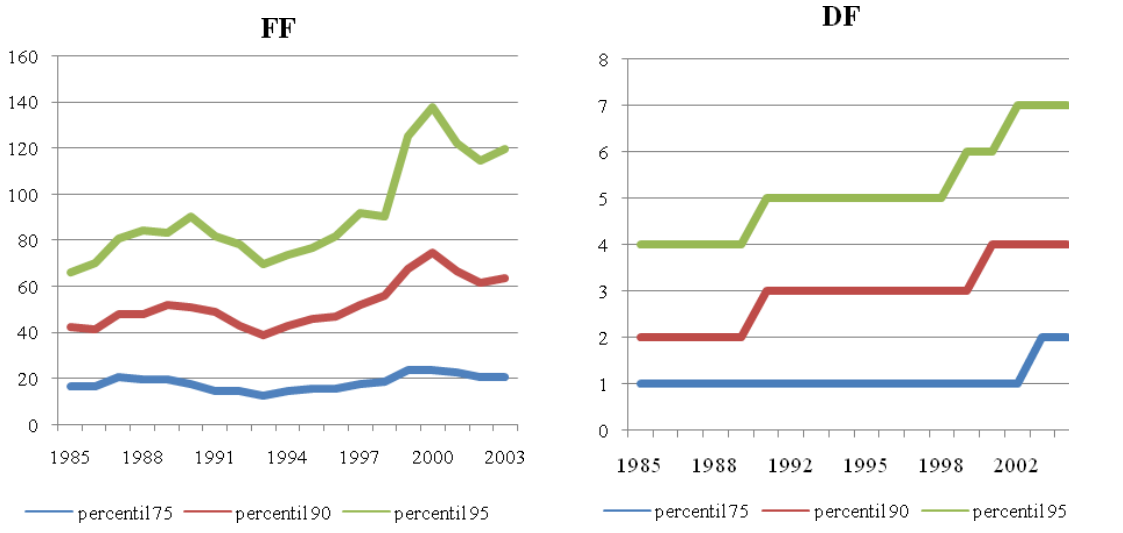
Inversely, subsidiaries of almost all percentiles always had a reasonable level of human capital, with positive trends (although figure 4 present only three percentiles for FF, more than three percentiles were analyzed).

Figure 3: Evolution of productivity distribution – analysis by percentiles



It is also visible that DF located in upper percentiles have still little human capital levels and the periods where no change have occurred were very large (about 10 years). All the preceding analyses confirm the clear superiority of FF against DF, as well as the empirical evidence present in the literature, showing that performance gaps are higher among the best FF and DF (located in upper percentiles) than among the worst FF and DF (Sabirianova, Svejnar and Terrel 2005).

Figure 4: Evolution of human capital distribution – analysis of percentiles



Next tables present some transition probability matrices, constructed by empirical observation of foreign and domestic evolutions, on turnover, employment, productivity and human capital. We specially used quartiles – percentiles of 25%, except to human capital matrices, where we just analyzed three percentiles due to very low values in DF's human capital. The period of time considered for transition was 5 years, so data in matrices gives us the probability of each group of firms to transit to an upper or lower percentile, or even stay in the same one, all over the period of 5 years. From the matrices presented below, we note a dynamic evolution of FF, against a more stable one in DF. FF show higher probabilities to transit to upper percentiles while DF present greater ones to stay in the same percentiles. However, it is also visible a cyclic movement of FF at turnover level: periods like 1985-90 and 1995-2000 were favorable and all FF located in the second percentile transited to two percentiles above in 5 years. However, the intermediate period (1990-95) was less positive, mainly to median foreign firms. All FF located in the second percentile in 1990 fell to first one in 5 years and no firm was present on third percentile. On the other hand, FF in higher positions remained steady during the period. At employment level, the conclusions are similar, as well as on productivity. FF reveal greater propensity to climb to upper stages and, in opposition, DF had always evidenced higher likelihood to keep on the same stages. On productivity evolution, we can see that in certain cases, the progress of FF was more modest than domestic one, with great propensity to remain in same stage. DF, however, had some setbacks over the time. Comparing the probabilities in matrices, DF are more likely to fall to lower percentiles than FF. Human capital matrices could only be constructed with three percentiles and since 1990, due to low values presented by DF. Besides these limitations, from matrices above we verify that it is easier for FF to overcome the current stage, while the probability of DF remain in the same percentile has been increasing over the time.

Hence, our analysis confirmed the positive evolution of FF located in Portugal and also the asymmetries between FF and DF and even within them. Other studies applied in other European countries corroborate these conclusions, showing that FF are more able to climb to higher levels of efficiency and less likely to fall (Sabirianova, Svejnar and Terrel 2005). In addition, the literature suggests that it is not surprising that FF are more efficient than DF, if they come from developed countries to countries in transition (Bellak 2004b). However, from Quadros de Pessoal database we cannot determine the origin of FF, a fact that limits our conclusions with respect to the relationship between the nationality of FF and their efficiency.

Table 15. Transition Probability Matrices – Turnover and Employment

TURNOVER

FF (1985-1990)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	22%	11%	0%	67%
2nd Perc	0%	0%	0%	100%
3rd Perc	12,5%	0%	0%	87,5%
4th Perc	4%	0%	0%	96%

DF (1985-1990)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	51,3%	23,2%	13,3%	12,2%
2nd Perc	11%	49%	33%	7%
3rd Perc	7,8%	7,5%	56%	28,7%
4th Perc	8,2%	1%	6,3%	84,5%

FF (1990-1995)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	33%	0%	11%	56%
2nd Perc	100%	0%	0%	0%
3rd Perc	0%	0%	0%	0%
4th Perc	0,5%	0,5%	0%	99%

DF (1990-1995)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	43%	14%	16%	27%
2nd Perc	10,5%	59%	26,5%	4%
3rd Perc	4%	10%	65%	21%
4th Perc	3,5%	1%	7,5%	88%

FF (1995-2000)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	40%	20%	0%	40%
2nd Perc	0%	0%	0%	100%
3rd Perc	0%	0%	100%	0%
4th Perc	3,2%	1,3%	0,6%	94,9%

DF (1995-2000)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	53,6%	16,7%	10,5%	19,2%
2nd Perc	9,5%	63%	24%	3,5%
3rd Perc	2,6%	9,8%	68,6%	19%
4th Perc	2,8%	0,6%	5,9%	90,7%

FF (2000-2005)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	29%	0%	0%	71%
2nd Perc	33,3%	33,3%	0%	33,3%
3rd Perc	50%	0%	50%	0%
4th Perc	3%	0%	1%	96%

DF (2000-2005)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	56%	16%	9%	19%
2nd Perc	13%	65%	19%	3%
3rd Perc	4%	14%	67%	15%
4th Perc	3,4%	1,1%	8,4%	87,1%

EMPLOYMENT

FF (1985-1990)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	25%	0%	50%	25%
2nd Perc	12,50%	0%	50%	37,50%
3rd Perc	0%	0%	33%	67%
4th Perc	0%	0%	0%	100%

DF (1985-1990)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	48%	34%	15%	3%
2nd Perc	11%	42%	39%	8%
3rd Perc	1,3%	7%	57%	34,7%
4th Perc	0,3%	0,7%	6%	93%

FF (1990-1995)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	100%	0%	0%	0%
2nd Perc	0%	0%	0%	100%
3rd Perc	25%	0%	50%	25%
4th Perc	0,7%	0%	0%	99,3%

DF (1990-1995)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	61%	30%	7%	2%
2nd Perc	18,5%	54,5%	22%	5%
3rd Perc	3,2%	17,3%	46%	33,5%
4th Perc	0,5%	1,4%	4,2%	93,9%

FF (1995-2000)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	75%	25%	0%	0%
2nd Perc	0%	0%	0%	0%
3rd Perc	0%	0%	50%	50%
4th Perc	0%	0%	4%	96%

DF (1995-2000)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	70,5%	16,5%	11%	2%
2nd Perc	22%	32%	42%	4%
3rd Perc	4%	8,5%	69,5%	18%
4th Perc	0,7%	0,8%	10%	88%

FF (2000-2005)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	100%	0%	0%	0%
2nd Perc	0%	0%	100%	0%
3rd Perc	0%	25%	37,5%	37,5%
4th Perc	0%	0,7%	0%	99,3%

DF (2000-2005)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	42,8%	48,7%	7%	1,5%
2nd Perc	13,1%	67,2%	17,5%	2,2%
3rd Perc	3,8%	26,5%	52,1%	17,6%
4th Perc	0,7%	2%	7,6%	89,7%

Table 16. Transition Probability Matrices – Productivity and Human Capital

PRODUCTIVITY

FF (1985-1990)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	23%	15,5%	15,5%	46%
2nd Perc	0%	62,5%	37,5%	0%
3rd Perc	10%	14%	55%	21%
4th Perc	4,5%	0%	4,5%	91%

DF (1985-1990)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	50%	27%	13%	10%
2nd Perc	13%	55%	27%	5%
3rd Perc	8%	21%	54%	17%
4th Perc	9%	3%	16%	72%

FF (1990-1995)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	36%	9%	9%	46%
2nd Perc	36%	64%	0%	0%
3rd Perc	0%	15%	66%	19%
4th Perc	0,9%	1,7%	3,4%	94%

DF (1990-1995)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	45%	22%	17%	16%
2nd Perc	15%	61%	20%	4%
3rd Perc	4,7%	24, 1%	57,7%	13,5%
4th Perc	4%	3%	19%	74%

FF (1995-2000)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	67%	0%	11%	22%
2nd Perc	7%	64%	7%	22%
3rd Perc	0%	5%	77%	18%
4th Perc	5,0%	0%	2,5%	92,5%

DF (1995-2000)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	53%	25%	11%	11%
2nd Perc	9,5%	63%	23,8%	3,7%
3rd Perc	3%	17%	60%	20%
4th Perc	3,5%	2,4%	13,4%	80,7%

FF (2000-2005)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	31%	23%	15%	31%
2nd Perc	10%	70%	10%	10%
3rd Perc	18%	0%	45,5%	36,5%
4th Perc	1,7%	0,8%	2,5%	95,0%

DF (2000-2005)

	1st Perc	2nd Perc	3rd Perc	4th Perc
1st Perc	55%	25%	10%	10%
2nd Perc	10,7%	62,3%	23,5%	3,5%
3rd Perc	4%	18%	60%	18%
4th Perc	4%	3%	15%	78%

HUMAN CAPITAL

FF (1990-1995)

	1st Perc	2nd Perc	3rd Perc
1st Perc	36,5%	36,5%	27%
2nd Perc	8%	15%	77%
3rd Perc	0%	2%	98%

DF (1990-1995)

	1st Perc	2nd Perc	3rd Perc
1st Perc	68%	16%	16%
2nd Perc	32%	26%	42%
3rd Perc	12%	12%	76%

FF (1995-2000)

	1st Perc	2nd Perc	3rd Perc
1st Perc	60%	20%	20%
2nd Perc	33%	11%	56%
3rd Perc	2%	0,7%	97,3%

DF (1995-2000)

	1st Perc	2nd Perc	3rd Perc
1st Perc	72%	15%	13%
2nd Perc	22%	37%	41%
3rd Perc	5%	8%	87%

FF (2000-2005)

	1st Perc	2nd Perc	3rd Perc
1st Perc	56%	22%	22%
2nd Perc	0%	67%	33%
3rd Perc	0,7%	1,3%	98%

DF (2000-2005)

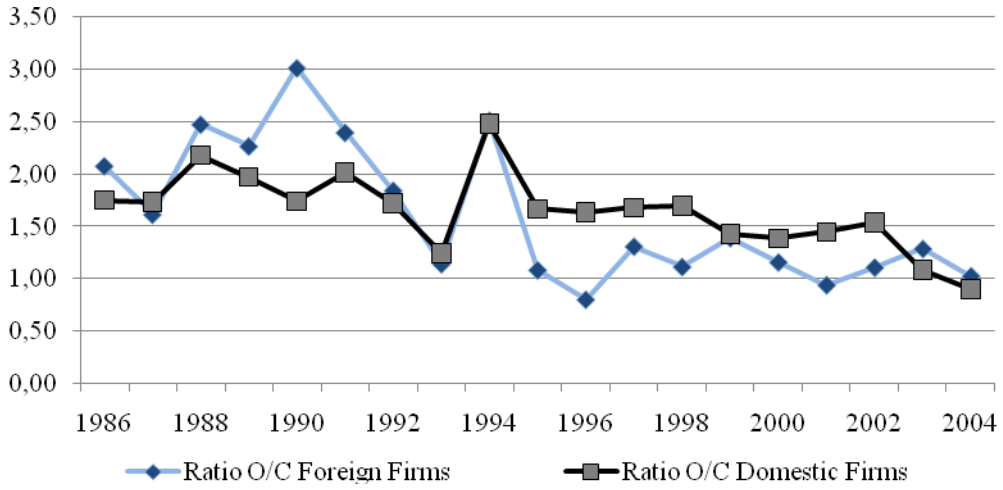
	1st Perc	2nd Perc	3rd Perc
1st Perc	77%	15%	8%
2nd Perc	27%	42%	31%
3rd Perc	5%	10%	85%

4. DYNAMICS OF FOREIGN FIRMS

Another point that differentiates FF from DF relates to demographical dynamics. From the whole sample of firms analyzed during the period 1985-2005, only 165 foreign firms (19%), from the group of 872 FF working in Portugal in 1985, had survived until 2005. In the case of DF, only 25.083 (25%) from the group of 101.329 registered in Quadros de Pessoal in 1985 remained active until 2005. The probability of survival between 1985 and 2005 was 18.92% for FF and 24.75% for domestic firms. Moreover, domestic firms survived, in average, 5,4 years, while FF stayed alive, in average, 14 years, during the period of two decades under study.

So, foreign supremacy is not limited to scale, efficiency and human capital, because FF reveal an average lifetime also greater than domestic one. Based on records in database, we also verified the number of investments and disinvestments, year after year. We calculated the open to closure ratio and the results are presented in Figure 5. We can immediately realise that the ratio has declined since the 1980's. By 2004 the number of openings is equal or less than the number of closures.

Figure 5. Ratio O/C



5. CONCLUSIONS AND DISCUSSION

The literature suggests that MNEs may be an important element in the strategies of countries' development and that foreign ownership is an important source of firm heterogeneity affecting productivity dynamics (Sleuwaegen and De Backer, 2003a, 2003b). However, the potential impact of FDI in host economies depends not only on the quality of the investments, but also on the way they evolve over time. For the Portuguese case, we found foreign subsidiaries to be relatively more productive, technologically more advanced and richer regarding human capital, which is likely to alter the patterns of jobs and thus expand specialized and highly skilled employment. Additionally, we demonstrate that foreign investments in Portugal have improved at those levels and the trends exposed above show that the attraction of FDI to MH/HT industries must be supported, as those industries already present a significant role on FDI composition. Moreover, the reduction of foreign activities in Portugal, mainly those focused on exports, may be a threat for a country where FDI has been crucial regarding industrial diversification and access to new markets. For Portugal, a country with huge needs to evolve at efficiency, technological intensity and human capital levels, in order to ease convergence process, inward FDI should be sustained, especially when new competitors seem to emerge (Galego et al., 2004; Oxerlheim and Ghauri, 2004; Ferragina and Pastore, 2007).

The longevity of foreign presence depends on the strategic choices of the firm (Li, 1995). Understanding the forces and the reasons that determine it is important to improving our

knowledge of the process of international expansion. This is not merely academic and should attract considerable attention from practitioners as well (Mata and Portugal, 1994), as managers considering going into international markets are interested in evaluating the chances of success.

We have proved that firms in the two samples differ widely in a number of aspects. Foreign firms are larger (in turnover and in the number of employees), employ a larger proportion of more educated workers, present higher probabilities to transit to upper percentiles and seem to have higher longevity. However, a deeper investigation, controlling firms by industry, size and capital intensity is needed, to give more robustness to these results. For Mata and Portugal (2001), being foreign does not decrease the chances of failure, does not imply different effects of the variables affecting survival and does not necessarily translate into different time patterns of survival. On the other hand, other surveys argue that foreign owned firms will be less likely to shut down, due to their superior productivity against domestic competitors (Falvey, Greenaway and Yu, 2007). So, conclusions about dynamics of firms' evolution should not be based on foreignness alone.

For an economy well integrated in the world market such as Portugal, those differences found between foreign and domestic firms may have strong implications for managers and policy-makers. For managers, they should not assume that new firms will stay in the market longer just because they are foreign. Foreignness may be a summary measure for a number of characteristics, but not a guarantee of success. For policy-makers, mainly those concerned with FDI, is important to evaluate the advantages of foreign presence in host countries to decide to encourage or not foreign investments.

For small open economies, as Portugal, efficient taxation of foreign and domestic capital depends on their relative mobility. If foreign capital is more mobile internationally, it will be optimal for policy-makers to fix lower taxes on capital owned by foreign firms, because higher taxes deter FDI. Moreover, there is empirical evidence that a more educated workforce and larger markets attract FDI, and MNEs seem to prefer high productivity and technologically more advanced industries (Hanson, 2001). The same studies also divulge that the relative technological superiority of MNEs make them a source of technological advancement for domestic firms, especially when host countries are relatively far from the technological frontier.

So, facing those evidences, the attitude of governments has changed and they are now actively providing generous investment incentives to attract FDI. Subsidizing FDI, governments typically assume that spillovers effects are present, so foreign subsidiaries are able to contribute to growth in host economies. However, the findings of a positive correlation between productivity growth of domestic firms and the increase of multinational presence should not be interpreted as a call of subsidizing FDI (Smarzynska, 2003). If foreign presence guarantees the existence of knowledge spillovers from foreign affiliates to local firms, policy-makers must act offering FDI incentives packages. If, on the other side, productivity spillovers occur from domestic firms to foreign ones, as Martins (2008) detected in Portuguese case, with foreign firms attracting the "best" workers of their domestic counterparts, by offering them large wage increases, there is no justification to give priority to foreign presence. In short, the existence of spillovers is closely related to a) local linkages between foreign and domestic firms, and b) the gap between them. Crespo, Proença and Fontoura (2007), with data from Dun&Bradstreet for the period 1996-2000, proved the existence of FDI spillovers in Portuguese economy. However, FDI impact seems to be positive only in vertical externalities, when firms operate in different sectors, though vertically related. Regarding the need of a gap, Flores, Fontoura and Santos (2007) confirmed that, for the period 1992-1995, spillovers only occur if there is a technological gap between the groups of firms. On the other hand, the surveys of Crespo and Fontoura (2007) and Pessoa (2007) refer that such gap must not be very large, otherwise domestic firms will have no absorption capacity. So, regarding this fact, specific programs aiming to transfer the best practices from foreign firms to domestic ones should be implemented. Such programs could focus, for instance, on promoting linkages between the firms and other mechanisms of sharing knowledge and managerial practices (Tavares, 2002; Cardoso, 2008).

Concluding, careful investigations on these effects seem to be in order for a deepening of our understanding on the advantages of the presence of foreign firms that seem so different from domestic ones. The competition between governments on FDI attraction may result in high

opportunity costs when compared to the alternative policy: support and encourage domestic firms. Inefficient attraction policies may result on huge social welfare losses, so governments should consider policies more focused on gaps and less on foreign ownership alone.

REFERENCES

AITKEN, B., HARRISON, A. (1994), "Do Domestic Firms Benefit from Foreign Direct Investment?". *Policy Research Working Paper Nº 1248*, Policy Research Department, Trade Policy Division, Washington D.C..The World Bank.

ALMEIDA, R. (2003), "The Effects of Foreign Owned Firms on the Labor Market", *IZA Discussion Paper* No. 785.

ANDERSSON, F., VEJSIU, A. (2001), "Determinants of Plant Closures in Swedish Manufacturing"*, Office of Labour Market Policy Evaluation.*

ANDERSSON, U., FORSGREN, M., PEDERSON, T. (2001), "Subsidiary Performance in Multinational Corporations: the Importance of Technology Embeddedness", *International Business Review*, 10, 3-23.

BANK OF PORTUGAL (2003), Annual Report of 2003.

BARBOSA, N. (2007), "Cross-border Acquisitions and Human Capital Intensity", *Paper Presented at the 3rd IIBC 2007,* Aveiro.

BARBOSA, N., GUIMARÃES, P., WOODWARD, D. (2004), "Foreign Entry in an Open Economy: The Case of Portugal", *Applied Economics*, 36, 5, 465-472.

BARBOSA, N., LOURI, H. (2005), "Corporate Performance: Does Ownership Matter? A Comparison of Foreign- and Domestic-Owned Firms in Greece and Portugal", *Review of Industrial Organization*, 27, 1, 73-102.

BATES, T. (1990), "Entrepreneur Human Capital Inputs and Small Business Longevity", *The Review of Economics and Statistics*, 62, 4, 551-559.

BELLAK, C. (2004a), "How Performance Gaps Between Domestic Firms and Foreign Affiliates Matter for Economic Policy?", *Transnational Corporations*, 13, 2, 29-55.

BELLAK, C. (2004b), "How Domestic and Foreign Firms Differ and Why Does it Matter?", *Journal of Economics Surveys*, 18, 4, 483-514.

BEN HAMIDA, L. (2007), *Inward Foreign Direct Investment and Intra-Industry Spillovers: The Swiss Case,* PhD Thesis presented to the Faculty of Economics and Social Sciences at the University of Fribourg: Switzerland.

BIRKINSHAW, J. (2001), "Strategy and Management in MNE Subsidiaries", in A.M. Rugman and T. Brewer (eds.), *The Oxford Handbook of International Business*. Oxford: Oxford University Press, 380-401.

BLOMSTRÖM, M., KOKKO, A. (1998), "Multinational Corporations and Spillovers", *Journal of Economic Surveys*, 12, 247-277.

BLOMSTRÖM, M., KOKKO, A. (2003), "Human Capital and Inward FDI", *CEPR Discussion Paper Nº 3762*. London.

BLOMSTRÖM, M., WOLFF, E. (1994), "Multinational Corporations and Productivity Convergence in Mexico". *NBER Working Paper Nº 3141*.

BOARDMAN, A., SHAPIRO, D., VINING, A. (1998), "The Role of Agency Costs in Explaining the Superior Performance of Foreign MNE Subsidiaries", *International Business Review*, 6, 3, 295-317.

BRENTON, P., DI MAURO, F., LUCKE, M. (1998), "Economic Integration and FDI: An Empirical Analysis of Foreign Investment in the EU and in Central and Eastern Europe", *Kiel Working Paper Nº 890*.

BUCKLEY, P., CASTRO, F. (1998), "The Investment Development Path: the Case of Portugal", *Transnational Corporations*, 7, 1, 1-15.

CABRAL, S. (1996), "Comparative Export Behaviour of Foreign and Domestic Firms in Portugal", *Banco de Portugal Working Paper Nº 8/96*.

CARDOSO, V. (2008), *Foreign Multinationals and Domestic Companies in Portugal: Are there Significant Performance Gaps?,* Master Thesis presented to Faculdade de Economia, University of Porto: Portugal.

CASTRO, FRANCISCO (2004), "Foreign Direct Investment in a Late Industrializing Country: The Portuguese IDP revisited", *FEP Working Paper* No. 147, University of Porto.

CAVES, R. (1974), "Multinational Firms, Competition and Productivity in Host-Country Markets", *Economica*, 41, 176-193.

CHUANG, Y., LIN, C. (1999), "Foreign Direct Investment, R&D and Spillover Efficiency: Evidence from Taiwan's Manufacturing Firms", *Journal of Development Studies*, 35, 4, 117-137.

CRESPO, N., FONTOURA, M. (2007), "Determinant Factors of FDI Spillovers - What Do We Really Know?", *World Development*, 34, 3, 410-425.

CRESPO, N., PROENÇA, I., FOUTOURA, M. P. (2007), "FDI Spillovers at Regional Level: Evidence from Portugal", *ISEG WP Nº 028/2007.*

DUNNING, J. (1981), *International Production and Multinational Enterprise*, George Allen and Unwin: London.

ERNST, C. (2005), "The FDI-Employment Link in a Globalizing World: the Case of Argentina, Brazil and Mexico", *Employment Strategy Papers,* Employment Strategy Department.

FALVEY, R., GREENAWAY, D., YU, Z. (2007), "Market Size and the Survival of Foreign-owned Firms", *Research Paper Series, University of Nottingham.*

FARINHA, L., MATA, J. (1996), "The Impact of Foreign Direct Investment in the Portuguese Economy", *Banco de Portugal Working Paper Nº 16/96.*

FERRAGINA, A., PASTORE, F. (2007), "FDI Potential and Shortfalls in the MED and CEECs: Determinants and Diversion Effects", *Paper presented at the CNR Conference on: Bridging the gap: The role of Trade and FDI in the Mediterranean,* Naples.

FLORES, R., FONTOURA, M., SANTOS, R. (2007), "Foreign Direct Investment Spillovers: Additional Lessons from a Country Study", *The European Journal of Development Research*, 19, 3, 372-390.

FREITAS, M. L., PAES MAMEDE, R. (2008), "Structural Transformations and the Role of Foreign Direct Investment in Portugal: a Descriptive Analysis for the Period 1990-2005", *GEE Papers, Working Paper Nº 9.*

GALEGO, A., VIEIRA, C., VIEIRA, I. (2004), "The CEECs as FDI Attractors: Are They a Menace to the EU Periphery?", *Emerging Markets Finance and Trade*, 40, 5, 74-91.

GEROSKI, P. A., MATA, J., PORTUGAL, P. (2003), "Founding Conditions and the Survival of New Firms", *Banco de Portugal Working Paper Nº 1/03.*

GLOBERMAN, S., RIES, J., VERTINSKY, I. (1994), "The Economic Performance of Foreign Affiliates in Canada", *Canadian Journal of Economics*, Canadian Economics Association, 27, 1, 143-156.

GONÇALVES, O., GUIMARÃES, P. (1997), "O Investimento Directo Estrangeiro na Indústria Transformadora Portuguesa: uma Abordagem Sectorial e Regional através do Emprego para o Período 1982-1992", *Estudos de Economia*, 16-17, 3, 333-358.

GRIFFITH, R. (1999), "Using the ARD Establishment Level Data to Look at Foreign Ownership and Productivity in the UK", *Institute for Fiscal Studies*, Working Paper Nº 99/6. London.

GRIFFITH, R., SIMPSON, H. (2001), "Characteristics of Foreign-Owned Firms in British Manufacturing", *Institute for Fiscal Studies*, Working Paper Nº 01/10. London.

HAMERMESH, D. (1988), "Plant Closings and the Value of the Firm", *Review of Economics and* Statistics, 70, 580-586.

HANSON, G. (2001), "Should Countries Promote Foreign Direct Investment?", *G24 Discussion Paper Series Nº 9*. New York and Geneva: United Nations.

HARRIS, R., ROBINSON, C. (2002), "Productivity Spillovers to Domestic Plants from Foreign Direct Investment: Evidence from UK Manufacturing, 1974-1995", *Royal Economic Society Annual Conference 2002*.

HASKEL, J., PEREIRA, S., SLAUGHTER, M. (2004), "Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms?", Working Paper Nº 452, Queen Mary, University of London.

HATZICHRONOGLOU, T. (1997), "Revision of the High-Technology Sector and Product Classification", *OECD Science, Technology and Industry Working Papers1997/2*.

HYMER, S. (1976), *The International Operations of National Firms: A Study of Direct Foreign Investment,* Cambridge, Mass: MIT Press.

JIMENO, J., CANTÓ, O., CARDOSO, A., IZQUIERDO, M., RODRIGUES, C. (2000), "Integration and Inequality: Lessons from the Accessions of Portugal and Spain to the EU", *FEDEA Working Paper 2000-10.*

KIMURA, F., KIYOTA, K. (2007), "Foreign-Owned Versus Domestically-Owned Firms: Economic Performance in Japan", *Review of Development Economics*, 11, 1, 31-48.

KUEMMERLE, W. (1999), "The Drivers of Foreign Direct Investment into Research and Development – An Empirical Investigation", *Journal of International Business Studies*, 30, 1, 1-24.

LI, J. (1995), "Foreign Entry and Survival: Effects of Strategic Choices on Performance in International Markets", *Strategic Management Journal*, 16, 5, 333-351.

MAMEDE, R., MOTA, D., GODINHO, M. M. (2007), "Are the Dynamics of Knowledge-Based Industries any different?", *GEE Papers,* WP No.3.

MARCHICK, D., SLAUGHTER, M. (2008), "Global FDI Policy: Correcting a Protectionist Drift", *Council Special Report 34*. New York: Council on Foreign Relations.

MARKUSEN, J. R. (1995), "The Boundaries of Multinational Enterprises and the Theory of International Trade", *Journal of Economic Perspectives*, 9, 2, 169-189.

MARKUSEN, J. R. (2002), *Multinational Firms and the Theory of International Trade*, Cambridge, Mass: The MIT Press.

MARTINS, P. (2008), "Paying More to Hire the Best?", *IZA Discussion Paper Nº 3607.*

MATA, J., PORTUGAL, P. (1999), "Closure and Divestiture by Foreign Entrants: the Impact of Entry and Post-Entry Strategies", *Banco de Portugal Working Paper Nº 6/99*.

MATA, J., PORTUGAL, P. (2001), "The Survival of New Domestic and Foreign Owned Firms", *Banco de Portugal Working Paper Nº 1/01*.

MATA, J., PORTUGAL, P. (2004), "Patterns of Entry, Post-Entry Growth and Survival: A Comparison Between Domestic and Foreign Owned Firms", *Small Business Economics*, 22, 3, 4, 283-298.

MATA, J., PORTUGAL, P (1994), "Life Duration of New Firms", *Journal of Industrial Economics*, 42, 227-246.

NARULA, R., MARIN, A. (2003), "FDI Spillovers, Absorptive Capacities, and Human Capital Development", *Research Memorandum Series*, 018, MERIT. Maastricht.

OXELHEIM, L., GHAURI, P. (2004), *European Union and the Race for Foreign Direct Investment in Europe.* Oxford: Elsevier.

PEREIRA, V. M. (2007), "O IDE em Portugal entre 1975 e 1995", *Departamento de Gestão e Economia*, Universidade da Beira Interior.

PESSOA, A. (2007), "FDI and Host Country Productivity: A Review", *FEP Working Paper Nº 251*.

RAMSTETTER, E. (1999), "Comparisons of Foreign Multinationals and Local Firms in Asian Manufacturing Over Time", *Asian economic Journal*, 12, 2, 163-203.

RIBEIRO, J. C., SANTOS, J. (2001), "Factores de Atracção da Empresa Estrangeira: A Experiência Portuguesa". Encontro de economistas de língua portuguesa. Universidade de Évora: Portugal.

SABIRIANOVA, K., SVEJNAR, J., TERRELL, K. (2005), "Foreign Investment, Corporate Ownership, and Development: Are Firms in Emerging Markets Catching Up to the World Standard?", *IZA Discussion Paper* No. 1457.

SLEUWAEGEN, L., DE BACKER, K. (2003a), "Foreign Ownership and Productivity Dynamics?",*Economic Letters,* 79, 2, 177-183.

SLEUWAEGEN, L., DE BACKER, K. (2003b), "Does foreign direct investment crowd out domestic entrepreneurship?", *Review of Industrial Organization*, 22, 67-24.

SMARZYNSKA, B. (2003), "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In search of Spillovers through Backward Linkages", *The William Davidson Institute, Working Paper* No. 548.

TAVARES, A. T. (2001), *Systems, Evolution and Integration: Modelling the Impact of Economic Integration on Multinationals' Strategies*. PhD Thesis presented to University of Reading: UK.

TAVARES, A. T. (2002), "Multinational Subsidiary Evolution and Public Policy: Two Tales from the European Periphery", *Journal of Industry, Competition and Trade*, 2, 3, 195-213.

TAYLOR, K., DRIFFIELD, N. (2004), "Wage Inequality and the Role of Multinationals. Evidence from UK Panel Data", *Labour Economics*. Elsevier Science Economics Articles Archive.

TEIXEIRA, A., TAVARES-LEHMANN, A. T. (2007), *Investimento Directo Estrangeiro, Capital Humano e Inovação. Uma Aplicação ao Caso Português*. Vida Económica: Porto.

TEIXEIRA, A. (2002), "On the Link Between Human Capital and Firm Performance", *FEP Working Paper* No. 121, University of Porto.

UNCTAD, United Nations Conference on Trade and Development, 2007. World Investment Report, 2007.

VERNON, R. (1966), "International Investment and International Trade in the Product Cycle", *Quarterly Journal of Economics*, LXXX, 190-207.

YASAR, M., PAUL, C. (2007), "Firm Performance and Foreign Direct Investment: Evidence from Transition Economies", *Economics Bulletin*, 15, 21, 1-11.

RETOS ESTRATÉGICOS DE LOS DEPARTAMENTOS DE COMPRAS EN ESPAÑA Y PORTUGAL

RICARDO ÚBEDA
ricardo.ubeda@iqs.es
CARLOS MOSLARES
moslares@iqs.es
Facultat d´Economia IQS
Universitat Ramon LLull

Resumen

Tradicionalmente el departamento de Compras ha servido para aprovisionar de bienes y servicios a la empresa. Este papel limita las posibilidades que Compras puede aportar a la empresa tanto en términos de ahorros como en aporte de valor. La conexión entre la cadena de valor de la empresa y toda su red de proveedores es Compras. En dicha red, el valor y la innovación a captar son enormes. El potencial que cuenta Compras para atraer y gestionar el valor exterior es una oportunidad que puede ayudar a las empresas a competir. El artículo describe las etapas por la cuales ha de pasar el desarrollo de Compras para lograr que sea un departamento que apoye la competitividad de la empresa.

Palabras clave: Competitividad, gestión compras, gestión de proveedores, innovación.

1. INTRODUCCIÓN

Recientemente, los resultados de la encuesta *"The CEO Challenge"*, realizada por *"The Conference Board"* a un centenar de Directores Generales, ha confirmado lo que muchas empresas ya sospechaban o estaban experimentando en estos tiempos. Dicha encuesta se basa en evaluar cuales son, siempre según los CEO's (*Chief Executive Officer*), los principales retos que afrontarán sus empresas en los próximos años. Parece haber un consenso en que estos retos serán afrontar la presión competitiva de un entorno cada vez más globalizado, gestionar la presión de los incrementos de costes y enfocar la empresa hacia la imperiosa necesidad de innovar.

Hasta aquí no parece que haya nada excesivamente nuevo o sorprendente, es, en si, un poco –o bastante- más de lo mismo. Sin embargo, en los últimos años, se están produciendo algunos cambios sobre todo en el modo de afrontar estas situaciones. Dichos cambios provienen del uso que las empresas hacen de sus recursos a la hora de afrontar estas o cualquier otras situaciones.

2. COMPRAS: EL INICIO DEL CAMBIO

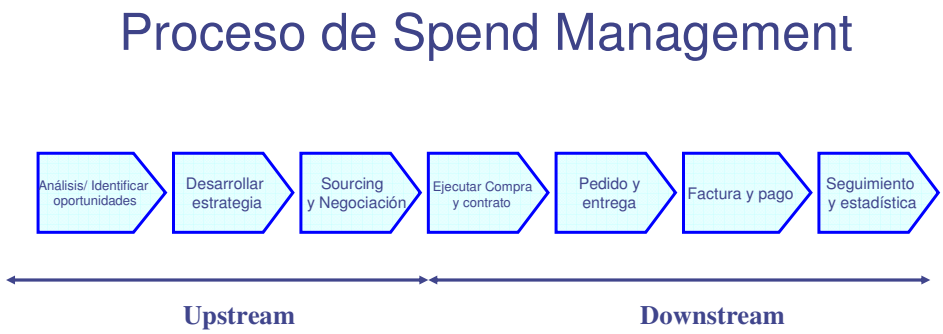
Durante los últimos veinte años se ha visto en la práctica, y se ha descrito en ámbitos académicos, la evolución del departamento de Compras desde su gestión como departamento operativo a departamento estratégico (Carter, 1996).

Si consideramos el proceso de *"spend management",* proceso por el cual la empresa se aprovisiona de bienes y servicios podemos diferenciar dos partes; el *upstream* y el *downstream*. (3 van weele o calvinato libro).

Tradicionalmente, el rol de Compras se había basado en ser un departamento transaccional, puramente operativo, que, tras negociar, optaba por el proveedor con el precio más reducido. Tras elegirlo se aseguraba que la mercancía o servicio estuvieran disponibles en el momento adecuado, en el lugar adecuado y en las condiciones pactadas. Posteriormente autorizaba el

pago del pedido. Esta función, basada fundamentalmente en el *downstream,* ha sido la visión operativa de Compras que ha estado vigente hasta hace unos 10 o 15 años atrás. (Freeman, Cavinato 1990, Morgan 1987). La gestión era simple; tener una base amplia de proveedores y negociar por igual hasta lograr las reducciones precios, que, en ocasiones, no lo eran de costes.

Figura 1: Proceso de Spend Management



Fuente: Elaboración propia

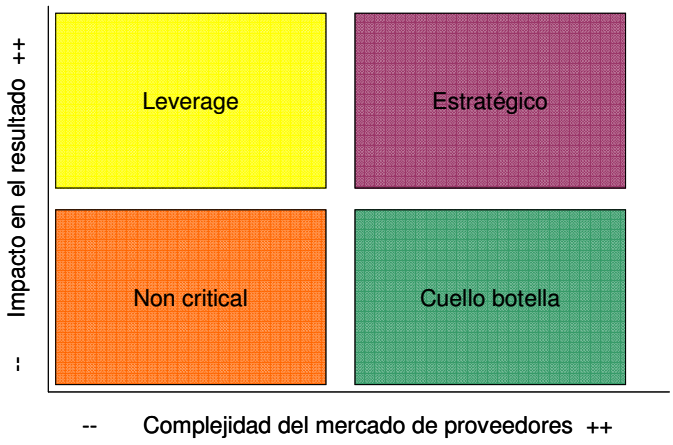
La transformación de la gestión exclusiva del *downstream* a la gestión del *upstream*, se produjo a raíz de la publicación de un artículo de Peter Kraljic en 1983.
El artículo recogía lo que han sido las bases de la gestión avanzada de Compras:

- La gestión estratégica de Categorías
- La gestión de la relación con los proveedores
- La gestión del suministro

A estos tres puntos, se le podría añadir la gestión del valor (Civit, J.2004) La gestión por categorías implica un cambio radical en la gestión del departamento de Compras. Se trata de agrupar las compras en categorías similares y cada una de ellas a su vez, posicionarlas en uno de los cuadrantes de la matriz de Kraljic. Dependiendo de cada uno de los cuadrantes se aplicarán estrategias diferentes en la gestión de las categorías y de los proveedores. De este modo aparecen cuatro cuadrantes básicos:

- Categorías estratégicas, con alto impacto en la cuenta de resultados y con un mercado complejo
- Categorías "cuello de botella" con un impacto bajo en la cuenta de resultados y un mercado complejo
- Categoría "leverage" con un impacto alto en la cuenta de resultados y un mercado de proveedores simple
- Categoría "non critical" con bajo impacto en la cuenta de resultados y un mercado simple

Figura 2: Matriz Kraljic



Fuente: Peter Kraljic, HBS , 1983

La gestión de las categorías incluidas en los diferentes cuadrantes es, lógicamente, distinta. Por ejemplo, en una fábrica de caramelos donde la glucosa está en manos de pocos proveedores, será una categoría estratégica, con lo cual tendrá una relación con el proveedor determinada (probablemente casi de socio o al menos muy a largo plazo). Sin embargo la compra de folios para las oficinas será probablemente "*non critical*" con lo que puede aplicar herramientas de *e-procurement* o subastas inversas a la hora de comprar. La relación con el proveedor no será tan intensa, e incluso será inexistente. Se buscará exclusivamente la reducción de costes, mientras que en el caso de la glucosa se buscará calidad, mejoras en costes y por que no, innovación.

Por ello las estrategias a aplicar en las compras no se basan en exclusiva en el ahorro de costes generados en negociaciones aplicadas por igual a todos los proveedores, si no que existen más factores a tener en cuenta.

Este es el principal cambio que originó el pasar de contar con unas Compras transaccionales u operativas a Compras Estratégicas. El conocimiento del mercado y las fuerzas que lo impulsan, la gestión de los proveedores y la medición de la aportación de valor de los mismos hizo que, hace unos quince o veinte años, determinadas industrias cambiaran el enfoque de sus Compras y lograran notables descensos de costes, mediante la gestión y la estrategia apropiada a la categoría apropiada.

Esta gestión por categorías se considera la gestión del *upstream,* es decir, conocer y analizar los mercados de proveedores, crear estrategias de compra en cada uno de ellos realizar *sourcings* y establecer la relación y gestión de los proveedores apropiada, en definitiva, gestionar las Compras estratégicamente.

3. COMPRAS COMO DEPARTAMENTO ESTRATÉGICO

Las diferentes presiones competitivas que sufren los diferentes sectores industriales llevó a que una serie de ellos fijaran, hace poco más de una década, su atención a la Gestión Estratégica de Compras.

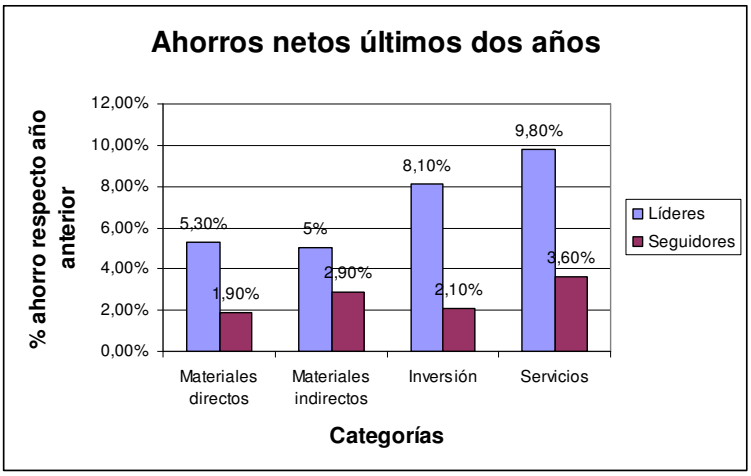
Los primeros indicadores y las primeras acciones estaban enfocados al ámbito financiero. De este modo, los primeros estudios acerca de la influencia de Compras en los resultados de la empresa, (Carr, Pearson, 2002) midieron el perímetro de Compras necesario para que el departamento fuera eficiente. Es decir, la centralización de todas las Compras en un solo departamento.

Aunque esto pueda parecer normal hoy en día, todavía sorprende la enorme cantidad de Compras que se realizan en los departamentos no dedicados a ello. Como ejemplo puede comentarse que atendiendo al Benchmark (Future purchasing alliance) realizado sobre 1.195 empresas en 2004, hasta un 41% de las empresas expuso gestionar menos del 60% del gasto en Compras de la empresa, lo cual apunta a una gestión ineficiente de las Compras realizadas, dado que con dicho volumen es muy difícil poder aplicar una gestión por categorías apropiada.

Teniendo en cuenta que Compras debería gestionar todo lo que la empresa compra, es decir, todo exceptuando salarios, impuestos e intereses, los primeros estudios señalaban que tanto en empresas industriales como comerciales este volumen (incluyendo inversiones) representa entre un 50% y un 65% de las ventas de la empresa en promedio. Un 65% del volumen de ventas es una cantidad de recursos muy importante que requiere una gestión profesional. No puede dejarse en manos de los diferentes departamentos y con criterios de gestión o no gestión heterogéneos.

Una vez logrado este perímetro y aplicando la gestión estratégica de categorías, herramientas de análisis de valor, gestión de proveedores estratégicos e incluso la integración del proveedor en el diseño preliminar del producto se lograron grandes avances en los ahorros. El estudio de AT Kearney en 2004 sobre 395 empresas en las que las llamadas líderes aplicaron la Gestión Estratégica de Compras y los seguidores no, o al menos no en todo su potencial, muestra la diferencia de ahorro entre ambos tipos de gestión:

Figura 3: Ahorros Bechmark 2004

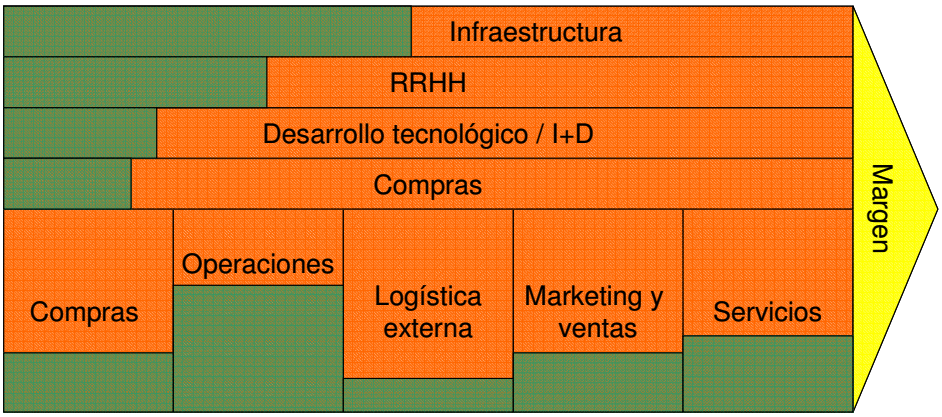


Fuente: AT Kearney 2004

Para lograr estos ahorros lógicamente deben emplearse recursos en el departamento. Lo curioso fue que los departamentos líderes representaron en promedio un gasto del 0,68% de las ventas y en las seguidoras un 0,85%.

A pesar que el ahorro en costes debe justificar de *per se* la existencia del la función de Compras analizando la cadena de valor se puede observar que el efecto de la misma puede ir más allá que el simple ahorro (Porter 1985). La siguiente figura muestra el impacto que los aprovisionamientos tienen en cada una de las actividades en la cadena de valor:

Figura 4: Impacto de Compras en la cadena de valor



Fuente: Elaboración propia a partir de la Cadena de Valor de Porter

La parte verde de cada actividad indica el impacto o la influencia que compras o los aprovisionamientos tienen en cada una de las actividades. Es lógico pensar que las operaciones o la logística o contar con la infraestructura apropiada se debe, en gran parte, a los aprovisionamientos y compras externas que se realizan.

Esta importancia de Compras sobre las actividades de valor de la empresa dio como resultado implicar la función en el logro y aseguramiento de:

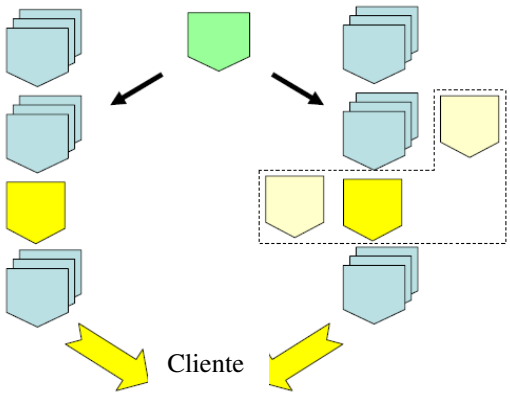
- Calidad cuantitativa y cualitativa de las adquisiciones
- Competitividad a través de reducciones de "*ownership cost*", no reducciones de precios
- Servicio interno

Este proceso de ampliación de las responsabilidades de Compras fue acompañado de cambios organizativos. Paulatinamente en las diferentes industrias los departamentos dejaron de depender funcionalmente de otros departamentos como el financiero o el de operaciones para pasar a depender de la dirección general. Por supuesto esta evolución no fue igual en todas las industrias.

4. ¿QUÉ MÁS SE ESPERA DE COMPRAS?

Captar el valor externo a la empresa a través de la gestión de la empresa extendida. Actualmente la competencia ya no se libra entre empresas que intentar suministrar valor a los clientes. Actualmente la competencia es entre cadenas de suministro.

Figura 5: Competencia entre cadenas de valor



Fuente: Elaboración propia

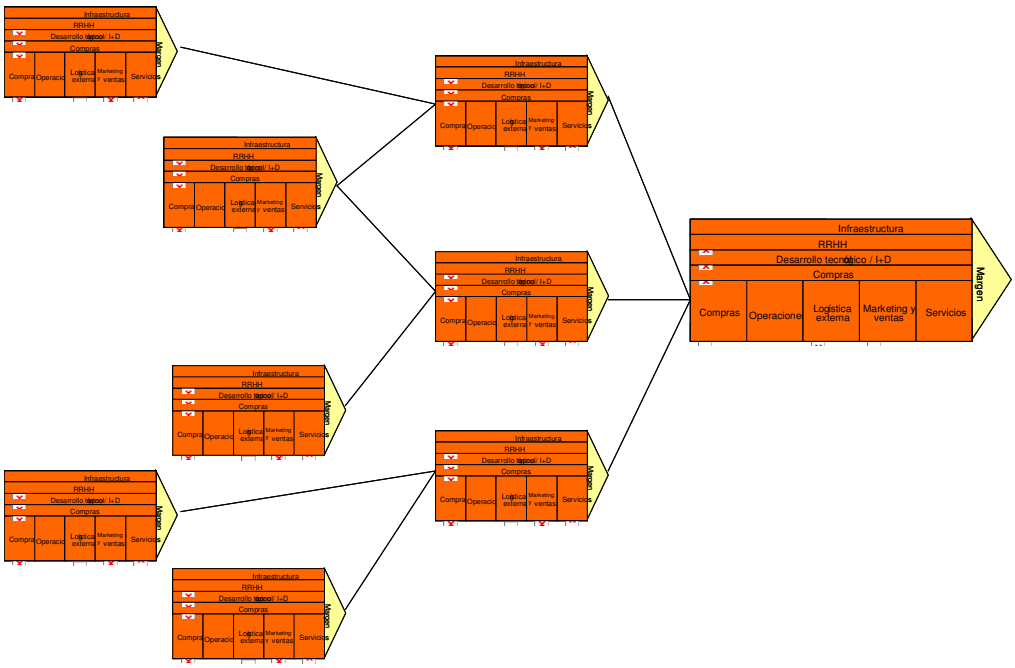
La innovación, el *"time to market"* y el *"speed to market"* no dependen en exclusiva de la capacidad de la empresa. La colaboración y el valor que puede obtener de sus proveedores es esencial para lograrlas. Los modelos *"open innovation"* (Chesbrough 2006) así lo atestiguan. Una empresa debe gestionar a sus proveedores, al menos a los estratégicos, sólo de este modo las compras pueden alinearse con la estrategia de la empresa y pueden aportar valor a las diferentes actividades.

Este concepto de "empresa extendida" es el futuro de Compras. A través de las diferentes encuestas realizadas a empresas con las "best practices" en Compras por parte de *CAPS (2006) y Future Purchasing Alliance (2006)* los retos que debían afrontar los departamentos de compras eran:

- Alinear las actuaciones de Compras con la estrategia de la empresa
- Compras debe basar su gestión en captar valor de la red de proveedores
- Liderar proyectos cross-funcionales para lograr ahorros
- A los anteriores retos cabría añadir uno nuevo detectado en el estudio de Aberdeen (2008) sobre competitividad y Compras. Para lograr competitividad una de las vías de actuación de numerosas empresas es *"do the best and outsource the rest"* (haz aquello que sepas hacer y subcontrata el resto). En este reto y en esta decisión sobre subcontratación es Compras quien toma un papel protagonista en cuanto es conocedora del mercado de proveedores y puede aportar gran valor al proceso de decisión.

En este largo proceso de maduración de la función de compras se ha pasado de la obtención de disminuciones de precios al aporte de valor. Este cambio ha implicado cambios organizativos enfocados, fundamentalmente, al SRM (supplier relationship management). La gestión de la relación con el proveedor, ha sido quizás el mayor y más notable de los cambios en la Gestión Estratégica de Compras. En este momento el Director de Compras ha de actuar y gestionar

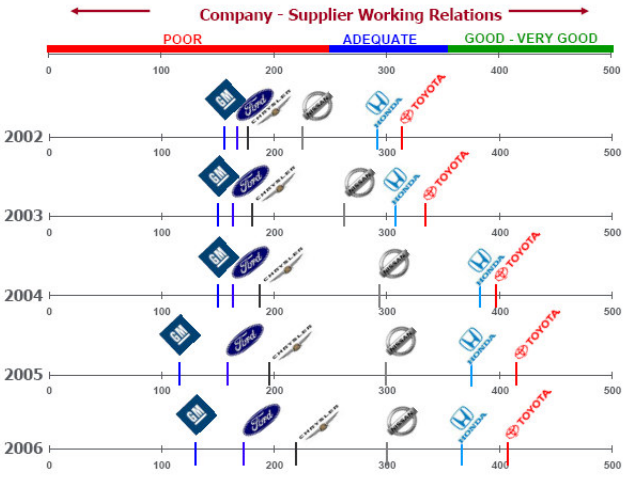
como un supuesto Director General de todos los proveedores, es el nexo de unión entre la cadena de valor de la empresa y el valor –que es mucho- que existe en la red de proveedores y está esperando ser captado.



Fuente: Elaboración propia

Las industrias más avanzadas cuentan con medidas de estas relaciones con el proveedor. Un ejemplo puede ilustrar este hecho. De todos es sabido que la industria automovilística norteamericana no pasa por buenos momentos, mientras que la japonesa está alcanzando éxitos hasta hace poco impensables. La respuesta de la industria norteamericana fue, en parte, presionar los costes de los proveedores, con lo que consiguió ciertos ahorros. Sin embargo los japoneses comenzaron una relación a largo plazo que se basó en la participación del proveedor desde el diseño del producto hasta su puesta en el mercado, lo cual dio como resultado ahorros mutuos, innovación conjunta y mejoras en los procesos. Esta colaboración se ve plasmada en la encuesta que la *North American Automotive Survey* realiza anualmente. Encuestando a los proveedores de la industria automovilística sobre la confianza y relación existente con los fabricantes se obtuvieron los siguientes resultados:

Figura 7: Grado de confianza relación clientes –proveedores industria del automóvil



Fuente: North American Automotive Survey

A juzgar por los resultados obtenidos, y en el supuesto que un proveedor tenga disponible alguna innovación, proceso o producto que pueda ser diferencial para el fabricante ¿a cuál se lo ofrecerá primero?; Parece lógico pensar que será a aquellos con los que tenga una buena relación.

5. MODELOS DE EVOLUCIÓN EN COMPRAS

La transformación desde un departamento operacional a un departamento estratégico ha sido estudiada durante las dos últimas décadas por teóricos y académicos. Existen varios modelos teóricos que muestran una serie de fases en la evolución de la función de Compras:

Figura 8: Modelos de evolución de Compras

Autor	Etapa 1	Estapa 2	Etapa 3	Etapa 4	Etapa 5
Reck & Long(1988)	Pasivo	Independiente	Soporte	Integrada	--
Syson (1989)	Transaccional	Comercial	Estratégica	--	--
Bhote (1989)	Confrotación	Negociación	Objetivos comunes	Partenariado	--
Cavinato (1990)	Bajos precios	Compras	Aprovisionamiento	Cadena suministro	Network
Cammish & Keough (1991)	Servir fábrica	Bajo coste unitario	Coordinación	Estratégico	--
Van Weele (1992)	Operacional	Comercial	Logística	Estratégica	---
Burt (1993)	Reactiva	Mecánica	Proactiva	---	---
Keough (1993)	Servir fábrica	Bajo coste unitario	Coordinada	---	---
AT Kearney (1994)	Transaccional	Control	Excelencia operativa	---	---
Monckza (1994)	Servir fábrica	Precio	Consolidación	---	---
Chadwick (1995)	Transaccional	Comercial	Soporte	---	---
Barry & Calvinato (1996)	Básico	Servicio mejorado	World class	---	---

Fuente: Weele, 2000 y elaboración propia

El modo de crearlos ha sido midiendo el grado de evolución de una serie limitada de parámetros, que, en la mayoría de de los casos, son comunes entre los modelos. Este hecho ha provocado que, a pesar del diferente número de etapas, las diferencias entre ellos sean más de forma que de fondo. Los parámetros empleados han sido:

- Estratégicos: En todos los autores se comprueba como el diseño, desarrollo, e implementación de una estrategia para compras y que esté vinculada a la estrategia de la empresa es un dato básico a la hora de medir el grado de maduración de la función. Obviamente a mayor formalización estratégica y mayor implicación en la estrategia de la empresa existirá un mayor grado de evolución de la función.
- Gestión de los proveedores: El modo de gestionar el panel de proveedores ha sido otro de los puntos estudiados en los modelos, visto desde un punto de vista de SRM (*supplier relationship management*).
- Actividades de Compras: El grado de implicación del departamento en funciones que originalmente no pertenecen a éste, como participación en el desarrollo de nuevos productos, gestión de contratos, evaluación y homologación de proveedores, o participación en decisiones sobre subcontratación muestra el tipo de actividad –más operacional o estratégica- que desarrolla la función.
- Apoyo de la dirección general: Mide el grado en el cual la dirección general implica a Compras en la formulación estratégica y, a su vez, como se implica en las actividades de Compras.
- Liderazgo: Tiene en cuenta el tipo de gestión interna que se aplica en el departamento.
- Recursos humanos: Es el grado de formación, desarrollo de carrera y habilidades que tienen los compradores
- *Reporting:* Este parámetro tiene en cuenta que consecución de objetivos mide el departamento –ahorros, calidad, innovación, fiabilidad de los proveedores- y a quien reporta estas mediciones.

Teniendo en cuenta el grado de desarrollo logrado en los anteriores parámetros en las diferentes industrias puede establecerse un ranking desde las menos evolucionadas en su función de compras a las más evolucionadas.

Figura 9: Grado de evolución en Compras por tipo de industria



Fuente: AT Kearney y Jordi Civit 2004

6. DESCRIPCIÓN DE LA EVOLUCIÓN

Considerando todos los puntos comunes de los modelos pueden establecerse de modo aproximado las características de cada una de las etapas. Suponiendo seis etapas, que coinciden con partes o totalmente con algunos modelos, las características de la función de Compras en cada una de ellas serían las siguientes:

6.1 ETAPA 1: GESTIÓN DE PEDIDOS

En esta fase de la evolución la función básica de Compras es encontrar suficientes proveedores para que la fábrica no pare su actividad. No existe una estrategia explícitamente formulada en Compras, los objetivos son intuitivos y el departamento depende por lo general de otro departamento como logística o producción. No se controlan todas las Compras que realiza la compañía, estando las indirectas en manos de los usuarios. Los sistemas de información están orientados a tareas administrativas, la formación del personal es baja y sus tareas son totalmente operacionales. La cultura de trabajo es reactiva ante los problemas y se gestionan las quejas que puedan tener los usuarios.

6.2 ETAPA 2: BAJOS PRECIOS UNITARIOS

En esta etapa los compradores están orientados a negociar bajos precios con los proveedores. Las estrategias de Compras están enfocadas a lograr los precios más bajos posibles de los proveedores, la independencia funcional es algo mayor que en la etapa anterior aunque se sigue rindiendo cuentas a otro departamento y no a la dirección general. La cultura de trabajo se basa en especializar a los compradores entorno a grupos de productos y en entablar duras negociaciones con los proveedores. Los éxitos se miden en ahorros conseguidos respecto al año anterior y en la fiabilidad de entrega del proveedor.

6.3 ETAPA 3: COORDINACIÓN DE LA FUNCIÓN

Aparece normalmente una centralización de la función y se buscan sinergias en los volúmenes de compras. Asimismo se crea una cierta coordinación en caso de existir múltiples centros y se establecen sistemas y políticas de compras. Con ellas aparece una incipiente formulación estratégica aunque no siempre se formaliza su proceso. Aparte de los costes se buscan acuerdos y beneficios globales para la empresa y se empieza a reconocer el papel de

compras en la organización por parte de la dirección general y el resto de la empresa. La atención al cliente interno empieza a ser una realidad. Todavía la organización de compras está focalizada sobre el producto aunque se introduce la gestión por categorías. Los sistemas informáticos incluyen la conexión y el compartir información entre diferentes unidades de negocio. Las habilidades de los compradores van más allá de la pura negociación y se inicia la formación en habilidades analíticas, gestión de la calidad total y comunicación.

La centralización de las compras, la búsqueda de sinergias entre diferentes centros y la incipiente gestión del panel de proveedores es el resultado final de esta etapa.

6.4 ETAPA 4: INTEGRACIÓN INTERNA CROSS-FUNCIONAL

En esta etapa la visión cross-funcional de Compras se pone e manifiesto. No se buscan reducciones de los componentes comprados si no la reducción de los costes totales del sistema. Para ello se acude a los proveedores clave no en términos de confrontación si no de asociación. Compras se organiza alrededor del servicio a prestar al cliente internos, orientada a los procesos no a los productos. Las compras indirectas que estaban en manos de los usuarios pasan a estar en manos de compras. La dirección general empieza a involucrar a compras en decisiones estratégicas especialmente de subcontratación. Los sistemas informáticos están integrados aunque no con los proveedores. Los equipos cross-funcionales y la medición de la satisfacción del cliente interno son dos características que aparecen en esta etapa. Las habilidades de los compradores pasan por crear y gestionar equipos de trabajo, comunicar al cliente interno y la los proveedores clave, participar en la formulación de las especificaciones de los productos y tener visión de negocio.

6.5 ETAPA 5: SUPPLY CHAIN MANAGEMENT, INTEGRACIÓN EXTERNA

En esta etapa Compras se hace responsable del la gestión del *outsourcing* de la empresa y además participa en la creación de nuevos productos junto a proveedores clave. Las compras indirectas –o de no producción como inversiones o publicidad- son gestionadas totalmente por Compras. Los sistemas informáticos están conectados a los de los principales proveedores. Los usuarios internos hacen los pedidos a través de herramientas de *e-procurement* o catálogos, especialmente en productos indirectos, ya que Compras invierta gran parte de su esfuerzo en facilitar los procesos a sus clientes internos. La gestión de los proveedores en esta etapa pasa a ser gestión de la cadena de suministro, involucrando a los proveedores en diferentes proyectos de innovación, mejora de procesos, lanzamiento de nuevos productos etc. No sólo compran a sus proveedores, si no que junto a equipos multidisciplinares internos y externos deciden de modo consensuado diferentes aspectos de la compras. Esto exige un estilo de liderazgo participativo, basado en el *coaching* y soporte a los colaboradores. El conocimiento del *"total cost ownership",* la cadena de suministro así como las habilidades en gestión y liderazgo son claves en esta etapa. La dirección mide la compras por los resultados obtenidos en estos procesos, no ya por los ahorros o la calidad lograda de los proveedores.

6.6 ETAPA 6: ORIENTACIÓN A LA CADENA DE VALOR

La última etapa se caracteriza en que Compras se orienta a entregar valor al cliente final. Para satisfacer las necesidades de los clientes finales Compras se sirve de todo el valor existente en la red de proveedores atrayéndolo mediante "reverse marketing" o marketing enfocado a los proveedores. En esta fase Compras debe hacer que sus proveedores estratégicos la vean como el "cliente favorito". Compras ha pasado de comprar a formar parte de la cadena de valor. Algunas de las herramientas utilizadas aquí son el análisis de valor y el análisis de especificaciones. La estrategia de Compras está completamente integrada en la de la empresa y participa activamente en ella. La cultura es emprendedora y compras se enfoca tanto "*aguas arriba*" como "*aguas abajo*". Gestiona y traslada valor desde los proveedores hasta el cliente final.

7. CONCLUSIÓN

Son muchos los retos que en una economía globalizada afrontan las empresas. Todos ellos se resuelven con competitividad, y, para lograrla, la empresa debe poder disponer y utilizar todos los recursos a su alcance. Compras es uno de ellos, y tal como se desprende de las encuestas de AT Kearney (2004), Ubeda, R. (2007), Future Purchasing Alliance (2006), el grado de evolución de las compras en España tiene todavía un gran potencial de mejora.
El artículo resume el camino teórico a través del cual compras puede ayudar al crecimiento sostenible a largo plazo de la empresa, a su posicionamiento estratégico y su competitividad, empleando compras para competir no tan sólo para aprovisionar.

BIBLIOGRAFÍA:

ABERDEEN GROUP, (2008) "The Purchasing Benchmark Report: Best Tactics to Increase Growth", *Aberdeen Group*

CAPS (2006) "ISM Services Group Purchasing Performance" CAPS, *Institute for supply chain management, School of Business Arizona State University*

CARR, A. PEARSON J. (2002) "The impact of Purchasing and supplier involvement on strategic purchasing and its impact on firm's performance" *International Journal of Operations & Production Management Vol. 22, nº9, págs. 1032.1053*

CARTER, J., NARASHIMHAN, R. (1996) "Is Purchasing Really Strategic?" *Journal of Supply Chain Management. Vol 32, nº 1, págs 20-28*

CAVINATO (1991) "Evolving procurement organizations: logistics implications" *Journal of Business logistics Vol 13, nº 1, págs 27-44*

CHAPMAN, T.L (1997) "Purchasing: No time for lone rangers", *The McKinsey Quarterly, number 2, págs 30-40*

CHESBROUGH, H. (2006) "Open Innovation"*Harvard Business Press, Boston MA*

CIVIT, J. (2004) "Proceso de transformación de la función de compras" *Compras y Existencias nº 133 Julio -agosto 2004*

CORPORATE AUTHORSHIP TCB (2007)"CEO Challenges 2007 Top 10 Challenges" *Referencia :R-1406-07-RR*

DYKEMA (2008) "Global automotive outlook survey" *Dykema Group*

FREEMAN, V.T., CAVINATO J.L (1990) "Fitting Purchasing to the strategic firm: framework, processes and values" *Journal of purchasing and materials management Winter, p.6-10*

FUTURE PURCHASING ALLIANCE (2004) "Benchmark 2004"*The Future Purchasing Research Foundation*

FUTURE PURCHASING RESEARCH FOUNDATION (2006) "Strategies for Purchasing transformation" *The Future Purchasing Research Foundation*

HEBERLING, M.E. (1993) "Reegineering the corporation" *International Journal of Purchasing and Materials Management, Fall, págs 48-52*

MORGAN - MCKINSEY QUARTERLY, (1987) " The purchasing revolution", *The McKinsey Quarterly, pp.49-55*

PORTER, M. (1985)"Competitive Advantage" Ch.1, págs 11-15, *The Free Press. NewYork*

UBEDA, R. (2007) "Benchmark estado de Compras en España", *Estudio no publicado*

VAN WEELE, A.J. and ROZEMEIJER, F. (1996) "Revolution in Purchasing: Building Competitive Power Through Pro-Active Purchasing" *European Journal of Purchasing & Supply Management, Vol.2, No.4, pp. 153-160*

ÁREA XVI

ESTRATEGIA ECONÓMICA

ESTRATÉGIA ECONÓMICA

ECONOMIA INFORMAL NO COMÉRCIO, ALOJAMENTO E RESTAURAÇÃO: MANIFESTAÇÕES E SOLUÇÕES

FRANCISCA GUEDES DE OLIVEIRA
foliveira@porto.ucp.pt
SOFIA SILVA
srsilva@porto.ucp.pt
VASCO RODRIGUES
vrodrigues@porto.ucp.pt
CEGEA – Centro de Estudos em Gestão e Economia Aplicada
Faculdade de Economia e Gestão
Universidade Católica Portuguesa

Resumo

O presente estudo pretende contribuir para uma análise dos mecanismos que incentivam a existência de economia informal no sector do comércio e restauração, contribuindo para a redução do fenómeno. Na primeira secção apresenta-se uma revisão de literatura sobre o tema com vista a delimitar-se o âmbito do conceito utilizado na posterior análise empírica. Na segunda secção faz-se uma breve contextualização dos sectores partindo-se na terceira secção para um levantamento do fenómeno da informalidade. Pela natureza do fenómeno a sua valorização é de extrema dificuldade. Optamos por entrevistar alguns actores relevantes e fazer um levantamento das sensibilidades dos agentes no terreno. O cruzamento dos resultados destas conversas com alguns dados existentes permitiram traçar um cenário provável da importância da economia informal no sector. O último capítulo consiste numa tentativa de levantamento de algumas propostas concretas para a diminuição do fenómeno. Estas propostas partem dos resultados das entrevistas, tendo como suporte a literatura existente.

Palavras Chave: Economia informal, comércio e restauração.
Área Temática: Estratégia Economica.

Abstract

The present study intends to contribute for the analysis of the mechanisms that stimulate the existence of informal economy in the sectors of commerce and restaurants, contributing for the reduction of the phenomenon. In the first section a literature revision is presented on the subject with sight to delimit the scope of the concept used in the subsequent empirical analysis. In the second section we present briefly the sectors and try, in section three, to infer about the phenomenon of informality. Due to the nature of the phenomenon its valuation is extremely difficult. We chose to interview some relevant agents to get a better grip of reality. The crossing of the results of these interviews with some existing data allowed tracing a probable scenario of the importance of informality in the sector. The last chapter consists of an attempt to assess some concrete proposals for the reduction of the phenomenon. These proposals depart from the results of the interviews, having as support the existing literature.

Keywords: Informal economy, commerce and restaurants.
Thematic Area: Economic Strategy.

1. O QUE É A ECONOMIA INFORMAL?

Este estudo deverá contribuir para analisar os mecanismos que incentivam a existência da economia informal, "sendo finalidade última da análise contribuir para a redução do fenómeno." Neste sentido, torna-se necessário clarificar o que é, para este propósito, a economia informal.

O único consenso que, nesta matéria, se encontra na literatura económica é o de que o tema é complexo: há múltiplas definições do que seja a economia informal. A existência de uma panóplia de outras designações que parecem referir-se ao mesmo fenómeno ou a fenómenos com ele estreitamente relacionados só agrava essa complexidade: economia paralela, economia subterrânea, economia sombra, economia escondida, mercado negro, fuga ao fisco, evasão fiscal, são apenas alguns exemplos dessas designações. Como nota a OCDE (2002, pág. 11) não há qualquer consenso sobre se todas estas expressões se referem à mesma coisa, nem, caso assim não seja, sobre as relações existentes entre elas.

A informalidade a que alude a designação remete para formas de actividade económica desenvolvidas fora do normal quadro jurídico-institucional para o seu exercício. Mas porque este quadro é extremamente vasto, esta noção não é, em si mesma, suficiente para guiar um trabalho com preocupações operacionais. Alguns autores têm procurado chegar a definições mais precisas atendendo aos aspectos do quadro jurídico-institucional a que as actividades escapam.

A economia informal é, muitas vezes, entendida como sendo aquela parcela da actividade económica que não é captada na Contabilidade Nacional, nomeadamente nos valores publicados do Produto Interno Bruto (Simon Johnson et al., 1998, David Dreyer Lassen, 2006). Feige (1990) chama a esta parcela da actividade económica "economia não registada". Sendo ou não esse o propósito dos seus autores, as estimativas da importância da economia informal que surgem, por vezes, nos *media* são normalmente entendidas desta forma: como valores que deveriam somar-se ao valor oficial para obter o "verdadeiro" valor do PIB.

Outro entendimento frequente é o de que a economia informal corresponde às actividades económicas em que não são cumpridas as obrigações fiscais, nomeadamente as obrigações de declaração. A esta parcela da actividade económica, Feige (1990) chama "economia não declarada". É igualmente frequente que as estimativas sobre a importância da economia informal sejam entendidas como estimativas da "economia não declarada".

Antes de avançar para outros entendimentos possíveis do conceito de economia informal importa, desde já, notar que a relação entre a economia não declarada e a economia não registada pode, pelo menos teoricamente, ser muito ténue (Vito Tanzi, 1999).

Desde logo porque, na construção das Contas Nacionais, os Institutos de Estatística utilizam uma multiplicidade de procedimentos destinados a quantificar a actividade económica que não é declarada pelos que a levam a cabo, como é analisado por Santos (1996) a propósito do Instituto Nacional de Estatística português. Existe até um manual da OCDE (OCDE, 2002) que propõem boas práticas na medição desta "economia não observada". Os números oficiais do PIB incluem, portanto, já a estimativa que as autoridades estatísticas fazem da actividade produtiva que não conseguem observar directamente, seja porque os que a levam a cabo a procuram ocultar do fisco, seja por outras razões.[1] A economia não declarada não é, portanto, necessariamente não registada.

Por esta mesma razão, as estimativas sobre a importância da economia informal não devem, em geral, ao contrário do que é prática corrente, ser entendidas como algo que acresce necessariamente às estatísticas oficiais. Só assim se deveria proceder se o estudo em causa estivesse a estimar exclusivamente actividades não observadas para além daquelas que as

[1] De acordo com o manual do OCDE, a não-observação pode resultar de cinco fenómenos diferentes. A "produção subterrânea" que se refere a actividades produtivas legais mas que são deliberadamente escondidas do Estado, para evitar o pagamento de impostos e contribuições sociais, para evitar o cumprimento de certos normativos legais (salário mínimo, horários de trabalho, normas de segurança, etc) ou para evitar o custo de certos procedimentos administrativos. A "produção ilegal" refere-se às actividades produtivas que são ilegais pela própria natureza do que é produzido (ex.: comércio de substâncias ilícitas) ou por serem realizadas por quem não está autorizado a fazê-lo (ex.: exercício da medicina por quem não tem as credenciais necessárias). A "produção do sector informal" refere-se a actividades produtivas desenvolvidas para dar emprego e rendimento a quem as desenvolve; esta produção do sector informal não incide sobre actividades que, em si mesmas, sejam ilegais e distingue-se da "produção subterrânea" por não haver um propósito deliberado de ocultação face ao Estado, sendo esta mero resultado do reduzido nível de organização da actividade. A "produção própria das famílias" refere-se a actividades desenvolvidas no seio da família para usufruto próprio (ex.: pequenas produções agrícolas e pecuárias, construção da casa própria, etc). Finalmente, alguma produção não é observada devido a deficiências dos procedimentos de recolha de informação estatística.

autoridades estatísticas já incluem nos números oficiais. Mas as estimativas publicadas não explicitam, em regra, se assim acontece.

Outra razão para que não haja relação estreita entre a economia não registada e a economia não declarada prende-se com os diferentes objectivos da Contabilidade Nacional e do Fisco. Na contabilização do PIB, e de outras grandezas semelhantes, as autoridades estatísticas estão a procurar medir a actividade produtiva do país. As actividades económicas de carácter redistributivo, nomeadamente as que envolvem aplicações de capitais, legais ou ilegais, não são, nem devem ser, tidas em conta no PIB, porque não correspondem a actividades produtivas. Logo, não integram a economia não registada. Mas podem gerar obrigações fiscais que, quando não cumpridas, as configurem como economia não declarada. Por outro lado, podem existir actividades produtivas não registadas que, por exemplo, por não atingirem certos limiares que sejam condição da obrigação de declaração ao fisco, não constituam economia não declarada.

Um terceiro entendimento de economia informal é o de que a informalidade decorre da ilegalidade das actividades económicas desenvolvidas: é a "economia ilegal", de acordo com a taxinomia de Feige (1990). Em princípio, o não cumprimento das obrigações fiscais configura, em si mesmo, uma ilegalidade, pelo que a economia não declarada seria parte da economia ilegal. No entanto, os autores que utilizam o conceito de "economia ilegal" têm, em geral, em mente a ilegalidade que decorre da natureza da actividade desenvolvida e não a que resulta do mero incumprimento as obrigações fiscais: estão a pensar em actividades como o tráfico de droga, o roubo, etc. Dependendo do ordenamento jurídico, estas actividades podem, ou não, gerar obrigações fiscais. Se essa obrigação existir, a economia ilegal é parte da economia não declarada. Mas se a obrigação não existir, não há qualquer sobreposição entre os dois conceitos.

A relação entre a economia ilegal e a economia não registada é também pouco clara. Novamente, na contabilização do PIB, as autoridades estatísticas procuram medir a actividade produtiva. Na medida em que sejam capazes de as elaborar, as autoridades estatísticas devem incluir nos valores oficiais do PIB as suas estimativas das actividades produtivas ilegais, como o tráfico de droga, a prática ilegal de actos médicos, etc. Mas não devem incluir as actividades ilegais não produtivas, como o roubo ou a obtenção fraudulenta de apoios públicos. No entanto, certos trabalhos sobre a economia informal dão grande destaque a este último tópico: por exemplo, no relatório sobre a economia informal que preparou para o Tesouro britânico, Lord Grabiner (2000) dá grande atenção a medidas que possam reduzir a fraude na obtenção indevida de benefícios públicos.

Como resulta do anteriormente exposto, parte da economia ilegal deve ser incluída na contabilização do PIB. A capacidade das instituições estatísticas para estimar o valor da produção ilegal é, no entanto, limitada, não sendo claro em que medida conseguem cumprir aquele objectivo, e varia de país para país. A interpretação e a comparação internacional de estimativas sobre a importância da economia ilegal e o seu peso no PIB são também muito dificultadas pelo facto de a noção de "ilegal" variar no espaço e no tempo. Por exemplo, mesmo no espaço da União Europeia, a prostituição ou o comércio de drogas leves são ilegais nuns países e legais noutros. E o estatuto legal de certas actividades altera-se ao longo do tempo, como recentemente aconteceu com a interrupção voluntária da gravidez em Portugal.

Depois de definir economia não registada, economia não declarada e economia ilegal, Feige (1990) apresenta ainda uma definição de economia informal, *strictu senso*: "(…) a economia informal compreende aquelas actividades económicas que evitam os custos e são excluídas dos benefícios e direitos inerentes às leis e regras administrativas que regulam as relações de propriedade, o licenciamento comercial, os contratos de trabalho, a responsabilidade civil, o crédito e os sistemas de segurança social." O autor não esclarece porém como é que esta última definição se relaciona com as anteriores, embora pareça que a economia informal, assim entendida, engloba a economia não declarada e a economia ilegal mas não necessariamente a economia não registada.

O conceito de economia informal é muito utilizado no âmbito da análise de questões de índole laboral. Em 1993, a Organização Internacional do Trabalho adoptou uma definição de economia informal, baseada na empresa, segundo a qual esta é constituída por todas as empresas não registadas abaixo de um determinado limiar de dimensão (a definir em cada país). Mais recentemente, em 2003, a OIT adoptou igualmente uma definição de "emprego informal",

baseada na pessoa que desenvolve o trabalho. Sinteticamente, esta definição considera informal todo o emprego na economia informal (de acordo com a definição anterior) bem como o emprego na economia formal que escapa à legislação de trabalho, fiscal e de protecção social (Ralf Hussmanns, 2005).

Esta última definição remete para o facto de que as fronteiras entre a economia formal e informal são fluidas. Parte substancial da economia informal é desenvolvida por empresas formalmente constituídas e que cumprem muitas, mas não todas, as suas obrigações legais: os mesmos agentes económicos estão, frequentemente, presentes na economia formal e na economia informal.

Em síntese, pode afirmar-se, como faz Schneider (2000) na recensão que publicou no *Journal of Economic Literature*, onde utiliza a designação economia sombra, que "(…) uma definição precisa parece bastante difícil, se não impossível (…)". A noção de economia informal pode ser entendida de múltiplas formas que têm implicações em termos da compreensão dos mecanismos que a viabilizam e da análise de possíveis medidas de política que visem reduzir a sua incidência.

Uma possibilidade, provavelmente não a mais relevante para o presente estudo, é entender a economia informal como aquela que não é registada nas contas nacionais e procurar identificar medidas que permitam que essas contas traduzam mais rigorosamente a actividade produtiva nacional.

Outra possibilidade é centrar a atenção no cumprimento ou incumprimento de obrigações fiscais e procurar identificar medidas que reduzam o incumprimento. Nesse caso, há ainda que decidir se pretende atender apenas ao incumprimento em actividades produtivas ou se se pretende igualmente ter em conta o incumprimento em actividades redistributivas. Evidentemente, a redução do incumprimento pode passar pelo desaparecimento das actividades incumpridoras ou, alternativamente, por levar a que elas passem a cumprir.

É igualmente possível focar a atenção nas actividades económicas ilegais, caso em que novamente haverá que decidir se se pretende, ou não, abranger actividades produtivas e redistributivas. Se a opção for esta, as medidas a estudar poderão visar reduzir determinadas actividades ou atraí-las para a legalidade.

Finalmente, é possível colocar a ênfase nas questões laborais, tendo eventualmente como preocupação fundamental a identificação de medidas que possam contribuir para melhorar o grau de protecção de que beneficiam aqueles que se encontram em situações de emprego informal. Neste caso, é importante decidir se se pretende incidir igualmente sobre emprego informal no seio de empresas formalmente constituídas e das que não o são.

De um modo geral tentamos utilizar um conceito de informalidade abrangente. Nas conversas tidas deixamos livre, ao critério do interlocutor, o conceito a ser utilizado no diagnóstico levantado. Posteriormente fizemos uma filtragem de alguns dos aspectos apresentados para garantir que os problemas levantados na secção 3 cairiam em alguma forma de informalidade tal como esta é balizada na literatura.

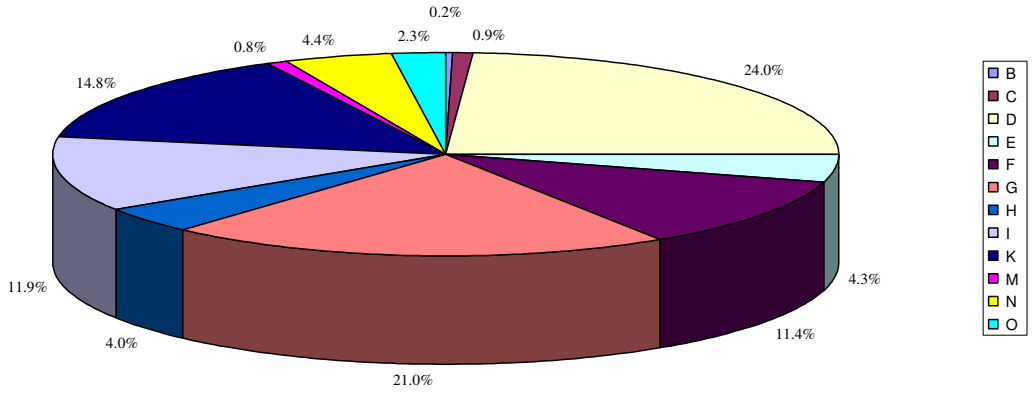
2. IMPORTÂNCIA DO COMÉRCIO E DO ALOJAMENTO E RESTAURAÇÃO

Nesta secção faremos uma breve apresentação dos sectores em análise. O objectivo é o de apresentar alguns números chave que permitam, rapidamente, apresentar os sectores em questão como sectores relevantes para a economia nacional. É este relevo, juntamente com a importância que a informalidade tem nas actividades do comércio e do turismo, que justificam a sua escolha para o presente estudo.

O sector G da CAE Rev. 2.1[2] contribuía em 2006 com quase 21 % do valor acrescentado bruto nacional. Dentro deste sector os comércios (por grosso e a retalho) eram responsáveis por si só por cerca de 18% do VAB nacional.

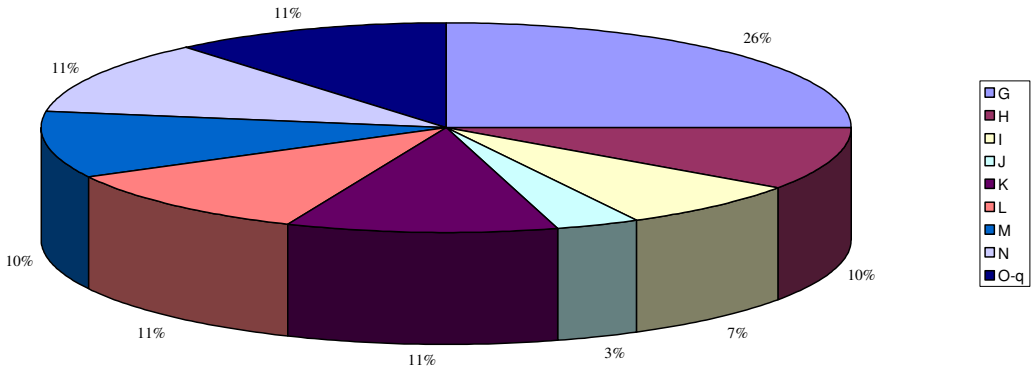
[2] A: Agricultura, produção animal, caça e silvicultura; B: Pesca; C: Indústrias extractivas; D: Indústrias transformadoras; E: Produção e distribuição de electricidade, gás e água; F: Construção; **G: Comércio por grosso e a retalho; reparação de veículos**

Gráfico 1 - Percentagem no VAB nacional dos sectores de actividade CAE Rev. 2.1 (2006)



O comércio absorvia, em 2007, ligeiramente mais do que 12% do emprego nacional. No total do sector terciário o ramo de actividade G correspondia a quase 26% do total do emprego, sendo que dentro deste ramo o comércio representa praticamente 84% da população empregada.

Gráfico 2 - População empregada no sector terciário segundo o ramo de actividade económica (2007)



Nesse ano a densidade de estabelecimentos comerciais por 1000 habitantes era de aproximadamente 18. Actualmente a estrutura empresarial caracteriza-se essencialmente por um

automóveis, motociclos e de bem de uso pessoal doméstico; H: Alojamento e restauração; I: Transportes, armazenagem e comunicações; J: Actividade financeiras; K: Actividade imobiliárias, alugueres e serviços prestados às empresas; L: Administração pública, defesa e segurança social obrigatória; M: Educação; N: Saúde e acção social; O: Outras actividades de serviços colectivos, sociais e pessoais; P Actividades das famílias com empregados domésticos e actividades de produção das famílias para uso próprio; Q: Organismos internacionais e outras instituições extra-territoriais.

elevado peso do comércio tradicional de menor dimensão e por uma concentração muito forte dos grandes formatos num número reduzido de operadores.

Em 2006 o total de remunerações pagas em Portugal ascendeu a 35 984 285 milhares de euros. Destas 7 790 227 foram pagas no comércio. Com cerca de 731 milhares de pessoas ao serviço, o comércio oferece uma remuneração média de mais de 10651 €, ligeiramente superior aos 10076 € pagos no total nacional.

O turismo é histórica e culturalmente um sector de eleição no nosso país. Com cerca de 11 300 milhares de entradas de turistas em 2006 e com mais de 269 000 estabelecimentos de hotelaria, este sector caracteriza-se por uma diversidade grande de formatos e por uma presença consolidada em quase toda a extensão do território nacional. Em 2007 houve um acréscimo de entradas de 7.5% de visitantes não residentes e um aumento de 11.2% de turistas. Com um aumento de cerca de 9.6% nos proveitos do sector o alojamento e restauração mostra-se um dos remos de actividade mais dinâmicos da economia nacional. Este aumento dos proveitos deve-se por um lado a um aumento dos preços acima da média nacional (3,5% em Janeiro de 2008, 0,6 pontos percentuais acima da taxa de inflação no mesmo período) mas em grande parte ao aumento de actividade (com o número de dormidas a aumentarem cerca de 6% no último ano).

Em termos de Valor Acrescentado Bruto o sector do alojamento e restauração contribui com cerca de 4% (gráfico 1) e em termos de emprego dentro do terciário atinge a quota dos 10%. Estes sectores registam também níveis de produtividade acima da média nacional:

Tabela 1 - Produtividade €/horas - 2004

Nº de trabalhadores	20-99	Mais 100
Total	12.15	19.2
Comércio	15.36	19.2
Restauração e Alojamento	20.52	35.3

A importância destes sectores na economia nacional conjugada com o peso que a informalidade tem no seu funcionamento exige uma atenção mais detalhada às suas manifestações, causas e consequências e a que se sugira formas de debelar o problema.

3. A INFORMALIDADE NO SECTOR

3.1 Principais manifestações

Em linha com o ponto inicial deste texto, o relatório Mckinsey (2003) considera que a informalidade é um fenómeno multifacetado, distinguindo três aspectos diferentes:

- Informalidade associada aos impostos – correspondente à evasão a impostos como o IRC ou o IVA;
- Informalidade relacionada com o mercado de trabalho – evasão a obrigações sociais e ao pagamento do salário mínimo, normalmente resultando da utilização de trabalhadores não declarados ou da não declaração da totalidade das horas de trabalho;
- Informalidade em relação a obrigações relacionadas com o mercado ou o produto – fuga ao cumprimento de requisitos de qualidade dos produtos, de normas de segurança e ambientais ou desrespeito de direitos de propriedade.

Nesta secção tentaremos fazer um levantamento das principais manifestações da informalidade no sector do comércio, alojamento e restauração. Por ser dificilmente observável, dada a natureza ilícita ou sombra destas actividades paralelas, foi considerado que a única forma de se conseguir entender a realidade efectiva e de se poder fazer um diagnóstico estratégico

seria através de conversas abertas com alguns agentes económicos relevantes. O anonimato das conversas foi sempre assegurado pelo que faremos uma apresentação genérica das principais conclusões a que se pode chegar. Os exemplos apresentados são aqueles considerados relevantes dentro dos muitos que foram mencionados pelos actores institucionais ou individuais entrevistados. Optou-se por organizar este levantamento por tipos de informalidade, de acordo com o relatório McKinsey (2003), para ficar mais claro o cenário que se pretende traçar.

Informalidade associada aos impostos

A informalidade pode manifestar-se por uma evasão a impostos como o IVA, IRC e IRS (aqui intimamente associado com situações de emprego menos formais e que discutiremos posteriormente). O facto de muitos estabelecimentos não passarem voluntariamente facturas e de frequentemente os clientes não as solicitarem torna fácil o não pagamento do IVA e uma distorção nos lucros declarados. Aqui a informalidade é tipicamente parcial. Na restauração as situações de "caixa aberta" (Exemplo 1) e consequentemente a não declaração de muitas transacções, são conhecidas e referidas por muitos como um exemplo claro de informalidade.

Exemplo 1:
São conhecidos casos recentes em que a DGCI detectou empresas de software que permitiam a correcção do montante facturado eliminando do sistema parte da facturação registada durante o período de actividade

Informalidade associada ao mercado de trabalho:

O sector caracteriza-se por uma predominância de formatos pequenos. Quase 90% das empresas são pequenas ou médias empresas e muitas vezes com características muito familiares. Não é invulgar neste sector (sobretudo nos formatos mais pequenos) existirem situações laborais informais. Esta informalidade traduz-se por trabalho não declarado ou apenas parcialmente declarado. Temos por um lado, um número de horas declaradas que não corresponde ao horário efectivo (havendo relato de jornadas de trabalho que chegam a atingir as 16 horas diárias) e por outro recibos de vencimento que ficam aquém das remunerações reais pagas (Exemplo 2). Estas situações derivam muitas vezes de trabalhadores não legalizados (nesta situação encontram-se trabalhadores imigrantes, familiares, sazonais que não têm particular interesse em declarar as contribuições respectivas). O resultado deste cenário é uma sistemática evasão às contribuições para a segurança social, declarações incompletas de IRS e uma sistemática fragilização da posição dos trabalhadores.

Exemplo 2:
Note-se que este tipo de actuação leva a um ciclo vicioso de difícil controlo: para poder pagar "por fora" aos trabalhadores é necessário receber "por fora" dos clientes o que implica necessariamente uma fuga ao fisco. Para se fugir ao pagamento de certos impostos sem declarar prejuízos absurdos e que facilmente levantariam suspeitas, há a necessidade de controlar os custos. Para se conseguir este controlo de custos não se pode declara todas as remunerações pagas

Uma outra fonte de informalidade prende-se com o tipo de contrato de trabalho negociado no sector. O recurso abusivo (e à margem de lei) de contratos a termo e de trabalho temporário ajuda a fragilizar a posição dos trabalhadores e a reforçar o carácter informal/ilegal das condições laborais do sector (Exemplo 3).

Exemplo 3:
Campanha contra a Precariedade Laboral no Sector
O sindicato fez recentemente um levantamento sobre a precariedade no sector e verificou o crescimento do trabalho ilegal e clandestino e do trabalho não declarado no sector, em particular nos estabelecimentos de restauração e de bebidas.

(…)

O levantamento feito a 320 empresas que empregam 3 936 trabalhadores, demonstra que 1 298 (33%) dos trabalhadores da restauração e bebidas são clandestinos.

(…)

A maior incidência de trabalho ilegal verifica-se nos estabelecimentos fora dos centros comerciais e nas zonas do litoral havendo concelhos onde a situação é assustadora.

Em Matosinhos, do levantamento resulta:

- 16 Estabelecimentos;
- 176 Trabalhadores;
- 14 Estabelecimentos têm ao serviço trabalho ilegal e clandestino no total de 82 trabalhadores ilegais;
- 59% dos trabalhadores ilegais são imigrantes;
- O valor pago em média mensalmente extra-recibo é de 296.92€;
- Apenas uma empresa cumpre o Contrato Colectivo de Trabalho no que toca ao descanso semanal com horário que garante dois dias de folga;
- Apenas duas empresas têm serviços de medicina no trabalho;
- Apenas duas empresas pagam as diuturnidades previstas no código do trabalho;
- Nenhuma empresa paga trabalho suplementar mas o horário médio diário é de 10,54; horas e o horário médio semanal é de 63.23 horas, havendo trabalhadores que cumprem horários de 13 horas diárias e 78 semanais; [3](…)

Informalidade associada às obrigações em termos de produto/serviço fornecido e licenciamento:

O não cumprimento de requisitos mínimos de qualidade, de segurança, de higiene e até de normas ambientais é uma das formas típicas que a informalidade assume nestes sectores de actividade. O cumprimento deste tipo de normas levaria a um aumento dos custos de actividade o que, associado a uma débil fiscalização, torna frequente a não observância de determinadas regras. Esta situação está contudo em clara transformação com a actuação da ASAE. Criticada por muitos com sendo excessiva esta actuação tem contribuído para uma alteração de comportamentos que pode levar a uma maior prática das regras instituídas. Há vozes críticas que apontam o comportamento da ASAE (a par com outros organismos da administração pública cujo o objectivo primeiro é de inspeccionar) como meramente punitivo e conducente à cobrança de coimas que beneficiem financeiramente o Estado. Alguns sugerem mesmo que comportamentos excessivos são um incentivo à fuga para a informalidade e não um bom mecanismo de combate da mesma.

As demoras na obtenção de licenças, as exigências processuais (quer em termos burocráticos quer em termos de custos) e ainda os requisitos de qualidade, higiene e segurança, incentivam a opção de operar sem o devido licenciamento. Esta situação conjugada com a suspeita de práticas mais ou menos generalizadas de subornos e corrupção torna este aspecto um dos mais complexos e difíceis de contornar.

O licenciamento comercial, por exemplo, tem um prazo de caducidade de 2 anos, se o estabelecimento não iniciar a actividade. Mesmo tendo esta licença os estabelecimentos comerciais necessitam, para poderem funcionar na legalidade, de uma licença de utilização que é atribuída pela câmara municipal. A falta de articulação entre os vários organismos envolvidos neste processo faz com que muitas vezes não se consiga a licença de utilização no prazo necessário para garantir a validade da licença comercial. Por esse motivo os proprietários dos estabelecimentos iniciam actividade ainda antes de conseguirem a tal licença de utilização. A falta de interlocução entre as entidades envolvidas atira alguns agentes para a economia informal. A agravar esta situação temos o facto de o licenciamento comercial ser atribuído por fases, sendo que, nalguns casos, apenas existem duas fases por ano.

[3] **Fonte:** Revista trimestral da FESAHT, n 84; Out/Nov/Dez 2007

3.2 A opção pela informalidade: a lógica económica

A existência da Economia Informal reflecte um desajustamento entre os interesses colectivos da sociedade, tal como definidos pelo Estado, e os incentivos individuais. O Estado define um modo normal, "formal", de desenvolver a actividade económica. No entanto, determinados agentes económicos, pesando os benefícios e os custos inerentes a esse modo de exercício da actividade, optam por não o respeitar. Quanto menores os benefícios da formalidade e quanto maiores os seus custos, mais provável se torna que os agentes económicos optem por exercer a sua actividade de modo informal (Lord Grabiner, 2000).

Os benefícios de desenvolver a actividade de modo formal são de múltipla ordem. O primeiro de todos pode ser definido, de forma muito lata, como o direito à protecção da lei. Quem exerce a sua actividade de modo formal pode, por exemplo, recorrer aos tribunais para defender os seus direitos, seja em caso de desrespeito pelos contratos que celebrou, de desrespeito pelos seus direitos de propriedade, etc. Pode igualmente obter protecção policial ou recorrer ao apoio de outro tipo de autoridades.

Estes direitos não existem, ou existem em medida mais limitada, para quem exerce a actividade de modo informal. E, aliás, quem o faz, não só não tem os mesmos direitos à protecção da lei, como pelo contrário está sujeito a ser por ela perseguido. Quanto maior a probabilidade de ser penalizado por exercício informal da actividade e quanto maiores as sanções aplicáveis em caso de penalização, menor o incentivo à informalidade.

Do ponto de vista empresarial, um dos grandes benefícios de actuar de modo formal é o acesso a um conjunto de mecanismos que permitem sustentar e expandir o negócio. O mais importante destes mecanismos é o sistema de crédito mas determinados sistemas públicos de apoio e incentivo empresarial podem igualmente ser relevantes. Em actividades de grande dimensão, que na Economia Informal tendem a ser a excepção, o mercado de capitais é também importante.

Quem actua de modo informal tem, em princípio, maior dificuldade no acesso a estes mecanismos. De notar que o que é aqui relevante é a diferença de facilidade de acesso para quem actua de modo formal e informal: se o acesso é muito difícil mesmo para quem actua de modo formal, o incentivo à formalização da actividade é diminuído; o mesmo acontece, se o acesso for facilitado a quem actua de modo informal.

Para os trabalhadores, a formalização da actividade traz importantes benefícios. Por um lado, a sujeição a normas de segurança no trabalho, de salário mínimo, de horários, etc. Por outro, o acesso a mecanismos de protecção social, nomeadamente em caso de desemprego ou doença.

Mas se tem estes e outros benefícios, a formalização tem também custos. Desde logo, a formalização da actividade implica custos com o cumprimento de obrigações administrativas e burocráticas que são evitados por quem actua de modo informal: registos, declarações anuais, etc. Por outro lado, custos que são a contrapartida dos benefícios proporcionados aos trabalhadores: pagamento de salários mais elevados, pagamento de contribuições para a Segurança Social, custos com o cumprimento de normas de segurança, etc. Analogamente, custos com o cumprimento de outros tipos de normas, nomeadamente no âmbito da qualidade dos produtos, do ambiente, etc. E, evidentemente, o custo do pagamento dos impostos, um dos mais mencionados na literatura (Takashi Kadokura, 2007, Friedrich Schneider, 2005).

Além disso, a formalização da actividade dá-lhe uma visibilidade que, pelo menos em determinados contextos, a pode expor a um maior risco de "assédio" por parte de funcionários públicos corruptos.

A comparação entre estes custos e benefícios tenderá a ditar a opção entre os modos formal e informal de exercer a actividade. Vale a pena, no entanto, notar alguns aspectos adicionais:

- A opção não é dicotómica; muitas vezes, o que está em causa é que parcela da actividade exercer de forma formal e não optar entre exercê-la integralmente de modo formal ou de modo informal;
- Diferentes agentes económicos avaliam de forma diferente estes custos e benefícios pelo que, nas mesmas circunstâncias, uns preferem a formalidade e outros a informalidade;

- A informação de que os agentes económicos dispõem condiciona as suas opções; nomeadamente, nem todos têm o mesmo conhecimento dos benefícios da formalidade e dos mecanismos para a ela aceder (Lord Grabiner, 2000);
- Finalmente, a opção entre formalidade e informalidade só se coloca se houver possibilidade de exercer a actividade da mesma forma; mas, muitas vezes, a informalidade parece ser o resultado de ausência de oportunidades para o exercício formal da actividade.

No sector em análise e do ponto de vista das empresas, talvez a primeira causa da informalidade seja a facilidade com que esta surge (o não pedir facturas, o não declarar todas as transacções, etc.) associada à redução de custos que esta implica.

O facto de, frequentemente, se efectuar muitas transacções de baixo valor ajuda a que a informalidade surja sem grande dificuldade.

Do lado dos trabalhadores será de salientar que a fragilidade da posição laboral força a aceitação de situações menos licitas. Acresce que, foi-nos sistematicamente referido, que o vencimento pago fora do recibo chega a atingir mais de 50% do salário declarado. Este facto, num sector com uma distribuição salarial muito desigual e em que um conjunto muito alargado de trabalhadores recebe salários médios bastante abaixo da média nacional, assume uma importância acrescida. Mais uma vez a fragilidade da posição negocial do trabalhador associada a alguma dificuldade em ler as verdadeiras consequências do vencimento não declarado, torna este tipo de situação muito mais vulgar do que o que seria desejado.

Um outro factor que condiciona fortemente a formalidade no sector (em particular na restauração) tem a ver com um sector de agro-pecuária e de pesca ainda muito tradicional (Ver caixa). É necessário fazer-se uma análise custos benefícios para se perceber se de facto há interesse ou não em levar a formalidade a estes sectores.

Exemplo 4:

Não é invulgar os proprietários dos restaurantes mais pequenos se queixarem da necessidade de terem um "saco azul" para suportar determinados custos com os fornecedores. Não é comum pedir a uma peixeira que traz o peixe fresco da lota numa canasta para passar um factura. Ou pedir ao agricultor que tem umas terras e que consegue legumes de boa qualidade que passe a factura das compras feitas. Na melhor das possibilidades nunca mais venderiam para esse restaurante. Se não há hipótese de declara os custos na totalidade, declarar as receitas todas seria permitir que se cobrasse mais imposto do que o devido.

3.3 Principais consequências

Desde logo ao falarmos de evasão fiscal (IVA, IRC, e IRS) e de fuga às contribuições da segurança social temos como consequência óbvia um contributo negativo para o Orçamento Geral de Estado e para o Orçamento da Segurança Social.

Do ponto de vista dos trabalhadores o não declarar do vencimento real conduz a que a base de incidência para cálculo das prestações sociais (subsídio de desemprego, de doença ou reforma) se torne bastante reduzida, descendo por isso o valor das respectivas prestações. Todo este cenário associado com a precariedade e sazonalidade do emprego conduz a uma degradação das condições de vida dos trabalhadores com particular incidência nos trabalhadores imigrantes.

Do ponto de vista do licenciamento as dificuldades burocráticas e os longos períodos inerentes à maioria dos processos conduzem muitas vezes a situações de suborno e corrupção. Acresce a tudo isto que a demora na obtenção das licenças (em média nove meses) é extraordinariamente punitivo quando se pensa em termos de incrementos de produtividade.

A existência de um sector informal forte tem também consequências do ponto de vista da competitividade e da produtividade. A informalidade é uma das maiores barreiras ao aumento da produtividade em Portugal. É responsável por cerca de 28% da diferença de produtividade não estrutural entre o nosso país e as melhores práticas da Europa, Mckinsey (2003), e tem um peso ainda maior quando consideramos o sector dos não transaccionáveis (em particular no comércio e restauração).

Ainda de acordo com o relatório McKinsey (2003) no sector do retalho alimentar, a informalidade permitiria uma vantagem, em termos de margem de lucro, de cerca de 6 pontos percentuais para os operadores informais. A evasão fiscal e às contribuições sociais é a forma de informalidade mais referida, permitindo a manutenção no mercado de operadores ineficientes.

As empresas que recorrem à redução dos custos via informalidade podem obter vantagens de preços e margens permitindo reter uma quota de mercado muitas vezes superior à, de facto, merecida. A informalidade funciona assim como um veículo que permite a distorção das vantagens competitivas.

Uma outra consequência intrinsecamente ligada com as anteriores prende-se com o facto de em processos de avaliação das empresas a informalidade nas operações levar a uma sobrevalorização das mesmas. Assim o EBITA (earings before interest taxes and amortizations) de uma empresa que não declara todas as transacções feitas ou que não faz os devidos descontos para segurança social será certamente muito superior ao EBITA dessa mesma empresa se esta operasse de forma completamente formal. Os custos de formalização associados por exemplo a processos de aquisição deste tipo de empresas são muitas vezes elevadíssimos. Durante o processo de negociação para possível aquisição de uma empresa que tenha algum grau de informalidade a diferença de valores de base com que as empresas trabalham pode inviabilizar a conclusão do negócio.

4. COMBATE À INFORMALIDADE

Atendendo à análise, apoiada na experiência, dos agentes envolvidos e suportando-nos em experiências concretas, o presente estudo identifica 6 recomendações de actuação no combate à informalidade nos ramos de actividade do comércio, restauração e alojamento.

O IRC é um dos impostos onde a informalidade terá um impacto relevante. Na perspectiva do comércio, restauração e hotelaria o facto de as transacções serem muitas vezes de pequeno valor e repetidas várias vezes por dia torna relativamente fácil a não contabilização real do volume e valor das transacções realizadas. Este facto tem também consequências importantes em termos de IVA declarado.

Recomendação I: Em vez de tributar os lucros e cobrar IVA com base nas declarações dos estabelecimentos utilizar outro tipo de medidas (nomeadamente aquelas que são propostas para medir a informalidade de um modo geral). Em alguns países da Europa já se recorre sistematicamente a indicadores físicos do tipo: consumo de electricidade ou água; área do estabelecimento; número de mesas; etc. Estes indicadores poderiam substituir ou complementar as declarações tradicionais numa tentativa de aproximar o pagamento do imposto do seu valor real (ver caixa).

Exemplo 5:

Em alguns países Europeus os restaurantes podem optar por uma de duas formas de declarações fiscais: ou têm contabilidade organizada e pagam impostos com base nos valores reais com que operam (e nesse caso estão sujeitos a inspecções mais ou menos regulares e com penalizações muito graves) ou então pagam impostos com base num indicador compósito que entra em linha de conta com: localização do espaço; preço médio da refeição; numero médio de refeições servidas; etc.. Nesta segunda possibilidade os restaurantes não possuem sequer facturas. Este indicador é revisto de x em x tempo e os proprietários do restaurante têm sempre a possibilidade de mudar de regime em qualquer momento que queiram.

O IVA é, em última instância, sempre pago pelo consumidor final ao vendedor do bem. O cliente de um restaurante paga sempre o IVA da sua refeição assim como o hóspede de um hotel paga o IVA da sua estadia. O não pedir factura tem como consequência o IVA não ser devidamente entregue ao estado.

Recomendação II: Campanhas de sensibilização massivas dos consumidores para pedirem facturas. Estas campanhas deveriam passar por fazer o consumidor perceber que ele paga o IVA de qualquer das formas, este apenas não chega às mãos do Estado ou seja deixa de ser dinheiro de todos nós para passar a ser dinheiro de apenas alguns (Ver caixa); As campanhas de sensibilização existentes deveriam ser reforçadas e dever-se-ia tornar a dar pequenos benefícios para quem pedir factura (por exemplo algum tipo de benefício no IRS).

O não pagamento das contribuições da segurança social por parte das empresas tem como consequência o peso nos cofres do estado e o degradar das condições de vida dos trabalhadores. É por isso necessário envolver os próprios trabalhadores na busca de uma solução para este problema.

Recomendação III: Campanhas de sensibilização aos trabalhadores para exigirem os seus direitos. Estas campanhas deveriam passar claramente a mensagem de que o custo dos vencimentos não declarados é todo suportado pelo trabalhador. A entidade patronal apenas retira benefícios desta situação. Eventualmente dar uma amnistia na parte das contribuições que são devidas pelos trabalhadores como incentivo para estes regularizarem ou pelo menos denunciarem situações de fuga.

Talvez um dos principais obstáculos para a resolução de algumas das questões referidas anteriormente prende-se com o envolver como parte da solução os "faltosos". A auto denuncia poderia servir como meio de combater o problema se fossem dadas algumas garantias aos sujeitos intervenientes.

Recomendação IV: Plano de "recuperação" para as empresas e/ou trabalhadores que estiverem disponíveis para regularizar as situações (tipo "plano Mateus"). Perdoar o pagamento de juros ou de coimas para as empresas que estiverem disponíveis para fazer pagamentos de contribuições passadas. Garantir planos com pagamentos faseados que permitam às entidades mais devedoras regularizarem as suas situações fiscais.

Recomendação V: Tornar os processos de licenciamento mais expeditos e transparentes. Coordenar os vários intervenientes no processo de licenciamento e permitir que este mesmo processo seja flexível no que diz respeito às datas de acesso ao mesmo.

Deve-se numa primeira fase considerar medidas que permitam controlar as causas para se evitar algumas das manifestações. Deve-se conseguir que os agentes e sujeitos envolvidos façam parte da solução mais do que do problema. Deve-se para isso fazer os intervenientes perceber a dimensão da questão e que se há quadros em que a informalidade é aceitável (e até talvez desejável) há outros onde ela é incomportável. No entanto este processo, ainda que possa vir a ser eficaz, não elimina o problema. É Por isso necessário continuar a fiscalizar e sempre que adequado penalizar.

Recomendação VI: Fiscalização e Penalização: Definir os alvos e fazer fiscalizações com penalizações severas. Estas fiscalizações devem ser bem publicitadas e anunciadas e as penalizações expeditas e muito severas. Neste ponto é necessário ter-se em atenção a conjuntura em que se está a trabalhar. Há um tipo de informalidade que é sempre necessário uma vez que a economia informal não absorve determinado tipo de actividade (ex: uma pessoa reformada que aluga quartos em casa própria). Em alturas de crise, mais difícil será absorver algumas das actividades e dos trabalhadores informais. Mais premente se torna ter cuidado com as fiscalizações.

5. CONCLUSÃO

O estudo apresentado tem como objectivo último analisar os mecanismos que incentivam a existência da economia informal no sector do comércio, alojamento e restauração e contribuir para a redução do fenómeno.

A metodologia utilizada partiu de uma revisão da literatura de forma a delimitar o âmbito da análise. Numa fase posterior, e dada a dificuldade de medir o fenómeno no terreno, levou-se a cabo um conjunto de entrevistas com agentes institucionais e individuais envolvidos nos ramos de actividade em estudo.

Do ponto de vista do conceito de informalidade a literatura aponta para um termo multifacetado que engloba diversas realidade. O conceito abarca desde economia ilegal, a economia não declarada, economia não registada e evasão fiscal. No estudo tentamos manter o conceito o mais aberto possível de forma a incorporar as diversas sensibilidades dos agentes entrevistados.

Das conversas tidas conseguiu-se fazer um levantamento das principais formas de economia informal presentes no sector em análise.

Temos então uma grande parte da informalidade associada aos impostos, da qual se destaca em particular o IVA. O facto de o tipo de transacção efectuada ser tipicamente de pequeno valor e em grande número justifica a dificuldade no controlo deste tipo de actividade não declarada. Também existe algum tipo de informalidade parcial associada ao IRC e até ao IRS. Este último caso está associado à informalidade existente no mercado do trabalho, de onde a não declaração parcial de IRS é apenas um exemplo. A este juntam-se a não declaração total de vencimento com a consequente diminuição das contribuições para a segurança social e a declaração parcial das jornadas de trabalho. Também como exemplo de não formalidade no mercado de trabalho temos o uso abusivo de contratos a termo que fragiliza de forma considerável a posição dos trabalhadores. Temos também toda uma face de economia sombra ou à margem que está intimamente ligada ao não cumprimento de obrigações em termos de produto/serviço. A não verificação de normas de higiene, segurança ou até de qualidade são exemplos desta situação. Por último as dificuldades burocráticas e processuais ligadas aos diversos licenciamentos de actividade justificam um funcionamento muitas vezes em condições ilegais ou no mínimo marginais.

Por serem claras as motivações para a informalidade e por serem graves as consequências da mesma achamos imperativo o levantamento de algumas linhas de actuação que possam permitir debelar o cenário do sector. Assim e após uma filtragem das entrevistas tidas, cruzando o resultado das mesmas com experiências concretas e ainda com algumas propostas do relatório McKinsey (2003), identificaram-se as seguintes recomendações:

1) Utilização de indicadores físicos para a tributação de impostos (nomeadamente IVA e IRC)

2) Campanha de sensibilização para a importância da factura junto de benefícios a quem efectivamente solicite factura

3) Campanhas de sensibilização e pedagogia junto dos trabalhadores para compreenderem a importância da formalidade da sua situação laboral.

4) Plano de recuperação (tipo plano Mateus) para que quiser regularizar situações, nomeadamente laborais. Este plano seria uma tentativa de envolver os infractores no combate à infracção.

5) Agilização dos processos de licenciamento de inicio de actividade ao mesmo tempo que se torna os mesmos mais transparentes. Homogeneização de prazos nos vários mecanismos de autorização e licenciamento exigidos.

6) Acompanhamento de todas estas medidas por acções de fiscalização e penalização em situações graves, escolhidas caso a caso e com publicitação clara.

O estudo pretende ser uma contribuição para a literatura existente tentando iniciar um processo de levantamento da informalidade na economia real para que se possa vir a chegar a um diagnóstico global. Acreditamos que as medidas enunciadas podem ajudar num combate difícil e necessário.

BIBLIOGRAFIA

ANTUNES, ANTÓNIO AND CAVALCANTI, TIAGO (2006). "Custos De Legalização, Cumprimento De Contratos E O Sector Informal," *Boletim Económico.* Banco de Portugal, 69-80.

FEIGE, EDGAR L. (1990) "Defining and Estimating Undergroud and Informal Economies: The New Institutional Economics Approach." *World Development*, *18*(7), pp. 898-1002.

FRIEDMAN, ERIC; JOHNSON, SIMON; KAUFMANN, DANIEL AND ZOIDO-LOBATON, PABLO (2000). "Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries." *Journal of Public Economics*, *76*(3), pp. 459-93.

HUSSMANNS, RALF. (2005)"Measuring the Informal Economy: From Employment in the Informal Sector to Informal Employment," *Working Papers.* Geneva: International Labour Office, 38.

JOHNSON, SIMON; KAUFMANN, DANIEL AND ZOIDO-LOBATON, PABLO (1998). "Corruption, Public Finances, and the Unofficial Economy," *Policy Research Working Paper.* Washington: The World Bank.

KADOKURA, TAKASHI. (2007) "What Is the Underground Economy?" *Japanese Economy*, *34*(2), pp. 7-19.

LASSEN, DAVID DREYER. (2006) "Ethnic Divisions, Trust, and the Size of the Informal Sector." *Journal of Economic Behavior & Organization*, *63*, pp. 423-38.

LORD GRABINER. (2000) "The Informal Economy," Lord Grabiner QC, 46.

MCKINSEY & COMPANY.(2003) *Portugal 2010: Acelerar O Crescimento Da Produtividade - Resumo De Conclusões*.

OCDE. (2002) "Measuring the Non-Observed Economy: A Handbook," Paris: OCDE - Organização para a Cooperação e Desenvolvimento Económico, 233.

SANTOS, DANIEL. (1996) "Avaliação Da Actividade Económica E Economia Não Registada - Contribuição Dos Inqueritos Às Famílias Para Medir O Sector Informal," *INE.* 24.

SCHNEIDER, FRIEDRICH. (2005) "Shadow Economies around the World: What Do We Really Know?" *European Journal of Political Economy*, *21*(3), pp. 598-642.

SCHNEIDER, FRIEDRICH AND ENSTE, DOMINIK H. (2000) "Shadow Economies: Size, Causes, and Consequences." *Journal of Economic Literature*, *38*(1), pp. 77-114.

TANZI, VITO. (1999) "Uses and Abuses of Estimates of the Underground Economy." *Economic Journal*, *109*(456), pp. 338-F47.

ASSOCIAÇÕES DE DESENVOLVIMENTO LOCAL: RESPONSABILIDADE SOCIAL *VERSUS* ESTRATÉGIA ECONÓMICA

FÁTIMA DAVID
E-mail: sdavid@ipg.pt
RUTE ABREU
E-mail: a@ipg.pt
Instituto Politécnico da Guarda
ODETE PINHEIRO
E-mail: pinheiroodete@gmail.com
Técnica Superior da ETL

Resumo

Apesar de ser reconhecido que a responsabilidade social e a estratégia económica das organizações são determinantes para o sucesso económico e o desenvolvimento social e ambiental da sociedade em que se encontram inseridas, constata-se que as práticas socialmente responsáveis se centram, muitas vezes, no estrito cumprimento das obrigações legais, esquecendo-se o real objectivo das mesmas. Neste contexto, esta investigação caracteriza as Associações de Desenvolvimento Local enquanto parceiras das novas estratégias económicas, em substituição das anteriores estratégias, assentes no paradigma do mercado de capitais. A investigação conclui que, a dinamização das economias locais passa pela fixação da população nos territórios rurais e pela realização de investimentos públicos e privados, que dotem as regiões de infra-estruturas que lhes permitam aproveitar as suas potencialidades e rentabilizar, de forma equilibrada, as suas riquezas, nomeadamente de recuperação do património cultural e ambiental e das actividades tradicionais.

Palavras-chave: Responsabilidade Social, Sustentabilidade, Desenvolvimento Económico.
Área Temática: Estratégia Económica

Abstract

Although social responsibility and economic strategy of organizations are recognized as crucial to economic success and social and environmental development of the society in which they are inserted, socially responsible practices focus, often, on strict compliance of the law, forgetting their real purpose. Thus, this research characterizes the Associations of Local Development as partners in the new economic strategy, replacing the previous strategies, which is based on capital market paradigm The research concludes that the promotion of local economies depends on settling the population in rural areas and of public and private investments, which equip the regions with infrastructures to enable them to harness their potential and, in a balanced way, promote their wealth, including the preservation of their cultural and environmental heritage and traditional activities.

Key words: Social Responsibility, Sustainability, Economic Development.
Thematic Area: Economic Strategy

1. INTRODUÇÃO

Se por um lado, se reconhece que a responsabilidade social e a estratégia económica das organizações são determinantes para o sucesso económico e o desenvolvimento social e ambiental da sociedade em que se encontram inseridas, sem as quais não seria possível às mesmas atingir a sua eficiência e eficácia económica. Por outro lado, constata-se que as práticas socialmente responsáveis se centram, muitas vezes, no estrito cumprimento das obrigações legais e fiscais, esquecendo-se o real objectivo das mesmas, ou seja, a sustentabilidade da sociedade, em geral, e da organização, em particular.

A organização socialmente responsável respeita os direitos dos seus trabalhadores, não recorrendo à exploração de mão-de-obra infantil, nem exercendo práticas discriminatórias e no caso de recorrerem a mão-de-obra localizada noutros países, nomeadamente, países em desenvolvimento, preocupam-se com as condições de vida destes trabalhadores (Abreu, 2005). Assim, a responsabilidade social da organização significa a integração das dimensões económica, social e meio ambiental na gestão, a qual enquanto gestão ética conduz à existência de comportamentos responsáveis face a todos os *stakeholders* (Bilbao, 2008; Dias, 2008). Tal como esgrimam Rego et al. (2007):

> "*os problemas e os escândalos éticos que têm sido mediaticamente propalados nos anos recentes apelam para a necessidade de se fomentar, nas empresas (…), uma cultura em que a honestidade, a integridade, a justiça, a confiança e o respeito sejam valores almejados*".

Adicionalmente, a organização começa, actualmente, a considerar como determinante para o seu sucesso empresarial o desenvolvimento local, sem o qual não é possível atingir a eficiência e eficácia económica. O desenvolvimento local pode entender-se como o:

> "*processo endógeno, equilibrado e sustentável de desenvolvimento territorial; e ainda um processo no qual todos os actores devem participar, intervir e investir de acordo com as próprias possibilidades*" (CIF, 2004: 4).

À semelhança do que defende o Centro Internacional de Formação da Organização Internacional do Trabalho (CIF, 2004), considera-se que a organização e o sistema económico local no seu conjunto são sustentáveis a longo prazo se toda a comunidade for envolvida num desenvolvimento que consiga melhorar os seus níveis de vida, que promova um crescimento harmonioso e equilibrado, e em que haja um ambiente de consenso e de paz social.

Face ao exposto, esta investigação caracteriza as Associações de Desenvolvimento Local (ADL) enquanto parceiras das novas estratégias económicas, em substituição das anteriores estratégias, assentes no paradigma do mercado de capitais. Assim, após esta introdução, seguem-se mais quatro partes claramente relacionadas. Na primeira parte (ponto 2) apresentam-se, no contexto da sua evolução histórica, os conceitos de responsabilidade social e sustentabilidade das organizações, tendo por base as suas linhas de orientação e princípios gerais. Na segunda parte (ponto 3) discute-se a importância da sustentabilidade das organizações no desenvolvimento económico, especificamente de nível local. Na terceira parte (ponto 4), analisam-se as boas práticas das Associações de Desenvolvimento Local, considerando dois projectos nacionais, um de cariz cultural e outro de natureza social. Por último, na quarta parte (ponto 5), apresentam-se as considerações finais que são o corolário das reflexões precedentes.

2. A RESPONSABILIDADE SOCIAL E A SUSTENTABILIDADE

A temática da responsabilidade social das empresas (RSE) foi introduzida na década de 50 na abordagem americana do ambiente empresarial, no domínio da ética nos negócios (Santos *et al.*, 2005). De facto, a globalização e a exigência por uma maior transparência nos negócios forçaram as organizações a adoptar uma postura mais responsável e a (re)considerem a sua forma de actuar numa sociedade em que o bem-estar dos cidadãos e o consumo desenfreado são paralelos.

É, neste contexto que, a responsabilidade social aparece defendida, a nível internacional e nacional, por inúmeras instituições e especialistas. A nível internacional, destaca-se a **União Europeia** (UE), que através da sua Comissão elaborou o *Green Paper - Promoting a European framework for corporate social responsibility* (COM, 2001)[1]. Ao qual, depois de um debate

[1] A UE através do Livro Verde procura promover e divulgar que as empresas e instituições adoptem uma estratégia de Responsabilidade Social, indo de encontro ao objectivo estratégico delineado na «Estratégia de Lisboa» de 2000, especificamente: "*tornar-se na economia baseada no conhecimento*

público, se seguiu a comunicação da Comissão sobre *Corporate Social Responsibility: A business contribution to Sustainable Development* (COM, 2002). Neste último documento, entende-se por RSE:

> "*o conceito através do qual, de forma voluntária, as empresas integram as preocupações sociais e ambientais nas suas operações empresariais e na interacção com os seus stakeholders*" (COM, 2002: 5).

Também a **Organização para a Cooperação e Desenvolvimento Económico** (OCDE) contribuiu para o debate em torno da RSE, ao apresentar os princípios do *Corporate Governance* (OECD, 1999) e as linhas orientadoras para as empresas multinacionais - *Guidelines for multinational enterprises* (OECD, 2000). A OCDE defende que um eficiente regime de *corporate governance* ajuda a assegurar que as empresas utilizam o seu capital de forma eficiente, tendo em conta o benefício da sociedade na qual operam e dos seus *shareholders* (OECD, 1999). Assim, cada empresa define a sua atitude na sociedade, de forma a contribuir para o equilíbrio do ambiente global, sendo o seu principal problema, uma vez aceite o conceito de responsabilidade social da empresa, a respectiva implementação (Jones, 1980).

Outras organizações também desempenharam um papel primordial, nomeadamente: a **Organização Internacional de Normalização** (do inglês, *International Organization for Standardization* - ISO), ao publicar vários sistemas de normas, de que se salientam as normas da série ISO 14000 sobre gestão ambiental; a ***Global Reporting Initiative*** (GRI), que desenvolveu linhas de orientação para a elaboração do relatório de sustentabilidade[2]; o ***Institute of Social and Ethical Accountability*** (ISEA), que publicou normas sobre certificação ética; o **Instituto ETHOS**, com a missão de mobilizar, sensibilizar e ajudar as empresas a gerir os seus negócios de forma socialmente responsável; e a **Organização das Nações Unidas** (ONU).

A **Organização das Nações Unidas**, na pessoa do seu Secretário-Geral Kofi Annan, analisou a problemática da RSE num contexto de globalização, com a *Global Compact Initiative* (ONU, 2000). Este documento teve por base a Declaração Universal dos Direitos Humanos (MNE, 1978), os princípios fundamentais da Organização Mundial do Trabalho, relativos ao direito do trabalho, e os *Rio Principles* relativos ao ambiente e desenvolvimento. Estes últimos foram definidos na *The United Nations Conference on Environment and Development,* ou «Cimeira da Terra», realizada no Rio de Janeiro em Junho de 1992. Esta iniciativa dirigida à comunidade empresarial internacional promoveu a adopção, por parte das mesmas, de um conjunto de valores e o cumprimento de nove princípios:

- Na área de Direitos Humanos (1. As empresas devem apoiar e respeitar a protecção de direitos humanos reconhecidos internacionalmente; e 2. Certificar-se de que suas próprias empresas não são cúmplices de abusos e violações de direitos humanos);
- Na área de Trabalho (3. As empresas devem apoiar a liberdade de associação e o reconhecimento efectivo do direito à negociação colectiva; 4. Eliminação de todas as formas de trabalho forçado ou compulsório; 5. Erradicação efectiva do trabalho infantil; e 6. Eliminar a discriminação com respeito ao empregado e ao cargo); e
- Na área de Meio Ambiente (7. As empresas devem adoptar uma abordagem preventiva aos desafios ambientais; 8. Desenvolver iniciativas para promover maior responsabilidade ambiental; e 9. Incentivar o desenvolvimento e difusão de tecnologias limpas que não prejudiquem o meio ambiente) (Abreu e David, 2004).

Adicionalmente, na «Cimeira da Terra» foram estabelecidos 5 acordos orientadores para um desenvolvimento sustentável (Santos *et al.*, 2005), especificamente:

---

mais dinâmica e competitiva do mundo, capaz de garantir um crescimento económico sustentável, com mais e melhores empregos, e com maior coesão social" (COM, 2000).

[2] A importância deste organismo é enfatizada por Carvalho et al. (2008) e Dias et al. (2008), ao considerarem que a GRI orienta e estabelece as directrizes para a divulgação de informação voluntária das organizações, aquando da elaboração dos seus relatórios anuais.

- **Agenda 21**: plano global de acção em prol do Desenvolvimento Sustentável, contendo mais de 100 áreas temáticas, indo do comércio internacional e do ambiente até à agricultura e à desertificação;
- **Declaração do Rio sobre Ambiente e Desenvolvimento**: conjunto de 27 princípios orientadores da integração do ambiente nas políticas de desenvolvimento económico (como por exemplo, o princípio do poluidor pagador);
- **Declaração de Princípios sobre as Florestas**: primeiro consenso global sobre gestão, conservação e desenvolvimento sustentável das florestas mundiais;
- **Convenção - Quadro sobre Mudança Climática**: acordo, com carácter de compromisso legal, relativo à estabilização dos gases com efeito de estufa em níveis que não perturbem o clima mundial;
- **Convenção sobre Diversidade Biológica**: acordo, com carácter de compromisso legal, relativo à conservação da diversidade mundial, em termos de genes, espécies e ecossistemas.

Também em 1992, na «Cimeira de Joanesburgo» realizada em Setembro desse ano, sobre Desenvolvimento Sustentável, foram tomadas decisões extremamente importantes no que dizia respeito à operacionalidade dos acordos já definidos na «Cimeira da Terra» (Santos *et al.*, 2005), nomeadamente:

- Reafirmar o desenvolvimento sustentável como prioridade central a nível internacional;
- Fixar objectivos de acção para melhor atingir o desenvolvimento sustentável, com metas e compromissos.
- Continuar a impulsionar a luta contra a pobreza e protecção do meio ambiente;
- Avançar na resolução das questões da energia e do saneamento básico;
- Apoiar a criação de um fundo mundial para a erradicação da pobreza;
- Dar mais atenção aos problemas específicos do continente africano.

Já a nível nacional, a Comissão de Normalização Contabilística (CNC) publicou em Junho de 2002, de acordo com as anteriores recomendações da Comissão Europeia (COM, 2001; 2002), a Directriz Contabilística (DC) 29 – Matérias Ambientais, a qual obrigou, a partir de Janeiro de 2003, todas as empresas abrangidas pelo Plano Oficial de Contabilidade a produzir e publicar um relatório ambiental, que:

> "*(…) diz respeito aos critérios para o reconhecimento, mensuração e divulgação relativos aos dispêndios de carácter ambiental, aos passivos e riscos ambientais e aos activos com eles relacionados (…) susceptíveis de afectar, a posição financeira e os resultados da entidade relatada*" (MF, 2005: 6181).

Assim, em Portugal, a regulamentação na área ambiental constitui um facto recente, pois até à aprovação da DC 29 não havia qualquer exigência específica, em termos legislativos e normativos, para que as empresas publicassem informação ambiental ou social da sua actividade, pelo que a pouca informação sobre matérias ambientais existente era apresentada de forma voluntária pelas mesmas (Abreu e David, 2004).

Por outro lado, também o Conselho Económico e Social (CES)[3], como órgão consultivo e de concertação no domínio das políticas económicas e sociais, emitiu em 2003 um parecer sobre a iniciativa da responsabilidade nas empresas, propondo-se:

> "*(…) construir uma base de dados e um sistema de tratamento e difusão de informação com respeito às boas práticas*" (CES, 2003: 26).

[3] Criado pela Lei nº 108/91, de 17 de Agosto (AR, 1991). Compõe-se por 64 membros de diferentes quadrantes da sociedade (representantes do Governo, das confederações sindicais, das organizações empresariais, do sector cooperativo, da área da ciência e tecnologia, profissionais independentes, entre outras).

Mais recentemente, a Norma Portuguesa (NP) 4469-1, com entrada em vigor em Fevereiro de 2008, veio responder às necessidades sentidas pelas organizações ao nível da definição de um sistema de gestão da responsabilidade social, especificando quais os:

> "(…) *requisitos de um sistema de gestão da responsabilidade social, que permite a uma organização desenvolver e implementar uma política, objectivos e acções coerentes, tendo em conta os requisitos legais, regulamentares e outros que a organização subscreva. Aplica-se aos aspectos da responsabilidade social que a organização identifica como sendo aqueles que pode controlar e aqueles que pode influenciar"* (IPQ, 2008: 11).

Desde logo, as organizações, em geral, e as portuguesas, em particular, têm toda a vantagem em implementar esta Norma, uma vez que a aplicação da mesma pode ser um factor de competitividade no mercado global, traduzindo-se numa mais-valia, ao permitir que as mesmas passem a ser certificadas como empresas socialmente responsáveis. Paralelamente, esta Norma insere-se no âmbito da Estratégia Nacional de Desenvolvimento Sustentável (PCM, 2007), que defende os seguintes objectivos:

- Preparar Portugal para a "sociedade de conhecimento";
- Crescimento sustentado, competitividade à escala global e eficiência energética;
- Melhor ambiente e valorização do património natural;
- Mais equidade, igualdade de oportunidade e coesão social;
- Melhor conectividade internacional do País e valorização equilibrada do território;
- Uma Administração Pública mais eficiente e modernizada.

Estes exemplos reflectem as crescentes preocupações e desafios que cada empresa enfrenta na actual sociedade, por um lado, porque os recursos são escassos e as suas necessidades infinitas, e, por outro, porque a responsabilidade social das empresas deve ser uma constante preocupação na realização da sua actividade económica, aumentando a sua competitividade empresarial e atingindo o bem-estar colectivo. Como afirma Mintzberg (1983), a responsabilidade social envolve um conjunto complexo e contraditório de necessidades ao longo do ciclo de vida da empresa, pelo que depende da sua influência responder, melhor ou pior, às necessidades sociais. Com efeito, cada vez mais, existe um largo consenso de que as empresas e os governos devem, em parceria, aceitar a sua responsabilidade moral em relação ao bem-estar social e aos interesses individuais nas transacções económicas (Amba-Rao, 1993).

Neste sentido, concorda-se com Díaz-Zorita e Gonzalo (1996) ao defenderem a RSE como uma consequência lógica do papel que as organizações desempenham na sociedade, considerando-se inquestionável a crescente responsabilidade social das mesmas, na medida em que da sua actividade depende, em larga medida, o bem-estar presente e futuro da sociedade em que se integram.

Uma estratégia de responsabilidade social numa organização baseia-se nos princípios e valores adoptados pela mesma e faz parte da sua estratégia global, sendo importante seguir uma linha de coerência entre a execução prática e o discurso teórico. Essa estratégia, centrada no exercício planeado e sistemático de acções, estratégias e implementação de canais de relacionamento entre a organização e o meio onde se insere, traduz-se, segundo Bueno (2003) *in* Viotto (2006), em:

- Contribuir para o desenvolvimento social, pelo respeito do ser humano, independentemente das suas opiniões e crenças, pela valorização da diversidade cultural e pela defesa da liberdade de pensamento e expressão;
- Proporcionar condições ideais de trabalho para os seus colaboradores, em concreto: remuneração justa; motivação profissional; realização pessoal; e estímulo ao diálogo e à participação no processo de tomada de decisão;
- Assumir a transparência e a ética como atributos fundamentais na condução dos negócios;
- Preservar o meio ambiente, privilegiando a gestão de recursos e a oferta de produtos não agressivos à natureza;
- Praticar a excelência na fabricação de produtos e na prestação de serviços, tendo em vista os interesses e expectativas dos consumidores e da sociedade;

- Implementar projectos que visem o desenvolvimento sustentável das comunidades onde se inserem.

Por conseguinte, concorda-se com Duarte (2008), ao considerar que a responsabilidade social da organização:

- Deve ser encarada como fazendo parte integrante da sua gestão;
- É uma forma de comportamento voluntariamente adoptado, que vai além das leis existentes;
- É uma forma de gestão que se define pela relação ética e transparente da empresa com todos os grupos com os quais se relaciona ou interage;
- Estabelece metas empresariais compatíveis com o desenvolvimento sustentável da sociedade, preservando recursos ambientais e culturais para as gerações futuras, respeitando a diversidade e promovendo a redução das desigualdades sociais.

Numa era de globalização, e as organizações não podem deixar de pautar as suas acções por princípios de eficiência e de rentabilidade económica, mas também não podem deixar de integrar nas suas actividades preocupações de ordem social, cultural e ambiental, enquanto condições necessárias para a sustentabilidade e sobrevivência das organizações.

Pelo exposto, pode concluir-se que existe uma estreita ligação entre RSE e sustentabilidade, cujo conceito não é novo, pois já em 1987 a Comissão Brundtland[4] definiu «Desenvolvimento Sustentável» enquanto:

> "*modelo de desenvolvimento que procura satisfazer as necessidades presentes sem comprometer a capacidade das gerações futuras de suprir as suas próprias necessidades*" (WCED, 1987 *in* Lourenço e Machado, 2008: 6).

Contudo, actualmente, este conceito, teórico e bastante genérico, tende a evoluir para conceitos e metas mais objectivas. Crowther e Rayman-Bacchus (2004: 239), consideram que a sustentabilidade:

> "*…diz respeito ao efeito produzido por uma acção adoptada no presente e as suas repercussões no futuro. Se os recursos são utilizados, no presente, então eles não estão mais disponíveis no futuro, e isto é particularmente preocupante quando os recursos são finitos, em quantidade*".

Neste mesmo contexto, a NP 4469-1, relativa ao sistema de gestão da responsabilidade social, considera o desenvolvimento sustentável como o desenvolvimento que procura satisfazer as necessidades do presente sem comprometer a capacidade das gerações futuras em satisfazerem as suas próprias necessidades (IPQ, 2008).

3. A SUSTENTABILIDADE E O DESENVOLVIMENTO ECONÓMICO

O sucesso de cada País, região ou localidade depende necessariamente da sua capacidade para se especializar naquilo que consiga estabelecer como vantagens comparativas, efectivas e dinâmicas, decorrentes de seu desenvolvimento e da capacidade local de inovação. Também, o sucesso das organizações depende da sua capacidade para implementar estratégias de RSE, com vista a cumprirem com a trilogia económica/social/ambiental (Giró, 2002). Como defende Heleno (2008), as organizações estão conscientes de que podem contribuir para o desenvolvimento sustentável, gerindo as suas operações de modo a consolidar o seu crescimento económico e aumentar a competitividade, ao mesmo tempo que asseguram a defesa do meio ambiente e provem a responsabilidade social.

Neste contexto, a responsabilidade social das organizações pelos impactos das suas decisões e actividades na sociedade em que se inserem, através da adopção de comportamentos éticos e

[4] Comissão Mundial das Nações Unidas para o Ambiente e Desenvolvimento, chefiada pelo Primeiro-Ministro Norueguês Gro Brundtland.

transparentes, deve ser consistente com o desenvolvimento sustentável e o bem-estar dessa mesma sociedade (Costa, 2008). Se, por um lado, a maior preocupação actual das organizações se centra na sustentabilidade (Lunardi e Clemente, 2008), por outro, a sustentabilidade deve reflectir políticas de desenvolvimento económico e social contínuo, sem prejuízo das políticas ambientais e dos recursos naturais, de cuja qualidade depende a continuidade da actividade humana (Duarte, 2008).

Associada aos modelos e políticas de desenvolvimento saídas do paradigma dominante encontra-se a delimitação de configurações territoriais dicotómicas, através da imposição de pólos de crescimento (geralmente urbanos e industriais, onde se concentram os processos de tomada de decisão de natureza política, económica e social) e da definição em simultâneo de áreas subdesenvolvidas (Figueiredo, 2003). Também, associada a esta dicotomia se encontra uma outra dupla dualidade: a dicotomia urbana/rural e a dicotomia indústria/agricultura. Os referenciais desta (dupla) dicotomia levaram a que as áreas rurais fossem, ainda que de forma selectiva e de acordo com o interesse económico que representavam para o processo global de crescimento, marginalizadas (Figueiredo, 2003).

Será por isso que, a problemática do desenvolvimento rural tem vindo a assumir uma relevância cada vez maior no quadro das políticas territoriais e sectoriais, estando-lhe associadas questões tão importantes e decisivas como o desenvolvimento económico e socialmente sustentado. O desenvolvimento rural sustentável só é possível tendo em atenção a conservação da natureza e em especial a manutenção da biodiversidade, a protecção da paisagem rural e manutenção das populações no espaço rural (IDRH, 2006). Para que exista um equilibrado desenvolvimento local torna-se necessário proporcionar às populações a manutenção de actividades agrícolas e silvícolas e florestais rentáveis, complementando-as e adequando-as, quanto possível, com a prestação de serviços.

A dinamização das economias locais pressupõe, assim, a fixação da população nos territórios rurais e logo a realização de investimento público e privado, ao nível: da melhoria dos índices de qualidade de vida das populações residentes; da recuperação do património e das actividades tradicionais; e da criação de infra-estruturas e de actividades que permitam a valorização e revitalização das zonas rurais (IDRH, 2006). Assim, as políticas de desenvolvimento rural devem ser desenhadas e dirigidas para as especificidades de cada território e implementadas em interacção com os actores públicos e privados que actuam no local.

Neste âmbito aparece o *Programa Leader,* apresentado como um programa da União Europeia (UE), que se preocupa com a acção integrada para o desenvolvimento de regiões essencialmente rurais, onde aparecem como princípios e orientações gerais o seu perfil integrador de políticas, actores, acções e projectos locais (Rover e Henriques, 2006). A sua abordagem de desenvolvimento endógeno baseia-se em iniciativas e competências locais, promovendo a aquisição do saber-fazer no domínio do desenvolvimento integrado local e divulgando esse saber-fazer a outras zonas rurais (COM, 2006). Deste modo, o *Programa Leader* sublinha a importância das populações locais tomarem consciência do valor dos seus territórios, da riqueza das suas identidades e da qualidade das suas paisagens e nesta perspectiva apontam para a necessidade de se criarem condições para a aquisição de competências e intervenção activa dos actores locais nos seus territórios (DGADR, 2000).

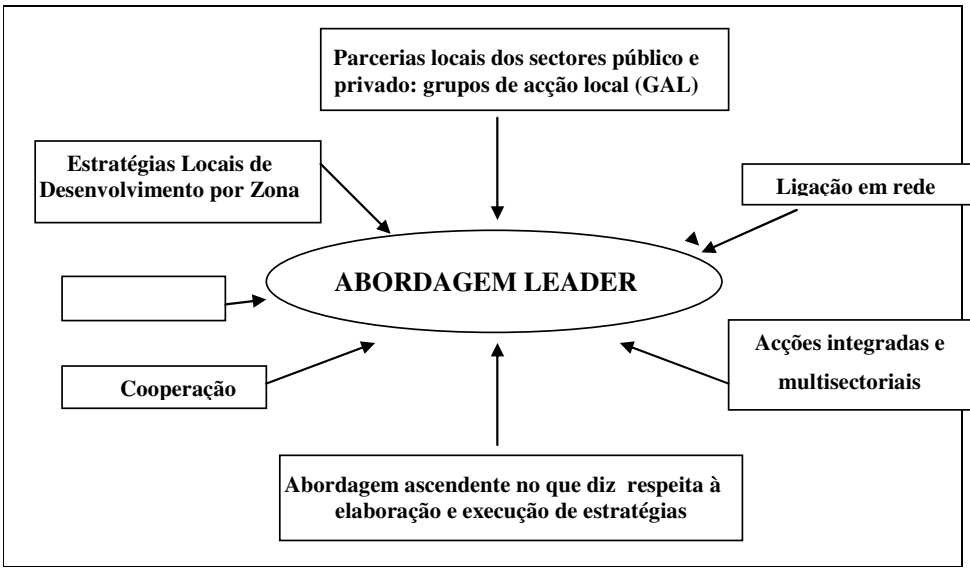
Como objectivos específicos, o *Programa Leader* procura: incentivar e apoiar os agentes rurais a reflectirem sobre as potencialidades das zonas rurais a longo prazo; e estimular a aplicação de estratégias de desenvolvimento sustentável integradas, de qualidade elevada e originais, destinadas a fomentar novas formas de:

- Valorizar o património natural e cultural;
- Reforçar e diversificar o ambiente económico, a fim de contribuir para a criação de empregos;
- Promover a igualdade de acesso a cultura a educação ao desporto e ao trabalho entre os grupos mais desfavorecidos;
- Melhorar as capacidades organizativas da comunidade (DGADR, 2000).

A abordagem do *Programa Leader*, representada pelo Esquema 1, implica a definição de uma política de desenvolvimento baseada na situação específica de cada território, tentando

valorizar os pontos fortes e minimizando os seus pontos fracos, através de uma intervenção directa e específica, envolvendo de uma forma directa todos os agentes de desenvolvimento, designadamente: poder local; empresas; associações; e população em geral (OETR, 2007)

Esquema 1. Abordagem do *Programa Leader*



Fonte: COM (2006: 8).

Da análise do Esquema 1, é possível concluir que existem sete características principais que resumem a abordagem do *Programa Leader*, especificamente: abordagem assente no nível local; abordagem ascendente; abordagem de parceria e grupo de acção local (GAL); inovação; abordagem integrada; criação de redes e cooperação entre zonas; e financiamento e gestão local (COM, 2006). Cada característica complementa e interage positivamente com as outras ao longo de todo o processo de implementação, com efeitos duradouros na dinâmica das zonas rurais e na capacidade de resolução dos seus próprios problemas (COM, 2006).

A nível nacional, o *Programa Leader*+, legislado pelo Regulamento (CE) nº 1260/1999 do Conselho, de 21 de Junho (CE, 1999), estabeleceu as disposições gerais sobre os fundos estruturais e definiu os objectivos prioritários do *Programa Leader*. Este Programa decorreu no período compreendido entre 1991 e 2008, tendo a sua implementação sido inserida nos fundos estruturais compreendidos entre 1991 e 2006 (COM, 2006). Respeitando a dimensão ambiental, económica, social e cultural do território nacional, o *Programa Leader*+ permitiu desenvolver estratégias de desenvolvimento baseadas na gestão dos recursos endógenos, em detrimento de outros modelos de desenvolvimento baseados na importação maciça de capital e de recursos exógenos para os territórios rurais (DGADR, 2000). Por conseguinte, este Programa ajuda as comunidades rurais a melhorarem a qualidade de vida e a prosperidade económica da sua zona, sendo objecto de um financiamento conjunto do Governo português e da EU, especificamente da secção Orientação do Fundo Europeu de Orientação e de Garantia Agrícola (FEOGA).

Mais recentemente, foi aprovada pelo Conselho das Comunidades Europeias a Decisão 2006/144/CE, de 20 de Fevereiro de 2006 (CE, 2006), relativa ao plano estratégico de desenvolvimento rural para o período 2007-2013, que sintetiza as medidas e acções de desenvolvimento rural a adoptar, com vista à promoção competitiva dos territórios, nomeadamente no sector agro-florestal, e cria territórios rurais sustentáveis.

Também, a nível nacional este plano estratégico teve a sua tradução na aprovação pelo Conselho de Ministros do Plano Estratégico Nacional para o período 2007-2013, o qual tem como objectivos estratégicos para os territórios rurais: aumentar a competitividade dos sectores agrícola e florestal; promover a sustentabilidade dos espaços rurais e dos recursos naturais; e revitalizar económica e socialmente as zonas rurais (MADRP, 2008). A par destes objectivos, foram ainda definidos objectivos horizontais ou transversais, designadamente: reforçar a coesão

territorial e social; e promover a eficácia da intervenção dos agentes públicos, privados e associativos na gestão sectorial e territorial (MADRP, 2008).

Contudo, de acordo com Matos e Tsuji (2002), embora se tenha feito um esforço no desenvolvimento rural, existem determinadas condicionantes que dificultam o progresso dos territórios rurais, tais como:

- **Desfasamento entre políticas, estratégias, instrumentos e organizações**, pois as organizações que executam as políticas públicas devem ter a capacidade de ajustar as políticas e os instrumentos ao desenvolvimento rural;
- **Dispersão institucional**, na medida em que a população se defronta com uma série complexa de agentes, representantes, funcionários, procedimentos e regulamentos, além de que cada entidade estabelece as suas próprias regras e existe pouca coordenação entre as entidades gestoras;
- **Falta de capacidade política de cada uma das entidades**, pois onde existe dispersão, há também o enfraquecimento da capacidade técnica e política de cada entidade, sendo que a dispersão implica igualmente uma competição irracional pelos recursos de investimento o que vai dificultar extraordinariamente a capacidade de coordenação;
- **Desertificação das áreas rurais**, pois a fraca produtividade, o empobrecimento, a ausência de alternativas e perspectivas de emprego, o esquecimento e fragilização das populações, conduzem ao abandono dos territórios;
- **Alterações Ambientais**, uma vez que actualmente se verificam alterações climatéricas acentuadas que transportam consigo doenças e ameaças ambientais, que vão alterar a paisagem florestal, para além das alterações provocadas pelo desleixo humano (por exemplo, os fogos florestais);
- **Acessibilidades e desenvolvimento tecnológico,** o sistema de transportes é um factor determinante da coesão social e territorial, e da competitividade do País;
- **Fraca qualificação dos recursos humanos**, na medida em que, no desenvolvimento global da sociedade, a qualificação é factor primordial para contrariar a debilitação dos territórios rurais. Para tal, a qualificação implica um investimento equilibrado em aspectos materiais e imateriais do desenvolvimento rural, o que vai capacitar as pessoas e organizações de forma a mobilizá-los e transformá-los em agentes activos e protagonistas de dinâmicas de desenvolvimento territorial.

Por conseguinte, defende-se que o desenvolvimento local deve fazer parte da missão da sociedade civil, não constituindo uma tarefa que o Estado deva assegurar isoladamente, devendo a intervenção ao nível da Administração Pública central e local passar, sobretudo, pela disponibilização de recursos e pelo estimulo à realização de iniciativas por parte da sociedade civil. Como defende Paula (2008), a sociedade civil tem um papel fundamental para exercer influência sobre as estruturas empresariais e políticas. Este tipo de desenvolvimento necessita, logicamente, de relações de forças (nomeadamente, os poderes locais) por meio das quais se estabelecem alianças e conflitos entre os diferentes actores locais, bem como da implementação de diversas organizações (designadamente, as Associações de Desenvolvimento Local) e práticas de gestão específicas, capazes de procederem à dinamização do desenvolvimento local sustentado.

4. AS BOAS PRÁTICAS DAS ASSOCIAÇÕES DE DESENVOLVIMENTO LOCAL

Para Oriol (2002) a teoria e a prática da responsabilidade social das empresas baseia-se nas prioridades, expectativas, atitudes e comportamentos dos indivíduos, sendo indiscutível que estes (enquanto consumidores, empregados ou investidores) premeiam as empresas socialmente responsáveis. Deste modo, para que as empresas desenvolvam uma prática coerente com a defesa teórica das boas práticas devem alinhar as suas actuações ao nível da responsabilidade social com as suas prioridades de negócio, devendo responder ao que o meio ambiente espera delas.

De encontro a esta abordagem vêm as práticas das Associações de Desenvolvimento Local, que, em Portugal, têm sido a forma preferencial da sociedade civil se organizar para assumir um papel activo no processo de desenvolvimento local. Constituídas como parceiras, públicas e/ou privadas, reunindo vários actores económicos e sociais ou, em alguns casos, constituídas em torno de uma ideia ou de um interesse mais específico, as ADL têm tido um papel crescente na intervenção na sociedade, quer como motores desse mesmo processo, no caso das intervenções do *Programa Leader+*, como também, no âmbito de outros programas para o desenvolvimento, ou simplesmente como beneficiárias desses programas, levando a cabo projectos específicos nas respectivas áreas de intervenção.

Por vezes, a exploração de um aspecto de um determinado recurso ou a sua exploração de uma forma inovadora, pode desenvolver sinergias, dando origem à reanimação ou criação de uma pequena economia baseada nesse recurso. Por exemplo, um lugar geográfico com notáveis atributos paisagísticos que uma vez descoberto por um pequeno projecto atrai outros investimentos. Ou, a recuperação, com fins meramente arquitectónicos, pelo menos na fase inicial, de um ou mais moinhos, que ao retomarem a sua produção, desaparecida há muito, voltam a estimular a produção agrícola.

Por conseguinte, a dinamização das economias locais passa, em primeira instância, pela fixação da população nos territórios rurais e pela realização de investimentos públicos e privados, que dotem as regiões de infra-estruturas que lhes permitam aproveitar as suas potencialidades e rentabilizar, de forma equilibrada, as suas riquezas, nomeadamente de recuperação do património cultural e ambiental e das actividades tradicionais. A confirmar esta realidade estão as Associações de Desenvolvimento Local, enquanto parceiras das novas estratégias económicas, em substituição das anteriores estratégias, assentes no paradigma do mercado de capitais. Assim, a título de exemplo, apresentam-se dois projectos que pretendem ser o corolário das boas práticas assumidas pelas Associações de Desenvolvimento Local em termos de desenvolvimento local.

O primeiro projecto, **«Começar pelo Princípio: formação em actividades teatrais no espaço rural»,** de âmbito cultural, foi desenvolvido através da cooperação de diversos agentes locais, designadamente: municípios; empresários; e estabelecimentos de ensino. O projecto surgiu da necessidade de promover mais intervenções culturais, identificadas pelos municípios de Melgaço, Monção, Paredes de Coura, Valença e Vila Nova de Cerveira como estratégicas para o desenvolvimento cultural da sua região e a melhoria da qualidade de vida da população local. Estes municípios, tendo por base os recursos locais, designadamente o património cultural da região, participaram activamente na criação de uma estratégia assente no desenvolvimento local, criando a Comunidade Intermunicipal do Vale do Minho para participar em reuniões estratégicas sobre o desenvolvimento cultural da região. Entre as principais áreas de intervenção figuram o desenvolvimento socioeconómico do Vale do Minho, a atracção do investimento, a criação de emprego e a diversificação das actividades económicas, a promoção da cultura local e a preservação do ambiente e dos recursos naturais (OETR, 2007).

O projecto «Começar pelo Princípio», no âmbito do *Programa Leader+*, para o período dos fundos estruturais entre 2000 e 2006 (mas que permitiu que os projectos fossem realizados até finais de 2008), teve um orçamento global de €251.330,00 e teve como fontes de financiamento: FEOGA - €94.229,00; fundos públicos nacionais - €31.416,00; e fundos privados - €125.665,00 (OETR, 2007). Este projecto constitui um exemplo de boas práticas, porque:

> *"Dá apoio à companhia de teatro local (Comédias do Minho) na realização de actividades de formação teatral na área abrangida. O acesso à informação, à formação e à cultura são factores importantes para melhorar a qualidade de vida nas zonas rurais. O projecto reforça as tradições e o património cultural local através da apresentação de peças sobre temas locais, como «As lendas do vale do Minho». A cultura também é importante para atrair mais visitantes, turistas e residentes à região. É, de um modo geral, reconhecido que a cultura contribui não só para o desenvolvimento social, mas também para o desenvolvimento económico local, através do incentivo ao crescimento do sector recreativo e do aumento da atractividade da zona"* (OETR, 2007: 58).

Este projecto, integrado numa estratégia de desenvolvimento cultural mais ampla para o Vale do Minho, permitiu a realização de alguns investimentos culturais importantes, nos últimos anos, nomeadamente: o Centro Cultural e o Museu Regional de Paredes de Coura; a Casa da Cultura e o Núcleo Museológico, em Melgaço; a Casa do Curro, o Cineteatro João Verde, em Monção; e a Biblioteca Municipal, em Valença. De facto, como realça o Observatório Europeu dos Territórios Rurais:

> "*(…) as actividades culturais locais foram reforçadas e estimuladas, tendo sido representadas várias peças nas freguesias locais. O projecto também animou actividades e estruturas culturais locais nos municípios participantes, tendo produzido um impacto particularmente importante nos jovens das zonas rurais. A cultura foi trazida para mais perto das populações locais e a sua compreensão das questões culturais melhorou*" (OETR, 2007: 58-59).

Adicionalmente, no âmbito deste projecto, foi realizada uma ampla gama de actividades de formação e culturais, que visaram sensibilizar e difundir a cultura teatral na zona, designadamente:

- Um curso de introdução à linguagem teatral, com o objectivo de familiarizar os alunos das escolas locais com esta linguagem e permitir a criação de grupos de teatro nas escolas;
- A iniciativa «Teatro na Escola», com o objectivo de formar professores das escolas dos cinco municípios sobre o modo de utilizar a linguagem teatral como instrumento pedagógico;
- A formação de actores amadores, com o objectivo de desenvolver actividades de teatro amador (OETR, 2007).

Por conseguinte, o valor acrescentado do projecto «Começar pelo Princípio» reside no facto de ter introduzido as actividades teatrais na região do Vale do Minho e por ter aproximado a cultura da população local. Se, por um lado, é difícil garantir a sustentabilidade financeira das actividades teatrais em zonas rurais, por outro, o financiamento do *Programa Leader+* permitiu realizar acções de formação e levar à cena várias produções teatrais, através do apoio dado à remuneração dos actores e à aquisição dos equipamentos utilizados. Donde, contribuiu para reconhecer a importância e o valor acrescentado da cultura no desenvolvimento social e na melhoria da qualidade de vida a nível local.

O segundo projecto, **«O Sonho da Criança - equipamentos para o funcionamento da creche»,** de âmbito económico, localiza-se no concelho de Mértola, distrito de Beja, e no âmbito do *Programa Leader+* tem um orçamento global de €50.060,50 e tem como fontes de financiamento: FEOGA com €37.500,00; e os fundos privados com €12.560,50. Este projecto constitui um exemplo de boas práticas, porque o Bairro da Esperança em Mértola é uma comunidade estigmatizada e excluída socialmente, por ser uma comunidade de etnia cigana. Este projecto visa melhorar a qualidade de prestação de serviços às famílias e crianças da comunidade:

- Permitindo o desenvolvimento integral das crianças;
- Facilitando a integração no mercado de trabalho das famílias; e
- Acompanhar e intervir precocemente em situações de risco, de forma, a atingir-se o objectivo de promover a inclusão social da comunidade global (CSCRBE, 2007).

Com o desenvolvimento social local desta comunidade e as dinâmicas locais geradas pelo envolvimento de parcerias público-privadas foi possível criar 5 postos de trabalho (1 educadora de infância, 3 auxiliares de educação e 1 auxiliar de serviços gerais), e, ainda, uma equipa de voluntariado vocacionada para a intervenção precoce de jovens e crianças em risco, constituída por um enfermeiro, uma educadora e um auxiliar da acção educativa (CSCRBE, 2007). Deste modo, também o projecto «O Sonho da Criança» contribui para promover a melhoria da qualidade de vida e reduzir os níveis de exclusão social da população, no estrito cumprimento dos princípios gerais de sustentabilidade do desenvolvimento social local.

A análise do projecto «O Sonho da Criança», enquanto promotor do desenvolvimento social do Bairro da Esperança em Mértola, permite concluir que do mesmo resultaram dois tipos de impactos positivos:

- Impactos directos, verificados através da criação de 5 postos de trabalho qualificado e do aumento do rendimento familiar, através da capacidade individual de cada indivíduo para procurar emprego, bem como pela qualificação profissional da comunidade;
- Impactos indirectos, verificados através da promoção e criação de condições de igualdade de oportunidades entre homens e mulheres, quer pela integração de um elemento do sexo masculino na equipa da creche, quer pelo aumento de vagas que permitiu acolher crianças em risco, dando-lhe a possibilidade de integração na sociedade, dando às mulheres uma maior facilidade ao acesso ao mercado de trabalho e à frequência de cursos de formação profissional.

Assim, a transmissibilidade destas experiências impõe o seu conhecimento e o seu intercâmbio. O desenvolvimento de redes ligadas ao desenvolvimento local, muito utilizadas na abordagem *Programa Leader*, parecem ser um instrumento eficaz a considerar no desenho de uma política de desenvolvimento. Intervenções inovadoras e interessantes em determinado local podem ter aplicabilidade noutro ou noutros locais, com características idênticas, mesmo que por vezes localizados a grande distância, talvez até noutros países.

5. CONSIDERAÇÕES FINAIS

Actualmente, a construção de alianças e parcerias, viabilizadas pela cooperação intersectorial e internacional, apresentam-se como uma forma mais adequada de promover o desenvolvimento rural em complementaridade vários programas e iniciativas comunitárias, onde se inclui o *Programa Leader*, no quadro dos interesses associados ao desenvolvimento das zonas rurais.

O desenvolvimento rural traz consigo novas perspectivas de adequação das políticas nacionais às peculiaridades regionais e locais, na sua dimensão económica, social, cultural e ambiental, respeitando todos os princípios do desenvolvimento sustentável (Senhoras, 2001). Nessa abordagem, reconhece-se o papel desempenhado pelas Associações de Desenvolvimento Local ao nível da realização de políticas sustentáveis de desenvolvimento económico, social, cultural e ambiental. De facto, o desenvolvimento local surgiu pela via prática, quando os agentes locais começaram a enfrentar problemas concretos de desertificação e de decréscimo económico dos territórios.

As ADL resultaram, por um lado, da vontade das pessoas para intervir em aspectos práticos da sua vida, de forma a melhorem a sua qualidade de vida. Por outro lado, procuraram proporcionar serviços na área social e cultural, que não existiriam localmente sem estas organizações. Neste âmbito, os projectos «Começar pelo Princípio» e «O Sonho da Criança» são exemplos do que estas Associações podem concretizar, reflectindo as crescentes preocupações e desafios que cada organização enfrenta na actual sociedade, por um lado, porque os recursos são escassos e as suas necessidades infinitas. Por outro, porque a responsabilidade social das organizações deve ser uma constante preocupação na realização da sua actividade, aumentando a sua competitividade empresarial e atingindo o bem-estar colectivo (David e Abreu, 2004).

Assim, as organizações de âmbito local representam a sociedade civil nos seus diversos aspectos e interesses, assumindo-se muitas vezes como parceiros privilegiados do Estado, consolidando os direitos individuais dos cidadãos face ao poder político e económico. Contudo, o reconhecimento desta responsabilidade social das organizações depende da sua postura ética e dos valores subjacentes ao quadro global das suas decisões.

Assim, à semelhança do que defende Beja (2003), considera-se que a aposta no desenvolvimento sustentável é uma das vantagens competitivas das organizações empresariais nacionais, em geral, e das associativas, em particular. Para isso, importa agir de forma pró-activa, dinamizando as vertentes de ordem económica, social e ambiental.

REFERÊNCIAS BIBLIOGRÁFICAS

ABREU, I. (2005). Responsabilidade Social das Empresas. Disponível em: http://www.naturlink.pt/canais/Artigo.asp?iArtigo=13109&iLingua=1.

ABREU, R. E DAVID, FÁTIMA (2004). *Responsabilidade Social: Um Modelo da Evidência Portuguesa*. Lisboa: X Congresso de Contabilidade, Novembro: 1-16.

AMBA-RAO, S. (1993). Multinational Corporate Social Responsibility, Ethics, Interactions and Third World Governments: An Agenda for the 1990s. *Journal of Business Ethics*, 12: 553-572.

ASSEMBLEIA DA REPÚBLICA (AR, 1991). Lei n.º 108/91, cria o Conselho Económico e Social. *Diário da República,* 188, Série I-A, de 17 de Agosto: 4199-4202.

BEJA, R. (2003). Desenvolvimento sustentável: o imperativo socio-económico e o posicionamento dos Revisores Oficiais de Contas. *Revisores & Empresas*, 22, Julho/Setembro: 7-14.

BILBAO, A. (2008). *Análisis de la Responsabilidad Social Corporativa en las Empresas Argentinas en Pos de Construir un País Sustentable*. Aveiro: XVIII Encuentro AECA, Setembro, 29h.

CARVALHO, F., QUIRÓS, J.T. E ALMEIDA, M.C. (2008). *A influencia da adopção da estrutura da Global Reporting Initiative nas empresas da Euronext de Lisboa.* Aveiro: XII Congresso de Contabilidade e Auditoria, Novembro, 1514-1545.

CENTRO INTERNACIONAL DE FORMAÇÃO (CIF, 2004). A Responsabilidade Social das Empresas e a sua Contribuição ao Desenvolvimento Local. *Working Papers - Documentos de Trabalho do CIF da OIT*, 30.

CENTRO SOCIAL, CULTURAL E RECREATIVO DO BAIRRO DA ESPERANÇA (CSCRBE, 2007). *Iniciativa Comunitária Leader+: Dinamização Rural do Alentejo – Projecto para Aquisição de Equipamentos para Funcionamento da Creche: O Sonho da Criança.* Beja: CSCRBE.

COMISSÃO EUROPEIA (COM, 2000). *Estratégia de Lisboa: Conclusões da Presidência Conselho Europeu de Lisboa 23 – 24 de Março de 2000.* Brussels: Official publications of the European Commission

COMISSÃO EUROPEIA (COM, 2001*). Livro verde - Promover um quadro europeu para a responsabilidade social das empresas, COM (2001) 366 final*. Brussels: Official publications of the European Commission, July 18.

COMISSÃO EUROPEIA (COM, 2002). *Responsabilidade Social das Empresas: Uma contribuição para o Desenvolvimento Sustentado, COM (2002) 347 final.* Brussels: Official publications of the European Commission, July 2.

COMISSÃO EUROPEIA (COM, 2006). *A abordagem Leader: Um guia básico.* Luxemburgo: Serviço das Publicações Oficiais das Comunidades Europeias.

COMUNIDADE EUROPEIA (CE, 1999). Regulamento (CE) nº 1260/1999 do Conselho de 21 de Junho de 1999, que estabelece disposições gerais sobre os Fundos estruturais. *Jornal Oficial das Comunidades Europeias*, L 161, de 26 de Junho.

COMUNIDADE EUROPEIA (CE, 2006). Decisão do Conselho 2006/144/CE, de 20 de Fevereiro de 2006, relativa às orientações estratégicas comunitárias de desenvolvimento rural (Período de programação 2007-2013). *Jornal Oficial das Comunidades Europeias*, L 55, de 25 de Fevereiro: 20-29.

CONSELHO ECONÓMICO E SOCIAL (CES, 2003). *Parecer de iniciativa sobre a Responsabilidade Social das Empresas*. Lisboa: CES.

COSTA, V.C. (2008). *Responsabilidade Social: NP 4469-1: 2008*. Lisboa: 3º Fórum da Responsabilidade Social das Organizações e Sustentabilidade.

CROWTHER, D. E RAYMAN-BACCHUS, L. (2004). The Future of Corporate Social Responsibility. *In* Crowther, D. & Rayman-Bacchus, L. (eds.). *Perspectives on Corporate Social Responsibility*. Aldershot: Ashgate, 229-249.

DAVID, FÁTIMA E ABREU, R. (2004). *Responsabilidade Social: Seu Impacto na Competitividade das Organizações*. Açores: XIV Jornadas Luso-Espanholas de Gestão Científica, Fevereiro, 1415-1424.

DIAS, A., CARVALHO, F. E ALMEIDA, M.C. (2008). *A Adopção da Estrutura da GRI na Divulgação da Informação Voluntária das Empresas Portuguesas*. Aveiro: XVIII Encuentro AECA, Setembro, 80h.

DIAS, A.C.G. (2008). *O Relato da Sustentabilidade Empresarial: Evidências das empresas cotadas em Portugal.* Aveiro: XII Congresso de Contabilidade e Auditoria, Novembro, 1201-1230.

DÍAZ-ZORITA, A. L. E GONZALO, J. A. (1996). *Las responsabilidades del Auditor: Responsabilidad social.* San Sebastián: XI Congreso de Auditores-Censores Jurados de Cuentas de España, Octubre, 3-40.

DIRECÇÃO-GERAL DE AGRICULTURA E DESENVOLVIMENTO RURAL (DGADR, 2000). *Programa Leader+*. Lisboa: DGADR.

DUARTE, M.M.R. (2008). *A Responsabilidade Social e Ambiental das Empresas como Factor Competitivo*. Aveiro: XVIII Encuentro AECA, Setembro, 41h (1-21).

FIGUEIREDO, E. (2003). *Sol na Eira e Chuva no Nabal: Que rural propõem as políticas de desenvolvimento?*. Bragança: V Colóquio Hispano-Português de Estudos Rurais, Outubro, 1-20.

GIRÓ, J. (2002). Globalización sostenible y responsabilidad social. *Dirección y Progreso*, 184: 58-60.

HELENO, M.L.L. (2008). *Avaliação do Impacto da Implementação da Responsabilidade Social através das Relações com os Stakeholders*. Coimbra: Dissertação de Mestrado em Estratégia Empresarial.

INSTITUTO DE DESENVOLVIMENTO RURAL E HIDRÁULICA (IDRH, 2006). *Programa Nacional Leader+*. Lisboa: IDRH.

INSTITUTO PORTUGUÊS DA QUALIDADE (IPQ, 2008). *Sistema de gestão da responsabilidade social. Parte 1: Requisitos e linhas de orientação para a sua utilização*. Lisboa: IPQ.

JONES, T. (1980). Corporate Social Responsibility: Revisited, Redefined. *California Management Review*, 23 (2, Spring): 59-67.

LOURENÇO, N. E MACHADO, C.R. (2008). Mudança Global e Desenvolvimento. Economia, ambiente e conflitos. *Mudança Global Notícias*, 8: 3-6.

LUNARDI, V.R. E CLEMENTE, A.G. (2008). *La Responsabilidad Social Corporativa por las Empresas Cotizadas en el Mercado de Valores de Brasil*. Aveiro: XVIII Encuentro AECA, Setembro, 84h.

MATOS, A.C. E TSUJI, T. (2002). Seminário Internacional "Desenvolvimento Rural e Agricultura Familiar" – Documento síntese. São Luís do Maranhão: Instituto Interamericano de Cooperação para a Agricultura.

MINISTÉRIO DA AGRICULTURA, DO DESENVOLVIMENTO RURAL E DAS PESCAS (MADRP, 2008). *Plano Nacional de Desenvolvimento Rural: 2007-2013. Objectivos estratégicos e princípios orientadores de programação*. Lisboa: MADRP.

MINISTÉRIO DAS FINANÇAS (MF, 2005). Despacho (extracto) n.º 8264/2005 (2.ª série), que homologada a Directriz Contabilística n.º 29, "Matérias ambientais", da Comissão de Normalização Contabilística. *Diário da República*, 75, II Série, 18 de Abril: 6181-6186.

MINISTÉRIO DOS NEGÓCIOS ESTRANGEIROS (MNE, 1978). Aviso que autoriza a publicação do texto em inglês e a respectiva tradução em português da Declaração Universal dos Direitos do Homem. *Diário da República*, 57, I Série, 9 de Março: 488-493.

MINTZBERG, H. (1983). The case for corporate social responsibility. *Journal of Business Strategy*, 4 (2, Fall): 3-15.

OBSERVATÓRIO EUROPEU DOS TERRITÓRIOS RURAIS (OETR, 2007). *Uma selecção das melhores práticas do Leader+*. Bruxelas: Comunidades Europeias.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD, 1999). *OECD Principles of Corporate Governance*. Paris: OECD Publications.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD, 2000). *The OECD Guidelines for Multinational Enterprises-Revision*. Paris: OECD Publications.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS (ONU, 2000). *Declaração do Milénio das Nações Unidas*. Lisbon: United Nations Information Centre.

ORIOL, F. (2002). Responsabilidad social y su impacto en el negocio. *Dirección y Progreso*, 184: 70-72.

PAULA, H.C. (2008). *Relatório de Sustentabilidade: Informações económicas e sociais do Banco do Brasil S.A.* Aveiro: XII Congresso de Contabilidade e Auditoria, Novembro, 653-667.

PRESIDÊNCIA DO CONSELHO DE MINISTROS (PCM, 2007). Resolução do Conselho de Ministros n.º 109/2007, que aprova a Estratégia Nacional de Desenvolvimento Sustentável, ENDS-2015 e Plano de Implementação. *Diário da República*, 159, I Série, 20 de Agosto: 5404-5478.

REGO, A., CUNHA, M.P., COSTA, N.G., GONÇALVES, H. E CABRAL-CARDOSO, C. (2007). *Gestão Ética e Socialmente Responsável: teoria e prática*. Lisboa: Editora RH.

ROVER, O.J. E HENRIQUES, M.A. (2006). A gestão democrática em debate: O programa LEADER e a sua relação com a sociedade civil local. *Revista Crítica de Ciências Sociais*, 75, Outubro: 117-137.

SANTOS, M.J.N. (coord.), Silva, J.L.A., Sampaio, J.J., Henriques, P.L. e Eusébio, C. (2005). *Desenvolvimento Sustentável e Responsabilidade Empresarial*. Lisboa: Celta.

SENHORAS, E.M. (2001). Políticas Públicas nos Anos Noventa: Um Ensaio sobre as Causas e Implicações da Crise do Estado Brasileiro de Bem-Estar Social. *Working Paper*: 1-12.

VIOTTO, P.J. (2006). *Comunicação Empresarial: qualidade de gestão, produtos e serviços - Asperbras, um estudo de caso*. São Paulo: Marília.

TRIBUTAÇÃO SOBRE O PATRIMÓNIO NO DISTRITO DA GUARDA: CONTRIBUTO PARA A DEFINIÇÃO DE UMA ESTRATÉGIA

FÁTIMA DAVID
E-mail: sdavid@ipg.pt
RUTE ABREU
E-mail: ra@ipg.pt
Instituto Politécnico da Guarda
GERMANO CARDOSO
E-mail: pituxa.car@sapo.pt
Direcção de Finanças da Guarda

Resumo

Esta investigação tem um duplo objectivo. Por um lado, analisar o quadro legal subjacente ao Imposto Municipal Sobre Imóveis (IMI) em Portugal, em conformidade com as alterações introduzidas pelo Decreto-Lei nº 287/2003, de 12 de Novembro. Por outro lado, realizar um diagnóstico prospectivo do Distrito da Guarda, com vista a identificar os principais problemas e potencialidades do Distrito e facilitar a definição de uma estratégia económica que permita o seu desenvolvimento. Assim, tendo subjacente que o IMI é um imposto cuja receita é do Município onde se localizam os imóveis e que o mesmo pode alterar as taxas a aplicar, ajustando-as à realidade socio-económica do seu concelho, esta investigação apresenta hipóteses de solução que pretendem contribuir para a sustentabilidade do Distrito da Guarda. De facto, o IMI, enquanto instrumento fiscal dos Municípios, promove a atractividade das populações para a sua área geográfica e contribui para a diminuição da desertificação dos concelhos do interior de Portugal.

Palavras-Chave: Tributação sobre o Património, Fiscalidade, Estratégia Económica.
Área Temática: Estratégia Económica.

Abstract

This research aims to examine the legal framework of the Municipal Property Tax (in Portuguese: *Imposto Municipal Sobre Imóveis* -IMI) in Portugal, with the amendments introduced by Decree-Law nº 287/2003 of 12 November. Additionally, the research provides a prospective diagnosis of the *Guarda district*, in order to identify key problems and potential of the district to facilitate the establishment of an economic strategy that will allow its development. Thus, considering that the IMI is a tax whose revenues go to the *município* where properties are located and that it has the ability to change the rates that will be applied, adjusting them to the socio-economic reality of their *concelho*, this research presents solutions to contribute to the sustainability of the *Guarda* District. The IMI, as a tax instrument, promotes the attractiveness of the geographical area and contributes to reducing the desertification of the inland *concelhos* of Portugal.

Key Words: Property tax, Taxation, Economic strategy.
Thematic Area: Economics of Strategy.

1. INTRODUÇÃO

Ao longo da presente investigação pretende analisar-se o quadro legal, os pressupostos e a aplicabilidade do Imposto Municipal Sobre Imóveis (IMI) em Portugal, em conformidade com as alterações introduzidas pelo Decreto-Lei nº 287/2003, de 12 de Novembro (MF, 2003a). Tendo subjacente que o IMI é um imposto cuja receita é do Município onde se localizam os imóveis e que o mesmo pode alterar as taxas a aplicar, ajustando-as à realidade socio-económica do seu concelho.

Neste sentido, revela-se essencial realizar um diagnóstico prospectivo do Distrito da Guarda, com vista a identificar os seus principais problemas e potencialidades e deste modo definir uma estratégia económica que permita o seu desenvolvimento. Este diagnóstico é dirigido às especificidades da realidade socio-económica de cada um dos catorze concelhos: Aguiar da Beira, Almeida, Celorico da Beira, Figueira de Castelo Rodrigo, Fornos de Algodres, Guarda, Gouveia, Manteigas, Mêda, Pinhel, Sabugal, Seia, Trancoso e Vila Nova de Foz Côa que integram o Distrito da Guarda, agregando as diferenças e destacando as similitudes. Esta investigação efectua uma análise a diversas variáveis, das quais se destacam as taxas do Imposto Municipal Sobre Imóveis (IMI), praticadas em cada um dos catorze concelhos do Distrito da Guarda, uma vez que tais taxas são fixadas anualmente pelas respectivas Assembleias Municipais e posteriormente comunicadas à Direcção Geral dos Impostos (DGCI), para aplicação aos imóveis inscritos em cada um dos concelhos.

A fusão das análises anteriores conduziu à apresentação de hipóteses que contribuem para a sustentabilidade do Distrito da Guarda, garantida a equidade a todos os agentes económicos. De facto, os Municípios através do IMI, enquanto instrumento fiscal destes, promovem a atractividade das populações para a sua área geográfica e desse modo contribuem para a diminuição da desertificação dos concelhos do interior de Portugal, servindo de referência a uma sólida e justa contribuição entre o Estado e esses agentes económicos.

Assim, em consideração aos objectivos desta investigação, a mesma estrutura-se em cinco pontos. À presente introdução seguem-se o ponto dois, relativo à análise histórica da tributação sobre o património, e o ponto três, onde se efectua o diagnóstico prospectivo do Distrito da Guarda, com especial incidência na distribuição da população por cada um dos catorze concelhos do Distrito. No ponto quatro são indicados os contributos para a definição de uma estratégia. Enquanto, no ponto cinco, e último, se tecem algumas considerações finais que pretendem ser o corolário da investigação realizada.

2. TRIBUTAÇÃO SOBRE O PATRIMÓNIO: ANÁLISE HISTÓRICA

Na década de 80 desenvolveu-se a reforma legal dos impostos sobre rendimentos e, ainda, sobre as transmissões de bens e serviços. Contudo, ao nível dos impostos sobre o património muito ficou por fazer. Este sistema tributário refere-se aos impostos sobre o património, especificamente: a contribuição autárquica, o imposto municipal de sisa e o imposto sobre as sucessões e doações, podendo, ainda, considerar-se integrado nesta classificação o imposto de selo incidente sobre actos relativos à transmissão de direitos reais sobre bens imóveis. Assim, com o Decreto-Lei n.º 287/2003 (MF, 2003a) iniciou a referida reforma da tributação sobre o património em Portugal considerando que os impostos sobre o património são: o IMI e o IMT, que incidiu sobre a riqueza evidenciada pela propriedade ou uso e fruição de bens patrimoniais, bem como sobre a transmissão da sua propriedade.

A **Contribuição Autárquica (CA)** era um imposto municipal incidente sobre o valor tributável dos prédios situados no território de cada Município e que recaía sobre os respectivos proprietários ou usufrutuários. A base de tributação era o valor patrimonial dos prédios, sobre os quais insidiam taxas de 0,8%, para os prédios rústicos, ou entre os limites de 0,7% a 1,3%, para os urbanos. Este valor era determinado por avaliações efectuadas de acordo com as regras estabelecidas no Código da Contribuição Predial, aprovado pelo Decreto-Lei n.º 45104/63 (MF, 1963), e, no tocante a terrenos para construção, no Código da Sisa, aprovado pelo Decreto-Lei n.º 41969/58 (MF, 1958) e tendo tido a última alteração pelo Decreto-Lei n.º 442-C/88 (MF, 1988). Na base do valor dos prédios rústicos encontrava-se a capitalização da renda fundiária e na base dos urbanos encontrava-se: a capitalização da renda anual efectiva no caso dos prédios arrendados; ou a que seria possível obter se estivessem arrendados. A avaliação dos terrenos para construção, em bom rigor, não tinha regras, limitando-se à determinação do seu valor transaccional.

Assim, a ausência de regras objectivas acentuou a discricionariedade e subjectivismo das avaliações. Por outro lado, a ausência de mecanismos determinantes da actualização periódica dos valores patrimoniais - sempre apurados por capitalização do rendimento dos prédios no momento em que foram avaliados - determinou que à data da reforma da tributação sobre o património, um «simples apartamento» num prédio recente de construção normal apresentasse um valor tributário várias vezes superior ao de um «palacete» avaliado há várias dezenas de anos, com reflexos, naturalmente, na Contribuição Autárquica suportada, em idêntica proporção, pelos respectivos proprietários. Este imposto foi revogado através do nº 1 do artigo 31º do Decreto-Lei n.º 287/2003 (MF, 2003a).

O **Imposto Municipal de Sisa (IMS)** era devido pelas pessoas que adquiriam prédios a título oneroso. A base de incidência era o preço da transacção, excepto se o valor patrimonial do prédio fosse superior ao valor da tributação, em que se fazia a operação pelo primeiro valor. O devedor do imposto era o comprador, que, além da Sisa, suportava ainda o imposto de selo da escritura. Este imposto foi aprovado Decreto-Lei n.º 41969 (MF, 1958) foi alterado pelo Decreto-Lei n.º 308/91 (MF, 1991) e, ainda, pelos Decreto-Lei n.º 140/92 (MF, 1992) e pela Lei n.º 30-C/92 (AR, 1992). Em 2002, as taxas gerais eram de 8% e 10% para os prédios rústicos e urbanos, respectivamente (AR, 2002). Contudo, os prédios habitacionais beneficiavam de uma isenção de base de € 61.216,00, seguida de vários escalões de taxas reduzidas até ao limite de € 169.376,00, sendo os de valor superior sujeitos à taxa única de 10% (AR, 2002). Nesta área, à semelhança da realidade constatada na Contribuição Autárquica, a desactualização dos valores patrimoniais e a ausência de regras objectivas de avaliação prestava-se à evasão fiscal, justificando-se a necessidade de reforma da tributação. Este imposto foi revogado através do nº 3 do artigo 31º do Decreto-Lei n.º 287/2003 (MF, 2003a).

O **Imposto sobre as Sucessões e Doações (ISD)** recaía sobre os valores de bens, móveis ou imóveis, transmitidos gratuitamente, por herança ou doação e era devido pelos beneficiários. Este imposto foi aprovado pelo Decreto-Lei n.º 41969 (MF, 1958) e sujeito às mesmas rectificações do Imposto Municipal de Sisa. Por outro lado, uma controversa modalidade de pagamento deste imposto era a da "avença". Tratava-se de uma tributação sobre rendimentos de títulos de crédito, de que se destacavam os dividendos distribuídos por sociedades residentes em território português. Estes rendimentos eram sujeitos a uma retenção na fonte pela entidade devedora a título de Imposto sobre o Rendimento das Pessoas Colectivas (IRC) ou de Imposto sobre o Rendimento das Pessoas Singulares (IRS), a que acrescia uma retenção de 5% a título de "imposto sobre as sucessões e doações por avença". A justificação desta tributação encontrava-se numa eventual futura transmissão gratuita desses títulos que, em todo o caso, não seriam novamente tributados quando tal ocorresse. Contudo, segundo o artigo 57.º do Estatuto dos Benefícios Fiscais (EBF), Decreto-Lei n.º 198/2001 (MF, 2001), ficaram isentas de "imposto sobre as sucessões e doações por avença" as obrigações emitidas durante o ano de 1999 até ao final do ano 2002. É, neste contexto que, desde há alguns anos, se vinha equacionando a abolição do ISD, o qual, pró exemplo, desde 1 de Janeiro de 2000, em conformidade com a Lei n.º 3-B/2000 (AR, 2000), deixou de ser aplicado às transmissões a favor de descendentes menores.

Da situação sumariamente descrita destaca-se que o elemento mais relevante na distorção da tributação do património era, justamente, o valor patrimonial dos prédios, tanto em consequência da sua desactualização como da ausência de regras objectivas de avaliação. Assim, uma medida fundamental da reforma teria, necessariamente, de passar pela formulação de regras relativas à avaliação dos prédios urbanos e, por outro lado, pela actualização dos valores patrimoniais dos prédios já avaliados. Um dos aspectos mais relevantes da reforma da tributação do património foi, pois, o que se prendeu com as avaliações prediais, em particular as dos prédios urbanos, tanto no que respeita às regras de determinação do valor patrimonial de cada imóvel como na actualização periódica desse valor. Este imposto foi revogado através do nº 3 do artigo 31º do Decreto-Lei n.º 287/2003 (MF, 2003a).

Por conseguinte, a reforma da tributação sobre o património deu-se com a introdução do **Imposto Municipal sobre Imóveis (IMI)** e registou a última alteração através da Lei n.º 64-A/2008 (AR, 2008). As regras da avaliação constam dos artigos 38.º a 46.º do Código do IMI, deixando-se a avaliação dos prédios urbanos de fazer-se por recurso à capitalização de um rendimento líquido, real ou presumido, passando a ser, tendencialmente, considerado o seu valor de mercado e a "*Direcção-Geral dos Impostos disponibiliza, por via electrónica, à câmara municipal da área da situação do imóvel, a informação relativa ao resultado da avaliação directa de prédios urbanos*" (AR, 2008. 9300-(42)). Enquanto, a actualização trienal dos valores patrimoniais dos prédios urbanos, com base em factores correspondentes a 75% dos coeficientes de desvalorização da moeda encontra-se prevista no artigo 138.º do referido Código.

Para a determinação do valor patrimonial de cada prédio urbano passam a ser considerados elementos objectivos, como sejam a estrutura do imóvel, com especial relevância para a sua arquitectura e materiais utilizados na construção, a sua área e envolvência, bem como a sua afectação a habitação, comércio, indústria ou serviços, a par de outros elementos correctivos como sejam os coeficientes de qualidade e conforto, passando, necessariamente, por um coeficiente de vetustez, relativo à idade do imóvel. A quantificação do valor patrimonial resultará de uma expressão que conjuga todos os elementos referidos. Ao aplicar estes elementos à determinação do valor patrimonial dos prédios urbanos, houve a intenção de que resultassem valores que se situassem entre 80% a 90% dos valores de mercado. Embora a avaliação continue a ser efectuada por uma comissão de avaliadores, será afastado, em grande medida, o subjectivismo que caracterizava o método anterior. Por exemplo, nas avaliações que sejam efectuadas na sequência de uma transmissão onerosa ou gratuita é notificado o alienante para, se assim o entender, requerer segunda avaliação. Naturalmente, o recurso a elementos objectivos dá ao contribuinte a possibilidade de, com alguma aproximação, conhecer o valor que previsivelmente será atribuído ao seu prédio.

Segundo o artigo 32.º do Decreto-Lei n.º 287/2003 (MF, 2003a), o novo regime de avaliações aplica-se a todos os prédios urbanos cujo pedido de inscrição na matriz predial foi entregue a partir do dia seguinte ao da sua publicação. Por outras palavras, este regime passou a ser imediatamente aplicável aos prédios novos, reconstruídos ou melhorados. Mas sempre que haja uma transmissão, gratuita ou onerosa, de um imóvel, é este avaliado de acordo com o novo regime. O valor obtido, além de relevar para a tributação em IMI é, igualmente, relevante para a tributação em IMT ou, se for o caso, para a tributação das transmissões gratuitas. Com particular relevância no plano da luta contra a fraude e a evasão fiscais é de assinalar, ainda, que o novo valor passou a ser considerado na determinação do rendimento tributável do alienante, ou seja, na determinação do lucro fiscal das empresas e na determinação de mais-valias dos particulares.

Desde então, passaram a coexistir valores patrimoniais tributários determinados por recurso às novas regras a par de valores desactualizados e resultantes de avaliações efectuadas há dezenas de anos atrás. Para ultrapassar as desigualdades daí resultantes, o legislador consagrou algumas medidas transitórias, que passaram pela actualização do valor dos prédios antigos e pela diferenciação das taxas aplicáveis a esses prédios e aos que foram avaliados já pelo novo regime.

Como medidas transitórias, o Decreto-Lei nº 287/2003 (MF, 2003a) prevê que os cerca de 6,5 milhões de prédios urbanos inscritos nas matrizes prediais estejam todos avaliados pelo novo regime no prazo máximo de 10 anos, a contar de 1 de Janeiro de 2004. Mas, desde logo foi efectuada uma actualização transitória do valor tributável dos imóveis não avaliados pelas regras do novo regime. Esta actualização, foi imediata e produziu efeitos logo na determinação do imposto em 2004, operando-se pela aplicação ao valor patrimonial tributário actual dos prédios de um coeficiente de correcção monetária fixado entre um máximo de 44,21 e um mínimo de 1, em conformidade com a Portaria n.º 1337/2003 (MF, 2003c).

Tratando-se, porém, de prédios arrendados, o valor a considerar é o que resulta da capitalização da renda anual pela aplicação do factor 12 sempre que o valor assim determinado

fosse inferior ao que resultava da actualização por aplicação do coeficiente de desvalorização da moeda. Para o efeito, os proprietários, usufrutuários ou superficiários, interessados, entregavam uma declaração, nos termos do artigo 17º do Decreto-Lei nº 287/2003 (MF, 2003a), conforme modelo constante da Portaria nº 1283/2003 (MF, 2003b).

A coexistência de valores patrimoniais tributários determinados por recurso a diferentes métodos de quantificação implicou que as taxas aplicáveis aos prédios urbanos fossem fixadas em percentagens variáveis. A competência para fixar as taxas aplicáveis em cada ano continuou a ser das Assembleias Municipais que, relativamente a esses prédios, podem escolher entre uma percentagem de 0,4% a 0,8%, aplicável aos prédios ainda não avaliados de acordo com as novas regras ou de 0,2% a 0,5% relativamente aos avaliados. Adicionalmente, podem, ainda, os Municípios estabelecer reduções e majorações das taxas de tributação por eles fixadas, em função de determinados parâmetros.

Na sequência da reforma levada a cabo, o imposto municipal de sisa passou a ser designado de **Imposto Municipal sobre Transmissões onerosas de imóveis (IMT)** e foi aprovado pelo Decreto-Lei n.º 287/2003, (MF, 2003a), mantendo-se o núcleo essencial de regras instrumentais do anterior diploma em vigor. As inovações relevantes foram as que se prendem com a fraude e evasão fiscais que caracterizavam este imposto, por outro lado, com um consequente abaixamento generalizado das taxas.

Desde logo, sempre que ocorra uma transmissão sujeita a este imposto, há lugar à avaliação do imóvel transmitido, por recurso às novas regras objectivas de avaliação, prevendo-se que o valor, assim, obtido se aproxime do valor efectivo de mercado. Recorda-se que o valor patrimonial, na economia deste tributo, é subsidiário do preço, relevando apenas quando este seja inferior àquele. O valor de avaliação passou, também, a ser considerado como valor mínimo da operação no que respeita à determinação do lucro das empresas vendedoras, podendo estas, contudo, provar que o valor efectivo da transacção foi inferior àquele. Com efeito, é de realçar que as taxas gerais do IMT desceram para 5% e 6%, face aos 8% e 10% aplicados no âmbito do IMS, consoante se trate de prédios rústicos ou urbanos, respectivamente. E, e elevaram-se os limites da isenção de base e dos diversos escalões, sobre os quais são aplicáveis as taxas, quando se trate de prédios habitacionais (MF, 2003a).

De entre todos os condicionalismos referidos e à luz do quadro legal, cada Município fixa as taxas dentro de critérios de razoabilidade. Porém, em relação as taxas praticadas há uma descida significativa dos seus valores nominais, simultaneamente com actualização de escalões logo implica uma diminuição da carga fiscal. È, neste espectro de relações que se pretende diagnosticar a situação económica de cada um dos 14 concelhos do Distrito da Guarda, o que será efectuado no ponto seguinte, de forma a podermos apurar qual é a tendência de cada Município, tendo em conta os indicadores sócio-económicos de cada um dos concelhos do Distrito da Guarda.

3. DISTRITO DA GUARDA: DIAGNÓSTICO PROSPECTIVO

O diagnóstico prospectivo do Distrito da Guarda desenvolve-se segundo grandes áreas de investigação, nomeadamente: geografia, organização administrativa, demografia e economia, com vista a identificar os seus principais problemas e potencialidades, e de forma a definir uma estratégia económica que permita o seu desenvolvimento.

Em termos geográficos, o Distrito da Guarda ocupa uma área de cerca de 5.535,5 Km^2, correspondendo a cerca de 6% da área total do País, sendo o sétimo maior Distrito do Continente em termos de dimensão conforme se constata no **Quadro I**. O Distrito da Guarda está localizado no centro interior de Portugal, tendo como limite a norte o Distrito de Bragança, a leste a Espanha, a sul o Distrito de Castelo Branco e a oeste os Distritos de Viseu e Coimbra, conferindo-lhe uma posição periférica relativamente às áreas mais desenvolvidas do País.

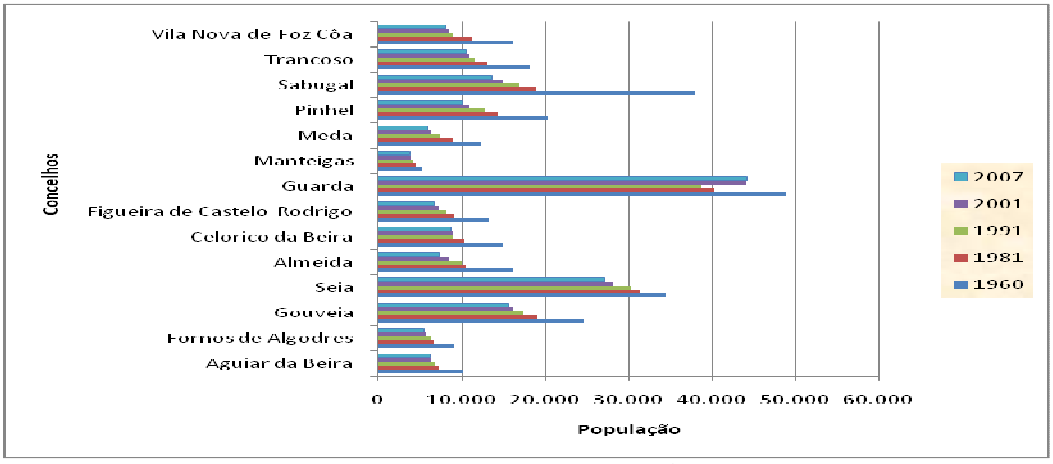
Quadro I. Área, perímetro, extensão máxima e altimétrica por Município, 2007

	Área	Perímetro	Comprimento máximo		Altitude	
			Norte-Sul	Este-Oeste	Máxima	Mínima
	km²	km			m	
Portugal	**92 090,1**	**4 071**	**1 400**	**2 200**	**2 351**	**0**
Continente	**88 967,1**	**2 731**	**576**	**281**	**1 993**	**0**
Centro	**28 200,4**	**1 320**	**235**	**234**	**1 993**	**0**
Distrito da Guarda	**5 535,3**	**1 584,4**	**-**	**-**	**1 993**	**84**
Aguiar da Beira	206,8	67	20	17	988	450
Fornos de Algodres	131,5	79	22	14	913	325
Gouveia	300,6	98	25	26	1 619	250
Seia	435,7	118	34	23	1 993	175
Almeida	518,0	119	34	22	844	500
Celorico da Beira	247,2	81	22	22	1 255	375
Figueira de Castelo Rodrigo	508,6	112	32	27	976	125
Guarda	712,1	162	36	37	1 286	442
Manteigas	122,0	65	17	16	1 993	519
Meda	286,1	100	24	21	944	225
Pinhel	484,5	156	43	26	925	150
Sabugal	822,7	182	33	47	1 215	450
Trancoso	361,5	121	26	24	982	425
Vila Nova de Foz Côa	398,2	126	29	31	813	84

Fonte: INE (2008a; 2008b).

Em termos de organização administrativa, o Distrito é constituído por 336 freguesias e 14 concelhos. Da descrição efectuada ressalta que, se tem um Distrito com catorze concelhos, que, por sua vez, ainda integra quatro sub-regiões de NUTS II, confirmando a sua dispersão geográfica e diferenciação regional. Esta última organização permitiu identificar que: o concelho de Aguiar da Beira integra a sub-região de Dão-Lafões; o concelho de Fornos de Algodres, Gouveia e Seia pertencem à sub-região designada de Serra da Estrela; os concelhos de Almeida, Celorico da Beira, Figueira de Castelo Rodrigo, Guarda, Manteigas, Meda, Pinhel, Sabugal e Trancoso englobam a sub-região Beira Interior Norte. Por último, o concelho de Vila Nova de Foz Côa faz parte da sub-região do Douro, que se situa ao Norte de Portugal.

Gráfico I. Evolução da População do Distrito da Guarda, 1960-2007



Fonte: INE (2008a; 2008b).

Em termos populacionais e de acordo com o **Gráfico 1**, o Distrito da Guarda regista continuadamente uma evolução demográfica recessiva, como resultado de elevadas percentagens do êxodo migratório. Segundo David (1996: 8), foi ainda nos anos 70, com a mudança de regime político nacional, que:

"*se assistiu ao declínio da emigração e ao aumento da imigração como resultado do regresso das ex-colónias, por via da mudança dos regimes políticos nesses países, e do retorno dos emigrantes, como resultado das alterações económico-sociais nos países europeus. Todavia, as consequências do fenómeno migratório continuaram-se a sentir, por efeito de arrastamento, nos anos seguintes, embora com menores quebras populacionais. Também como consequência da emigração houve a alteração das características sociais da população*".

O narrado fenómeno migratório no Distrito da Guarda encontra a sua explicação na escassa oferta do mercado de trabalho, devido à estagnação do investimento estatal nas regiões do interior do País, por via da alteração do investimento para as áreas metropolitanas que então começavam a surgir (David, 1996). Assim, o fluxo migratório foi caracterizado por dois movimentos: um interno para o litoral e principais centros urbanos, particularmente para Lisboa; e outro externo para a Europa, concretamente para França e Alemanha (David, 1996). Praticamente ao longo dos anos, a anterior evolução demográfica em não se alterou, excepto no que se refere ao concelho da Guarda. Segundo os dados de 2007, o peso relativo da população residente aumentou para 25,7% do total de população residente do Distrito desse ano. Enquanto que, em 1960, o peso relativo da população residente registava apenas 17,4% do total de população residente do Distrito nesse ano.

Quadro II. Peso relativo da população Residente no Distrito da Guarda, 1960-2007

Distrito da Guarda	**1960**	**1981**	**1991**	**2001**	**2007**
Aguiar da Beira	3,62%	3,54%	3,57%	3,47%	3,61%
Fornos de Algodres	3,20%	3,21%	3,33%	3,13%	3,09%
Gouveia	8,74%	9,26%	9,25%	8,96%	9,00%
Seia	12,21%	15,25%	16,14%	15,64%	15,72%
Almeida	5,71%	5,12%	5,34%	4,68%	4,19%
Celorico da Beira	5,29%	4,99%	4,72%	4,93%	5,02%
Figueira de Castelo Rodrigo	4,69%	4,44%	4,31%	3,98%	3,85%
Guarda	17,37%	19,63%	20,60%	24,50%	25,65%
Manteigas	1,87%	2,18%	2,23%	2,13%	2,16%
Meda	4,39%	4,36%	3,95%	3,47%	3,36%
Pinhel	7,20%	6,97%	6,75%	6,09%	5,81%
Sabugal	13,50%	9,20%	8,99%	8,26%	7,85%
Trancoso	6,46%	6,37%	6,10%	6,05%	6,05%
Vila Nova de Foz Côa	5,75%	5,47%	4,72%	4,72%	4,66%
População Total do Distrito	**282.037**	**205.631**	**188.165**	**179.962**	**172.304**

Fonte: INE (2008a; 2008b).

Observando as importantes disparidades entre os diversos concelhos, expressas no **Quadro II**, verifica-se que os concelhos de fronteira com Espanha, em regra, apresentam densidades populacionais inferiores a metade da densidade dos concelhos mais ocidentais do Distrito da Guarda (David, 1996). De acordo com os censos de 2001, que confirma a realidade diagnosticada nos censos de 1991, o concelho de Figueira de Castelo Rodrigo detém menor densidade populacional, enquanto os concelhos da Guarda e de Seia são os que apresentam maiores densidades populacionais e, no seu conjunto, concentram cerca de 40% da população do Distrito (**Quadro II**). Ainda, de acordo com os censos de 2001, a população residente era de 179.962 habitantes (INE, 2002), representando cerca de 1,8% do total da população de Portugal continental. A mesma relação baixou para 1,7%, em 2007 (INE, 2008). É o quinto Distrito menos populoso do País, com uma densidade populacional em 2007 de 31,12 habitantes por Km^2, situando-se, em termos populacionais, entre os Distritos de Castelo Branco e de Évora. Tendo em conta os dados dos anuários estatísticos da região Norte e da região Centro, referentes ao ano de 2007, a população do Distrito da Guarda totalizava 172.304 habitantes (INE, 2008).

Assim, no período 2001-2007 regista-se uma diminuição de cerca de 4,2%, em apenas 6 anos. Esta tendência tem vindo sucessivamente a agravar-se no tempo, tal como se demonstra pelo **Quadro II**. No período 1991-2001 alcançou uma diminuição de 4,4%, em 10 anos de análise.

Quadro III. Distribuição por escalão etário da população Residente nos concelhos do Distrito da Guarda, 2007

	Total	0 a 14 anos	15 a 24 anos	25-64 anos	65-74 anos	75 e mais anos
Portugal	**10.617.575**	**1.628.852**	**1.236.004**	**5.902.888**	**1.004.160**	**845.671**
Continente	**10.126.880**	**1.538.369**	**1.163.561**	**5.637.606**	**969.412**	**817.932**
Centro	**2.385.911**	**335.161**	**272.769**	**1.292.583**	**253.081**	**232.317**
Distrito da Guarda	**172.304**	**20.058**	**20.114**	**88.929**	**20.676**	**22.527**
Aguiar da Beira	6.225	750	858	3.133	731	753
Fornos de Algodres	5.317	632	649	2.619	678	739
Gouveia	15.505	1.598	1.785	7.823	2.070	2.229
Seia	27.082	3.143	3.276	15.031	3.032	2.600
Almeida	7.213	676	755	3.642	1.012	1.128
Celorico da Beira	8.643	1.061	1.009	4.303	977	1.293
Figueira de Castelo Rodrigo	6.638	688	758	3.202	894	1.096
Guarda	44.191	6.229	5.238	23.976	4.237	4.511
Manteigas	3.714	422	467	2.052	443	330
Meda	5.788	598	676	2.792	756	966
Pinhel	10.009	1.056	1.055	5.047	1.373	1.478
Sabugal	13.533	1.137	1.390	6.197	1.991	2.818
Trancoso	10.422	1.226	1.310	5.135	1.372	1.379
Vila Nova de Foz Côa	8.024	842	888	3.977	1.110	1.207

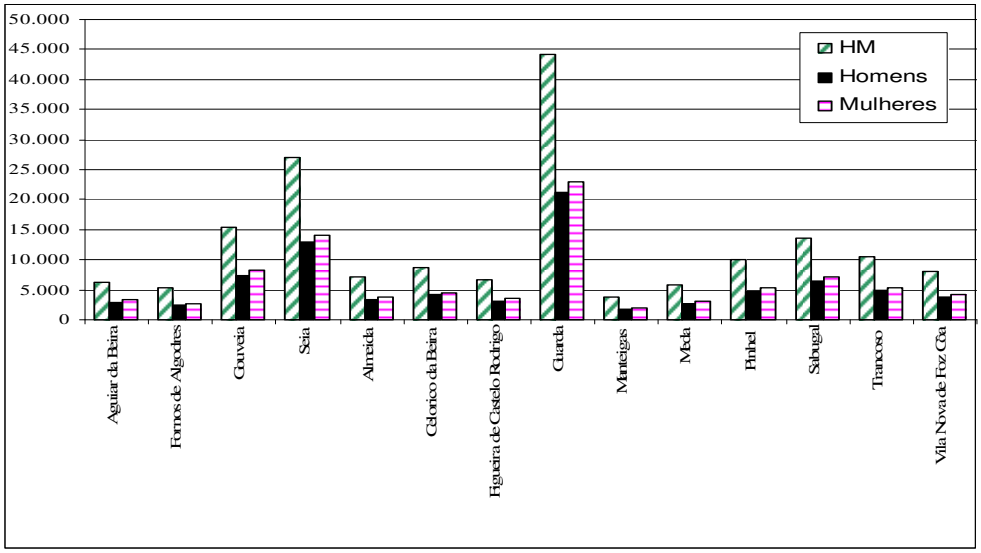
Fonte: INE (2008a; 2008b).

Em paralelo ao referido e tal como pode ser confirmado pelo **Quadro III**, na década de 80, assistiu-se ao fenómeno do envelhecimento da população, apontando-se como argumento justificativo a redução da natalidade ao invés do prolongamento da esperança de vida ou da redução da mortalidade. Esta última também incide sobre a mortalidade infantil, permitindo ganhos de vida simultâneos na infância e na velhice (David, 1996). Por exemplo, em Portugal Continental registou-se, por um lado, a redução da percentagem de jovens com menos de 14 anos de 19,7% em 1991 para 16,0% em 2001 e 15,2% em 2007; e, por outro lado, a um aumento da percentagem de idosos com mais de 65 anos de 13,7%, para 16,4% e para 17,6%, respectivamente, nos mesmos períodos (INE, 1993; 2002; 2008). Esta mesma realidade, de redução da população jovem e aumento da população idosa, ocorreu no período 1991-2001 nos concelhos do Distrito da Guarda, com excepção de ligeiras oscilações nos concelhos de Almeida, Gouveia, Guarda, Manteigas e Seia. Assim, é possível concluir que a quebra da natalidade e o crescimento natural estão actualmente ligados à modernização das sociedades, ou seja, ao aumento dos rendimentos económicos, às oportunidades de mobilidade social e à elevação do nível de instrução (David, 1996).

A distribuição por escalão etário da população residente no Distrito da Guarda, no final de 2007, expressa no **Quadro III**, será conjugada com a subdivisão por género, conforme se apresenta no **Gráfico II.** Assim, conclui-se que em todos os concelhos do Distrito da Guarda a média da população residente feminina (52,4%) é superior à masculina (47,6%). Contudo, nos concelhos de Manteigas (53,8%), Almeida (53,2%), Sabugal (52,9%), Mêda (52,6%) e Vila Nova de Foz Côa (52,5%) são registados valores médios da população residente feminina superiores à masculina. Segundo os grandes grupos etários retira-se que a população até aos 24 anos representa, em termos médios, 23,4% do total contra a população com mais de 65 anos

25,3% do total, retirando como consequência directa o seu envelhecimento. Com os concelhos do Sabugal (35,6%), Almeida (30,0%) e Figueira Castelo Rodrigo (29,8%) a registarem valores muito superiores à média da população com mais de 65 anos e Guarda (25,9%), Aguiar da Beira (25,8%) e Trancoso (24,3%) a registarem valores muito superiores à média da população com menos de 24 anos.

Gráfico II. Distribuição por concelho da população residente segundo o género



Fonte: INE (2008a; 2008b).

A apresentar variações entre os diversos concelhos encontra-se o coeficiente de dependência. Este representa a importância da distribuição da população por grupos etários, uma vez que quanto maior for este índice maior desequilíbrio se regista entre os conjuntos de pessoas que produzem e não produzem riqueza. Este rácio resulta do somatório da população com menos de 14 anos e da população com mais de 65 anos pela população com idade compreendida entre os 15 e os 64 anos (CCRC, 1986b).

Assim, os concelhos em que os escalões menos produtivos apresentam valores mais elevados correspondem aos de Sabugal e de Celorico da Beira, sendo considerados como demograficamente mais dinâmicos, por possuírem os índices mais baixos, os concelhos de Manteigas e Guarda, quer no ano de 1991, quer no ano 2001 (INE, 1993; 2002). Efectuando-se a mesma análise com referência a 2007, nos concelhos mais dinâmicos mantêm-se Manteigas (47%) e aparece agora Seia (48%) e nos menos produtivos continua o Sabugal (78%) e entra Figueira de Castelo Rodrigo (68%).

Em termos de desenvolvimento económico, no **Quadro IV** observa-se que, no ano de 2006, a média da densidade de empresas no Distrito da Guarda (2,5 empresas/km^2) é, substancialmente, inferior à média da região Centro (8,4 empresas/km^2), do Continente (11,7 empresas/km^2) e de Portugal (11,8 empresas/km^2). Por outro lado, a dimensão dessas empresas concentra-se nas micro-empresas com um valor médio de 96,8%, sendo superior quer à região Centro (95,8%), quer ao Continente (95,5%) e, ainda, a Portugal (95,5%), tendo como consequência directa um valor médio baixo de pessoas ao serviço (2,3) e um valor médio reduzido do volume de negócios (115,7 milhares de euros).

A repercussão de todos estes indicadores comprova o reduzido desenvolvimento económico, a severidade e a persistência de dificuldades económicas neste Distrito, para além do impacto negativo da economia portuguesa. A Comissão de Coordenação da Região Centro (CCRC, 1986a: 361) considera que este facto deriva de uma:

"*estrutura produtiva ineficaz onde a interioridade da área e a sua consequente má acessibilidade ao todo nacional são o principal bloqueio ao desenvolvimento, à inovação, à captação do investimento e à existência paralela de elevados índices de carências sociais que, no seu conjunto, não permitem atingir os padrões mínimos de qualidade de vida de uma população*".

Quadro IV. Distribuição por concelho do Distrito da Guarda de diversos Indicadores das empresas, 2006

	Densidade de empresas	Proporção de micro-empresas	Proporção de pequenas e médias empresas	Pessoal ao serviço por empresa	Volume de negócios por empresa	Indicador de concentração do volume de negócios das 4 maiores empresas
	N.º/km²	%		N.º	Milhares de euros	%
Portugal	**11,8**	**95,4**	**4,5**	**3,4**	**305,5**	**5,6**
Continente	**11,7**	**95,5**	**4,5**	**3,4**	**305,4**	**5,9**
Centro	**8,4**	**95,8**	**4,1**	**2,9**	**213,8**	**4,1**
Distrito da Guarda						
Aguiar da Beira	2,4	97,4	2,6	2,1	134,8	25,8
Fornos de Algodres	3,1	95,8	4,2	2,4	99,0	16,1
Gouveia	3,8	96,0	4,0	2,5	95,9	19,4
Seia	4,9	95,5	4,4	2,8	156,2	34,4
Almeida	1,1	96,7	3,3	2,2	111,4	25,0
Celorico da Beira	2,4	96,1	3,9	2,6	109,4	23,6
Figueira de Castelo Rodrigo	0,8	98,1	1,9	2,1	95,3	27,2
Guarda	6,1	96,6	3,3	2,6	172,8	22,2
Manteigas	2,3	96,4	3,6	2,4	72,1	34,7
Meda	1,3	96,5	3,5	2,1	109,6	27,6
Pinhel	1,8	97,5	2,5	2,0	125,9	24,6
Sabugal	1,3	97,7	2,3	2,0	95,9	25,6
Trancoso	2,3	96,3	3,7	2,5	140,3	27,4
Vila Nova de Foz Côa	1,5	98,1	1,9	1,8	92,5	25,3

Fonte: INE (2008a; 2008b).

Relativamente à actividade económica, constata-se que para a média dos concelhos do Distrito da Guarda, quer em 1991, quer em 2001, o sector mais importante em termos de emprego foi o terciário nos concelhos de Almeida, Celorico da Beira e Guarda. Os serviços, em especial os públicos, têm vindo a constituir, particularmente nos centros urbanos, uma fonte crescente de emprego e, previsivelmente, de riqueza. De facto, a evolução dos diferentes sectores da actividade económica tem apontado para uma mudança na estrutura produtiva do Distrito, isto é, a quebra do sector primário em detrimento do aumento do sector secundário e terciário.

Esta situação torna claro que houve um abandono das actividades agrícolas para a indústria e para os serviços sob a forma de operários não especializados (David, 1996). No entanto, a imagem do Distrito da Guarda, enquanto uma região subdesenvolvida, ressalta como consequência do peso que o sector primário possui e da insignificância do secundário, sendo mesmo abaixo do terciário.

Quadro V. Distribuição por concelho do Distrito da Guarda de número de empresas por sector económico agregado segundo a CAE-Rev 2.1, 2006

	Total	A	C	D	E	F	G	H	I	K	M	N	O	Pessoal ao serviço	Volume de negócios
	N.º														milhares de euros
Portugal	**1 085 435**	**4 984**	**1 565**	**97 958**	**704**	**122 070**	**298 593**	**87 478**	**29 554**	**223 549**	**59 195**	**74 959**	**84 826**	**3 738 983**	**331 631 797**
Continente	**1 044 450**	**4 386**	**1 517**	**94 980**	**687**	**116 457**	**288 334**	**83 860**	**27 393**	**214 717**	**57 934**	**72 506**	**81 679**	**3 593 213**	**318 938 679**
Centro	**237 907**	**1 085**	**608**	**22 932**	**150**	**36 762**	**69 689**	**19 000**	**6 762**	**39 145**	**13 186**	**13 605**	**14 983**	**700 261**	**50 868 556**
Distrito da Guarda		**6**	**72**	**1 237**	**12**	**2 168**	**4 245**	**1 664**	**577**	**1 717**	**965**	**610**	**787**	**34 244**	**1 908 372**
Aguiar da Beira	493	0	3	51	1	79	174	62	40	32	14	9	28	1 015	66 459
Fornos de Algodres	409	0	0	45	0	66	118	54	35	42	20	6	23	988	40 503
Gouveia	1 130	0	3	116	3	153	356	132	37	139	93	35	63	2 829	108 311
Seia	2 122	0	3	240	1	299	692	211	65	260	143	93	115	5 952	331 439
Almeida	553	0	10	39	0	75	179	81	46	50	19	18	36	1 238	61 622
Celorico da Beira	586	0	2	44	1	73	177	98	28	72	33	23	35	1 529	64 101
Figueira de Castelo Rodrigo	411	1	4	51	1	61	128	62	14	34	10	16	29	862	39 174
Guarda	4 366	1	4	262	2	624	1 175	408	131	737	453	295	274	11 344	754 445
Manteigas	280	0	0	26	0	45	74	55	8	30	13	9	20	659	20 188
Meda	375	0	3	50	0	60	115	54	25	30	13	11	14	790	41 100
Pinhel	876	1	28	82	0	173	288	82	28	90	38	33	33	1 781	110 265
Sabugal	1 037	1	2	99	0	265	283	160	52	74	34	28	39	2 105	99 420
Trancoso	833	0	3	76	1	135	281	106	44	77	50	17	43	2 071	116 874
Vila Nova de Foz Côa	589	2	7	56	2	60	205	99	24	50	32	17	35	1 081	54 472

Fonte: INE (2008a; 2008b).

No **Quadro V** é possível aferir qual é o sector da actividade económica dominante no Distrito da Guarda, segundo a Classificação da Actividade Económica, revisão 2.1. Assim, no ano de 2006, o Comércio por grosso e a retalho, reparação de automóveis e de bens de uso pessoal e doméstico (G) registou 30,2% do total do Distrito, seguido pelo sector da Construção (F) com 15,4% do total do Distrito e, ainda, do Alojamento, Restauração e Similares (H) com 11,8% do total do Distrito.

Esta tendência foi influenciada pelo contexto de incerteza que caracteriza a actual conjuntura macroeconómica e por perspectivas pouco favoráveis para a actividade económica, com as respectivas repercussões no mercado de trabalho (MFAP, 2009). Os concelhos de Guarda, Seia e Gouveia foram os mais representativos no sector do comércio e os concelhos da Guarda, Seia e Sabugal foram os mais numerosos em empresas com sede no Distrito nos restantes dois sectores. A turbulência no sector da construção identificou como os menos numerosos em empresas os concelhos da Mêda, Vila Nova de Foz Côa e Figueira de Castelo Rodrigo.

Quadro VI. Distribuição por concelho do Distrito da Guarda dos Contratos de Compra e Venda segundo a natureza dos Prédios, 2007

	Total de prédios		Prédios urbanos				Prédios rústicos		Prédios mistos	
			Total		Em propriedade horizontal					
	N.º	milhares de euros	N.º	milhares de euros	N.º	milhares de euros	N.º	milhares de euros	N.º	milhares de euros
Portugal	**281 365**	**29 630 074**	**210 892**	**26 236 033**	**145 245**	**16 708 444**	**66 173**	**2 340 664**	**4 300**	**1 053 377**
Continente	**265 314**	**28 323 769**	**201 736**	**25 149 445**	**139 904**	**16 007 698**	**59 619**	**2 165 760**	**3 959**	**1 008 564**
Centro	**71 927**	**4 132 245**	**39 125**	**3 552 302**	**20 806**	**1 962 774**	**31 588**	**433 692**	**1 214**	**146 251**
Distrito da Guarda	**5 192**	**126 174**	**2 198**	**104 190**	**625**	**49 608**	**2 929**	**17 948**	**65**	**4 036**
Aguiar da Beira	219	2 333	66	1 878	19	1 208	153	456	0	0
Fornos de Algodres	161	2 364	61	1 614	15	770	97	657	3	93
Gouveia	366	7 650	140	5 226	28	2 039	195	677	31	1 747
Seia	601	19 786	273	14 456	84	5 768	317	4 399	11	932
Almeida	424	3 079	90	2 596	4	432	333	483	1	ə
Celorico da Beira	307	7 201	153	5 885	45	2 989	150	1 126	4	190
Figueira de Castelo Rodrigo	421	3 404	129	2 558	10	282	292	845	0	0
Guarda	929	50 868	583	46 690	314	28 752	344	3 945	2	233
Manteigas	64	1 737	40	1 663	2	137	22	65	2	10
Meda	248	3 254	91	2 564	6	600	155	643	2	48
Pinhel	393	6 245	153	5 687	28	2 295	240	558	0	0
Sabugal	372	4 227	141	3 413	11	834	230	763	1	50
Trancoso	343	6 193	160	5 403	37	1 689	180	442	3	348
Vila Nova de Foz Côa	344	7 833	118	4 557	22	1 811	221	2 890	5	387

Fonte: INE (2008a; 2008b).

No **Quadro VI** é apresentada a distribuição por concelho do Distrito relativa aos contratos de compra e venda de prédios, em número e valor. Segundo a sua natureza, a nível de Portugal e do Distrito da Guarda registam-se valores diferentes, respectivamente, 75,0% e 42,3% nos prédios urbanos, 23,5% e 56,4% nos prédios rústicos e 1,5% e 1,3% nos prédios mistos. Estas tendências justificam-se na característica de urbanidade no caso de Portugal e de interioridade no caso do Distrito da Guarda.

Complementarmente, no Distrito da Guarda foram transaccionados 5.192 prédios através de contratos de compra e venda, com um valor total de 126.174 milhares de euros, implicando um valor médio de 24,30 milhares de euros por cada prédio transaccionado. Este valor é muito inferior ao valor médio apurado para o Continente, que é de 106,76 milhares de euros por cada prédio transaccionado. Além disso, por cada prédio transaccionado, o concelho da Guarda regista o maior valor de 54,76 milhares de euros e o concelho de Almeida regista o menor valor com 7,26 milhares de euros. Note-se, contudo, que parece existir evidencia que "*os preços no mercado imobiliário para habitação não estão sobrevalorizados*" (BP, 2007: 47). Muito interessante é o caso do concelho de Seia que apresenta em todas as variáveis (número de prédios e valor) e para todas as naturezas de prédios (urbanos, rústicos e mistos) os valores mais significativos. Similar tendência é verificada no concelho da Guarda, exceptuando-se os prédios mistos. Em oposição, o concelho de Manteigas inscreveu os valores menos significativos, talvez por se encontrar no centro do Parque Natural da Serra da Estrela, rodeado pela vegetação frondosa da Beira Interior.

Numa perspectiva similar, no **Quadro VII** são analisados os contratos de mútuo com hipoteca voluntária, para todas as naturezas de prédios (urbanos, rústicos e mistos). No ano de 2007, o volume desses contratos ascenderam a 2.411 prédios, com um valor total de 224.221 milhares de euros, correspondendo um valor médio de 93 milhares de euros por cada prédio transaccionado em mútuo, com hipoteca voluntária. Além disso, por cada prédio transaccionado, o concelho da Almeida regista o maior valor de 115,73 milhares de euros e o

concelho de Mêda regista o menor valor com 48,12 milhares de euros. Todos os valores são muito inferiores ao valor apurado para Portugal que é de 125,69 milhares de euros por cada prédio transaccionado em mútuo.

Quadro VII. Distribuição por concelho do Distrito da Guarda dos Contratos de mútuo, com hipoteca voluntária, segundo a natureza dos Prédios, 2007

	Total de prédios		Prédios urbanos				Prédios rústicos		Prédios mistos	
			Total		Em propriedade horizontal					
	N.º	milhares de euros	N.º	milhares de euros	N.º	milhares de euros	N.º	milhares de euros	N.º	milhares de euros
Portugal	**301 564**	**37 903 696**	**285 520**	**34 770 708**	**199 651**	**20 288 825**	**10 312**	**1 629 955**	**5 732**	**1 503 033**
Continente	**287 405**	**35 862 261**	**272 233**	**32 960 487**	**193 071**	**19 632 093**	**9 775**	**1 456 752**	**5 397**	**1 445 023**
Centro	**60 970**	**6 355 606**	**53 397**	**5 553 555**	**28 973**	**2 511 244**	**5 526**	**385 367**	**2 047**	**416 684**
Distrito da Guarda	**2 411**	**224 221**	**2 066**	**196 121**	**818**	**68 160**	**259**	**16 054**	**86**	**12 046**
Aguiar da Beira	72	6 006	52	4 172	14	853	19	1 614	1	220
Fornos de Algodres	67	5 100	53	4 109	16	1 161	10	423	4	568
Gouveia	197	20 952	164	16 898	41	2 943	13	1 439	20	2 615
Seia	429	43 814	369	37 008	135	10 606	42	4 068	18	2 739
Almeida	62	7 175	52	5 800	1	70	7	765	3	610
Celorico da Beira	137	12 406	113	9 716	45	3 587	13	866	11	1 825
Figueira de Castelo Rodrigo	73	6 337	68	6 054	8	655	3	105	2	178
Guarda	768	76 798	743	73 509	437	39 181	16	2 014	9	1 276
Manteigas	46	4 085	44	3 845	4	226	1	225	1	15
Meda	102	4 908	45	4 208	8	622	57	700	0	0
Pinhel	118	10 708	106	9 215	35	2 804	9	1 056	3	437
Sabugal	95	6 553	80	5 374	12	1 023	10	753	5	426
Trancoso	95	8 020	85	7 244	29	1 870	7	367	3	410
Vila Nova de Foz Côa	150	11 356	92	8 969	33	2 559	52	1 659	6	728

Fonte: INE (2008a; 2008b).

Dos dados constantes dos **Quadros VI e VII** poder-se-á concluir que os valores médios dos contratos são, completamente, dispares. Resulta assim que, muito embora possa existir alguma subvalorização dos contratos de compra e venda em relação aos contratos de mútuo com hipoteca voluntária, a realidade evidencia a problemática do endividamento das famílias bastante elevado. De acordo com os dados anteriores, no ano de 2007, 56,2% do volume dos contratos de compra e venda são representados por contratos de mútuo com hipoteca voluntária a avaliar pelo volume de contratos realizados. Além disso, apesar de aparentemente o grau de risco poder ser elevado, na verdade o reembolso do endividamento está, em grande medida, garantido pela hipoteca do imóvel.

Em sequência do anterior volume contratual, no **Quadro VIII**, apresenta-se a distribuição por concelho do Distrito da Guarda relativa às receitas provenientes de vários impostos: IMI, IMT e IRS. Por conseguinte, assiste-se, por um lado, a uma bipolarização das receitas na Guarda e em Seia, representando mais de 50% do total de receitas do Distrito, qualquer que seja o imposto em referência, em resultado da elevada densidade populacional e do número de residentes, levando a que existam mais domicílios e alojamentos, e logo um maior número de imóveis para tributar em sede de IMI e IMT. Por outro lado, verifica-se um equilíbrio relativo nos restantes concelhos, com excepção do concelho de Manteigas, cujas receitas fiscais representam apenas 1% do total de receitas de IMI e de IMT e cerca de 2% do total de IRS apurado no Distrito.

Quadro VIII. Distribuição por concelho do Distrito da Guarda das Receitas do IMI, IMT e IRS, 2006

Concelho	IMI (Euro)	IMT (Euro)	IRS (Euro)
Aguiar da Beira	161.100	50.450	1.044.842
Almeida	241.123	76.907	2.489.790
Celorico da Beira	323.854	107.962	2.113.321
Figueira de Castelo Rodrigo	160.485	123.277	1.614.365
Fornos de Algodres	117.874	60.316	1.335.942
Gouveia	508.604	213.260	4.272.092
Guarda	2.603.374	953.816	25.369.197
Manteigas	71.449	30.038	888.304
Meda	139.362	75.451	1.502.317
Pinhel	292.677	84.842	2.653.935
Sabugal	224.497	140.394	2.980.643
Seia	1.037.838	418.875	8.346.991
Trancoso	574.537	92.878	2.598.668
Vila Nova de Foz Côa	168.133	242.537	2.138.177

Fonte: Adaptado de ANMP (2008).

Analisados os valores recebidos do IRS por cada concelho, verifica-se que os concelhos com maior volume de receita são os concelhos da Guarda, Seia e Gouveia, na medida em que serão aqueles onde é obtido maior rendimento. Também, quando analisada a receita em sede de IMI e de IMT, os concelhos de Guarda e Seia são os que arrecadam maior valor de receita fiscal. Em clara oposição, ao concelho de Manteigas e de Aguiar da Beira que registam os valores mais reduzidos, logo detectam-se situações de vulnerabilidade económica, sobretudo no IRS, o que não deixa margem de manobra para se adaptarem e se desenvolverem economicamente.

Quadro IX. Distribuição por concelho do Distrito da Guarda do ganho médio mensal dos trabalhadores por conta de outrem nos estabelecimentos, segundo o nível de habilitações, 2006

	Total	Inferior ao 1º ciclo do ensino básico	1º ciclo do ensino básico	2º ciclo do ensino básico	3º ciclo do ensino básico	Ensino secundário	Bacharelato	Licenciatura	Mestrado	Doutoramento
Portugal	**933,96**	**595,66**	**683,09**	**691,82**	**803,89**	**1 027,69**	**1 655,82**	**1 944,48**	**1 942,51**	**2 260,12**
Continente	**935,97**	**590,54**	**679,88**	**689,90**	**803,63**	**1 029,54**	**1 653,58**	**1 946,46**	**1 943,07**	**2 268,83**
Centro	**805,06**	**574,94**	**663,83**	**682,93**	**750,00**	**865,04**	**1 367,40**	**1 544,73**	**1 550,79**	**1 916,73**
Distrito da Guarda										
Aguiar da Beira	651,66	534,48	576,77	525,54	625,40	800,94	1 022,97	1 610,57	...	//
Fornos de Algodres	593,40	430,22	525,30	521,15	596,26	757,21	1 307,09	1 171,78	//	//
Gouveia	617,68	519,49	533,86	527,11	623,75	672,80	988,38	1 209,97	1 534,57	//
Seia	692,10	531,77	583,64	588,33	666,50	757,20	1 669,32	1 426,85	1 128,28	...
Almeida	651,26	419,76	561,35	621,37	641,33	729,59	1 183,47	1 063,92	...	//
Celorico da Beira	637,27	522,67	551,47	576,81	567,97	770,69	1 193,13	1 195,24	985,27	...
Figueira de Castelo Rodrigo	641,26	455,15	561,51	551,21	715,50	903,82	1 046,81	1 113,45	//	//
Guarda	751,69	510,98	619,45	643,94	687,85	827,45	1 257,67	1 341,73	1 093,38	...
Manteigas	703,41	500,48	546,35	620,56	666,67	661,62	1 063,84	1 613,85	//	//
Meda	678,94	702,33	579,47	629,75	632,23	873,49	996,22	1 057,18	1 266,75	//
Pinhel	665,28	547,18	572,48	585,10	623,94	825,36	1 345,29	1 367,95	...	//
Sabugal	618,12	551,40	540,73	556,10	548,79	730,36	1 182,25	1 118,76	...	//
Trancoso	693,94	574,08	594,27	599,85	732,07	741,92	1 167,93	1 254,58	1 131,53	//
Vila Nova de Foz Côa	655,06	616,16	566,74	593,49	612,31	831,46	1 111,41	1 072,43	...	//

Fonte: INE (2008a; 2008b).

Não podem os Municípios ser indiferentes ao ganho médio mensal dos trabalhadores por conta de outrem dos seus residentes, como mostra o **Quadro IX**. Destaca-se com naturalidade o concelho da Guarda, mas ao contrário do que poderia parecer, Manteigas apesar de ser o concelho mais pequeno do Distrito, nem por isso deixa de ser dos que evidência o ganho médio mensal mais elevado. Contudo, nem por isso deixa de ser o concelho que gera menor receita em termos de IRS, IMI e IMT, conforme evidenciado no Quadro IX.

Os baixos valores de IRS apurados resultam dos baixos salários auferidos pelos trabalhadores, na medida em que no Distrito predominam as agro-indústrias, têxteis e vestuário, incapacitadas, pelo elevado preço dos factores de produção, de competir no mercado internacional. Ora, uma estrutura produtiva assente, na generalidade, em Pequenas e Médias Empresas (PME), que, podendo apresentar algumas vantagens, pois as suas estruturas flexíveis permitem uma maior capacidade de adaptação às alterações de mercado, tem como desvantagem a incapacidade de praticar remunerações ajustadas às necessidades dos seus trabalhadores.

Neste quadro de referência, revela-se essencial a promoção do desempenho do concelho através de análises baseadas na avaliação crítica da sua situação económico-financeira gerada pelas acções do passado, bem como de análises ao contexto sócio-económico no qual se integra, para assim fundamentar a tomada de decisões futuras, com vista a dar resposta às necessidades da sociedade em que se integra e, deste modo, promover o Município como serviço público que impulsiona, efectivamente, a solução de problemas locais, regionais e nacionais, para além de propor outras soluções inovadoras (Abreu e David, 2008).

Por exemplo, aconselha-se e sugere-se que as deliberações da Assembleia Municipal sobre as Taxas de IMI possam evidenciar as regras elementares de economia, eficiência, eficácia e equidade, tendo Costa (2007) afirmado que a nível nacional:

> *"…só dez por cento das câmaras cobram a taxa mínima de IMI. (…) A competitividade fiscal é uma expressão que não parece fazer parte do vocabulário dos autarcas portugueses. Endividados e com grandes dificuldades financeiras, quer seja ao nível das empresas, quer seja ao nível dos imóveis detidos pelos munícipes, as Câmaras Municipais não usam as poucas possibilidades que a legislação lhes concede para atrair empresas ou famílias para os seus concelhos.".*

A este nível e face à análise realizada, segundo Abreu e David (2008), a existência de um sistema de informação contabilístico devidamente estruturado fornece aos "gestores" públicos, de forma oportuna, informações sobre a real situação dos Municípios.

4. CONTRIBUTO PARA A DEFINIÇÃO DE UMA ESTRATÉGIA

Na agenda de qualquer Município como entidade da Administração Pública é possível encontrar ligações entre a sustentabilidade, a responsabilidade social e a transparência, quer de modo explícito, quer de modo implícito, socorrendo-se de mecanismos ou práticas de gestão que promovem o desenvolvimento sustentável da mesma (Jallow, 2008). Além disso, a ambição de criar uma diferença estratégica tem de respeitar constrangimentos e falhas do próprio Município e, principalmente, do seu meio envolvente, pelo que se revela uma decisão consciente a definição pelo Plano Oficial de Contabilidade Pública (POCP) dos seus próprios objectivos, nomeadamente:

> *"a) A tomada de decisões estratégicas no domínio orçamental, designadamente no âmbito da orçamentação plurianual, face ao acompanhamento dos compromissos com reflexos em anos futuros;*
>
> *b) Disponibilizar informação para apoiar a actividade de controlo da actividade financeira da Administração Pública pelas entidades com competência legal nesse domínio e reforçar a transparência da situação financeira e patrimonial, bem como das relações financeiras do Estado."* (MF, 1997).

Como se observa, o POCP gera um sistema de informação contabilístico que estimula a expressão da discriminação, quer pela positiva, quer pela negativa, dos esforços e realizações do Município. Neste sentido, a definição de estratégias na área da tributação sobre o património deve orientar-se, fundamentalmente, para três áreas fundamentais: sustentabilidade, transparência e responsabilidade social, confirmando o efeito benéfico do seu confronto com as áreas referidas, especialmente quando estas últimas saem fora de controle e se tornam problemas, o que pode exigir a protecção contra imprevistos, incertezas, falhas, constrangimentos e logo serem retiradas vantagens económicas, sociais e profissionais (Hartmann, 2005).

Na realidade, a própria dinâmica económica a nível nacional ou local está intimamente associada à criação de novos contextos empresariais, onde a sustentabilidade, a transparência e a responsabilidade social permitem o desenvolvimento da inovação e respectiva investigação aplicada, tendo em vista a promoção de boas práticas, bem como a definição de políticas em áreas prioritárias. Os autores consideram fundamental a aproximação entre os diferentes actores e entidades de modo a respeitar o desenvolvimento sustentado e, deste modo, disseminar estratégias de bem-estar colectivo através da consolidação do Estado Social, que só é possível de manter através da arrecadação de receita pública, nomeadamente através da tributação sobre o património.

Contudo, o conceito de sustentabilidade é difícil de definir, na medida em que é difícil de medir (Likierman, 1993; Bevan e Hood, 2006; Hoque 2008), mas podem desenvolver-se abordagens e métodos para a sua avaliação. Por exemplo, ainda na área da tributação sobre o património e numa procura sistematizada da literatura, verificou-se a inexistência de estudos comportamentais do IMI, a nível municipal e a nível nacional, que, de forma sustentável, justificassem os níveis de receitas necessários para que uma determina autarquia salvaguardasse e cumprisse, acima de tudo, regras elementares de prudência na contabilidade orçamental. Pode, por isso, facilmente concluir-se que o conjunto de estratégias que devem conduzir à sustentabilidade do Município se deve enquadrar, sempre, no papel que essa entidade tem na sociedade, enquanto elemento da Administração Pública.

Este compromisso deve justificar-se, segundo Korhonen (2003a), em abordagens e métodos suportados nos três desempenhos: económico (produtividade e rentabilidade); ecológico (protecção e preservação do meio ambiente) e social (bem-estar para todos). A combinação do desempenho económico e ecológico poder criar valor para a entidade e também para toda a comunidade, mas tem contudo menos impacto no desempenho social[1] (Stigson, 2000). Os autores concordam com Korhonen (2003b: 301) ao afirmar que:

> *"… a noção de sustentabilidade deve, no entanto, tornar-se a base de discussão na sociedade. Talvez nós não possamos determinar a direcção da sustentabilidade, mas nós podemos saber o sentido geral de distância do desenvolvimento sustentável".*

Face ao exposto, importa salientar que o novo modelo de gestão pública deve ser transparente, com orientação para o cidadão e as empresas, eliminando todas as ineficiências que não agregam valor aos serviços prestados. Além disso, e como já anteriormente referido, é o Município, através da sua Assembleia Municipal, que fixa as taxas do IMI, dentro do intervalo permitido por lei, tendo esta entidade da administração pública responsabilidades acrescidas, dado a sua pertença ao sector público, bem como pelas exigências subjacentes à aplicação dos novos Códigos legais. No caso do CIMI, este permite adoptar diferentes estratégias para a promoção do bem-estar social das populações da sua área de influência. Segundo Costa (2007),

[1] Por exemplo, a desigualdade de impostos pagos em situações de imóveis que na realidade têm um valor igual. Uma fracção destinada à habitação, virada a norte, sem vista para o mar, mas localizada numa zona boa de costa, registou € 750,00 /m2 como custo médio da construção acrescido de mais 25% para o custo do terreno de implantação. A área de construção é de 150m^2 e o coeficiente de qualidade e conforto é igual a 1,2; o coeficiente de vetustez é igual a 1; e o coeficiente de localização é igual a 2; o valor fiscal é igual a € 270.000,00. No entanto, se a zona em causa vier a ser considerada de elevado valor de mercado imobiliário e for fixado para o coeficiente de localização o valor de 2,8 (próximo do máximo previsto) então o valor fiscal passa para € 378.000,00. O seu valor de mercado é, de cerca, € 300.000,00. Contudo, se o mesmo coeficiente for aplicado a outra fracção com a mesma área e no mesmo edifício, mas localizada num andar superior, não virada a norte e com excelente vista de mar, então o valor fiscal já não se afastaria muito do valor de mercado.

que desenvolveu uma análise nacional às taxas do IMI a cobrar em 2007 sobre imóveis detidos até 2006:

> *"…só uma pequena minoria das Câmaras Municipais é que utiliza o instrumento fiscal como factor de atractividade. Os dados disponíveis na página de Internet da Direcção-Geral dos Impostos não deixam grande margem para dúvidas: apenas cerca de 10% dos municípios portugueses é que cobram a taxa mínima de IMI prevista na legislação. Nas restantes câmaras municipais, mais de um terço aplica a taxa máxima de imposto e as restantes distribuem-se pelas taxas intermédias. Não é assim de estranhar que desde a entrada em vigor da reforma de tributação do património em 2003, as receitas das autarquias provenientes da tributação dos imóveis não pare de crescer: em 2006 o crescimento atingiu quase 10% e em 2007 a variação será ainda mais significativa".*

Perante esta realidade nacional, no **Quadro X** apresenta-se a competitividade fiscal dos Municípios do Distrito da Guarda, nomeadamente dos seus 14 concelhos. Deste modo, os valores a nível nacional quando comparados com os valores para o Distrito da Guarda, permitem concluir uma inversão de comportamento, já que se fala da aplicação da taxa mínima, esta é aplicada a nível nacional por 10% dos Municípios, enquanto que a nível do Distrito da Guarda, esta foi aplicada em 2007 por 35,7% dos Municípios.

Conforme reflecte o **Quadro X,** as taxas incidentes sobre os prédios urbanos, "novos" e "antigos", detidos no final do ano de 2006, cujo imposto foi pago no ano de 2007 registaram em cinco Municípios o valor mínimo (0,2), o que representa 35,7% do total do Distrito. Por sua vez, Guarda e Seia aplicaram a taxa máxima (0,5), o que representa 14,2% e os restantes Municípios aplicaram taxas intermédias.

Quadro X: Taxas do IMI nos Concelhos do Distrito da Guarda, 2006 (%)

Concelho	Prédios Urbanos	Prédios Urbanos Avaliados nos Termos do CIMI	Prédios Rústicos
Aguiar da Beira	0,6	0,3	0,8
Almeida	0,4	0,2	0,8
Celorico da Beira	0,7	0,4	0,8
Figueira de Castelo Rodrigo	0,4	0,2	0,8
Fornos de Algodres	0,7	0,4	0,8
Gouveia	0,8	0,5	0,8
Guarda	0,8	0,5	0,8
Manteigas	0,4	0,2	0,8
Meda	0,7	0,4	0,8
Pinhel	0,6	0,3	0,8
Sabugal	0,4	0,2	0,8
Seia	0,78	0,47	0,8
Trancoso	0,5	0,3	0,8
Vila Nova de Foz Côa	0,4	0,2	0,8

Fonte: Adaptado de DGCI (2008).

Face à análise realizada, os Municípios do Distrito da Guarda evidenciam comportamentos contrários ao panorama nacional, denotando, um conjunto significativo de Municípios, a existência de competitividade fiscal. No entanto, desse grupo, não fazem parte os Municípios maiores (Guarda e Seia), concluindo-se que tal preocupação está patente nos Municípios mais pequenos. Também é certo que nos Municípios de menor dimensão, em termos de tecido urbano, a perda de receita é menor, mas não se pode ignorar que os Municípios maiores para

não prescindirem da receita proveniente de tal imposto, deixam de utilizar o instrumento fiscal de que dispõem como factor de atractividade.

5. CONSIDERAÇÕES GERAIS

Esta investigação analisou a reforma da tributação sobre o património, com especial incidência no IMI. A primeira iniquidade surge em termos do próprio imposto, porque os autores defendem que embora se tenha procurado criar um sistema de tributação mais objectivo, porque o valor de referência é um "suposto" valor de mercado, verifica-se que a determinação do valor patrimonial dos prédios urbanos fica assente em coeficientes como localização, conforto e outros que podem envolver discricionariedade na aplicação. Assim, na prática, o que se tem verificado é que o sistema conduziu ao pagamento de impostos superiores ao valor real dos imóveis no mercado e completamente díspares. Apesar de terem passado poucos anos após a reforma da Tributação do Património, nesta altura já estão criadas novas iniquidades, que conduzem a injustiças entre os contribuintes.

A segunda iniquidade é para o próprio Município e a sua sustentabilidade. A redução generalizada no valor máximo das taxas coloca em causa as receitas que justificam as previsões orçamentais que deveriam ser mais adequadas, já que aqueles Municípios que tinham fixado taxas máximas, vêem agora a sua receita afastar-se para menos dos valores orçamentados. Assim, se em devido tempo cada Município tivesse tomado decisões tendo em conta a capacidade contributiva e a sustentabilidade dos seus munícipes, o mesmo teria tido oportunidade de procurar outras fontes de receita, ou adequar as despesas às receitas, evitando o endividamento que se tem vindo a revelar excessivo (Carvalho *et al*., 2009).

A terceira iniquidade surge em termos populacionais. Assim, a redução da natalidade, o envelhecimento da população, o prolongamento da esperança de vida, o êxodo migratório, entre outros padrões de vida exigem do Governo Nacional e Local adopção de reformas que sejam potenciadoras e correctas dos desequilíbrios entre os concelhos do Distrito da Guarda. A este nível agrega-se a problemática da responsabilidade social do Município, porque ao não ajustar as taxas do IMI à sua realidade socio-económica, o que provavelmente prejudica o crescimento das receitas provenientes deste imposto, também não contribui para o seu desenvolvimento sustentado. Muito ainda muito falta fazer, para que a tributação sobre o património seja um reflexo da transparência informativa dos Municípios, ao mesmo tempo que é um motor da sua sustentabilidade.

A quarta iniquidade surge em termos da organização administrativa. A este nível, a Administração Financeira do Estado parece apresentar algumas reformas potenciadoras das estruturas regionais, mas na verdade prossegue uma agenda de diferenciação negativa. Esta diferenciação também ocorre no Distrito da Guarda, ao ter 4 sub-regiões de NUTS II, 14 concelhos e 336 freguesias. Mais que uma diversidade de entidades, deveria exigir importantes reconfigurações nas formas organizativas para assim serem encontradas soluções em benefício dos agentes económicos com melhoria da eficiência, eficácia e económica e exigindo critérios de equidade fiscal para evitar fraudes e abusos. Muito ainda falta fazer, para que a tributação sobre o património seja um exemplo da atractividade regional através de benefícios fiscais, porque dois concelhos do mesmo Distrito podem concorrer entre si promovendo a atractividade das populações para a sua área geográfica, mas à custa da redução de receitas fiscais, causando descontentamento entre os contribuintes.

A quinta, e última, iniquidade surge em termos económicos. As adversidades da conjuntura económica global e nacional influenciam negativamente os Municípios, as actividades económicas, os agentes económicos e a sociedade, em geral. Neste contexto, a reforma da tributação do património deverá no futuro potenciar a criação de condições favoráveis: ao desenvolvimento da competitividade empresarial para mudar a estrutura produtiva do Distrito da Guarda; à luta contra a degradação e obsolescência dos prédios urbanos, rústicos e mistos, justificando uma intervenção integrada entre o Estado e a Sociedade; à responsabilidade social da requalificação do espaço urbano e rural com modernização de infra-estruturas e equipamentos; e, ainda, à dinamização de um conjunto de outras actividades económicas e

sociais que promovam em cada Município critérios socio-económicos capazes de se constituírem como verdadeiras soluções que pretendem contribuir para a sustentabilidade do Distrito da Guarda.

REFERÊNCIAS BIBLIOGRÁFICAS

ABREU, R. E DAVID, F. (2008). *Desenvolvimento Sustentável das Contas da Administração Pública: da perspectiva contabilística à perspectiva financeira*. Lisboa: 6º Congresso Nacional da Administração Pública.

ASSEMBLEIA DA REPÚBLICA (AR, 1992). Lei n.º 30-C/92. *Diário da República*, 298, Série I-A, 1º Suplemento, 28 de Dezembro, 5980-(43)-5980-(302).

ASSEMBLEIA DA REPÚBLICA (AR, 2000). Lei n.º 3-B/2000. *Diário da República,* 80, Série I-A, 2º Suplemento, 4 de Abril, 1490-(102)-1490-(631).

ASSEMBLEIA DA REPÚBLICA (AR, 2002). Lei n.º 32-B/2002. *Diário da República*, 301, Série I-A, 2º Suplemento, 30 de Dezembro, 8186-(136)-8186-(603).

ASSEMBLEIA DA REPÚBLICA (AR, 2008). Lei n.º 64-A/2008. *Diário da República*, 252, Série I-A, 1º Suplemento, 31 de Dezembro, 9300-(2)-9300-(389).

ASSOCIAÇÃO NACIONAL DE MUNICÍPIOS PORTUGUESES (ANMP, 2008). *Finanças Locais - Indicadores de Receita por Município – Ano 2006.* Disponível em: [http://www.anmp.pt/], acedido em 24 de Outubro de 2008.

BANCO DE PORTUGAL (BP, 2007). *Relatório de Estabilidade Financeira 2007*. Lisboa: BP.

BEVAN, G. E HOOD, C. (2006). What's measured is what matters: targets and gaming in the English public health care system, *Public Administration*, 84 (3), 517-38.

CARVALHO, J., FERNANDES, M., CAMÕES, P. & JORGE, S. (2009) *Anuário Financeiro dos Municípios Portugueses 2007.* Lisboa: CTOC

COMISSÃO DE COORDENAÇÃO DA REGIÃO CENTRO (CCRC, 1986a). *Programa de Desenvolvimento da Região Centro: Separata do Volume III -Raia*. Coimbra: CCRC/Ministério do Plano e da Administração do Território.

COMISSÃO DE COORDENAÇÃO DA REGIÃO CENTRO (CCRC, 1986b). *Programa de Desenvolvimento da Região Centro: Separata do Volume III - Serra da Estrela e Alto Mondego*. Coimbra: CCRC/MPAT.

COSTA, V. (2007). Só dez por cento das Câmaras cobram a taxa mínima de IMI. *Jornal Público*, 29 de Janeiro.

DAVID, M.F. (1996). *Contributos para uma Estratégia de Desenvolvimento dos Recursos Humanos visando a Inovação e a Qualidade no Distrito da Guarda - O Instituto Politécnico da Guarda* (Tese de Mestrado). Covilhã: Universidade da Beira Interior.

DIRECÇÃO GERAL DOS IMPOSTOS (DGCI, 2008). Taxas de IMI. Disponível em: [http://www.e-financas.gov.pt/de/jsp-dgci/main.jsp], acedido em 29 de Outubro de 2008.

HARTMANN, F. (2005). The effects of tolerance for ambiguity and uncertainty on the appropriateness of accounting performance measures, *Abacus*, 41 (3), 241-66

HOQUE, Z. (2008). Measuring and reporting public sector outputs/outcomes: Exploratory evidence from Australia. *International Journal of Public Sector Management***,** 21 (5), 468-493.

INSTITUTO NACIONAL DE ESTATÍSTICA (INE, 1993). Censos *1991 Recenseamento Geral da População e da Habitação*. Lisboa: INE.

INSTITUTO NACIONAL DE ESTATÍSTICA (INE, 2002). Censos *2001-Recenseamento Geral da População e da Habitação*. Lisboa: INE.

INSTITUTO NACIONAL DE ESTATÍSTICA (INE, 2008a). *Anuário Estatístico da Região Centro*. Lisboa: INE.

INSTITUTO NACIONAL DE ESTATÍSTICA (INE, 2008b). *Anuário Estatístico da Região Norte*. Lisboa: INE.

JALLOW, K. (2008). Sustainability and its place in CSR research. In Crowther, D. & Capaldi, N. (eds.) *The Ashgate research companion to corporate social responsibility,* Aldershot: Ashgate, 31-46.

KORHONEN, J. (2003a). Should we measure corporate social responsibility, *Corporate Social Responsibility and Environmental Management,* 10 (1): 25-39.

KORHONEN, J. (2003b). On the Ethics of Corporate Social Responsibility – Considering the Paradigm of Industrial Metabolism, *Journal of Business Ethics,* 48 (4): 301-315.

LIKIERMAN, A. (1993). Performance indicators: 20 early lessons from managerial use, *Public Money and Management*, 13 (4), 15-22.

MINISTÉRIO DAS FINANÇAS (MF, 1958). Decreto-Lei n.º 41969/58. *Diário da República*, Série I, 24 de Novembro.

MINISTÉRIO DAS FINANÇAS (MF, 1963). Decreto-Lei n.º 45104/63. *Diário da República,* 153, Série I, 1 de Julho, 868-902.

MINISTÉRIO DAS FINANÇAS (MF, 1988). Decreto-Lei n.º 442-C/88. *Diário da República*, 277, Série I, 3º Suplemento: 4754-(74)-4754-(81).

MINISTÉRIO DAS FINANÇAS (MF, 1991). Decreto-Lei n.º 308/91. *Diário da República,* 188, Série I-A, 17 de Agosto, 4212-4213.

MINISTÉRIO DAS FINANÇAS (MF, 1992). Decreto-Lei n.º 140/92. *Diário da República,* 163, Série I-A, 17 de Julho de 1992, 3353-3354.

MINISTÉRIO DAS FINANÇAS (MF, 1997). Decreto-Lei nº 232/97. *Diário da República,* 203, I Série-A, 3 de Setembro.

MINISTÉRIO DAS FINANÇAS (MF, 2001). Decreto-Lei n.º 198/2001. *Diário da República*, 152, Série I-A, 3 de Julho, 3923-4012.

MINISTÉRIO DAS FINANÇAS (MF, 2003a). Decreto-Lei nº 287/2003. *Diário da República*, 262, Série I-A, 12 de Novembro, 7568-7647.

MINISTÉRIO DAS FINANÇAS (MF, 2003b). Portaria n.º 1283/2003. *Diário da República*, 263, Série I-B, 13 de Novembro, 7664-7666.

MINISTÉRIO DAS FINANÇAS (MF, 2003c). Portaria n.º 1337/2003. *Diário da República*, 281, Série I-B, 5 de Dezembro, 8241-8241.

MINISTÉRIO DAS FINANÇAS E DA ADMINISTRAÇÃO PÚBLICA (MFAP, 2009). *Relatório ao Investimento e ao Emprego. Regime Fiscal de Apoio ao Investimento 2009 e Alteração à Lei nº 64-4/2008.* Lisboa: MFAP.

STIGSON, B. (2000). Sustainable development and business, Organisation for Economic Cooperation and Development. *The OECD Observer*, 221/222: 36-37.

EVOLUÇÃO DO SECTOR DO TURISMO EM PORTUGAL: RELEVÂNCIA ECONÓMICA E ESTRATÉGICA

ERMELINDA OLIVEIRA
E-mail: ermelindaol@ipg.pt
Instituto Politécnico da Guarda
JOSE PIRES MANSO
E.mail : pmanso@ubi.pt
Universidade da Beira Interior

Resumo

Esta investigação analisa a evolução do sector do turismo em Portugal. Este tem assistido a importantes alterações na forma de actuar dos seus agentes económicos, com vista ao seu desenvolvimento sustentável. Este sector assume o papel de motor do desenvolvimento regional e, simultaneamente, de factor de expansão económica. A nível nacional, dado o aumento do número de turistas, por parte da procura, e do número de agentes económicos, por parte da oferta, a importância estratégica do turismo, traduzida nas receitas que proporciona, no seu contributo para a taxa de emprego da população, bem como nos efeitos multiplicadores que induz nas mais variadas áreas, tem levado à adopção de medidas inovadoras e dinamizadoras do crescimento económico. Assim, face à importância estratégica deste sector, a presente investigação discute sobre a expansão significativa da oferta turística, que tem sido acompanhada de uma maior diversificação, bem como de uma maior aposta na qualidade dos serviços prestados. Para além de concluir que, em consequência dessa rápida expansão, o sector enfrenta problemas cada vez maiores quanto às dimensões naturais, culturais e socioeconómicas.

Palavras-chave: Turismo, Desenvolvimento Sustentável, Estratégia Económica.
Área Temática: Estratégia Económica.

Abstract

This research examines the evolution of the Portuguese Hospitality Sector. It has been affected by changes of operators acting, because of its sustainable development perspective. This sector assumes to promote strategies of regional development, despite it is classify as the economic expansion factor. At national level, the increasing of demand in number of tourists, and, on the supply side, the number of operators shows the strategic importance of Hospitality Sector, reflected in the revenue it provides, in its contribution to the employment. The effects, in various areas, have led to the adoption of innovative and dynamic economic growth. Thus, given the strategic importance of this sector, this research discusses the significant expansion of tourism, which has been accomplished by greater diversification, as well as, greater service quality. Apart from that, in consequence of such rapid expansion, the sector faces increasing problems as the natural, cultural and socioeconomic dimension.

Key words: Tourism, Sustainable development, Economic strategy.
Thematic Area: Economics of Strategy.

1 -INTRODUÇÃO

Apesar de desde sempre o homem se deslocar de um lado para o outro, a natureza, a dimensão e os fins das suas viagens, foram-se alterando ao longo do tempo. Como actividade económica, o turismo é um fenómeno do século XVIII, devido ao progresso económico e social verificado neste século. Contudo, foi no século XX que o turismo se transformou num fenómeno da sociedade, alcançando uma dimensão económica nunca vista. A partir desta época, dada a importância que o turismo já assumia, quase todos os países da Europa criaram instituições governamentais e privadas com o fim de promover e organizar o turismo.

Apesar de em Portugal, sempre existirem condições que favorecem o desenvolvimento do turismo, a coexistência de outros factores nefastos ao desenvolvimento do turismo enquanto actividade económica, fez com que este se tenha desenvolvido com um atraso temporal considerável, comparativamente a outros países europeus. Segundo Cunha (2006), em termos temporais, existem quatro etapas marcantes e distintas quando em questão está a análise do desenvolvimento do turismo português. Por outro lado, apesar de Portugal oferecer, quer em termos naturais, patrimoniais e ambientais, em todo o país condições impares para atracão turística, o seu desenvolvimento tem sido diferenciado de região para região.

Pelo facto de o turismo ter uma importância verdadeiramente estratégica para a economia mundial, bem como para a economia portuguesa, na presente comunicação será feita uma análise da evolução histórica do turismo quer em termos mundiais, quer em termos nacionais. De seguida, e porque a hotelaria é considerada a espinha dorsal da actividade turística, será efectuada uma análise critica da evolução do número de estabelecimentos hoteleiros existentes em Portugal, por NUT`s II, ao longo dos últimos vinte anos. Antes de uma breve consideração final, será abordada a importância do turismo de desenvolver de uma forma sustentável, pelo facto de cada vez mais este sector enfrentar problemas cada vez maiores nas dimensões naturais, culturais e socioeconómicas.

2 – PERSPECTIVA HISTÓRICA DO TURISMO

O desejo de conhecer e estabelecer relações com outras civilizações, foi sempre uma constante na história do homem. Pelos mais variados motivos, nomeadamente, por razões comerciais, politicas, religiosas, de incremento do próprio território, a história da humanidade está estreitamente ligada às deslocações e às viagens. Contudo, segundo Cunha (2006: 35) "a natureza das viagens, os seus fins, a sua dimensão e extensão, bem como a forma que assumiram, foram adquirindo características muito diferentes de época para época até alcançarem aquilo a que se chama «Idade da Viagem», tal como se fala da Idade da Agricultura, da Industria ou da Electrónica." Segundo o mesmo autor, e tendo em conta "as características da evolução registada ao longo dos tempos podemos identificar três épocas históricas do turismo: a idade clássica, a idade moderna e a idade contemporânea" (2006:35).

2.1 - Idade Clássica do Turismo

A idade clássica do turismo, que vai desde os primórdios das primeiras civilizações e se prolonga até aos princípios do século XVIII, caracteriza-se pelo facto das viagens serem individuais e se realizarem, predominantemente, por necessidades fundamentais como o comércio, as peregrinações religiosas, a saúde ou por razões políticas e de estudo (Cunha, 2006). Apesar de não ser possível, em termos temporais, delimitar o início das primeiras viagens, podemos atribuir aos Sumérios o mérito de terem proporcionado as primeiras condições que possibilitaram a realização de viagens. A eles se deve a invenção da moeda, o desenvolvimento do comércio, bem com a invenção da roda que veio impulsionar a construção de estradas.

Foram os Romanos (150 a.C.) que instruíram a maior rede de estradas até então existente. Segundo cunha (2006:36), "provavelmente, a primeira viagem feita, com objectivos que hoje consideraríamos como turísticos, foi realizada pela rainha Hatshepsut cerca de 1500 a. C., tal como se descreve no templo Dar el Bahari situado no deserto próximo de Luxor." Já naquela época os Romanos e os Gregos, viajavam com a finalidade de conhecer templos e as sete maravilhas mediterrânicas, nomeadamente as pirâmides e os monumentos do Egipto. Com os romanos desenvolveram-se os *Hospes* (estalagens), os *Hospitium* (Hotéis) e os *Hospitalia* (Estalagens Públicas), bem como o espírito da hospitalidade, que passou a ser um acto honroso, e institui-se a obrigação de receber com benevolência os estrangeiros que chegassem a uma cidade. Em Roma, a 25 a.C., com o desenvolvimento das instalações termais, desenvolveram-se verdadeiros centros de turismo, que se prologaram até aos nossos dias.

Com os Cristãos, a hospitalidade continuou a ser um dever e um direito sagrado que devia ser gratuitamente concedida, tendo-se no século IV aberto as primeiras casas de refúgio para os viajantes. De acordo com Cunha (2006) nesta época, as viagens tinham como principal razão as peregrinações sendo célebres as que se dirigiam a Santiago de Compostela (Espanha), a Canterbury (Inglaterra), à terra Santa (Palestina) e a Meca (Arábia). Já nesta época existam guias de viagens para fornecer aos peregrinos informações sobre as regiões, alojamentos e o caminho a percorrer.

No século XIII, é de referir a grande viagem de Marco Pólo que, durante 24 anos, percorreu o Oriente até a China. Já no século XV, o século dos descobrimentos portugueses, os portugueses percorreram toda a costa de África e mar Vermelho, chegaram à Índia, China, Japão, Malaca, Timor e descobriram o Brasil. Posteriormente aos portugueses, seguiram-se as grandes viagens dos Espanhóis (chegaram às Caraíbas, Antilhas e América Central e Sul), dos Ingleses (descobriram a América do Norte), Franceses e Holandeses. Com estas descobertas, iniciadas pelos portugueses, chegou uma nova era para o mundo das viagens.

2.2 - Idade Moderna do Turismo

Ao longo do período histórico que abarca a Antiguidade e a Idade Média, o formato e os motivos das viagens conservaram as mesmas particulares, não se conseguindo diferenciar com clareza estas duas épocas. Foi a partir de meadas do século XVIII que as grandes mudanças verificadas em termos económicos, sociais, culturais bem como a nível tecnológico, vieram provocar modificações significativas no âmbito das viagens. Em toda a Europa, desenvolveu-se a construção de estradas e das redes de comunicação. Em Inglaterra, com os primeiros passos da revolução industrial, incrementaram-se as relações internacionais, pois com a invenção da máquina a vapor, promoveu-se uma revolução nos meios de transportes.

Nesta época, estudantes, diplomatas e membros das famílias ricas inglesas, começaram a fazer o *Grand Tour,* viajando pela Europa, durante um período de normalmente três anos. "Com a *Grand Tour* nasce o conceito de turismo e, pela primeira vez, começam a designar-se as pessoas que viajam por turistas" (Cunha, 2006: 41). Nesta época verificou-se um incremento no número de livros dedicados às viagens, bem como no número dos guias turísticas. Por outro lado, o movimento dos ingleses para a Europa influenciou muito o desenvolvimento dos transportes, da hotelaria e da restauração

No século XIX, a proliferação das trocas, o desenvolvimento dos transportes, bem como a propagação de ideias com a difusão de publicações, vieram dar "um novo impulso às viagens que começam a encontrar a sua verdadeira identidade: um meio de as pessoas se interessarem pelas particularidades de cada povo, pelas tradições, pelo exotismo e por outros modos de vida e novas culturas" (Cunha, 2006: 42).

Em 1841, nasceu o turismo organizado com Thomas Cook, quando este organizou a primeira viagem colectiva com a duração de um dia, destinada a 570 passageiros num congresso de médicos. As iniciativas de Cook (em 1864 organizou a primeira excursão acompanhada no regime «tudo incluído»; em 1872 organizou a primeira viagem à volta do mundo que durou 222 dias), marcaram "uma das mais importantes etapas na história do turismo e estão na origem do turismo dos nossos dias, continuando a agência por ele criada a ser uma das maiores organizações turísticas do mundo" (Cunha, 2006: 42).

No decorrer do século XIX, popularizaram-se as viagens, deu-se a criação de inúmeras agências de viagem e o turismo começa a surgir como uma actividade organizada. Em 1883 o austríaco Stadner falou pela primeira vez na «indústria do turismo» e em 1896 Frenlu publicou o documento «Contribuições para uma Estatística do Turismo».

A primeira década do século XX caracterizou-se por transformações e inovações que alteraram profundamente os modos de vida das populações, nomeadamente (Cunha, 2006).

- A descoberta do telégrafo e do telefone;
- O alargamento da rede de caminhos de ferro;
- O aumento das extensões das redes de estradas;
- O grande desenvolvimento industrial;

- A racionalização do trabalho e as reivindicações sindicais, que levaram a uma maior democratização das sociedades e a novos conceitos de vida;
- O tempo de trabalho diminui e alcança-se o direito ao repouso semanal, dai o conceito de lazer surgir com uma nova noção;
- A Organização Internacional de Trabalho (OIT) estabelece numa convenção o princípio das férias pagas, posteriormente reconhecido pela Declaração Universal dos Direitos do Homem.

Assim, ao longo de todo o século XX, o turismo transformou-se num fenómeno da sociedade, alcançou uma dimensão económica até então nunca vista, influenciando os comportamentos das pessoas na sociedade. A partir desta época, dada a importância que o turismo já assumia, em quase todos os países da Europa criaram-se instituições governamentais e privadas com o fim de promover e organizar o turismo. De entre muitas organizações, destaca-se a primeira organização internacional do turismo, de nome Federação Franco-Hispano-Portuguesa de Sindicatos de Iniciativa e Propaganda, transformando-se depois na União Internacional dos Organismos Oficiais de Turismo (UIOOT), e que mais tarde está na origem da actual Organização Mundial do Turismo (OMT). Perante a análise de todos os factos, podemos concluir que nesta época, já existiam condições para o desenvolvimento do turismo como actividade económica constante, bem como a sua universalização. Porém, a II Guerra Mundial veio atrasar o seu desenvolvimento e a sua universalização.

2.3 - Idade Contemporânea do Turismo

Apesar das ocorrências que atemorizaram o mundo (I Guerra Mundial, Grande Crise de 1229 e a Guerra Civil de Espanha), até ao início da II Guerra Mundial o turismo atingiu grandezas significativas. Contudo, durante o período que decorreu da II Guerra Mundial à época da recuperação económica, o turismo sofreu grandes dificuldades, entrando numa fase de declínio acentuado. Só mais tarde, a partir dos anos 50 e aliado à fase de progresso económico e social é que entrou novamente numa fase de desenvolvimento e consolidação. Com o objectivo de analisar com profundidade os factores que influenciaram o turismo, o perfil dos turistas, bem como o lado da procura turística, Cunha (2006) subdivide a época posterior a 1945, em três partes distintas: de 1945 a 1973, de 1973 a 1990 e finalmente, após 1990.

1º Período: 1945 a 1973

Neste período, das mais significativas alterações ocorridas destacam-se (Cunha, 2006):
- a ascensão de um grande número de países à independência;
- a produção mundial aumentou à média anual de 5%;
- um crescimento do rendimento real por habitante de 3%;
- um incremento muito significativo das trocas internacionais;
- os créditos internacionais, o capital, a tecnologia e a mão-de-obra registaram uma enorme mobilidade;
- verificou-se a emergência de grandes multinacionais bem como a constituição de grupos económicos (CEE);
- um aumento dos dias de férias pagas dos trabalhadores;
- o lançamento dos primeiros satélites, bem como a primeira viagem à Lua;

Os efeitos produzidos ao nível do turismo, por todas estas alterações, fizeram-se sentir em duas vertentes, no lado da procura e no lado da oferta turística. No lado da procura, registaram-se alterações por intermédio de modificações produzias ao nível do rendimento, da duração do tempo livre das pessoas e das suas motivações. Ao nível do rendimento é de registar o aumento verificado no rendimento considerado discricionário, na adopção de medidas sociais, bem como no incremento verificado na facilidade de compra de viagens.

No que concerne ao tempo livre, decorrente de uma diminuição do tempo de trabalho semanal, este aumentou e por outro lado, nos países desenvolvidos verificou-se a generalização

das férias pagas. Ao nível das motivações existiram alterações derivadas da necessidade de diversificação e diferenciação, bem como pela crescente necessidade pela fuga ao meio, pois as pessoas sentem cada vez mais a necessidade de reparar os desequilíbrios psicológicos decorrentes da sua vida profissional. No lado da oferta turística, salienta-se um incremento nas viagens aéreas e, por outro lado, os operadores turísticos desenvolveram a produção em série de produtos de massa. Este período caracteriza-se por uma procura turística de sol e mar. No lado da oferta, todas as preocupações se centraram no desenvolvimento do turismo internacional, sendo o turismo interno um subproduto do turismo internacional no seio das orientações das politicas turísticas.

2º Período: 1973 a 1990

Após um acelerado desenvolvimento da economia mundial, verificado no período compreendido entre 1945 a 1973, após este período registou-se:

- uma acentuada diferença entre o nível de vida dos países em vias de desenvolvimento e os países industrializados;
- um afrouxamento do crescimento económico mundial, derivado de uma diminuição da produção, acompanhado da estagflação e do aumento da taxa de desemprego;
- variações velozes das taxas de câmbio e uma crise de confiança no sistema monetário mundial, acompanhada com endividamento externo na generalidade dos países de tal nível que abalou os fundamentos do sistema financeiro internacional;
- os problemas mundiais tomaram uma tal amplitude que foi necessário criar condições para uma melhor cooperação mundial.
- pela primeira vez, os problemas ambientais (poluição da atmosfera, das águas, bem como derivados da utilização de produtos tóxicos), são encarados como um grande perigo para a humanidade. Tal constatação, levou a novas atitudes e comportamentos por parte dos consumidores.

Fruto destas ocorrências, o turismo mundial não diminuiu, mas o seu ritmo de crescimento abaixou e sofreu grandes alterações estruturais. Do lado da oferta, verificou-se uma diversificação, designadamente pela oferta de mais meios de animação e da prática de desporto. Por outro lado, e ainda ao nível da oferta, em detrimento do turismo internacional, neste período o turismo interno passou a adquirir maior importância. Ao nível da procura, a duração e a distância das viagens diminuíram e os alojamentos de mais baixo custo passaram a ser os mais procurados.

Segundo Cunha (2006: 48), neste segundo período "a nível conceptual, passou a enfatizar-se menos o papel económico do turismo no qual se tinha, até então, insistido em excesso para, igualmente, se atribuir importância ao seu papel social, politico, ecológico, cultural e educativo o que levou a passar a considerá-lo como uma das componentes essenciais da vida do homem. Deixou de ser unidimensional para passar a ser multidimensional, na medida em que responde a uma multiplicidade de necessidades humanas e não apenas à melhoria do bem-estar material. Os valores económicos do turismo, sem deixar de estar presentes nas preocupações do seu desenvolvimento, foram relegados para um plano de menor evidência dando lugar aos valores da identidade e valorização do homem."

3º Período: após 1990

Apesar de nos anos 80, a economia mundial ter registado taxas de crescimento, que vieram permitir a retoma dos mercados, a partir de 1992 a economia mundial entrou numa fase de recessão generalizada, considerada por muitos analistas como a maior crise económica posterior à II Guerra mundial. Contudo, num cenário de crise económica, com altas taxas de desemprego em todos os países, um clima de grande tensões politicas em termos internacionais e apesar dos grandes avanços sentidos para uma maior integração económica, o turismo resistiu-se mas mantêm a sua tendência de crescimento ainda que a um ritmo menos acelerado.

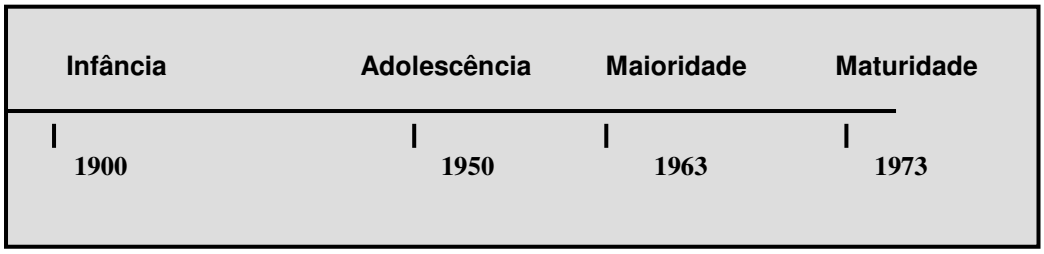
Entrando no século XXI, ao nível do turismo saliente-se os movimentos verificados em termos de reorganização empresarial (formação de grandes aglomerados económicos e de novos agentes), novas formas de turismo (turismo social), uma maior diversificação da oferta (novas formas de alojamento, novas formas de atracção turística, em suma, novos produtos e novos destinos.

3 – PERSPECTIVAS EVOLUTIVAS DO TURISMO PORTUGUÊS

Apesar de em Portugal, sempre existirem condições que favorecem o desenvolvimento do turismo (condições climatéricas, naturais, sentido de hospitalidade, peculiaridade da sua cultura, entre outras) a coexistência de outros factores nefastos ao desenvolvimento do turismo enquanto actividade económica (atraso nos meios de transporte e nas vias de comunicação, a inexistência de uma iniciativa privada e informada, condão pelo isolamento internacional), fez com que este se tenha desenvolvido com um atraso temporal considerável, comparativamente a outros países europeus.

Segundo Cunha (2006), quando em questão está a análise do desenvolvimento do turismo português, em termos temporais existem quatro etapas marcantes e distintas. À primeira etapa, que vai de 1900 a 1950 (50 anos), o autor chamou-lhe a «Fase da Infância». Ao segundo período, compreendido entre 1950 a 1963 (com uma duração de 13 anos), denominou-o por «Adolescência», seguindo-se a etapa da «Maioridade» entre1963 e 1973 e, por fim, a quarta e última etapa onde o turismo português atingiu a sua fase de «Maturidade» (**Quadro 1**).

Quadro 1. Etapas da Evolução do Turismo em Portugal



Fonte: Retirado de Cunha (2006: 88).

3.1 - Primeira Etapa – Fase da Infância

Até aqui, o turismo interno português assentava basicamente nas instâncias termais e o turismo internacional na Madeira e em Lisboa. O desenvolvimento do turismo interno e do turismo internacional português verificado nesta época, em muito se deve à Sociedade de Propaganda de Portugal (SPP), uma sociedade de iniciativa privada que foi criada em 1906. Esta sociedade tinha como objectivos, organizar e divulgar todo o património cultural e natural de Portugal, com a finalidade de promover quer o turismo interno, quero turismo internacional. Para o efeito, em 1920 a SPP abriu cerca de 143 representações, espalhadas por todo o país.

No seguimento da realização em Lisboa do IV Congresso Internacional de turismo, organizado pela Federação Franco-Hispano-Portuguesa de Iniciativa e Propaganda, em 1911 surgiu, em Portugal, a primeira organização oficial do turismo português, denominada por «Repartição do Turismo» e integrada no Ministério do Fomento.

Nesta primeira etapa, a desabrochamento da I Guerra Mundial (1914 – 1918), a Guerra Civil de Espanha (1936) e a eclosão da II Guerra Mundial (1936-1945), bem como os seus efeitos quer a nível económico, politico e também social, contribuíram para o reduzido numero de visitas de estrangeiros e para que o desenvolvimento do turismo português de adiasse. Apesar disso, neste período (Infância do Turismo) algumas iniciativas foram tomadas, nomeadamente:

- o Plano de aproveitamento do Estoril e o programa de construção de pousadas, que ainda nos nossos dias são uma imagem de marca do nosso país;

- em 1920 foi reforçada a estrutura da organização oficial do turismo com a criação do Comércio e Comunicações;
- subordinadas à Repartição do Turismo foram criadas em certas localidades de Portugal comissões, com o objectivo de estas protegerem os monumentos nacionais e cooperarem em estudos do turismo;
- em 1928 o jogo ficou ligado ao turismo e criou-se a Repartição de Jogos e Turismo, repartição que ficou integrada no Ministério do Interior;
- em 1930 criou-se a Comissão de Propaganda do turismo de Portugal, e nos anos seguintes foram criadas as Casas de Portugal em vários países, tendo-se assim lançado as primeiros representações de Portugal no estrangeiro.

3.2 - Segunda Etapa – Fase da Adolescência

No início dos anos 50, os países europeus saíram da guerra e entraram em plena recuperação económica e social, recomeçando-se assim uma nova era para o turismo. Portugal, apesar de ser um país detentor de condições climatéricas favoráveis e rico em termos paisagísticos, não acompanhou a nova era do turismo da Europa, pois encontrava-se numa situação de atraso em relação a um conjunto de meios materiais exigidos pela vida moderna (rede de estradas, capacidade hoteleira bem como um conjunto de atractivos turísticos).

Contudo, perante obstáculos e dificuldades, nesta época começou a aparecer uma nova concepção do turismo, passando este a ser entendido como uma verdadeira actividade económica. Segundo Cunha (2006: 82), nesta etapa "considerava-se que o turismo não deve ser tido como divertimento, mas sim como uma verdadeira indústria não podendo, por isso, ser obra de amadores, mas antes tem de se subordinar a princípios de organização administrativa semelhante aos de qualquer actividade económica."

Esta nova abordagem do turismo contribui para que nesta fase temporal:
- a rede de estradas aumentasse consideravelmente;
- em 1954 se publica-se a Lei n.º 2073 referente ao exercício da industria hoteleira, e que veio criar o estatuto de utilidade turística, permitindo a um conjunto de empresas hoteleiras o acesso a um conjunto de regalias, como forma de fomentar a capacidade hoteleira no nosso país;
- em 1956 se publica-se a Lei n.º 2082 que passou a ser a lei base do turismo que criou o Fundo de Turismo;
- em 1962 se criasse uma subsecretaria de Estado da Presidência do Conselho, com competências especificas na área do turismo. Foi aqui que pela primeira vez o turismo passou a ser tratado a nível governamental.
- se criassem zonas de turismo nos concelhos em que existiam praias, estâncias hidrológicas ou climáticas, estâncias de repouso e de recreio ou monumentos.

3.3 - Terceira Etapa – Fase da Maioridade

Apesar de ter sido a partir da década de 30 do século passado que se verificou o despertar da actividade turística, em Portugal, "por razões ligados a factos de conjuntura político-económica, as Grandes Guerra desse século, bem como à condução da vida portuguesa pouco inclinada à mudança, só a partir da década de 60, o turismo começou a ter expressão significativa no conjunto das actividades nacionais" (Alexandre, 2003:5). A partir de 1964 surgiu uma nova era para o turismo europeu e para o turismo português. Depois da recuperação económica dos países industrializados afectados pela guerra, estes estavam agora perante um boom económico, que veio permitir a vulgarização do automóvel, o desenvolvimento do trânsito aéreo bem como, a generalização das férias pagas.

Nesta época, em Portugal assistiu-se a uma expansão dos empreendimentos turísticos, de uma oferta hoteleira em maior número e mais diversifica, bem como ao desenvolvimento de novos centros turísticos (Algarve, Madeira e Tróia). Fruto do desenvolvimento, desordenado e

sem apoio das infra-estruturas indispensáveis, dos empreendimentos turísticos, surgiram os primeiros impactos nefastos em termos ambientais, bem como sinais de assimetrias regionais. Foi nesta época que surgiu a Direcção-Geral do Turismo e o turismo, que até aqui nunca tinha sido objecto de qualquer planeamento, passou a ser enquadrado nos Planos de Fomento.

A busca do sol e mar era dominante ao nível da procura, em detrimento do turismo do interior. Por outro lado, "é neste período que, pela primeira vez, as dormidas de estrangeiros na hotelaria ultrapassam as dormidas de nacionais alterando-se, simultaneamente, a estrutura da procura externa" (Cunha, 2006: 87).

3.4 - Quarta Etapa – Fase da Maturidade

Portugal, com a revolução de 1974 sofreu grandes transformações económicas, políticas e sociais. Fruto da revolução de 25 de Abril, um grande número de desalojadas das ex-colónias passaram a ocupar os alojamentos hoteleiros e a procura externa sofreu uma quebra acentuada ao longo da década de 70. Porém, após o início da década de 80 o sector reagiu e a procura turística externa tomou valores nunca vistos, provocando em paralelo um clima de euforia que levou a que decisões empresariais de campo imobiliário substituíssem e subestimassem os princípios do próprio turismo. "Apesar do extraordinário crescimento da procura, o turismo português passou a viver em estado de crise latente ou explícita derivada do excesso da oferta, da inadequação das infra-estrutura, do desordenado aproveitamento dos espaços, do desequilíbrio entre as várias componentes da oferta e de outros factores ligados às condições de recepção internas. É uma crise originada pelo lado da oferta e não pelo lado da procura" (Cunha, 2006: 91).

Se até aos anos 70 do século passado, o turismo era considerado como uma "indústria limpa", daí não lhe serem imputados efeitos ou externalidades negativas quer de ordem social, quer de ordem ambiental, a partir da década de 80 começaram a despontar vozes denunciantes dos efeitos adversos que a actividade turística exercia sobre o meio ambiente e sobre a população residente e receptora do turismo (Burnay, 2000). Esses efeitos negativos traduziam-se "na perda da qualidade de vida, nos desajustes das economias locais, no desmembramento da organização social, na ocultação dos povos e consequente perda de identidade" (Burnay, 2000: 26).

Segundo José Alexandre, (2003: 56) "no plano teórico, o turismo parecia encerrar qualidades susceptíveis de gerar dinâmicas de desenvolvimento endógeno, designadamente nas regiões de acolhimento. O turismo fazia o uso de recursos locais, em muitos casos sub-aproveitados, com fracas potencialidades de utilização económica, pelo que permitia uma exploração mais eficaz das potencialidades da região. Paralelamente, constituía um factor de diversificação da estrutura da região, multiplicando as oportunidades de emprego e aumentando, consequentemente, o produto e o nível de vida das populações locais. O turismo poderia ainda produzir importantes efeitos multiplicadores na economia regional, promovendo a dinamização e modernização do tecido económico em que se inseria, com reflexos no domínio cultural e garantindo os limiares mínimos necessários à criação de infra-estruturas e equipamentos". Contudo, pelo facto de o chamado turismo de massa, escorado pela procura externa, ser orientado por operadores turísticos, fugiu do alcance de controlo das autoridades locais, não se geraram dinâmicas de desenvolvimento nas regiões necessitadas. Efectivamente, o que se veio a verificar, foi que a actividade turística, orientada e implementada da forma como estava a ser feita, provocou graves efeitos no domínio social, cultural, ainda maiores desequilíbrios espaciais em termos de desenvolvimento e, paralelamente, produziram-se graves problemas ambientais.

Nesta época, e segundo Cunha (2006) o turismo português peca quer em termos de concentração territorial, quer em termos de concentração em motivações e atractivos. Concentração territorial pelo facto da região do Algarve em conjunto com a região de Lisboa, absorverem mais de metade das dormidas de estrangeiros e ser nestas regiões onde esta situado a grande maioria dos alojamentos hoteleiros. Concentração em atractivos e motivações por ser o sol e a praia o grande atractivo da oferta turística portuguesa. Consciente destas vulnerabilidades do turismo, e tendo como objectivos atenuar os desequilíbrios regionais, valorizar o património cultural, bem como, conseguir uma melhoria na qualidade de vida dos

portugueses, na década de 80 foi lançado um Plano Nacional de Turismo. Segundo Cunha (2006: 93) a filosofia adoptada neste Plano Nacional de Turismo "foi a transformação qualitativa do turismo com base na defesa da qualidade, da diversidade e da harmonização do aproveitamento do território".

3.5 – Após a década de noventa

Segundo Cunha (2006: 97) "a evolução do turismo português, a partir dão início da década de 90, indicia o esgotamento do modelo de desenvolvimento iniciado na década de 60 baseado na excessiva exploração do «sol e mar», no aproveitamento indiscriminado dos recursos naturais e na concentração dos investimentos junto ao litoral orientados, na perspectiva do negócio imobiliário".

Como consequência do esgotamento do modelo de desenvolvimento do turismo até aqui adoptado, no início do século XXI começaram a surgir esforços no sentido da diversificação e da desconcentração do turismo. Do lado da oferta têm-se assistido a um conjunto de medidas dinamizadoras, nomeadamente uma expansão significativa da oferta turística, acompanhada por uma maior diversificação (por exemplo, turismo rural e turismo de habitação) bem como uma maior aposta na qualidade dos serviços prestados. Paralelamente, temos assistido a uma aposta num desenvolvimento regional mais equilibrado ao nível das infra-estruturas necessárias para o desenvolvimento do turismo.

Em termos governamentais, existe uma grande aposta no sentido de dinamizar e diversificar o turismo, por todo o território nacional pelo facto de o turismo ter uma importância verdadeiramente estratégica para a economia portuguesa em virtude da sua capacidade em criar riqueza e emprego. "Trata-se de um sector em que temos vantagens competitivas claras como sucede com poucos outros" (MEI, 2007: 1). Tendo por fim último a promoção e o desenvolvimento do turismo em Portugal, de um turismo sustentável, o Plano Estratégico Nacional do Turismo (PENT), estabeleceu como objectivo até 2015 desenvolver o sector através das seguintes onze linhas de desenvolvimento estratégico (MEI, 2007):

1 - Produtos, destinos e pólos;
2 - Intervenções em Urbanismo, ambiente e Paisagem;
3 – Desenvolvimento de conteúdos distintos e inovadores;
4 – Eventos;
5 – Acessibilidade aérea;
6 – Marcas, promoção e distribuição;
7 – Programa da qualidade;
8 – Excelência no capital humano;
9 – Conhecimento e inovação;
10 – Eficácia no relacionamento Estado-empresa;
11 – Modernização empresarial.

Apesar da emergência de novos destinos, que têm afastado turistas dos mercados tradicionais, nomeadamente de Portugal, e ao contrário do que se verifica com muitos dos seus concorrentes, Portugal tem conseguido manter a sua participação a nível mundial. No ano de 2005, Portugal encontrava-se posicionado no 18º lugar no ranking dos principais destinos turísticos.

4 – ALGUNS INDICADORES DO TURISMO EM PORTUGAL

Apesar de Portugal oferecer, quer em termos naturais, patrimoniais e ambientais, em todo o país condições impares para atracão turística, o seu desenvolvimento tem sido diferenciado de região para região. Fruto dessa diferenciação é o número de estabelecimentos hoteleiros e do número de hotéis que existem nas várias regiões de Portugal.

A hotelaria é considerada por muitos a espinha dorsal da actividade turística. Como se pode observar através do **Quadro 2** e do **Gráfico 1**, de 1998 a 2007 (período de 20 anos), em

Portugal Continental registou-se uma evolução positiva do número de estabelecimentos hoteleiros (estabelecimento cuja actividade principal consiste na prestação de serviços de alojamento e de outros serviços acessórios ou de apoio, com ou sem fornecimento de refeições, mediante pagamento: hotéis, pensões, estalagens, pousadas, móteis, hotéis-apartamentos, aldeamentos turísticos e apartamentos turísticos).

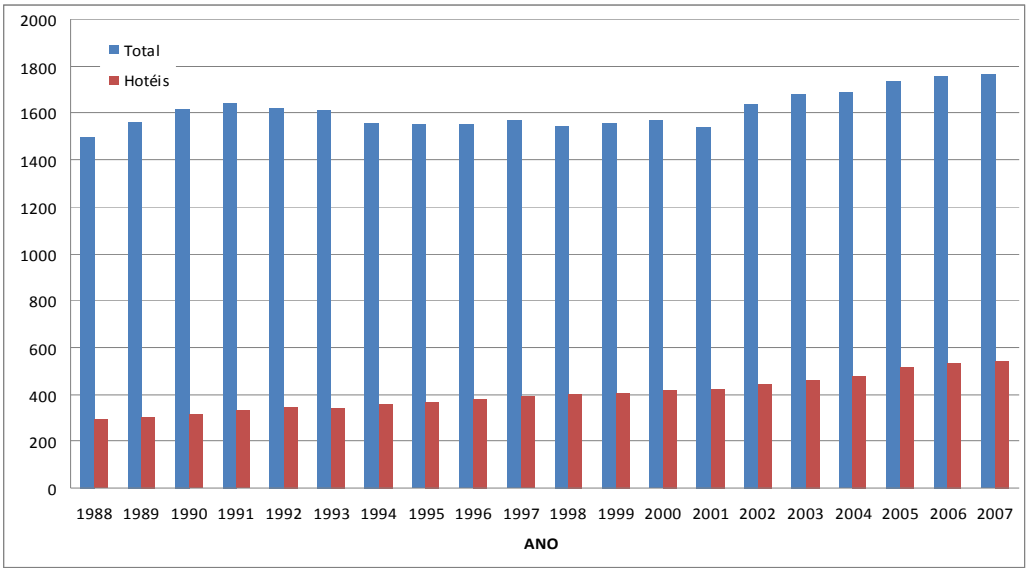
Quadro 2. Evolução do número total de estabelecimentos hoteleiros e do número de Hotéis por NUT`s II (1988 -2007)

ANO	1988		1989		1990		1991	
	Total	Hotéis	Total	Hotéis	Total	Hotéis	Total	Hotéis
Continente	1499	296	1564	306	1620	316	1646	330
Norte	373	78	402	79	417	82	422	85
Centro	283	56	288	59	294	61	298	66
Lisb. V T.	456	104	461	107	464	111	464	115
Alentejo	74	7	76	7	81	7	83	7
Algarve	313	51	337	54	364	55	379	57
	1992		1993		1994		1995	
	Total	Hotéis	Total	Hotéis	Total	Hotéis	Total	Hotéis
Continente	1625	347	1612	345	1561	360	1553	369
Norte	413	88	418	88	392	86	379	85
Centro	285	65	287	64	281	68	272	68
Lisb. V T.	461	125	434	119	428	127	436	132
Alentejo	79	8	89	10	82	12	88	13
Algarve	387	61	384	64	378	67	378	71
	1996		1997		1998		1999	
	Total	Hotéis	Total	Hotéis	Total	Hotéis	Total	Hotéis
Continente	1553	379	1569	392	1548	403	1561	408
Norte	383	88	391	90	390	92	395	95
Centro	274	75	277	77	267	78	264	74
Lisb. V T.	425	133	413	137	407	144	412	151
Alentejo	92	13	103	16	100	15	102	16
Algarve	379	70	385	72	384	74	388	72
	2000		2001		2002		2003	
	Total	Hotéis	Total	Hotéis	Total	Hotéis	Total	Hotéis
Continente	1568	422	1544	427	1640	447	1682	464
Norte	393	99	392	100	436	107	435	107
Centro	264	78	254	82	370	116	393	125
Lisb. V T.	414	154	410	154	288	123	293	126
Alentejo	105	17	104	17	118	22	131	24
Algarve	392	74	384	74	428	79	430	82
	2004		2005		2006		2007	
	Total	Hotéis	Total	Hotéis	Total	Hotéis	Total	Hotéis
Continente	1689	479	1738	518	1758	532	1763	547
Norte	435	106	450	118	452	122	461	127
Centro	398	129	418	141	425	151	427	155
Lisb. V T.	301	135	303	136	304	135	306	139
Alentejo	130	24	134	30	142	30	154	35
Algarve	425	85	433	93	427	94	415	91

Fonte: INE (1988 a 2007).

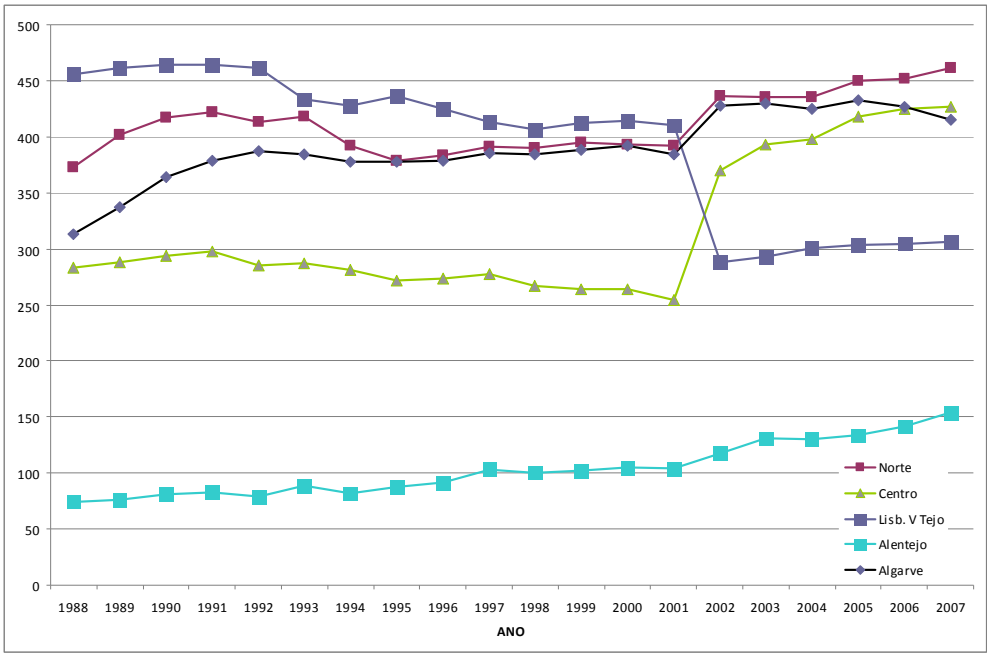
Dos vários tipos de estabelecimentos hoteleiros existentes em Portugal, os hotéis são os estabelecimentos que existem em maior número. Observando o **Quadro 2** e o **Gráfico 1**, podemos observar que ao longo dos últimos 20 anos, o número este tipo de estabelecimento aumentou consideravelmente no continente.

Gráfico 1. Evolução do número total de estabelecimentos de alojamento e do número de Hotéis (1988 - 2007)



Fazendo uma análise da evolução do número de estabelecimentos hoteleiros por NUT`s II, através da análise do gráfico nº 2 podemos contactar, que para o período em análise (1998 a 2007) e no conjunto das cinco NUT´s II, apenas na região de Lisboa e Vale do Tejo se registou uma evolução negativa. Ao contrário do que aconteceu na região de Lisboa e Vale do Tejo, a partir do ano de 2001 foi na região Centro onde se registou um maior aumento no número de estabelecimentos hoteleiros. É de salientar que apesar da sua tendência sempre crescente é na região do Alentejo onde ao longo destes vinte anos sempre existiu um menor número de estabelecimentos hoteleiros.

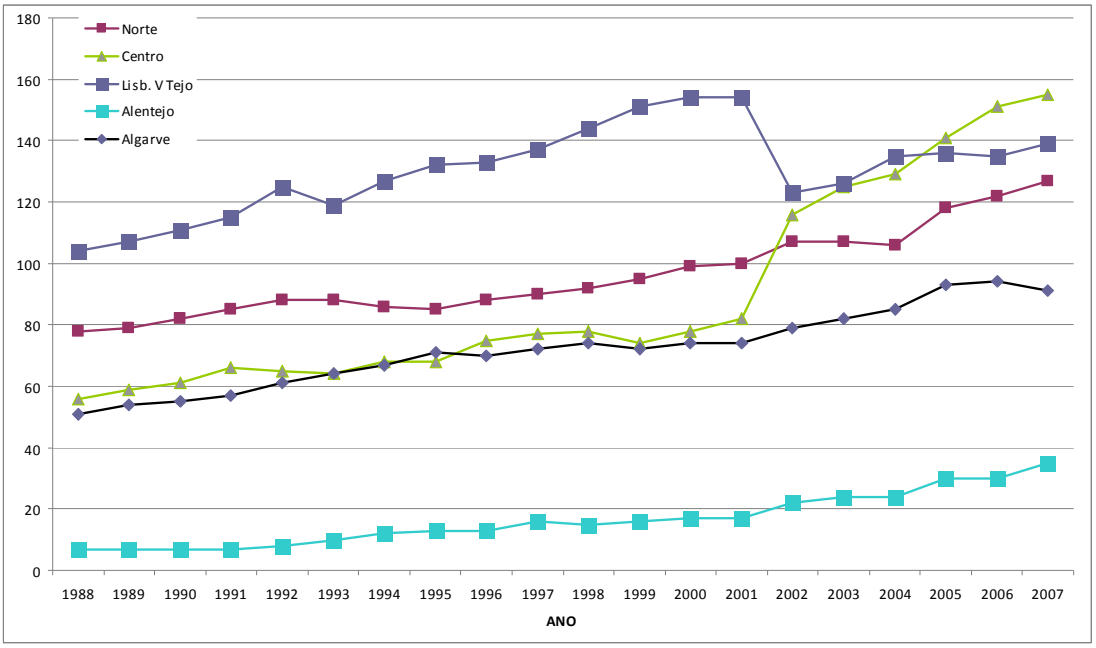
Gráfico2. Evolução do número total de estabelecimentos de alojamento por NUT's II (1988 - 2007)



Tal como acontece para Portugal Continental, ao nível das NUT`s II também são os hotéis o tipo de alojamento que existe em maior número. Comparando o ano de 1998 com o ano de 2007, em todas as NUT`s II se registou uma variação positiva do número de hotéis. Contudo, é

de salientar o facto que em todas as NUT`sII, ao longo do período em análise a tendência for sempre crescente, excepto na região de Lisboa e Vale do Tejo onde de 2001 para 2002 se registou uma diminuição do número destes estabelecimentos. Contudo, para essa região a partir de 2003 a tendência foi novamente crescente. Mas uma vez, sem qualquer surpresa é na região do Alentejo onde existe um menor número de hotéis ao longo deste vinte anos.

Gráfico 3. Evolução do número total de Hotéis por NUT's II (1988 - 2007)



Se analisarmos o número de estabelecimentos hoteleiros que existem em cada NUT`s II, mas agora tendo em conta a área de cada região, contacta-se que ao longo do território continental existe uma grande discrepância em termos da concentração destes estabelecimentos.

Quadro 3. Número de estabelecimentos hoteleiros e de hotéis por100 Km^2 por NUT`s II, em 2007

	Área Km^2	Nº Estabelecimentos Hoteleiros por 100 Km^2	Nº Hotéis em por 100 Km^2
Continente	88 633	1,98	0,612
Norte	21 278	2,16	0,59
Centro	28 405	1,50	0,54
Lisboa e V. Tejo	2 802	10,9	4,96
Alentejo	31 152	0,49	0,11
Algarve	4 996	8,31	0,18

Fonte: Elaboração própria.

Conforme se pode ver através do **Quadro 3**, é na Região de Lisboa e Vale do Tejo e na região do Algarve onde existe uma maior concentração, quer do número de estabelecimentos hoteleiros, quer do número de hotéis por km^2 . Em 2007, em cada 100 km^2, na região do Algarve existiam oito estabelecimentos hoteleiros, na região de Lisboa e Vale do Tejo onze, enquanto que na região do Alentejo aproximadamente um estabelecimento hoteleiro. Quando analisamos o número de hotéis, a realidade mantém-se, sendo a região de Lisboa, onde existe uma maior concentração.

Perante tais assimetrias regionais, é imprescindível a necessidade de desenvolver e dotar todas as regiões de Portugal em termos turísticos. Cada região possui características específicas, mas todas elas com grandes potencialidades que podem e devem contribuir para o

desenvolvimento económico e social do nosso país, como um todo. Tendo como objectivo último promover e valorizar a actividade turística, em 2008, com a publicação do Decreto-Lei n.º67/2008, procedeu-se a uma reorganização das Entidades Públicas Regionais com responsabilidades na área do turismo (MEI, 2008).

5 – PERSPECTIVAS REGIONAIS DO TURISMO PORTUGUÊS

No âmbito do Programa de Reestruturação da Administração Central do Estado (PRACE), e tendo presente a necessidade de dotar os organismos públicos na área do turismo de competências indispensáveis à afirmação de Portugal como um dos principais destinos turísticos europeus, o Governo aprovou, através da Resolução do Conselho de Ministros n.º 39/2006, de 21 de Abril, as orientações para a reestruturação dos vários ministérios, nas quais assume a necessidade de criar um organismo central do turismo único, este responsável pela prossecução da política de turismo nacional (o Turismo de Portugal, I.P.), e a descentralização das 19 regiões de turismo.

O Decreto-Lei n.º141/2007, de 27 de Abril definiu a missão e atribuições do Instituto do Turismo de Portugal, abreviadamente designado por Turismo de Portugal, I.P., autoridade turística nacional, concretizando-se assim o objectivo de criar uma única estrutura pública que tem por missão promover a valorização e a sustentabilidade da actividade turística nacional. O Turismo de Portugal, I.P., integrado no Ministério da Economia e da Inovação, cabe definir a estratégia nacional, contribuir para o desenvolvimento do sector e promover e valorizar a actividade turística, cuja missão consiste em:

- qualificar e desenvolver as infra-estruturas turísticas; - desenvolver a formação de recursos humanos;
- apoiar o investimento no sector;
- coordenar a promoção interna e externa de Portugal como destino turístico e,
- regular e fiscalizar os jogos de fortuna e azar.

Em 2008, com a publicação do Decreto-Lei n.º67/2008, de 10 de Abril, procedeu-se a uma reorganização das Entidades Públicas Regionais com responsabilidades na área do turismo, passando-se de dezanove[1] regiões do turismo para apenas 5 áreas regionais que reflectem as unidades territoriais utilizadas para fins estatísticos NUTS II – Norte, Centro, Lisboa e Vale do Tejo, Alentejo e Algarve. Para além das cinco áreas regionais, de acordo com o artigo 4º do Decreto-Lei n.º 67/2008 (MEI, 2008: 2172), "nas áreas regionais de turismo definidas no artigo 2.º são criados os pólos de desenvolvimento turístico" (Douro, Serra da Estrela, Leiria-Fátima, Oeste, Alentejo Litoral e Alqueva). Assim, para além das duas direcções regionais do turismo (regiões autónomas dos Açores e da Madeira), actualmente, temos onze entidades regionais de turismo que asseguram o desenvolvimento do turismo no território continental (**Quadro 4**).

Quadro 4. Entidades Públicas Regionais com responsabilidade na área do turismo

Regiões Turísticas	Pólos de desenvolvimento turístico
Turismo do Porto e Norte de Portugal	Turismo do Douro
Turismo do Centro de Portugal	Turismo da Serra da Estrela
Turismo de Lisboa e Vale do Tejo	Turismo de Leiria – Fátima
Turismo do Alentejo, E.R.T	Turismo do Oeste
Turismo do Algarve	Turismo Terras do Grande Lago Alqueva – Alentejo
	Turismo do Alentejo Litoral

Fonte: MEI(2008).

[1] As antigas dezanove regiões do turismo: Algarve, Alto Douro (Costa Verde), Alto Tâmega e Barroso, Centro, Dão Lafões; Douro Sul, Évora, Leiria-Fátima, Nordeste Transmontano, Oeste, Planície Dourada, Ribatejo, Rota da Luz, São Mamede (Norte Alentejo), Serra da Estrela, Serra do Marrão, Setúbal (Costa Azul), Templários (Floresta Central e Albufeiras) e Verde Minho.

Com o objectivo promover externamente a actividade turística, existem sete Agências Regionais de Promoção Turística (ARPT). Estas agências são associações de direito privado, sem fins lucrativos, constituídas por representantes dos agentes económicos do turismo e por empresas privadas com actividade turística. Existe uma agência em cada uma das 7 áreas promocionais (Norte, Centro, Lisboa Região, Alentejo, Algarve, Açores e Madeira) e são responsáveis pela elaboração, apresentação e execução dos respectivos planos regionais de promoção turística externa.

Pretendendo-se um País mais equilibrado e mais coeso territorialmente, para Silva (2000: 54), é ponto assente que "o desenvolvimento racional e equilibrado do turismo é fundamental para ao nosso Pais, já que, tratando-se de um sector gerador de infra-estruturas, de equipamentos, de actividades e emprego, pode proporcionar uma verdadeira cadeia de riqueza com repercussões directas nas economias regionais e na economia nacional".

6 - TURISMO SUSTENTÁVEL

È certo que o turismo se tornou num fenómeno de alcance mundial, ganhando cada vez mais relevo em termos económicos, sociais e ambientais. Estamos todos convictos de que apesar de o seu desenvolvimento trazer grandes benefícios para os países, o desenvolvimento desta actividade promotora pode conduzir a um vasto conjunto de problemas de ordem ambiental, económico e social.

Segundo a Comissão do Desenvolvimento Sustentável (CES, 2001: 3) "em consequência de uma rápida expansão do sector turístico, tanto nos tradicionais como nos novos destinos turísticos, enfrenta-se uma pressão cada vez maior sobre as dimensões naturais, culturais e socioeconómicas. Admite-se que um crescimento incontrolado de um turismo destinado a obter benefícios a curto prazo tende a ter consequências negativas, na medida em que danifica o meio ambiente, as sociedades e destrói as bases sobre as quais o turismo assenta e se desenvolve".

São vários os exemplos onde um desenvolvimento descontrolado do turismo acarretou impactos ambientais negativos, perda da identidade local, bem como outras ameaças ao meio envolvente. Perante tais factos, tem existido uma maior consciência por parte dos governos, instituições públicas, organismos não governamentais, bem como dos vários intervenientes da actividade turística, que têm levado a que o desenvolvimento sustentável esteja no centro das atenções. Contudo, devido á elevada complexidade de funcionamento da actividade turística, os múltiplos agentes que nela interferem, a escassez de tecnologias, falta de recursos e de conhecimentos por parte de alguns agentes económicos, ainda nos deparamos com grandes dificuldades de acção para a prática de um desenvolvimento sustentável no turismo.

Pretende-se que no futuro, Portugal seja um País competitivo e moderno "mas também um País com mais qualidade ambiental, mais equilibrado e coeso territorialmente, onde as populações encontrem qualidade de vida e as actividades económicas se desenvolvam no respeito pelos valores ambientais e de sustentabilidade" (MOPTC, 2005: 55). O conceito de desenvolvimento sustentável, isto é, o desenvolvimento que responde às necessidades das gerações presentes sem comprometer a possibilidade de satisfazer essas mesmas necessidades às gerações futuras, estende-se ao turismo. Segundo a Organização Mundial do Turismo, o desenvolvimento do turismo sustentável vai de encontro às necessidades das regiões turísticas, no sentido em que protege e aumenta as oportunidades de desenvolvimento dessas mesmas regiões.

O futuro do sector do turismo deve passar por uma perspectiva de sustentabilidade ambiental, económica e social, privilegiando-se a qualidade e a prática de acções socialmente responsáveis. Além disso, e de forma a maximizar os benefícios do turismo e minimizar as mudanças adversas e os impactos negativos "devem ser definidas medidas sociais e culturais como parte dos recursos turísticos, incorporadas no planeamento, no desenvolvimento e nos processos de gestão do turismo, dado que a gestão efectiva dos impactos sociais e dos impactos culturais é vital para obter uma indústria turística sustentável" (Baptista, 1997: 507).

7 – CONSIDERAÇÕES FINAIS

O turismo é, por um lado, uma actividade económica complexa, na medida em que agrega em si um conjunto de actividades, produtos e recursos de variada natureza. Por outro lado, é uma actividade económica que envolve uma multiplicidade de agentes que reproduzem diversos efeitos, que se fazem sentir transversalmente por toda a economia. Assim, dada a dimensão das receitas geradas, a sua importância na criação de postos de trabalho, bem como os efeitos multiplicadores que induz em vários sectores da economia, faz com que actualmente o turismo seja considerado como um sector estratégico na promoção do desenvolvimento de Portugal.

Não é só pelo facto de o turismo ter uma importância verdadeiramente estratégica para a economia portuguesa, mas em termos governamentais, temos assistido a um incremento de esforços no sentido de dinamizar e diversificar essa actividade económica por todo o território nacional. Contudo, o turismo, tal como acontece com outros sectores da economia, consome recursos e gera desperdícios de natureza económica, social e ambiental, devendo assim contabilizar-se à sua conta, quer através dos seus efeitos de natureza positiva, quer dos seus efeitos de natureza negativa. A este nível pouco ou nenhuma investigação foi desenvolvida, o que dificulta a focalização das políticas e estratégias governativas.

Assim, as preocupações, em particular, com os valores, a ética, o meio ambiente, o emprego, a utilização eficiente dos recursos disponíveis e, em geral, com o desenvolvimento sustentável desta actividade económica, contribuem para que, cada vez mais, a responsabilidade social empresarial se esteja a tornar um fenómeno de gestão empresarial. Efectivamente, as empresas de turismo ao presenciarem os efeitos da era da globalização económica sentem-se "obrigadas" a adoptar diferentes políticas, incorporando nas suas estratégias novos valores, novos padrões de comportamento e novas visões de negócio que conduzem a práticas de responsabilidade social que promovem, não só o seu volume de negócios, mas contribuem de forma equilibrada para um melhor ambiente.

Este ambiente com a focalização no conceito de desenvolvimento sustentável, isto é, o desenvolvimento que responde às necessidades das gerações presentes sem comprometer a possibilidade de satisfazer essas mesmas necessidades às gerações futuras, não pode deixar de entender-se e difundir-se também ao nível do turismo, com todas as vantagens que lhe estão inerentes. A preocupação crescente com a sustentabilidade da actividade turística, torna esta actividade económica num "filão de ouro" que importa descobrir.

BIBLIOGRAFIA

ALEXANDRE, J.A. (2003). *O turismo em Portugal: evolução e distribuição*. Tese de Mestrado, Universidade de Aveiro.

BAPTISTA, M. (1997). *Turismo Competitividade Sustentável*. Porto: Editora Verbo.

BURNAY, M.J. (2000). *Turismo e Natureza*. Livro de Actas do Ciclo de Debates do IFT. Instituto de Turismo de Portugal, 25-34.

COMISSÃO DAS COMUNIDADES EUROPEIAS (CCE, 2006). *Uma política de turismo europeia renovada: Rumo a uma parceria reforçada para o turismo na Europa, COM (2006) 134 final*. Bruxelas: Publicações Oficiais das Comunidades Europeias.

CONSELHO ECONÓMICO E SOCIAL (CES, 2001). *Sustainable Development of Tourism*. E/CN.17/2001/PC/21.

CUNHA, L. (2006). *Economia e Política do Turismo*. Lisboa: Editorial Verbo.

INSTITUTO NACIONAL DE ESTATÍSTICA (1988 a 2007). Anuários estatísticos das NUT`S II. Lisboa: INE:

MINISTÉRIO DA ECONOMIA E DA INOVAÇÃO (MEI, 2008). Decreto-Lei n.º67/2008, aprova o regime jurídico das áreas regionais de turismo de Portugal Continental e dos Pólos de Desenvolvimento Turístico, a delimitação e características. *Diário da República*, 1.ª Série, 71, 10 de Abril.

MINISTÉRIO DA ECONOMIA E DA INOVAÇÃO (MEI, 2007). *Plano Estratégico Nacional do Turismo*. Lisboa: Turismo de Portugal.

MINISTÉRIO DAS OBRAS PÚBLICAS, TRANSPORTES E COMUNICAÇÕES (MOPTC, 2005). *Grandes Opções do Plano 2005-2009*. Lisboa: MOPTC.

EVOLUÇÃO DO SECTOR DO TURISMO EM PORTUGAL: RELEVÂNCIA ECONÓMICA E ESTRATÉGICA

ERMELINDA OLIVEIRA
E-mail: ermelindaol@ipg.pt
Instituto Politécnico da Guarda
JOSE PIRES MANSO
E.mail : pmanso@ubi.pt
Universidade da Beira Interior

Resumo

Esta investigação analisa a evolução do sector do turismo em Portugal. Este tem assistido a importantes alterações na forma de actuar dos seus agentes económicos, com vista ao seu desenvolvimento sustentável. Este sector assume o papel de motor do desenvolvimento regional e, simultaneamente, de factor de expansão económica. A nível nacional, dado o aumento do número de turistas, por parte da procura, e do número de agentes económicos, por parte da oferta, a importância estratégica do turismo, traduzida nas receitas que proporciona, no seu contributo para a taxa de emprego da população, bem como nos efeitos multiplicadores que induz nas mais variadas áreas, tem levado à adopção de medidas inovadoras e dinamizadoras do crescimento económico. Assim, face à importância estratégica deste sector, a presente investigação discute sobre a expansão significativa da oferta turística, que tem sido acompanhada de uma maior diversificação, bem como de uma maior aposta na qualidade dos serviços prestados. Para além de concluir que, em consequência dessa rápida expansão, o sector enfrenta problemas cada vez maiores quanto às dimensões naturais, culturais e socioeconómicas.

Palavras-chave: Turismo, Desenvolvimento Sustentável, Estratégia Económica.
Área Temática: Estratégia Económica.

Abstract

This research examines the evolution of the Portuguese Hospitality Sector. It has been affected by changes of operators acting, because of its sustainable development perspective. This sector assumes to promote strategies of regional development, despite it is classify as the economic expansion factor. At national level, the increasing of demand in number of tourists, and, on the supply side, the number of operators shows the strategic importance of Hospitality Sector, reflected in the revenue it provides, in its contribution to the employment. The effects, in various areas, have led to the adoption of innovative and dynamic economic growth. Thus, given the strategic importance of this sector, this research discusses the significant expansion of tourism, which has been accomplished by greater diversification, as well as, greater service quality. Apart from that, in consequence of such rapid expansion, the sector faces increasing problems as the natural, cultural and socioeconomic dimension.

Key words: Tourism, Sustainable development, Economic strategy.
Thematic Area: Economics of Strategy.

1 -INTRODUÇÃO

Apesar de desde sempre o homem se deslocar de um lado para o outro, a natureza, a dimensão e os fins das suas viagens, foram-se alterando ao longo do tempo. Como actividade económica, o turismo é um fenómeno do século XVIII, devido ao progresso económico e social verificado neste século. Contudo, foi no século XX que o turismo se transformou num fenómeno da sociedade, alcançando uma dimensão económica nunca vista. A partir desta época, dada a importância que o turismo já assumia, quase todos os países da Europa criaram instituições governamentais e privadas com o fim de promover e organizar o turismo.

Apesar de em Portugal, sempre existirem condições que favorecem o desenvolvimento do turismo, a coexistência de outros factores nefastos ao desenvolvimento do turismo enquanto actividade económica, fez com que este se tenha desenvolvido com um atraso temporal considerável, comparativamente a outros países europeus. Segundo Cunha (2006), em termos temporais, existem quatro etapas marcantes e distintas quando em questão está a análise do desenvolvimento do turismo português. Por outro lado, apesar de Portugal oferecer, quer em termos naturais, patrimoniais e ambientais, em todo o país condições impares para atracão turística, o seu desenvolvimento tem sido diferenciado de região para região.

Pelo facto de o turismo ter uma importância verdadeiramente estratégica para a economia mundial, bem como para a economia portuguesa, na presente comunicação será feita uma análise da evolução histórica do turismo quer em termos mundiais, quer em termos nacionais. De seguida, e porque a hotelaria é considerada a espinha dorsal da actividade turística, será efectuada uma análise critica da evolução do número de estabelecimentos hoteleiros existentes em Portugal, por NUT`s II, ao longo dos últimos vinte anos. Antes de uma breve consideração final, será abordada a importância do turismo de desenvolver de uma forma sustentável, pelo facto de cada vez mais este sector enfrentar problemas cada vez maiores nas dimensões naturais, culturais e socioeconómicas.

2 – PERSPECTIVA HISTÓRICA DO TURISMO

O desejo de conhecer e estabelecer relações com outras civilizações, foi sempre uma constante na história do homem. Pelos mais variados motivos, nomeadamente, por razões comerciais, politicas, religiosas, de incremento do próprio território, a história da humanidade está estreitamente ligada às deslocações e às viagens. Contudo, segundo Cunha (2006: 35) "a natureza das viagens, os seus fins, a sua dimensão e extensão, bem como a forma que assumiram, foram adquirindo características muito diferentes de época para época até alcançarem aquilo a que se chama «Idade da Viagem», tal como se fala da Idade da Agricultura, da Industria ou da Electrónica." Segundo o mesmo autor, e tendo em conta "as características da evolução registada ao longo dos tempos podemos identificar três épocas históricas do turismo: a idade clássica, a idade moderna e a idade contemporânea" (2006:35).

2.1 - Idade Clássica do Turismo

A idade clássica do turismo, que vai desde os primórdios das primeiras civilizações e se prolonga até aos princípios do século XVIII, caracteriza-se pelo facto das viagens serem individuais e se realizarem, predominantemente, por necessidades fundamentais como o comércio, as peregrinações religiosas, a saúde ou por razões políticas e de estudo (Cunha, 2006). Apesar de não ser possível, em termos temporais, delimitar o início das primeiras viagens, podemos atribuir aos Sumérios o mérito de terem proporcionado as primeiras condições que possibilitaram a realização de viagens. A eles se deve a invenção da moeda, o desenvolvimento do comércio, bem com a invenção da roda que veio impulsionar a construção de estradas.

Foram os Romanos (150 a.C.) que instruíram a maior rede de estradas até então existente. Segundo cunha (2006:36), "provavelmente, a primeira viagem feita, com objectivos que hoje consideraríamos como turísticos, foi realizada pela rainha Hatshepsut cerca de 1500 a. C., tal como se descreve no templo Dar el Bahari situado no deserto próximo de Luxor." Já naquela época os Romanos e os Gregos, viajavam com a finalidade de conhecer templos e as sete maravilhas mediterrânicas, nomeadamente as pirâmides e os monumentos do Egipto. Com os romanos desenvolveram-se os *Hospes* (estalagens), os *Hospitium* (Hotéis) e os *Hospitalia* (Estalagens Públicas), bem como o espírito da hospitalidade, que passou a ser um acto honroso, e institui-se a obrigação de receber com benevolência os estrangeiros que chegassem a uma cidade. Em Roma, a 25 a.C., com o desenvolvimento das instalações termais, desenvolveram-se verdadeiros centros de turismo, que se prologaram até aos nossos dias.

Com os Cristãos, a hospitalidade continuou a ser um dever e um direito sagrado que devia ser gratuitamente concedida, tendo-se no século IV aberto as primeiras casas de refúgio para os viajantes. De acordo com Cunha (2006) nesta época, as viagens tinham como principal razão as peregrinações sendo célebres as que se dirigiam a Santiago de Compostela (Espanha), a Canterbury (Inglaterra), à terra Santa (Palestina) e a Meca (Arábia). Já nesta época existam guias de viagens para fornecer aos peregrinos informações sobre as regiões, alojamentos e o caminho a percorrer.

No século XIII, é de referir a grande viagem de Marco Pólo que, durante 24 anos, percorreu o Oriente até a China. Já no século XV, o século dos descobrimentos portugueses, os portugueses percorreram toda a costa de África e mar Vermelho, chegaram à Índia, China, Japão, Malaca, Timor e descobriram o Brasil. Posteriormente aos portugueses, seguiram-se as grandes viagens dos Espanhóis (chegaram às Caraíbas, Antilhas e América Central e Sul), dos Ingleses (descobriram a América do Norte), Franceses e Holandeses. Com estas descobertas, iniciadas pelos portugueses, chegou uma nova era para o mundo das viagens.

2.2 - Idade Moderna do Turismo

Ao longo do período histórico que abarca a Antiguidade e a Idade Média, o formato e os motivos das viagens conservaram as mesmas particulares, não se conseguindo diferenciar com clareza estas duas épocas. Foi a partir de meadas do século XVIII que as grandes mudanças verificadas em termos económicos, sociais, culturais bem como a nível tecnológico, vieram provocar modificações significativas no âmbito das viagens. Em toda a Europa, desenvolveu-se a construção de estradas e das redes de comunicação. Em Inglaterra, com os primeiros passos da revolução industrial, incrementaram-se as relações internacionais, pois com a invenção da máquina a vapor, promoveu-se uma revolução nos meios de transportes.

Nesta época, estudantes, diplomatas e membros das famílias ricas inglesas, começaram a fazer o *Grand Tour,* viajando pela Europa, durante um período de normalmente três anos. "Com a *Grand Tour* nasce o conceito de turismo e, pela primeira vez, começam a designar-se as pessoas que viajam por turistas" (Cunha, 2006: 41). Nesta época verificou-se um incremento no número de livros dedicados às viagens, bem como no número dos guias turísticas. Por outro lado, o movimento dos ingleses para a Europa influenciou muito o desenvolvimento dos transportes, da hotelaria e da restauração

No século XIX, a proliferação das trocas, o desenvolvimento dos transportes, bem como a propagação de ideias com a difusão de publicações, vieram dar "um novo impulso às viagens que começam a encontrar a sua verdadeira identidade: um meio de as pessoas se interessarem pelas particularidades de cada povo, pelas tradições, pelo exotismo e por outros modos de vida e novas culturas" (Cunha, 2006: 42).

Em 1841, nasceu o turismo organizado com Thomas Cook, quando este organizou a primeira viagem colectiva com a duração de um dia, destinada a 570 passageiros num congresso de médicos. As iniciativas de Cook (em 1864 organizou a primeira excursão acompanhada no regime «tudo incluído»; em 1872 organizou a primeira viagem à volta do mundo que durou 222 dias), marcaram "uma das mais importantes etapas na história do turismo e estão na origem do turismo dos nossos dias, continuando a agência por ele criada a ser uma das maiores organizações turísticas do mundo" (Cunha, 2006: 42).

No decorrer do século XIX, popularizaram-se as viagens, deu-se a criação de inúmeras agências de viagem e o turismo começa a surgir como uma actividade organizada. Em 1883 o austríaco Stadner falou pela primeira vez na «indústria do turismo» e em 1896 Frenlu publicou o documento «Contribuições para uma Estatística do Turismo».

A primeira década do século XX caracterizou-se por transformações e inovações que alteraram profundamente os modos de vida das populações, nomeadamente (Cunha, 2006).

- A descoberta do telégrafo e do telefone;
- O alargamento da rede de caminhos de ferro;
- O aumento das extensões das redes de estradas;
- O grande desenvolvimento industrial;

- A racionalização do trabalho e as reivindicações sindicais, que levaram a uma maior democratização das sociedades e a novos conceitos de vida;
- O tempo de trabalho diminui e alcança-se o direito ao repouso semanal, dai o conceito de lazer surgir com uma nova noção;
- A Organização Internacional de Trabalho (OIT) estabelece numa convenção o princípio das férias pagas, posteriormente reconhecido pela Declaração Universal dos Direitos do Homem.

Assim, ao longo de todo o século XX, o turismo transformou-se num fenómeno da sociedade, alcançou uma dimensão económica até então nunca vista, influenciando os comportamentos das pessoas na sociedade. A partir desta época, dada a importância que o turismo já assumia, em quase todos os países da Europa criaram-se instituições governamentais e privadas com o fim de promover e organizar o turismo. De entre muitas organizações, destaca-se a primeira organização internacional do turismo, de nome Federação Franco-Hispano-Portuguesa de Sindicatos de Iniciativa e Propaganda, transformando-se depois na União Internacional dos Organismos Oficiais de Turismo (UIOOT), e que mais tarde está na origem da actual Organização Mundial do Turismo (OMT). Perante a análise de todos os factos, podemos concluir que nesta época, já existiam condições para o desenvolvimento do turismo como actividade económica constante, bem como a sua universalização. Porém, a II Guerra Mundial veio atrasar o seu desenvolvimento e a sua universalização.

2.3 - Idade Contemporânea do Turismo

Apesar das ocorrências que atemorizaram o mundo (I Guerra Mundial, Grande Crise de 1229 e a Guerra Civil de Espanha), até ao início da II Guerra Mundial o turismo atingiu grandezas significativas. Contudo, durante o período que decorreu da II Guerra Mundial à época da recuperação económica, o turismo sofreu grandes dificuldades, entrando numa fase de declínio acentuado. Só mais tarde, a partir dos anos 50 e aliado à fase de progresso económico e social é que entrou novamente numa fase de desenvolvimento e consolidação. Com o objectivo de analisar com profundidade os factores que influenciaram o turismo, o perfil dos turistas, bem como o lado da procura turística, Cunha (2006) subdivide a época posterior a 1945, em três partes distintas: de 1945 a 1973, de 1973 a 1990 e finalmente, após 1990.

1º Período: 1945 a 1973

Neste período, das mais significativas alterações ocorridas destacam-se (Cunha, 2006):
- a ascensão de um grande número de países à independência;
- a produção mundial aumentou à média anual de 5%;
- um crescimento do rendimento real por habitante de 3%;
- um incremento muito significativo das trocas internacionais;
- os créditos internacionais, o capital, a tecnologia e a mão-de-obra registaram uma enorme mobilidade;
- verificou-se a emergência de grandes multinacionais bem como a constituição de grupos económicos (CEE);
- um aumento dos dias de férias pagas dos trabalhadores;
- o lançamento dos primeiros satélites, bem como a primeira viagem à Lua;

Os efeitos produzidos ao nível do turismo, por todas estas alterações, fizeram-se sentir em duas vertentes, no lado da procura e no lado da oferta turística. No lado da procura, registaram-se alterações por intermédio de modificações produzias ao nível do rendimento, da duração do tempo livre das pessoas e das suas motivações. Ao nível do rendimento é de registar o aumento verificado no rendimento considerado discricionário, na adopção de medidas sociais, bem como no incremento verificado na facilidade de compra de viagens.

No que concerne ao tempo livre, decorrente de uma diminuição do tempo de trabalho semanal, este aumentou e por outro lado, nos países desenvolvidos verificou-se a generalização

das férias pagas. Ao nível das motivações existiram alterações derivadas da necessidade de diversificação e diferenciação, bem como pela crescente necessidade pela fuga ao meio, pois as pessoas sentem cada vez mais a necessidade de reparar os desequilíbrios psicológicos decorrentes da sua vida profissional. No lado da oferta turística, salienta-se um incremento nas viagens aéreas e, por outro lado, os operadores turísticos desenvolveram a produção em série de produtos de massa. Este período caracteriza-se por uma procura turística de sol e mar. No lado da oferta, todas as preocupações se centraram no desenvolvimento do turismo internacional, sendo o turismo interno um subproduto do turismo internacional no seio das orientações das politicas turísticas.

2º Período: 1973 a 1990

Após um acelerado desenvolvimento da economia mundial, verificado no período compreendido entre 1945 a 1973, após este período registou-se:

- uma acentuada diferença entre o nível de vida dos países em vias de desenvolvimento e os países industrializados;
- um afrouxamento do crescimento económico mundial, derivado de uma diminuição da produção, acompanhado da estagflação e do aumento da taxa de desemprego;
- variações velozes das taxas de câmbio e uma crise de confiança no sistema monetário mundial, acompanhada com endividamento externo na generalidade dos países de tal nível que abalou os fundamentos do sistema financeiro internacional;
- os problemas mundiais tomaram uma tal amplitude que foi necessário criar condições para uma melhor cooperação mundial.
- pela primeira vez, os problemas ambientais (poluição da atmosfera, das águas, bem como derivados da utilização de produtos tóxicos), são encarados como um grande perigo para a humanidade. Tal constatação, levou a novas atitudes e comportamentos por parte dos consumidores.

Fruto destas ocorrências, o turismo mundial não diminuiu, mas o seu ritmo de crescimento abaixou e sofreu grandes alterações estruturais. Do lado da oferta, verificou-se uma diversificação, designadamente pela oferta de mais meios de animação e da prática de desporto. Por outro lado, e ainda ao nível da oferta, em detrimento do turismo internacional, neste período o turismo interno passou a adquirir maior importância. Ao nível da procura, a duração e a distância das viagens diminuíram e os alojamentos de mais baixo custo passaram a ser os mais procurados.

Segundo Cunha (2006: 48), neste segundo período "a nível conceptual, passou a enfatizar-se menos o papel económico do turismo no qual se tinha, até então, insistido em excesso para, igualmente, se atribuir importância ao seu papel social, politico, ecológico, cultural e educativo o que levou a passar a considerá-lo como uma das componentes essenciais da vida do homem. Deixou de ser unidimensional para passar a ser multidimensional, na medida em que responde a uma multiplicidade de necessidades humanas e não apenas à melhoria do bem-estar material. Os valores económicos do turismo, sem deixar de estar presentes nas preocupações do seu desenvolvimento, foram relegados para um plano de menor evidência dando lugar aos valores da identidade e valorização do homem."

3º Período: após 1990

Apesar de nos anos 80, a economia mundial ter registado taxas de crescimento, que vieram permitir a retoma dos mercados, a partir de 1992 a economia mundial entrou numa fase de recessão generalizada, considerada por muitos analistas como a maior crise económica posterior à II Guerra mundial. Contudo, num cenário de crise económica, com altas taxas de desemprego em todos os países, um clima de grande tensões politicas em termos internacionais e apesar dos grandes avanços sentidos para uma maior integração económica, o turismo resistiu-se mas mantêm a sua tendência de crescimento ainda que a um ritmo menos acelerado.

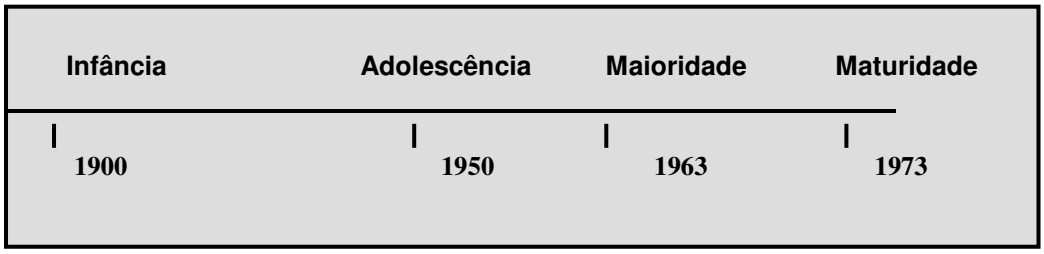
Entrando no século XXI, ao nível do turismo saliente-se os movimentos verificados em termos de reorganização empresarial (formação de grandes aglomerados económicos e de novos agentes), novas formas de turismo (turismo social), uma maior diversificação da oferta (novas formas de alojamento, novas formas de atracção turística, em suma, novos produtos e novos destinos.

3 – PERSPECTIVAS EVOLUTIVAS DO TURISMO PORTUGUÊS

Apesar de em Portugal, sempre existirem condições que favorecem o desenvolvimento do turismo (condições climatéricas, naturais, sentido de hospitalidade, peculiaridade da sua cultura, entre outras) a coexistência de outros factores nefastos ao desenvolvimento do turismo enquanto actividade económica (atraso nos meios de transporte e nas vias de comunicação, a inexistência de uma iniciativa privada e informada, condão pelo isolamento internacional), fez com que este se tenha desenvolvido com um atraso temporal considerável, comparativamente a outros países europeus.

Segundo Cunha (2006), quando em questão está a análise do desenvolvimento do turismo português, em termos temporais existem quatro etapas marcantes e distintas. À primeira etapa, que vai de 1900 a 1950 (50 anos), o autor chamou-lhe a «Fase da Infância». Ao segundo período, compreendido entre 1950 a 1963 (com uma duração de 13 anos), denominou-o por «Adolescência», seguindo-se a etapa da «Maioridade» entre1963 e 1973 e, por fim, a quarta e última etapa onde o turismo português atingiu a sua fase de «Maturidade» (**Quadro 1**).

Quadro 1. Etapas da Evolução do Turismo em Portugal



Fonte: Retirado de Cunha (2006: 88).

3.1 - Primeira Etapa – Fase da Infância

Até aqui, o turismo interno português assentava basicamente nas instâncias termais e o turismo internacional na Madeira e em Lisboa. O desenvolvimento do turismo interno e do turismo internacional português verificado nesta época, em muito se deve à Sociedade de Propaganda de Portugal (SPP), uma sociedade de iniciativa privada que foi criada em 1906. Esta sociedade tinha como objectivos, organizar e divulgar todo o património cultural e natural de Portugal, com a finalidade de promover quer o turismo interno, quero turismo internacional. Para o efeito, em 1920 a SPP abriu cerca de 143 representações, espalhadas por todo o país.

No seguimento da realização em Lisboa do IV Congresso Internacional de turismo, organizado pela Federação Franco-Hispano-Portuguesa de Iniciativa e Propaganda, em 1911 surgiu, em Portugal, a primeira organização oficial do turismo português, denominada por «Repartição do Turismo» e integrada no Ministério do Fomento.

Nesta primeira etapa, a desabrochamento da I Guerra Mundial (1914 – 1918), a Guerra Civil de Espanha (1936) e a eclosão da II Guerra Mundial (1936-1945), bem como os seus efeitos quer a nível económico, politico e também social, contribuíram para o reduzido numero de visitas de estrangeiros e para que o desenvolvimento do turismo português de adiasse. Apesar disso, neste período (Infância do Turismo) algumas iniciativas foram tomadas, nomeadamente:

- o Plano de aproveitamento do Estoril e o programa de construção de pousadas, que ainda nos nossos dias são uma imagem de marca do nosso país;

- em 1920 foi reforçada a estrutura da organização oficial do turismo com a criação do Comércio e Comunicações;
- subordinadas à Repartição do Turismo foram criadas em certas localidades de Portugal comissões, com o objectivo de estas protegerem os monumentos nacionais e cooperarem em estudos do turismo;
- em 1928 o jogo ficou ligado ao turismo e criou-se a Repartição de Jogos e Turismo, repartição que ficou integrada no Ministério do Interior;
- em 1930 criou-se a Comissão de Propaganda do turismo de Portugal, e nos anos seguintes foram criadas as Casas de Portugal em vários países, tendo-se assim lançado as primeiros representações de Portugal no estrangeiro.

3.2 - Segunda Etapa – Fase da Adolescência

No início dos anos 50, os países europeus saíram da guerra e entraram em plena recuperação económica e social, recomeçando-se assim uma nova era para o turismo. Portugal, apesar de ser um país detentor de condições climatéricas favoráveis e rico em termos paisagísticos, não acompanhou a nova era do turismo da Europa, pois encontrava-se numa situação de atraso em relação a um conjunto de meios materiais exigidos pela vida moderna (rede de estradas, capacidade hoteleira bem como um conjunto de atractivos turísticos).

Contudo, perante obstáculos e dificuldades, nesta época começou a aparecer uma nova concepção do turismo, passando este a ser entendido como uma verdadeira actividade económica. Segundo Cunha (2006: 82), nesta etapa "considerava-se que o turismo não deve ser tido como divertimento, mas sim como uma verdadeira indústria não podendo, por isso, ser obra de amadores, mas antes tem de se subordinar a princípios de organização administrativa semelhante aos de qualquer actividade económica."

Esta nova abordagem do turismo contribui para que nesta fase temporal:
- a rede de estradas aumentasse consideravelmente;
- em 1954 se publica-se a Lei n.º 2073 referente ao exercício da industria hoteleira, e que veio criar o estatuto de utilidade turística, permitindo a um conjunto de empresas hoteleiras o acesso a um conjunto de regalias, como forma de fomentar a capacidade hoteleira no nosso país;
- em 1956 se publica-se a Lei n.º 2082 que passou a ser a lei base do turismo que criou o Fundo de Turismo;
- em 1962 se criasse uma subsecretaria de Estado da Presidência do Conselho, com competências especificas na área do turismo. Foi aqui que pela primeira vez o turismo passou a ser tratado a nível governamental.
- se criassem zonas de turismo nos concelhos em que existiam praias, estâncias hidrológicas ou climáticas, estâncias de repouso e de recreio ou monumentos.

3.3 - Terceira Etapa – Fase da Maioridade

Apesar de ter sido a partir da década de 30 do século passado que se verificou o despertar da actividade turística, em Portugal, "por razões ligados a factos de conjuntura político-económica, as Grandes Guerra desse século, bem como à condução da vida portuguesa pouco inclinada à mudança, só a partir da década de 60, o turismo começou a ter expressão significativa no conjunto das actividades nacionais" (Alexandre, 2003:5). A partir de 1964 surgiu uma nova era para o turismo europeu e para o turismo português. Depois da recuperação económica dos países industrializados afectados pela guerra, estes estavam agora perante um boom económico, que veio permitir a vulgarização do automóvel, o desenvolvimento do trânsito aéreo bem como, a generalização das férias pagas.

Nesta época, em Portugal assistiu-se a uma expansão dos empreendimentos turísticos, de uma oferta hoteleira em maior número e mais diversifica, bem como ao desenvolvimento de novos centros turísticos (Algarve, Madeira e Tróia). Fruto do desenvolvimento, desordenado e

sem apoio das infra-estruturas indispensáveis, dos empreendimentos turísticos, surgiram os primeiros impactos nefastos em termos ambientais, bem como sinais de assimetrias regionais. Foi nesta época que surgiu a Direcção-Geral do Turismo e o turismo, que até aqui nunca tinha sido objecto de qualquer planeamento, passou a ser enquadrado nos Planos de Fomento.

A busca do sol e mar era dominante ao nível da procura, em detrimento do turismo do interior. Por outro lado, "é neste período que, pela primeira vez, as dormidas de estrangeiros na hotelaria ultrapassam as dormidas de nacionais alterando-se, simultaneamente, a estrutura da procura externa" (Cunha, 2006: 87).

3.4 - Quarta Etapa – Fase da Maturidade

Portugal, com a revolução de 1974 sofreu grandes transformações económicas, políticas e sociais. Fruto da revolução de 25 de Abril, um grande número de desalojadas das ex-colónias passaram a ocupar os alojamentos hoteleiros e a procura externa sofreu uma quebra acentuada ao longo da década de 70. Porém, após o início da década de 80 o sector reagiu e a procura turística externa tomou valores nunca vistos, provocando em paralelo um clima de euforia que levou a que decisões empresariais de campo imobiliário substituíssem e subestimassem os princípios do próprio turismo. "Apesar do extraordinário crescimento da procura, o turismo português passou a viver em estado de crise latente ou explícita derivada do excesso da oferta, da inadequação das infra-estrutura, do desordenado aproveitamento dos espaços, do desequilíbrio entre as várias componentes da oferta e de outros factores ligados às condições de recepção internas. É uma crise originada pelo lado da oferta e não pelo lado da procura" (Cunha, 2006: 91).

Se até aos anos 70 do século passado, o turismo era considerado como uma "indústria limpa", daí não lhe serem imputados efeitos ou externalidades negativas quer de ordem social, quer de ordem ambiental, a partir da década de 80 começaram a despontar vozes denunciantes dos efeitos adversos que a actividade turística exercia sobre o meio ambiente e sobre a população residente e receptora do turismo (Burnay, 2000). Esses efeitos negativos traduziam-se "na perda da qualidade de vida, nos desajustes das economias locais, no desmembramento da organização social, na ocultação dos povos e consequente perda de identidade" (Burnay, 2000: 26).

Segundo José Alexandre, (2003: 56) "no plano teórico, o turismo parecia encerrar qualidades susceptíveis de gerar dinâmicas de desenvolvimento endógeno, designadamente nas regiões de acolhimento. O turismo fazia o uso de recursos locais, em muitos casos sub-aproveitados, com fracas potencialidades de utilização económica, pelo que permitia uma exploração mais eficaz das potencialidades da região. Paralelamente, constituía um factor de diversificação da estrutura da região, multiplicando as oportunidades de emprego e aumentando, consequentemente, o produto e o nível de vida das populações locais. O turismo poderia ainda produzir importantes efeitos multiplicadores na economia regional, promovendo a dinamização e modernização do tecido económico em que se inseria, com reflexos no domínio cultural e garantindo os limiares mínimos necessários à criação de infra-estruturas e equipamentos". Contudo, pelo facto de o chamado turismo de massa, escorado pela procura externa, ser orientado por operadores turísticos, fugiu do alcance de controlo das autoridades locais, não se geraram dinâmicas de desenvolvimento nas regiões necessitadas. Efectivamente, o que se veio a verificar, foi que a actividade turística, orientada e implementada da forma como estava a ser feita, provocou graves efeitos no domínio social, cultural, ainda maiores desequilíbrios espaciais em termos de desenvolvimento e, paralelamente, produziram-se graves problemas ambientais.

Nesta época, e segundo Cunha (2006) o turismo português peca quer em termos de concentração territorial, quer em termos de concentração em motivações e atractivos. Concentração territorial pelo facto da região do Algarve em conjunto com a região de Lisboa, absorverem mais de metade das dormidas de estrangeiros e ser nestas regiões onde esta situado a grande maioria dos alojamentos hoteleiros. Concentração em atractivos e motivações por ser o sol e a praia o grande atractivo da oferta turística portuguesa. Consciente destas vulnerabilidades do turismo, e tendo como objectivos atenuar os desequilíbrios regionais, valorizar o património cultural, bem como, conseguir uma melhoria na qualidade de vida dos

portugueses, na década de 80 foi lançado um Plano Nacional de Turismo. Segundo Cunha (2006: 93) a filosofia adoptada neste Plano Nacional de Turismo "foi a transformação qualitativa do turismo com base na defesa da qualidade, da diversidade e da harmonização do aproveitamento do território".

3.5 – Após a década de noventa

Segundo Cunha (2006: 97) "a evolução do turismo português, a partir dão início da década de 90, indicia o esgotamento do modelo de desenvolvimento iniciado na década de 60 baseado na excessiva exploração do «sol e mar», no aproveitamento indiscriminado dos recursos naturais e na concentração dos investimentos junto ao litoral orientados, na perspectiva do negócio imobiliário".

Como consequência do esgotamento do modelo de desenvolvimento do turismo até aqui adoptado, no início do século XXI começaram a surgir esforços no sentido da diversificação e da desconcentração do turismo. Do lado da oferta têm-se assistido a um conjunto de medidas dinamizadoras, nomeadamente uma expansão significativa da oferta turística, acompanhada por uma maior diversificação (por exemplo, turismo rural e turismo de habitação) bem como uma maior aposta na qualidade dos serviços prestados. Paralelamente, temos assistido a uma aposta num desenvolvimento regional mais equilibrado ao nível das infra-estruturas necessárias para o desenvolvimento do turismo.

Em termos governamentais, existe uma grande aposta no sentido de dinamizar e diversificar o turismo, por todo o território nacional pelo facto de o turismo ter uma importância verdadeiramente estratégica para a economia portuguesa em virtude da sua capacidade em criar riqueza e emprego. "Trata-se de um sector em que temos vantagens competitivas claras como sucede com poucos outros" (MEI, 2007: 1). Tendo por fim último a promoção e o desenvolvimento do turismo em Portugal, de um turismo sustentável, o Plano Estratégico Nacional do Turismo (PENT), estabeleceu como objectivo até 2015 desenvolver o sector através das seguintes onze linhas de desenvolvimento estratégico (MEI, 2007):

1 - Produtos, destinos e pólos;
2 - Intervenções em Urbanismo, ambiente e Paisagem;
3 – Desenvolvimento de conteúdos distintos e inovadores;
4 – Eventos;
5 – Acessibilidade aérea;
6 – Marcas, promoção e distribuição;
7 – Programa da qualidade;
8 – Excelência no capital humano;
9 – Conhecimento e inovação;
10 – Eficácia no relacionamento Estado-empresa;
11 – Modernização empresarial.

Apesar da emergência de novos destinos, que têm afastado turistas dos mercados tradicionais, nomeadamente de Portugal, e ao contrário do que se verifica com muitos dos seus concorrentes, Portugal tem conseguido manter a sua participação a nível mundial. No ano de 2005, Portugal encontrava-se posicionado no 18º lugar no ranking dos principais destinos turísticos.

4 – ALGUNS INDICADORES DO TURISMO EM PORTUGAL

Apesar de Portugal oferecer, quer em termos naturais, patrimoniais e ambientais, em todo o país condições impares para atracão turística, o seu desenvolvimento tem sido diferenciado de região para região. Fruto dessa diferenciação é o número de estabelecimentos hoteleiros e do número de hotéis que existem nas várias regiões de Portugal.

A hotelaria é considerada por muitos a espinha dorsal da actividade turística. Como se pode observar através do **Quadro 2** e do **Gráfico 1**, de 1998 a 2007 (período de 20 anos), em

Portugal Continental registou-se uma evolução positiva do número de estabelecimentos hoteleiros (estabelecimento cuja actividade principal consiste na prestação de serviços de alojamento e de outros serviços acessórios ou de apoio, com ou sem fornecimento de refeições, mediante pagamento: hotéis, pensões, estalagens, pousadas, móteis, hotéis-apartamentos, aldeamentos turísticos e apartamentos turísticos).

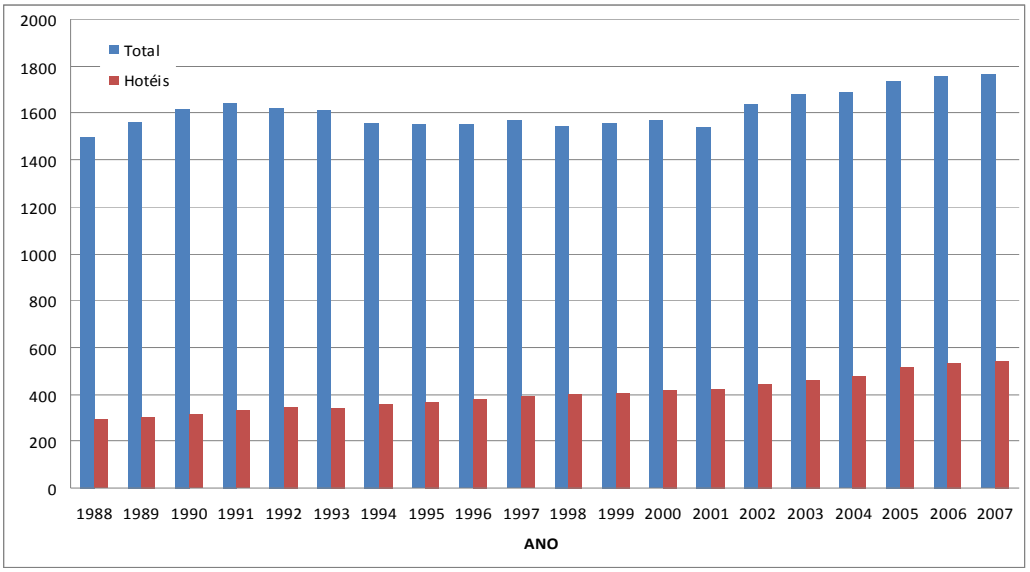
Quadro 2. Evolução do número total de estabelecimentos hoteleiros e do número de Hotéis por NUT`s II (1988 -2007)

ANO	1988		1989		1990		1991	
	Total	Hotéis	Total	Hotéis	Total	Hotéis	Total	Hotéis
Continente	1499	296	1564	306	1620	316	1646	330
Norte	373	78	402	79	417	82	422	85
Centro	283	56	288	59	294	61	298	66
Lisb. V T.	456	104	461	107	464	111	464	115
Alentejo	74	7	76	7	81	7	83	7
Algarve	313	51	337	54	364	55	379	57
	1992		1993		1994		1995	
	Total	Hotéis	Total	Hotéis	Total	Hotéis	Total	Hotéis
Continente	1625	347	1612	345	1561	360	1553	369
Norte	413	88	418	88	392	86	379	85
Centro	285	65	287	64	281	68	272	68
Lisb. V T.	461	125	434	119	428	127	436	132
Alentejo	79	8	89	10	82	12	88	13
Algarve	387	61	384	64	378	67	378	71
	1996		1997		1998		1999	
	Total	Hotéis	Total	Hotéis	Total	Hotéis	Total	Hotéis
Continente	1553	379	1569	392	1548	403	1561	408
Norte	383	88	391	90	390	92	395	95
Centro	274	75	277	77	267	78	264	74
Lisb. V T.	425	133	413	137	407	144	412	151
Alentejo	92	13	103	16	100	15	102	16
Algarve	379	70	385	72	384	74	388	72
	2000		2001		2002		2003	
	Total	Hotéis	Total	Hotéis	Total	Hotéis	Total	Hotéis
Continente	1568	422	1544	427	1640	447	1682	464
Norte	393	99	392	100	436	107	435	107
Centro	264	78	254	82	370	116	393	125
Lisb. V T.	414	154	410	154	288	123	293	126
Alentejo	105	17	104	17	118	22	131	24
Algarve	392	74	384	74	428	79	430	82
	2004		2005		2006		2007	
	Total	Hotéis	Total	Hotéis	Total	Hotéis	Total	Hotéis
Continente	1689	479	1738	518	1758	532	1763	547
Norte	435	106	450	118	452	122	461	127
Centro	398	129	418	141	425	151	427	155
Lisb. V T.	301	135	303	136	304	135	306	139
Alentejo	130	24	134	30	142	30	154	35
Algarve	425	85	433	93	427	94	415	91

Fonte: INE (1988 a 2007).

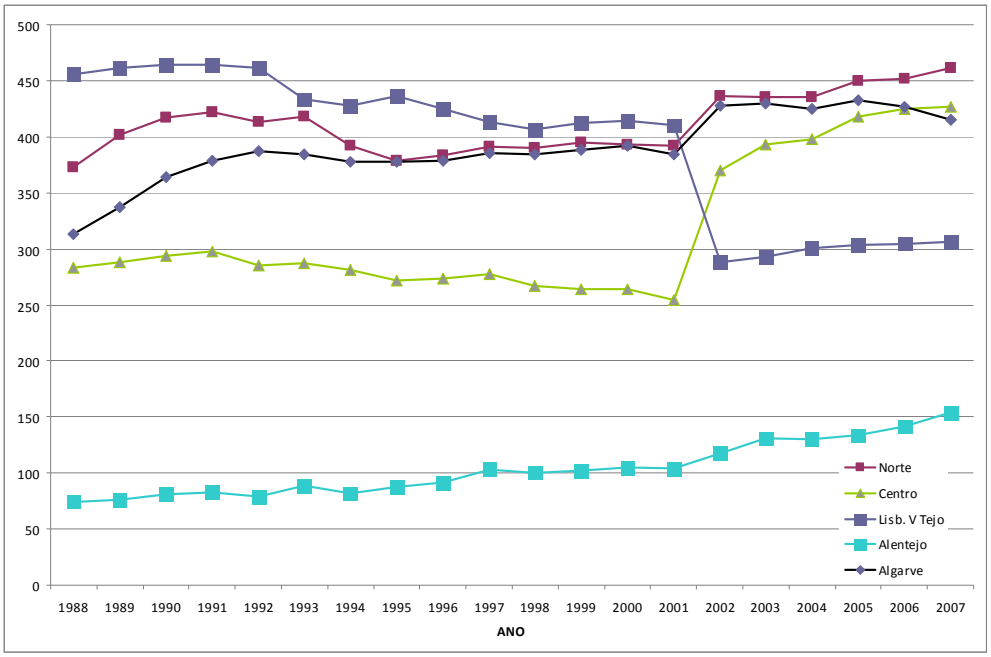
Dos vários tipos de estabelecimentos hoteleiros existentes em Portugal, os hotéis são os estabelecimentos que existem em maior número. Observando o **Quadro 2** e o **Gráfico 1**, podemos observar que ao longo dos últimos 20 anos, o número este tipo de estabelecimento aumentou consideravelmente no continente.

Gráfico 1. Evolução do número total de estabelecimentos de alojamento e do número de Hotéis (1988 - 2007)



Fazendo uma análise da evolução do número de estabelecimentos hoteleiros por NUT`s II, através da análise do gráfico nº 2 podemos contactar, que para o período em análise (1998 a 2007) e no conjunto das cinco NUT´s II, apenas na região de Lisboa e Vale do Tejo se registou uma evolução negativa. Ao contrário do que aconteceu na região de Lisboa e Vale do Tejo, a partir do ano de 2001 foi na região Centro onde se registou um maior aumento no número de estabelecimentos hoteleiros. É de salientar que apesar da sua tendência sempre crescente é na região do Alentejo onde ao longo destes vinte anos sempre existiu um menor número de estabelecimentos hoteleiros.

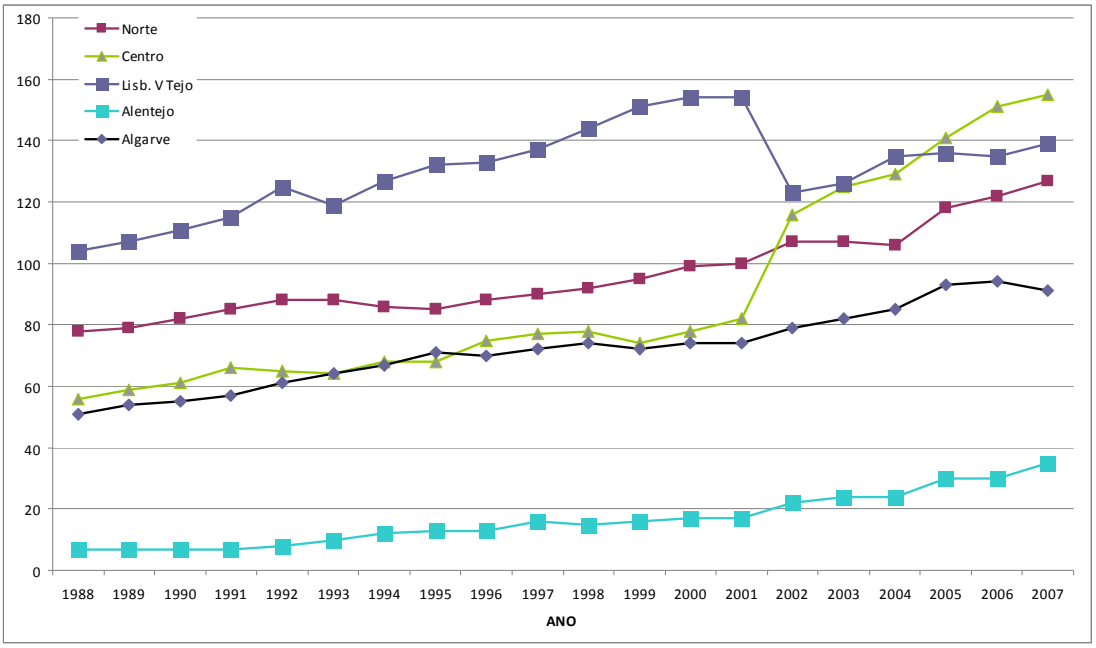
Gráfico2. Evolução do número total de estabelecimentos de alojamento por NUT's II (1988 - 2007)



Tal como acontece para Portugal Continental, ao nível das NUT`s II também são os hotéis o tipo de alojamento que existe em maior número. Comparando o ano de 1998 com o ano de 2007, em todas as NUT`s II se registou uma variação positiva do número de hotéis. Contudo, é

de salientar o facto que em todas as NUT`sII, ao longo do período em análise a tendência for sempre crescente, excepto na região de Lisboa e Vale do Tejo onde de 2001 para 2002 se registou uma diminuição do número destes estabelecimentos. Contudo, para essa região a partir de 2003 a tendência foi novamente crescente. Mas uma vez, sem qualquer surpresa é na região do Alentejo onde existe um menor número de hotéis ao longo deste vinte anos.

Gráfico 3. Evolução do número total de Hotéis por NUT's II (1988 - 2007)



Se analisarmos o número de estabelecimentos hoteleiros que existem em cada NUT`s II, mas agora tendo em conta a área de cada região, contacta-se que ao longo do território continental existe uma grande discrepância em termos da concentração destes estabelecimentos.

Quadro 3. Número de estabelecimentos hoteleiros e de hotéis por100 Km^2 por NUT`s II, em 2007

	Área Km^2	Nº Estabelecimentos Hoteleiros por 100 Km^2	Nº Hotéis em por 100 Km^2
Continente	88 633	1,98	0,612
Norte	21 278	2,16	0,59
Centro	28 405	1,50	0,54
Lisboa e V. Tejo	2 802	10,9	4,96
Alentejo	31 152	0,49	0,11
Algarve	4 996	8,31	0,18

Fonte: Elaboração própria.

Conforme se pode ver através do **Quadro 3**, é na Região de Lisboa e Vale do Tejo e na região do Algarve onde existe uma maior concentração, quer do número de estabelecimentos hoteleiros, quer do número de hotéis por km^2 . Em 2007, em cada 100 km^2, na região do Algarve existiam oito estabelecimentos hoteleiros, na região de Lisboa e Vale do Tejo onze, enquanto que na região do Alentejo aproximadamente um estabelecimento hoteleiro. Quando analisamos o número de hotéis, a realidade mantém-se, sendo a região de Lisboa, onde existe uma maior concentração.

Perante tais assimetrias regionais, é imprescindível a necessidade de desenvolver e dotar todas as regiões de Portugal em termos turísticos. Cada região possui características específicas, mas todas elas com grandes potencialidades que podem e devem contribuir para o

desenvolvimento económico e social do nosso país, como um todo. Tendo como objectivo último promover e valorizar a actividade turística, em 2008, com a publicação do Decreto-Lei n.º67/2008, procedeu-se a uma reorganização das Entidades Públicas Regionais com responsabilidades na área do turismo (MEI, 2008).

5 – PERSPECTIVAS REGIONAIS DO TURISMO PORTUGUÊS

No âmbito do Programa de Reestruturação da Administração Central do Estado (PRACE), e tendo presente a necessidade de dotar os organismos públicos na área do turismo de competências indispensáveis à afirmação de Portugal como um dos principais destinos turísticos europeus, o Governo aprovou, através da Resolução do Conselho de Ministros n.º 39/2006, de 21 de Abril, as orientações para a reestruturação dos vários ministérios, nas quais assume a necessidade de criar um organismo central do turismo único, este responsável pela prossecução da política de turismo nacional (o Turismo de Portugal, I.P.), e a descentralização das 19 regiões de turismo.

O Decreto-Lei n.º141/2007, de 27 de Abril definiu a missão e atribuições do Instituto do Turismo de Portugal, abreviadamente designado por Turismo de Portugal, I.P., autoridade turística nacional, concretizando-se assim o objectivo de criar uma única estrutura pública que tem por missão promover a valorização e a sustentabilidade da actividade turística nacional. O Turismo de Portugal, I.P., integrado no Ministério da Economia e da Inovação, cabe definir a estratégia nacional, contribuir para o desenvolvimento do sector e promover e valorizar a actividade turística, cuja missão consiste em:

- qualificar e desenvolver as infra-estruturas turísticas; - desenvolver a formação de recursos humanos;
- apoiar o investimento no sector;
- coordenar a promoção interna e externa de Portugal como destino turístico e,
- regular e fiscalizar os jogos de fortuna e azar.

Em 2008, com a publicação do Decreto-Lei n.º67/2008, de 10 de Abril, procedeu-se a uma reorganização das Entidades Públicas Regionais com responsabilidades na área do turismo, passando-se de dezanove[1] regiões do turismo para apenas 5 áreas regionais que reflectem as unidades territoriais utilizadas para fins estatísticos NUTS II – Norte, Centro, Lisboa e Vale do Tejo, Alentejo e Algarve. Para além das cinco áreas regionais, de acordo com o artigo 4º do Decreto-Lei n.º 67/2008 (MEI, 2008: 2172), "nas áreas regionais de turismo definidas no artigo 2.º são criados os pólos de desenvolvimento turístico" (Douro, Serra da Estrela, Leiria-Fátima, Oeste, Alentejo Litoral e Alqueva). Assim, para além das duas direcções regionais do turismo (regiões autónomas dos Açores e da Madeira), actualmente, temos onze entidades regionais de turismo que asseguram o desenvolvimento do turismo no território continental (**Quadro 4**).

Quadro 4. Entidades Públicas Regionais com responsabilidade na área do turismo

Regiões Turísticas	Pólos de desenvolvimento turístico
Turismo do Porto e Norte de Portugal	Turismo do Douro
Turismo do Centro de Portugal	Turismo da Serra da Estrela
Turismo de Lisboa e Vale do Tejo	Turismo de Leiria – Fátima
Turismo do Alentejo, E.R.T	Turismo do Oeste
Turismo do Algarve	Turismo Terras do Grande Lago Alqueva – Alentejo
	Turismo do Alentejo Litoral

Fonte: MEI(2008).

[1] As antigas dezanove regiões do turismo: Algarve, Alto Douro (Costa Verde), Alto Tâmega e Barroso, Centro, Dão Lafões; Douro Sul, Évora, Leiria-Fátima, Nordeste Transmontano, Oeste, Planície Dourada, Ribatejo, Rota da Luz, São Mamede (Norte Alentejo), Serra da Estrela, Serra do Marrão, Setúbal (Costa Azul), Templários (Floresta Central e Albufeiras) e Verde Minho.

Com o objectivo promover externamente a actividade turística, existem sete Agências Regionais de Promoção Turística (ARPT). Estas agências são associações de direito privado, sem fins lucrativos, constituídas por representantes dos agentes económicos do turismo e por empresas privadas com actividade turística. Existe uma agência em cada uma das 7 áreas promocionais (Norte, Centro, Lisboa Região, Alentejo, Algarve, Açores e Madeira) e são responsáveis pela elaboração, apresentação e execução dos respectivos planos regionais de promoção turística externa.

Pretendendo-se um País mais equilibrado e mais coeso territorialmente, para Silva (2000: 54), é ponto assente que "o desenvolvimento racional e equilibrado do turismo é fundamental para ao nosso Pais, já que, tratando-se de um sector gerador de infra-estruturas, de equipamentos, de actividades e emprego, pode proporcionar uma verdadeira cadeia de riqueza com repercussões directas nas economias regionais e na economia nacional".

6 - TURISMO SUSTENTÁVEL

È certo que o turismo se tornou num fenómeno de alcance mundial, ganhando cada vez mais relevo em termos económicos, sociais e ambientais. Estamos todos convictos de que apesar de o seu desenvolvimento trazer grandes benefícios para os países, o desenvolvimento desta actividade promotora pode conduzir a um vasto conjunto de problemas de ordem ambiental, económico e social.

Segundo a Comissão do Desenvolvimento Sustentável (CES, 2001: 3) "em consequência de uma rápida expansão do sector turístico, tanto nos tradicionais como nos novos destinos turísticos, enfrenta-se uma pressão cada vez maior sobre as dimensões naturais, culturais e socioeconómicas. Admite-se que um crescimento incontrolado de um turismo destinado a obter benefícios a curto prazo tende a ter consequências negativas, na medida em que danifica o meio ambiente, as sociedades e destrói as bases sobre as quais o turismo assenta e se desenvolve".

São vários os exemplos onde um desenvolvimento descontrolado do turismo acarretou impactos ambientais negativos, perda da identidade local, bem como outras ameaças ao meio envolvente. Perante tais factos, tem existido uma maior consciência por parte dos governos, instituições públicas, organismos não governamentais, bem como dos vários intervenientes da actividade turística, que têm levado a que o desenvolvimento sustentável esteja no centro das atenções. Contudo, devido á elevada complexidade de funcionamento da actividade turística, os múltiplos agentes que nela interferem, a escassez de tecnologias, falta de recursos e de conhecimentos por parte de alguns agentes económicos, ainda nos deparamos com grandes dificuldades de acção para a prática de um desenvolvimento sustentável no turismo.

Pretende-se que no futuro, Portugal seja um País competitivo e moderno "mas também um País com mais qualidade ambiental, mais equilibrado e coeso territorialmente, onde as populações encontrem qualidade de vida e as actividades económicas se desenvolvam no respeito pelos valores ambientais e de sustentabilidade" (MOPTC, 2005: 55). O conceito de desenvolvimento sustentável, isto é, o desenvolvimento que responde às necessidades das gerações presentes sem comprometer a possibilidade de satisfazer essas mesmas necessidades às gerações futuras, estende-se ao turismo. Segundo a Organização Mundial do Turismo, o desenvolvimento do turismo sustentável vai de encontro às necessidades das regiões turísticas, no sentido em que protege e aumenta as oportunidades de desenvolvimento dessas mesmas regiões.

O futuro do sector do turismo deve passar por uma perspectiva de sustentabilidade ambiental, económica e social, privilegiando-se a qualidade e a prática de acções socialmente responsáveis. Além disso, e de forma a maximizar os benefícios do turismo e minimizar as mudanças adversas e os impactos negativos "devem ser definidas medidas sociais e culturais como parte dos recursos turísticos, incorporadas no planeamento, no desenvolvimento e nos processos de gestão do turismo, dado que a gestão efectiva dos impactos sociais e dos impactos culturais é vital para obter uma indústria turística sustentável" (Baptista, 1997: 507).

7 – CONSIDERAÇÕES FINAIS

O turismo é, por um lado, uma actividade económica complexa, na medida em que agrega em si um conjunto de actividades, produtos e recursos de variada natureza. Por outro lado, é uma actividade económica que envolve uma multiplicidade de agentes que reproduzem diversos efeitos, que se fazem sentir transversalmente por toda a economia. Assim, dada a dimensão das receitas geradas, a sua importância na criação de postos de trabalho, bem como os efeitos multiplicadores que induz em vários sectores da economia, faz com que actualmente o turismo seja considerado como um sector estratégico na promoção do desenvolvimento de Portugal.

Não é só pelo facto de o turismo ter uma importância verdadeiramente estratégica para a economia portuguesa, mas em termos governamentais, temos assistido a um incremento de esforços no sentido de dinamizar e diversificar essa actividade económica por todo o território nacional. Contudo, o turismo, tal como acontece com outros sectores da economia, consome recursos e gera desperdícios de natureza económica, social e ambiental, devendo assim contabilizar-se à sua conta, quer através dos seus efeitos de natureza positiva, quer dos seus efeitos de natureza negativa. A este nível pouco ou nenhuma investigação foi desenvolvida, o que dificulta a focalização das políticas e estratégias governativas.

Assim, as preocupações, em particular, com os valores, a ética, o meio ambiente, o emprego, a utilização eficiente dos recursos disponíveis e, em geral, com o desenvolvimento sustentável desta actividade económica, contribuem para que, cada vez mais, a responsabilidade social empresarial se esteja a tornar um fenómeno de gestão empresarial. Efectivamente, as empresas de turismo ao presenciarem os efeitos da era da globalização económica sentem-se "obrigadas" a adoptar diferentes políticas, incorporando nas suas estratégias novos valores, novos padrões de comportamento e novas visões de negócio que conduzem a práticas de responsabilidade social que promovem, não só o seu volume de negócios, mas contribuem de forma equilibrada para um melhor ambiente.

Este ambiente com a focalização no conceito de desenvolvimento sustentável, isto é, o desenvolvimento que responde às necessidades das gerações presentes sem comprometer a possibilidade de satisfazer essas mesmas necessidades às gerações futuras, não pode deixar de entender-se e difundir-se também ao nível do turismo, com todas as vantagens que lhe estão inerentes. A preocupação crescente com a sustentabilidade da actividade turística, torna esta actividade económica num "filão de ouro" que importa descobrir.

BIBLIOGRAFIA

ALEXANDRE, J.A. (2003). *O turismo em Portugal: evolução e distribuição*. Tese de Mestrado, Universidade de Aveiro.

BAPTISTA, M. (1997). *Turismo Competitividade Sustentável*. Porto: Editora Verbo.

BURNAY, M.J. (2000). *Turismo e Natureza*. Livro de Actas do Ciclo de Debates do IFT. Instituto de Turismo de Portugal, 25-34.

COMISSÃO DAS COMUNIDADES EUROPEIAS (CCE, 2006). *Uma política de turismo europeia renovada: Rumo a uma parceria reforçada para o turismo na Europa, COM (2006) 134 final*. Bruxelas: Publicações Oficiais das Comunidades Europeias.

CONSELHO ECONÓMICO E SOCIAL (CES, 2001). *Sustainable Development of Tourism*. E/CN.17/2001/PC/21.

CUNHA, L. (2006). *Economia e Política do Turismo*. Lisboa: Editorial Verbo.

INSTITUTO NACIONAL DE ESTATÍSTICA (1988 a 2007). Anuários estatísticos das NUT`S II. Lisboa: INE:

MINISTÉRIO DA ECONOMIA E DA INOVAÇÃO (MEI, 2008). Decreto-Lei n.º67/2008, aprova o regime jurídico das áreas regionais de turismo de Portugal Continental e dos Pólos de Desenvolvimento Turístico, a delimitação e características. *Diário da República*, 1.ª Série, 71, 10 de Abril.

MINISTÉRIO DA ECONOMIA E DA INOVAÇÃO (MEI, 2007). *Plano Estratégico Nacional do Turismo*. Lisboa: Turismo de Portugal.

MINISTÉRIO DAS OBRAS PÚBLICAS, TRANSPORTES E COMUNICAÇÕES (MOPTC, 2005). *Grandes Opções do Plano 2005-2009*. Lisboa: MOPTC.

CARACTERÍSTICAS DESEABLES EN LOS ESTUDIOS DE ADE-DERECHO: UNA PERSPECTIVA DESDE LA VISIÓN DEL ALUMNADO.

JESÚS ESTEBAN
E-mail: estebanjesus2004@yahoo.es
CONCEPCIÓN ROJO
FÉLIX RUIZ
Facultat d'Economia
Universitat de València (España)

Resumen

Constituye un auténtico reto la implantación de las nuevas titulaciones así como la adaptación de los curricula a los requerimientos derivados del proceso Bolonia. Tanto en la evaluación institucional como la que corresponde a las titulaciones subyace la metodología EFQM. Este trabajo se centra en la visión que el cliente tiene de los aspectos deseables que deberían tener los profesores, las materias que imparten, el desarrollo de las clases, los recursos y los procesos de evaluación docente. De la misma manera se presentan las características consideradas inapropiadas por los estudiantes que han participado en la encuesta.

Palabras clave: visión cliente, evaluación calidad, mejora continua, innovación.
Área Temática: Estrategia Económica.

Abstract

New degrees and adapting curricula to the new requirements coming from Bolonia process constitute a real challenge. Subjacent methodology for institution and degree evaluation is inspired in EFQM methodology. This paper focuses on the customer side about the main aspects which should conform on professors, subjects, class development, resources and docent evaluation process. The paper also deals with the characteristics which are considered inappropriate on the same items by the students in the survey.

Thematic Area: Economics of Strategy.

1. INTRODUCCIÓN

Las universidades españolas se encuentran inmersas en el proceso de Convergencia al Espacio Europeo de Educación Superior. La implantación de nuevos planes de estudio ha ido acompañada de una actividad que pretende concluir en un cambio radical en las metodologías docentes. Este cambio se centra, en síntesis, en conseguir un papel más activo de los alumnos y en el uso de un mayor número de "herramientas" en el proceso de transmisión del conocimiento con el pretendido objetivo de conseguir un nivel de homologación de las titulaciones en el marco europeo y a su vez, unas mejores cotas de eficiencia del sistema educativo universitario.

Durante todo este periodo, cada universidad ha impulsado iniciativas tendentes a experimentar en sus titulaciones todos estos cambios en la estructura de la docencia, que traía aparejado el denominado "Proceso de Bolonia". Innumerables Proyectos de Innovación Educativa (PIE) han sido desarrollados y sus resultados han sido y siguen siendo debatidas en

numerosos congresos, jornadas o reuniones, con el fin de intercambiar experiencias y enriquecer al profesorado en esta ardua tarea que nos espera.

La Facultat d'Economia de la Universitat de València fue una de las pioneras, recogió el testigo rápidamente y un conjunto de entusiastas profesores se involucró con distintos proyectos, tanto individuales como colectivos, iniciando un camino por el que todos hemos de transitar. En un primer paso, se aplicaron nuevas metodologías docentes en algunas asignaturas dentro de lo que podríamos considerar como iniciativas personales. Posteriormente, coincidiendo con la implantación de la doble titulación de Licenciado en Administración y Dirección de Empresas y Licenciado en Derecho (ADE-Derecho), desde los equipos decanales de ambos centros, se tomó la decisión de que esta titulación se desarrollase dentro de un PIE auspiciado por la Universitat de València. Era la forma de añadir a los aspectos de las nuevas metodologías docentes, ya desarrollados en los proyectos individuales, todo aquello relacionado con la coordinación horizontal y vertical de los contenidos y la necesaria estructura de los estudios. El proyecto era ambicioso pues hay que añadir a todo lo anteriormente citado, el hecho de que se encuentren implicados los dos centros más grandes de nuestra universidad, con metodologías docentes tradicionalmente distintas.

Esta experiencia, que este curso cumple su sexto año, ha sido el detonante de otras experiencias interesantes como la implantación de un PIE en la Licenciatura de Economía, acogido a un proyecto del Ministerio de Educación, con el fin de analizar los resultados de las nuevas metodologías aplicadas a un grupo más numeroso.

La oferta de la doble titulación de ADE-Derecho supuso un éxito y la demanda, año tras año, supera con mucho la oferta de plazas, que es de 120. La nota de acceso suele superar el 7'5 y el bloque de estudiantes que cursa esta doble titulación presenta, por tanto, unas características peculiares dentro del alumnado matriculado en las distintas titulaciones que actualmente oferta la Facultat d'Economia. Hasta el curso 2007-08 el número de estudiantes que accedían por cada titulación era el mismo (60) y la nota de corte difería mínimamente. Sin embargo, en el curso 2007-08 la diferencia entre la nota de corte por cada centro difería sustancialmente lo que produjo agravios comparativos que desde los centros y el rectorado se intentó solucionar con el incremento, ese mismo curso, de 20 plazas que se distribuyeron por nota de acceso en un listado único de ambos centros. Ya en el 2008-09 se modificó el reparto de alumnos, accediendo 40 desde cada centro y 40 por nota según listado conjunto.

La Facultat d'Economia es, en estos momentos, el centro más grande de todas las universidades de la Comunidad Valenciana, con unos 9200 alumnos aproximadamente. Cada año acceden más de 1700 alumnos, se reciben más de 400 alumnos Erasmus y más de 350 de nuestros estudiantes se desplazan a alguna de las universidades extranjeras con las que se mantiene convenio.

La descripción, a grandes rasgos, de este escenario se considera conveniente para entender que los alumnos de la doble titulación ADE-Derecho constituyen un grupo homogéneo y claramente diferenciado con unas características particulares.

Un grupo de profesores implicados, desde el principio, en este proyecto publicaron en 2007 el libro "Innovación educativa en la Universidad: ADE-Derecho" donde se recogen, con gran profusión, los resultados académicos de los alumnos y se establece un estudio comparativo con los resultados de un grupo "generalista" de ambas titulaciones, es decir un grupo no sometido a ningún PIE.

La peculiar situación, ya comentada, del acceso de los alumnos durante el curso 2007-08 significó un escenario un poco distinto sobre el existente hasta ese momento. Las notas de acceso por la vía de Derecho habían descendido y el número de alumnos que accedieron por ADE se incrementó, generando algunas cuestiones de debate sobre la posible existencia de diferencias entre los rendimientos de unos y otros. Esto y la necesidad de abordar algunas cuestiones complementarias al trabajo anteriormente citado es lo que nos ha animado a iniciar un proyecto que trate de abordar diferentes cuestiones relacionadas con estos planes y/o con los agentes implicados en ellos.

En este congreso se presentan dos trabajos, muy iniciales, del proyecto comentado. El trabajo "ADE-Derecho en la Universitat de València: estudio de una experiencia piloto" (Antonia Ivars, Maria Isabel López y Félix Ruiz), ha dirigido la atención a dar respuesta a diversas cuestiones que se han ido planteando en el transcurso de estos cursos y parecían hacerse más patente en los estudiantes del curso 2007-08, y que deberían ayudar a conocer más, tanto las peculiaridades de estos estudiantes como las posibles relaciones entre diversas características y su rendimiento académico. Sin embargo, en el presente trabajo el objetivo ha ido dirigido a conocer, de los mismos alumnos del trabajo anterior, los aspectos que desean y no desean de algunos ítems relevantes del proceso de enseñanza.

El interés del estudio del presente trabajo se justifica por la creencia de que el conocimiento del cliente resulta necesario para el establecimiento de una estrategia adecuada de la gestión de todos los recursos que el profesor pueda tener, así como la de desvirtuar, si fuera el caso, el valor de determinados instrumentos, o herramientas. Otra consecuencia importante reside en la identificación de aquellos factores críticos que, de no evitarse, llevarían al traste cualquier acción valiosa.

Somos conscientes que estamos dando los primeros pasos y que, por ello, en fases posteriores se deberá recabar, como se está haciendo, nueva información que permita concretar y precisar los resultados obtenidos.

2. LA INFORMACIÓN OBTENIDA.

Dado el carácter inicial del estudio se pasó una encuesta de carácter abierto en la que cada alumno debería dar un máximo de tres contestaciones en cada ítem. Los ítems sobre los que se pedía información fueron:

1. El profesor
2. La asignatura o materia
3. El desarrollo de las clases
4. Los recursos
5. La evaluación de los alumnos
6. El sistema de evaluación de los profesores

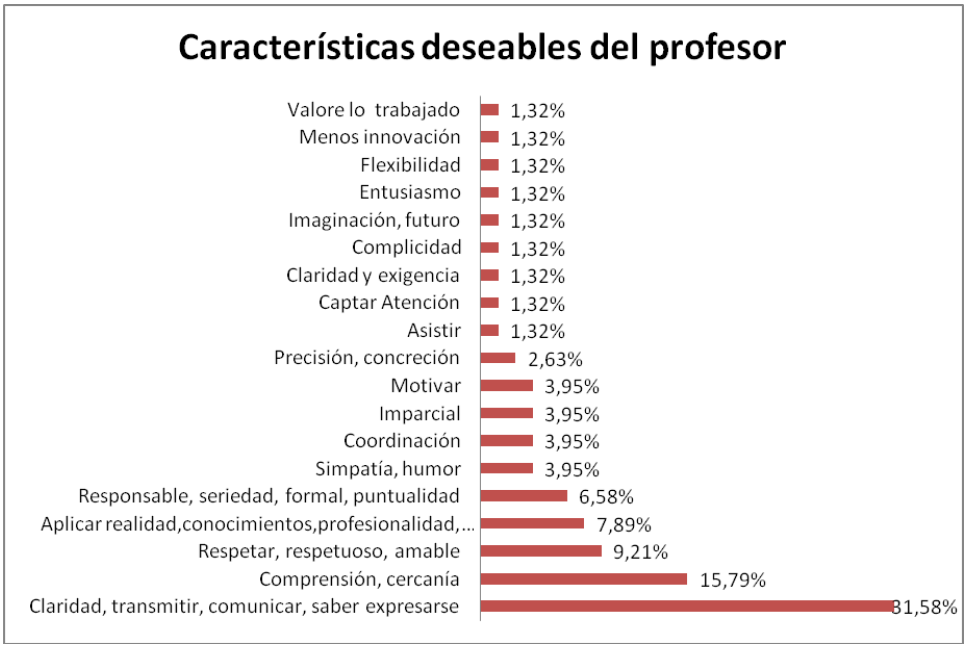
De los ítems 1,2,3 y 4 se les pedía que identificaran hasta un máximo de tres aspectos deseables de cada ítem así como tres aspectos no deseables. Del ítem 5 se les pidió tres aspectos deseables, en tanto que del ítem 6 se requería solamente los no deseables.

Resulta obvio que lo que se pretendía con esta primera encuesta no era sino conocer el parecer del cliente objetivo y conocerlo con sus propias palabras dejándolo exento, en la medida de lo posible, de cualquier influencia provocada por alguna definición o concepto previo, dejando para fases posteriores la depuración de los conceptos que se identifiquen en esta fase.

3. ANÁLISIS DE LOS DATOS.

3.1. Características o aspectos deseables del profesor.

La característica que sobresale es sin duda la de que sea claro, transmita, comunique y sepa expresarse. Dicho esto, el alumno valora el que el profesor sea cercano y comprensivo, respetuoso y amable. Con ello se ha alcanzado más de un 60%. El cuarto lugar es ocupado por la profesionalidad, el nivel de sus conocimientos y su vinculación con la realidad a través de la aplicabilidad que es capaz de transmitir. El quinto lugar lo ocupa la responsabilidad, seriedad y puntualidad características con las que se alcanza el 70%. El 30% restante se distribuye en un amplio elenco de características positivas.



3.2. Características deseables de la materia o asignatura.

El estudiante desea que los contenidos de la asignatura sean actuales y aplicables, prácticos reales y útiles. Este conjunto de conceptos supone el 54%. Se llega al 75% con otro conjunto de conceptos que se refieren a que la materia sea ajustada, básica, comprensible y clara, fácil y con ejemplos.

Hay lugar para los contenidos magistrales, innovadores, metódicos y sistemáticos. Se señalan, asimismo, lo interesante de que tengan fondo y futuro, y también se alude a que guarde relación con los estudios que sean amenos y entretenidos.

La petición a que versen de economía y de derecho parece lógico, dada la doble titulación, pero debe remarcarse que quien ha requerido una no lo ha hecho con la otra, por lo que cabría suponer que algunos estudiante perciben menos contenidos de derecho, o economía, de los que ellos desean, o quizás pudiera tener más sentido en su enunciado negativo, esto es, perciben más contenidos de economía o derecho de lo que desean.



3.3. Características deseables en el desarrollo de las clases.

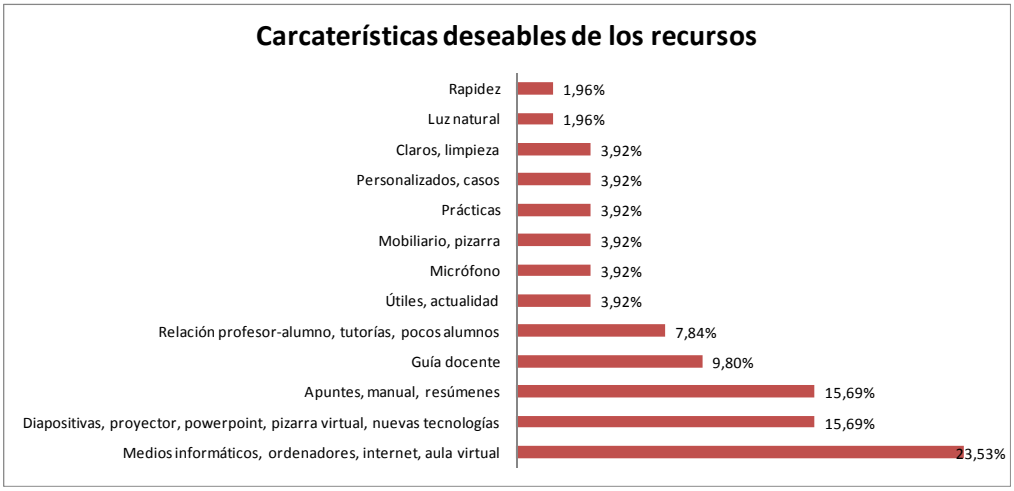
Como puede observarse, en este ítem existen atributos deseables que han aparecido en el epígrafe anterior; si bien cabe señalar que aquí se refieren al desarrollo de las clases y en aquel a los contenidos de la materia.

En este ítem la dispersión de conceptos es mayor. Las características que más veces se repiten se refieren a un desarrollo ameno, interesante y entretenido de la clase, seguido por los atributos de claridad, precisión, actuales y realistas. Estas últimas serían próximas al tercer conjunto compuesto por ser prácticas, útiles y que motiven.



Se desea brevedad, equilibrio en los tiempos, participación e interrelación profesor-alumno, que guarde relación con el examen, … Conviene llamar la atención en la matización que se hace relativa al carácter de "experto" del profesor y que éste introduzca sus experiencias personales.

3.4. Características deseables de los recursos.

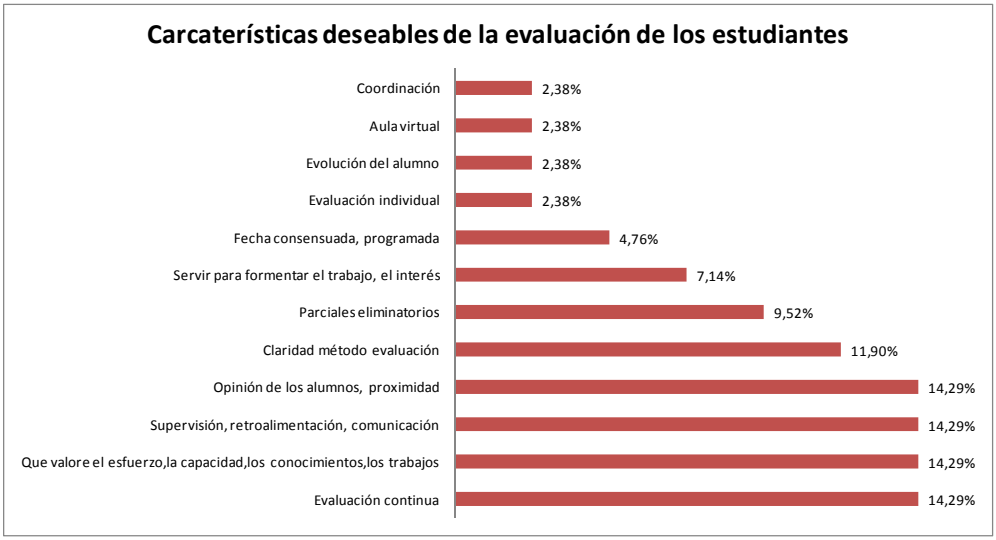


En este ítem las contestaciones se han encaminado a enunciar los recursos deseables. Así el 40% se relaciona con medios informáticos, ordenadores, los recursos de internet,… incluyendo, asimismo, apuntes, manuales y resúmenes, guía docente y prácticas.

Como atributos deseables de los recursos se enuncian la utilidad, actualidad, personalizados y "casos".

3.5. Características deseables de la evaluación de los estudiantes.

El tipo de evaluación señalada es la evaluación continua existiendo una llamada a la evaluación individual, cuestión que se entenderá mejor en conjunción con los atributos que se predican.



En efecto, se desean la valoración del esfuerzo, de los trabajos y de los conocimientos adquiridos, claridad en el método de evaluación, … todo lo cual nos lleva a pensar que la llamada a la evaluación individual se relaciona con evitar las sombras que produce algunas evaluaciones por trabajos colectivos en donde el esfuerzo de unos oculta la ignorancia o pasividad de otros.

El carácter eliminatorio de los exámenes, la valoración de la evolución del alumno y la utilización del aula virtual en la realización de pruebas o evaluaciones, entrega de trabajos, … son otras características que merecen observar.

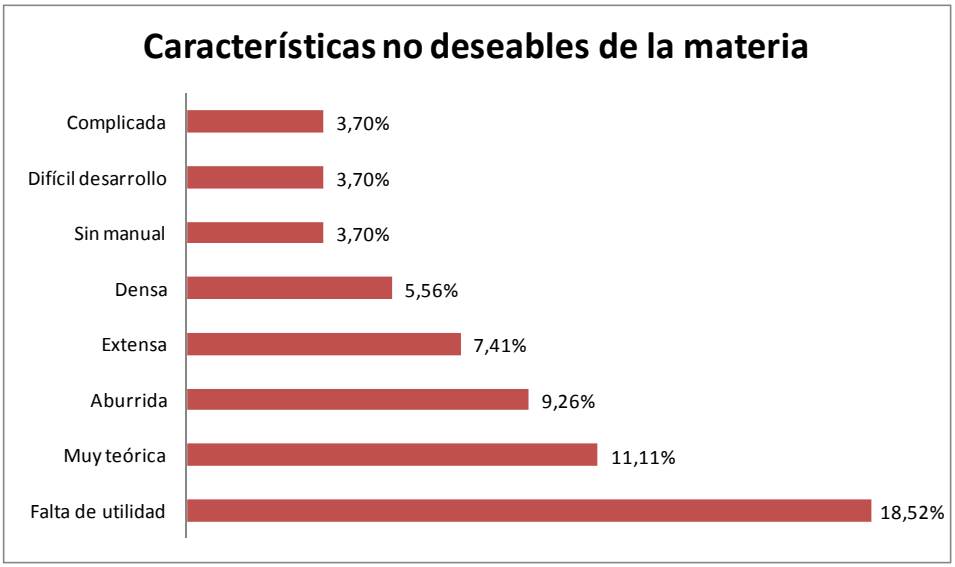
3.6. Características no deseables del profesor.

Son muchos los atributos no deseables que citan los estudiantes. Solamente se han representado el 60% de ellos y nótese que el que más se repite, la "prepotencia", no llega al 10%, si bien es cierto que si se agrega con el de "superioridad", suman un 15%.

En efecto, la prepotencia-superioridad, constituye el defecto más veces repudiado por los estudiantes, ocupando un segundo plano de importancia la falta de respeto y la mala educación. A continuación la lista se extiende, mal humor, pasividad, autoritario, no comprensivo, …

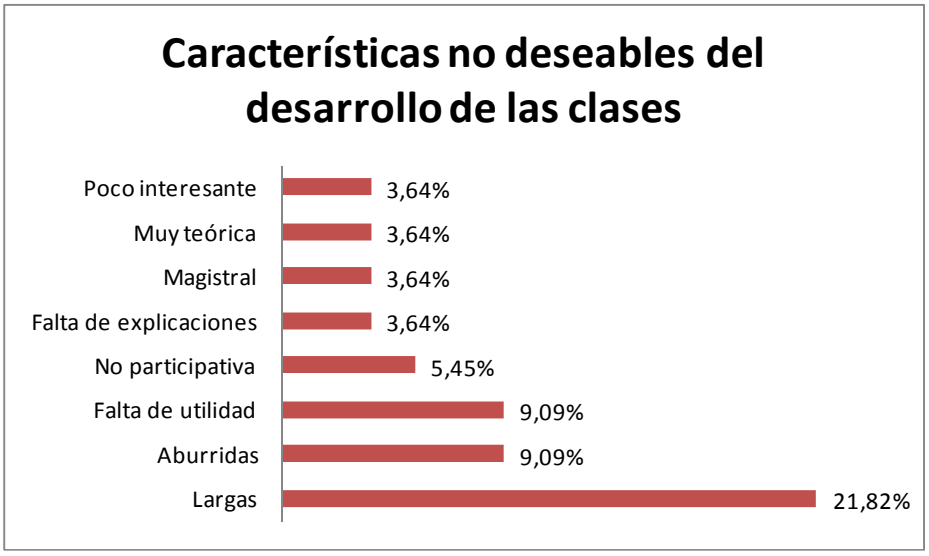


3.7. Características no deseables de la asignatura.



Aquí también se han se han señalado las características que constituyen el 62% de las señaladas. Sin embargo existe una mayor coincidencia que en el caso anterior. La falta de utilidad, el exceso teórico, lo aburrido, extenso, denso, complicado constituyen los atributos indeseables junto con la ausencia de manual.

3.8. Características no deseables del desarrollo de las clases.



Al igual que las precedentes se señalan las que constituyen el 60%. Es obvio que el atributo indeseable se focaliza sobre la extensión del desarrollo, lo aburrido y la ausencia de percepción de utilidad.

3.9. Características no deseables de los recursos.

En este caso se ha agrupado lo no deseable en cuatro categorías, según se relacionen con el:

- Aula y los instrumentos

- Material de trabajo y bibliografía
- Organización
- Utilización



Entre los primeros se citan: aula en malas condiciones, clase pequeña, ausencia de tarima, mala ubicación del ordenador, poca luminosidad, salubridad y falta de ventilación, lentitud de los medios informáticos.

En cuanto al material de trabajo y la bibliografía, se menciona la existencia de mucha bibliografía no directamente relacionada con el tema, abundancia de material complementario, exceso de manuales y ausencia de un auténtico manual y falta de apuntes así como manuales complicados. En definitiva se repite la apelación a la sencillez, agilidad y concreción señalada de forma reiterada en anteriores ítems.

El grupo que se ha denominado organización incluye aspectos referentes a disposición en horarios poco adecuados, exceso de exposiciones de trabajo, mal uso del aula virtual en el sentido de no dar tiempo suficiente para que el alumno acceda a los contenidos dispuestos en ella.

Por último, se da una queja explicita al abuso del Power Point.

3.10. Características no deseables de la evaluación de los profesores.

Como puede observarse existe una gran coincidencia en relación con la opinión de lo ineficaz de la evaluación del profesorado. No se trata de que no sea deseable la evaluación sino que la percepción que los mismos alumnos tienen es que no se realiza adecuadamente y/o no genera consecuencias acordes con ella.

Es de destacar como ellos mismos manifiestan la falta de seriedad y compromiso de los alumnos a la hora de la evaluación lo que parece justificarse, al menos en parte siempre y cuando la percepción generalizad es de ineficacia, inutilidad, incompletas o no adecuadas a la evaluación que se pretende.



4. CONCLUSIONES.

No parece que deban sorprender los resultados descritos. Los profesores con una mínima experiencia universitaria coincidirán, al menos en gran parte, con las conclusiones expuestas. Ciertamente que se trata de un primer paso con el único objetivo de identificar la percepción de los alumnos en torno a los atributos que deben de informar unos ítems que se han considerado relevantes en el proceso docente.

En el frontispicio del templo de Apolo los sabios griegos habían inscrito el "conócete a ti mismo". No cabe duda de que se trata de una cuestión previa y que "conocer a tus alumnos" constituye el complemento necesario para el diseño de una adecuada estrategia que asegure el éxito de la misión docente.

La seguridad que dan los años pueden, en alguna ocasión, llevar a olvidar que todo cambia y que para cada situación existe una combinación de acciones que pueden mantener la eficiencia en el logro del objetivo requerido. Adaptabilidad, flexibilidad, anticipación a los cambios y ausencia de dogmatismo son cualidades que deben estar presentes en el ser universitario.

La universidad se podrá configurar como servicio público o privado, se podrá reglamentar más o menos, se podrá uniformar, homologar,… pero sin duda la dualidad proveedor-cliente termina configurándose en una dualidad profesor-alumno que, por tanto, delimita el alcance de excesos en la uniformidad y coordinación. No solo la libertad de cátedra es un derecho, se trata de una institución prevista para garantizar la gestión que el docente experto deba diseñar como más adecuada a la dualidad profesor-alumno que tenga que resolver.

5. BIBLIOGRAFÍA.

DASÍ, A.; GARCÍA, J.; HUGUET, A.; JUAN, R.; MONTAGUD, M.D.; GOLLNERT, G. (2007) "Innovación educativa en la Universidad: ADE-Derecho", PUV Universitat de València, Valencia.

ESTEBAN, J; BACHERO, J.M.; BLASCO, O.; COLL, V.; DÍEZ, R.; IVARS, A.; LÓPEZ, M.I.; ROJO, C.; RUIZ, F. (2005) "Estadística Descriptiva y nociones de Probabilidad", Thomson, Madrid.

DESENVOLVIMENTO DOS MERCADOS FINANCEIROS: PAÍSES E CARACTERÍSTICAS

WILLIAM EID JUNIOR
E-mail: william.eid@fgv.br
FGV/EAESP – Escola de Administração de Empresas de São Paulo
JOSÉ ROBERTO SECURATO
E-mail: securato@usp.br
FEA/USP – Faculdade de Economia e Administração

Resumo

O artigo analisa o nível de desenvolvimento dos mercados financeiros numa amostra de 42 países e compara esse nível com características econômicas, de integração internacional e culturais dos diferentes países. Os resultados indicam que características como maior integração internacional, menor inflação, maior direcionamento para o empreendedorismo e maior liberdade de imprensa, dentre outras, são presentes nos países com mercados mais desenvolvidos. Ao mesmo tempo concluímos que o Brasil está ainda longe de possuir as mencionadas características.

Palavras chave: Brasil, mercados financeiros, desenvolvimento.
Área Temática: Estratégia Económica.

Abstract

The article analyses the level of financial market's development in a sample of 42 different countries. We compare the development of financial markets with economic, international integration and cultural characteristics from the countries. The results show that great international integration, lower inflation, great focus on entrepreneurship and great press freedom, among others, are characteristics that exist in countries with more developed financial market. At the same time we conclude that Brazil is far from having these characteristics.
Key words: Brazil, financial markets, development

Thematic area: economic strategy.

1 – INTRODUÇÃO

Não há dúvidas de que mercados financeiros desenvolvidos são o suporte para o crescimento econômico das nações. Desde o trabalho de Goldsmith em 1969 esse fato é bem documentado na literatura. No entanto não é apenas o grau de desenvolvimento dos mercados que é importante. Sua arquitetura também, e como arquitetura entende-se o direcionamento maior da economia para um sistema bancário ou um sistema de mercado. Exemplos desse direcionamento são de um lado os Estados Unidos e Inglaterra, com economias mais voltadas para os mercados acionários, e Alemanha e Japão, com economias mais voltadas para o sistema bancário. Diversos autores, como citados em Dermigue-Kunt e Levine (1999) indicam vantagens e desvantagens nos dois sistemas. De um lado, segundo os defensores dos sistemas baseados em bancos, eles têm um papel positivo no sentido de identificar projetos adequados, mobilizar os capitais necessários, monitorar gestores e riscos. Além disso, em países com estrutura legal incipiente, bancos poderosos podem forçar empresas a revelar informações e pagar seus débitos. Os defensores do sistema de mercado apontam as desvantagens nas economias dominadas por bancos, sendo que a primeira delas é o viés para projetos de baixo risco, retardando o desenvolvimento econômico. Já as vantagens do sistema de mercado,

segundo esses mesmos defensores, são a maior difusão de informações, maior facilidade para diversificação dos investimentos e padronização dos instrumentos de gestão de riscos.

Dada a importância dos mercados financeiros para o crescimento econômico dos países surge um questionamento natural: quais são as características que diferenciam os países com diferentes graus de desenvolvimento dos mercados financeiros? Melhor dizendo, quais são as características dos países em que os mercados financeiros são mais desenvolvidos, e quais são as características dos países em que eles têm um desenvolvimento menor.

Nesse sentido, nos últimos anos do século XX uma nova abordagem passou a ser utilizada pelos acadêmicos para explicar o desenvolvimento dos diferentes mercados. No lugar da análise neoclássica baseada em ganhos decrescentes do capital passamos a lidar com características institucionais nos modelos, sobretudo com o sistema cultural/legal dos diferentes países. Diversos autores, dentre eles La Porta, dedicaram-se a esse tipo de estudo, separando os países pela origem do sistema cultural/legal em *common law* e *civil* (ou *code*) law. Ou alternativamente em sistemas derivados da tradição britânica (*common law*) e da tradição continental européia, francesa, alemã ou escandinava. Uma pequena digressão pode auxiliar no entendimento desta questão. De um lado temos os países onde os códigos legais foram construídos e impostos num determinado momento ou em períodos relativamente curtos. De outros países onde a legislação nasceu dos costumes e tradições. No primeiro grupo vemos os países escandinavos que organizam seu código legal nos séculos XVII e XVIII, e esses códigos permaneceram inalterados ao longo dos séculos, e imunes às influências germânicas ou francesas. Na França é Napoleão que em 1804 organiza o Código Civil Francês, e obriga sua adoção nos países conquistados como Itália, Polônia, Países Baixos e império Austro-Húngaro. A influência desse código se estenderá, ao longo dos anos, para a Indochina, norte da África, região sub-sahariana da África, Oceania e Guiana Francesa. Além disso, ele serviu como base para o desenvolvimento do sistema legal da Espanha e Portugal, levando a sua influência à América Latina, incluindo o Brasil. Já o código alemão, completado no final do século XIX, estendeu sua influência para a Áustria, Suíça, China, Taiwan, República Tcheca, Eslováquia, Grécia, Hungria, Itália, países componentes da ex-república Iugoslávia, Japão e Coréia.

Em sentido contrário ao da elaboração de um código rígido de conduta, o sistema dito de *civil law*, ou inglês, tem a origem de suas leis na solução individual de casos pelos juízes. Esse sistema tem sua influência observada no Caribe, América do Norte, partes da Ásia e África.

Além da forma de desenvolvimento original, os dois sistemas se diferenciam pelo papel dos juízes. No sistema de *common law* eles têm liberdade de ação, decidindo caso a caso e conforme as circunstâncias. No sistema de *code law* essa liberdade não existe, estando as decisões associadas diretamente à legislação. Essa diferença leva o sistema de *common law* a ser muito mais adaptativo e rápido do acompanhamento da evolução dos costumes. Por conseqüência favorece o desenvolvimento dos mercados financeiros. Uma outra visão sobre esse processo mostra que o sistema de *code law* protege o estado, enquanto que o de *common law* defende a propriedade privada.

Os defensores da influência absoluta do sistema legal sobre o nível de desenvolvimento dos mercados, como La Porta e al (1998), alegam que os sistemas continentais europeus são muito rígidos, impedindo a formação de uma classe de empreendedores ao mesmo tempo em que não protegem adequadamente os investidores.

Após alguns anos de domínio absoluto da conexão entre origem do sistema legal e desenvolvimento dos mercados, mais recentemente surgiram vozes, como Modigliani e Perotti (1997) em prol da ampliação do rol dos determinantes do desenvolvimento dos mercados financeiros. Aspectos mais relacionados ao dia a dia dos agentes econômicos, como proteção aos investidores, papel do sistema legal e nível de corrupção, por exemplo, passaram a ser considerados na análise. E isso em função da descoberta, por esses autores, de que não há exatamente uma dicotomia entre economias baseadas em bancos ou em mercados financeiros. Na verdade, ao menos em parte dos países, os dois sistemas convivem e o melhor a fazer para analisá-los é identificar métricas que considerem a existência dos dois sistemas, e não simplesmente separar os países num universo binomial puro.

Neste trabalho vamos analisar a existência de características que expliquem, ao menos parcialmente, as diferenças no grau de desenvolvimento dos mercados financeiros ao redor do

mundo dando especial atenção ao caso do Brasil. Inicialmente serão descritos os principais trabalhos desenvolvidos sobre o tema, para posteriormente descrevermos as novas análises. A sessão II traz uma revisão bibliográfica sobre o tema, a III apresenta dos dados utilizados, a IV a metodologia e resultados e finalmente a V mostra as principais conclusões.

2 – REVISÃO BIBLIOGRÁFICA

La Porta, Lopez-de-Silanez, Schleifer e Vishny (1997 e 1998) concluem que os países que têm seu sistema legal baseado em *code law* têm menor nível de desenvolvimento econômico que os outros. Em particular esses países apresentam menor nível de valor de mercado das empresas abertas sobre o PIB, menor valor de crédito sobre o PIB e também menor número de ofertas públicas iniciais. Os autores também mostram que países baseados em *code law* oferecem menos direitos aos acionistas. Além disso, esses países apresentam menores exigências em termos de difusão de informações. Já os países com *common law* oferecem também maior proteção aos credores em geral. É interessante notar aqui que maior proteção aos credores leva a um maior desenvolvimento da intermediação financeira e maior proteção aos acionistas maior desenvolvimento do mercado acionário. Podemos inferir que países com *common law* terão mercados, tanto de crédito como acionários, mais desenvolvidos.

Corroborando essa hipótese, Dermigue-Kunt e Levine (1999) analisaram a estrutura financeira de 150 países abordando três aspectos: o desenvolvimento econômico vis a vis o desenvolvimento dos bancos e mercados de ações, desenvolvimento econômico e sistemas baseados em mercado versus sistemas baseados em bancos e finalmente os determinantes da estrutura financeira, observando aspectos legais, regulatórios, tributários e macroeconômicos. Sua principal conclusão diz que países que têm o sistema legal baseado em *common law*, forte proteção aos acionistas minoritários, bons padrões contábeis, baixo nível de corrupção e sem seguro para depósitos bancários tendem a ser mais orientados para mercado. Por outro lado, os países que tem o sistema legal baseado no sistema francês, ou *code law*, baixa proteção aos acionistas minoritários, baixa garantia para contratos, altos níveis de corrupção, padrões contábeis ruins e altos níveis de inflação tendem a ter sistemas financeiros menos desenvolvidos, mesmo quando controlados para o nível de renda dos países.

No entanto outra vertente derivou desta primeira, advogando que não apenas a origem do sistema cultural/legal é determinante do desenvolvimento dos mercados financeiros, mas outros aspectos institucionais também contam, e muito. Aspectos como proteção aos investidores e acionistas, corrupção, risco político e outros têm sido alvo de diversos estudos. Em particular La Porta et al (2000), complementando seus artigos anteriores, argumentam que a eficiência do sistema legal é o principal provedor de condições que determinarão inovação e crescimento econômico. Isto é, não é a origem do sistema legal que é determinante das características do sistema econômico e do desempenho da economia, mas sim a eficiência do sistema legal.

Na mesma linha Beck e Levine (2003) descrevem a chamada Teoria da Lei e Finanças, que foca no papel do sistema legal de cada país como explicação para as diferenças entre os mercados financeiros. A primeira parte da teoria afirma que em países onde as instituições legais dão apoio à propriedade privada, aos contratos e protege os direitos dos investidores, os poupadores são mais dispostos a financiar os empreendimentos e os mercados financeiros florescem. Em contraste, em países onde as instituições legais não dão esse apoio, há uma maior inibição do financiamento dos empreendimentos e pouco desenvolvimento financeiro. A segunda parte da teoria diz que as diferentes tradições legais, que surgiram na Europa e se espalharam pelo mundo via colonização, conquista ou imitação, explicam parte das diferenças de desenvolvimento financeiro observadas hoje em dia.

Tadesse (2002) relaciona a arquitetura financeira dos diferentes países, isto é, o grau em que o sistema financeiro é mais orientado para bancos ou para o mercado, com o desempenho econômico das economias. O autor advoga a tese de que não é apenas a origem do sistema legal que vai definir o desempenho da economia, mas um conjunto de fatores. Ele ressalta uma vantagem dos sistemas mais orientados para o mercado: os preços nele formados têm um conteúdo informacional importante. O autor usa indicadores de arquitetura relacionados ao

tamanho, atividade, e eficiência do mercado de ações em relação ao setor bancário, agregando essas medidas para efeito dos testes. Tamanho é medido pelo quociente entre o volume de capitalização das empresas listadas em bolsa pelo volume de créditos bancários do setor privado. Atividade é medida pelo quociente entre valor total negociado de ações e o volume de créditos bancários. Finalmente a eficiência é medida pelo produto do valor total negociado de ações, medido como percentual do PIB, pelos custos indiretos do setor bancário divididos pelos ativos totais do mesmo setor. Retiradas as médias desses três indicadores, a variável arquitetura será a média aritmética das delas. Outra medida utilizada no estudo é a referente ao nível de desenvolvimento dos mercados. O autor considera subdesenvolvidos os países abaixo da média em dois indicadores: desenvolvimento dos bancos e dos mercados. Estudando dados de 36 países relativos à arquitetura financeira, desenvolvimento financeiro e desempenho da indústria, o autor conclui que há uma correlação positiva entre desempenho econômico e grau de orientação para o mercado em países com sistemas financeiros desenvolvidos. Já em países com sistemas financeiros pouco desenvolvidos, a correlação é negativa.

Chinn e Ito (2005) analisam o quanto à liberalização financeira contribui para o desenvolvimento dos mercados acionários. Os autores indicam diversos benefícios da liberalização financeira, como a redução do custo de capital e redução da artificialidade na determinação da taxa de juros. Eles utilizam três medidas de desenvolvimento de mercado: capitalização bursátil, valor total das ações negociadas e giro do mercado acionário. Já para abertura ou liberalização dos mercados foi desenvolvido um índice baseado em quatro variáveis binárias: existência de múltiplas taxas de câmbio, existência de restrições para transações com moedas, existência de restrições para transações com capitais e por último, existência de restrições para utilização das receitas de exportação. Eles concluem que a liberalização financeira contribui para o desenvolvimento do mercado acionário, mas apenas quando há um desenvolvimento do sistema legal e institucional do país. Outro resultado interessante diz que a abertura do mercado de bens de consumo e de capital é uma precondição para a abertura financeira. Um último resultado mostra que o desenvolvimento de um setor bancário forte é também precondição para o desenvolvimento do mercado acionário.

Levine e Zervos (1996) analisam a relação empírica entre desenvolvimento do mercado acionário e o crescimento econômico dos diferentes países no longo prazo. Utilizando medidas como capitalização bursátil sobre PIB-Produto Interno Bruto, liquidez dada pela relação entre valor médio dos negócios sobre o PIB, e também giro do mercado, e finalmente integração internacional baseada num modelo multifator IAPM – International Arbitrage Pricing Model desenvolvido por Korajczyk (1996), os autores constroem uma medida agregada de desenvolvimento dos mercados calculando a média das medidas anteriormente citadas. Os autores também estimam a profundidade do mercado através do índice valor das dívidas do sistema financeiro sobre PIB. Com dados de 24 países para o período compreendido entre 1976 e 1996 os autores mostram uma relação positiva e forte entre desenvolvimento do mercado acionário e crescimento econômico. Mais que isso, eles demonstram uma conexão forte entre os determinantes do desenvolvimento dos mercados acionários e o crescimento econômico.

Outro estudo relacionado ao tema, não diretamente, mas indiretamente, é o de Fan e Titman (2003). Nele os autores estudam a estrutura de capital e as alternativas de maturidade para a dívida corporativa em 39 países. Eles encontram evidências de que o sistema legal, comomm law ou *code law*, e o sistema tributário afetam diretamente a escolha de estrutura de capital. Também indicam o impacto positivo da existência de fundos de investimentos destinados a investidores institucionais sobre o desenvolvimento dos sistemas bancários e mercados acionários.

Alfaro et al (2002) examinam a relação entre o investimento estrangeiro direto, mercados financeiros e crescimento econômico. Com uma amostra de mais de 50 países, eles concluem que mercados financeiros não desenvolvidos restringem os benefícios advindos do investimento direto. Os autores destacam que, em que pese que sua definição de mercados financeiros não desenvolvidos esteja associada às restrições de circulação de capitais de curto prazo, os benefícios do fluxo de investimentos estrangeiros direto, que é constituído de recursos de longo prazo, também são reduzidos.

Ainda dentro dos estudos relacionando arquitetura financeira e crescimento econômico, Beck et al (2000) utilizam três metodologias. Inicialmente avaliam se países com sistema financeiro baseado em bancos crescem mais que os com sistema baseado em mercado. Em seguida analisam se indústrias que dependem fortemente de financiamento externo crescem mais em sistemas baseados em mercado ou em sistemas baseados em bancos. Por último analisam se as empresas nos diferentes sistemas crescem abaixo da taxa predita pela geração de recursos internos e financiamento de curto prazo em sistemas baseados em bancos ou em sistemas baseados em mercados. As três metodologias apontam os mesmos resultados. O primeiro mostra que a estrutura financeira não é um mecanismo adequado para distinguir sistemas financeiros, o segundo indica que a estrutura financeira não explica o crescimento econômico, ou o desempenho das diferentes indústrias ou ainda a expansão das empresas. Por último os autores indicam que os resultados são inconsistentes com a tradicional visão de separação entre economias baseadas em mercado ou bancos. O elemento que distingue as economias é a existência de um sistema legal que efetivamente proteja os investidores externos.

Ndikumana (2001) analisa o impacto da arquitetura do sistema financeiro sobre o investimento doméstico. Com uma amostra de 99 países no período entre 1965 e 1997, ele classifica os países em quatro grupos: subdesenvolvidos com sistema financeiro baseado em mercado ou sistema financeiro baseado em bancos, desenvolvidos com sistema financeiro baseado em mercado ou sistema financeiro baseado em bancos. O autor busca responder duas questões. A primeira diz respeito à relação entre desenvolvimento dos mercados financeiros e o nível de investimento doméstico. A segunda pergunta se há relação entre a arquitetura dos mercados e o nível de investimento doméstico. A evidência empírica suporta a tese de que há uma relação positiva entre desenvolvimento dos mercados financeiros e nível de investimento doméstico. Isso sugere que um mercado financeiro mais desenvolvido oferecerá capital em maior abundância, mais barato e alocado de forma mais eficiente. Quanto à segunda questão, a evidência empírica sugere que não há impacto da arquitetura do sistema financeiro sobre o nível de investimento doméstico. Esse resultado é inconsistente com as afirmações referentes à superioridade de uma ou outra arquitetura dos sistemas financeiros. Segundo o autor, os resultados são consistentes com a idéia de mercados e bancos complementares, não excludentes.

Francis, Khudara e Pereira (2001) analisam as diferenças entre os sistemas contábeis de países com comomm law e com *code law*. Eles analisam dados de 1990 para 31 países, testando duas hipóteses. A primeira relaciona países com *code law* com baixo nível de proteção ao investidor e consequentemente sistemas contábeis de pouca transparência. A segunda relaciona países com *code law* com baixa demanda por auditoria. Os resultados empíricos suportam as duas hipóteses. Os países com *common law* têm sistemas contábeis mais transparentes e maior demanda por auditoria. Os autores também identificaram diferenças entre os países com *code law*. Dentre eles, os que contam com mais transparência contábil têm mais demanda por auditoria.

Fuchs-Schündeln e Funke (2001) analisam os efeitos da liberalização nos mercados acionários em 27 países sobre o desenvolvimento econômico e financeiro. Os dados analisados cobrem o período entre 1975 e 2000, com dados de países que implementaram programas de liberalização entre 1980 e 1995. Os resultados indicam que a liberalização está associada a um aumento nos investimentos privados no curto prazo da ordem de 14% nos quatro primeiros anos a partir da liberalização, e de 4% no renda per capita no mesmo período. O crescimento tende a ser maior se a liberalização foi precedida por reformas institucionais. Os autores também encontraram evidências de um efeito permanente no crescimento de 0,4%, usando agora uma amostra de 72 países.

Boutchkova e Megginson (2000) estudam o impacto das privatizações sobre os mercados acionários. Analisando uma amostra de 31 paises com dados de 1983 a 1999 os autores constatam um declínio na participação dos bancos nos sistemas financeiros de suporte às corporações ao redor do mundo. A fração do crédito doméstico detido pelos bancos nos países componentes da amostra manteve-se estável, por volta de 125% durante o período de análise. Em compensação a valor de capitalização das bolsas subiu de 52% para 82% do PIB dos países analisados. As ações emitidas nos programas de privatização contribuíram significativamente para o incremento de mais de 11 vezes no total de capitalização dos mercados e também no

imenso crescimento no volume negociado nas bolsas de valores, que passou de US$ 1,23 trilhões em 1983 para US$ 37,5 trilhões em 1999. A liquidez dos mercados também foi afetada pelas privatizações. Os autores indicam que elas são responsáveis por um incremento de 2,3% no primeiro ano após a privatização e de 1,7% no segundo ano. As empresas privatizadas são as mais valiosas em seis dos dez maiores mercados acionários do mundo, à exceção dos EUA. As práticas dos bancos de investimento foram transformadas, já que 23 das 25 maiores emissões de ações foram de empresas privatizadas. Os autores ainda citam a evidência empírica favorável às privatizações do ponto de vista dos investidores. Além de excesso de retorno para os investidores, as privatizações trouxeram um imenso contingente de novos investidores para os diferentes mercados. Quase 2/3 das empresas não americanas com mais de 500.000 acionistas foram fruto de privatização. Mais de uma dúzia delas tem mais de 1.000.000 de acionistas. No entanto os autores ressaltam que essa base de acionistas não é estável, tendo se reduzido de 1/3 nos primeiros anos após a privatização. Os autores concluem que as privatizações foram muito benéficas para os mercados acionários.

Modigliani e Perroti (1997), usam como *proxy* para qualidade do sistema legal e jurídico o prêmio nas ações com direito a voto sobre as que não têm esse direito. Em paises onde este prêmio é grande, o financiamento das empresas é dominado por bancos e o mercado acionário é pequeno. Os autores mostram que onde o direito dos investidores é pouco protegido, há dificuldades para as empresas emitirem ações, e como resultado, novos empreendimentos lucrativos não são desenvolvidos.

Herrero et al (2002) fazem uma análise descritiva de alguns determinantes do desenvolvimento dos mercados financeiros latino americanos. Depois de descrever métricas de desenvolvimento de mercados financeiros, os autores classificam os diferentes mercados em cluster a partir da análise de duas dimensões: a profundidade dos mercados e a renda per capita. Eles concluem que os mercados latino americanos estão menos desenvolvidos que os asiáticos e se comparam desfavoravelmente também com os países do leste europeu. Localmente a análise de clusters apontou quatro agrupamentos que foram estudados através de diferentes indicadores. Os autores ressaltam que não buscaram relações de causalidade no estudo, mas que há claras indicações de existência de um círculo virtuoso, com relação direta entre profundidade financeira e eficiência dos sistemas financeiros.

Claessens, Djankov e Nenova (2000) analisam os diferentes padrões de risco financeiro das empresas numa amostra cobrindo 46 países e 11.000 empresas para o período de 1995 a 1996, utilizando 12 indicadores financeiros. Eles também utilizam três indicadores de desempenho concluindo que há grandes diferenças entre os países no que tange ao risco e desempenho das empresas. E que essas diferenças estão relacionadas ao ambiente institucional de cada país, em grande parte oriundos do seu sistema legal. Mas além do sistema legal, *common law* ou *code law*, eles também indicam a relação entre esses aspectos e a arquitetura do sistema financeiro com o nível de risco das empresas.

Caprio e Honohan (2004) estudam o efeito dos diferentes instrumentos de disciplina dos mercados em mercados não desenvolvidos. As conclusões do estudo, que abrange 50 países em desenvolvimento, indicam que não há uma relação direta entre baixo nível de renda e falta de disciplina nos mercados. A análise dos autores indica que em função da presença de bancos estrangeiros em países de baixo nível de renda, a disciplina é exercida nos padrões da matriz desses bancos. Segundo eles, há mais disciplina financeira nesses países do que se pensava anteriormente. Eles sugerem aos países que, em busca de maior disciplina, reduzam a participação do estado nos bancos, eliminem garantias de depósitos, e não confiando totalmente na supervisão bancária como indutora de disciplina. Maior transparência é uma das indicações para melhoria da disciplina e desenvolvimento dos mercados.

Barth, Caprio e Levine (2000) analisam a regulação e propriedade de bancos em 60 países relacionando esses indicadores com o desenvolvimento e estabilidade dos mercados. Suas principais conclusões são: não há relação direta entre restrições impostas pela regulamentação e o nível de desenvolvimento dos mercados financeiros; não há relação entre propriedade dos bancos (ou restrição à) e desenvolvimento dos mercados financeiros. Eles também concluem

que países com maiores níveis de restrições à ação dos bancos no mercado de títulos têm maior probabilidade de sofrer uma crise no sistema bancário, isto é, maiores restrições regulatórias trazem maior instabilidade ao sistema. Também concluem que não permitir que bancos possuam empresas não financeiras é maléfico para a estabilidade do sistema financeiro. A evidência indica que a maior integração entre bancos e outras atividades é positiva para o desenvolvimento dos mercados. Por último a evidência empírica por eles encontrada indica que a propriedade estatal dos bancos é prejudicial ao desenvolvimento dos mercados, seja em relação ao setor bancário seja em relação ao mercado acionário. O trabalho mostra as implicações negativas da imposição de maiores restrições à atividade bancária.

A tabela abaixo mostra um resumo dos trabalhos elencados nessa revisão:

Tabela 1: Estudos e variáveis relevantes

Autor	Data	Estudo	Variáveis Relevante
La Porta et al	1998 1999	Desenvolvimento econômico	Origem do sistema legal Proteção ao credor
Dermigue-Kunt e Levine	1999	Mercados financeiros	Origem do sistema legal Outras características culturais/legais
La Porta et al	2000	Mercados financeiros	Eficiência do sistema legal
Beck e Levine	2003	Mercados financeiros	Proteção à propriedade privada Origem do sistema legal
Tadesse	2002	Desenvolvimento econômico	Arquitetura financeira
Chinn e Ito	2005	Mercado acionário	Liberalização financeira
Levine e Zervos	1996	Desenvolvimento econômico	Mercado acionário Integração internacional
Alfaro et al	2002	Desenvolvimento econômico	Investimento estrangeiro direto
Beck et al	2000	Desenvolvimento econômico	Estrutura financeira Proteção aos investidores
Ndikumana	2001	Investimento doméstico	Desenvolvimento dos mercados financeiros
Francis et al	2001	Sistemas contábeis	Origem do sistema legal
Fuchs-Schündeln e Funke	2001	Desenvolvimento econômico	Liberalização dos mercados
Boutchkova e Megginson	2000	Mercados acionários	Privatizações
Modigliani e Perroti	1997	Mercados financeiros	Proteção ao investidor
Herrero et al	2002	Mercados financeiros	Profundidade dos mercados Renda per capita
Claessens	2000	Risco das empresas	Origem do sistema legal
Caprio e Honohan	2004	Disciplina nos mercados	Nível de renda
Barth et al	2000	Desenvolvimento e estabilidade dos mercados	Regulação Propriedade dos bancos

Temos então um panorama apresentando os principais estudos relativos ao desenvolvimento dos mercados financeiros. Nosso próximo passo é, à luz do que já foi estudado e também com a inserção de novas variáveis, analisar os determinantes do desenvolvimento dos mercados.

3 – DADOS

Vamos analisar esse desenvolvimento de três ângulos complementares. O desenvolvimento do mercado acionário é medido pelo quociente entre o valor de capitalização das empresas listadas em cada país por seu Produto Interno Bruto (MKCAP) , como proposto por Tadesse (2002). Já o desenvolvimento do mercado bancário pelo quociente entre o volume de débito de

longo prazo e o Produto Interno (LTD), também como em Tadesse (2002). A fonte dos dados é o Relatório Anual do Banco Mundial (World Bank). Uma última variável dependente indica a orientação do mercado de cada país: se é mais orientado para mercado acionário ou mais orientado para mercado de crédito. Utilizaremos a variável também desenvolvida por Tadesse (2002) e chamada por ele de Arquitetura do Mercado (ARQ) e discutida na revisão bibliográfica. Obtivemos dados de 42 países. São eles:

Tabela 2: Países componentes da amostra

África do Sul	Dinamarca	Irlanda	Polônia
Alemanha	Egito	Israel	Portugal
Austrália	Espanha	Itália	Reino Unido
Áustria	Estados Unidos	Japão	Singapura
Bélgica	Filipinas	Jordânia	Suécia
Brasil	Finlândia	Malásia	Suíça
Canadá	França	México	Tailândia
Chile	Grécia	Noruega	Turquia
China	Holanda	Nova Zelândia	Venezuela
Colômbia	Índia	Paquistão	
Coréia	Indonésia	Peru	

Os gráficos a seguir mostram as três variáveis dependentes do estudo.

Gráfico 1: Valor das empresas listadas em bolsa

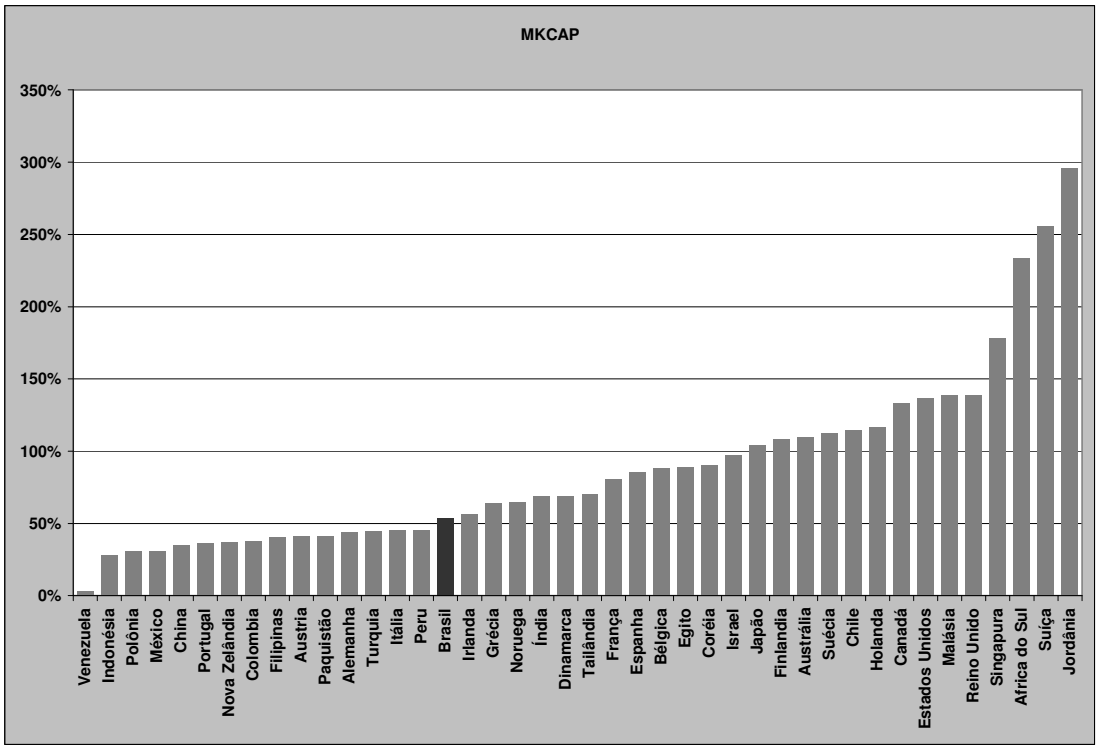


Gráfico 2: Valor total dos créditos de longo prazo

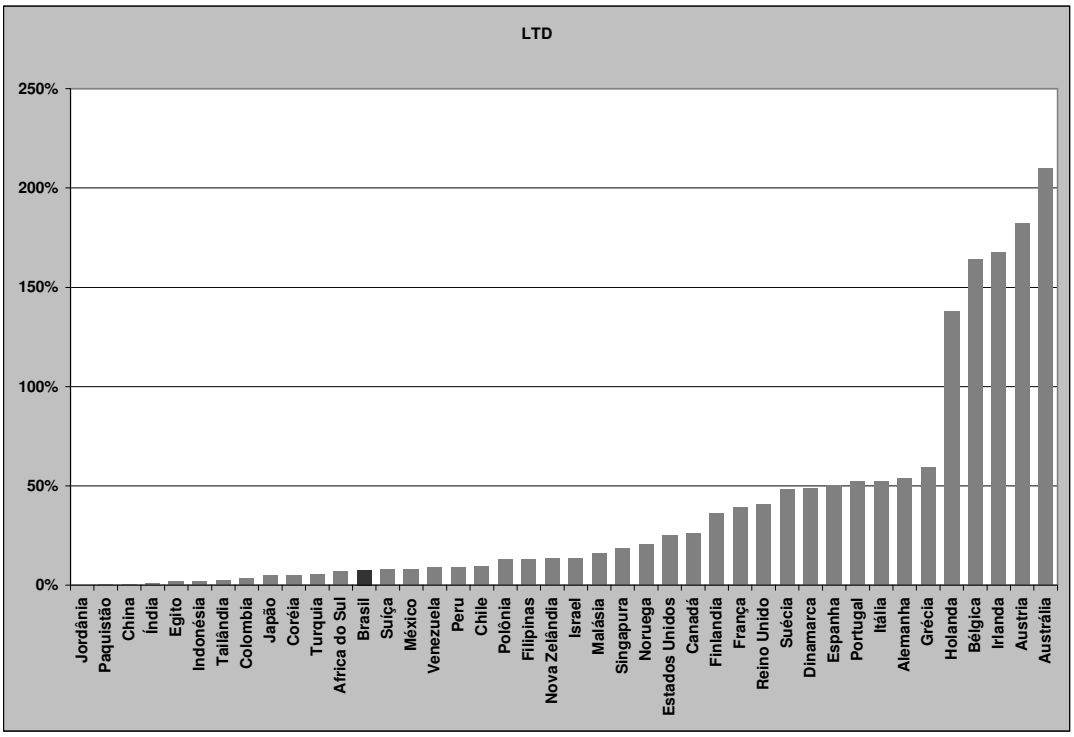
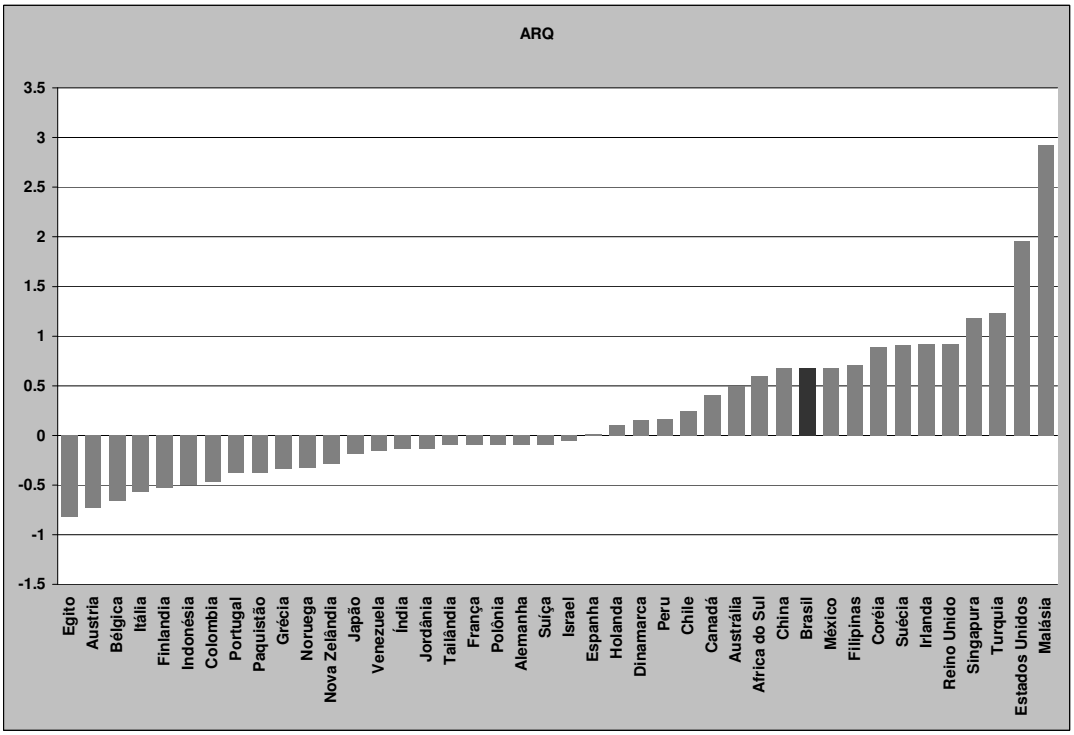


Gráfico 3: Variável Arquitetura dos mercados



O Brasil aparece em 27º lugar em desenvolvimento do mercado acionário, e em 30º em desenvolvimento do mercado de crédito. Em ambos os casos uma indicação de que estamos distantes dos países mais desenvolvidos nesses dois aspectos. Já quando a análise é feita pela variável ARQ, o Brasil sobe para a 11ª posição dentre os países da amostra, indicando que estaríamos entre os que estão mais voltados para o mercado acionário.

Já os determinantes, ou variáveis independentes do nosso modelo, são 17. Elas podem ser classificadas como relativas a condições gerais da economia do país em estudo, seu grau de inserção no mercado internacional, ou globalização, e culturais. A tabela a seguir apresenta as variáveis, suas fontes, período de apuração dos dados, métrica e sigla, classificação e indicação de relação.

Tabela 3: As variáveis do estudo: origem, período, classificação e indicação.

Variável	Descrição	Fonte	Período	Classificação	Indicação
MKCAP	Valor das empresas listadas no país / PIB	World Bank	2005		
ARQ	Arquitetura do mercado financeiro	Tadesse (2002)	2000		
LTD	Dívida de longo prazo / PIB	World Bank	2005		
SUP	Superávit ou déficit nas contas pública / PIB	World Bank	2005	Economia	Equilíbrio nas contas públicas
EXPOR	Exportações / PIB	World Bank	2005	Internacionalização	Grau de globalização
IDE	Investimento Estrangeiro Direto / PIB	World Bank	2005	Internacionalização	Grau de globalização
GNIC	Renda per Capita / Total da Renda per Capita da amostra	World Bank	2005	Economia	Riqueza individual
GCF	Formação bruta de capital / PIB	World Bank	2005	Economia	Formação de capital
IMP	Importações / PIB	World Bank	2005	Internacionalização	Grau de globalização
IVA	Valor agregado pela indústria / PIB	World Bank	2005	Economia	Sofisticação do parque industrial
INF	Inflação anual	World Bank	2005	Economia	Estabilidade
POP	Crescimento População	World Bank	2005	Economia	Estabilidade
UNIV	% de alunos no terceiro grau	World Bank	2005	Cultura	Presença nas universidades
DAYS	Tempo requerido para abrir um negócio (dias)	World Bank	2005	Cultura	Empreendedorismo
LI	Liberdade de Imprensa	Reportes Sans Frontieres	2005	Cultura	Liberdade
CORR	Corrupção	Transparencia INternacional	2005	Cultura	Corrupção
DI	Disseminação da informação	Bushman et all (2004)	2003	Cultura	Informação
LEC	Liberdade Econômica	Heritage Foundation - Wall Street Journal	2005	Cultura	Empreendedorismo
OSL	Origem do Sistema Legal	Beck et all (2000)	1999	Cultura	Sistema legal

Observamos que em três casos os períodos de apuração dos dados são diversos. São as variáveis ARC, DI e OSL. A última é atemporal por natureza, e para as duas outras podemos supor que têm uma velocidade de modificação bastante pequena, já que estão muito mais associadas a características perenes das economias. Isso nos permite utilizar dados de diferentes períodos na análise, sem comprometer a robustez dos resultados.

Somando-se as 15 variáveis listadas as duas informações sobre os mercados financeiros, temos um total de 966 dados possíveis (42 países x 23 informações). Não obtivemos 11 informações, sendo que seis da variável DI – Disseminação da Informação. Os dados faltantes foram substituídos pelas médias das respectivas variáveis, como proposto por Gujarati (2003).

A tabela a seguir apresenta o resumo estatístico das diferentes variáveis.

Tabela 4: Variáveis – Estatísticas Descritivas

	MKCAP	ARC	LTD	SUP	EXPOR	IDE	GNIC	GCF	IMP	IVA	INF	POP	UNIV	DAYS	LI	CORR	DI	LEC	OSL
Média	0.88	0.25	0.38	-0.01	0.44	0.03	0.02	0.22	0.42	0.32	0.04	0.01	0.52	36.33	17.80	6.06	73.11	68.28	0.60
Mediana	0.69	0.07	0.14	0.00	0.37	0.03	0.02	0.21	0.33	0.30	0.03	0.01	0.56	25.50	8.50	6.40	77.15	67.55	1.00
Máximo	2.96	2.93	2.10	0.16	2.44	0.17	0.07	0.43	2.15	0.52	0.29	0.02	0.92	152.00	83.00	9.60	96.72	85.70	1.00
Mínimo	0.03	-1.10	0.00	-0.16	0.10	-0.15	0.00	0.15	0.11	0.21	-0.02	0.00	0.05	2.00	0.50	2.10	29.51	47.70	0.00
Devio Padrão	0.62	0.81	0.54	0.05	0.39	0.04	0.02	0.05	0.34	0.08	0.05	0.01	0.23	37.04	19.80	2.51	18.98	9.69	0.50
Skewness	1.58	1.15	2.00	0.42	3.64	-0.47	0.34	2.01	3.37	1.00	3.33	0.37	-0.12	2.06	1.41	-0.07	-0.50	-0.05	-0.39
Kurtosis	5.49	4.63	5.91	8.30	18.67	9.37	1.84	8.08	16.79	3.11	17.01	2.28	2.10	6.66	4.40	1.52	2.20	2.01	1.15
Jarque-Bera	28.30	13.90	42.72	50.36	522.74	72.54	3.16	73.43	412.47	6.98	421.12	1.84	1.51	52.97	17.28	3.88	2.88	1.72	7.04
Prob.	0%	0%	0%	0%	0%	0%	21%	0%	0%	3%	0%	40%	47%	0%	0%	14%	24%	42%	3%

Alguns detalhes emergem da análise dessas estatísticas. Em primeiro lugar o maior volume no valor das empresas em relação ao débito. Em que pese que o valor das empresas incorpore o valor do débito, mesmo assim temos um valor de capital próprio (diferença entre MKCAP e LTD), superior ao do endividamento, talvez sugerindo uma subutilização da alavancagem pelas empresas.

Outro dado interessante diz respeito às importações e exportações. Os valores médios são muito similares, indicando um equilíbrio nos países analisados. A análise de mínimos e máximos de outros dados, como o DAYS - número de dias necessário para a abertura de um negócio, CORR – Corrupção, LEC – Liberdade Econômica mostram a grande diversidade entre os países.

A análise de correlações entre as variáveis, apresentada a seguir, indicou a existência de multicolinearidade.

Tabela 5: Variáveis - Correlações

	EXPOR	IDE	GNIC	GCF	IMP	IVA	INF	POP	UNIV	DAYS	LI	CORR	DI	LEC	OSL
SUP	19%	0%	40%	-2%	7%	34%	11%	-10%	33%	-2%	-22%	42%	34%	37%	1%
EXPOR		42%	10%	-6%	96%	26%	-14%	30%	8%	-20%	17%	27%	15%	33%	-22%
IDE			-17%	-19%	50%	-3%	-8%	10%	-13%	-8%	29%	6%	-6%	-1%	5%
GNIC				-19%	3%	-35%	-41%	-47%	72%	-51%	-68%	86%	90%	74%	18%
GCF					-4%	27%	-2%	-6%	-6%	0%	26%	-18%	-22%	-16%	-33%
IMP						14%	-22%	34%	6%	-26%	19%	24%	10%	31%	-27%
IVA							52%	27%	-28%	37%	36%	-31%	-29%	-30%	-20%
INF								37%	-38%	65%	26%	-50%	-40%	-54%	4%
POP									-50%	21%	51%	-36%	-49%	-21%	-41%
UNIV										-50%	-67%	71%	81%	60%	15%
DAYS											23%	-60%	-52%	-59%	11%
LI												-64%	-71%	-54%	-20%
CORR													88%	90%	5%
DI														76%	11%
LEC															-16%

Seguindo Gujarati (2003), que sugere existência de multicolinearidade quando a correlação entre as variáveis for igual ou superior a 80%, identificamos indicações de existência de multicolinearidade entre as variáveis GNIC e CORR e DI, bem como entre UNIV e DI e EXP e IMP. Essas correlações estão assinaladas com fundo cinza na tabela. Em que pese que esta não seja a única indicação possível de multicolinearidade, ela já nos indica que algumas dessas variáveis não devem ser consideradas em conjunto.

4 – METODOLOGIA E RESULTADOS

A influência das diferentes variáveis sobre o desenvolvimento dos mercados financeiros será estimada através da seguinte regressão múltipla:

$$D_i = \beta_i + \beta_{1i}V_1 + \beta_{2i}V_2 + ...+\beta_{14i}V_{14} + \varepsilon_i$$

Onde D_i representa as três variáveis dependentes que serão utilizadas no estudo (ARQ, MKCAP, LTD,) e os V_{ni} representam as diferentes variáveis explicativas. A correção para heterocedasticidade foi realizada seguindo os procedimentos propostos por White e constantes em Gujarati (2003).

Analisamos diversas configurações possíveis de modelos, com diferentes transformações das variáveis, e selecionamos os três melhores para cada uma das variáveis dependentes através da análise dos critérios de Akaike e Schwarz.

Para a variável dependente MKCAP obtivemos o seguinte modelo:

Tabela 6: Modelo MKCAP

Variable	Coefficient	Std. Error	t-Statistic	Prob.	Sig.
IMP	0.449	0.254	1.769	8.53%	*
IDE	4.916	2.275	2.161	3.74%	**
EXP(GNIC)	0.514	0.153	3.368	0.18%	***
1/POP	5.6E-05	2.1E-05	2.656	1.17%	**
(LI)^2	-1.2E-04	4.7E-05	-2.526	1.61%	**
1/DAYS	1.296	0.476	2.725	0.99%	***
Significância:	* 10%	** 5%	*** 1%		

Ele nos indica que o maior desenvolvimento do mercado acionário está associado a uma maior globalização, seja através da variável IMP, seja através da variável IDE, a uma maior renda per capita, uma menor taxa de crescimento da população, uma maior liberdade de imprensa, aliás, de forma acelerada como mostra a transformação da variável e finalmente a uma maior facilidade para o desenvolvimento de negócios. Destaque-se que os coeficientes para POP e LI são pequenos. O R2 ajustado deste modelo é de 33,63%. Eram os sinais esperados, que mostram a influência dos diferentes fatores no desenvolvimento dos mercados acionários.

Já quando a variável dependente é o débito de longo prazo, LTD, o modelo obtido foi:

Tabela 7: Modelo LTD

Variable	Coefficient	Std. Error	t-Statistic	Prob.	Sig.
IMP	0.493	0.152	3.245	0.25%	***
IDE	-4.740	1.792	-2.645	1.20%	**
EXP(GNIC)	4.886	2.556	1.911	6.39%	*
1/POP	-3.82E-05	1.82E-05	-2.094	4.34%	**
LI	-0.007	0.003	-2.126	4.05%	**
EXP(INF^2)	-4.549	2.577	-1.765	8.61%	*
Significância:	* 10%	** 5%	*** 1%		

Dois dos sinais aqui são similares aos obtidos no modelo referente ao mercado acionário. Quanto maior o nível de globalização do país, dado pela variável IMP, maior o desenvolvimento do mercado de crédito. O mesmo para a variável GNIC que representa a renda per capita. Já o indicador de inflação mostra um resultado interessante. O tamanho do mercado de crédito é sensível a altas taxas de inflação, mas não às pequenas. Assim, países com elevada taxa de inflação não têm um mercado de crédito desenvolvido. Também um resultado esperado. Interessante o resultado referente à população, maiores mercados de crédito estão associados a maiores taxas de crescimento. Mas temos que levar em conta que o coeficiente é muito pequeno, mostrando uma influência restrita desse indicador. Talvez como contraposição ao menor desenvolvimento do mercado acionário como fonte de receita das empresas. O coeficiente da variável IDE, Investimento Estrangeiro Direto, mostra que países com menor ingresso de capitais estrangeiros tem maior desenvolvimento do mercado de crédito, indicando uma reação das economias à falta de uma alternativa importante para o financiamento das empresas. Finalmente a variável LI, Liberdade de Imprensa apresenta um sinal negativo, indicando que países com menor liberdade de imprensa têm maior desenvolvimento do mercado de crédito, ou bancário. Aqui uma clara contraposição ao resultado obtido para o mercado acionário onde a sensibilidade do tamanho do mercado é claramente associada a uma maior liberdade de imprensa.

O modelo para a variável dependente ARQ – arquitetura do mercado é:

Tabela 8: Modelo ARQ

Variable	Coefficient	Std. Error	t-Statistic	Prob.	Sig.
EXP(EXPOR)	0.114	0.059	1.951	5.91%	*
1/(IDE)	-3.73E-04	1.79E-04	-2.086	4.43%	**
1/DAYS^2	1.350	0.654	2.065	4.64%	**
EXP(DI)	1.58E-42	2.04E-43	7.722	0.00%	***
1/(INF)^2	2.82E-06	1.04E-06	2.696	1.07%	**
1/GNIC^2	-4.11E-07	1.70E-07	-2.414	2.11%	**
SUP	-4.718	2.272	-2.077	4.52%	**
Significância:	* 10%	** 5%	*** 1%		

A variável ARQ mostra o quão um país é mais orientado para o mercado acionário ou mercado de crédito. Quanto maior, mais voltado para o mercado acionário. Com isso em mente vemos que países com maior nível de globalização, dado pela variável EXPOR, são mais voltados para o mercado acionário. Maior volume de IDE - Investimento Estrangeiro Direto também é uma característica presente nos países com maior desenvolvimento dos mercados acionários, assim como DI – Difusão da Informação. Um resultado interessante refere-se à variável DAYS. O coeficiente associado à transformação 1/DAYS^2 mostra que ter uma economia voltada para o empreendedorismo, aqui representada por poucos dias necessários para a abertura de um novo negócio, é fundamental para o desenvolvimento dos mercados acionários. O mesmo resultado foi obtido para a variável INF, indicando que altas taxas de inflação são muito prejudiciais para o desenvolvimento do mercado de ações, levando o país a ser mais voltado para o mercado de crédito. Também o resultado da variável GNIC não surpreende. Há um forte impulso aos mercados financeiros, em particular o mercado acionário, dado por altos níveis de renda per capita. Finalmente a variável SUP é interessante, mostrando que os mercados acionários seriam mais desenvolvidos em países com maior déficit público. Isso talvez se explique pela presença de países com mercados acionários bastante desenvolvidos, como EUA e Reino Unido, dentre outros, com elevados níveis de déficit.

5 – CONCLUSÕES

Diversas conclusões emergem do estudo. E outras tantas dúvidas. A tabela a seguir resume os resultados obtidos:

Tabela 9: Resumo dos coeficientes obtidos

MKTCAP		**LTD**		**ARQ**	
				EXP(EXPOR)	0.114
IMP	0.449	IMP	0.493		
IDE	4.916	IDE	-4.740	1/(IDE)	-3.73E-04
EXP(GNIC)	0.514	EXP(GNIC)	4.886	1/GNIC^2	-4.11E-07
1/POP	5.6E-05	1/POP	-3.82E-05		
(LI)^2	-1.2E-04	LI	-0.007		
1/DAYS	1.296			1/DAYS^2	1.350
		EXP(INF^2)	-4.549		
				EXP(DI)	1.58E-42
				1/(INF)^2	2.82E-06
				SUP	-4.718

O fundo cinza nos coeficientes indica sinal oposto ao encontrado nos modelos das outras variáveis dependentes. Assim podemos concluir que países que têm mercado de ações desenvolvido são países mais globalizados, com maior presença de capital estrangeiro, maior renda per capita, maior liberdade de imprensa, mais voltados para o empreendedorismo e finalmente com menor taxa de crescimento populacional. Boa parte dessas afirmações pode ser repetida para o mercado como um todo. Países que o tem mais desenvolvido são também mais globalizados, com maior renda per capita, mais voltados para o empreendedorismo e com baixas taxas de inflação. Os resultados para o mercado de crédito indicam também que países com maior volume de crédito relativo têm também maior nível de globalização, maior liberdade de imprensa, maior renda per capita e menores taxas de inflação. Eles também apresentam maior crescimento populacional e menor volume de investimentos estrangeiros diretos. No primeiro caso há uma relação direta entre taxa de crescimento populacional e sistema legal de origem francesa. E nos países de origem francesa, segundo a literatura, há um menor desenvolvimento do mercado acionário, sendo ele substituído como fonte de recursos para as empresas pelo mercado de crédito. No segundo caso podemos explicar o sinal do coeficiente por um efeito

substituição. Não há recursos externos, de alguma forma se desenvolvem as possibilidades internas.

Na comparação com estudos anteriores, temos resultados interessantes. No caso da variável IDE – investimento estrangeiro direto, em que Alfaro e al (2002) concluem pela sua pouca relevância na análise do desenvolvimento dos mercados, nós temos uma conclusão distinta. Nos três modelos estimados obtivemos coeficientes signficativos a 5%, sendo que nos modelos em que as variáveis dependentes eram MKCAP e ARC há clara indicação de que países com mercados mais desenvolvidos e mais voltados para o mercado acionário tem maiores volumes de investimento estrangeiro direto. Esta mesma variável está negativamente associada ao desenvolvimento de mercados de crédito, sugerindo uma contraposição deste em relação aos mercados acionários.

Outra variável que foi utilizada neste trabalho e também em trabalho anterior, foi GNIC – nível de renda per capita. Caprio e Honohan (2004) concluem que não há relação entre nível de renda e disciplina nos mercados. Estudamos outras variáveis dependentes e concluímos que a renda per capita tem relação com o nível de desenvolvimento dos mercados. Nossa análise indica que há uma relação exponencial entre o nível de desenvolvimento dos mercados nos diferentes países e a renda per capita dos mesmos.

A continuidade da análise dos resultados passa necessariamente pela análise de eficiência de mercado, já que uma das variáveis importantes identificada é a liberdade de imprensa. Esta variável pode ser considerada uma *proxy* para a fluidez da informação, que é fundamental para o desenvolvimento de mercados eficientes. Também é importante destacar o papel da maior globalização nesse contexto. Países com maior índice de globalização tendem a ter mercados mais eficientes até pela criação de vasos comunicantes maiores com o exterior, o que os leva a ter maior transparência nas suas ações. Isso fica claro na análise de risco conduzida pelas agências de classificação de risco, onde é dado grande peso para o indicador chamado de exposição internacional, que é o resultado da divisão da soma de exportações com importações pelo total do PIB.

Outro fator que deve ser destacado é o viés pelo empreendedorismo. O número de dias para abrir um novo negócio foi aqui usado como *proxy* para esse viés. E há claras indicações de que países onde o ambiente institucional é mais favorável aos negócios têm mercados mais desenvolvidos, em particular mercados acionários. Esse é um resultado muito interessante para reguladores, já que há caminhos para a alteração desse padrão nos países onde o ambiente não é tão favorável aos negócios.

Também a renda per capita é um fator determinante do nível do desenvolvimento dos mercados. Podemos então esperar que, juntamente com outros indicadores, o crescimento deste fator traga maior desenvolvimento aos mercados.

A próxima questão é: e como o Brasil está em relação às variáveis que caracterizam os países com mercados mais desenvolvidos? A tabela a seguir mostra a colocação do nosso país nos diferentes indicadores, considerando sempre que a 1ª colocação no ranking é a melhor.

Tabela 10: Brasil e as variáveis

	IMP	IDE	GNIC	POP	LI	DAYS	INF	EXP	DI
Brasil	41	31	32	28	27	42	37	40	36

A tabela mostra que o Brasil ainda está muito distante de ter as características básicas para apresentar um bom desenvolvimento dos mercados financeiros. Estamos nos últimos lugares quando classificamos a amostra de 42 países pela variável DAYS, que indica quantos dias são necessários para abrir um novo negócio, também em IMP – importações e EXP – exportações. A situação não é muito distinta em DI – difusão da informação, LI – liberdade de imprensa, INF - inflação, GNIC – renda per capita POP – taxa de crescimento populacional. Podemos concluir que estamos ainda muito distantes dos países com mercados mais desenvolvidos e que um esforço bastante direcionado deve ser produzido para que em alguns anos nos posicionemos melhor.

Mas também é importante ressaltar que há um esforço grande no sentido de tornar os nossos mercados financeiros mais próximos dos mercados desenvolvidos. Esses esforços, notadamente no campo regulatório, incluída aqui a auto-regulação, têm surtido efeitos positivos. Um primeiro efeito que podemos destacar refere-se ao mercado de crédito. A ampliação das fontes de crédito de longo prazo, antes restritas quase que exclusivamente ao BNDES, é um fato recente. Também a ampliação dos prazos de financiamentos para bens de consumo em níveis impensáveis há poucos anos é um sinal positivo. Do lado do mercado acionário observamos uma depuração nas empresas listadas. Ao mesmo tempo em que cem novas empresas foram listadas na Bovespa desde 2004, outras 103 se deslistaram a pedido entre 2000 e 2005, conforme Horng e Eid Jr (2005). Segundo esses autores essas empresas têm como característica comum a falta de crescimento, a baixa liquidez e baixo free-float. Aparentemente houve uma troca, com o ingresso de empresas de melhor qualidade no pregão.

Mas ainda no mercado acionário podemos destacar o crescimento dos investidores individuais. O número deles cresceu a taxas anuais próximas de 20% entre 2004 e 2007, segundo dados da Bovespa. Mas também é importante ressaltar que outros fatores vem tornando o brasileiro mais próximo ao mercado acionário. O crescimento da indústria de fundos de investimentos e também de fundos de pensão é fator fundamental. Temos hoje aproximadamente 6,5 milhões de brasileiros participando de alguma forma dos 300 fundos de pensão existentes. E todos eles com aplicações no mercado acionário. Há dez anos esse número era infinitamente menor. Também do lado dos fundos de investimentos os números são surpreendentes. Apenas os dois maiores fundos de ações somam 2 milhões de cotistas. E novamente há dez anos esse número era muito menor. Podemos hoje estimar em aproximadamente 10 milhões o número de brasileiros que têm investimentos em ações. E esse número deve crescer mais nos próximos anos. Será interessante refazer esse estudo em 2017.

Cabe aqui destacar que algumas das variáveis selecionadas inicialmente não se mostraram significativas quando do processamento das análises. Algumas foram eliminadas já no início das análises por serem substituíveis por outras, dadas suas altas correlações. Foram os casos das variáveis CORR – corrupção e DI - difusão de informações. As variáveis IMP - importação e EXP - exportação foram utilizadas alternadamente também dada sua elevada correlação. Possivelmente temos casos de multicolineridade entre LEC – liberdade econômica, UNIV – percentual de alunos nas universidades e outras variáveis como GNIC – renda per capita, já que ambas apresentam correlação elevada, de 74%. Mas das variáveis que não se mostraram significativas nas regressões a que mais surpreendeu foi OSL – origem do sistema legal. Em outros estudos, como o de Dermigue-Kunt e Levine (1999), essa variável se mostrou significativa em relação ao desenvolvimento dos mercados financeiros. No nosso trabalho não. Mesmo numa regressão simples os resultados foram não significativos. Por vezes significativos, mas com R^2 ajustado negativo ou nulo. Podemos conjecturar que sua construção não foi a mais adequada, já que se trata de uma variável *dummy* dentre outras tantas com características que as aproximam muito de variáveis contínuas. Também podemos imaginar efeitos de multicolineridade, que a análise de correlação não capturou. E aí podemos ter a indicação da origem do sistema legal oculta em alguma das variáveis. De toda forma foi surpreendente não encontrarmos uma relação significativa entre a origem do sistema legal e o desenvolvimento dos mercados.

Nesse estudo fizemos a análise de características de paises e o grau de desenvolvimento dos mercados financeiros. Fizemos praticamente uma foto do momento e comparamos o desenvolvimento dos mercados com diferentes indicadores. Um caminho natural para próximas pesquisas é a elaboração de um estudo de painel buscando analisar a evolução das características e dos mercados. No entanto é preciso ressaltar que há grande dificuldade para acessar dados de longo prazo para a maioria dos países analisados. Também a reconstrução de algumas variáveis e mesmo a inclusão de outras deve ser um caminho de continuidade deste trabalho.

REFERÊNCIAS BIBLIOGRÁFICAS

ALFARO, L. et al. (2002), "FDI and economic growth: The role of local financial markets". ssrn-id305762.

BARTH, J.R.; CAPRIO, C.; LEVINE, R. (2000), "Banking systems around the globe: Do regulations and ownership affect performance and stability?". World Bank Policy Research Working Paper n. 2325, ssrn-id 630697.

BECK, T.; LEVINE, R. (2003), "Legal institutions and financial development". NBER Working Paper nº 10126, ssrn-id636556.

BECK, T.; LEVINE, R.; DEMIRGÜÇ-KUNT, MAKSIMOVIC, V. (2000), "Financial structure and economic development: Firm, industry, and country evidence". World Bank Policy Research Working Paper n. 2423.

BOUTCHKOVA, M.K.; MEGGINSON, W.L. (2000), "Privatization and the rise of global capital markets, Financial Management", vol. 29, issue 4, p 31-75.

CAPRIO, C.; HONOHAN, P. (2004), "Can the unsophisticated market provide discipline?". World Bank Policy Research Working Paper n. 3364, ssrn-id 625241.

CHINN, M.D.; ITO, H. (2005), "What matters for financial development? Capital controls, institutions and interactions". Working Paper University of Wisconsin and Portland State University, ssrn-id677506.pdf.

CLAESSENS, S.; DJANKOV, S.; NENOVA, T. (2000), "Corporate risk around the world". World Bank Policy Research Working Paper n. 2271, ssrn-id629121.

DERMIGUE-KUNT, A.; LEVINE, R. (1999), "Bank based and market based financial systems: Cross-country comparisons". World Bank Working Paper, ssrn-id569255.

FAN, J.P.H.; TITMAN, S. (2003), "An international comparison of capital structure and debt maturity choices". Working Paper Hong Kong University of Science and University of Texas at Austin, ssrn-id423483.

FRANCIS, J.R.; KHURANA, I.K.; PEREIRA, R. (2001), "Investor protection laws, accounting and auditing around the world". Working Paper University of Missouri, ssrn-id287652.

FUCHS-SCHÜNDELN, N.; FUNKE, N. (2001), "Stock Market liberalizations: Financial and macroeconomic implications". IMF Working Paper WP/01/193, ssrn-id296509.

GUJARATI, D. (2003), "Basic Econometrics". 4º Ed. McGraw-Hill Irwin.

HERRERO, A.G. (2003), "Latin American Financial Development in Perspective". Working Paper Banco de España.

Horng, W. J.; Eid Jr, W. (2005), "A Saída – Uma análise da deslistagem na Bovespa" – Manuscrito – EAESP/FGV.

KORAJCZYK, R.A. (1996), "A measure of stock market integration". World Bank Economic Review, vol. 10, nº. 2. pp 267-289.

LA PORTA, R.; LOPEZ-DE-SILANEZ, F.; SHLEIFER, A.; VISHNY, R.W. (2000), "Investor protection and corporate governance". Journal of Financial Economics.

LA PORTA, R.; LOPEZ-DE-SILANEZ, F.; SHLEIFER, A.; VISHNY, R.W. (1998), "Law and finance". Journal of Political Economy, vol 52, pp 1131-1150.

LA PORTA, R.; LOPEZ-DE-SILANEZ, F.; SHLEIFER, A.; VISHNY, R.W. (1999), "The quality of government". Journal of Law and Economics Organization, vol 15 pp 222-279.

LEVINE, R.; ZERVOS, S. (1996), "Stock Market development and long-run growth". The World Bank, Policy Research Working Paper 1582, ssrn-id614946.

MODIGLIANI, F.; PEROTTI, E. (1997), "Protections of minority interest and the development of sercurity markets". Managerial and Decision Economics, vol 18, nº 7/8, p. 519-528.

NDIKUMANA, L. (2001), "Financial development, financial structure and domestic investment: International Evidence". Working Paper nº 16 PERI - University of Massachusetts.

TADESSE, S. (2002), "Financial architeture and economic performance: International Evidence". Working Paper, The University of South Carolina, ssrn-id307223.pdf.

WORLD BANK – "World Development Indicators" – (2005).

Apêndice 1: Dados

País	MKCAP	LTD	ARQ	SUP	EXP	IDE	GNIC	GCF	IMP	IVA	INF	POP	UNIV	DAYS	LI	CORR	DI	LEC	OSL
Australia	110%	210%	2.10	1.2%	18.4%	-4.7%	3.7%	25.7%	21.2%	27.1%	4.7%	1.2%	71.8%	2	6.5	8.8	89.25	82.7	0
Austria	41%	182%	1.82	-2.0%	53.2%	3.0%	4.2%	21.2%	47.8%	30.7%	2.0%	0.7%	50.3%	29	2.5	8.7	87.53	71.3	1
Belgium	88%	164%	1.64	-0.1%	87.1%	8.6%	4.0%	21.4%	84.8%	24.0%	2.2%	0.6%	63.0%	34	4.0	7.4	86.73	74.5	1
Brazil	54%	8%	0.08	0.0%	15.1%	1.7%	0.4%	16.0%	11.5%	30.3%	7.5%	1.4%	23.8%	152	14.5	3.7	56.14	60.9	1
Canada	133%	26%	0.26	1.7%	38.6%	3.1%	3.7%	21.0%	34.4%	31.4%	3.2%	1.0%	62.3%	3	4.5	8.4	93.37	78.7	0
Chile	115%	10%	0.10	4.6%	41.0%	5.6%	0.7%	22.4%	32.5%	46.8%	7.9%	1.1%	47.8%	27	11.8	7.3	62.46	78.3	1
China	35%	0%	0.00	-2.1%	37.3%	3.5%	0.2%	43.3%	31.7%	47.3%	4.4%	0.6%	20.3%	48	83.0	3.2	50	54	0
Colombia	37%	3%	0.03	3.9%	21.4%	8.4%	0.3%	20.4%	22.2%	34.4%	5.7%	1.4%	29.3%	43	40.2	4.0	58.13	60.5	1
Denmark	69%	49%	0.49	1.8%	48.7%	2.0%	5.4%	20.9%	43.9%	24.7%	2.6%	0.3%	80.5%	5	0.5	9.5	95.52	77.6	1
Egypt	89%	2%	0.02	-5.8%	30.3%	6.0%	0.1%	18.0%	32.6%	36.1%	6.2%	1.9%	33.9%	22	52.0	3.4	60	53.2	1
Finland	108%	36%	0.36	3.2%	38.7%	2.1%	4.2%	20.2%	35.2%	29.6%	1.6%	0.3%	91.7%	14	0.5	9.6	94.82	76.5	1
France	80%	39%	0.39	-2.8%	26.1%	3.3%	3.9%	20.2%	27.1%	20.9%	1.9%	0.6%	56.4%	8	6.3	7.5	86.14	66.1	1
Germany	44%	54%	0.54	-2.3%	40.1%	1.1%	3.9%	17.2%	35.1%	29.7%	0.5%	-0.1%	60.0%	24	4.0	8.2	90.99	73.5	1
Greece	64%	59%	0.59	-5.1%	20.8%	0.3%	2.2%	23.8%	28.0%	20.8%	3.7%	0.4%	89.1%	38	4.0	4.3	72.07	57.6	1
India	69%	1%	0.01	-3.6%	20.3%	0.8%	0.1%	33.4%	23.3%	27.3%	4.5%	1.4%	11.4%	71	27.0	2.9	29.51	55.6	0
Indonesia	28%	2%	0.02	-1.1%	32.9%	1.8%	0.1%	21.8%	28.6%	44.9%	14.8%	1.4%	17.1%	151	26.0	2.2	60	55.1	0
Ireland	57%	168%	1.68	1.0%	83.2%	-14.7%	4.6%	25.0%	67.8%	37.3%	3.5%	2.2%	59.3%	24	0.5	7.4	83.34	81.3	0
Israel	97%	14%	0.14	-2.6%	45.9%	4.5%	2.1%	18.5%	51.3%	30.0%	0.6%	1.8%	57.9%	34	10.0	6.3	82.47	68.4	0
Italy	45%	52%	0.52	-3.5%	26.3%	1.1%	3.4%	20.9%	26.4%	26.9%	2.1%	0.7%	65.6%	13	8.7	5.0	78.98	63.4	1
Japan	104%	5%	0.05	0.0%	13.4%	0.1%	4.4%	22.7%	11.4%	30.2%	-1.8%	0.0%	55.3%	31	8.0	7.3	91.79	73.6	1
Jordan	296%	0%	0.00	-4.7%	51.9%	12.1%	0.3%	23.7%	93.4%	29.6%	4.0%	2.3%	39.2%	18	24.0	5.7	60	64	0
Korea	91%	5%	0.05	0.8%	42.3%	0.5%	1.8%	30.1%	39.9%	40.3%	-0.2%	0.4%	89.9%	22	7.5	5.0	83.5	68.6	0
Malaysia	139%	16%	0.16	-4.3%	123.4%	3.0%	0.6%	19.9%	99.9%	51.8%	4.6%	1.8%	32.0%	30	33.0	5.1	63.83	65.8	0
Mexico	31%	8%	0.08	-1.2%	30.0%	2.4%	0.8%	21.8%	31.5%	26.0%	5.5%	1.0%	24.0%	58	45.5	3.5	59.95	65.8	1
Netherlands	117%	138%	1.38	0.0%	71.2%	6.5%	4.4%	19.2%	63.0%	24.4%	1.6%	0.2%	60.7%	11	0.5	8.6	92	77.1	1
New Zealand	37%	13%	0.13	4.6%	29.0%	1.8%	2.9%	24.6%	30.1%	24.9%	2.5%	0.9%	81.9%	12	2.0	9.6	90	81.6	0
Norway	65%	21%	0.21	16.3%	45.3%	1.1%	6.8%	20.6%	28.1%	43.4%	8.4%	0.7%	79.7%	13	0.5	8.9	95.31	70.1	1
Pakistan	41%	0%	0.00	-3.2%	15.5%	2.0%	0.1%	18.1%	19.3%	26.5%	8.7%	2.4%	4.7%	24	60.8	2.1	32.47	58.2	0
Peru	45%	9%	0.09	-0.8%	24.8%	3.2%	0.3%	18.6%	19.2%	34.8%	3.4%	1.5%	33.5%	102	33.3	3.5	40.33	62.1	1
Philippines	41%	13%	0.13	-3.0%	47.3%	1.2%	0.1%	15.1%	52.0%	32.3%	6.2%	1.8%	28.1%	48	50.0	2.5	44.26	57.4	1
Poland	31%	13%	0.13	-2.3%	37.2%	3.2%	0.8%	19.2%	37.5%	30.8%	2.6%	0.0%	63.4%	31	12.5	3.4	60	58.8	1
Portugal	37%	52%	0.52	-5.8%	28.5%	1.7%	1.9%	22.3%	37.4%	24.7%	2.7%	0.5%	56.5%	54	4.8	6.5	70.59	66.7	1
Singapore	179%	18%	0.18	4.1%	244.3%	17.2%	3.0%	19.0%	214.8%	33.8%	0.4%	2.4%	60.0%	6	50.7	9.4	83.72	85.7	0
South Africa	234%	7%	0.07	0.2%	26.8%	2.6%	0.5%	18.0%	28.3%	30.3%	4.8%	1.1%	15.3%	35	6.5	4.5	59.56	64.1	0
Spain	85%	50%	0.50	1.5%	25.4%	2.0%	2.8%	29.7%	30.6%	29.5%	4.4%	1.7%	66.9%	47	8.3	7.0	75.31	70.9	1
Sweden	113%	48%	0.48	2.1%	48.6%	3.0%	4.6%	17.1%	40.9%	28.2%	1.2%	0.4%	81.7%	16	2.0	9.2	95.47	72.6	1
Switzerland	256%	8%	0.08	0.6%	46.2%	4.2%	6.2%	20.4%	38.9%	28.3%	0.6%	0.6%	47.4%	20	0.5	9.1	93.78	79.1	1
Thailand	70%	3%	0.03	2.5%	73.8%	2.6%	0.3%	31.7%	75.4%	44.2%	4.5%	0.8%	43.0%	33	28.0	3.8	52.55	65.6	0
Turkey	44%	5%	0.05	-15.5%	27.4%	2.7%	0.5%	24.8%	34.0%	23.7%	5.4%	1.3%	31.2%	9	25.0	3.5	58.55	59.3	1
United Kingdom	139%	41%	0.41	-2.9%	26.1%	7.2%	4.2%	16.8%	30.0%	26.2%	2.0%	0.7%	59.7%	18	5.2	8.6	90.81	81.6	0
United States	137%	25%	0.25	-2.9%	10.1%	0.9%	4.9%	19.2%	15.4%	22.0%	3.0%	1.0%	82.7%	5	9.5	7.6	96.72	82	0
Venezuela	3%	9%	0.09	2.2%	39.7%	2.0%	0.6%	21.0%	17.4%	51.6%	29.0%	1.7%	41.2%	141	23.0	2.3	62.64	47.7	1

A IMPORTÂNCIA DA RESPONSABILIDADE SOCIAL DAS EMPRESAS E DOS CONSUMIDORES EM TEMPOS DE CRISE

PAULA CRISTINA LOPES RODRIGUES
E-mail: pcrislr@netcabo.pt
Investigadora do CEPESE
Universidade do Porto
SUSANA FERREIRA GABRIEL
Email:sferreiragabriel@gmail.com
Departamento de Gestão e Economia
Universidade da Beira Interior

Resumo

A Responsabilidade Social das Empresas (RSE) consiste na integração voluntária de preocupações sociais e ambientais, nas operações quotidianas das organizações, na sua interacção com todos os públicos-alvo. A RSE deve ser correspondida pela Responsabilidade Social do Consumidor (RSC), que por sua vez versa sobre o impacto dos comportamentos de consumo nas escolhas preocupadas com a sociedade, o meio ambiente e os direitos humanos.
A linha de investigação é centrada na relação entre a percepção dos consumidores sobre empresas com responsabilidade social e a intenção de compra dos seus produtos.
Interessa averiguar como é que em épocas de crise se justifica a pertinência da tomada desse tipo de medidas pelas empresas, que acarreta custos mas também pode trazer benefícios na sua interligação com o consumidor, e como tal, deve merecer o apoio governamental.
Para isso, recorre-se a um inquérito junto dos consumidores, no sentido de analisar até que ponto, as empresas devem promover políticas de responsabilidade social, com a cooperação do sector público, em resultado da percepção positiva que os consumidores demonstrem deste tipo de práticas.

Palavras-Chave: Responsabilidade Social, Empresas, Consumidores, Modelo de Equações Estruturais.
Área Temática: Estratégia Económica.

Abstract

Corporate Social Responsibility (CSR) is the voluntary integration of social and environmental concerns in the daily operations of organizations, in their interaction with all target audiences. CSR must be matched by the Consumer Social Responsibility (CSR), which in turn is about the impact of the behavior of consumption choices in a society concerned about the environment and human rights.
The line of research is focused on the relationship between the perception of consumers about companies with social responsibility and intention to purchase their products.
Interested investigate how in times of crisis justified the relevance of such measures taken by companies, which entails costs but can also bring benefits in its interconnection with the consumer, and as such, deserves the support of governments.
For this, it is a survey of consumers, to examine the extent to which companies must promote policies of social responsibility, with the cooperation of the public sector as a result of the positive perception that consumers are showing this kind of practice.

Keywords: Corporate Social Performance/ Responsibility; Conscientious Consumption, Structural Equation Modelling.
Thematic Area: Economics of Strategy.

INTRODUÇÃO

De acordo com os principais resultados apresentados na literatura sobre a Responsabilidade Social, os consumidores e o público em geral estão cada vez mais interessados no comportamento socialmente responsável das empresas. O aumento desse interesse justifica os investimentos empresariais e os apoios governamentais, que podem ser enfatizados em épocas de crise económica, pelo conjunto de vantagens que daí possam resultar.

A importância actual do tema está evidenciada na comunicação feita pela Comissão Europeia a propósito da publicação do seu "Relatório de Competitividade Europeia de 2008": "When re-launching the Lisbon Strategy in 2005, the Commission stated that Corporate Social Responsibility "can play a key role in contributing to sustainable development while enhancing Europe's innovative potential and competitiveness". The importance of CSR cannot be overestimated, not least since one lesson from the current financial crisis is that socially responsible entrepreneurs and CEOs are of utmost importance for the wellbeing of our societies".

O objectivo deste *paper* é investigar a relação entre a percepção dos consumidores sobre as empresas com responsabilidade social e a intenção de compra dos seus produtos. Interessa investigar como é que em épocas de crise essa ligação se comporta e se é recomendável um apoio expresso do sector público. Para isso são desenvolvidos os conceitos de responsabilidade social das empresas (RSE) e responsabilidade social dos consumidores (RSC), realçando a sua importância em momentos de crise económica, justificando o amparo governamental. Para isso, a ideia original é desenvolver um modelo de medida que permita estabelecer a relação entre a responsabilidade social do consumidor, traduzido em comportamentos de compra tais como a lealdade e a disposição em pagar um preço prémio pelo produto, e as estratégias empresariais de responsabilidade social, apoiadas, em época de crise pelo sector público.

RESPONSABILIDADE SOCIAL DAS EMPRESAS E DOS CONSUMIDORES:

Segundo a Comissão Europeia (2001), a RSE consiste na "integração voluntária de preocupações sociais e ambientais por parte das empresas nas suas operações e na sua interacção com outras partes interessadas". Esta entidade tem promovido o seu debate com a intenção de se definirem formas de actuação no âmbito da responsabilização cívica de todos os intervenientes económicos. Nesse sentido, apresentou em Julho de 2001, o Livro Verde "Promover um quadro europeu para a responsabilidade social das empresas" e apoiou a criação em 2002, do Fórum Multilateral sobre a RSE, que apresenta um relatório final sobre conclusões e recomendações neste tema.

Também o Instituto para a Segurança, Higiene e Saúde no Trabalho português apoia as medidas de que a RSE deve promover o "desenvolvimento de melhores práticas de responsabilidade social que deverá ter tradução na melhoria das condições de trabalho, nomeadamente de saúde, segurança e bem-estar, não pelo simples cumprimento da lei, mas por acções voluntárias de criação de ambientes mais saudáveis e gratificantes para quem neles trabalha e potenciadores de desenvolvimento profissional e pessoal."

Assim, a RSE envolve uma combinação de agentes internos e externos da empresa, e deve ser integrada nos seus objectivos estratégicos, abrangendo todo o processo produtivo desde a invenção, produção, armazenamento, distribuição e comercialização do produto. Autores como Porter, M. e Kramer, M.(2006) chamam à atenção para o carácter não inteiramente voluntário da adopção de medidas de RSE. Afirmam que muitas das empresas despertaram para a RSE pressionados por boicotes dos consumidores depois dos *media* e organizações activistas divulgarem práticas suas menos respeitáveis. E alertam que muitos dos relatórios de responsabilidade social demonstram que as suas práticas não foram tomadas de forma estratégica ou operacional no seio da actividade da empresa, mas antes como medidas de cosmética, o que pode levar à sua total inocuidade.

Um maior aprofundamento do conceito de RSE na pesquisa académica evidencia que, por um lado é necessário priorizar o envolvimento das empresas numa rede de relações sociais

(consumidores, fornecedores, administração pública, comunidade em que se insere, etc.) na identificação dos seus instrumentos estratégicos de gestão, por outro lado, no esforço das políticas de mercado e dos stakeholders na tradução da ideia de um comportamento social responsável (Perrini et. al., 2006).

As vantagens da RSE são inúmeras, apontadas por vários autores na determinação dos factores de competitividade das empresas. Kotler (1997) defende que com a globalização, a forma de uma empresa conseguir vantagens competitivas perduráveis no tempo, já não assenta nos factores físicos diferenciadores do seu produto, facilmente imitáveis pelos concorrentes, mas antes no carácter cívico da empresa: "[…] Portanto como uma empresa pode conquistar uma diferenciação mais sustentável na opinião pública? Eu argumentaria que uma base honestamente sustentável para a diferenciação é o carácter cívico da empresa. […] a imagem de uma empresa costuma persistir durante longo tempo. Essas imagens em geral revelam forte carga emocional capaz de criar vínculos com o cliente ou, no extremo oposto, afastá-lo. E bem pode vir a acontecer que à medida que os produtos forem se tornando mais iguais, a imagem cívica de uma empresa venha a ser um dos mais fortes factores de influência na preferência do cliente." Assim, fica defendida a ideia de que através de acções de RSE é possível obter um diferencial estratégico no mercado internacional.

Outra consequência imediata que recompensa empresas com preocupações sociais é o aumento do seu capital reputacional. Fombrum et al. (2000) destacam que não se pode estabelecer qualquer relação positiva entre medidas de RSE e o desempenho financeiro da empresa, mas antes com o valor dos activos intangíveis. Destaca a maior capacidade que uma empresa com reputação possui no recrutamento dos melhores profissionais, interligando este facto com um forte fomento à inovação interna da empresa.

Um terceiro benefício atribuído às medidas de RSE é o de consolidar a sustentabilidade da empresa. A necessidade das empresas desenvolverem práticas de gestão que lhes assegurem competitividade não só hoje mas também amanhã, faz com que o foco de atenção de muitas delas se dirija para as medidas de RSE, embora tais práticas coloquem o desafio de balancear os objectivos de longo-prazo com os custos de curto-prazo.

Do ponto de vista racional, se os stakeholders perceberem a empresa com um comportamento social responsável, consideram essa empresa como preferida para se transaccionar. Pelo lado da oferta, as empresas socialmente responsáveis conseguem atrair e reter os consumidores ou conseguem fazer-se cobrar por um preço prémio pelos seus produtos (Barnett, 2007). Existem vários estudos onde se defende a ideia que os consumidores estão dispostos a pagar um preço superior pelos produtos que derivem de empresas com um perfil de responsabilidade social, contudo, não conseguem relacionar as percepções dos consumidores com os seus comportamentos de compra (Page e Fearn, 2005). Também Boulstridge e Carrigan (2000) defendem que existe uma diferença entre as atitudes dos consumidores e o seu comportamento, isto é, apesar dos consumidores estarem dispostos a pagar um preço superior por um produto de uma empresa com responsabilidade social, a responsabilidade social da empresa não é um critério que os consumidores tenham em atenção quando vão às compras.

Contudo, na óptica do consumidor, as atitudes também têm evoluído no sentido do respeito pelo ambiente, em que a consciencialização pela cidadania colectiva tem ganho terreno, fortalecendo posturas de RSC. A necessidade dum modelo de consumo sustentável e o seu conceito é sublinhado pelo PNUD em 1998: "O consumo sustentável significa o fornecimento de serviços e de produtos correlatos, que preencham as necessidades básicas e dêem uma melhor qualidade de vida, ao mesmo tempo em que se diminui o uso de recursos naturais e de substâncias tóxicas, assim como as emissões de resíduos e de poluentes durante o ciclo de vida do serviço ou do produto, com a ideia de não se ameaçar as necessidades das gerações futuras." É necessário que o consumidor altere os seus hábitos de consumo, requerendo-se campanhas de sensibilização e mobilização social nesse processo de reaprendizagem. Quando percebe a existência de consciência social, o consumidor identifica-se com as empresas que compartem o mesmo exercício de cidadania, criando vínculos de fidelidade difíceis de ocorrer com entidades com valores contrários. Assim se demonstra que a RSE está fortemente interligada à RSC.

De acordo com os principais resultados da literatura sobre a responsabilidade social das empresas (RSE), os consumidores estão interessados no seu comportamento social e são

influenciados nas suas compras por esses comportamentos. Frequentemente, os consumidores defendem que o perfil de uma empresa sobre a responsabilidade social é importante na escolha e na compra de determinadas marcas (Castaldo et al, 2009).

Se o impacto da RSE nos consumidores fosse universal e significativo, dever-se-ia verificar um impacto claro no "bottom line" das empresas com uma forte performance social (Castaldo et al., 2009). Estes autores sugerem que a RSE tem impacto nas decisões dos consumidores se o produto que as empresas vendem contém uma proposta ética e se a reputação de RSE incluir um compromisso na protecção dos direitos e interesses dos consumidores. Apesar das condições parecerem restritas, hoje em dia muitas empresas fazem publicidade às suas práticas de responsabilidade social como forma de se distinguir da concorrência.

Existem alguns estudos que mostram que os consumidores são influenciados pela RSE (Smith, 2003), apesar de não existir pesquisas empíricas para suporte dessas conclusões (Bhattacharya e Sem, 2004; Castaldo e Perrini, 2004). A falta de suporte empírico reflecte-se na dificuldade notória de encontrar uma ligação percepcionada entre a performance social da empresa e a sua performance financeira.

Partindo deste ponto de vista, pretende-se analisar a ligação entre a responsabilidade social das empresas e o comportamento do consumidor socialmente responsável, através de um inquérito passado junto dos consumidores portugueses e desenvolvendo um modelo de equações estruturais de acordo com as percepções dos consumidores. E ainda verificar se esse comportamento sofre alterações em épocas de crise económica, justificando os apoios governamentais a este tipo de acções.

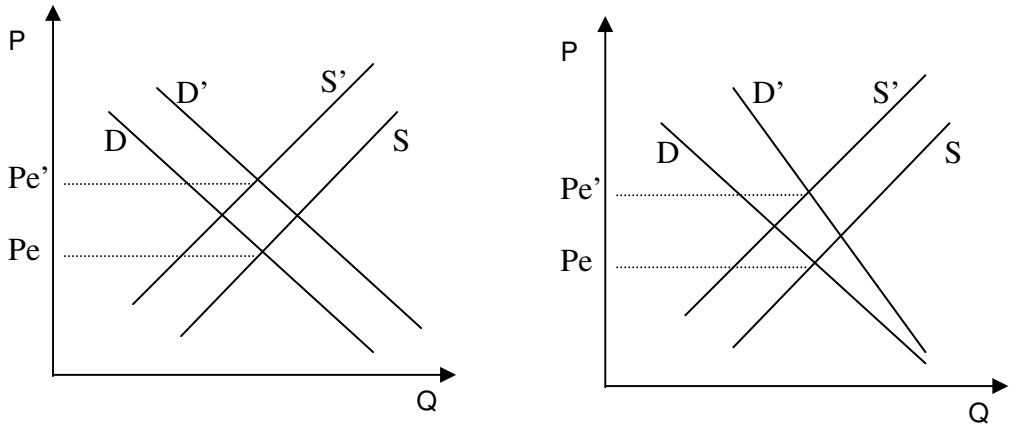
Um terceiro participante na cooperação entre a RSE e a RSC é o sector público. Um relatório do Banco Mundial dedicado ao tema (2002) destacou quatro funções governativas fulcrais no desenvolvimento das mesmas. No seu papel mandatário, o Sector Público Administrativo (SPA) deve legislar no sentido de incluir penalizações e recompensas fiscais ou legais a estas práticas ou à ausência delas. Na sua posição de simplificação, pode estimular o compromisso entre os vários actores da responsabilidade social, patrocinando e divulgando esse tipo de medidas. Na sua missão participativa, pode combinar os seus recursos com os dos restantes agentes económicos de forma a incentivar o mercado na aquisição do compromisso da responsabilidade social, dando ele próprio o exemplo. Por fim, o SPA pode desempenhar um papel patrocinador, dando suporte político a essas práticas, publicitando-as através da criação de prémios ou assinando acordos internacionais nesse âmbito.

O enquadramento legal que a RSE recebeu a nível internacional, já data de 1997. Baseando-se em 12 convenções da Organização Internacional do Trabalho, na declaração dos direitos do Homem das Nações Unidas e na Convenção das Nações Unidas dos direitos das Crianças, a norma SA 8000 [Responsabilidade Social] surgiu através da Social Accountability International [SAI], em colaboração com outras organizações internacionais. Destinada a auxiliar as empresas, a SA 8000, especifica requisitos de RSE, contribuindo para uma uniformização internacional de padrões. O Instituto Português da Qualidade editou em 2008, a NP 4469-1 relativa a "Sistemas de gestão da responsabilidade social", adaptando a norma internacional ao contexto empresarial português.

MODELO MICRO E MACRO ECONÓMICO DE MEDIDA:

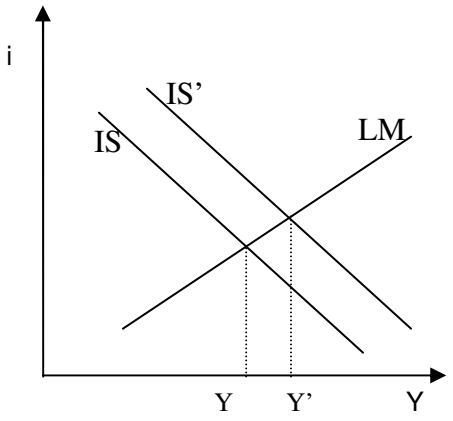
Para se ponderar as vantagens e inconvenientes que as medidas de RSE podem trazer a uma economia deprimida, utilizando instrumentos simples de análise económica como o modelo de oferta e procura, consegue-se adivinhar os resultados deste fenómeno.

Inegavelmente, a RSE constitui um aumento de custo para a empresa que a leva a cabo, provocando a contracção da sua curva da oferta. Ao mesmo tempo, com aquelas medidas a empresa pretendeu ganhar posição no mercado ao atrair novos e ao fidelizar anteriores consumidores, que são mais conscientes das suas responsabilidades sociais e que preferem comprar a produtores que partilham dos mesmos valores éticos. Assim, será de esperar que também uma alteração da procura se dê, em um de dois sentidos: ou como deslocação paralela para a direita ou como rotação da curva para a direita.



Se averiguarmos o resultado destas alterações no mercado em causa, há algo em comum às duas: o novo preço de equilibro é sempre mais elevado que na situação inicial de inexistência de RSE. Isto confirma a ideia do preço prémio que muita da bibliografia afirma ser a compensação mais imediata que o produtor recebe. Quanto às quantidades transaccionadas no novo equilíbrio nada se pode afirmar, pois estas vão depender da amplitude com que se dão os ajustamentos do lado da oferta ou da procura à incorporação da RSE. O efeito mais vasto é que determina a extensão nas quantidades. No caso destas se terem expandido, verificar-se-ia também um aumento do excedente social, resultado do beneficio acumulado por consumidores e produtores por terem intercambiado no mercado.

Numa perspectiva macroeconómica, utilizando uma ferramenta simples como é o modelo keynesiano de curto prazo, poderemos inferir ilações curiosas também. Consideremos que na globalidade da economia, as acções de RSE se traduzem em aumentos do agregado investimento e também numa redução das transferências administrativas para os privados. A sua fundamentação está, em que as acções de RSE por integrarem maioritariamente bens e serviços de carácter público (Ex: plano de saúde, formação profissional, etc.), passam do fornecedor Sector Público para o Sector Privado. O efeito multiplicador duma expansão do investimento é mais acentuado do que a contracção verificada ao nível das transferências, provocando uma expansão da produção da economia, a uma taxa de juro mais elevada que precedeu a tomada de medidas de RSE. Não estamos aqui a incluir deliberadamente as repercussões de atitudes de RSE no alcance de vantagens competitivas para as empresas que as tomam, pois isso também ajudaria a que o saldo com exterior fosse um ponto a favor e nada desdenhável do crescimento interno do país de origem dessas empresas responsáveis.



Com os instrumentos de análise económica utilizados, consegue-se provar que atitudes de responsabilidade social podem ter impacto positivo na economia, que devem receber o apoio da Administração Pública, e que ajudarão em momentos de incerteza económica. Nesses momentos, a redução das vendas em quase todos os sectores da economia é evidenciada, ao

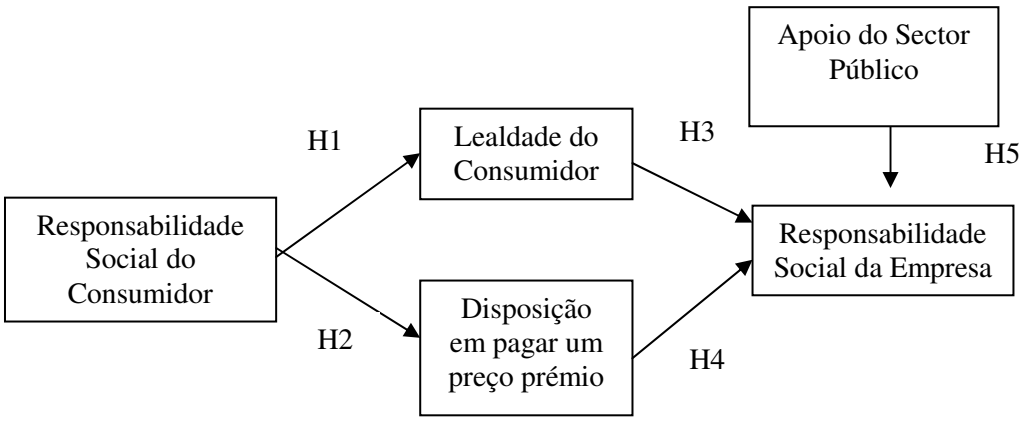
mesmo tempo que a dilatação das elasticidades preço e rendimento da procura. Certamente os valores éticos dos consumidores são bastante mais irrevogáveis e independentes da situação económica vivida. Unindo estes dois factos, poderia adivinhar-se que através da fidelização de clientes que provocou as medidas de RSE, estas empresas vissem menos reduzidas as suas vendas em momentos depressivos.

Em épocas de crise é também mais fácil consciencializar o consumidor para a poupança de recursos e a reeducá-los no sentido da sua responsabilização ética. Se existir o acompanhamento de medidas governamentais anti-crise neste sentido, faz-se pressão quer nos consumidores, quer nos produtores para que sejam mais responsáveis socialmente e assim se promova o crescimento do país. Portanto, os momentos de crise são reais oportunidades para a posta em prática de planos relativos à RSE e RSC, que em si já são fontes de inovação, crescimento e vantagens competitivas. Medidas levadas a cabo pelo nosso actual governo para combater a crise que vivemos, tem mostrado uma preocupação na responsabilidade social dos agentes económicos: apoiam-se projectos e dão-se incentivos ligados ao sector das energias renováveis, financiam-se formações de trabalhadores desempregados que conduzirão a mais produção e consumo ético futuros.

MODELO EMPÍRICO:

O modelo de medida das atitudes de responsabilidade social é assim justificado por tudo o que foi mencionado. Desta forma, vai ser utilizada uma metodologia quantitativa para analisar a percepção dos consumidores face a esta problemática, sabendo que o sentido deverá ser o indicado na figura.

Modelo de Mensuração da Percepção do Consumidor sobre Responsabilidade Social e os seus Impactos:



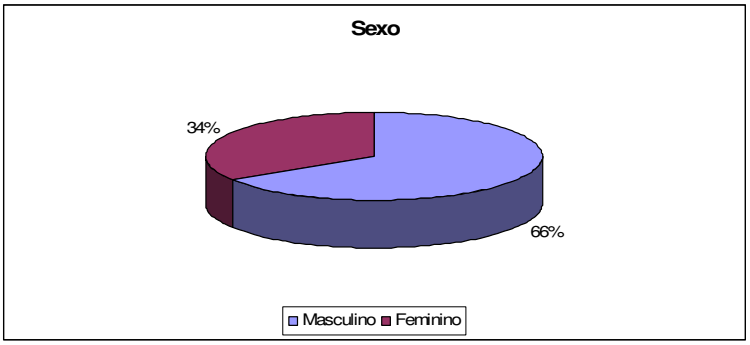
A recolha da informação é feita através de um inquérito juntos aos consumidores, utilizando escalas de responsabilidade social de Roberts (1996), de lealdade de Yoo e Donthu (2001) e de preço prémio de Chauduri e Halbrook (2001).

O inquérito tem como base um questionário individual (apresentado em anexo) que comporta duas partes principais, uma para a recolha da informação sócio-demográfica dos consumidores e a outra parte comporta as questões que vão permitir medir os *constructos* (variáveis quantitativas intangíveis) apresentados no modelo: RSC, RSE, Lealdade, Preço Prémio, Apoio Governamental. A recolha desta informação é feita com uma Escala de Likert de cinco pontos, que traduzem diferentes graus de concordância.

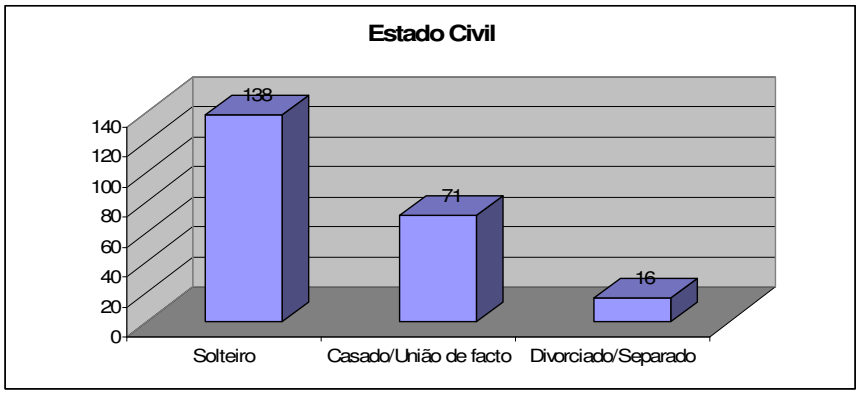
A estimação do modelo de medida é realizada com base na metodologia de Equações Estruturais desenvolvidas por Hair et. al. (2006), Anderson e Gerbing (1988) e Bollen (1989).

RESULTADOS:

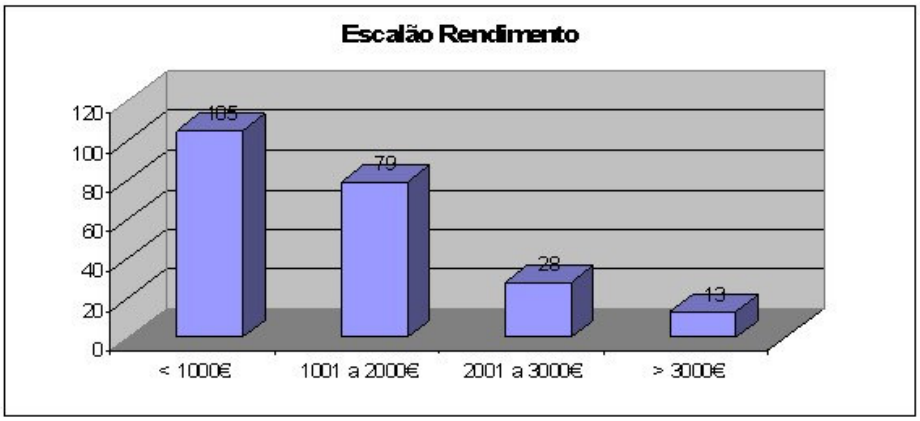
Para a técnica de estimação que se utiliza neste trabalho, recomenda-se amostras cuja dimensão esteja compreendida entre 200 a 250 indivíduos (Hair et. al., 2006; Bollen, 1989). Participaram no estudo 225 indivíduos, da região Norte de Portugal, realizada durante o mês de Abril, cuja caracterização sócio-demográfica é apresentada a seguir. Assim, dos 225 indivíduos, 148 eram do sexo feminino e 77 do sexo masculino.



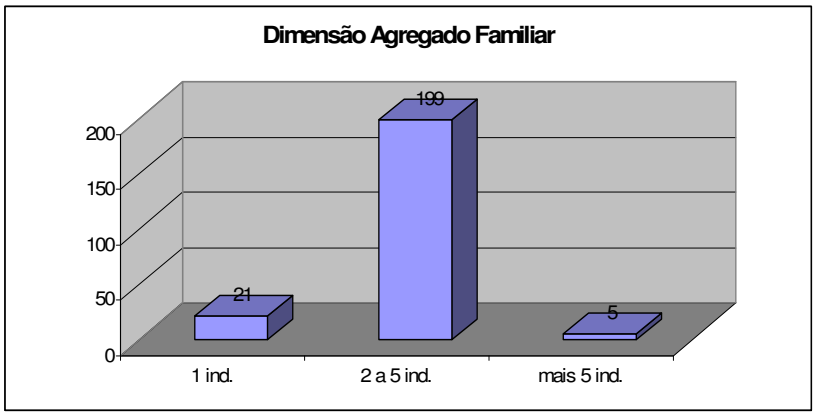
Por outro lado, 61,3% dos inquiridos eram solteiros e 31,5 % casados/união de facto.

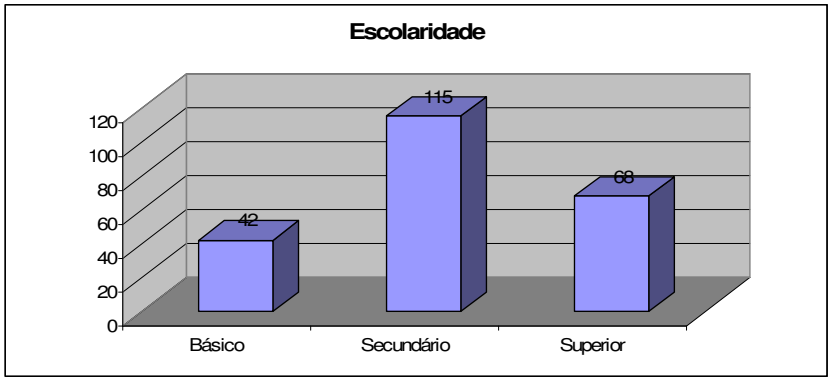


Dos 225 inquirido 105 responderam que tinham um rendimento mensal da família inferior ou igual a 1000€, enquanto 79 indivíduos responderam que tinham um rendimento compreendido entre os 1000 e os 2000€.



A grande maioria dos inquiridos (88,4%) fazem parte de um agregado familiar composto por 2 a 5 indivíduos e a maioria dos inquiridos (51,1%) tem o ensino secundário.





A utilização de equações estruturais é a metodologia adequada para medir e avaliar as relações propostas pelo modelo apresentado. Esta metodologia permite confirmar as relações pretendidas entre constructos endógenos e exógenos, e admite inter-relações entre os próprios constructos exógenos. Na utilização desta metodologia parte-se de um quadro teórico, para que as especificações, modificações ou avaliações das relações de causa-e-efeito nos modelos, possam ser fundamentadas. Assim, cada constructo ou variável latente é medido por um conjunto de questões, incluídas num questionário, aplicado a uma amostra do universo em estudo. Esse conjunto de questões resultam da escolha de escalas de medida, desenvolvidas e validadas em variados estudos e que se encontram especificadas, por constructo.

A segunda parte do questionário contém as questões ligadas com a problemática do estudo. O modelo proposto para analisar a percepção dos consumidores sobre a Responsabilidade Social tem por base o seguinte conjunto de hipóteses:

H1: A RSC evidencia-se através de um comportamento leal do consumidor.
H2: A RSC conduz à disposição em pagar um preço prémio pelo consumidor.
H3: Um consumidor leal explica as práticas de RSE.
H4: Um consumidor disposto a pagar um preço superior pelo produto justifica a RSE.
H5: Em épocas de crise deve-se acentuar as medidas de apoio à RSE.

Para confirmar estas hipóteses, foi utilizada uma escala com 21 itens, cujo objectivo é avaliar a percepção dos consumidores sobre os conceitos em causa. Os itens foram respondidos através de uma escala de Likert de 5 pontos, que se apresenta a seguir.

Discordo fortemente	Discordo	Não concordo nem discordo	Concordo	Concordo fortemente
1	2	3	4	5

Primeiro realiza-se uma análise factorial exploratória a todas as variáveis manifestas, no sentido de verificar se, para a realidade em causa, faz sentido manter as escalas sem nenhum tipo de ajustamento. Isto em resultado do conceito da unidimensionalidade que envolve o estabelecimento de um conjunto de indicadores empíricos relacionados com apenas um único constructo. Depois é feita uma análise da fiabilidade aos itens através da análise do coeficiente Alfa de Cronbach.

Foi realizada uma análise factorial exploratória, através da análise das componentes principais, à escala de responsabilidade social utilizada tendo-se confirmado a existência de três factores distintos, F1 – que traduz a preocupação ambiental, F2 – que traduz a preocupação com a reciclagem e F3 – que traduz a preocupação social dos inquiridos. Estes três factores explicam 65,32% da variância.

Análise Factorial à Escala de Responsabilidade Social

Factores	Total do Valor Próprio (Eigenvalue)	% da Variância do Valor Próprio (Eigenvalue)
F1	5,626	37,504
F2	3,08	30,531
F3	1,093	7,284

Logo de seguida foi realizada a mesma análise para as outras duas escalas, a lealdade do consumidor e a disposição de pagar um preço prémio pelo produto da empresa com responsabilidade social.

Análise Factorial à Lealdade

Factores	Total do Valor Próprio (Eigenvalue)	% da Variância do Valor Próprio (Eigenvalue)
F1	1,914	63,807

Análise Factorial do Preço Prémio

Factores	Total do Valor Próprio (Eigenvalue)	% da Variância do Valor Próprio (Eigenvalue)
F1	2,529	84,3

De seguida é realizada uma análise à fiabilidade das escalas. Para isso, analisa-se o coeficiente Alfa de Cronbach (α), para o conjunto de indicadores manifestos de cada escala utilizada para medir o constructo teórico. Este coeficiente vai resultar directamente das hipóteses do modelo amostral a estudar. Um baixo coeficiente Alfa de Cronbach (α) indica que o conjunto dos itens actua pobremente na captação da construção que motivou a medida. Ao contrário, um alto valor para o coeficiente Alfa de Cronbach (α) indica que os itens estão altamente correlacionados com as verdadeiras classificações. A fiabilidade é a medida da consistência interna de um conjunto de itens de um constructo. Está inversamente relacionada com o grau em que a medida se encontra contaminada com erros.

O grau de aceitação de um indicador de fiabilidade é função da proposta de pesquisa, isto é, depende se é uma pesquisa exploratória ou aplicada. Para além disso, segundo Peterson (1994), nenhuma recomendação para os valores dos indicadores tem uma base empírica, uma justificação teórica ou qualquer processo analítico racional. A recomendação mais seguida para os valores dos indicadores de fiabilidade é a de Nunnally (1978). Este autor recomenda como valores mínimos para os indicadores de fiabilidade, nomeadamente o coeficiente Alfa de Cronbach (α) em pesquisas aplicadas, os valores de 0,7 são tidos como um indicador razoável,

os valores de 0,8 são tidos como um bom indicador e os valores de 0,9 são considerados como um indicador excelente. Contudo, tratando-se de outro tipo de pesquisas os valores alteram-se, isto é, um valor para o coeficiente Alfa de Cronbach (α) de 0,7 será um indicador excelente numa pesquisa preliminar e o valor de 0,8 será considerado excelente para uma pesquisa básica.

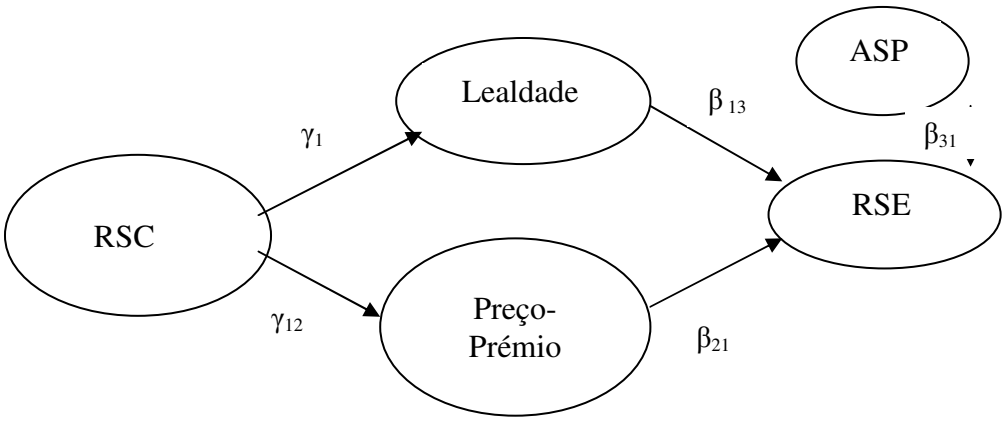
Constructo	Nº Itens	Coeficiente Alfa de Cronbach
Responsabilidade Social	15	0,850
Lealdade	3	0,720
Preço-Prémio	3	0,906

Considerando os três factores da escala de responsabilidade social obtidos pela análise factorial, procedeu-se a análise da sua fiabilidade tendo-se obtido os seguintes resultados.

Constructo	Nº Itens	Coeficiente Alfa de Cronbach
Total da Escala	21	0,881
Preocupação Ambiental	5	0,897
Preocupação com a Reciclagem	4	0,791
Preocupação Social	6	0,813

O modelo estrutural descreve as relações entre as variáveis latentes entre si e indica a quantidade de variância que não é explicada. São vários os constructos teóricos a medir no modelo proposto. Têm em comum, o facto de serem construções de variáveis que não são directamente observáveis. A este tipo de variáveis chama-se variáveis latentes. Os indicadores ou os itens de medida das escalas são denominadas de variáveis manifestas que reflectem as variáveis latentes.

Modelo Estrutural Estimado:



Os modelos de equações estruturais podem ser utilizados para testar e desenvolver teorias, através do modelo de factores comuns, ou podem ser utilizados para fazer previsões através do modelo das componentes principais. Esta distinção e as suas implicações dizem respeito à escolha do método de estimação subjacente ao modelo. Essa escolha vai recair sobre o método de máxima verosimilhança (ML) ou o método de mínimos quadrados generalizados (GLS), se se utilizar o modelo de factores comuns ou, recair sobre a estimação por mínimos quadrados parciais (PLS) se se utilizar o modelo das componentes principais. Para testar e desenvolver teoria, a abordagem através do método da máxima verosimilhança (ML) ou o método de mínimos quadrados generalizados (GLS) é preferível (Hair et. al. (2006), Bollen (1987)).

Modelo Estrutural Estimado

Hipóteses Principais	Parâmetro	Estimativa	p-Value	Conclusão
RSC → Lealdade (+)	γ_{11}	0,812	0,004	Suportada
RSC → Preço-Prémio (+)	γ_{12}	0,720	0,002	Suportada
Lealdade → RSE (+)	β_{13}	0,163	0,109	Não Suportada
Preço-Prémio → RSE (+)	β_{21}	0,036	0,142	Não Suportada
ASP → RSE (+)	β_{31}	0,038	0,110	Não Suportada
Indicadores do Ajustamento				
χ^2 Normalizado		2,880		
RMSEA		0,070		
GFI		0,844		
AGFI		0,797		

Os indicadores escolhidos são os sugeridos por Hair et. al. (2006) como os melhores indicadores absolutos sobre a bondade do ajustamento, Qui-Quadrado, RMSEA (*Root Mean Square Error of Approximation*) e GFI (*Goodness-of-Fit*).

Utiliza-se o Qui-Quadrado Normalizado para diminuir a sensibilidade do teste à dimensão da amostra (n = 225). O Qui-Quadrado Normalizado apresenta valores aceitáveis quando estes se encontrem compreendidos entre 1 a 3 (Hair, et. al., 2006). É utilizado o RMSEA (*Root Mean Square Error of Approximation*) em vez do RMSSR (*Root Mean Square Residual*) porque o modelo é estimado com base na matriz de covariâncias dos dados. Este indicador deve estar situado entre os valores compreendidos de 0,05 (bom ajustamento) a 0,08 (ajustamento aceitável). O GFI (*Goodness-of-Fit*) é um índice da bondade do ajustamento que representa o grau total do ajustamento, sem uma correcção relativamente aos graus de liberdade dos modelos. Altos valores deste indicador traduzem um bom ajustamento, mas não se encontram estabelecidos níveis mínimos aceitáveis.

CONCLUSÕES:

Confirmadas as hipóteses levantadas, o modelo de mensuração pode ser utilizado pelas empresas e entidades públicas nas suas estratégias de gestão e de política económica.

Conclui-se pela análise dos resultados que a responsabilidade social dos consumidores faz com que estes sejam leais e estejam dispostos a pagar um preço mais alto por produtos provenientes de empresas com o mesmo sentido ético. Contudo, eles não conseguem perceber que as empresas levam a cabo este tipo de medidas, e ainda, que estas têm o apoio estatal. É o mesmo que dizer que não se verificam as relações entre a lealdade dos consumidores e o preço-prémio com a responsabilidade social da empresa.

Assim, este trabalho alerta para a urgente necessidade de se activarem, quer por parte das empresas quer do sector público, formas de divulgação das acções de responsabilidade sociais que são empreendidas, em épocas de crise como as que estamos a viver, já que os consumidores poderão reconhecê-las e recompensá-las.

O estudo deve ser desenvolvido no sentido da obtenção de um novo modelo explicativo destas relações, que a nível micro e macro económico fazem todo o sentido.

REFERÊNCIAS BIBLIOGRÁFICAS:

ANDERSON, J. C. E D. W. Gerbing (1988), "Structural Equation Modelling in Practice: A Review and Recommended Two-Step Approach", *Psychological Bulletin*, Vol. 103, Nº 3, pp. 411-423.

BARNETT, M. L. (2007), "Stakeholder Influence Capacity and the Variability of Financial Returns to Corporate Social Responsibility", *Academy of Management Review*, Vol. 32, Nº 3, pp. 794-816.

BHATTACHARYA, C. B. AND S. SEN (2004), "Doing Better at Doing Good: When, Why and How Consumers Respond to Corporate Social Initiatives", *California Management Review*, Vol. 47, Nº 1, pp. 9-24.

BOLLEN, K. (1989), *A Structural Equation with Latent Variables*, New York: Wiley.

BOULSTRIDGE, E. AND M. CARRIGAN (2000), "Do Consumers Really Care About Corporate Responsibility? Highlighting the Attitude-Behaviour Gap", *Journal of Communication Management*, Vol. 4, Nº 4, pp. 355-368.

CASTALDO, SANDRO, FRANCESCO PERRINI, NICOLA MISANI AND ANTONIO TENCATI (2009), "The Missing Link Between Corporate Social Responsibility and Consumer Trust: The Case of Fair Trade Products", *Journal of Business Ethics*, Vol. 84, pp. 1-15.

CASTALDO, S. AND F. PERRINI (2004), "Corporate Social Responsibility, Trust Management and Value Creation", presented at EGOS 2004 "Trust in Hybrids", Ljubljana, Slovenia.

CHAUDHURI, A. E M. B. HOLBROOK (2001), "The Chain of Effects from Brand trust and Brand Affect to Brand Performance: The Role of Brand Loyalty", *Journal of Marketing*, Vol. 65, pp. 81-93.

COMISSÃO EUROPEIA (2001), *Livro Verde – Promover um Quadro Europeu para Responsabilidade Social das Empresas*, COM 366, Julho.

COMISSÃO EUROPEIA (2008), *Communication from the Commission on the European Competitiveness Report 2008*, Brussels 28.11.2008.

FOMBRUN, C. ET AL.(2000), "Opportunity platforms and safety nets: corporative citizenship and reputational risk", *Business and Society Review*, Vol.105, Nº1, pp.85-106.

FOX, T., WARD, H. E HOWARD, B. (2002), "Public sector roles in strengthening corporate social responsibility: a baseline study", Publicação do Banco Mundial.

HAIR, J. F., R. E. ANDERSON, R. L. TATHAM E W. C. BLACK (2006), *Multivariate Data Analysis*, 6th Ed., Pearson, Prentice-Hall, Upper Saddle River, N. J..

INSTITUTO PORTUGUÊS DA QUALIDADE, integrado no Ministério da Economia e da Inovação (2008) http://www.ipq.pt/custompage.aspx?pagid=4290

KOTLER, PHILIP (1997), "Competitividade e Caráter Cívico" in Frances Hesselbein et al.(editores), Peter F. Drucker Foudation (organização), "A organização do futuro: Como preparar hoje as empresas de amanhã", Ed.Futura.

NUNNALLY, J. C. (1978), *Psychometric Theory*, 2nd Ed., New York: McGraw-Hill.

PAGE, G. AND H. FEARN (2005), "Corporate Reputation: What do Consumers Really Care About?", *Journal of Advertising Research*, Vol. 45, Nº 3, pp. 305-313.

PERRINI, F., S. POGUTZ AND A. TENCATI (2006), *Developing Corporate Social Responsibility. A European Perspective*, Edward Elgar Publishing Ltd, Cheltenham, UK.

PETERSON, ROBERT A. (1994), "A Meta-analysis of Cronbach's Coefficient Alpha", *Journal of Consumer Research*, Vol. 21, Nº 2, pp. 381-391.

PORTER, M. AND KRAMER, M.(2006), "Strategy and Society: the link between competitive advantage and corporate social responsibility", *Harvard Business Review*, December 2006.

PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO (1998), *Consumo sustentável*, Trad. Admond Ben Meir. São Paulo: Secretaria do Meio Ambiente / IDEC / CONSUMERS INTERNATIONAL.

ROBERTS, J. A. (1996), "Will the Real Socially Responsible Consumer Please Step Forward?", *Business Horizon*, Vol. 39, pp. 79-83.

SMITH, N. C. (2003), "Corporate Social Responsibility: Whether or How?", *California Management Review*, Vol. 45, Nº 4, pp. 52-76.

YOO, B. AND N. DONTHU (2001), "Developing and Validating a Multidimensional Consumer-Based Brand Equity Scale", *Journal of Business Research*, Vol. 52, pp. 1-14.

Anexo:
Inquérito: Responsabilidade Social do Consumidor e da Empresa em Época de Crise

Objectivo: recolha de dados para um trabalho de Investigação sobre a percepção da Responsabilidade Social do Consumidor e da Empresa.
Agradecemos muito a sua colaboração no preenchimento deste questionário.

Caracterização do Inquirido: assinale com uma cruz					
01	Idade:				
02	Sexo:	Masculino □	Feminino □		
03	Estado Civil	Solteiro(a) □	Casado/União de facto □	Divorciado/Separado □	
04.	Escalão Rendimento Mensal da Família	≤ 1000€ □	1001 a 2000 € □	2001 a 3000€ □	> 3000 € □
05.	Dimensão do Agregado Familiar	1 indivíduo □	2 a 5 indivíduos □	Mais de 5 indivíduos □	
06.	Escolaridade	Básico □	Secundário □	Superior □	

Tenha presente a escala de 1 a 5 aqui representada. A seguir são apresentadas diversas afirmações. Indique com uma cruz o seu grau de concordância sobre cada uma. Não existem respostas certas ou erradas. Queremos apenas a sua opinião sincera.

Discordo fortemente	Discordo	Não concordo nem discordo	Concordo	Concordo fortemente
1	2	3	4	5

	1	2	3	4	5
Quando posso escolher, dou preferência a produtos que causem menos poluição.					
Tenho cuidado na compra de produtos menos poluidores.					
Quando posso escolher entre dois produtos semelhantes, compro o que é menos prejudicial para as outras pessoas e para o meio ambiente.					
Escolho os produtos em função da menor poluição que causem.					
Tenho escolhido determinados produtos por razões ecológicas.					
Tento comprar produtos que podem ser reciclados.					
Faço o possível para comprar produtos de papel reciclado.					
Sempre que possível compro produtos empacotados em embalagens recicláveis.					
Utilizo algum meio de reciclagem para o meu lixo de casa.					
Não compro produtos de empresas que não aceitam funcionários negros, homossexuais ou mulheres.					
Não compro produtos de empresas que utilizam mão-de-obra infantil.					
Não compro produtos de empresas onde a publicidade mostra mulheres de forma negativa.					
Não compro produtos de empresas onde a publicidade mostra grupos minoritários de forma negativa.					
Não compro produtos de empresas cujos trabalhadores moram e trabalham em más condições.					
Não compro um produto de uma empresa que é socialmente responsável.					
Produtos de empresas socialmente responsáveis são sempre a minha primeira escolha.					
Se produtos de empresas socialmente responsáveis não estão disponíveis nas lojas, não escolho outro produto.					
Considero-me leal a produtos de empresas socialmente responsáveis.					
Comprar produtos de empresas socialmente responsáveis parece-me uma boa ideia apesar de serem mais caros.					
Estou disposto a pagar um preço mais elevado por produtos de empresas socialmente responsáveis.					

Continuarei a comprar produtos de empresas socialmente responsáveis mesmo que produtos idênticos de outras empresas não socialmente responsáveis sejam mais baratos.					

Pensa que as crises económicas alteram o comportamento dos consumidores e das empresas, no que diz respeito, às práticas de responsabilidade social? Sim____Não____

Muito Obrigada pela sua participação!